13

Size-Based Models

13.1 Introduction

13.1.1 Stock Assessment Modelling Options

Commercial fisheries for abalone and rock lobster (and many other invertebrates) often suffer from the fact that the species concerned are difficult or impossible to age using readily available technology. Nevertheless, many of these species are the basis of valuable fisheries, and thus require an assessment of some kind to assist with the adequate management of each stock. The use of age-structured models for assessing these species is compromised, so alternatives must be considered. It would be possible to use a surplus production model, which does not require age-structured information (see Example Boxes 8.4 and 8.5 for an example with abalone). However, an alternative that permits the use of more than just catch and catch rate data would be to use a size- or stage-structured model, with the basic form of these models described by Sullivan et al. (1990) and Sullivan (1992); also see Caswell (2001), and a fully developed model for abalone is described by Breen et al. (2003). Such models follow the fate of the numbers in a set of size classes, which contrasts with age-structured models that follow numbers in each age class or cohort through time. A major difference is that the size classes are not related to specific cohorts, and so the growth of individuals passing from one size class into another is not automatic as the years pass. The transition of animals from size class to size class requires an adequate description of the growth of the species concerned. In particular, this would need to be in terms of the expected growth increment of given sizes of animals, which preadapts the method to use growth estimates derived from tagging studies. In addition, one requires data on catches and catch rates, but one can also include data describing the size distribution of the commercial catch (commonly collected and known as shed or market sampling), as well as fishery-independent surveys of abundance and field surveys of the size distribution of abalone after fishing. As with the age-structured integrated analysis, it is possible to include different forms of ancillary data once the basic size-based stock assessment model is developed. Size-structured models are relatively complex, and the examples in this chapter will need to be developed in stages (Figure 13.1); we will use
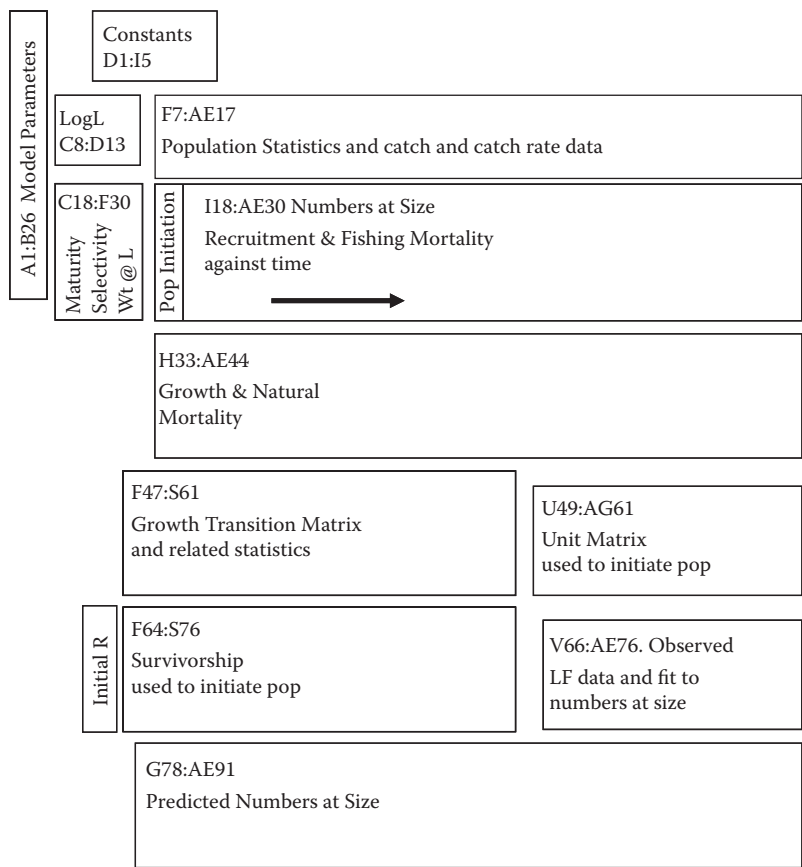
381

Fig ur e 13.1
Worksheet map for the size-based model to be developed in this chapter; it will be developed in five stages. There are two selectivity curves to reflect a change of legal minimum length from 132 mm to 140 mm in 1990.

the same western zone abalone data from Tasmania, Australia, as was used in Chapter 8 (Table 8.1). The examples in this chapter use only twelve 10 mm size classes, which is a major simplification. With Tasmanian blacklip abalone (Haddon, 2009), seventy-six 2 mm size classes are used to describe the populations being assessed.

13.2 The Model Structure

The population being modelled is described by a vector, **Nt**, of the numbers in a defined set of size classes at time *t*:

© 2011 by Taylor & Francis Group, LLC

$$\mathbf{N_t} = \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ . \\ N_{n,t} \end{bmatrix} \tag{13.1}$$

where **Nt** is a vector of n size classes indicating the numbers in length class i at time t. At each time step, the individuals in the population in **Nt** either grow or not, which implies that they either stay in their original size class or move into a larger one (negative growth is usually disregarded, though in principle it could be included). Mathematically, this growth is described by using a transition or projection matrix containing the probabilities of shifting from one size class into others in whatever time step the model is designed around. The population vector is repeatedly multiplied by the growth transition matrix to describe the changing population size structure brought about by growth through time. The transition matrix contains the probability that the individuals in length class j (matrix columns) at time t have grown into length classes $i*$ (matrix rows) during a single time increment. The growth transition matrix can be combined with survivorship probabilities and recruitment relationships (Caswell, 2001); however, it is often simpler and more flexible to keep them separate. The proportions of animals growing from length class j into length classes $i*$ are estimated using a probability density function whose parameters are the average expected growth increment (from each initial length class j) surrounded by some measure of variation. Any probability density function of a suitable shape could be used to describe the growth increments, although, naturally, the distribution used should reflect the observed distribution of growth increments; the normal distribution has been used (Punt et al., 1997; Breen et al., 2003; Haddon et al., 2008), as has the gamma distribution (Sullivan, 1992). In matrix notation the growth transition matrix would be represented thus:

$$\mathbf{G} = \begin{bmatrix} G_{1,1} & 0 & . & 0 \\ G_{1,2} & G_{2,2} & & \\ . & & . & 0 \\ G_{1,n} & . & G_{n-1,n} & G_{n,n} \end{bmatrix} \tag{13.2}$$

where **G** is the square transition or projection matrix in which the $\mathbf{G}_{i,j}$ are the probabilities that an animal in size class j (along the columns) will grow into size class i (along the rows). The upper triangle of zeros indicates that negative growth is not being modelled (if it were, then some of these matrix cells would be nonzero). The generation of the individual $\mathbf{G}_{i,j}$ entails using a probability distribution to describe the different likelihood of growing from one size to another, rather than using some deterministic growth trajectory.

© 2011 by Taylor & Francis Group, LLC

Generally, a growth curve, such as the von Bertalnaffy curve, is used to estimate the average length increment for a given size class, and then the probability distribution chosen is used to describe how individuals would be distributed around the average expected increments. Using a normal distribution to describe the transition probabilities, we need to estimate the proportion of the expected distribution of growth increments to be found between the limits of each size class. With the focus on the bottom of each size class (L_i) of width CW, the expected probabilities of growing from size class j into size class i ($G_{i,j}$) would be defined as

$$G_{i,j} = \int_{-\infty}^{L_i+CW} \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\left[\frac{\left(L_i-\bar{L}_j\right)^2}{2\left(\sigma_j\right)^2}\right]} dL \qquad L_i = L_{Min}$$

$$G_{i,j} = \int_{L_i}^{L_i+CW} \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\left[\frac{\left(L_i-\bar{L}_j\right)^2}{2\left(\sigma_j\right)^2}\right]} dL \qquad L_{Min} < L_i < L_{Max}$$

$$G_{i,j} = \int_{L_i}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\left[\frac{\left(L_i-\bar{L}_j\right)^2}{2\left(\sigma_j\right)^2}\right]} dL \qquad L_i = L_{Max} \qquad (13.3)$$

where σ_j is the standard deviation of the normal distribution of growth increments for size class j. $\bar{L}_j$ is the expected average size for size class j after the growth expected in one time interval, that is, $L_j + \Delta\hat{L}_j$, where $\Delta\hat{L}_j$ is the average expected growth increment for size class j. Summing the smallest size class to $-\infty$ and the largest size class to $+\infty$ effectively makes both these size classes plus groups that ensure that the transition probabilities for all n size classes sum to 1. Using the Fabens version of the von Bertalanffy growth equation means that

$$\Delta\hat{L}_j = \left(L_\infty - L_j\right)\left(1 - e^{-K}\right) \qquad (13.4)$$

where L_∞ is the length at which growth increments are zero and K is the rate at which the maximum length, L_∞, is approached. However, any other growth curve that can predict growth increments for a given initial length could be used. For example, in Tasmania, the growth of blacklip abalone (*Haliotis rubra*) appears to be best described with a growth trajectory that reflects an inverse logistic curve (Haddon et al., 2008).

Multiplying the numbers-at-size vector, $\mathbf{N}_t$, by the transition matrix leads to

$$\mathbf{N_{t+1}} = \mathbf{GN_t} \tag{13.5}$$

which describes growth in the population size structure without recruitment or mortality (Example Box 13.1).

The full stage-structured model uses the transition matrix along with descriptions of the processes of mortality and recruitment to describe the calculation of **Nt+1**. The population dynamics can be described by repeated application of these processes. Natural mortality can be represented in this model by the annual survivorship in year t in each length class i, $S_{i,t}$, and is easily included in the model either as a vector of individual terms or as a zero square matrix with only the diagonal elements filled:

$$\mathbf{S_t} = \begin{bmatrix} S_{1,t} & 0 & & 0 \\ 0 & S_{2,t} & & \\ & & \cdot & 0 \\ 0 & & 0 & S_{n,t} \end{bmatrix} \tag{13.6}$$

Annual survivorship, accounting for natural mortality only, would be

$$S_{i,t} = e^{-M_{i,t}} \tag{13.7}$$

where $M_{i,t}$ refers to the instantaneous rate of natural mortality for size class i during period t; generally this is held constant across size classes and through time. If there is also some size-selective fishing mortality, then the annual survivorship would include

$$S_{i,t} = e^{-(M_{i,t}+s_i F_t)} \tag{13.8}$$

where s_i is the selectivity that applies to length class i and F_t is the fully selected instantaneous fishing mortality in year t. Using matrix notation, the model now becomes

$$\mathbf{N_{t+1}} = \mathbf{GSN_t} \tag{13.9}$$

Finally, the expected recruitment in each time step can be literally added as a vector of the numbers of recruits into each size class (Example Box 13.2):

$$\mathbf{N_{t+1}} = \mathbf{GSN_t} + \mathbf{R_t} \tag{13.10}$$

Care must be taken concerning the order of multiplication of the growth and survivorship matrices with the numbers-at-size vector. It is the case that

EXAMPLE BOX 13.1

The generation of the size transition matrix, using abalone as an example. The placement allows space for later developments. Cells H1:I4 are constants. SigmaG is the standard deviation of the growth residuals, and SigmaR will be used in the penalty that limits the variation of the recruitment residuals. Note the jump from 170 to 200 in G61; this group acts as a plus group so each column sums to 1.0. Into H47 put =If((H49+H48)<H49,H49,H49+H48), which precludes negative growth. Put =($I$1–H49)*(1–EXP(–$I$2)) into H48. Select H47:H48 and copy over to column S. Now the complicated bit: put =if(H$49>$F50 ,0,if(H$49=$F50,normdist($G50,H$47,$I$3,true),normdist($G50,H$47, $I$3,true)–normdist($F50,H$47, $I$3,true))) into H50. Check all brackets and $ signs. Copy down to H61 and across to column S; this is the transition matrix. Select H50:S61 and name it "Grow" in the name box just above the "A" column to simplify later use. To illustrate the use of the transition matrix, in G19:G30 put the size classes 60 to 170, and in H18 to AE18 put the years 1985 to 2008. Put zeros into H20:H30 and 1e6 into H19. Select I19:I30 and type =mmult(Grow,H19: H30), then press <Ctrl><Shift><Enter> all at once to enter the matrix multiplication array function. Select I19:I30 and copy across to column AE. The effect of growth without mortality or recruitment should be clear. Vary K and Linf to see the effect.

	F	G	H	I	J		R	S
1			Linf	172				
2			K	0.36				
3			SigmaG	5.714				
4			SigmaR	0.6				
47		PredL	93.860	100.837	107.814		163.628	170.605
48		ExptDL	33.860	30.837	27.814		3.628	0.605
49		Grow	60	70	80		160	170
50	60	70	0.0000	0.0000	0.0000		0.0000	0.0000
51	70	80	0.0076	0.0001	0.0000		0.0000	0.0000
52	80	90	0.2420	0.0288	0.0009		0.0000	0.0000
53	90	100	0.6090	0.4128	0.0848		0.0000	0.0000
54	100	110	0.1389	0.5038	0.5633		0.0000	0.0000
60	160	170	0.0000	0.0000	0.0000		0.8676	0.0000
61	170	200	0.0000	0.0000	0.0000		0.1324	1.0000

© 2011 by Taylor & Francis Group, LLC

EXAMPLE BOX 13.2

Simple size-based model dynamics. To include natural and fishing mortality plus recruitment add the constants as indicated in D1:I5, and copy I17 to AE17. Maturity and weight-at-length will be used later, but selectivity will be used now. In C19 put Equation 13.21 =1/(1+(exp($E$3+$E$4*G19))^–1), and put Equation 13.20 into D19: =1/(1+exp(–Ln(19)*(G19–$G$1)/($G$2–$G$1))). The legal minimum length changed in 1990, so put =1/(1+exp(–Ln(19)*(G19–$G$3)/($G$4–$G$3))) into E19, and put Equation 13.19 into F19 =($E$1*G19^$E$2)/1000. Select C19:F19 and copy down to row 30. Fill H20:H30 with zeros and put 1e6 in H19. Select H33:H44 and type =mmult(Grow,H19:H30)*$G$5, and then press <Ctrl><Shift><Enter> all at once to enter the matrix multiplication array function; this is growth and natural mortality (Equation 13.9). Select H33:H44 and copy over to column AE. To include fishing mortality and recruitment, put =(1000000)+H33 into I19 (i.e., constant recruitment of 1e6), and in I20 put =H34*(1–H$17*$D20) and copy down to I30. Select I19:I33 and copy across to column M. Because the selectivity changed in 1990, alter M20 to become =L34*(1–L$17*$E20), then copy down to M30 across to column AE. Compare the size distribution in column AE if the harvest rate in H17:AE17 is set at 0.2 or at 0.0 (Figure 13.2).

	C	D	E	F	G	H	I
1		Wta	5.62E–05	SelL50_85	127	Linf	172
2		Wtb	3.1792	SelL95_85	132	K	0.36
3		SaMa	–16.8	SelL50_90	139	Sigma	5.714
4		SaMb	0.14	SelL95_90	146	SigmaR	0.6
5		M	0.2	NatSurv	=exp(–E5)	TAC	900
17				Harvest		0.2	=H17
18	Maturity	Sel to 89	Sel 90 on	Weight at L	SizeClass	1985	1986
19	0.0002	0.0000	0.0000	0.0253	60	1000000	1000012
20	0.0009	0.0000	0.0000	0.0413	70	0	6243
30							
	0.9991	1.0000	1.0000	0.6931	170	0	0
33					60	12	12
34					70	6243	6244
35					80	198144	198442
44					170	0	0

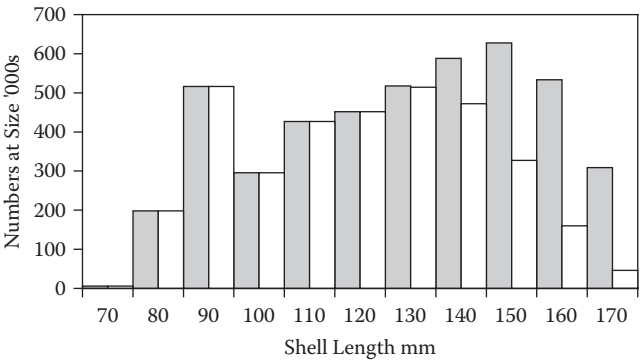© 2011 by Taylor & Francis Group, LLC

Fig ur e 13.2
The size structure obtained in a hypothetical abalone population with a constant recruitment of 1 million individuals into the 60 mm size class each year and a constant harvest rate of either 0.2 (20%; the open columns) or 0.0 (the grey columns), as in Example Boxes 13.1 and 13.2. The difference shows the effect of fishing on a virgin stock.

G(SNt) (survive first, then grow) is only the same as **S(GNt)** (growth first, then survive) if the survivorship is constant across all sizes, which will not be the case if there is size-selective fishing. Which order is chosen in practice will depend on when in the year the fishery concerned tends to operate. One way of reducing the impact of when mortality is applied, relative to when growth is applied, is to break each year up into a number of seasons and generate a transition matrix for each season. This is the strategy used when generating transition matrices to describe the growth of the southern rock lobster, *Jasus edwardsii*, around Tasmania, Australia; Punt et al. (1997) used twelve monthly periods, while Haddon and Gardner (2009) used eight periods (six monthly and two three-monthly). Combining such partial year matrices in the appropriate order means the recruitment and mortality processes that occur in a particular fishery may be represented more realistically; however, here in our examples we will restrict ourselves to an annual model.

The dynamics of a population can thus be described if one has details of growth, an estimate of natural and fishing mortality, and an average annual recruitment. The recruitment can be represented by a vector of zeros except for the first size class, though in reality we might expect new recruits to be spread across a few size classes, especially if the size classes were small enough. By repeatedly applying Equation 13.10, that is, multiplying the vector of numbers-at-size by the survivorship matrix, the growth transition matrix, and then adding the recruitment vector, with or without a constant fishing mortality, eventually a stable size structure and population size becomes established (Sullivan, 1992; Example Box 13.2, Figure 13.2). Such asymptotic behaviour is also useful for determining the expected size structure (and hence the biomass, given weight-at-length) in the absence of fishing. This unfished or virgin biomass is part of various

biological reference points often used as targets and limits when managing a resource.

A complete history of each fishery from its start to the current day is often not known, and so instead of starting the modelling from an unfished state, it would be useful if we could initiate the model by estimating the size structure and state of depletion in the first year when data of sufficient quality and completeness are available. It would be possible to do this in the same way as estimating the unfished state by using an average recruitment and an initial harvest rate and repeatedly applying these until an equilibrium arises (see Example Box 13.2). However, in an assessment context using asymptotic behaviour is inefficient, as it can take many iterations to achieve equilibrium. Instead, it is possible to use relatively simple matrix algebra to derive the initial equilibrium size structure in an analytical fashion. At equilibrium, $\mathbf{Nt+1}$ is the same as $\mathbf{Nt}$ (represented by $\mathbf{N^*}$):

$$\mathbf{N}^* = \mathbf{GSN}^* + \mathbf{R} \tag{13.11}$$

$$\mathbf{N}^* - \mathbf{GSN}^* = \mathbf{N}^*\left(\mathbf{I} - \mathbf{GS}\right) = \mathbf{R} \tag{13.12}$$

and thus

$$\mathbf{N}^* = \left(\mathbf{I} - \mathbf{GS}\right)^{-1}\mathbf{R} \tag{13.13}$$

where $\mathbf{I}$ is the identity or unit matrix (see Figure 13.1), $\mathbf{S}$ is a square survivorship matrix describing the combination of both natural mortality and size-selective fishing mortality, and $\mathbf{R}$ is a vector of recruitment values in each size class (Example Box 13.3). Explicitly this is denoted

$$
\begin{bmatrix} N_1 \\ N_2 \\ \cdot \\ N_n \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & \\ & & \cdot & 0 \\ 0 & & 0 & 1 \end{bmatrix} - \begin{bmatrix} G_{1,1} & 0 & & 0 \\ G_{1,2} & G_{2,2} & & \\ \cdot & \cdot & \cdot & 0 \\ G_{1,n} & \cdot & \cdot & G_{n,n} \end{bmatrix} \cdot \begin{bmatrix} S_1 & \mathbf{0} & & 0 \\ 0 & S_2 & & \\ & & \cdot & 0 \\ 0 & & 0 & S_n \end{bmatrix} \right)^{-1} \begin{bmatrix} \overline{R} \\ 0 \\ \cdot \\ 0 \end{bmatrix} \tag{13.14}
$$

In practice, instead of generating matrices for the survivorship terms, to speed the computations, the survivorship following fishing mortality may be added to the noninitiation dynamics independently of the survivorship from natural mortality. Equation 13.10 thus becomes

$$\mathbf{N_{t+1}} = \left(\left(\left(\mathbf{GN_t}\right)e^{-M}\right) \otimes \mathbf{e^{-sF}}\right) + \mathbf{R_t} \tag{13.15}$$

where the numbers-at-size is matrix multiplied by the growth transition matrix, and then the elements of the resulting numbers-at-size are multiplied by the scalar describing the survivorship following natural mortality,

© 2011 by Taylor & Francis Group, LLC

EXAMPLE BOX 13.3

The survivorship and unit matrices. Extend the recruitment deviates as model parameters from B6:B26 as rdev88 to rdev08; set these at 1.0. Put =exp(average (H15:AE15)) into B1 to estimate the closed form of *q* for later (Equation 13.25). Fill E66:E76 with zeros and copy F67 down to F76. Extend the size classes from G65 to 170 in G76, and from H64:S64. Fill H65:S76 with zeros and then put =($G$5–($D19*$F65)) into H65, making sure the $ are put in correctly. Copy H65 and paste it diagonally to fill I66, J67, and so on to S76; this makes a square survivorship matrix used in the initiation. Select H65:S76 and name it "Surv" in the name box. Copy G64:S76 and paste as values in U49. Replace the diagonal values with 1s to form a unit matrix. Select V50:AG61 and name it "Unit." To generate the initial vector of numbers-at-size select H19:H30 and type =mmult(minverse(Unit–mmult (Grow,Surv)),E65:E76), then press <Ctrl><Shift><Enter> all at the same time to enter the array functions used. Replace I19 with =($B$2*$B4)+H33 and copy across to column AE. Then in J19 replace the reference to $B4 with a =($B$2*$B5)+I33 so that it references the recruitment deviate for 1987. In K19 have it point to B6. This change needs to be made right across to AE19, which should read =($B$2*$B26)+AD33. These will be used when we fit the model to data. If you change B2:B3 the initial size distribution will also change.

	A	B		E	F	G	H	I
1	q	#DIV/0!		5.62E–05	SelL50_85	127	Linf	172
2	AvRec	11112562		3.1792	SelL95_85	132	K	0.36
3	InitH	0.26632		–16.8	SelL50_90	139	Sigma	5.714
4	rdev86	1		0.14	SelL95_90	146	SigmaR	0.6
5	rdev87	1		0.2	NatSurv	0.8187	TAC	900
64				InitR	Harvest	Survival	60	70
65				=$B$2	=$B$3	60	0.8187	0
66				0	=F65	70	0	0.8187
67				0	=F66	80	0	0
68				0	0.2663	90	0	0

e^{-M}, which is constant across size classes and time, to form a new intermediary numbers-at-size, which is then multiplied element by element by the vector representing survivorship following fishing mortality, $\mathbf{e}^{-s\mathbf{F}}$, which represents nonconstant survivorship across size classes due to the selectivity vector **s** (the symbol $\otimes$ represents element-by-element multiplication). Finally, the vector of recruitments for the given year (the average recruitment multiplied by the lognormal recruitment deviate; $\bar{R} \times e_t^{N(0,\sigma_R^2)}$) is added to the

© 2011 by Taylor & Francis Group, LLC

numbers-at-size to generate the final numbers-at-size for time $t + 1$. Given that the harvest rate, H_t, can easily be estimated,

$$H_t = C_t / B_t^E \qquad (13.16)$$

where C_t is the catch taken in year t and B_t^E is the exploitable biomass in year t. Equation 13.22, the complement of this, adjusted for selectivity, would be the survivorship following fishing mortality; thus, Equation 13.15 can be represented as

$$\mathbf{N_{t+1}} = \left(\left((\mathbf{GN_t}) e^{-M} \right) \otimes (1 - \mathbf{s_t} H_t) \right) + \mathbf{R_t} \qquad (13.17)$$

which is what is used in Example Boxes 13.2 and 13.3. In addition, the predicted vector of numbers-at-size in the commercial catch, $\mathbf{N_t^C}$, which can be compared with the observed catch at size, can then be described by

$$\mathbf{N_t^C} = \left(\left((\mathbf{GN_t}) e^{-M} \right) \otimes (\mathbf{s_t} H_t) \right) \qquad (13.18)$$

where the vector of numbers-at-size after natural mortality and growth is multiplied element by element by the vector of selectivity, $\mathbf{s_t}$, for that year, t, after it has been multiplied by the scalar representing the harvest rate, H_t, in year t.

There are various other simple relationships that are required to connect the population dynamics model to a real fishery. Just as with the age-structured models described in Chapter 12, we require a formal description of any selectivity that applies to the fishing (determined by the legal minimum length and diver behaviour); in addition, a length-to-weight relationship is also required. A simple power relationship can be used between length and weight:

$$w_i = aL_i^b \qquad (13.19)$$

where w_i is the expected weight of an animal in length class i, a and b are curve parameters, and L_i is the length of abalone in length class i. Selectivity is described using a logistic curve:

$$s_i = \frac{1}{\left(1 + \exp\left(-Ln(19) \dfrac{L_i - L_{50}}{L_{95} - L_{50}} \right) \right)} \qquad (13.20)$$

where L_i is the length of length class i, L_{50} is the length at which 50% of abalone are selected, and L_{95} is the length at 95% selection. Size-at-maturity is described using an alternative version of the logistic:

$$m_i = \frac{1}{\left[1 + \left(\exp\left(a + bL_i\right)\right)^{-1}\right]} \qquad (13.21)$$

where m_i is the proportion mature at length L_i of length class i, and a and b are the logistic parameters ($-b/a$ is the size at 50% maturity). To generate predicted catch rates and harvest rates we need to be able to estimate the exploitable biomass:

$$B_t^E = \sum_{i=1}^{n} N_{t,i} w_i s_i \qquad (13.22)$$

where B_t^E is the exploitable biomass at time t, n is the number of size classes, $N_{t,i}$ is the numbers in size class i at time t, w_i is the weight of size class i, and s_i is the selectivity of the size class i (if s_i were omitted, this would generate the total biomass). To estimate the degree of depletion of the spawning stock, we need to be able to estimate the spawning biomass:

$$B_t^S = \sum_{i=1}^{n} N_{t,i} w_i m_i \qquad (13.23)$$

where B_t^S is the spawning or mature biomass at time t, n is the number of size classes, $N_{t,i}$ is the numbers in size class i at time t, w_i is the weight of size class i, and m_i is the proportion mature within size class i.

Just as with surplus production and age-structured models, we use the concept of catchability to make predictions of the expected catch rates, $\hat{I}_t$, given a particular exploitable biomass in a given year t:

$$\hat{I}_t = q B_t^E \qquad (13.24)$$

where q is the catchability (defined as the proportion of the exploitable biomass taken with one unit of effort). As with surplus production and age-structured models, a closed-form estimate of catchability can be obtained using the observed catch rates and the estimates of exploitable biomass:

$$\hat{q} = e^{\frac{1}{n}\Sigma Ln\left(\frac{I_t}{\hat{B}_t^E}\right)} \qquad (13.25)$$

Finally, when fitting the model to observations from a fishery, we need some objective function to optimize the model fit. We could use the least sum of squared residuals for the catch rates (using lognormal errors), but for

© 2011 by Taylor & Francis Group, LLC

ease of expansion later, we will use maximum likelihood methods, as was done with the surplus production modelling in Chapter 8:

$$L\left(\text{data}|\bar{R}, H^I, q\right) = \prod_t \frac{1}{I_t \sqrt{2\pi}\hat{\sigma}} e^{\frac{-\left(\text{Ln}\,I_t - \text{Ln}\,\hat{I}_t\right)^2}{2\hat{\sigma}^2}} \tag{13.26}$$

where $L\left(\text{data}|\bar{R}, H^I, q\right)$ is the likelihood of the data given the parameters (for example, as here: average recruitment, initial harvest rate, and catchability), the product is overall years (t) for which CPUE data are available, and

$$\hat{\sigma}^2 = \sum_t \frac{\left(\text{Ln}\,I_t - \text{Ln}\,\hat{I}_t\right)^2}{n} \tag{13.27}$$

where n is the number of observations (maximum likelihood estimate of the variance uses n rather than $n-1$; Neter et al., 1996, p. 34). A simplified version of the same maximum likelihood equation is

$$LL^{CE} = -\frac{n}{2}\left(Ln(2\pi) + 2Ln(\hat{\sigma}) + 1\right) \tag{13.28}$$

(see Appendix 11.3 for its derivation). If only the three parameters shown in Equation 13.26 are used, this equation can stand as it is; however, instead of using only an average recruitment level for each year, it is possible to estimate recruitment residuals that modify the level of recruitment in each year:

$$R_t = \bar{R}.e_t^{N(0,\sigma^2)} \tag{13.29}$$

The model can estimate the lognormal recruitment residuals $e^{\varepsilon_R} = e^{N(0,\sigma^2)}$ for each year t. However, to prevent a perfect fit from overparameterization and to avoid the model generating biologically meaningless estimates, a penalty term is required as well as additional data where available. The penalty term on the recruitment residuals, P_R, includes an estimate of the expected variation of the recruitment residuals:

$$P_R = \sum_{t=1}^{n} Ln(e^{\varepsilon_R})^2 / 2\sigma_R^2 \tag{13.30}$$

where σ_R^2 is the expected variance of the recruitment residuals (Example Box 13.5). The smaller the value attributed to σ_R^2 (SigmaR), the less variable the recruitment deviates can be, and as a consequence, the fit to the catch rate time series declines (Figures 13.3 to 13.5).

© 2011 by Taylor & Francis Group, LLC

With only two parameters directly estimated, the fit of the model to the observed catch rates is not as good as was obtained with the surplus production model (compare Figure 13.3 with Figure 8.5). Obviously, by estimating the recruitment deviates (Figure 13.4), the fit of the model increased substantially, but we have also increased the number of parameters. While we haven't included any further direct data, we have included information in the form of the selectivity, maturity, and weight-at-length relationships, as well as the information in the growth transition matrix. The last recruitment deviate (for 2008) does not differ from 1.0, and that for 2007 is close to 1.0. This is because there is little information about the recruitment strengths in 2007 and 2008 in the catch rate data. It will take at least two to three years for animals to grow through the Legal Minimum Length (LML) from recruitment in sufficient numbers to influence catch rates, and so these parameters stay near 1.0. They need not be fitted, which would reduce the number of fitted parameters by two. However, to improve the balance between fitted parameters and number of observations, we could include data that the surplus production model could not include. If we had the size distributions of commercial catches, these could be compared to the predicted proportions in the different size classes that can be derived from the model (see Equation 13.18). The comparison of the relative proportional distribution of numbers across a set of size classes is best made using multinomial likelihoods:

$$LL^{LF} = \sum_{i=140}^{170} N_{t,i}^C Ln\left(\frac{\hat{N}_{t,i}^C}{\sum \hat{N}_{t,i}^C}\right) \tag{13.31}$$

where LL^{LF} is the negative log-likelihood for the length frequency data, $N_{t,i}^C$ is the observed numbers in the commercial catch in size class i at time t, and $\hat{N}_{t,i}^C$ is the predicted numbers in the commercial catch in size class i at time t. This is combined with the negative log-likelihood for the catch rate data, and its penalty term, using a weighting term, λ, that scales the two values to become approximately the same order of magnitude (in the example λ was about 0.001; we divided the sum in D9 of Example Box 13.4 by 1,000).

$$LL = LL^{CE} + \lambda LL^{LF} - PenR \tag{13.32}$$

The penalty is subtracted from the negative log-likelihood, as the aim is to maximize Equation 13.32 in order to optimize the model fit (Example Box 13.5). Observed length frequencies from the commercial catch need to be summarized into the same size classes as in the model (Table 13.1). Unfortunately, the 10 mm size classes used in the examples are rather too coarse for these data to have much effect on the model fit. This is the case because the wide size classes tend to smooth over the effects of size-selective

© 2011 by Taylor & Francis Group, LLC

EXAMPLE BOX 13.4

Initial fit of the size-structured model. Extend the years from H7 to 2008 in AE7. The catch and catch rate data in H12:AE13 are from Table 8.1 in Chapter 8. Alter the harvest rates in H17:AE17 as shown. Put =100*H10/$G$7 into H8. In H9 put =sumproduct(H33:H44,$F$19:$F$30)/1000, in H10 put =sumproduct(H33:H44,$F$19:$F$30,$C$19:$C$30)/1000 for the mature biomass, and in H11 put =sumproduct($F$19:$F$30,H33:H44,$D$19 :$D$30)/1000. Select H8:H11 and copy across to column AE. In H15 put =LN(H13/H11); the lognormal catch rate residuals are in H16:AE16. Select H14:H17 and copy across to column AE. That summarizes the dynamics of the model. In D8 put =count(H13: AE13) to get the number of years of data. In D9, for later, put =SUM(W66:AE69)/1000, while in D10 put Equation 13.28 =−(D8/2)*(Ln(2*pi()) +2*Ln (D11) +1). To calculate the sigma in D11 put =sqrt(sumxmy2(Ln(H13:AE13), Ln (H14:AE14))/D8) (look up the function help if needed). The recruitment deviate penalty is =10*sumproduct(Ln(B4:B26)* Ln(B4:B26))/(2*$I$4 *$I$4) in D12 (Equation 13.32). Finally, in D13 put =D10+D9−$D$12. The model can now be fitted by using the solver. First, maximize D13 by changing B2:B3. Plot the years against the catch rates and predicted catch rates, as well as against the residuals to obtain something similar to Figure 13.3. Plot the recruitment residuals against the years 1986–2008 to get something like Figure 13.4. Now try using the solver maximizing D13 by changing B2:B26. The plots should now look more like Figure 13.5. Once the optimum fit has been obtained set the InitialH to 0.0 and copy the mature biomass in H10 as a value into G7, then put 0.28624 back into B3. Vary the value in SigmaR; the smaller this value, the smoother the recruitment deviates become, and so the fit to the catch rates declines (Figure 13.6).

|	C	D	E	F	G	H	I
7				BzeroSp	12898	1985	1986
8	n	24		Depletion SpawnB		43.42	44.90
9	LF_LL	0.0000		Total biomass		8185.390	8206.561
10	CE_LL	70.2138		Mature biomass		5600.913	5790.924
11	sigma	0.0130		Exploitable Biomass		5098.539	5289.981
12	PenaltyR	11.2365		Catch		1018.884	742.347
13	LL	58.9773		Catch Rates		1	1.0957
14				Predicted Catch Rates		=$B$1*H11	1.0586
15				q_contrib		−8.5367	−8.4822
16				ResidualCE		=H13/H14	1.0351
17				Harvest		=H12/H11	0.1403

EXAMPLE BOX 13.5

Inclusion of predicted catch at size. Extend the years from W71 to 2008 in AE71. Enter the observed length frequency data and totals from Table 13.1 into W72:AE76. In W79 put =W33*W$17*$D19 (Equation [13.18]) and copy down to W90. In W91 put =sum(W87:W90), which sums the total count in size classes 140:170. Select W79:W91 and copy across to column AE. For completeness, select W79:W90 and copy across to column H (=1985, alter the selectivity from $E19 to $D19 from 1989 to 1985). Finally, in W66 put =(W72)*LN(W87/W$91), and copy down to W69. Select W66:W69 and copy across to column AE. This should fill the log-likelihood component for length frequency data in D9. Solving for the maximum of D13 by changing B2:B25 should generate a log-likelihood of about –2.39, with a weight of 0.001 given to the length frequency log-likelihood. In this case the addition of the length frequency data only makes a minor difference (a slight improvement in the fit; Table 13.2).

|	V	W	X	Y	Z	AA	AB
71		2000	2001	2002	2003	2004	2005
72	140	567	1870	1029	1431	590	1198
73	150	959	2941	1478	2164	1027	1448
74	160	815	2115	1259	1748	948	1053
75	170	566	832	620	715	467	489
76		2907	7758	4386	6058	3032	4188
77							
78	1999	2000	2001	2002	2003	2004	2005
79	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0
88	408965	609693	591705	559054	567949	543900	533408
89	243686	387614	398092	396101	417293	404826	395641
90	86254	140751	150633	157440	174779	177995	180887
91	Totals ~>	1751891	1714120	1645940	1713613	1670758	1662564

---

fishing; nevertheless, the principles remain the same. In an operational model many more size classes would be used, and these increase the sensitivity of the dynamics to the available data.

In 1985 the Tasmanian western zone stock was in a relatively depleted state, which led to large catch reductions. We can compare the size distribution of the stock in 1985 with that expected in an unfished population to see the effect of fishing. As might be expected, once depleted away from the unfished state, many of the larger animals are selectively removed from the stock and the size structure is greatly modified (Figure 13.6).
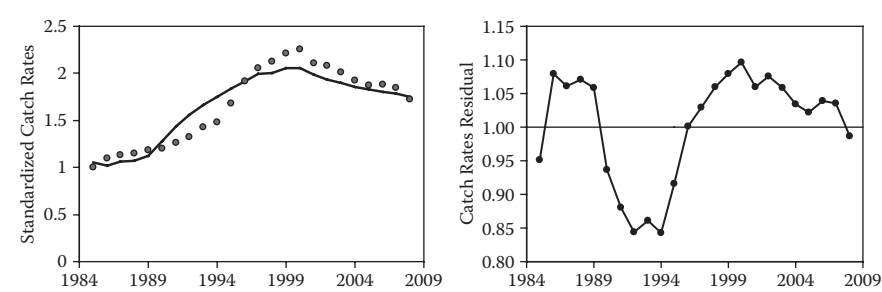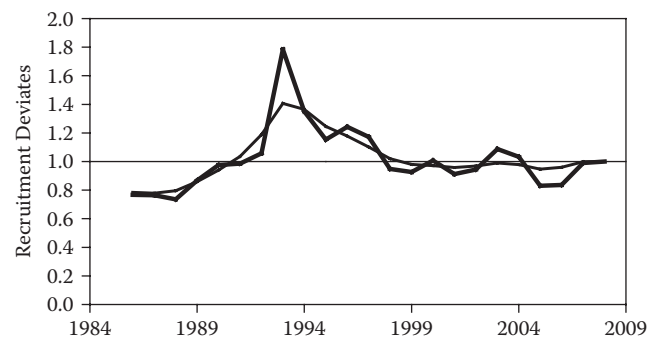
Fig ur e 13.3
The fit to observed standardized catch rates as expressed by the values themselves (left panel) and by the lognormal residuals (right panel) when only using the average recruitment and initial harvest rate parameters. This is not as good a fit as that obtained with the surplus production model in Chapter 8.
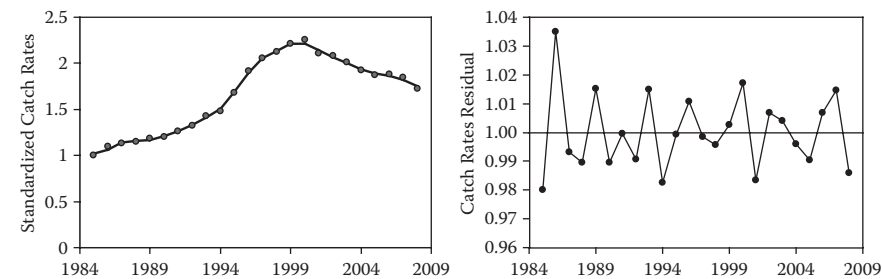
Fig ur e 13.4
The fit to observed standardized catch rates as expressed by the values themselves (left panel) and by the lognormal residuals (right panel). Not surprisingly, because of the increased number of parameters, this is rather a better fit than obtained with the surplus production model in Chapter 8.

Fig ur e 13.5
The recruitment deviates from the initial fit to catch rate data in Example Box 13.4. The thick spiky line relates to a SigmaR value of 0.6, while the finer, smoother line relates to a SigmaR of 0.3. With less variation in the recruitment deviates, the ability to fit the catch rate data is reduced.
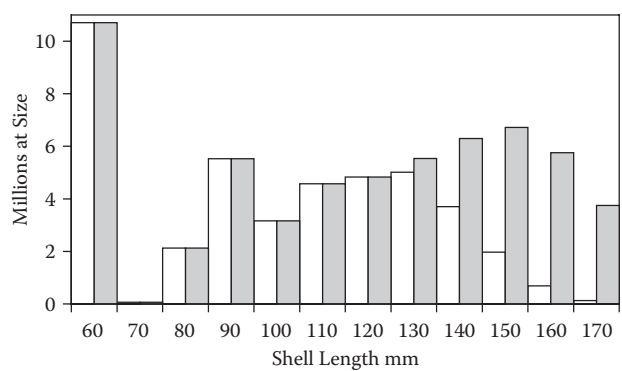
© 2011 by Taylor & Francis Group, LLC

Fig ur e 13.6
The equilibrium initial size structure of the population from Example Box 13.5. The shaded bars relate to an average recruitment set at 10,405,973 with a zero initial harvest, while the open bars relate to an initial harvest rate of 0.28738. The legal minimum length in 1985 was 132 mm; the depletion of the larger size classes away from the unfished exploitable biomass is clear.

TABLe 13.1

Observed Frequencies of Different 10 mm Length Classes in the Commercial Blacklip Abalone (*Haliotis rubra*) Catch Taken in the Western Zone of Tasmania from 2000 to 2008

Year	140	150	160	170	Total
2000	567	959	815	566	2,907
2001	1,870	2,941	2,115	832	7,758
2002	1,029	1,478	1,259	620	4,386
2003	1,431	2,164	1,748	715	6,058
2004	590	1,027	948	467	3,032
2005	1,198	1,448	1,053	489	4,188
2006	1,483	1,844	1,392	631	5,350
2007	646	882	599	214	2,341
2008	3,073	3,790	2,415	1,008	10,286

The data suggest that the Tasmanian western zone stock exhibited lower than average recruitment prior to about 1992, but then was fortunate enough to receive well above average recruitment between 1993 and 1997 (Figure 13.5), which, along with the reduced catches at that time, significantly rebuilt the stock. Fishing on the eastern coast of Tasmania is much simpler than the more exposed western coast, and from about 1989 to 2000 a large proportion of the available total allowable catch (TAC) was being taken from the east. In an attempt to distribute the effort more evenly across the resource, zonation was introduced in 2000, which significantly increased the catches expected to be taken from the western zone each year. The immediate effect of this was for catch rates to begin to decline from very high levels in the west (Figure 13.4). These changes to the fishery and management led

© 2011 by Taylor & Francis Group, LLC

TABLe 13.2

The Main Parameters and Likelihood Components of the Two Models

	Example Box 13.4	Example Box 13.5
q	0.000200	0.0002011
AvRecruit	10436714	10405978
Initial H	0.286243	0.287383
LF_LL	0	–61.3695
CE_LL	70.2138	70.17077
Sigma	0.012978	0.01300
PenaltyRec	11.23654	11.19408
Log likelihood	58.97725	–2.39281

Note: Example Box 13.4 is without the length frequency data and Example Box 13.5 includes that data. The fit to the catch rates is slightly reduced, the average recruitment is slightly decreased, and the initial harvest rate is slightly increased. The weighting given to the length frequency data was 0.001, which scales it closer to the log-likelihood for the catch rates.

to the fishery data having remarkable contrast. Initially, in the 1980s, large catches were taken out and the stock was depleted, but then the TAC was greatly reduced and most fishing moved to the east coast so that catches in the west declined markedly and the stock rebuilt, then catches in the west were again increased. Such treatment, heavy fishing followed by reduced fishing followed again by heavy fishing, can provide excellent information about how the resource responds to fishing pressure. Such high contrast is one of the reasons why the model fit is remarkably good (even the surplus production model in Chapter 8 provides an excellent fit). Of course, in a real fishery assessment, more size classes would be used and more specific details from the fishery included so that one would have more confidence in the outcomes of the modelling.

There is naturally a close relationship between the catches removed and the annual harvest rate. In 1997 the catches increased from about 450 t to about 650 t and then back down to about 450 t, and this is apparent in the resulting harvest rate (Figure 13.7). The impact of zonation and the forcing of catch back onto the west coast is also apparent.

The assessment model as it stands provides an adequate summary of what has happened in this fishery. But to make it useful for management, we need to be able to project its dynamics forward to see the expected effects of different levels of catch and the influence of any other management changes. Besides projecting the effects of different catch levels, which is simple to do (Example Box 13.6), as with the surplus production models in Chapter 11 and the age-structured models in Chapter 12, it would be possible to bootstrap the residuals to the catch rate fit and to the length frequency fit and characterize some of the uncertainty around any projections. However, capturing all of the uncertainty in these assessments would be difficult. Abalone are notoriously
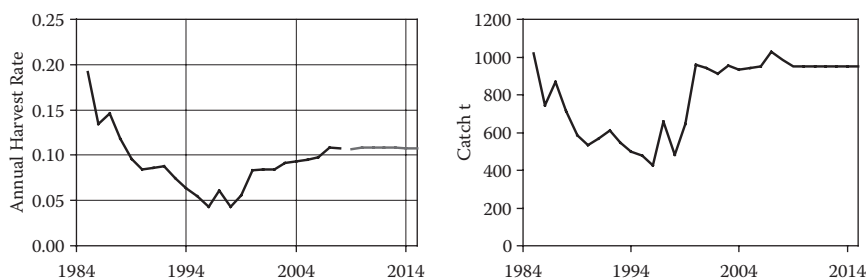
Fig ur e 13.7
The annual catch in the western zone (right panel) and the resulting harvest rate (left panel).
Note the effect of the increase in catches in 1997 on the harvest rates. The right-hand extension
of both graphs is where the model is being used for projecting given catch levels forward.

variable (and seemingly those around Tasmania especially so). There is varia-
tion in the weight-at-length relationships, the selectivity relationships, the
size-at-maturity relationships, and especially in the description of growth.
This is one occasion where it might be better to sample vectors of parameters
from the posterior distribution of a Markov chain Monte Carlo (MCMC) and
use them as the basis of projections rather than use simple bootstrapping.

13.3 Concluding Remarks

Integrated size or length-based assessment models offer the possibility of
producing detailed stock assessments for difficult to age species that could
be used in formal management systems. The most valuable fisheries in
Tasmania include rock lobster (*Jasus edwardsii*) and abalone (*Haliotis rubra*),
and the management of rock lobster is very dependent on length-based
models (Haddon and Gardner, 2009), and the abalone fishery is beginning
to use them (Haddon, 2009), though characterizing the variation around the
coast is slowing their implementation. They are also used in other abalone
and rock lobster fisheries elsewhere (Breen et al., 2003). Not surprisingly,
size-based models are very sensitive to having an adequate description of
the growth process; the growth transition matrix is a vital component. Using
abalone in the example may not have been optimal, as these species tend to
be highly variable in their biology. The new western zone is the bottom half
of the west coast of Tasmania, and this extensive area contains many popula-
tions of abalone producing about 1,000 tonnes a year. An assessment model
that attempts to capture the dynamics of this large area averages over much
of the variation present between populations within the zone. Nevertheless,
the potential value of size-based models should be apparent, as important
management questions can be answered. The expected population trajectory

EXAMPLE BOX 13.6

Projecting the size-based model. Extend the years from AE7 to 2015 in AL7 and from AF18 to AL18 (you could go further is wished). Select AE8:AE11 and copy across to column AL. Copy AH12 across to AL12. Select AE14:AE17 and copy across to column AL (delete AF15:AL16). Finally, select AE19:AE44 and copy across to column AL, and that is all that is needed to conduct deterministic projections. Plot the spawning biomass or the catch rates or harvest rates, treating the data from 1985:2008 as one data series and the projection data as another, akin to Figure 13.7. Then try altering the TAC in I5 and consider the impact; do not expect rapid effects. You could copy the row labels from F7:F17 into AM7:AM17 to make things more clear.

	AF	AG	AH	AI	AJ	AK	AL
7	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
8	69.2	69	69	69.3	69.6	69.9	70.2
9	11498	11494	11521	11561	11606	11648	11687
10	8930	8893	8905	8939	8980	9021	9058
11	8475	8427	8441	8479	8522	8564	8603
12	=I5	=AF12	=AG12	900	900	900	900
13							
14	1.7045	1.6948	1.6977	1.7053	1.714	1.7225	1.7302
15							
16							
17	0.1062	0.1068	0.1066	0.1061	0.1056	0.1051	0.1046
18	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
19	10406099	10406099	10406099	10406099	10406099	10406099	10406099
20	64973	64973	64973	64973	64973	64973	64973

under different catch regimes can be determined, or the effect of a change in legal minimum length can be identified, along with other, more site-specific questions. Age-structured models naturally follow the fate of cohorts, and the annual incrementing of ages is an automatic clock in such models. Size-based models are not so constrained by their internal dynamics, and the growth of organisms can be highly variable. The intuition is that such models are going to be less certain in their predictions. Age-based models now have a long history and their behaviour is well known. Size-based models are relatively new, and until more experience with using them in real situations has been accrued and reported, especially where the biology of the modelled species is highly variable, it would be well to be cautious with their use.