12

Age-Structured Models

12.1 Types of Models

12.1.1 introduction

Surplus production models (Chapter 11) ignore sexual, size-, and age-based differences by treating a stock as undifferentiated biomass. Even a superficial consideration of the ecological differences to be found within and between the members of a population suggest that this assumption may be leaving out important influences on the population dynamics. An obvious example would be the time delays present in the dynamics of populations that have a number of years between biological recruitment as juveniles and sexual maturity. In a favourable year, leading to a strong year class, there will be a major increase in stock biomass, but it may take a few years before that biomass starts to contribute to reproduction. By lumping growth, reproduction, and mortality into one production function, dynamic interactions between these processes are ignored.

The obvious solution is to differentiate a stock's biomass into component parts. To describe the dynamics of the population, we will still need to account for recruitment, growth, and mortality, but these processes will either have to be dealt with separately or be included in however the stock biomass is subdivided. Thus, we could generate production models of a stock in which the two sexes are differentiated. Similarly, one could generate a length-structured model of stock dynamics, as briefly discussed in Chapter 8 on growth (Section 8.3.5; Sullivan et al., 1990). The most commonly used option, however, is to subdivide the population into age classes or cohorts and follow the dynamics of each cohort separately, combining them when inputs to the dynamics, such as recruitment, and outputs, such as yield, are being considered. Of course, to utilize age-structured methods, it must be possible to age a species accurately, or at least with a known error rate. There are methods for accounting for ageing error, but ideally the ageing should be highly accurate and with low bias.

The age-structured models to be discussed in this chapter are relatively complex and have more parts than the models presented in earlier chapters.

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central,
http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

Indeed, the example boxes will need to be generated in sections to account for the model complexity. In fact, rather than using Excel to work with these models, it will usually be more efficient to adopt a programming language and write custom programs to conduct the model fitting. If a modeller wishes to utilize such complex models, then eventually the need for expertise in a programming language becomes essential. Which language is used is not an issue, as there are active modellers using Pascal, Fortran, C, C++, Visual Basic, APL, and others. Fortran used to be the language of choice among fishery modellers (the new Visual Fortran 90 and 95 is a nicely versatile and very fast language), but C++ is now gaining headway. Personally, I find C++ to be more of a computer programmer's language requiring a higher level of programming skill than that needed in, say, Pascal or Fortran. However, as with all software, the choice is up to particular users, and generally all languages permit the necessary speed and complexity.

Despite this need for custom programs we will continue to use Excel in the examples in this chapter. We will only be considering relatively simple age-structured models, but even so in one of the models we will attempt to estimate twenty-nine parameters. This chapter only aims to introduce some of the more important ideas behind age-structured models. It does not investigate all of their intricacies and will only attempt to describe a fraction of the range of possibilities. Detailed reviews of age-structured models, such as those by Megrey (1989) and Quinn and Deriso (1999), provide a great deal more information and detail about the complexities and other developments possible with this class of models. To treat these models with the same detail as that given to the simpler models in this book would require a much larger book.

In this chapter we will first generalize the standard catch curve (Beverton and Holt, 1957) into a dynamic system of equations simulating the dynamics of an age-structured population. We will then provide a very brief introduction to cohort analysis (Gulland, 1965, cited in Megrey, 1989; Pope, 1972), followed by an introduction to statistical catch-at-age methods (Doubleday, 1976; Methot, 1989, 1990). In this latter section we will also consider the algorithms necessary to conduct bootstrap estimates of uncertainty around parameter estimates and model outputs. There have been developments in the public provision of generalized catch-at-age software (SS3; originating from Methot, 1990, now based in AD-Model Builder), and any serious user should consider this option (Methot, 2009).

We have already considered some of the fundamentals of the dynamics of cohorts in the coverage of yield-per-recruit in Chapter 2 and in the Monte Carlo simulation of catch curves in Chapter 7. Following the dynamics of each cohort has the advantage that after each cohort has recruited, assuming there is no immigration or emigration, the numbers present in the population can only decline. Age-structured models are founded on the basis that a careful examination of this decline can provide information on the total mortality being experienced. If there is an estimate of the natural mortality

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

TABLE 12.1

Number of North Sea Plaice in Each of Nine Age Classes, Landed at Lowestoft per 100 Hours Fishing by British First-Class Steam Trawlers from 1929 to 1938

Year/Age	2+	3+	4+	5+	6+	7+	8+	9+	10+	Effort
29/30	328	2,120	2,783	1,128	370	768	237	112	48	5.81
30/31	223	2,246	1,938	1,620	302	106	181	58	18	5.84
31/32	95	2,898	3,017	1,150	591	116	100	82	33	4.97
32/33	77	606	4,385	1,186	231	138	42	21	51	4.91
33/34	50	489	1,121	4,738	456	106	80	27	18	5.19
34/35	44	475	1,666	1,538	2510	160	50	43	14	4.94
35/36	131	1,373	1,595	1,587	1,326	883	144	30	28	4.63
36/37	38	691	2,862	1,094	864	382	436	27	15	4.44
37/38	138	1,293	1,804	1,810	426	390	163	228	26	4.39

Source: Data from Table 13.1 in Beverton, R. J. H., and Holt, S. J., *U.K. Ministry of Agriculture and Fisheries, Fisheries Investigations* (Series 2), 19, 1–533, 1957.

Note: The fishing year is April 1 to March 31. The + symbol after the age indicates that a fish of t+ years is in its $t + 1^{\text{th}}$ year, somewhere between t and $t + 1$ years old. Note there are relatively strong year classes, such as the one highlighted, which would have arisen in 1928–1929. Despite being standardized to the same levels of effort the 5+ group in 33/34 is larger than the 4+ group in 32/33. This suggests that the early ages are not fully selected (Figure 12.1). Effort is in millions of steam trawler hours fishing (from Table 14.15 in Beverton and Holt, 1957).

then, clearly, deductions can be made concerning the exploitation rates of each cohort.

As with surplus production models, we need to have data on the total catch (fishing mortality can include discarding mortality) as weight, but we also need data on the numbers-at-age in the catch. Ideally, for each year of the fishery, there will be an estimate of the relative numbers caught in each age class (e.g., Table 12.1). Effort information or some index of relative abundance is also required to obtain an optimum fit in all of the models. Beyond this minimum there are many other forms of information that can be included in such stock assessment models. Such information might include estimates of annual recruitment, stock biomass estimates, catch rates, and the mean length and weight of animals in the catch.

A commonly observed phenomenon in age-structured data is the progression of year classes (Table 12.1, Figure 12.1). This also provides evidence that the ageing of the animals concerned is at least consistent through time, and also that the ring counts used to age the animals actually relate to yearly increments. Note also that the absolute numbers and proportions caught in each year do not necessarily decline through time as one might expect. If the availability of each age class changes from year to year, for example, for reasons of selectivity, or through different fishing gear being used, or different people doing the fishing, then a simple progression of the cohorts will not be observed. The North Sea plaice data (Table 12.1) are standardized to a constant amount of the same kind of effort so, after the species is fully selected

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

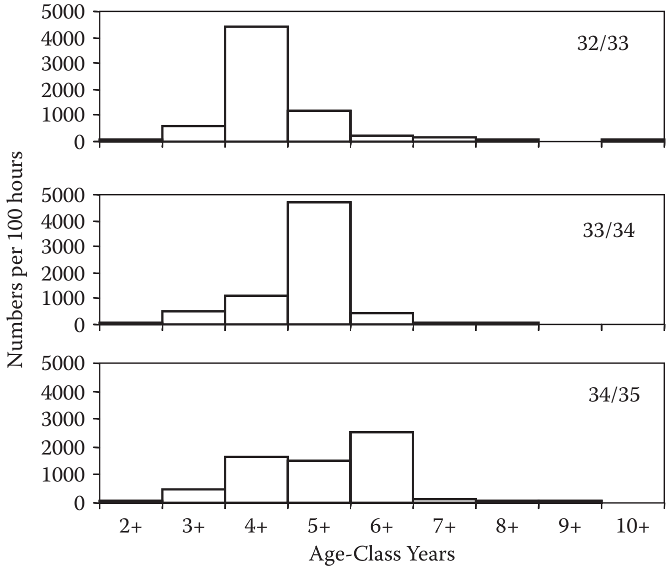

Fig ur e 12.1
Progression of a relatively strong year class (recruited into the population as 0+ in 1928–1929) through the population of North Sea plaice over the three years from 1932–1933 to 1934–1935 (Table 12.1). The fishery is primarily imposed upon three- to seven-year-old fish. (Data from Beverton, 1957.)

by the gear (age 5+ and older), the numbers always decline from one year and age to the next (Figure 12.1).

Data on catch and catch-at-age will provide information regarding the population dynamics. However, for stock assessment purposes, some index of relative abundance through time is required to strengthen the attachment of the model to changes in stock size through time or to provide robust estimates of fishing mortality. A suitable index could use catch rates or the effort imposed to obtain the standardized catches, or even fishery-independent survey estimates. In summary, the minimum data requirements are the commercial catches, the catch-at-age as numbers, and effort or some index of relative abundance through time.

12.1.2 Age-Structured Population Dynamics

As demonstrated in Chapters 2 and 7, if there is no emigration or immigration, then, after recruitment, the numbers in any cohort will decline exponentially through time with a rate of decline equal to the instantaneous total mortality rate:

$$N_y = N_0 e^{-Zy} \tag{12.1}$$

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

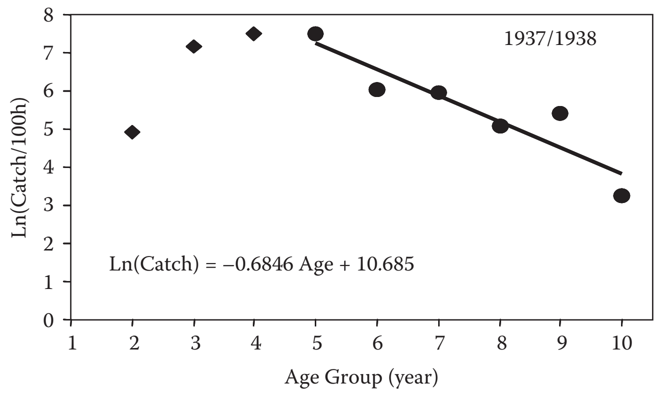

Figure 12.2
Natural logarithm of the average number of North Sea plaice in each age group caught per one hundred hours fishing by a first-class steam trawler in 1937–1938. (Data from Beverton and Holt, 1957.) The negative gradient of the slope estimates the total mortality $Z = 0.685$. Age groups 2 to 4 are omitted from the regression because they are not fully selected by the fishing gear. The elevated value in the ninth age group derives from the strong year class that arose in 1928–1929 (Table 12.1).

where N_y is the numbers in year y, N_0 is the initial recruitment into the cohort, and Z is the instantaneous rate of total mortality (M is natural mortality and F is fishing mortality; $Z = M + F$). A log transformation leads to the familiar

$$\text{Ln}\left(N_y\right) = \text{Ln}\left(N_0\right) - Zy \tag{12.2}$$

which has the form of a linear relationship and is the basis of the classical catch curve (Figure 12.2; Beverton and Holt, 1957). There are two kinds of catch curves possible. The first samples all age classes present in a particular year, assumes the system is in equilibrium, and treats the proportions of the different age classes as if they were the product of a single cohort (Figure 12.2). The major assumptions here are that there has been a constant recruitment level in all years, that all ages have been exposed to the same history of fishing mortality (fishing mortality is constant across years), and that after age 4+ all animals are fully selected and have the same catchability. One problem with using this snapshot approach is that standardizing the data collected to a constant amount of effort using a particular gear does not necessarily remove all occasions where an age class is more numerous than the previous age class (e.g., age groups 8 and 9 in Figure 12.2). This is a combination of a failure of all or some of the assumptions of this approach to catch curves and also may reflect the impact of sampling error. The assumption of constant recruitment is particularly unlikely, as demonstrated by the strong year classes evident in Table 12.1.

The second kind of catch curve follows the fate of single cohorts (Figure 12.3). This approach is more difficult because it relies on standardizing the catch

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

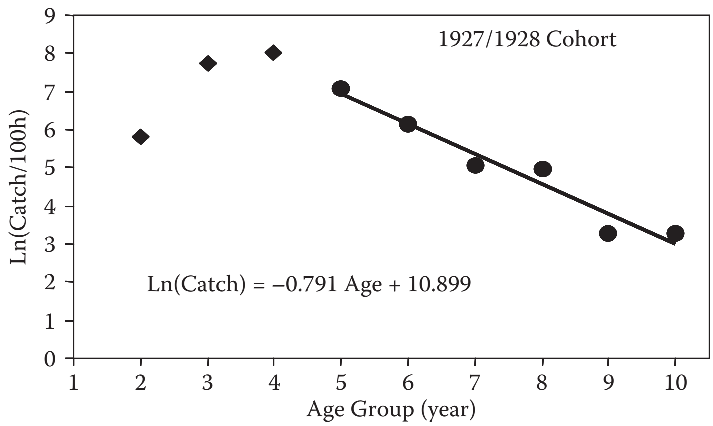

Fig ur e 12.3
A catch curve following the average number of North Sea plaice from a single cohort caught per one hundred hours fishing by a first-class steam trawler over the years 1929 to 1938 (Table 12.1). (Data from Beverton and Holt, 1957.) The negative gradient of the slope estimates the total mortality $Z = 0.791$. Age groups 2 to 4 are omitted from the regression because they are not fully selected by the fishing gear (despite, in this instance, age 4 appearing to be selected as well as later ages).

of the cohort being followed to a given amount of effort with a particular type of fishing gear. This is necessary because if catchability varied between years, then cohort numbers would not always decline in the steady fashion required by the catch curve methodology. Also, of course, many more years of comparable sampling are required to generate a single catch curve.

In Chapter 7, a Monte Carlo simulation of an age-structured population was produced that included variable recruitment and sampling error (Section 7.4.3; Example Boxes 7.6 and 7.7). This used the approach of sampling the whole population in any one year and treating the combination of many cohorts as a single pseudocohort to be analyzed.

The next step would be to extend Equation 12.1, which deals solely with ages within a given year, to add sequential years to the system. Obviously, the system has to start somewhere, and there are two alternatives commonly used when conducting stock assessments of age-structured populations.

The first is to assume the modelling starts at the beginning of exploitation and to ascribe an equilibrium age structure to the starting population. The second is to directly estimate the starting numbers-at-age in the model (this has the obvious disadvantage of adding the same number of parameters as there are age classes to the estimation problem). At equilibrium, in an unexploited population, the age structure would be the result of natural mortality acting alone upon the average virgin levels of recruitment. Thus, the relative numbers in each age class in the initial year would be equivalent to the single-year catch curve where all of the assumptions have been met:

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

$$N_{a,1} = \begin{cases} R_{0,1} & a = 0 \\ N_{a-1,1}e^{-M} & 1 \le a \le a_{\max} - 1 \\ N_{a_{\max}-1,1}e^{-M} / \left(1 - e^{-M}\right) & a = a_{\max} \end{cases} \tag{12.3}$$

where $N_{a,1}$ is the numbers of age a, in year 1, $a_{\max}$ is the maximum age modelled (the plus group), and M is, as usual, the instantaneous rate of natural mortality. Recruitment variability has been omitted from Equation 12.3, but this will be addressed later. In a preexploitation population there is no fishing mortality. The final component of Equation 12.3, where $a = a_{\max}$, is referred to as the plus group because it combines ages $a_{\max}$ and all older ages that are not modelled explicitly. The inclusion of the $(1 - e^{-M})$ divisor forces the equation to be the sum of an exponential series (Example Box 12.1).

After initial conditions are defined (e.g., Example Box 12.1), the population will grow approximately in accord with some stock recruitment relationship and the mortality imposed naturally and via fishing pressure. In yearly steps, the numbers in each age class in each year will depend on the numbers surviving from the preceding age class from the preceding year:

$$N_{a+1,y+1} = N_{a,y}e^{-\left(M + s_a F_y\right)} \tag{12.4}$$

which is the numbers at age a in year y, multiplied by the survivorship (e^{-Z}) after natural mortality M, and $s_a F_y$, the selectivity for each age a times the fully selected fishing mortality in year y. This is equivalent to the catch curve equation and could be transformed into a linear format. Not all age classes are necessarily fully selected; thus, the fishing mortality term must be multiplied by the selectivity associated with the fishing gear for age a, s_a. We use a lowercase s for selectivity to leave the uppercase S for the survivorship term. We also need a term for the recruitment in each new year, and this is assumed to be a function of the spawning biomass of the stock in the previous year y, B^s_y (Example Box 12.2); thus,

$$N_{a,y+1} = \begin{cases} f\left(B^s_y\right) = R_{0,y+1} & a = 0 \\ N_{a-1,y}e^{-\left(M + s_{a-1}F_y\right)} & 1 \le a \le a_{\max} - 1 \\ N_{a_{\max}-1,y}e^{-\left(M + s_{a-1}F_y\right)} + N_{a_{\max},y}e^{-\left(M + s_a F_y\right)} & a = a_{\max} \end{cases} \tag{12.5}$$

where the plus group is modelled by applying the total mortality to the preceding age class and the maximum age class in the preceding year. These two cannot be combined prior to multiplication by the survivorship term if the selectivities of the two age classes are different (Example Box 12.3).

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.1

An equilibrium age-structured simulation model illustrating the initial age structure (using Equation 12.3). The selectivities in row 11 are constants, but an equation could be used. In B9 put =$E$1*(1–exp(–$E$2*(B12–$E$3))) and in B10 put =$E$4*B9^$E$5. Select B9:B10 and copy across to column K. Copy C13 across to column J, and in K13 put =J13*exp(–M)/(1–exp(–M)) to obtain the plus group. Copy C7 across to column K and modify K7 to be =J7*exp(–M)/(1–exp(–M)) to obtain the per-recruit numbers at age. Copy C8 into B8 and into D8:K8 to obtain the biomass per recruit, needed for the estimation of the stock recruitment relationship. To test the plus-group calculation, record the value in K13 (133) and extend the ages across to age 25 in column AA; then copy J13 across to column AA and in K6 put =sum(K13:AA13). This summation should also equal 133. B13 is the first term, C13:J13 is the second, and K13 is the third term from Equation 12.3 (see Example Boxes 12.2 and 12.3 to fill column O).

	A	B	C	D	E	~	J	K
1	**Steepness**	0.75		**Linf**	200			
2	**Nat Mort**	0.4		**K**	0.45			
3	**Bzero**	1000		**t0**	–0.02			
4	**a Recruit**	=1/O2		**a_Wt**	0.0002			
5	**b Recruit**	=O1/O2		**b_Wt**	3			
6	**R0**	=O3		**q**	0.0005			
7	**Rec_Nage**	1	=B7*exp(–$B$2)	0.4493	0.3012	~	0.0408	0.0829
8	**Rec_Biom**	1E–06	=C7*C10/1000	0.153	0.1977	~	0.0601	0.1259
9	**Length**	1.8	73.6	119.4	148.6	~	194.6	196.5
10	**Weight**	0.0012	79.793	340.57	656.49	~	1473.5	1518.5
11	**Selectivity**	0	1	1	1	~	1	1
12	**Year/Age**	0	1	2	3	~	8	9
13	**1967**	=B6	=B13*exp(–$B$2)	720	482	~	65	133

Compare Equation 12.5 with Equation 12.3 to see the impact of including fishing mortality.

The spawning biomass is defined as

$$B_y^s = \sum_{a=c}^{a_{\max}} w_a N_{a,y} \tag{12.6}$$

where w_a is the average weight of an animal of age a, c is the age of sexual maturity, and $a_{\max}$ is the maximum age class. The weight can be determined

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.2

The derivation of the recruitment parameters from the growth and mortality rates. This section of the worksheet should be added to that given in Example Box 12.1. The explanations behind these equations and their relationships are given in Section 10.9.3 and Example Boxes 10.7 and 10.8. The alpha and beta in column O are used to calculate the Beverton and Holt stock recruitment parameters in B4:B6.

	M	N	O
1	**Alpha**	=N4*(1–B1)/(4*B1*N3)	=O4*(1–B1)/(4*B1*O3)
2	**Beta**	=(5*B1–1)/(4*B1*N3)	=(5*B1–1)/(4*B1*O3)
3	**R0**	=N4/N5	=O4/O5
4	**B0**	=N5	=B3
5	**A0**	=sum(E8:K8)*exp(–B2)	=N6
6	**A0/Rec**	=N5/B7	

empirically by market measuring or from a growth curve derived independently of any model fitting (Example Box 12.1)

$$w_a = a\left[L_\infty\left(1 - e^{-K(a-a_0)}\right)\right]^b \tag{12.7}$$

where the constants a and b alter the von Bertalanffy growth curve into a curve relating to mass (see Chapter 9). The fully selected fishing mortality rate can be determined if there is an estimate of the catchability coefficient and measures of the relative effort imposed each year:

$$\hat{F}_y = \hat{q}E_y e^{\varepsilon} \tag{12.8}$$

These estimated fishing mortality rates could be used when fitting the model.

One way of defining the recruitment terms was described in Chapter 10 on recruitment (Section 10.9.3 and Example Boxes 10.7 and 10.8; Example Box 12.2).

12.1.3 Fitting Age-Structured Models

The series of equations represented in Example Boxes 12.1 to 12.3 provide for a relatively simple simulation of an age structure population starting from equilibrium. The end product is a matrix of numbers-at-age for the population after it has been exposed to natural and fishing mortality. Many of the features used to illustrate the simulation of an age-structured population are also used when the objective is to fit an age-structured model to observations

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.3

Extension of Example Boxes 12.1 and 12.2. The years extend down to 1979 in row 25. Effort is arbitrary values between 50 and 1,200. Row 13 is explained in Example Box 12.1. To generate the spawning biomass from each year, put =sumproduct(C13:K13,$C$10:$K$10)/1000 into L13 and copy down to row 25. Put =$E$6*O14 into N14 and copy down to row 25. To reflect Equation 12.5, generate the expected recruitment from the spawning biomass and put =$B$4*$L13/($B$5+$L13)*loginv(rand(), 0,M14) into B14. Into C14, put =B13*exp(–($B$2+B$11*$N14)) and copy across to J14. Finally, into K14 put =J13*exp(–($B$2+J$11*$N14))+K13*exp(–($B$2+K$11*$N14)) to generate the plus group. Select B14:K14 and copy down to row 25 to complete the matrix of numbers at age. The sig_r introduces random variation into the stock recruitment relationship. To remove its effects make the values in column M very small; this, approximately, enables the deterministic behaviour to be exhibited. If you set all the effort values to zero, only natural mortality will occur and the population will stay in equilibrium except for the effects of recruitment variability on the lower half of the matrix. If you set sig_r to 0.000001 and have a constant effort, then the whole population will attain an equilibrium. By plotting each year's numbers as a series of histograms vertically above each other (as in Figure 12.1), relatively strong year classes should be visible through time (press F9).

	A	B	C	D	~	J	K	L	M	N	O
12	**Yr\Age**	**0**	**1**	**2**	**~**	**8**	**9**	**Bs**	**sig_r**	**F**	**Effort**
13	**1967**	=B6	1074	720	~	65	133	1822.6			
14	**1968**	2118	1074	720	~	40	81	1235.6	0.5	0.5	1000
15	**1969**	1914	1420	720	~	25	51	1014.7	0.5	0.45	900
16	**1970**	2380	1283	952	~	13	25	847.0	0.5	0.7	1400
17	**1971**	2078	1595	860	~	11	22	1080.8	0.5	0.15	300
18	**1972**	2516	1393	1069	~	10	19	1315.8	0.5	0.125	250

made on a fishery. As stated before, the minimum data required comprise the relative catch-at-age for a number of years of the fishery, plus some estimate of effort imposed upon the fishery through time or some index of relative abundance. The aim, when fitting the model, will be to attempt to back-calculate a matrix of numbers-at-age that would have given rise to the observed catches, given the imposed fishing effort (or catch rate) and estimates of the catchability coefficient and the selectivity coefficients. We will introduce the two main strategies used for conducting such analyses.

Virtual population analysis (VPA) was the analytical strategy developed first, and this relies on the idea that if one has records of the catch-at-age

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

of a set of cohorts until the cohorts all die off, then one should be able to literally back-calculate what the numbers-at-age must have been. In this way the numbers in the population can be projected backwards until estimates are obtained of the original recruitments. This requires an estimate of the natural mortality and estimates of the fishing mortality by age and year.

The data requirements for a VPA are fairly stringent in that there can be no years of missing information if the calculations are to continue uninterrupted. VPA models are sometimes referred to as cohort analysis, although Megrey (1989) indicates some differences. The number of parameters estimated equals the number of data points available, so neither approach is fitted to fisheries data using an objective function, such as by minimizing a sum of squared residuals or a negative log-likelihood. Instead, solutions to the model equations are determined through iterative, analytical methods. There are many alternative VPA fisheries models (Megrey, 1989), but we will only consider one of the simplest.

The second analytical strategy appears to be the reverse of the first and goes under a number of names. Methot (1989) termed this approach the synthetic analysis, but others have referred to this analytical strategy as integrated analysis (Punt et al., 2001). Given equations describing the population dynamics, such that if we know $S_{a,y}$, the survivorship of age a to age $a + 1$ from year y to $y + 1$, and that $N_{a+1,y+1} = S_{a,y}N_{a,y}$, then, given knowledge of the initial population in year 1 and of the recruitments in each subsequent year, the numbers-at-age for the full population can be calculated. This approach requires the model to be fitted to data using some form of minimization routine. The parameter estimates include the initial population age structure, the recruitment levels in each year, the selectivity by age class, and often other parameters as well. It is common to estimate tens of parameters, and models exist that estimate hundreds of parameters. Of course, the number of parameters that can be estimated efficiently in any model will be at least partly determined by the number of independent data points available. The point is, however, that although integrated analyses share some equations with cohort analysis, the methods differ fundamentally from each other (Figure 12.4). A knowledge of both methodologies is helpful in fisheries stock assessment.

12.2 Cohort Analysis

12.2.1 introduction

Given that we are dealing with cohorts, the change in the numbers in each cohort can be derived each year from the numbers at age a at the start of year y, $N_{a,y}$, and the survivorship during that year, as in Equation 12.4 (where

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

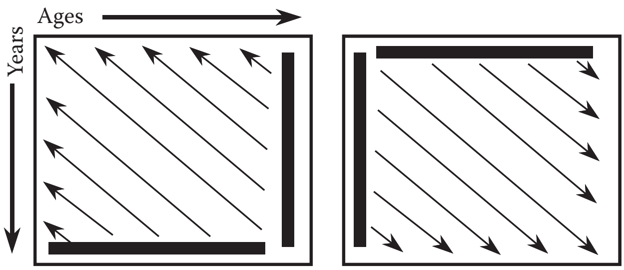

Fig ur e 12.4
Alternative analytical strategies adopted in efforts to model commercial fisheries using age-structured information. The general objective is to estimate a vector of numbers-at-age or fishing mortalities-at-age for each year of the fishery. The left-hand panel relates to cohort analysis/VPA, while the right-hand panel reflects integrated analysis or statistical catch-at-age. In cohort analysis, the calculations proceed given knowledge of all ages in the last year and the last age class in all years, and projecting backwards through time and ages. In statistical catch-at-age, knowledge is assumed of all ages in the first year and the first age class in all years (recruitment), projecting the cohorts forward through time and ages.

M and *F* are the instantaneous rates of natural and fishing mortality, and s_a is the age-specific selectivity of the fishing gear used):

$$N_{a+1,y+1} = N_{a,y}e^{-\left(M+s_aF_y\right)} = N_{a,y}e^{-M}e^{-s_aF_y} \tag{12.9}$$

Unfortunately, the number of fish present in early years is unknown, so forward projections cannot be made. However, there will be an age at which the number of fish remaining in the cohort is effectively zero. We can rearrange Equation 12.9 to start at a known number of fish in the oldest age class and back-project the population until we reach the age of recruitment (which need not be age 0). Thus, if we have an estimate of natural mortality and an age class in which the number of survivors is trivially small, then the fishing mortality for each age in each year can be estimated by back-calculating from the catches, C_y, and the natural mortality; this implies that there are as many parameter estimates as there are data points. Given

$$N_{a,y} - C_{a,y} = N_{a,y}e^{-s_aF_y} \tag{12.10}$$

rearrange Equation 12.9 to give

$$N_{a,y} = \frac{N_{a+1,y+1}}{e^{-M}} + C_{a,y} \tag{12.11}$$

While Equation 12.11 provides a mechanism for calculating the relative numbers at age in each year, it is limited to complete cohorts (those that reach the maximum age class considered; Figure 12.5).

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

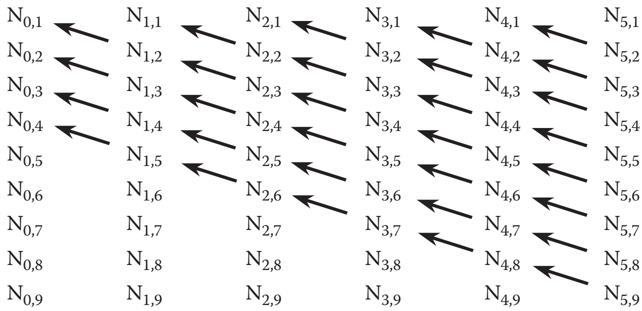

Fig ur e 12.5
A table of numbers-at-age in each year showing the backward progression of cohorts from their last age class. The columns represent age classes (first subscript), while the rows represent nine years of data (second subscript). The arrows represent the direction of calculation using Equation 12.11. Only completed cohorts, which have attained an age at which negligible animals remain, can be back-projected validly. This means that we are most uncertain about the most recent cohorts, the ones of most current interest.

One of the ways in which the many variants of VPA/cohort analysis differ is in how they address this problem of incomplete cohorts (Megrey, 1989; in Edwards and Megrey, 1989). Given a_{MAX} age classes, in the last year of data there will be $a_{\mathrm{MAX}} - 1$ incomplete year classes; these will be along the bottom of the numbers at age matrix but will not have reached the maximum age. Some way of estimating the fishing mortality rate in these age classes in the final year of data is required to complete the table of estimates of numbers-at-age in each year. These fishing mortality rates are referred to descriptively as terminal *F* estimates.

The data available include the catch-at-age and effort data. There is more than one way to fit a cohort analysis model (referred to by Megrey, 1989, as sequential population analysis (SPA)) to these data. We can either generate a fit directly to the fishing mortality rates for each age in each year, or fit to the numbers-at-age in each year. It is an odd fact that some of the seminal papers in fisheries modelling are published in obscure places (e.g., Gulland, 1965, cited in Megrey, 1989). Fortunately, alternative listings of these developments are available (Gulland, 1983; Megrey, 1989; Hilborn and Walters, 1992).

12.2.2 The equations

The basic equation relating the numbers in a cohort in one year to those in the previous year is

$$N_{y+1} = N_y e^{-(M+F_y)} \tag{12.12}$$

where N_y is the population size in year *y*, and *M* and *F* are the instantaneous rates of natural and fishing mortality, respectively. This implies that $e^{-(M+F)}$ is the survivorship, the proportion of a population that survives from year to

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

year. The survivorship itself can be determined simply as the ratio of numbers from one year to those in the year previous (starting from the last age class whose numbers are known from the catch data):

$$\frac{N_{y+1}}{N_y} = e^{-(M+F_y)} = e^{-M}e^{-F_y} \tag{12.13}$$

The complement of survivorship would be the total loss from year to year, and this can be represented a number of ways:

$$N_y - N_{y+1} = N_y - N_y e^{-(M+F_y)} = N_y\left(1 - e^{-(M+F_y)}\right) \tag{12.14}$$

The catch each year would be the proportion of this total loss due to fishing:

$$C_y = \frac{F_y}{M+F_y} N_y\left(1-e^{-Z_y}\right) = \frac{F_y}{M+F_y}\left[N_y - N_{y+1}\right] \tag{12.15}$$

which is the total loss by the proportion of mortality due to fishing. It would be possible to remove explicit mention of F_y from Equation 12.15 if we solved Equation 12.13 for F_y:

$$\frac{1}{e^{-F_y}} = \frac{N_y}{N_{y+1}} e^{-M} \tag{12.16}$$

remembering that $1/e^{-F}{}_y = e^{F}{}_y$, we can log-transform Equation 12.16 to give F_y:

$$F_y = \text{Ln}\left(\frac{N_y}{N_{y+1}}\right) - M \tag{12.17}$$

This puts us into a position of being able to estimate the catch-at-age from knowledge of the catches, the final N_{y+1}, and a given value for M, and these estimates can be used to fit the model to the data:

$$C_y = \frac{\text{Ln}\left(N_y/N_{y+1}\right) - M}{\text{Ln}\left(N_y/N_{y+1}\right) - M + M}\left[N_y - N_{y+1}\right] \tag{12.18}$$

Equation 12.18 simplifies to

$$C_y = \left(1 - \frac{M}{\text{Ln}\left(N_y/N_{y+1}\right)}\right)\left[N_y - N_{y+1}\right] \tag{12.19}$$

or

$$0 = \left(1 - \frac{M}{\text{Ln}\left(N_y/N_{y+1}\right)}\right)\left[N_y - N_{y+1}\right] - C_y \tag{12.20}$$

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

There is no direct analytical solution to Equation 12.20, but there are two ways in which the equation may be solved to produce a matrix of numbers-at-age that would balance Equation 12.20.

12.2.3 Pope's and MacCall's Approximate Solutions

An approximate solution to Equation 12.20 can be used to give the required matrix of numbers-at-age. Pope (1972) produced the following approximation (note that the exponents are positive):

$$N_y = N_{y+1}e^M + C_y e^{M/2} \tag{12.21}$$

which derives from Equation 12.12:

$$\frac{N_{y+1}}{e^{-F}e^{-M}} = N_y = N_{y+1}e^M e^F \tag{12.22}$$

Pope's (1972) advance was to introduce the discrete approximation of using the addition of $C_y e^{M/2}$ to be equivalent in effect to the multiplication by e^F. Being discrete changes the multiplication in Equation 12.22 into an addition but assumes that all fishing occurs instantaneously in the middle of the year (hence the $e^{M/2}$ to account for natural mortality acting before the fishery operated). Pope (1972) showed that his approximation was usable with values of M up to 0.3 and F of 1.2 over the time periods used in the model. Thus, if M or F is greater than these limits, cohort analysis can still be used if the catch-at-age data can be divided into shorter intervals than one year. MacCall (1986) provided an alternative approximate solution that was an improvement over Pope's (1972) equation:

$$N_y = N_{y+1}e^M + C_y\left(\frac{M}{1-e^{-M}}\right) \tag{12.23}$$

MacCall's equation behaves rather better at higher values of M and is also less sensitive to the assumption that all the catch is taken halfway through the fishing year (Example Box 12.4).

It is no longer necessary to use the approximations by Pope (1972) and MacCall (1986), but they can still be used as the starting point for the second approach that can be used to solve Equation 12.20.

12.2.4 Newton's Method

Classical analytical methods can be used to find values of N_y and N_{y+1} for each cohort that will solve Equation 12.20 so that it approximates to zero. Newton's method provides a simple and powerful method for finding the roots of a function of the form $f(N_y) = 0$ (Jeffrey, 1969). This is an iterative

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.4

Two alternative approximate calculations of the matrix of numbers-at-age (Equations 12.21 and 12.23) by Pope (1972) and MacCall (1986). This worksheet will become much larger, so care is needed in its construction. Extend the column labels in row 5 across to column J with 4+, 5+, …, 10+. Fill in the catch-at-age data from Table 12.1. Copy the column and row labels across into L5:U5 and L6:L14, respectively. Name cell J17 as *M*, the natural mortality rate, or use $J$17 wherever we need *M*. In U6 put =V7*exp(M)+J6*exp(M/2) and copy it down to U14 to obtain Pope's approximation (column V is deliberately empty). Copy U6:U14 back across to column M to create the matrix of values. Select L5:U14 and copy the contents, select A19 and paste as values (edit/paste special/values; <Alt>ESV). Fishing mortalities and other statistics may be calculated from this matrix. This new matrix, denoting an approximation to the numbers-at-age, in B20:J28, will contribute to the extension of this worksheet in Example Box 12.5, as will I15:J16. In A18 put the text "Numbers-at-Age." To compare the results with those obtained from MacCall's (1986) approximation, create another table with row labels in W5:W14 and column labels in X5:AF5. Column V must be left clear or else the calculations in column U will receive interference (trace the precedents of U6). In AF6 put =AG7*exp(M)+J6*(M/(1–exp(–M))) and copy down to AF14. Copy AF6:AF14 across to column X to generate the matrix. Compare this with the results from Pope's (1972) approximation. Notice that the values from MacCall's method tend to be slightly smaller than those from Pope's method. The absolute values are of less interest than the relative proportions by age, and while these still differ the differences are less marked. Consider the values in row 14 from column M to T. They also have no preceding data (no N_{y+1}) and yet are not the end of a cohort. Is this important?

	A	**B**	**C**	**~**	**J**	**~**	**L**	**M**	**~**	**R**	**S**	**T**	**U**
4	**Catch-at-Age**						**Approximation to Numbers-at-Age**						
5	Year\Age	2+	3+	~	10+	~	Y/A	2+	~	7+	8+	9+	10+
6	29/30	328	2120	~	48	~	29/30	15283.6	~	1409.2	414.6	153.1	54.4
7	30/31	223	2246	~	18	~	30/31	38952.0	~	348.0	419.7	113.7	20.4
8	31/32	95	2898	~	33	~	31/32	13196.1	~	276.6	177.5	167.1	37.4
9	32/33	77	606	~	51	~	32/33	11922.0	~	420.3	113.0	50.0	57.8
10	33/34	50	489	~	18	~	33/34	7664.2	~	284.9	205.5	51.0	20.4
11	34/35	44	475	~	14	~	34/35	11736.4	~	503.6	128.3	89.5	15.9

continued

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.4 (continued)

	A	B	C	~	J	~	L	M	~	R	S	T	U
12	35/36	131	1373	~	28	~	35/36	4524.2	~	2060.9	251.0	55.8	31.7
13	36/37	38	691	~	15	~	36/37	1924.4	~	670.0	825.8	68.4	17.0
14	37/38	138	1293	~	26	~	37/38	156.4	~	441.9	184.7	258.4	29.5
15	Use Terminal F												
16			Limit										
17			M		0.25								

method in which the solution is modified by the ratio of the function and the first differential of the function:

$$N_y^{Updated} = N_y^{Orig} - \frac{f\left(N_y^{Orig}\right)}{f'\left(N_y^{Orig}\right)} \tag{12.24}$$

Starting from the first iteration, N_y^{Orig} would be the individual elements of the numbers-at-age matrix derived from one of the approximations listed earlier. The modifier is made up of $f(N_y^{Orig})$, which would be Equation 12.20 for each particular age and year being considered. This means that $f'(N_y^{Orig})$ is the differential of Equation 12.20 with respect to N_y:

$$f'\left(N_y^{Orig}\right) = 1 - \frac{M}{\mathrm{Ln}\left(N_y/N_{y+1}\right)} + \frac{\left(1 - N_{y+1}/N_y\right)M}{\left(\mathrm{Ln}\left(N_y/N_{y+1}\right)\right)^2} \tag{12.25}$$

There are no N_{y+1} when the N_y in the last age class are being considered (e.g., in the North Sea plaice example there is no 11+ age class), so Equation 12.25 would fail if applied without modification. Thus, when N_{y+1} is zero one can use

$$f\left(N_y^{Orig}\right) = \left(1 - \frac{M}{\mathrm{Ln}\left(N_y\right)}\right)N_y - C_y = 0 \tag{12.26}$$

and

$$f'\left(N_y^{Orig}\right) = 1 - \frac{M}{\mathrm{Ln}\left(N_y\right)} + \frac{M}{\left(\mathrm{Ln}\left(N_y\right)\right)^2} \tag{12.27}$$

Obviously, if N_y is zero, there can be no older animals in that cohort, so one should return a zero.

When Equation 12.24 is completed to include both Equations 12.20 and 12.25, along with the option of Equations 12.26 and 12.27, the result looks

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

dauntingly long and complex. Fortunately, the elements are simple, and with care they are easily implemented, even in Excel (see Example Box 12.5).

In operation, the initial guess at the numbers-at-age matrix (from one of the approximations) is updated using the function to be solved and its differential, as in Equation 12.24. Then the original values are replaced by the updated, which leads to a new set of updated values. This iterative process is repeated until no perceptible difference between the original and the updated values is observed. Usually, a stopping rule needs to be defined that stops the iterations once some threshold similarity has been reached (see Example Box 12.5, Figure 12.6).

12.2.5 Terminal *F* estimates

In the last age class the assumption is that the cohort will all die out. It is this termination of the cohort in the fishery that permits the backwards calculation of the numbers-at-age from the catch data and the natural mortality estimate. However, in the last year of data not all age classes will represent completed cohorts. If there are A year classes, then there will always be an $A - 1 \times A - 1$ triangle of incomplete cohorts arrayed along the bottom left of the numbers-at-age matrix. If we are to calculate their relative abundance in a valid manner, we must find a way to estimate the numbers-at-age in the unknown lower-left triangle of incomplete cohorts. Unfortunately, these incomplete cohorts are of most interest because they will affect future stock numbers; i.e., we know least about the year classes we need to know the most about.

Direct survey estimates of numbers- or fishing mortality-at-age, perhaps using tagging, are possible but rarely made. The most common approach is to produce independent estimates of the fishing mortality experienced by the cohorts being fished. Given a value of F_y and the catch-at-age in the final year, the numbers-at-age for the incomplete cohorts may be estimated using a rearrangement of Equation 12.15:

$$N_y = \frac{C_y}{1 - e^{-Z_y}} \frac{F_y + M}{F_y} \tag{12.28}$$

The independent estimates of F_y are known as terminal F estimates and, used with Equation 12.28, can give rise to improved estimates of the numbers-at-age in the bottom row of the numbers-at-age matrix.

The terminal F estimates may be obtained through some survey method or from the standard equation $F_y = q_a E_y$. The catchability coefficients for each age class a are commonly obtained by calculating the fishing mortality rate for complete cohorts, and then, using effort data in the years in which they were fished, the catchability by age can be determined from $q_a = F_y/E_y$. Given q_a and effort in the final year, the terminal numbers-at-age may be determined from Equation 12.28 (Example Box 12.6).

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.5

Gulland's cohort analysis focused on numbers-at-age. This is an extension to the worksheet generated in Example Box 12.4. It assumes that the catch-at-age data are in B6:J14 and that Pope's approximation to the numbers-at-age, as values, is in B20:J28. The matrix described below describes the modifier in Equation 12.24, i.e., it is the $f(N_y)/f'(N_y)$. The Excel equation looks terrific and great care is needed to put in the correct number of brackets. In U20 put =IF(J20>0,IF(K21>0, ((((1–(M/Ln(J20/K21)))*(J20–K21))–J6)/(1–(M/Ln(J20/K21)) + (((1–K21/J20)*M)/(Ln(J20/K21)^2)))), (((1–(M/Ln(J20)))*(J20)–J6)/(1–(M/Ln(J20)) + (M/(Ln(J20)^2))))),0). If you have the data entered correctly and the Pope's approximation is correct, the value in U20 should be 3.14. If this is the case, then copy U20 down to U28 and copy U20:U28 across to column M. Remember that cell J17 is named M. The equation inside the if statement has two options: first is Equation 12.20 divided by Equation 12.25, second is Equation 12.26 divided by Equation 12.27. The first option is for when both N_y and N_{y+1} are available, while the second option is for when no N_{y+1} are available. Thus, option 2 will be used in column U and in row 28. Copy the column labels down into L31:U31 and the row labels from L20:L28 into L32:L40. In U32 put =J20–U20 and copy down to U40 and copy U23:U40 across to column M. This completes Equation 12.24 and provides an updated version of the numbers-at-age. Notice that if the modifier matrix is negative in any cell this will increase the corresponding cell of the numbers-at-age matrix. Copy M32:U40 and paste as values over the top of the original numbers-at-age matrix in B20:J28. This updates the modifier, which updates the derived numbers-at-age. If this copying as values onto the original numbers-at-age matrix is repeated enough times the cells in the modifier matrix will tend to zero. In J16, put =sum(M20:U28) to enable a watch on the sum of the modifier matrix. Repeat the copying of the updated M32:U40 onto the original B20:J28 until the J16 limit is very small. At this point the numbers-at-age matrix should have stabilized to a solution. Does it appear to be closer to Pope or MacCall's approximation?

Once the spreadsheet is constructed and working it can be improved by automating the copying and pasting required in the Newton's method iterations. Buttons can be added that set up the original numbers-at-age matrix as a set of values from either Pope's or MacCall's approximations. A final button can be added that automates the copying of the updated numbers-at-age matrix as values onto the original numbers-at-age. A map of the worksheet developed in Example Box 12.5 should appear to be very similar to Figure 12.6.

continued

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.5 (continued)

	L	M	N	O	P	Q	R	S	T	U
19	**Y\Age**	**2+**	**3+**	**4+**	**5+**	**6+**	**7+**	**8+**	**9+**	**10+**
20	29/30	1.16	22.12	43.02	25.91	6.68	18.87	6.24	4.57	3.14
21	30/31	0.69	19.53	32.07	33.47	5.54	1.35	3.30	1.30	0.75
22	31/32	0.30	16.35	54.47	28.28	12.55	2.05	2.58	1.75	1.91
23	32/33	0.24	2.77	40.53	20.89	3.54	1.89	0.65	0.37	3.39
24	33/34	0.16	2.15	8.22	75.47	7.48	1.64	1.30	0.64	0.75
25	34/35	0.13	2.46	17.58	20.62	49.09	2.12	0.81	0.89	0.46
26	35/36	0.49	10.04	25.31	28.06	33.97	16.00	3.81	0.72	1.51
27	36/37	0.13	6.23	60.53	30.06	21.99	10.01	10.25	0.45	0.53
28	37/38	11.08	125.6	178.2	178.9	38.50	34.99	13.37	19.45	1.36
~										
31	**Y\Age**	**2+**	**3+**	**4+**	**5+**	**6+**	**7+**	**8+**	**9+**	**10+**
32	29/30	15282	8677	7461	2148	859	1390	408	=I20–T20	=J20–U20
33	30/31	38951	11594	4872	3355	692	347	416	=I21–T21	=J21–U21
34	31/32	13196	30123	7008	2081	1197	275	175	I22–T22	=J22–U22
35	32/33	11922	10191	20874	2817	624	418	112	50	54

12.2.6 Potential Problems with Cohort Analysis

The name *virtual population analysis* or *cohort analysis* refers to a class of models each of which relies upon knowledge of final fishing mortality rates and the back-calculation of numbers-at-age in the fished population. The method is clearly sensitive to the estimates of the terminal F values that permit the analysis to be extended to the incomplete cohorts. If these estimates are flawed, the analysis will be biased. Other sources of potential bias include the presence of ageing error in the determination of the catch-at-age. This is especially a problem if recruitment is highly variable. A 10% ageing error will not affect a small year class too badly, but a large year class could inflate the apparent numbers of smaller year classes around it. Thus, ageing errors will tend to obscure recruitment variability (Richards et al., 1992). A further problem relates to the idea of following particular cohorts. If there is significant immigration to a region, then the numbers in a cohort may increase through time, and thus lead to an overestimate of the cohort size. This effect would be greatest if older fish are the ones that migrate the most. Finally, errors in the natural mortality estimate can have relatively complex impacts on the estimation of fishing mortality (Mertz and Myers, 1997).

12.2.7 Concluding r emarks on Cohort Analysis

Only a very introductory treatment of cohort analysis is presented here. Virtual population analysis remains the preferred method of stock

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

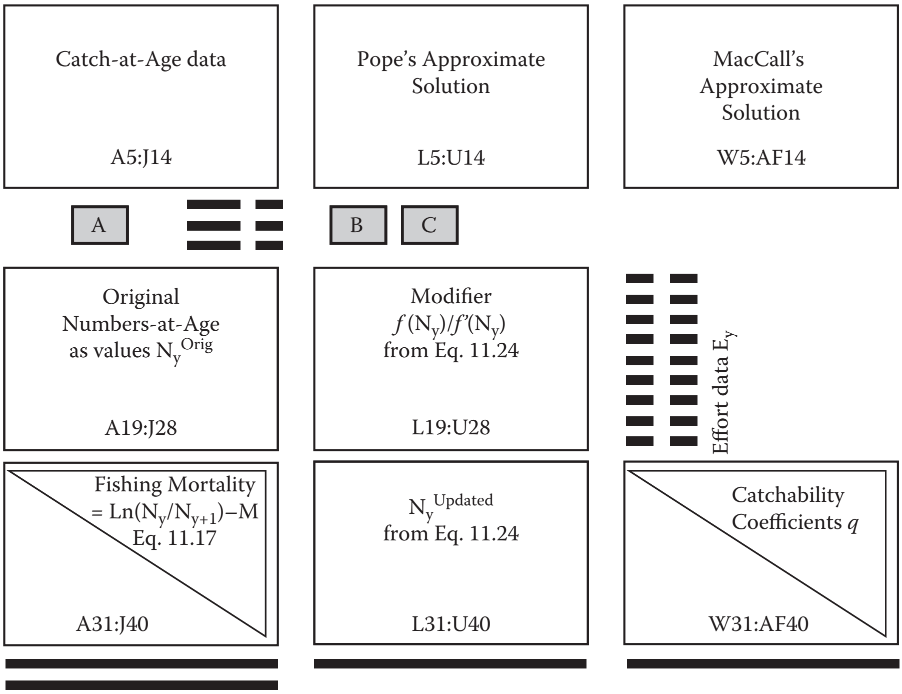

Figure 12.6
A schematic map of the worksheet developed in Example Boxes 12.4 and 12.5. Each block contains a matrix and its column and row labels. The contents are identified along with the cell positions on the sheet. The three grey objects are buttons attached to macros. The three sets of lines are text labels and numbers, including natural mortality, *M*. The rectangles with unbolded text will be added later. The triangles imply that these are only upper-triangular matrices. The lines below MacCall's solution represent the effort data. The lines below the lower boxes are summary data, such as average age-specific fishing mortality (see later). The effort data are used to calculate the terminal *F* values.

assessment in the European Community and other parts of the Atlantic. Not surprisingly, there is an enormous literature, both grey and formal, dealing with the various techniques, ways to improve their performance, and how to "tune" the VPA to information from the fishery. The aim of this chapter was only to introduce the reader to the methodology so that a more detailed investigation of the primary literature would be more understandable. Megrey (1989) and Quinn and Deriso (1999) provide excellent reviews of the recent developments in the techniques used.

Without using an optimization technique (such as minimizing the sum of squared residuals), model selection and determination of uncertainty levels along with model projections for risk assessments present their own challenges. The alternative analytical strategy of statistical catch-at-age or catch-at-age with ancillary information, or integrated analysis, permits these options in a very straightforward manner.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.6

Calculation of the instantaneous fishing mortality rates for complete cohorts. This extends Example Boxes 12.4 and 12.5. Using Equation 12.17, and checking for zeros, put =if(T32>0,if(U33>0,Ln(T32/U33)–M,"—"),"—") into I32. Copy this down to row 39, then copy I32:I39 across to column B. Delete the subdiagonal elements. In B41 put =average(B32:B39) and copy across to column I to provide the estimates of average fishing mortality rate on each age class. One might be tempted to use the average fishing mortality in Equation 12.28, but this would ignore variations in fishing effort in each year. Put the effort data into X20:X28 (from Table 12.1; it is best to label the column as effort in X19 and put year labels in W20:W28). Then, in B42 put =X41*$X$28 to estimate F_a as $q_y E_y$ and copy across to column I. Note that these only go across to where the incomplete cohorts apply. The effort data in the final year (X28) and average catchability for each age class are described below as estimates of the catchability coefficient by age class and year. The column labels, below, extend across to AF31, but that column does not produce numbers and is suppressed here for brevity.

	A	B	C	D	E	F	G	H	I	J
31	**Year\Age**	**2+**	**3+**	**4+**	**5+**	**6+**	**7+**	**8+**	**9+**	**10+**
32	29/30	0.025	0.325	0.547	0.885	0.664	0.964	1.050	1.766	—
33	30/31		0.249	0.596	0.781	0.676	0.428	0.681	0.881	—
34	31/32			0.662	0.958	0.807	0.647	1.020	0.835	—
35	32/33				0.643	0.541	0.470	0.551	0.657	—
36	33/34					0.590	0.550	0.590	0.922	—
37	34/35						0.452	0.587	0.803	—
38	35/36							1.055	0.945	—
39	36/37								0.607	—
40	37/38	—	—	—	—	—	—	—	—	—
41	Average F	0.025	0.287	0.602	0.817	0.656	0.585	0.791	0.927	
42	Terminal F	0.019	0.217	0.482	0.669	0.541	0.485	0.674	0.789	

In AE32 put =if(isnumber(I32),I32/$X20,"—") and copy down to row 40. Copy AE32:AE40 across to column X (and into AF for completeness). In X41 put =average(X32:X40) and copy across to column AE to generate the average catchability coefficient for each age class. These get used, along with the last year's effort in X28, to generate the age-specific fishing mortalities for the incomplete cohorts (across in B42:I42). In X42 put =(B14/(1–exp(–(M+B42))))*((B42+M)/B42), which is Equation 12.28, and

continued

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.6 (continued)

copy across to column AE to generate the terminal N_y values. These are then ready for pasting into the left-hand side of the bottom row of the original numbers-at-age matrix. The algorithm is now to paste as values the approximation of the terminal Ns onto the estimated numbers-at-age matrix. Therefore, copy X42:AE42 and paste as values into B28:I28. This will produce a new updated numbers-at-age matrix along with a new estimate of the terminal F_y and N_y. This row copying procedure is iterated until the limit of precision selected has been reached. When using the terminal *F* estimates the limit is found by summing the modifier matrix except the bottom row (i.e., =sum(M20:U27) instead of =sum(M20:U28) in J16). To ensure no errors it is best to write a few short macros to do the copying and pasting of values. As in Figure 12.6, create two buttons with attached macros, one of which copies and pastes Pope's approximation and the other MacCall's approximation into B20:J28. The third button should conduct the copy/pasting relating to the Newton's method iterations. If J15 is given an integer value greater than zero, it would be possible to turn the use of the terminal *F* values on and off using a macro similar to that printed next. In this way it would be possible to easily compare the results obtained from the different starting points. How different are the answers if the terminal *F* values are used? Do the years with complete cohorts change or is it just the numbers-at-age for the incomplete cohorts?

	W	X	Y	Z	AA	AB	AC	AD	AE
30	Catchability Coefficient								
31	Y\Age	2+	3+	4+	5+	6+	7+	8+	9+
32	29/30	0.0043	0.0560	0.0942	0.1523	0.1142	0.1660	0.1807	0.3039
33	30/31	—	0.0427	0.1021	0.1337	0.1157	0.0733	0.1166	0.1508
34	31/32	—	—	0.1333	0.1928	0.1625	0.1301	0.2052	0.1680
35	32/33	—	—	—	0.1309	0.1103	0.0957	0.1122	0.1337
36	33/34	—	—	—	—	0.1137	0.1060	0.1136	0.1776
37	34/35	—	—	—	—	—	0.0914	0.1189	0.1625
38	35/36	—	—	—	—	—	—	0.2280	0.2041
39	36/37	—	—	—	—	—	—	—	0.1367
40	37/38	—	—	—	—	—	—	—	—
41	Av_q	0.0043	0.0494	0.1098	0.1524	0.1233	0.1104	0.1536	0.1797
42	Term N	8363	7468	5277	4135	1139	1136	370	465
43	Min N	370.4							

VBA macro to perform the Newton's method iterations with or without terminal F_y estimates leading to terminal N_y estimates for the incomplete cohorts. By placing 1 in J15 the terminal *F* estimates are made. If J15 is zero, then the original approximations are used.

continued

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central,
http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.6 (continued)

```
Sub do_vpa()
Dim termf As Integer
  termf = Range("J15").Value
  If termf > 0 Then
    Range("J16").Select       ' Alter the limit summation
    ActiveCell.FormulaR1C1 = "=SUM(R[4]C[3]:R[11]C[11])"
  Else
    Range("J16").Select
    ActiveCell.FormulaR1C1 = "=SUM(R[4]C[3]:R[12]C[11])"
  End If
  Range("M32:U40").Select     ' copy and paste the updated
  Selection.Copy              ' numbers-at-age
  Range("B20").Select
  Selection.PasteSpecial Paste:=xlValues
  Range("B19").Select
  If termf > 0 Then           ' If the terminal F option
    Range("X42:AE42").Select  ' Copy and paste terminal N
    Selection.Copy
    Range("B28").Select
    Selection.PasteSpecial Paste:=xlValues
    Range("B12").Select
  End If
  Application.CutCopyMode = False
End Sub
```

12.3 Statistical Catch-at-Age

12.3.1 introduction

Statistical catch-at-age will be referred to here as integrated analysis (Punt et al., 2001). Unlike VPA, integrated analysis estimates fewer parameters than the available number of data points, although one can still be estimating tens of parameters. An objective function (least squares or maximum likelihood) is used to optimize the fit of the model to the available data. It requires catch-at-age data along with some information to tie the model to the stock size (either catch rates, or effort, or independent population estimates). The numbers-at-age at the start of the first year in the population being modelled are model parameters along with recruitment levels in each year of the fishery. With further parameters describing age-specific selectivity, it is possible to project each cohort forward to generate a matrix of numbers-at-age. From this it is possible to generate a matrix of predicted catch-at-age that can

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central,
http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

be compared with the observed data and the fit optimized. Catch-at-age data alone are usually insufficient to tie the model to reality, so a further connection, either through effort or catch rates, is necessary.

12.3.2 The equations

The equations behind integrated analysis are remarkably simple; the complex part is organizing the information and calculations. The fully selected fishing mortality rate in year y, F_y, is one of the foundations of the analysis, and values for each year are treated as model parameters in the fitting process. The fishing mortality rate for each age, a, in each year y, $F_{a,y}$, is

$$F_{a,y} = s_a \hat{F}_y \tag{12.29}$$

where F_y hat is the fitted fishing mortality rate in year y and s_a is the selectivity of age a (Figure 12.7). The fishing mortalities are combined with the natural mortality, M, to generate the age- and year-specific survivorships, which are used to complete the matrix of numbers-at-age:

$$N_{a+1,y+1} = N_{a,y} e^{-\left(M + s_a \hat{F}_y\right)} = N_{a,y} e^{-M} e^{-s_a \hat{F}_y} \tag{12.30}$$

where, as before, $N_{a,y}$ is the numbers of age a in year y (Figure 12.8).

Selectivity can be estimated either directly for each age or the parameters of an equation describing the shape of the selectivity curve can be estimated. A logistic equation is often used to describe selectivity. Given a as age, a_{50} as

Fig ur e 12.7
Selectivity curve generated by Equation 12.31. It is not smooth because the different ages have been treated as integers. By varying the parameters a_{50} and a_{95} the steepness of the curve and its location along the age axis can be altered.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

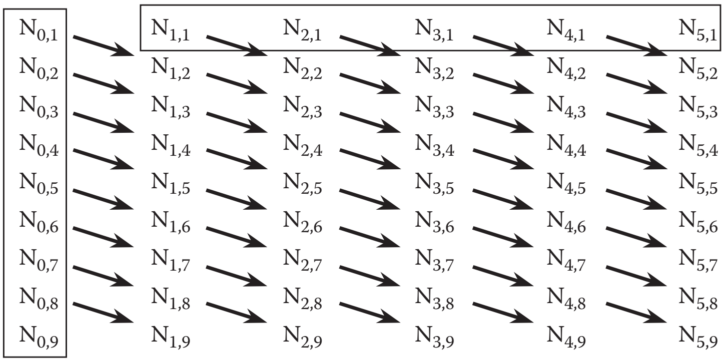

Fig ur e 12.8
In integrated analysis the initial population age structure and each year's recruitment (boxed off in the diagram) are estimated parameters. The survivorship of each age class in each year is calculated ($e^{-(M+sF)}$), and these survivorships are used to complete the numbers-at-age matrix using Equation 12.30.

the age at which selectivity is 50%, a_{95} as the age at which selectivity is 95%, and s_a *a*s the selectivity at age *a* (Figure 12.7), age-specific selectivity is

$$s_a = \frac{1}{1 + e^{-\mathrm{Ln}(19)\frac{(a-a_{50})}{(a_{95}-a_{50})}}} \tag{12.31}$$

Fitting the model involves estimating the $N_{a,y}$ for all ages in year 1, and for the first age class in all subsequent years (Figure 12.8).

Despite estimating a large number of parameters, the calculations for an integrated analysis can still be conducted successfully on an Excel worksheet. We will construct an example of an integrated analysis using the data from Beverton and Holt (1957) that was used in the cohort analysis. Once again, this will be a complex worksheet and care is needed in its construction. The map in Figure 12.9 indicates the broad structure.

The algorithm begins by calculating the age-specific fishing mortality for each year from the selectivity equation parameters and the fishing mortality parameters (Equation 12.29, Figure 12.9). These are combined with the natural mortality rate to generate the age-specific survivorships, which are used, in turn, to complete the numbers-at-age matrix (Equation 12.30).

12.3.3 Fitting to Catch-at-Age Data

Once the predicted numbers-at-age are calculated, the predicted catch-at-age can be generated, which provides the first opportunity to generate an objective function for use when fitting the model by comparing the observed catch-at-age with the predicted. Predicted catch-at-age is

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

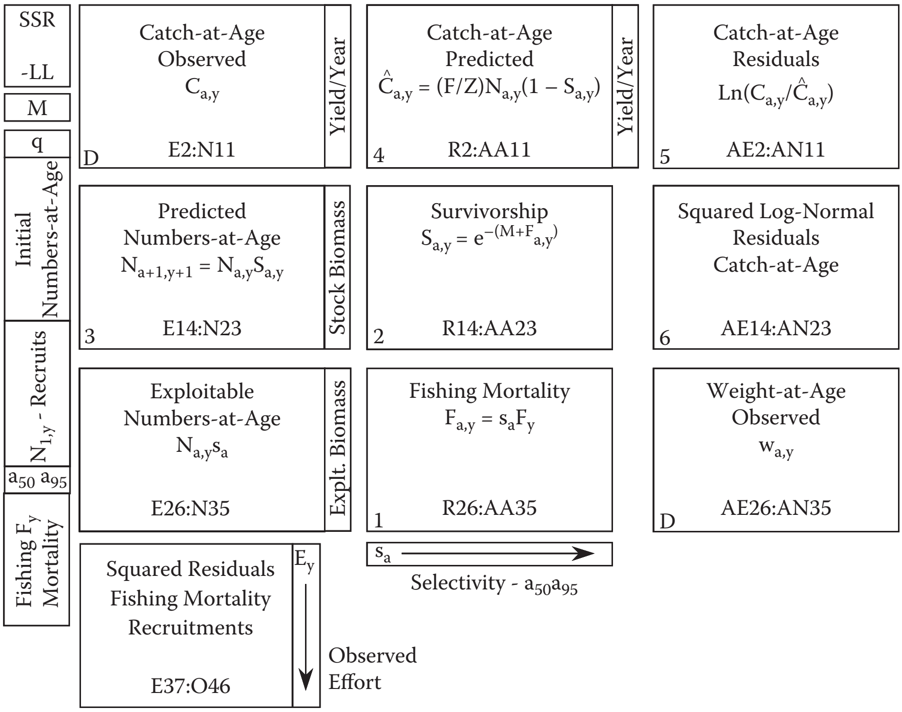

Figure 12.9
Schematic map of an Excel worksheet illustrating a possible layout for an integrated analysis. Each of the nine main boxes would have its upper edges labeled with ages and left-hand edges labeled with years. The small, upper, left-most box is where the minimizations occur. The model parameters are listed down the left-hand side of the worksheet below the natural mortality *M*. Each of the estimated parameters is natural log transformed, which scales all parameters to similar sizes and makes all parameter changes proportional changes. In the worksheet they are back-transformed as appropriate. The catch-at-age residuals in the top right are there for ease of plotting as a diagnostic relating to the quality of fit (Figure 12.11). The heavy numerals in the lower left of some of the boxes relate to the order of calculation in the algorithm. The capital D, in the same place, implies these are data matrices. See Example Boxes 12.7 and 12.8.

$$\hat{C}_{a,y} = \frac{F_{a,y}}{M + F_{a,y}} N_{a,y} \left(1 - e^{-(M+F_{a,y})}\right) \tag{12.32}$$

Doubleday (1976) suggested using lognormal residual errors:

$$SSR_C = \sum_a \sum_y \left(\mathrm{Ln}\left(C_{a,y} / \hat{C}_{a,y}\right)\right)^2 = \sum_a \sum_y \left(\mathrm{Ln}C_{a,y} - \mathrm{Ln}\hat{C}_{a,y}\right)^2 \tag{12.33}$$

With the Beverton and Holt (1957) North Sea plaice data, an optimum fit gives rise to a relatively even spread of residuals (according to Equation 12.33, Figures 12.10 and 12.11, Example Box 12.7).

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

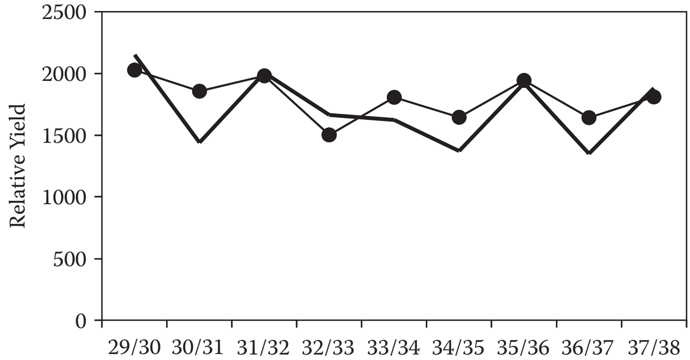

Fig ur e 12.10
A plot of observed yield (fine line with dots) and predicted yield (thick line) against fishing year (April to March) deriving from a fit on catch-at-age alone (see Example Box 12.7). The pattern matches well but is exaggerated in 30/31, and is less pronounced in 33/34.

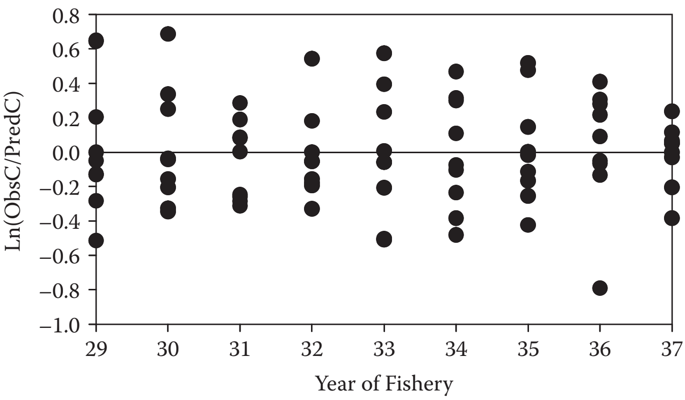

Fig ur e 12.11
Catch-at-age residuals against fishing year when fitting the model only to catch-at-age data (see Example Box 12.7). The overall fit is good; a regression through these residuals is essentially flat (residual = 0.0156–0.00047 year).

Doubleday (1976) was able to show that if one only fits the stock assessment model using catch-at-age data, as with Equation 12.33, then correlations between some of the parameter estimates could be so extreme that some of the parameter estimates can be effectively linear combinations of others. While this may be acceptable for estimating changes in the relative abundance, at least over short periods, catch-at-age data alone are insufficient to estimate absolute abundance. It is for this reason that integrated analysis is sometimes referred to as catch-at-age analysis with auxiliary data (Deriso et al., 1985, 1989). A common addition is to include observed effort into the model, combined with an extra parameter, the catchability coefficient.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.7

An integrated analysis of the North Sea plaice data (Beverton and Holt, 1957; see Table 12.1 and earlier example boxes). In preparation for later developments put =sum(AF15:AN23) in B2 and =sum(K38:K46) into B3. In B5 put =C2*B2+C3*B3 to get the total weighted sum of squared residuals. The log of the initial numbers-at-age estimates are given in B9:B16 (e.g., Ln_N_29_10 is the log of numbers-at-age 10 in 1929), and the log recruitments in each of the nine years are given in B17:B25 (e.g., Ln_R_30_2 is the log recruitment into age 2+ in year 1930). The logged parameters of the selectivity equation are given in B26:B27. Finally, the log fishing mortality rates are given in rows 28 down to 36. Initiate all parameters (the logs of the first row and column of Pope's approximation will fill B9:B25; you can get the values from Example Box 12.6; see Appendix 12.1), and then enter the catch-at-age data (from Table 12.1) in F3:N11 along with appropriate column and row labels (as in E3:E11 and E2:N2). Equivalent weight-at-age data should be entered into AF27:AN35 (with labels in row 26 and column AE); the weight data are in Appendix 12.1. In R37 put the label "Age," in S37:AA37 put the numbers 2 to 10, and in S38 put =1/(1+exp(–Ln(19)*(S37–exp($B$26))/(exp($B$27)–exp($B$26)))), which is Equation 12.31. Copy S38 across to column AA to calculate selectivity (label as "Selectivity" in R38). In R25 put the label "Fishing Mortality by Age by Year" and label ages in S26:AA26 and years in R27:R35 as before. To calculate these fishing mortalities put =S$38*exp($B28) into S27, being sure to get the single $ signs in the correct places. Copy S27 down to row 35, then copy S27:S35 across to column AA to generate the required matrix. Put =exp(–(S27+$B$6)) in S15 and copy down to row 23, then copy S15:S23 across to column AA to generate the survivorships by age and year. Again label for rows and columns appropriately.

	A	B	C	D	E	F	G	H
1	**Source**	**Value**	**Wt**		**Catch-at-Age—Observed**			
2	**SSRC**	7.6363	1		**Year\Age**	2+	3+	4+
3	**SSRE**	0.3789	1		**29/30**	328	2120	2783
4					**30/31**	223	2246	1938
5	**SSRT**	8.0152			**31/32**	95	2898	3017
6	**Natural M**	0.2			**32/33**	77	606	4385
7	**Parameter**	**Ln(value)**			**33/34**	50	489	1121
8	**Ln_q**	–2.0981			**34/35**	44	475	1666
9	**Ln_N_29_10**	4.4570			**35/36**	131	1373	1595

continued

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.7 (continued)

	A	B	C	D	E	F	G	H
10	**Ln_N_29_9**	5.0806			**36/37**	38	691	2862
~	~	~			~	~	~	~
17	**Ln_R_29_2**	9.3607			**31/32**	10807	25748	6614
18	**Ln_R_30_2**	10.3659			**32/33**	10684	8747	17685
19	**Ln_R_31_2**	9.2879			**33/34**	7478	8680	6359
~	~	~			~	~	~	~
25	**Ln_R_37_2**	9.5997			**Exploitable Numbers-at-Age**			
26	**Ln_Sel50**	1.2232			**Year\Age**	**2+**	**3+**	**4+**
27	**Ln_Sel95**	1.4784			**29/30**	177.6	1468.7	6398.4
28	**Ln_F_29**	–0.0739			**30/31**	485.2	2195.2	3545.5
29	**Ln_F_30**	–0.4475			**31/32**	165.1	6024.9	5671.0

From the worksheet map (Figure 12.9), you should now have in place the parameters, both D (data) matrices and the fishing mortality (1) and survivorship (2) matrices. The predicted numbers-at-age matrix is constructed in two parts. In F15 put =exp(B17) and copy down to F23 to obtain the estimated recruitments. In G15 put =exp(B16), in H15 put = exp(B15), and similarly across to N15 with =exp(B9), to list the initial numbers-at-age. The matrix can then be completed by putting =M15*Z15 into N16 and copying down to N23, and then copying N16:N23 across to column G. Finally, to obtain the total biomass estimates the numbers in each age class must be multiplied by their respective average weight and summed. Put =sumproduct(F15:N15,AF27:AN27)/1000 into O15 and copy down to O23. The division by 1,000 is to keep the numbers manageable. The variation in recruitment from year to year is obvious, as are the strong year classes passing through the fishery. The stock biomass can be used as a diagnostic to indicate annual trends in relative abundance. By putting =sumproduct(F3:N3,$AF27:$AN27)/1000 into O3 and copying down to O11, an equivalent column can be added to the observed catch-at-age matrix; these values would be observed annual yield. From the matrix of numbers-at-age we calculate the predicted catch-at-age by putting =(S27/(S27+$B$6))*F15*(1–S15) into S3 and copying down to S11 and across to column AA. This is Equation 12.32, with S15 being the survivorship calculated previously. You should label the rows and columns appropriately. Copy O3:O11 and paste into AB3:AB11 to obtain the predicted annual yield. The final two matrices are made by putting =Ln(F3/S3) into AF3, copying down to AF11 and

continued

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.7 (continued)

across to column AN. These are just the residuals and can be plotted against the year of the fishery to act as a diagnostic relating to how well the fitting process is proceeding (cf. Figure 12.11). To obtain the squared residuals put =AF3^2 into AF15 and copy across to AN15 and down to row 23. This has the effect of completing the sum of squared residuals for the catch-at-age held in B2. With 8 in B9:B25, 1.25 in B26, 1.75 in B27, 0 in B28, and –0.4 in B29:B36 the SSR_C is about 176.938. Use the solver to minimize B2 by changing B9:B36 (twenty-eight parameters). If it does not converge to about 7.265 try a different starting point (cf. Figures 12.10 and 12.11). The values below are close to those obtained once Example Box 12.8 is implemented and completed.

	E	F	G	H	I	J	K	L	M	N	O
14	**Yr\Age**	**2+**	**3+**	**4+**	**5+**	**6+**	**7+**	**8+**	**9+**	**10+**	**Biomass**
15	**29/30**	11623	6277	7462	2274	1087	1539	441	161	86	6721
16	**30/31**	31755	9382	4136	2755	741	352	498	143	52	9048
17	**31/32**	10805	25746	6614	1957	1197	320	152	215	62	8404
18	**32/33**	10685	8746	17684	2845	761	463	124	59	83	7483
19	**33/34**	7478	8680	6359	9370	1408	375	228	61	29	7034
20	**34/35**	11660	6076	6336	3419	4717	706	188	114	31	6967
21	**35/36**	9599	9480	4469	3501	1776	2442	366	97	59	6957
22	**36/37**	7889	7780	6648	2075	1488	751	1032	154	41	5777
23	**37/38**	14765	6407	5629	3460	1006	718	362	498	75	7093

12.3.4 Fitting to Fully Selected Fishing Mortality

Including the observed fishing effort (E_y) adds a further nine data points, which offsets the fact that an estimate of the catchability coefficient (q) must be made. The two are combined to generate a semiobserved fishing mortality (F_y), which is compared with the fully selected fishing mortality parameters (Example Box 12.8, Figures 12.12 and 12.13). Take note that this fishing mortality is not the instantaneous fishing mortality rate but is rather the annual harvest rate:

$$F_y = \hat{q}E_y \tag{12.34}$$

and thus:

$$SSR_E = \sum_y \left[\text{Ln}\left(\hat{F}_y\right) - \left(\text{Ln}\left(\hat{q}\right) + \text{Ln}\left(E_y\right)\right)\right]^2 \tag{12.35}$$

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.8

Include observed effort data into O38:O46 in the worksheet developed in Example Box 12.7 (the data are in Table 12.1). Copy J38:N38, as below, down to row 46. Column N is Equation 12.34, column J is the inner part of Equation 12.35, deriving from the parameter estimate of F_y and the semiobserved in column M. Column K is the residual squared, which generates the sum of squared residuals in B3. Columns L and N are for plotting as a visual diagnostic of the quality of fit (Figure 12.13), though one could equally well plot column J. If the worksheet has been solved for an optimum fit to catch-at-age and Ln_q = –2, then B3, the sum of squared residuals against effort, should be about 4.55. If both contributions to the total sum of squared residuals in B5 are given equal weight (i.e., 1 in C2:C3), then one can minimize B5 using the solver to modify cells B8:B36 (twenty-nine parameters). Given equal weight, the optimum balance between the two sources of squared residuals appears to be where $SSR_C = 7.636$ and $SSR_E = 0.378$. The parameters that give rise to this result are given in Appendix 12.1. It may be necessary to run the solver more than once, or from different starting points (the results from a VPA, or one of the approximations, would be a reasonable starting point—try the values obtained from Example Box 12.6). How robust are the answers when each SSR is given equal weighting? How sensitive are the results to the relative weighting of the two sets of squared residuals? The values below are for the optimum fit with equal weights.

	I	J	K	L	M	N	O
37	**Yr**	**F_res**	**SSq**	**Pred_F**	**ObsLnF**	**ObsF**	**Effort**
38	29	=B28–M38	=J38^2	=exp(B28)	=Ln(N38)	=exp($B$8)*O38	5.81
39	30	=B29–M39	=J39^2	=exp(B29)	=Ln(N39)	=exp($B$8)*O39	5.84
40	31	0.2080	0.0432	0.7506	–0.4949	0.6096	4.97
41	32	–0.1712	0.0293	0.5075	–0.5070	0.6023	4.91
42	33	–0.2610	0.0681	0.4904	–0.4516	0.6366	5.19
43	34	–0.2787	0.0777	0.4586	–0.5009	0.6060	4.94
44	35	0.1522	0.0232	0.6613	–0.5658	0.5679	4.63
45	36	–0.0304	0.0009	0.5283	–0.6077	0.5446	4.44
46	37	0.2305	0.0531	0.6781	–0.6190	0.5385	4.39

12.3.5 Adding a Stock r ecruitment r elationship

There is very little information relating to the last few cohorts, with the extreme being the single catch data point for the very latest recruits. Fitting the model to the catch-at-age and estimates of fully selected fishing mortality could generate relatively uncertain estimates of the status of the affected years and age classes. One suggested solution (Fournier and Archibald, 1982)

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

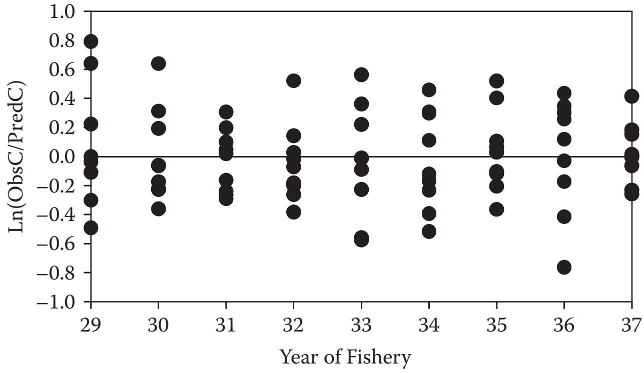

Figure 12.12
A plot of the residuals between the log of the observed catch-at-age and the log of the expected catch-at-age, from the North Sea plaice data (Beverton and Holt, 1957), fitted to catch-at-age and fishing mortality (Example Boxes 12.7 and 12.8). A regression line fitted to these data, as shown by the fine line, is essentially flat along the zero line (residual = 0.00389–0.00012 year). Note the differences between this residual plot and that shown in Figure 12.11.

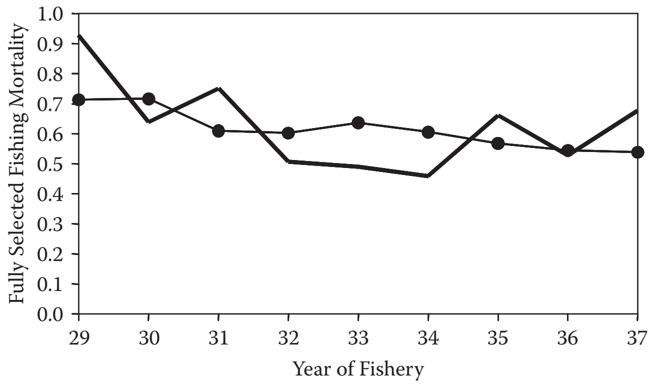

Figure 12.13
Estimated fully selected fishing mortality (thick line) vs. observed fishing mortality rates (as in Equation 12.35).

is to impose a stock recruitment relationship (see Chapter 10) to add extra constraints to the last few years of recruitment. This will also be necessary if risk assessment projections are to be made. Spawning stock size may be defined as the mature biomass, or the stock size times the relative fecundity at age, or some other available measure. In addition, one can use a variety of stock recruitment relationships (e.g., Beverton and Holt or Ricker). In general, as discussed in Chapter 10, the residual errors used would tend to be lognormal. Of course, if the age at recruitment is not 0+, then the implied time lag between spawning stock size and subsequent recruitment must be accounted for:

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

$$\hat{N}_{r,y+r} = \frac{\alpha B_y^S}{\beta + B_y^S} e^{\varepsilon} \tag{12.36}$$

where r is the age at recruitment, B^S_y is the spawning stock size in year y, and $y + r$ is the year plus the time lag before the recruits join the fishery. The parameters α and β are from the Beverton and Holt stock recruitment relationship. Whether the predicted number of recruits in year $y + r$ are derived from a Beverton and Holt, a Ricker, or any other stock recruitment relationship does not affect the form of the squared residuals:

$$SSR_R = \left(\mathrm{Ln}N_{r,y+r} - \mathrm{Ln}\hat{N}_{r,y+r}\right)^2 \tag{12.37}$$

which compares the observed model parameters for recruitment in the rth year onwards (i.e., the fitted recruitments in B17:B25) with those expected from the stock recruitment relationship (Example Box 12.9, Figure 12.14).

12.3.6 Other Auxiliary Data and Different Criteria of Fit

We have considered fitting the catch-at-age stock assessment model through using catch-at-age data, using relative effort (fishing mortality) data, and adding a stock recruitment relationship. Other possible sources that could be added include fishery-independent surveys of stock size and fishery-dependent catch/effort rates. If a series of fishery-independent surveys is available, then either they should derive from a standardized design or some measure of the relative efficiency of each survey (a relative catchability coefficient) would be required. If commercial catch rates are to be used, then, ideally, these should be standardized to remove noise unrelated to changes in stock size (Kimura, 1981, 1988; Klaer, 1994).

With catch rates, lognormal residuals tend to be used. The expected catch rates derive from the simple relation

$$\hat{I}_y = qB_y^E = q\sum_a w_a s_a N_{a,y} \tag{12.38}$$

where q is the catchability coefficient, B^E_y is the exploitable biomass in year y, w_a is the average weight of fish of age a, s_a is the age-specific selectivity, and $N_{a,y}$ is the numbers-at-age a in year y. The sum of squared residuals for this potential component of the total would be (Example Box 12.10)

$$SSR_I = \sum_y \left(\mathrm{Ln}\left(I_y\right) - \mathrm{Ln}\left(\hat{I}_y\right)\right)^2 \tag{12.39}$$

A similar arrangement can be used for fishery-independent surveys, although a separate q estimate would be required for the survey estimates of exploitable biomass.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.9

Implementation of a stock recruitment relationship into the worksheet developed in Example Boxes 12.7 and 12.8. Copy F40:G40 down to row 46. In A37 put "Alpha" and A38 "Beta," adding the logged initial value ready for fitting the model ($\alpha = 9$ and $\beta = -5$, are reasonable beginnings for the Beverton and Holt relationship, in B37:B38). If a Ricker relationship is preferred then put =exp($B$37)*O15*exp(–exp($B$38)*O15) into F40 and copy down to row 46. In this case starting values would need to be something like ($\text{Ln}(\alpha) = 1.6$ and $\text{Ln}(\beta) = -8.6$). Plot the recruitment parameters against those predicted from the stock recruitment relationship. Which appears best, the Beverton and Holt or the Ricker? F38:F39 are empty because of the lag of *r* years. There are now thirty-one parameters for the solver, B8:B38.

	E	F	G
37	**Year\Age**	**Beverton and Holt**	**Sum Squared Residual**
38	29/30		
39	30/31		
40	31/32	=exp($B$37)*O15/(exp($B$38)+O15)	=(B19–Ln(F40))^2
41	32/33	=exp($B$37)*O16/(exp($B$38)+O16)	=(B20–Ln(F41))^2
42	33/34	9625	0.059
43	34/35	9625	0.027
44	35/36	9625	0.002
45	36/37	9625	0.003
46	37/38	9625	0.038

The sum of squared residuals for all the additions to the integrated analysis. SSR_C relates to fitting the model to the catch-at-age data, SSR_E relates to fitting the model to the fully selected fishing mortality estimates, and SSR_R relates to the inclusion of the stock recruitment relationship to the model fitting. The total sum of squared residuals, SSR_T, is the sum of each of the separate contributions after each is weighted according to the predefined weightings. As with all weighted least squares methods, ideally the weighting should relate to the variability in the estimates of the statistics involved; unfortunately, this is often unknown. Experiment with different relative weightings to see the impact upon the final model fit. If it is desired to turn off a particular component, simply give it a weighting of zero. Does the impact of adding the stock recruitment relationship differ depending upon whether a Beverton and Holt or a Ricker relationship is used?

continued

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.9 (continued)

	A	B	C
1	**Source**	**Value**	**Weighting**
2	**SSRC**	=sum(AF15:AN23)	1
3	**SSRE**	=sum(K38:K46)	1
4	**SSRR**	=sum(G40:G46)	1
5	**SSRT**	=C2*B2+C3*B3+C4*B4	
6	**Natural M**	0.2	

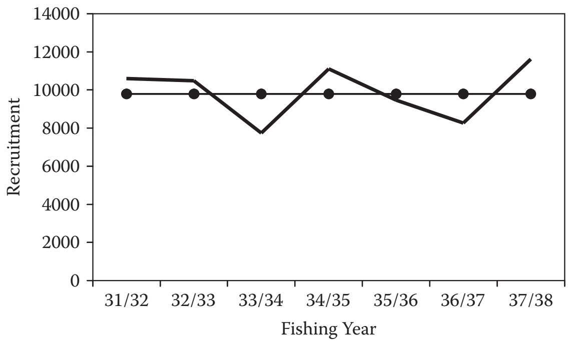

Fig ur e 12.14
Fitted recruitment levels (thick line) vs. expected recruitment levels from a Beverton-Holt stock recruitment relationship (Ln_Alpha = 9.1920 and Ln_Beta = –5.000). The stock recruitment relationship predicts effectively constant recruitment.

While we have consistently been using the sum of squared residuals as the criterion of model fit, we could equally well have used maximum likelihood methods and their extensions into Bayesian methods. Using the same weightings and the lognormal residuals, we would expect to obtain essentially the same answers, but if any of the contributions fit very closely, the log-likelihoods for that component could go negative, which will distort the impact of that component and a different result may occur. If the model is fitted to only one component, then the same parameters will be found as with the least squares method.

The multinomial distribution has been suggested as an alternative likelihood function for fitting the catch-at-age data (Deriso et al., 1985, 1989), especially where measurement errors are primarily due to ageing errors and sampling error in the catch sampling:

$$\mathrm{LL} = \sum_{y} \sum_{a} n_{a,y} \mathrm{Ln}\left(\hat{p}_{a,y}\right) \tag{12.40}$$

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.10

Exploitable biomass and expected catch rates. In E25 put the label "Exploitable Numbers-at-Age" and then label the columns as age classes in F26:N26 and the rows as years in E27:E35. Put =F15*S$38 into F27 and copy down to row 35 and across to column N to generate the required matrix. Label O26 as "Biomass." Put =sumproduct(F27:N27,AF 27:AN27)/1000 into O27 and copy down to row 35 to generate the exploitable biomass in each year of the fishery. Put =O27*exp($B$8)/1000 into P27, and copy down to row 35 to generate the expected catch rates (as per Equation 12.38). Label P37 Obs(CE) and in P38 put AB3/(O38*1000), which divides the expected yield by the scaled effort. Then a column of the squared residuals (as per Equation 12.39) can be placed somewhere convenient on the worksheet and the weighted sum can be added to the total sum of squared residuals in B5.

where LL is the multinomial log-likelihood, $n_{a,y}$ is the number of fish of age a, aged in year y, and $\hat{p}_{a,y}$ is the expected proportion of animals age a in year y. Equation 12.40 does not actually involve the amount of catch, so it has been suggested that it might be more stable to add a penalty term along the lines of $\Sigma(C'_y - C_y)^2$, weighted so that the expected catches (C'_y) are close to the observed (Quinn and Deriso, 1999). An alternative approach has been suggested by Schnute and Richards (1995) that does not require this extra term. If the use of multinomial likelihoods is relatively unstable, it is also an option to solve the model initially using lognormal likelihoods, and then, when near the optimum fit, begin using multinomial likelihoods.

12.3.7 r elative Weight to Different Contributions

Irrespective of whether maximum likelihood methods, Bayesian methods, or sum of squared residuals are used as the criterion of optimum fit, when there are a number of categorically different sources contributing to the overall likelihood or sum of squares, then each contribution will receive a relative weighting. Providing no explicit weighting implies an equal weighting of 1 for each component of the total. With weighted least squares, and maximum penalized likelihood, it is usual to weight each component in relation to the degree of uncertainty associated with the data or statistic used. Thus, if some estimate of the variance or coefficient of variation for each data source is available, then the inverse of these would be used to ascribe relative weights to each component; the greater the variability, the less the relative weight.

If each of the sources of information is consistent with the others, then fitting the model should not prove difficult. For example, in Example Box 12.8, where a stock recruitment relationship was added to the model, the impact

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

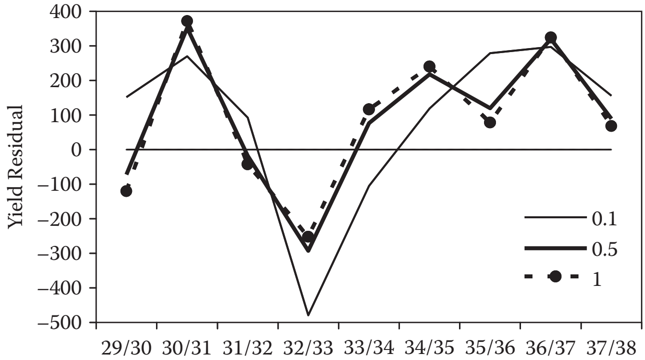

Fig ur e 12.15
The impact of different relative weights (1.0, 0.5, and 0.1) being ascribed to the sum of squares contribution from the catch-at-age (SSR_C) when the model is fitted only to catch-at-age and effort data (i.e., only the optimum stock recruitment relationship included). The fine dotted line reflects the weights of observed catches, the fine solid line is where a weighting of 0.1 is used, the thicker line is where a weighting of 0.5 is used, and the dashed line with dots has a weighting of 1.0. The heavier weighting improves the fit in the first, third, fifth, seventh, and ninth years, but the relative improvements in fit decline past a weighting of 0.5.

of the Beverton and Holt relationship was minimal. At least it could be said not to be contradictory to the other data sources.

Problems can arise when the separate contributions to the overall criterion of fit are inconsistent or contradictory (Richards, 1991; Schnute and Hilborn, 1993). Under the schema described earlier, if one had contradictory data sets, then the final outcome of the analyses would be largely influenced by the relative weights ascribed to the various components of the fit. The result will reflect the weighted average of the different conclusions deriving from the different components. As a minimum, it is a good idea to conduct a series of sensitivity trials with different sets of weightings to determine the implications of whether one source of data is more reliable than the others (Figure 12.15; Richards, 1991).

Schnute and Hilborn (1993) go somewhat further and suggest that, as with robust statistics (Huber, 1981), where individual data points that are really noninformative outliers may still have a nonzero likelihood, whole data sets may be noninformative in an analogous fashion. In effect, they recommend a more formal, structured investigation of the impacts on the quality of the overall fit of giving different emphases to the different contradictory data sets.

The issue of what weightings to use is one that will not go away and must be treated explicitly in any formal assessment. If a definite selection of relative weights is made instead of conducting a sensitivity analysis, then these weights require a formal justification. Further examples of the impacts of these weightings should be published. Further work along the lines suggested by Schnute and Hilborn (1993) may also prove helpful.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

12.3.8 Characterization of u ncertainty

As with all other stock assessment analyses, one obtains a specific optimum model fit that will have a variety of management implications. From the integrated analysis one can obtain estimates of stock biomass, fully selected fishing mortality, and average fishing mortality, along with other, possibly related, performance indicators. Although the analyses generate specific values for all of these parameters and outputs, a further step is required to obtain an indication of the level of uncertainty associated with each estimate. Depending upon how the minimization is conducted, some software will provide asymptotic standard errors around each parameter estimate. However, these standard errors rely on linear statistical theory, and in all cases will be symmetrical (unless results are transformed). Confidence intervals derived from such standard errors are recognized as being only approximate.

A better way of obtaining approximate confidence intervals or of generating likelihood profiles around model estimates would be to use either Bayesian methods or bootstrapping. As described in Chapter 3 on parameter estimation and Chapter 8 on characterizing uncertainty, Bayesian methods retain the original data and describe uncertainty by determining how well different combinations of parameters fit the available data. Bootstrapping techniques recognize that the available data are only a sample of what was possible. By generating bootstrap samples from the data, refitting the model, and collating the resulting sets of parameter estimates, it is possible to generate percentile confidence intervals. In this chapter we will only be considering bootstrap methods. See Punt and Hilborn (1997) for a description of Bayesian methods as applied to integrated analysis stock assessment models. The methods described in Chapter 8 for generating posterior distributions using the Gibbs sampler are very general, but applying such methods to a twenty-nine-parameter model in Excel would be remarkably inefficient. Even bootstrapping will be slow, but it is at least manageable (Example Box 12.11, Figure 12.16).

With catch-at-age data there are many serial correlations between age classes and years. When generating the required bootstrap samples it is best to resample the residuals from the optimum model fit and combine them with the expected catch-at-age data to form the bootstrap catch-at-age sample. Thus, the bootstrap samples (C^b) would be

$$C^b_{a,y} = \hat{C}_{a,y}\left(\frac{C_{a,y}}{\hat{C}_{a,y}}\right)^{boot} \tag{12.41}$$

where C hat is the expected catch-at-age and the residual, $\left(C_{a,y}/\hat{C}_{a,y}\right)^{boot}$, is a randomly selected residual from those available (Figure 12.16).

The bootstrapping highlights the level of uncertainty in the various parts of the analysis and permits the investigator to make stronger statements

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.11

A macro for conducting bootstraps on the catch-at-age worksheet (see Example Boxes 12.7 to 12.10 and Figure 12.16). Once you have an optimum model fit, put =F3/S3 into AR15, copy down and across to AZ23. Select AR15:AZ23, copy and save as values onto itself; these are the residuals to be bootstrapped. Copy the original data (F3:N11) into AR27:AZ35, and copy the optimum predicted catch-at-age from S3:AA11 into AR3:AZ11. Put =AR3*offset($AQ$14,trunc(rand()*9)+1,trunc(rand()*9)+1) into AR39 and copy down to row 47 and across to column AZ to create bootstrap samples (see Figure 12.16). Put =O15 into B39 and copy down to row 47, to make storing model results slightly easier. Copy the optimum parameter values and outputs B8:B47, and paste special/transpose into BE7:CR7. This will be used as the starting point for fitting each bootstrap sample to the model. The results will be pasted underneath these optimum values by the macro. It took just under two hours to conduct one thousand bootstraps on a 2.5 Ghz dual-core computer (Figures 12.17 and 12.18), when solving for twenty-nine parameters. Time how long it takes to conduct ten bootstraps on your own machine before setting it off on a marathon.

```
Sub Do_Boot()
Dim i As Integer
Application.ScreenUpdating = False
  For i = 1 To 1000
    Range("AR39:AZ47").Select    ' Replace original data
    Selection.Copy
    Range("F3").Select
    Selection.PasteSpecial Paste:=xlValues,
      Transpose:=False
    Range("BE7:CG7").Select      ' Paste in Optimal
    Selection.Copy
    Range("B8").Select
    Selection.PasteSpecial Paste:=xlValues,
      Transpose:=True
    Application.CutCopyMode = False
    SolverOk SetCell:="$B$5", MaxMinVal:=2, ValueOf:="0",
      ByChange:="$B$8:$B$36"
    SolverSolve (True)           ' Run the solver twice to
    SolverOk SetCell:="$B$5", MaxMinVal:=2, ValueOf:="0",
      ByChange:="$B$8:$B$36"
    SolverSolve (True)           ' in the solution
```

continued

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

EXAMPLE BOX 12.11 (continued)

```
      Range("B8:B47").Select
      Selection.Copy
      Range("BE8").Select
      ActiveCell.Offset(i, 0).Range("A1").Select
      Selection.PasteSpecial Paste:=xlValues,
        Transpose:=True
   Next i
   ActiveWorkbook.Save            ' just in case
   Application.ScreenUpdating = True
End Sub
```

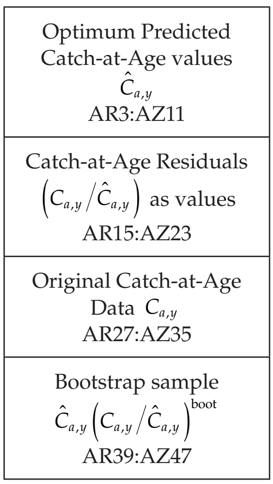

Fig ur e 12.16
Schematic map of a worksheet arrangement for conducting bootstraps on the catch-at-age data and model. The cell ranges depict where the actual values go; it is assumed that the rows and columns will be labeled appropriately. The top three matrices are filled with values; only the bottom matrix contains equations. A macro is used to run the bootstrap (Example Box 12.11).

about the management implications of the results. In this instance it also confirms the suspicions (Figure 12.14) that the estimate of the fully selected fishing mortality in the fishing year 1929–1930 is biased (Figure 12.16). The ability of the bootstrapping procedure to identify bias in the parameter estimates is a real advantage, although, as discussed in Chapter 6, once bias is detected, it is difficult to know specifically what to do about it.

In this case, it is only the first year that is exhibiting bias, and this suggests a closer consideration of the first year's data. Perhaps the effort data used are

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

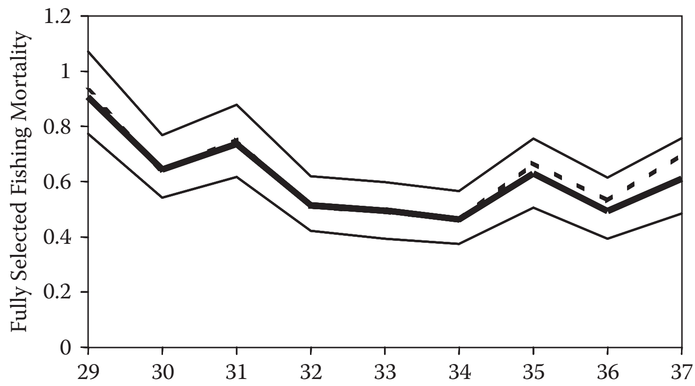

Fig ur e 12.17
Bootstrap 95% confidence intervals around the fully selected fishing mortality through the years of the fishery. The heavy line is the bootstrap average while the dotted line is the optimum fit. There is little evidence of bias except in the last three years, where the optimum fit appears to be biased high (5.9% in 35, 6.9% in 36, and 12.8% in 37). That the most bias occurs in the final year is not surprising because that year has the least information available.

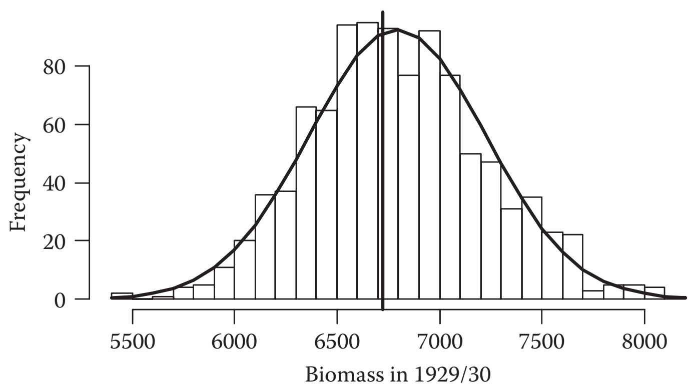

Fig ur e 12.18
Bootstrap 95 percentile confidence intervals around the stock biomass estimate for the 1929–1930 fishing year. The optimum index, using fixed optimum Beverton and Holt recruitment parameters, was 67,721*t* and the fitted normal curve has a mean of 6,794*t*. The small amount of bias early becomes greater as the years advance. Possibly, using the Ricker would be less biased; alternatively, one could solve for the recruitment parameters in each bootstrap, which would increase the variation but may reduce the bias.

more error prone than imagined. However the case may be, awareness of the problem is the first step to a better understanding of our perception of the fishery. Once again, the stock assessment model synthesizes a wide range of information and provides a more convenient and defensible statement about the status of the stock. Standard diagnostic tests, such as plotting residuals

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

and other visual indicators (Richards et al., 1997) should always be made on the analysis results.

12.3.9 Model Projections and r isk Assessment

As with surplus production models, characterization of the uncertainty in an assessment model is only the first step. Ideally, it should be possible to decide on a management strategy (be it a certain catch level or fishing mortality level), impose that on the model fishery, and project the population forward in time to determine the consequences of different strategies. This would be the basis behind a formal risk assessment (Francis, 1992). With the integrated analysis model this can be implemented relatively easily.

If a particular fully selected fishing mortality rate has been selected as the management strategy, this might also entail imposing a particular selectivity curve. However, in the following discussion we will only consider imposing a particular fishing mortality rate. Assuming the model has been optimally fitted to catch-at-age, effort, and the stock recruitment relationship has been included, the algorithm would be the following:

1. Calculate age-specific fishing mortality rates—from $s_a F_y$.
2. Calculate age-specific survivorship rates—from $S_{a,y}=\exp(-(M + s_a F_y))$.
3. Calculate the predicted numbers-at-age for age classes above the age of recruitment—from $N_{a+1,y+1}=N_{a,y}S_{a,y}$, Equation 12.30.
4. Calculate predicted numbers-at-age for recruits from the stock recruitment relationship and the biomass from r years previous using Equation 12.36, including a stochastic residual term.
5. Calculate the predicted catch-at-age as per Equation 12.32.
6. Calculate the fishery performance measures.
7. Repeat for the next year of projection.
8. Repeat a sufficient number of times to obtain summary information from the stochastic nature of the projections.

The predicted recruitment should have some intrinsic variation away from the deterministic recruitment value predicted from the spawning stock biomass from the requisite number of years prior to the recruitment year. This stochasticity could either be selected at random from the residuals available (analogous to a bootstrap sample) or, if the time series of residuals is short, should be selected at random from under a probability density function used to describe the recruitment residuals (usually a lognormal pdf would be used). What this implies is that the projections would be in the nature of a Monte Carlo simulation requiring numerous replicates to obtain the necessary summary information.

Copyright © 2011. CRC Press LLC. All rights reserved.

© 2011 by Taylor & Francis Group, LLC

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

12.4 Concluding Remarks

In this chapter we have considered both of the main analytical strategies adopted for assessing age-structured fishery data. Which approach best suits a particular situation will depend upon circumstances. The fact that integrated analyses are less stringent in their data requirements means that they are a more useful method for many fisheries in countries that do not have long traditions of collected detailed age-structured information.

Both VPA (in all its forms) and integrated analyses are large fields of endeavour with many examples in the literature and many developments not covered in this chapter. Each of the analysis strategies would form the basis of a significant book by themselves. This introduction covers a number of the important issues, but these models continue to be developed and articulated (Myers and Cadigan, 1995; Schnute and Richards, 1995; Methot, 2009).

In real-life situations there are likely to be different stocks present within a single fishery, and possibly there will be more than one fishing fleet exploiting the various stocks (e.g., trawl and nontrawl), each with their own selectivity characteristics. These would require separate treatments, which implies parallel analyses for each fleet (the fleets are separated by their relative effort, their selectivity curves, and the particular stocked fished) and for each stock (recruitment being kept separate). These multiple analyses would imply the necessity for more extensive data sets relating to catch-at-age and related effort (Punt et al., 2001).

The analyses in this chapter were conducted inside Excel. The fact that it took under two hours to conduct one thousand bootstraps in the integrated analysis means that the need for custom computer programs seems to have lessened. However, implementing a more sophisticated multifleet, multistock, catch-at-age model in Excel would be stretching the abilities of the Excel solver, although a stepped approach to solving subsets of the model would be possible. In the end, custom computer programs to conduct these stock assessments are still the optimum approach (Richards et al., 1997). The ability to mimic such models in Excel is, however, a handy double-check on the defensibility and reality of the results obtained from more sophisticated programs. Excel is remarkably good, but its limits must be recognized, and when dealing with significantly complex problems, the analyst's life is made easier by using some other software platform.

Once the basics have been absorbed, the best teacher relating to these methods is experience with different fisheries and problems. There are large numbers of options available when implementing one of these models. A good strategy is not to restrict oneself to a single model but to implement different versions of an assessment to investigate the implications of the different data sets and the different model structures.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

Appendix 12.1: Weight-at-Age Data and Optimum Fit to Catch-at-Age Model

TABLE A12.1

Average Weight in Grams of North Sea Plaice in Each Age Class Measured from Market Samples Taken in Lowestoft and Grimsby during 1929–1938

Year\Age	2+	3+	4+	5+	6+	7+	8+	9+	10+
29/30	167	190	218	270	289	392	574	665	802
30/31	136	190	257	340	441	498	582	740	961
31/32	127	152	221	349	440	503	593	675	809
32/33	132	143	182	297	421	535	641	721	827
33/34	146	165	189	251	383	528	645	787	818
34/35	157	189	202	229	277	521	711	798	819
35/36	154	178	214	280	313	383	643	789	876
36/37	149	171	198	255	338	377	467	781	885
37/38	160	173	222	310	383	490	519	624	845

Source: Data from Table 16.2 in Beverton and Holt, 1957.
Note: In the original table the weights went up to age 20, but are truncated here at age 10 to match the ageing data. In addition, the weights were given for calendar years while the ageing was for fishing years (April 1 to March 31); weights from 1929 were assumed to hold for the 1929–1930 fishing year, and equivalently for later years. Data used in Example Box 12.7.

TABLE A12.2

Parameter Values and Sum of Squared Residuals for the Optimum Fit from Example Boxes 12.7 and 12.8

Source	Value	Weight				
SSQC	7.6366	1	SSQT	8.0152115		
SSQE	0.3786	1	Natural M	0.2		
Parameter	Ln(value)					
Ln_q	–2.09828	N_30_2	10.3658		LnF_29	–0.0740
N_29_10	4.4570	N_31_2	9.2878		LnF_30	–0.4475
N_29_9	5.0807	N_32_2	9.2766		LnF_31	–0.2869
N_29_8	6.0900	N_33_2	8.9197		LnF_32	–0.6782
N_29_7	7.3389	N_34_2	9.3640		LnF_33	–0.7126
N_29_6	6.9909	N_35_2	9.1696		LnF_34	–0.7796
N_29_5	7.7294	N_36_2	8.9735		LnF_35	–0.4135
N_29_4	8.9176	N_37_2	9.6004		LnF_36	–0.6380
N_29_3	8.7447	Sel50	1.2232		LnF_37	–0.3885
N_29_2	9.3607	Sel95	1.4785			

Note: SSQC is the sum of squared residuals from the catch-at-age data, SSQE is from the fully selected fishing mortality rate comparison, and SSQT is simply the sum of both sources weighted by their respective weights. Natural mortality is assumed to be constant.

© 2011 by Taylor & Francis Group, LLC

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.

Copyright © 2011. CRC Press LLC. All rights reserved.

Haddon, Malcolm. Modelling and Quantitative Methods in Fisheries, CRC Press LLC, 2011. ProQuest Ebook Central,
http://ebookcentral.proquest.com/lib/ucsc/detail.action?docID=1648249.
Created from ucsc on 2021-05-04 16:33:06.