

A General Theory for Analysis of Catch and Effort Data

Jon Schnute

Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, B.C.V9R 5K6

Schnute, J. 1985. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.* 42: 414–429.

This paper presents a general theory for analysis of catch and effort data from a fishery. Almost all previous methods are shown to be special cases, including those of Schaefer, Pella and Tomlinson, Schnute, and Deriso, as well as the stock reduction analysis technique of Kimura and Tagart and Kimura, Balsiger, and Ito. Like that of Deriso, the theory here is based on natural equations for an age structured population. However, instead of a fixed single model, this paper gives a general model that can be tailored to any particular fishery. The problem of determining the appropriate special case is conceptually identical to the model identification problem described by Box and Jenkins in the context of time series analysis. Identification necessarily begins with a suitable class of models. This paper defines such a class, unique to fisheries, complete with mathematical proofs and biological explanations of all important equations.

Le présent article porte sur une théorie générale pour l'analyse des données sur les prises et l'effort. L'auteur démontre que presque toutes les anciennes méthodes sont des cas spéciaux, y compris celles de Schaefer, de Pella et Tomlinson, de Schnute et de Deriso, ainsi que la technique d'analyse de la réduction des stocks de Kimura et Tagart et de Kimura, Balsiger et Ito. Comme celle de Deriso, la présente théorie est basée sur des équations naturelles pour une population à structure d'âge. Toutefois, au lieu d'un seul modèle fixe, le présent article décrit un modèle général qui peut être adapté à une pêche donnée. La détermination du cas spécial approprié pose un problème essentiellement identique au problème de détermination d'un modèle que mentionnent Box et Jenkins relativement à l'analyse de séries chronologiques. Le processus commence forcément avec une classe adéquate de modèles. Le présent article définit une telle classe, particulière aux pêches, avec les preuves mathématiques et les explications biologiques de toutes les équations importantes.

Received June 1, 1984

Accepted November 20, 1984

(J7816)

Reçu le 1 juin 1984

Accepté le 20 novembre 1984

A fundamental problem of fishery management is to deduce the state of a stock from the historical record of catches and fishing efforts. The literature contains numerous proposed methods for solving this problem, including those of Baranov (1925), Thompson and Bell (1934), Schaefer (1954, 1957), Pella and Tomlinson (1969), Walter (1973), and Schnute (1977). Concurrent with these developments, various critics have debated the validity of the results. For example, Skud (1975) documented a lengthy debate prompted by Burkenroad's (1948) claim that the fishery was not the dominant influence on halibut (*Hippoglossus stenolepis*) stocks. In a more theoretical vein, Uhler (1980), Ludwig and Walters (1981), Walters and Ludwig (1981), and Ludwig and Hilborn (1983) pointed out problems with the statistical analysis of models based on data measured with error and suggested various methods to overcome these problems.

Deriso (1980) gave a remarkably fresh approach to this basic problem by showing that a composite model for catch and effort data can be derived mathematically from an age structured model of the Beverton–Holt type (Beverton and Holt 1957). Since fish populations, at least in northern and temperate climates, are composed of distinct annual cohorts, Deriso's result ties together biology and theory in a very neat way. Practitioners became enthusiastic; for example, Walters (1980) publically complimented Deriso for "what may be the most

important contribution of fisheries population theory in the last two decades." As people began to apply the method, however, enthusiasm waned somewhat. For example, Roff (1983) investigated a simple ad hoc model which gave predictions as good or better than the methods of either Deriso (1980) or Schnute (1977).

Even if Deriso's method performs poorly in practice, it has considerable value as a theoretical and conceptual tool. Earlier models, such as those of Schaefer (1957) and Schnute (1977), are based on theoretical properties of how a population limited in size by its environment ought to perform. Models of the Deriso type have a more solid footing. Among other things, they reflect the undeniable reality that the population consists of cohorts that get one year older each year. Thus, even if they serve only as a guide to intuition, they have value to the practitioner. Part of my intent in this paper is to enlarge the framework of understanding such models so that they serve as better tools for conceptual thinking.

The diverse and conflicting literature on this subject raises a fundamental question: "Which model is right?" In fact, fisheries scientists are not the first to ask such questions. Catch and effort data fall into the general class of time series data, and statisticians have long pondered the problem of determining useful time series models. Box and Jenkins (1970) not only proposed a very general model, but also described a philosophy

for tackling the fundamental question just cited. Their strategy (1970, p. 18–19) involves four steps, paraphrased below:

Step 1. From the interaction of theory and practice, postulate a useful class of models appropriate to the situation. A typical model in the class may potentially involve a large number of parameters.

Step 2. Use data and a knowledge of the system to identify a particular tentative model from the class. This involves restricting the model parameters in a meaningful way.

Step 3. Estimate the parameters in the tentative model of step 2.

Step 4. Perform diagnostic checking. Is the tentative model adequate? If not, return to step 2 and revise accordingly. If so, use the model for forecasting and control.

Much of the debate on fisheries models stems from a failure to come to terms with step 1. The important concept here is a *class* of models. If we are to decide which model is right, then we must be able to compare the candidates; consequently, the candidates themselves must all be special cases of some general class. For example, Box and Jenkins (1970) proposed the class of ARMA (autoregressive moving average) models for time series. Their step 2 consists in deciding how many lagged terms should be included in the autoregression and moving average. Revision in step 4 involves either increasing or reducing the number of terms in the model. At the end of the process, all models entertained are comparable because they fit into the unified class postulated in step 1.

By contrast with the situation for ARMA models, consider, for example, the models of Schnute (1977), Deriso (1980), and Roff (1983). Schnute's model is based on one theory, Deriso's on another theory, and Roff's on an ad hoc assumption. To compare them is, in the proverbial sense, to compare apples and oranges. Of the three models, however, Deriso's clearly best follows the spirit of step 1 above. More than the others, it is based on biological theory and practice; consequently, it is most likely to be useful for making decisions related to fish biology.

In this paper, I derive a class of models for analysing catch and effort data, based on assumptions of the Deriso type. My main object is to make the model building strategy of Box and Jenkins (1970) at least potentially achievable by completing the first step: postulating a useful class of models. I do not believe that the general ARMA model is appropriate in this context because it is not derived from biological theory. Furthermore, ARMA models are linear, although recruitment is an inherently nonlinear process (otherwise populations would be either infinite or extinct). The biologically motivated class described here turns out to include the models of Schaefer (1957), Pella and Tomlinson (1969), Schnute (1977), Deriso (1980), and Roff (1983) as special cases. Consequently, it becomes possible to compare them, apples to apples, as it were. Furthermore, many of the statistical problems described by various authors can be precisely identified in this new formulation.

In summary, the ideas given here have practical implications for the whole technique of fishery modelling. The modeller does not simply apply a universal model for all catch and effort data. Instead, he or she (step 1) uses a general framework for constructing such models, with enough flexibility to accommodate a wide range of biological assumptions. The modeller's task is then (step 2) to identify those particular assumptions that apply to the case at hand, (step 3) to fit these into the general framework, and thus to build an appropriate model for the stock in question. Since biological assumptions are often tenuous, it

may typically be reasonable (step 4) to build several related models and to compare the results. The advantage of a universal framework is that the various results are then easily comparable, since all can be considered special cases of a larger whole.

1. The Model

1A. Notation and Assumptions

A consistent notation is important for stating the model and its consequences properly, particularly since the model involves numerous mathematical objects. There are two basic measures of time (in years): age a and year t . It is assumed that there is a fixed age k at which fish are recruited to the fishery. Deriso (1980) called this "knife-edged" recruitment. In order not to detract from a few central features of the model, I have devoted this paper entirely to the knife-edged case, although a more general "incomplete recruitment" model similar to that discussed by Deriso could also be developed. Notice that a fish of age a in year t was born in year $t - a$ and recruited in year $t - a + k$. Some minor clarification may be needed here because fish (unlike humans) are not usually aged from their birthdate; furthermore, fish age is measured from the time of birth as an egg. In any particular fishery, there is a convention that defines the "start of the year." It may be January 1, but it may also be the time when the fishing season begins in the spring. A fish born the previous fall, for example, automatically becomes 1 yr old at this time, even though chronologically it is much less than 1 yr old. As a result of this convention, a fish of age a in year t has parents belonging to the mature stock in year $t - a$. In particular, the mature stock in year $t - k$ determines recruitment in year t .

The model involves three main functions of a and t : the total number of available fish (the population) N_{at} aged a at the start of year t , the number C_{at} caught at age a during year t , and the weight w_{at} of a single fish aged a in year t . Corresponding to the numbers N_{at} and C_{at} are the biomasses

$$(1.1) \quad \begin{aligned} N_{at}^* &= w_{at} N_{at} \\ C_{at}^* &= w_{at} C_{at}. \end{aligned}$$

Throughout this paper, a superscript asterisk distinguishes biomasses of fish from numbers of fish. The model also involves two functions of t alone: the recruitment R_t of fish of age k at the start of year t , and the fishing effort E_t during year t . Again, there is a related biomass

$$(1.2) \quad R_t^* = w_{kt} R_t.$$

Furthermore, corresponding to age structured populations, are the totals

$$(1.3) \quad \begin{aligned} N_t &= \sum_{a=k}^{\infty} N_{at} \\ C_t &= \sum_{a=k}^{\infty} C_{at} \end{aligned}$$

and related biomass totals

$$(1.4) \quad \begin{aligned} N_t^* &= \sum_{a=k}^{\infty} N_{at}^* \\ C_t^* &= \sum_{a=k}^{\infty} C_{at}^*. \end{aligned}$$

Definitions (1.3)–(1.4) reflect a convention used throughout this paper: when the first subscript a is dropped from a doubly

subscripted quantity such as N_{at} , the resulting function of t refers to the sum over recruited ages.

In addition to the above functions of a and t , there are numerous parameters that drive the model. One of these is the recruitment (age k) weight V_t associated with the cohort born in year t and recruited in year $t + k$. This same cohort is assumed to have an extrapolated prerecruitment (age $k - 1$) weight v_t , and a Ford growth coefficient ρ . By definition of v_t and V_t :

$$(1.5) \quad \begin{aligned} v_t &= w_{k-1, t+k-1} \\ V_t &= w_{k, t+k} \end{aligned}$$

In (1.5) the prerecruitment weight v_t should be interpreted as a parameter, rather than the actual weight at age $k - 1$; for this reason it is called an "extrapolated" weight. The model simply uses v_t , V_t , and ρ as three parameters to define the growth curve for ages $a \geq k$. Typically, fish younger than age k have weights that do not conform to the mature growth curve. The model also involves a mortality parameter M_t and a catchability parameter q_t associated with natural mortality and fishing, respectively. These allow the definition of three fractions: natural survival σ_t , fishing survival ϕ_t , and total survival τ_t . Finally, there are three parameters α_t , β , and γ related to recruitment. This notation is summarized in Tables 1 and 2. Various parameters, such as M_t , are presumed to depend on t , rather than being constants, to allow stochastic behavior to be introduced systematically in the model. For example, variation in natural mortality can be accounted for by assuming that M_t is a random variable with mean M . Similarly, q_t , V_t , and α_t can be used to introduce random fluctuations in catchability, growth, and recruitment, respectively.

Using the above notation, the model consists of eight assumptions:

$$(1.6) \quad N_{a+1, t+1} = \tau_t N_{at}$$

$$(1.7) \quad C_{at} = [1 - \phi_t] N_{at}$$

$$(1.8) \quad N_{kt} = R_t \{N_{t-k}^* - C_{t-k}^*\}$$

$$(1.9) \quad w_{a+1, t+1} - w_{at} = \rho(w_{at} - w_{a-1, t-1})$$

$$(1.10) \quad \tau_t = \sigma_t \phi_t$$

$$(1.11a) \quad \sigma_t = \exp(-M_t)$$

$$(1.12a) \quad \phi_t = \exp(-q_t E_t)$$

$$(1.13a) \quad R_t \{S\} = \alpha_t S(1 - \beta \gamma S)^{1/\gamma}$$

The first three equations describe fish population dynamics, and the fourth describes growth dynamics. The next three give the particulars of survival from total, natural, and fishing mortalities, respectively. Finally, the last equation details the recruitment process. Equations (1.6), (1.7), and (1.9) all apply only to recruited ages $a \geq k$, with the understanding in (1.9) that, when $a = k$, the prerecruitment weight $w_{k-1, t-1}$ is extrapolated. In (1.8) and (1.13a), braces designate functional dependence; thus " $R_t \{S\}$ " means that R_t is a function of S .

The model is constructed so that the last three equations can readily be modified without altering some of the main conclusions. For example, three alternatives to the pair (1.11a)–(1.12a) are

$$(1.11b) \quad \sigma_t = \frac{M_t + q_t E_t}{M_t \exp(M_t + q_t E_t) + q_t E_t}$$

$$(1.12b) \quad \phi_t = \frac{M_t + q_t E_t \exp(-M_t - q_t E_t)}{M_t + q_t E_t}$$

TABLE 1. Notation for the primary quantities that define the model (1.6)–(1.13). A dimensionless quantity is indicated by a dash in the units column.

Notation	Meaning	Typical units
<i>Time</i>		
a	Fish age	yr
t	Year of the fishery	yr
<i>Functions</i>		
N_{at}	Population of age a fish at the start of year t	Fish
C_{at}	Catch of age a fish during year t	Fish
w_{at}	Weight of a fish at age a in year t	kg
R_t	Recruitment at the start of year t	Fish
E_t	Fishing effort during year t	Boat-days
<i>Parameters</i>		
k	Age of recruitment	yr
M_t	Instantaneous natural mortality in year t	—
q_t	Catchability in year t	1/(boat-days)
v_t	Prerecruitment weight of a fish born in year t	kg
V_t	Recruitment weight of a fish born in year t	kg
ρ	Ford's growth coefficient	—
α_t	Recruitment productivity parameter	Fish/kg
β	Recruitment optimality parameter	1/kg
γ	Recruitment limitation parameter	—

or

$$(1.11c) \quad \sigma_t = \{1 + \exp(q_t E_t)[\exp(M_t) - 1]\}^{-1}$$

$$(1.12c) \quad \phi_t = 1 - \exp(-M_t)[1 - \exp(-q_t E_t)]$$

or

$$(1.11d) \quad \sigma_t = \exp(-M_t)$$

$$(1.12d) \quad \phi_t = 1 - q_t E_t$$

An alternative to (2.13a) is

$$(1.13b) \quad R_t = R$$

where R is a constant. The significance of each model assumption, including the various alternatives, is described in the paragraphs following.

1B. Population Dynamics

The first assumption (1.6) summarizes the process of aging one year. For age $a \geq k$, the cohort population N_{at} at the start of year t is simply multiplied by the total survival fraction τ_t to obtain number of fish remaining in this cohort at the start of the next year. Similarly, if ϕ_t represents the fraction of the population that survives fishing, then (1.7) expresses catch C_{at} as the complementary fraction $1 - \phi_t$ of the cohort population. The most important aspect of both these assumptions is that survival does not depend on age after the recruitment age k . This will be true if both natural and fishing mortality are age-independent. Although constant natural mortality is a rather common assumption, the requirement that catchability should not vary with age is more severe. As mentioned earlier, the model assumes that fishing gear has knife-edged selectivity for recruited versus nonrecruited fish. Thus, the model does not include a more general selection ogive (Beverton and Holt 1957, p. 75–82; Fournier and Archibald 1982, p. 1198, equations 2.1 and 2.2), a restriction that is perhaps reasonable for the model's intended use on catch data without age information. At the price of

TABLE 2. Notation defined in terms of the model's primary quantities listed in Table 1. The number of the defining equation from the text is shown, along with typical units of measurement.

Notation	Equation	Meaning	Units
<i>Functions</i>			
N_{at}^*	(1.1)	Population biomass at age a , start of year t	kg
C_{at}^*	(1.1)	Biomass at age a , caught in year t	kg
N_t	(1.3)	Total available population, start of year t	Fish
N_t^*	(1.4)	Total available biomass, start of year t	kg
C_t	(1.3)	Number of fish caught during year t	Fish
C_t^*	(1.4)	Biomass of fish caught during year t	kg
R_t^*	(1.2)	Biomass of fish recruited at the start of year t	kg
<i>Parameters</i>			
ϕ_t	(1.12)	Survival from fishing mortality in year t	—
σ_t	(1.11)	Survival from natural mortality during year t	—
τ_t	(1.10)	Survival from all mortality during year t	—
W_t	(1.18)	Asymptotic weight of a fish born in year t	kg
K	(1.19)	Brody's growth coefficient	1/yr
$a_{0,t}$	(1.20)	Extrapolated age of 0 weight for birth year t	yr
R_m	(1.31)	Maximum possible recruitment	Fish
S_m	(1.31)	Stock for maximum recruitment	Fish
S_0	(1.32)	Carrying capacity	Fish

greater mathematical complexity, this assumption can be partly avoided (Deriso 1980, p. 271–273), but for reasons cited earlier, incomplete recruitment is not considered here.

The third equation (1.8) describes recruitment of age k fish in year t as a function of the post-fishing biomass in year $t - k$. The particular form of the function $R_t\{S\}$, where S is the stock biomass, is given in (1.13), one of the flexible assumptions identified earlier. The most important aspects of (1.8) are, first, that recruitment occurs at age k and, second, that recruitment depends on available biomass. If, for example, recruitment were to depend instead on available population (remove the two asterisks from (1.8)), then some of the model's main conclusions would change, as explained later.

1C. Growth

The fourth model assumption (1.9) is essentially Ford's (1933) growth model (Ricker 1975, p. 222, equation 9.17). Ricker refers to ρ as "Ford's growth coefficient." Typically, ρ is a fraction, $0 < \rho < 1$; thus, (1.9) states that the annual weight increment decreases by the factor ρ each year. Values of $\rho \geq 1$ are also possible, and the growth increment is constant when $\rho = 1$. When $\rho \neq 1$, it follows that

$$(1.14) \quad w_{at} = v_{t-a} + (V_{t-a} - v_{t-a}) \frac{1 - \rho^{1+a-k}}{1 - \rho}.$$

This is obtained by solving the difference equation (1.9), using the initial conditions (1.5) rewritten with different indices as

$$(1.15) \quad \begin{aligned} w_{k-1,t-1} &= v_{t-k} \\ w_{kt} &= V_{t-k}. \end{aligned}$$

Schnute and Fournier (1980, appendix A) give a formal derivation. The reader can verify that (1.14) satisfies both conditions (1.15), and a proof also follows by mathematical induction. Consistent with earlier discussion, (1.14) applies only to recruited ages $a \geq k$; the prerecruitment weight v_{t-a} is merely a convenient parameter.

The growth equation (1.14) can be written in numerous

alternative forms, using different choices of parameters. Two such alternatives are

$$(1.16) \quad w_{at} = W_{t-a} - (W_{t-a} - V_{t-a})\rho^{a-k}$$

$$(1.17) \quad w_{at} = W_{t-a}[1 - e^{-K(a-a_{0,t-a})}]$$

where

$$(1.18) \quad W_{t-a} = (V_{t-a} - \rho v_{t-a}) / (1 - \rho)$$

$$(1.19) \quad K = -\log \rho$$

$$(1.20) \quad a_{0,t-a} = k - 1 - \frac{\log [(V_{t-a} - v_{t-a}) / (V_{t-a} - \rho v_{t-a})]}{\log \rho}.$$

The parameters in (1.18)–(1.20) are, respectively, the extrapolated asymptotic weight W of a fish born in year $t - a$, Brody's growth coefficient K (Ricker 1975, p. 221), and the extrapolated age a_0 when a fish born in year $t - a$ has weight 0.

The three growth curves (1.14), (1.16), and (1.17) are equivalent; the following table lists the parameters in each:

Equation	Parameters
(1.14)	v, V, ρ
(1.16)	W, V, ρ
(1.17)	W, K, a_0

Perhaps the most common form of the curve is (1.17), the von Bertalanffy equation (Ricker 1975, p. 221, equation 9.9). However, the form (1.14) is preferable here for two reasons. First, equations of the general theory below take the simplest form when (1.14) is used. Second, (1.14) is closely related to the growth model adopted by Deriso (1980, equation 1), namely

$$(1.21) \quad w_{at} = w_{k,t-a+k} \frac{1 - \rho^{1+a-k}}{1 - \rho}.$$

Since $w_{k,t-a+k} = V_{t-a}$ by (1.5), this model is the same as (1.14) with the restriction

$$(1.22) \quad v = 0.$$

From (1.20), (1.22) is equivalent to the requirement

$$(1.23) \quad a_0 = k - 1$$

in the von Bertalanffy equation (1.17). Thus, Deriso's growth model extrapolates back to weight 0 at the prerecruit age $k - 1$, as can be seen directly by substituting $a = k - 1$ in (1.21).

Since growth is just one component of the general theory, Deriso quite reasonably decided to impose (1.22) (or, equivalently, (1.23)) for the sake of parsimony. In such a large model, the number of parameters quickly gets out of hand. Also, without (1.22), it seems impossible to complete the formulation of the incomplete recruitment model. Here, I am avoiding the latter problem, and I do not assume (1.22). This gives the general theory somewhat better symmetry, as shown below. Also, in at least some cases, (1.22) may be too restrictive. For example, consider a Pacific herring (*Clupea harengus pallasii*) sample taken from Yellow Point on the east coast of Vancouver Island, B.C., in March 1982. Fish are recruited dominantly at age $k = 3$ and mean weights (grams) for ages 3–8 turn out to be 95.27, 113.45, 130.99, 146.19, 170.43, and 175.38, respectively (Stocker et al. 1982, p. 102). Parameter estimates for a least squares fit to these data are listed below:

Model	v	V	ρ
(1.14)	73.4	94.5	0.925
(1.21)	0.0	76.9	0.553

Figure 1 shows the two growth curves in relation to the data. The constraint (1.23) that the curve should pass through 0 at age 2 (i.e. age $k - 1$) makes it difficult for the model (1.21) to capture the nearly linear trend in the data, particularly for the dominant ages 3–6. By contrast, notice that ρ is nearly 1 in the model (1.14).

1D. Mortality Details

The fifth model assumption (1.10) states that the total survival fraction is the product of survivals from fishing and natural mortality. Given the earlier equations (1.6)–(1.7) involving τ_t and ϕ_t , this apparently reasonable assumption really amounts to a definition of σ_t . A short algebraic calculation (Appendix A) based on (1.6), (1.7), and (1.10) shows that

$$(1.24) \quad N_{a+1,t+1} = \sigma_t(N_{at} - C_{at});$$

that is, σ_t represents the fraction of fish *not caught* in the fishery that survives to the next year. Fishing and natural mortality may occur concurrently; however, from (1.7), ϕ_t is the fraction of the original population that survives the fishery, and from (1.24), σ_t is the fraction of the remaining population that survives natural mortality.

The model is flexible as to the precise details whereby fishing and natural mortality occur. If fishing occurs first, in a pulse immediately following recruitment, then (1.11a)–(1.12a) describe the relevant survivals based on mortality M_t , catchability q_t , and effort E_t . If fishing and natural mortality take place simultaneously throughout the year (as assumed, for example, by Beverton and Holt 1957; Ricker 1975; Fournier and Archibald 1982; Kimura and Tagart 1982), then (1.11b)–(1.12b) apply. If pulse fishing occurs at the end of the year, just prior to recruitment, then (1.11c)–(1.12c) are correct. Appendix A justifies all three of these formula pairs. They have the common feature that the product $\sigma_t \phi_t$, the total mortality, in each case is

$$(1.25) \quad \tau_t = e^{-(M_t + q_t E_t)}.$$

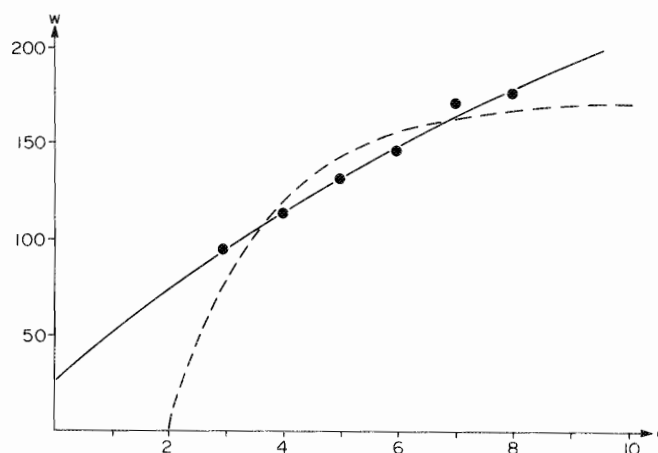


FIG. 1. Two growth curves for Pacific herring (*C. harengus pallasii*) from least-squares fits to data listed in the text. Recruitment is at age $k = 3$. Data points are shown as enlarged dots. The solid curve is based on the three-parameter model (1.14) with $(v, V, \rho) = (73.4 \text{ g}, 94.5 \text{ g}, 0.925)$. The broken curve is Deriso's two-parameter model (1.21) with $(V, \rho) = (76.9 \text{ g}, 0.553)$, constrained to weight 0 at the prerecruitment age 2.

Deriso (1980) implicitly assumed none of the above three options, but rather (1.11d)–(1.12d). The expression (1.12d) for ϕ_t has the advantage that, from (1.7), it implies

$$C_{at}/E_t = q_t N_{at};$$

that is, catch per unit effort (CPUE) is proportional to the population. However, unlike (1.12a), (1.12d) also has the rather unreasonable feature that a finite effort, namely

$$E_t = 1/q_t,$$

will remove the whole population ($\phi_t = 0$). Consequently, (1.12d) is avoided throughout the rest of this paper in favor of one of the other assumptions. This means that CPUE is not strictly an index of population in the discussion here, a limitation that turns out to be irrelevant for practical application of the model. Indeed, the model involves only catch and effort data, CPUE is never used directly. Notice, incidentally, that (1.12a) and (1.12d) are approximately the same for low levels of effort E_t .

1E. Recruitment Details

The final model assumption (1.13a) is essentially the recruitment function proposed by Deriso (1980); however, it is written here with somewhat different parameters to allow greater flexibility. Specific values of the parameter γ determine four cases of special interest:

$$(1.26) \quad \gamma = -\infty: \quad R_t\{S\} = \alpha_t S$$

$$(1.27) \quad \gamma = -1: \quad R_t\{S\} = \alpha_t S / (1 + \beta S)$$

$$(1.28) \quad \gamma = 0: \quad R_t\{S\} = \alpha_t S e^{-\beta S}$$

$$(1.29) \quad \gamma = 1: \quad R_t\{S\} = \alpha_t S (1 - \beta S).$$

Each case above is either a valid mathematical limit for (1.13a) or simply the result of substituting the given value of γ . The first case (1.26) is a constant productivity model; α_t recruits are produced for each unit of stock biomass, regardless of stock size. The result (1.26) also follows from (1.13a) by setting $\beta = 0$. The next three cases (1.27)–(1.29) correspond to classical recruitment models, according to the following table:

Equation	Reference
(1.27)	Beverton and Holt 1957 (p. 49, equation 6.10)
(1.28)	Ricker 1954, 1958 (p. 238, equation 11.7)
(1.29)	Schaefer 1954 (bottom of p. 34)

Because of his slightly different expression for $R_i\{S\}$, Deriso (1980) noted the limiting versions (1.27)–(1.28), but he did not mention (1.29) as an additional possibility.

To complete the discussion of (1.13a), it is convenient to drop the subscripts:

$$(1.30) \quad R = \alpha S(1 - \beta\gamma S)^{1/\gamma}.$$

The parameters α and β should always be positive, although γ can have any sign. The curve (1.30) in the SR plane always passes through the origin (0, 0). When $\gamma > -1$, it has a maximum point (S_m, R_m) with coordinates (Appendix B)

$$(1.31) \quad \begin{aligned} S_m &= [\beta(1 + \gamma)]^{-1} \\ R_m &= \alpha\beta^{-1}(1 + \gamma)^{-(1+\gamma)/\gamma}. \end{aligned}$$

Furthermore, when $\gamma > 0$, the curve also has a second point S_0 (in addition to $S = 0$) where $R = 0$ (Appendix B):

$$(1.32) \quad S_0 = (\beta\gamma)^{-1}.$$

The coordinates (1.31)–(1.32) can be used as new parameters to rewrite (1.30) in the following two forms (Appendix B):

$$(1.33) \quad R = R_m(S/S_m)[1 + \gamma - \gamma(S/S_m)]^{1/\gamma}$$

$$(1.34) \quad R = (1 + \gamma)^{(1+\gamma)/\gamma} \gamma^{-1} R_m(S/S_0)[(S_0 - S)/S_0]^{1/\gamma}.$$

Figures 2–4 represent recruitment curves for the three different formulations (1.30), (1.33), and (1.34), respectively, for various values of γ . Figure 2 includes all four cases (1.26)–(1.29); notice that the curves do not have a maximum for $\gamma < -1$. Figure 3 illustrates the role of the maximum for $\gamma > -1$ and shows that the intercept S_0 does not exist for $\gamma \leq 0$. Finally, Fig. 4 includes only curves with a maximum and two R -intercepts when $\gamma > 0$.

In Figure 3, the shape of the particular curve with $\gamma = -0.8$ suggests the mathematical fact that, as γ tends to -1 in (1.33), R takes the limiting form

$$(1.35) \quad R = R_m$$

for $S > 0$, that is, the constant recruitment model (1.13b). Thus, (1.13b) is really just another special case of (1.13a) when the latter is written in the form (1.33). Notice, however, that the limit (1.27) as γ tends to -1 with α and β fixed is different from the limit (1.35) as γ tends to -1 with S_m and R_m fixed.

From (1.31)–(1.32), it follows that

$$(1.36) \quad S_m = [\gamma/(1 + \gamma)]S_0.$$

When $\gamma = 1$, the factor $\gamma/(1 + \gamma)$ is $1/2$, and the high point of the curve is midway between the intercepts at 0 and S_0 . This fact corresponds to the symmetry of the Schaefer curve. In that context the quantity S_0 is called the “carrying capacity” (Schnute 1977), and the stock size S_m for optimum recruitment is well-known to be half the carrying capacity. To avoid this possibly unrealistic symmetry, Pella and Tomlinson (1969) added a parameter to the Schaefer model to produce a family of curves similar to that shown in Fig. 4. Here, the factor $\gamma/(1 + \gamma)$ serves essentially the same purpose, as shown in (1.36).

The parameters α , β , and γ are given names in Table 1. The term “productivity parameter” for α comes from the limiting case (1.26), as well as the fact that (1.26) is an approximate

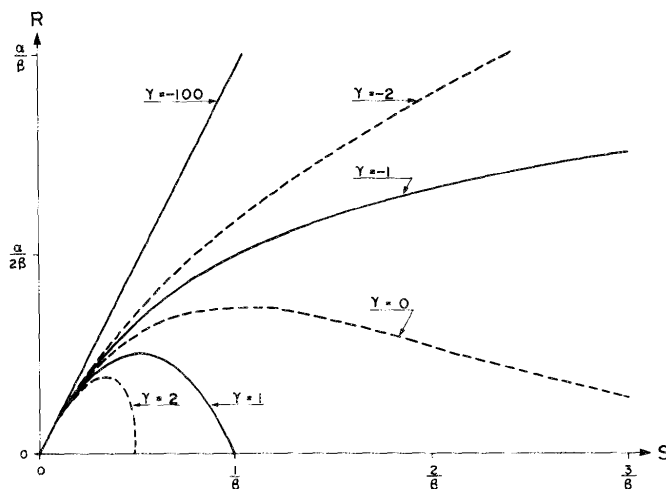


FIG. 2. Plot of recruitment R vs. stock S for six values of γ in equation (1.30) with fixed α and β . Cases of special interest include the models of Beverton and Holt ($\gamma = -1$), Ricker ($\gamma = 0$), and Schaefer ($\gamma = 1$). A large negative value of γ (e.g. $\gamma = -100$) gives essentially the constant productivity model $R = \alpha S$.

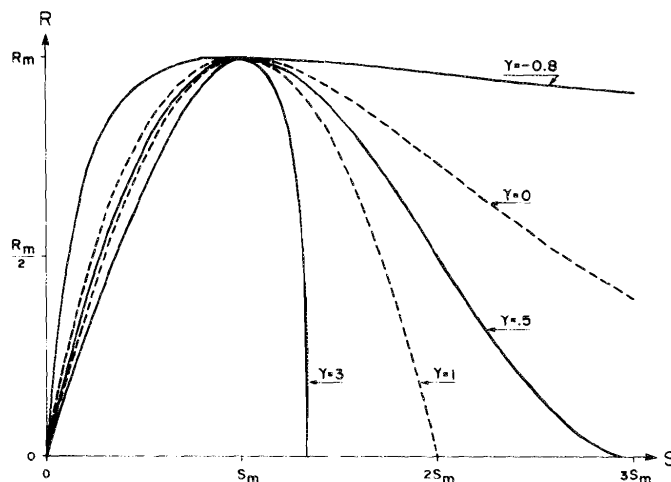


FIG. 3. Plot of recruitment R vs. stock S for five values of γ in equation (1.33) with a fixed maximum point (S_m, R_m). The models of Ricker ($\gamma = 0$) and Schaefer ($\gamma = 1$) are again special cases. As γ approaches -1 (e.g. $\gamma = -0.8$), R tends to the constant value R_m with a sudden jump from 0 at $S = 0$. When $\gamma > 0$, there is an intercept $S_0 > 0$ related to S_m by (1.36).

version of (1.30) when S is small. Since the coordinates S_m and R_m of the optimal point (1.31) are both inversely proportional to β , this parameter is termed an “optimality parameter.” Finally, γ is called a “recruitment limitation parameter” for its role in determining whether or not S_0 and R_m exist, that is, whether or not arbitrary stock and recruitment sizes are possible. The term “skewness parameter” would also be appropriate for γ , due to its significance in Fig. 4 and equation (1.36) above. The curve is symmetric when $\gamma = 1$, skewed left when $0 < \gamma < 1$, and skewed right when $\gamma > 1$.

2. Difference Equations

The previous section completely describes an age structured model for a fish stock. It consists of the eight assumptions (1.6)–(1.13). These are quite general, although their main limitations are, first, that fishing and natural mortalities are

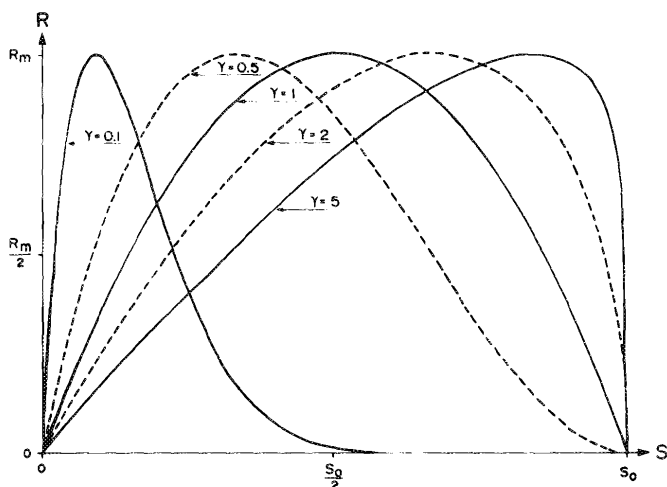


FIG. 4. Plot of recruitment R vs. stock S for five values of γ in equation (1.34) with fixed S_0 and R_m . The Schaefer case ($\gamma = 1$) is symmetric about $S_0/2$. Otherwise the curve is skewed left for $\gamma < 1$ and right for $\gamma > 1$.

presumed age-independent and, second, that recruitment to the fishery occurs at a distinct age k . To apply this model to total annual catch and effort data, it is necessary to derive equations that involve these data only. Such equations are called difference equations, because they typically involve changes, or differences, in functions of time t . In particular, this section proves mathematically and explains biologically the following eight consequences of the first five assumptions (1.6)–(1.10):

- (2.1) $N_{a+1,t+1} - \tau_t N_{at} = 0$
- (2.2) $N_{a+1,t+1}^* - \tau_t N_{at}^* = \rho \tau_t (N_{at}^* - \tau_{t-1} N_{a-1,t-1}^*)$
- (2.3) $N_{t+1} - \tau_t N_t - R_{t+1} = 0$
- (2.4) $N_{t+1}^* - \tau_t N_t^* - R_{t+1}^* = \rho \tau_t (N_t^* - \tau_{t-1} N_{t-1}^* - v_{t-k} R_t^* / V_{t-k})$
- (2.5) $N_t = C_t / (1 - \phi_t)$
- (2.6) $N_t^* = C_t^* / (1 - \phi_t)$
- (2.7) $N_{t+1}^* = (1 + \rho) \tau_t N_t^* - \rho \tau_t \tau_{t-1} N_{t-1}^* + V_{t+1-k} R_{t+1} \{ \phi_{t+1-k} N_{t+1-k}^* \} - \rho \tau_t v_{t-k} R_t \{ \phi_{t-k} N_{t-k}^* \}$
- (2.8) $C_{t+1}^* / (1 - \phi_{t+1}) = (1 + \rho) \tau_t C_t^* / (1 - \phi_t) - \rho \tau_t \tau_{t-1} C_{t-1}^* / (1 - \phi_{t-1}) + V_{t+1-k} R_{t+1} \{ \phi_{t+1-k} C_{t+1-k}^* / (1 - \phi_{t+1-k}) \} - \rho \tau_t v_{t-k} R_t \{ \phi_{t-k} C_{t-k}^* / (1 - \phi_{t-k}) \}$

Equation (2.1) is a minor variation of assumption (1.6). It can be regarded as a conservation statement; those fish at age a that survive year t are precisely the fish at age $a + 1$ that begin the next year $t + 1$. Similarly, (2.2) is a related conservation principle for biomass; however, the principle is not quite so simple because, due to growth, the weights w_{at} and $w_{a+1,t+1}$ are not the same. To prove (2.2), notice first from (1.6) that for $a \geq k + 1$

$$N_{a+1,t+1} = \tau_t N_{at} = \tau_t \tau_{t-1} N_{a-1,t-1}.$$

Subscripts of N in each of these three equal terms match the various subscripts of w in (1.9). Multiplying the above terms by

their counterparts in (1.9) gives

$$N_{a+1,t+1}^* - \tau_t N_{at}^* = \rho (\tau_t N_{at}^* - \tau_t \tau_{t-1} N_{a-1,t-1}^*),$$

that is, (2.2). Notice that (2.2) and (1.9) are very similar mathematically, particularly when τ_t and τ_{t-1} are both 1. Biologically, (2.2) expresses conservation, not of biomass itself, but rather of biomass increments.

Equation (2.3) is similar to (2.1), except that (2.3) describes the entire population, rather than just one age class. Again, (2.3) is a conservation principle; the population in year $t + 1$ consists of survivors and recruits. To prove (2.3), sum (2.1) over ages a from k to ∞ , and notice that

$$\sum_{a=k}^{\infty} N_{a+1,t+1} = \sum_{a=k+1}^{\infty} N_{a,t+1} = N_{t+1} - R_{t+1}.$$

This completes the proof of (2.3).

Just as (2.3) is the counterpart of (2.1) for the total population, (2.4) is the counterpart of (2.2) for the total biomass. Furthermore, the proof is similar. Since (2.2) is valid for $a \geq k + 1$, the proof begins by summing (2.2) over ages from $k + 1$ to ∞ . Notice that

$$\begin{aligned} \sum_{a=k+1}^{\infty} N_{a+1,t+1}^* &= N_{t+1}^* - N_{k+1,t+1}^* - N_{k,t+1}^* \\ &= N_{t+1}^* - w_{k+1,t+1} \tau_t R_t - w_{k,t+1} R_{t+1} \\ \sum_{a=k+1}^{\infty} N_{at}^* &= N_t^* - w_{kt} R_t \\ \sum_{a=k+1}^{\infty} N_{a-1,t-1}^* &= N_{t-1}^*. \end{aligned}$$

Applying these three results to the sum of (2.2) and collecting terms gives

$$(2.9) \quad N_{t+1}^* - \tau_t N_t^* - w_{k,t+1} R_{t+1} = \rho \tau_t (N_t^* - \tau_{t-1} N_{t-1}^* - [w_{kt} + \rho^{-1}(w_{kt} - w_{k+1,t+1})] R_t).$$

Here, the expression in square brackets reduces simply to $w_{k-1,t-1}$ as a result of the growth equation (1.9) when $a = k$. Consequently, (2.9) becomes (2.4), given the definitions (1.2) and (1.5) for R^* , v , and V . This completes the proof of (2.4).

Like (2.2), (2.4) expresses conservation of biomass increments, in this case for the total population. Equation (2.4) is the counterpart of a similar equation in Deriso (1980, p. 269, equation 4), except that the earlier equation excludes the final term involving v_{t-k} because Deriso's growth model (1.21) has the parameter $v_t = 0$ for every t . The new term gives (2.4) an interesting symmetry in the role of R^* on each side of the equation.

Equation (2.5) is a simple consequence of (1.7); it follows as usual on summing over ages a from k to ∞ . Since ϕ_t represents survival, $1 - \phi_t$ refers to mortality. Thus, $C_t / (1 - \phi_t)$ might be called the "catch per unit mortality" (CPUM) in comparison with the more conventional catch per unit effort, CPUE. Even though CPUE may not index population in this model, (2.5) shows that the CPUM does. Similarly, (2.6) shows that the same concept applies to biomass. The straightforward proof is based on multiplying (1.7) by w_{at} and summing over ages a .

It follows from (2.6) that

$$N_t^* - C_t^* = \phi_t N_t^*;$$

consequently, from (1.2), (1.15), and (1.8):

$$(2.10) \quad R_t^* = V_{t-k} R_t \{ \phi_{t-k} N_{t-k}^* \}$$

where it is understood that the function $R_t \{ \cdot \}$ is specified by an

assumption such as (1.13a) or (1.13b). By combining (2.4) and (2.10), one obtains (2.7), which is a recursion formula for calculating the biomass N_{t+1}^* at time $t + 1$ based on the biomass at two previous times t and $t - 1$. Furthermore, (2.6) allows one to convert (2.7) into (2.8), a similar recursion formula for the catch C_{t+1}^* . Even though (2.7) and (2.8) look rather complex, their biological meaning is identical to (2.4), that is, a conservation equation for biomass increments.

Equation (2.8) satisfies the criterion mentioned at the start of this section: it involves only the catch and effort data, C_t^* and E_t , for various years t . The dependence on C_t^* is shown explicitly, and E_t enters the equation through τ_t and ϕ_t via (1.10) and one of the formula pairs (1.11)–(1.12). Practical application of (2.8) involves making specific assumptions about the various parameters in the equation. For example, suppose that all parameters are independent of time and that survival and recruitment are given by (1.11a), (1.12a), and (1.13a). Then (2.8) becomes

$$(2.11) \quad (1 - e^{-qE_{t+1}})^{-1} C_{t+1}^* \\ = (1 + \rho) e^{-M} (e^{qE_t} - 1)^{-1} C_t^* \\ - \rho e^{-2M - qE_t} (e^{qE_{t-1}} - 1)^{-1} C_{t-1}^* \\ + \alpha V (e^{qE_{t+1-k}} - 1)^{-1} C_{t+1-k}^* \\ \times [1 - \beta \gamma (e^{qE_{t+1-k}} - 1)^{-1} C_{t+1-k}^*]^{1/\gamma} \\ - \rho \alpha v e^{-M - qE_t} (e^{qE_{t-k}} - 1)^{-1} C_{t-k}^* \\ \times [1 - \beta \gamma (e^{qE_{t-k}} - 1)^{-1} C_{t-k}^*]^{1/\gamma}.$$

Although (2.11) is rather complex, it can be regarded as a predictive equation for C_{t+1}^* based on the data

$$(2.12) \quad C_{t-k}^*, C_{t+1-k}^*, C_{t-1}^*, C_t^*, E_{t-k}, E_{t+1-k}, E_{t-1}, E_t, E_{t+1}$$

and the seven parameters

$$(2.13) \quad M, q, \rho, \alpha v, \alpha V, \beta, \gamma$$

where the eighth parameter k is presumed known. (Here, α , v , and V are confounded because these parameters occur only in the products αv and αV .) In principle, then, (2.11) could be used to estimate the parameters (2.13) by, say, nonlinear regression.

For many data sets, the estimation problem suggested by (2.11)–(2.13) may not be completely solvable; also, there may be problems related to the type of statistical error in (2.11). However, this example illustrates a general method to construct a model for catch and effort data. The key equation is (2.8), termed the *catch equation* in the rest of this paper. This follows from the first five model assumptions (1.6)–(1.10). The modeller's initial task is to incorporate into the catch equation a set of additional assumptions appropriate to the fishery in question. These always concern the parameters that drive the model. For example, (2.11) is based on time-independent parameters, as well as the last three model assumptions (1.11a), (1.12a), and (1.13a). This is only one of a very large number of possibilities. Other models might be based on (1.11c)–(1.12c). Furthermore, the parameters might be allowed to depend on time; for example, survivals might be related to environmental observations. Some parameters, such as those associated with growth and mortality, might be known from survey data. Fishing mortality might be independent of effort, or, as in (1.13b), recruitment might be independent of stock biomass. The modeller must apply his or her knowledge of the fishery to tailor the model to the situation. The advantage of the method here is that it provides a single framework in which numerous models can be tried and compared easily with one another.

Thus, the catch equation serves as a starting point for the iterative model-building of Box and Jenkins (1970); it specifies a useful *class* of fisheries models. The next section compiles some of the possibilities.

3. Special Cases

The model (1.6)–(1.13) involves three flexible biological components: mortality, growth, and recruitment. To obtain an orderly list of variations in the catch equation, it is necessary to define the alternatives available for each of these three components. In this section, consider the case of constant parameters discussed above in connection with (2.11). Table 3 lists some of the main ways in which the model can be specialized. There are two options for mortality, three for growth, and four for recruitment, with an additional set of four possible recruitment submodels. All these cases have been singled out because they fundamentally affect the number of time lags and parameters in the catch equation, as explained later. In the theory here, these assumptions serve to guide the biologist in identifying a tentative model, along the lines of "step 2" in Box and Jenkins' (1970) model-building strategy. The next few paragraphs describe in detail the options listed in Table 3.

Mortality

Model M1 is simply the general case (1.11a)–(1.12a) in which fishing occurs at the start of the year; the associated parameters are M and q . The alternatives (1.11b)–(1.12b) and (1.11c)–(1.12c) are not particularly important here, as they do not fundamentally affect the structure of the catch equation. A more significant alternative is M2, the assumption that fishing mortality is constant. This model would be valid in fisheries where effort is adjusted by fishermen or by regulation to take a fixed annual fraction of available stock. For example, effort might be held at a fixed level from year to year, regardless of stock size; or effort might increase in years of low stock, but with a corresponding loss of efficiency. The resulting version of (2.8),

$$(3.1) \quad C_{t+1}^* = (1 + \rho) \tau C_{t+1}^* - \rho \tau^2 C_{t-1}^* \\ + V(1 - \phi) R_{t+1} \{ \phi C_{t+1-k}^* / (1 - \phi) \} \\ - \rho \tau v (1 - \phi) R_t \{ \phi C_{t-k}^* / (1 - \phi) \},$$

involves only the catch C^* , not the effort E_t . Here, ϕ and τ are used as the associated mortality parameters.

Growth

Model G1 is the general case (1.14), with parameters ρ , V , and v . Model G2 is Deriso's (1980) restriction (1.22) that the extrapolated prerecruitment weight v is 0; this leaves only the two parameters ρ and V . Model G3 is the case of constant weight

$$(3.2) \quad v = V;$$

by (1.14), (3.2) implies that $w_{at} = V$ for all a and t . In this case the value of ρ is irrelevant in (1.14); furthermore, if

$$(3.3) \quad \rho = 0,$$

then the difference equation (1.9) also implies constant weight. Thus, (3.2) and (3.3) are companion conditions; either implies constant weight for ages $a \geq k$. When (3.3) is true, the catch equation (2.8) assumes the much simpler form

$$(3.4) \quad C_{t+1}^* / (1 - \phi_{t+1}) = \tau_t C_t^* / (1 - \phi_t) \\ + V R_{t+1} \{ \phi_{t+1-k} C_{t+1-k}^* / (1 - \phi_{t+1-k}) \}.$$

TABLE 3. Special cases of the general model, corresponding to particular assumptions on mortality, growth, and recruitment. Each case is assigned a label and a brief description. The model parameters for each case are also listed.

Label	Parameters	Defining assumptions	Description
<i>Mortality</i>			
M1	M, q	(1.11a) and (1.12a)	General mortality model
M2	τ, ϕ	(1.11a) and $\phi = \text{constant}$	Constant fishing mortality
<i>Growth</i>			
G1	ρ, V, v	(1.14) with $v \neq V$	General growth model
G2	ρ, V	(1.14) with $v = 0$	Deriso's growth model
G3	V	(1.14) with $v = V$	Constant weight
<i>Recruitment</i>			
R1	α, β, γ	(1.13a) with $k > 2$	General recruitment model
R2	α, β, γ	(1.13a) with $k = 2$	Two-year recruitment
R3	α, β, γ	(1.13a) with $k = 1$	One-year recruitment
R4	R	(1.13b)	Constant recruitment
<i>Recruitment submodels for $i = 1, 2$, or 3</i>			
R1a	α, β	Model R1 with $\gamma = 1$	Schaefer recruitment
R1b	α, β	Model R1 with $\gamma = 0$	Ricker recruitment
R1c	α, β	Model R1 with $\gamma = -1$	Beverton-Holt recruitment
R1d	α	Model R1 with $\beta = 0$	Constant productivity

This equation can also be obtained another way. When the weight w_{at} is constant, there is no essential mathematical distinction between N_t and N_t^* ; these quantities differ only by the fixed factor V . Similar remarks apply to R_t^* and R_t . Consequently, the result (2.3) is valid with superscript asterisks on N and R ; call this revised version (2.3*). Then (3.4) follows from (2.3*) exactly as the earlier equation (2.8) follows from (2.4). Notice that (2.3*) is much simpler than (2.4); (2.3*) is a first-order difference equation, while (2.4) is second order. Biologically, the assumption of constant weight allows conservation of biomass increment to be replaced by conservation of biomass itself. Indeed, (3.4) is just a biomass conservation equation.

The constant weight equation (3.4) has one other potential application. Suppose that recruitment depends on population numbers rather than biomass and that the catch data are in numbers rather than weight. Mathematically, one can represent this situation simply by letting $V = 1$ and dropping the superscript asterisks. Then (3.4) becomes

$$(3.5) \quad C_{t+1}/(1 - \phi_{t+1}) = \tau_t C_t/(1 - \phi_t) + R_{t+1}$$

which is the catch equation for C_t (as opposed to C^*) when recruitment depends on population numbers or, as in (1.13b), is independent of population variables.

Recruitment

The model R1 is simply the general recruitment model (1.13a) with parameters α, β , and γ . The case $k > 2$ (R1) is distinguished from $k = 2$ and $k = 1$ (R2 and R3, respectively) because the recruitment age k influences the lags appearing in the catch equation. Model R4 is the constant recruitment case (1.13b). As the discussion leading to (1.35) shows, this can also be considered a limiting form of (1.13a) with only one remaining parameter R . In this case the catch equation (2.8) reduces to

$$(3.6) \quad C_{t+1}^*/(1 - \phi_{t+1}) = (1 + \rho)\tau_t C_t^*/(1 - \phi_t) - \rho\tau_t \tau_{t-1} C_{t-1}^*/(1 - \phi_{t-1}) + (V - \rho\tau_t)R.$$

Finally, models R1, R2, and R3 all have submodels corresponding to special values of γ , as shown in Table 3. For example, R1a is Schaefer recruitment with $k > 2$, and R3b is Ricker recruitment with $k = 1$.

The two mortality models, three growth models, and four recruitment models give a total of $2 \times 3 \times 4 = 24$ possible variations to the catch equation. These are listed in Table 4, along with the corresponding model parameters and lags in catch and effort used to predict C_{t+1}^* . (For example, C_t^* has lag 1.) The full set of 24 possible equations need not all be listed here, but Table 5 shows all those obtained by combinations of (1) M1 or M2, (2) G1 or G3, and (3) R1 or R4. In general, the catch equation for a model follows from (2.8) or its counterpart (2.11) with constant parameters. The cases M2, G3, and R4 can be obtained from the specialized versions of the catch equation (3.1), (3.4), and (3.6), respectively.

It is convenient to devise a compact notation to refer to a model. Let " $\langle mgr \rangle$ " refer to the combination of components Mm, Gg, and Rr. For example, (111) is the full model (2.11) with seven parameters (2.13). Similarly, (214) is the constant effort and recruitment model obtained by setting R_t constant in (3.1), or, equivalently, setting τ_t and ϕ_t constant in (3.6). Various historical models can, at least approximately, be represented in this notation as follows:

Model	Historical reference
$\langle 133a \rangle$	Schaefer 1957; Schnute 1977
$\langle 133 \rangle$	Pella and Tomlinson 1969
$\langle 121 \rangle$	Deriso 1980
$\langle 134 \rangle$	Roff 1983

I will not give a detailed discussion of the approximations necessary to demonstrate the validity of the above table, except to comment briefly on Roff's (1983) model

$$(3.7) \quad C_{t+1}^* = A + BE_{t+1}C_t^*/E_t$$

with parameters A and B . Roff justified (3.7) with the ad hoc

TABLE 4. Models obtained by combining assumptions on mortality, growth, and recruitment. Labels are consistent with Table 3. Each model is listed with its associated catch lags, effort lags (if any), parameters, and parameter count.

Model, MGR	Lags		Parameters	
	Catch lags	Effort lags	Parameter list	Count
111	1, 2, $k, k + 1$	0, 1, 2, $k, k + 1$	$M, q, \rho, \alpha V, \alpha v, \beta, \gamma$	7
112	1, 2, 3	0, 1, 2, 3	$M, q, \rho, \alpha V, \alpha v, \beta, \gamma$	7
113	1, 2	0, 1, 2	$M, q, \rho, \alpha V, \alpha v, \beta, \gamma$	7
114	1, 2	0, 1, 2	M, q, ρ, RV, Rv	5
121	1, 2, k	0, 1, 2, k	$M, q, \rho, \alpha V, \beta, \gamma$	6
122	1, 2	0, 1, 2	$M, q, \rho, \alpha V, \beta, \gamma$	6
123	1, 2	0, 1, 2	$M, q, \rho, \alpha V, \beta, \gamma$	6
124	1, 2	0, 1, 2	M, q, ρ, RV	4
131	1, k	0, 1, k	$M, q, \alpha V, \beta, \gamma$	5
132	1, 2	0, 1, 2	$M, q, \alpha V, \beta, \gamma$	5
133	1	0, 1	$M, q, \alpha V, \beta, \gamma$	5
134	1	0, 1	M, q, RV	3
211	1, 2, $k, k + 1$	—	$\tau, \rho, \phi \alpha V, \phi \alpha v, \phi \beta / (1 - \phi), \gamma$	6
212	1, 2, 3	—	$\tau, \rho, \phi \alpha V, \phi \alpha v, \phi \beta / (1 - \phi), \gamma$	6
213	1, 2	—	$\tau, \rho, \phi \alpha V, \phi \alpha v, \phi \beta / (1 - \phi), \gamma$	6
214	1, 2	—	$\tau, \rho, (1 - \phi)R(V - \rho \tau v)$	3
221	1, 2, k	—	$\tau, \rho, \phi \alpha V, \phi \beta / (1 - \phi), \gamma$	5
222	1, 2	—	$\tau, \rho, \phi \alpha V, \phi \beta / (1 - \phi), \gamma$	5
223	1, 2	—	$\tau, \rho, \phi \alpha V, \phi \beta / (1 - \phi), \gamma$	5
224	1, 2	—	$\tau, \rho, (1 - \phi)RV$	3
231	1, k	—	$\tau, \phi \alpha V, \phi \beta / (1 - \phi), \gamma$	4
232	1, 2	—	$\tau, \phi \alpha V, \phi \beta / (1 - \phi), \gamma$	4
233	1	—	$\tau, \phi \alpha V, \phi \beta / (1 - \phi), \gamma$	4
234	1	—	$\tau, (1 - \phi)RV$	2

TABLE 5. All possible catch equations for the combined mortality models M1 or M2, growth models G1 or G3, and recruitment models R1 or R4. Model numbers are labelled as the triple (MGR).

(111)	$(1 - e^{-qE_{t+1}})^{-1} C_{t+1}^*$ $= (1 + \rho)e^{-M}(e^{qE_t} - 1)^{-1} C_t^* - \rho e^{-2M - qE_t}(e^{qE_{t-1}} - 1)^{-1} C_{t-1}^*$ $+ \alpha V(e^{qE_{t+1-k}} - 1)^{-1} C_{t+1-k}^*[1 - \beta\gamma(e^{qE_{t+1-k}} - 1)^{-1} C_{t+1-k}^*]^{1/\gamma}$ $- \rho\alpha v e^{-M - qE_t}(e^{qE_{t-k}} - 1)^{-1} C_{t-k}^*[1 - \beta\gamma(e^{qE_{t-k}} - 1)^{-1} C_{t-k}^*]^{1/\gamma}$
(114)	$\frac{C_{t+1}^*}{1 - e^{-qE_{t+1}}} = \frac{(1 + \rho)e^{-M}C_t^*}{e^{qE_t} - 1} - \frac{\rho e^{-2M - qE_t}C_{t-1}^*}{e^{qE_{t-1}} - 1} + RV - Rv\rho e^{-M - qE_t}$
(131)	$(1 - e^{-qE_{t+1}})^{-1} C_{t+1}^* = e^{-M}(e^{qE_t} - 1)^{-1} C_t^*$ $+ \alpha V(e^{qE_{t+1-k}} - 1)^{-1} C_{t+1-k}^*[1 - \beta\gamma(e^{qE_{t+1-k}} - 1)^{-1} C_{t+1-k}^*]^{1/\gamma}$
(134)	$\frac{C_{t+1}^*}{1 - e^{-qE_{t+1}}} = \frac{e^{-M}C_t^*}{e^{qE_t} - 1} + RV$
(211)	$C_{t+1}^* = (1 + \rho)\tau C_t^* - \rho\tau^2 C_{t-1}^*$ $+ \alpha V\phi C_{t+1-k}^*[1 - \beta\gamma\phi(1 - \phi)^{-1} C_{t+1-k}^*]^{1/\gamma}$ $- \rho\alpha v\tau\phi C_{t-k}^*[1 - \beta\gamma\phi(1 - \phi)^{-1} C_{t-k}^*]^{1/\gamma}$
(214)	$C_{t+1}^* = (1 + \rho)\tau C_t^* - \rho\tau^2 C_{t-1}^* + (1 - \phi)RV - (1 - \phi)\rho\tau Rv$
(231)	$C_{t+1}^* = \tau C_t^* + \alpha V\phi C_{t+1-k}^*[1 - \beta\gamma\phi(1 - \phi)^{-1} C_{t+1-k}^*]^{1/\gamma}$
(234)	$C_{t+1}^* = \tau C_t^* + (1 - \phi)RV$

assumption that CPUE should be roughly similar from one time step to the next. For comparison, see model (134) in Table 5. When qE_t is small for all t ($qE_t \ll 1$), then the catch equation for (134) takes the form

$$(3.8) \quad C_{t+1}^* = (qVR + e^{-M}C_t^*/E_t)E_{t+1}.$$

This is quite similar to Roff's model with the added advantage

that (unlike Roff's) it predicts zero catch in year $t + 1$ if there is no effort that year. Furthermore, (3.8) is based on three clearly stated hypotheses: (1) constant weight, (2) constant recruitment, and (3) low fishing mortality ($qE \ll 1$).

In any model (mgr), it is important to take note of the number of estimable parameters. As Table 4 indicates, this is not always the complete list of parameters for the separate model compo-

TABLE 6. A comparison of the Schaefer (R3a) and general (R3) recruitment models. The table lists estimable parameters for all models obtained by combining R3a with a mortality model (M1 or M2) and a growth model (G1, G2, or G3). The parameter count for this list is then compared with the corresponding parameter count from Table 4. Theoretically, the counts should differ by only one parameter, since R3a is obtained from R3 by setting $\gamma = 1$.

Model, MGR	Parameters	R3a	R3
113a	$q, (1 + \rho)e^{-M} + \alpha V, \rho e^{-M}(e^{-M} + \alpha v), \alpha V\beta, \rho\alpha v\beta$	5	7
123a	$q, (1 + \rho)e^{-M} + \alpha V, \rho e^{-2M}, \alpha V\beta$	4	6
133a	$q, e^{-M} + \alpha V, \alpha V\beta$	3	5
213a	$(1 + \rho)\tau + \alpha V\phi, \alpha V\phi^2\beta/(1 - \phi), \rho\tau(\tau + \alpha v\phi), \rho\tau\alpha v\phi^2\beta/(1 - \phi)$	4	6
223a	$(1 + \rho)\tau + \alpha V\phi, \alpha V\phi^2\beta/(1 - \phi), \rho\tau^2$	3	5
233a	$\tau + \alpha V\phi, \alpha V\phi^2\beta/(1 - \phi)$	2	4

nents in Table 3. For example, the full model (111) theoretically involves eight parameters in Table 3: $M, q, \rho, V, v, \alpha, \beta$, and γ . However, Table 4 (and the earlier discussion of (2.11)–(2.13)) shows that the three parameters α, V , and v are confounded, so that (111) involves only seven parameters. This is not surprising biologically. If only the catch biomass C_t is available, the model does not know if there are a large number of small fish or a smaller number of large fish. Consequently, the number recruited (proportional to α) is confounded with fish weight (proportional to V and v). As the model becomes simpler, the list of estimable parameters becomes smaller. For example, (234) involves only two parameters: τ and $(1 - \phi)RV$. Again this is reasonable biologically. Since the model assumes constant fishing mortality, weight, and recruitment, it cannot distinguish among fishing survival $1 - \phi$, weight V , and recruitment numbers R ; raising the value of one of these parameters can be compensated by lowering another.

It is often difficult to tell just by looking at a model precisely how many parameters are estimable. Although confounding is obvious in certain cases (such as those above) when parameters appear in products or sums, other cases involve more complex analytical expressions. Table 6 illustrates this difficulty by listing parameters for models with Schaefer recruitment at age $k = 1$ (model R3a). In this case, the model (mg3a) consistently involves two fewer parameters than (mg3), even though the submodel is obtained by restricting only one parameter, $\gamma = 1$. Thus, the apparent number of parameters in (mg3) may be wrong, at least for values of γ near 1. Mathematically, the problem occurs because, when $\gamma = 1$, a recruitment term and survival term combine. For example, the classical Schaefer model (133a) has the catch equation

$$(3.9) \quad C_{t+1}^* = \frac{1 - e^{-qE_{t+1}}}{e^{qE_t} - 1} \left(e^{-M} + \alpha V - \frac{\alpha V\beta C_t^*}{e^{qE_t} - 1} \right) C_t^*$$

with the three parameters

$$(3.10) \quad q, e^{-M} + \alpha V, \alpha V\beta.$$

Consistent with Schnute (1977), these correspond to a catchability, production rate, and carrying capacity, respectively. The production rate $e^{-M} + \alpha V$ results from the combination of survival (e^{-M}) and recruitment (αV).

The examples of Table 6 are not the only ones to be aware of. For instance, model (134) in Tables 4 and 5 apparently has three parameters, q, M , and RV . However, as shown above, the model takes the form (3.8) with only two parameters, M and qRV , when qE is small. Thus, the practitioner should anticipate problems with confounded parameters when attempting to find

estimates from actual data. In real applications, some parameters are usually available independently and should not be estimated from catch and effort data. This is particularly true for the growth parameters ρ, V , and v . A compelling advantage to the approach here is that it allows the modeller to build on prior biological knowledge.

4. Equilibrium

After estimating a model's parameters, the modeller often wishes to use them to determine an optimal harvest policy. Typically, the analysis begins by studying how the fishery would perform if all parameters remained fixed forever at their estimated values, along with a constant fishing effort E . The resulting situation is a theoretical equilibrium in which none of the model variables depend on time. This section describes equilibrium for model (111), which is the same as that for (211), since at equilibrium ϕ is constant:

$$(4.1) \quad \phi = e^{-qE}.$$

The catch equation for model (211) in Table 5 becomes an equilibrium equation if catch C^* is presumed independent of the year t . Dropping the subscripts in this equation simplifies it considerably because the catch C^* can then be cancelled from all terms. One can then solve the simplified equation for C^* to obtain

$$(4.2) \quad C^* = \frac{1 - \phi}{\beta\gamma\phi} \left\{ 1 - \left[\frac{(1 - \sigma\phi)(1 - \rho\sigma\phi)}{\phi(\alpha V - \rho\alpha v\sigma\phi)} \right]^\gamma \right\}$$

where $\sigma = e^{-M}$ is a parameter independent of E . Thus, (4.2) describes equilibrium catch C^* as a function of equilibrium fishing survival ϕ and, by (4.1), implicitly as a function of E .

Although the function (4.2) is rather complex, it is possible to derive some of its main features. For example, notice that $C^* = 0$ when $\phi = 1$, that is, by (4.1), when $E = 0$. This, of course, simply means that no fishing gives no catch. It seems reasonable that some high level of effort E (and low survival ϕ) would also give $C^* = 0$, because the population would be driven to extinction. This turns out to be true, as illustrated in Fig. 5. There are three critical values of ϕ : a lower value ϕ_l where $C^* = 0$, a higher value ϕ_m where C^* takes a maximum value C_m^* , and a highest possible value $\phi = 1$ where $C^* = 0$ again. These conclusions, proved in Appendix C, are based on the following assumptions for the model parameters:

$$(4.3) \quad 0 < \sigma < 1$$

$$(4.4) \quad 0 < \rho < 1$$

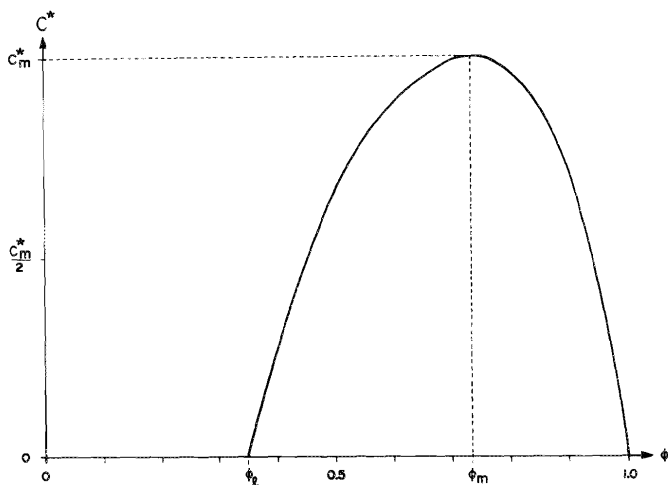


FIG. 5. Plot of equilibrium catch biomass C^* vs. fishing survival ϕ from equation (4.2). This figure, typical of the general case, is based on the parameters $\alpha = 0.8$, $\rho = 0.7$, $v = 0.5$ kg, $V = 0.9$ kg, $\alpha = 2.0$ fish/kg, and $\gamma = -1$ (i.e. Beverton-Holt recruitment). From (4.8) and (4.10) it turns out that $\phi_l = 0.357$ and $\phi_m = 0.729$, respectively; also, $C_m^* = 1.157/\beta$.

$$(4.5) \quad 0 < v < V$$

$$(4.6) \quad 0 < \beta$$

$$(4.7) \quad 1 < \alpha V + \sigma.$$

Conditions (4.3)–(4.6) are precisely the biological restrictions necessary to define σ , ρ , v , V , and β in Section 1. The final condition (4.7) is required to guarantee that the combined effects of survival and recruitment are adequate to maintain the stock when there is no fishery.

An explicit formula for ϕ_l turns out to be (Appendix C)

$$(4.8) \quad \phi_l = 2\{\sigma + \alpha V + \rho\sigma + [(\sigma + \alpha V - \rho\sigma)^2 + 4\rho\sigma(\alpha V - \alpha v)]^{1/2}\}^{-1}.$$

Notice in particular that ϕ_l does not depend on β or γ ; biologically, this reflects the fact that extinction depends only on $R\{S\}$ for low values of S , where the approximation

$$R\{S\} \approx \alpha S$$

is valid regardless of β and γ . In fact, ϕ_l is determined by the four parameters

$$(4.9) \quad \sigma, \rho, \alpha v, \alpha V.$$

Unfortunately, a formula like (4.8) for ϕ_m is not possible, in general; however, it can be shown (Appendix C) that ϕ_m satisfies

$$(4.10) \quad \frac{\alpha v(1 - \rho\sigma\phi_m)^2 + (\alpha V - \alpha v)(1 - \rho\sigma^2\phi_m^2)}{(1 - \sigma\phi_m)(1 - \rho\sigma\phi_m)(\alpha V - \rho\sigma\alpha v\phi_m)} = \frac{1}{\gamma(1 - \phi_m)} \left\{ \left[\frac{(\alpha V - \rho\sigma\alpha v\phi_m)\phi_m}{(1 - \sigma\phi_m)(1 - \rho\sigma\phi_m)} \right]^\gamma - 1 \right\}.$$

This equation in ϕ_m must be solved numerically. A simple method for doing so was suggested by Schnute (1982, p. 48, section 8.5), although essentially the same method can be used to maximize C^* in (4.2) directly. Theoretical considerations (Appendix C) prove that (4.10) always has a solution ϕ_m between ϕ_l and 1, that is,

$$(4.11) \quad \phi_l < \phi_m < 1.$$

The midpoint $(\phi_l + 1)/2$ of the interval (4.11) is a reasonable first guess for numerical iteration toward the solution ϕ_m of (4.10). Notice that ϕ_m depends on the four parameters (4.9), plus the fifth parameter γ . Once ϕ_m is found, the optimal effort and catch, E_m and C_m^* , can be calculated from (4.1) and (4.2), respectively.

Equations (4.2), (4.8), and (4.10) simplify greatly in the constant weight case, $v = V$; explicitly, they reduce, respectively, to

$$(4.12) \quad C^* = \frac{1 - \phi}{\beta\gamma\phi} \left[1 - \left(\frac{1 - \sigma\phi}{\alpha V\phi} \right)^\gamma \right]$$

$$(4.13) \quad \phi_l = 1/(\sigma + \alpha V)$$

$$(4.14) \quad \frac{1 - \phi_m}{1 - \sigma\phi_m} = \frac{1}{\gamma} \left[\left(\frac{1 - \sigma\phi_m}{\alpha V\phi_m} \right)^{-\gamma} - 1 \right].$$

Furthermore, for Schaefer recruitment with $\gamma = 1$, (4.14) can be solved explicitly. This, together with (4.1)–(4.2), gives

$$(4.15) \quad \phi_m = 2/(1 + \alpha V + \sigma)$$

$$(4.16) \quad C_m^* = (\sigma + \alpha V - 1)^2/(4\alpha V\beta)$$

$$(4.17) \quad E_m = q^{-1} \log [(1 + \alpha V + \sigma)/2].$$

The expression (4.17) for E_m exhibits the biologically reasonable property that sustainable effort increases with increasing productivity αV and survival σ .

Incidentally, ϕ_m in (4.15) satisfies the condition

$$(4.18) \quad \frac{1}{\phi_m} = \frac{1}{2} \left(\frac{1}{\phi_l} + \frac{1}{1} \right);$$

that is, ϕ_m is the harmonic mean (reciprocal average of reciprocals) of the end points of the interval (4.11). This condition is the counterpart of the traditional observation for the Schaefer model that the optimum is midway between extremes; that is, the apex of a parabola lies above the midpoint of the axis crossings. When $\gamma = 1$ in (4.12), C^* is a quadratic function of $1/\phi$.

5. Stochastic Models

The practical application of any model requires a method of estimating its parameters. This method, in turn, hinges on assumptions related to statistical error. The general model (1.6)–(1.13) is defined with time-varying parameters (Table 1) M_t , q_t , v_t , V_t , and α_t , although these parameters are frequently presumed constant in later discussion. It is important, however, to show as in Section 2 that the main conclusions (2.1)–(2.8) are valid even with parameters dependent on time t , so that statistical variation can be introduced systematically. Explicitly, the model admits at least four possible sources of random behavior: (1) natural mortality M_t , or, equivalently, survival σ_t , (2) fishing efficiency q_t , or, equivalently, survival ϕ_t , (3) cohort weight v_t and V_t , and (4) recruitment productivity α_t . Incidentally, item (2) in this list is also equivalent to a random error in measuring the effort E_t .

Just as there are many possible ways to specialize the deterministic aspects of the general model, so too are there many ways to specialize its stochastic aspects. In fact, the number of potential stochastic versions is so great that I make no attempt here to organize them as I have done for deterministic models in Section 3. The modeller must decide in a particular case where the most important sources of variation lie. The general plan is to extend the deterministic form of the model by

adding error terms at appropriate places and then to estimate parameters by minimizing error. As the number of parameters increases, the apparent error can, of course, be reduced. More parameters give the model more flexibility, although, as explained below, this is not necessarily an advantage.

An extreme possibility is to include enough parameters so that the model equations can be fit exactly. This approach, called "stock reduction analysis" (SRA), was taken by Kimura and Tagart (1982) and Kimura et al. (1984). Essentially, they considered a model for catch data in which ϕ_t is an unknown parameter for every t . This gives a large number of equations with a large number of unknowns, and these can be solved numerically on a computer. Among other things, SRA gives fishing mortalities that precisely explain the catch data, in the spirit of "virtual population analysis" (VPA; Gulland 1955, Pope 1972) for catch at age data. Typically, SRA includes several more parameters than equations and involves locating those solutions that best correspond to the modeller's conception of the fishery.

One advantage of SRA is that it completely avoids the problem of statistical definition of the model. A disadvantage of any such method is that it will fail to locate trends in the data. Statistics is the science of data reduction, of locating a signal in the presence of noise. Without statistical assumptions, a model can only echo back the noise (errors) in the data. In the context of this paper, SRA is just one extreme in which the number of parameters becomes large. Indeed, a parameter estimation technique based on minimizing error will also work for SRA because one possibility is that the minimum error is precisely zero.

The first problem in introducing model error is to decide which of the various possible errors dominate. For example, consider the constant weight model (131) in Table 5. It has the form

$$(5.1) \quad F_{t+1} = e^{-M}F_t + \alpha G_{t+1-k}$$

where, for convenience, we define

$$F_t = (1 - e^{-qE_t})^{-1}C_t^* \\ G_t = V(e^{qE_t} - 1)^{-1}C_t^*[1 - \beta\gamma(e^{qE_t} - 1)^{-1}C_t^*]^{1/\gamma}.$$

If we assume that survival e^{-M} and recruitment productivity α are both subject to multiplicative errors of about the same magnitude, then (5.1) could reasonably be written in the stochastic form

$$(5.2) \quad \log(F_{t+1}) - \log(e^{-M}F_t + \alpha G_{t+1-k}) = \epsilon_t$$

where ϵ_t is normal with mean 0. Parameters might then be estimated by minimizing the sum of squares of errors ϵ_t in (5.2). Alternatively, one might hope to maximize the likelihood associated with (5.2), but an exact likelihood function is extremely difficult to calculate because (unlike classical autoregressive processes) the data C_t^* do not enter the model linearly.

The above idea seems temptingly appropriate to all models discussed in this paper. These typically have a left-hand side with data from year $t+1$ and a right-hand side with data from earlier years. Let D_t represent either the difference between left-hand and right-hand sides or, as in (5.2), the difference of logarithms. The former choice is appropriate if the modeller believes that errors are additive, and the latter if errors are presumed multiplicative. Then, D_t is a function of observed data (represented as a vector X_t) and various parameters (represented as a vector θ). A general stochastic model, analogous to

(5.2), is

$$(5.3) \quad D_t(X_t, \theta) = \epsilon_t.$$

For example, in (5.2)

$$X_t = (C_{t+1-k}^*, C_t^*, C_{t+1}^*, E_{t+1-k}, E_t, E_{t+1})$$

$$\theta = (M, q, \alpha V, \beta, \gamma)$$

$$D_t = \log(F_{t+1}) - \log(e^{-M}F_t + \alpha G_{t+1-k}).$$

Although (5.3) seems reasonable, it may be inadequate for all situations. To see this, consider the general model (111), equation (2.11), obtained from (2.8) by assuming constant parameters. Notice that (2.8) involves both τ_t and τ_{t-1} , that is (from (1.25)), both M_t and M_{t-1} . Consequently, if we assume that variation in natural survival is important, the stochastic version of (2.11) should perhaps contain *two* independent errors ϵ_t and ϵ_{t-1} associated with years t and $t-1$. This suggests replacing (5.3) with the more general stochastic model

$$(5.4) \quad D_t(X_t, \theta) = \epsilon_t + \lambda\epsilon_{t-1}$$

where λ is a constant with

$$(5.5) \quad -1 < \lambda < 1.$$

Theoretical reasons for the restriction (5.5) were given by Box and Jenkins (1970, p. 284, appendix A7.6).

The fundamental difference between (5.3) and (5.4) is that errors in (5.4) are correlated. In both models, the errors ϵ_t for various years t are presumed normal (with mean 0) and independent; however, the combined errors

$$\eta_t = \epsilon_t + \lambda\epsilon_{t-1}$$

in (5.4) are correlated. Explicitly,

$$\text{corr}(\eta_t, \eta_{t-1}) = \lambda/(1 + \lambda^2).$$

In particular, (5.3) is the special case of (5.4) obtained when $\lambda = 0$. Parameters in (5.4), including λ , can be estimated by minimizing the sum of squares of errors

$$(5.6) \quad S = \sum \epsilon_t^2$$

for the given data; however, ϵ_t is harder to compute in (5.4) than in (5.3). In addition, an unknown initial value of ϵ_t , say ϵ_0 , is needed to start the process. Here is a recipe for calculating S in (5.6):

Step 1. Let a parameter vector θ be given, along with values of λ and ϵ_0 . Also, let there be enough data to determine the required data vector X_t for $t = 1, \dots, n$. For example, in (5.2) this would mean that C_t^* and E_t are known for $t = 2 - k$ and $t = 1, \dots, n + 1$.

Step 2. Begin with $t = 1$.

Step 3. Compute ϵ_t from (5.4) as follows:

$$(5.7) \quad \epsilon_t = D_t(X_t, \theta) - \lambda\epsilon_{t-1}.$$

Step 4. Replace t by $t + 1$, and repeat step 3. Continue until step 3 is completed with $t = n$.

Step 5. Compute S in (5.6) from $\epsilon_0, \epsilon_1, \dots, \epsilon_n$. Notice that, since ϵ_t in (5.7) depends on θ and λ , S in (5.6) also depends on all these parameters, as well as the initial ϵ_0 :

$$S = S(\theta, \lambda, \epsilon_0).$$

To minimize S as a function of its various parameters, one needs an algorithm that does not require explicit derivatives of S . The simplex method, as implemented by Schnute (1982), should perform well for the small number of parameters

typically involved. Constraints, such as (4.3)–(4.6) and (5.5), may play a significant role. Schnute (1982, p. 47, section 8.4) gave an easy method for including such constraints. Also, certain parameters, such as the growth parameters ρ , v , and V , may be known from studies external to the data. Schnute (1982, p. 43, section 8.2) described how to involve prescribed parameters easily. The minimum value of S may prove highly insensitive to the value of ϵ_0 , and it may simply be best to prescribe $\epsilon_0 = 0$. In any case, both λ and ϵ_0 should be set equal to 0 as initial values for the search.

The stochastic model (5.4) is called a moving average (MA) model with lag 1 because the error is a combination (weighted average) of independent errors associated with time steps t and $t - 1$. The choice of lag 1 here is based on the biological arguments in the paragraph preceding (5.4); essentially, the general catch equation (2.8) is a second-order difference equation involving two time steps. To be consistent with this point of view, one might choose $\lambda = 0$ in a model based on the constant weight version (3.4), although there may always be some interest in seeing how efficiently the additional parameter λ removes error correlation. It is, in any case, important to base parameter estimates on a sum of squares of *uncorrelated* errors. Deriso (1980, p. 274) cited a comment by Box and Jenkins (1970, p. 213) that least squares estimates usually provide close approximations to maximum likelihood estimates. That is true, but the reader should note that Box and Jenkins (1970) referred not to the model errors (denoted η_t above), but their uncorrelated counterparts.

As suggested earlier, the possibilities for incorporating stochastic error into the deterministic models here are almost endless. In general, one could include a random influence anywhere in the difference equation and then solve for ϵ_0 to obtain an equation conceptually the same as (5.3) above. One could also include more lags in the moving average term on the right side of (5.4). Furthermore, there could be more than one source of error in catchability q_t with variable survival and recruitment. This would give rise to an errors-in-variables problem for which at least the relative magnitudes of errors would need to be known a priori. A full explication of all these possibilities is beyond the scope of this paper.

6. Model Identification

The technology developed here makes it possible to return with fresh insight to the question posed in the introduction: "Which model is right?" For example, we have seen that Roff's (1983) model (3.7) is essentially model (134) with constant weight and recruitment. Table 4 shows that (134) involves a minimal number of lags in the catch and effort data. Thus, Roff's model primarily raises the question: can we detect evidence of recruitment in our data? If so, then variations of the constant recruitment model (3.6), including Roff's, should be less effective than more general models which involve lag k recruitment terms. It may be that most fisheries have such noisy recruitment that no predictive values accrues from the lag k terms; the main information may come from lag 1 survival terms as Roff's example suggests. Whichever conclusion proves true in any particular case, the methods here make it possible to relate the model to biological reality. This is an important point because the model may have many purposes beyond simple prediction.

As described in the introduction, the process of iterating toward the best model (that is, model identification) has its roots

in the general theory of time series analysis. Even there, the correct techniques are still a matter of current debate. For example, the recent paper by Tsay and Tiao (1984) rigorously develops a new method of ARMA model identification. If identification is difficult for linear ARMA models, it must be harder still for fisheries models with nonlinear recruitment. I doubt that a solid theoretical procedure is possible or even appropriate in this context, particularly since biological knowledge independent of catch and effort data can play a valuable role in the process.

This paper is devoted primarily to the first step of Box and Jenkins (1970) model building strategy: formulating an appropriate class of models. Once the modeller has a time series of catch and effort data, I think the next step is to accumulate all possible additional information, particularly growth parameters v , V , and ρ and natural mortality M , if possible. To select a tentative model, consider the main apparent sources of error. For example, is recruitment primarily driven by the environment, as appears true for many shellfish? If so, then recruitment can perhaps best be modelled as a constant plus a stochastic term. Is effort measured reasonably well or are the effort data almost meaningless? In the latter case, it may be best to treat the fishing mortality as noise around a constant level.

In any application, it is important to consider the model's sensitivity to the parameters involved. For this reason, Section 3 includes frequent reference to possibilities for parameters that are confounded or drop out of the model in a special case. For instance, if β is small in the expression

$$\alpha S(1 - \beta\gamma S)^{1/\gamma},$$

then the value of γ may be irrelevant because the factor in parentheses is nearly 1. Also, if ρ is small in a second-order difference equation of the form

$$x_{t+1} - ax_t = \rho(x_t - bx_{t-1}),$$

then the equation collapses to first order and the parameter b may be inestimable, particularly when the error dominates the term on the right.

Bearing in mind numerous considerations, such as those mentioned in the two previous paragraphs, the modeller should be able to contrive a tentative model specifically tailored to the fishery in question. It will involve some parameters known a priori and others to be estimated from the data by minimizing a function such as S in (5.6). Attention now shifts to these estimates. How well determined are they? Can some parameters be dropped, thus simplifying the model, with no loss of predictive power? It is an element of folklore among most contemporary fisheries modellers that only two or three meaningful parameters can be estimated from any real catch-effort time series.

This paper is obviously only part of a potentially long discussion and debate, some of it already in the literature. In my view there is a need for many worked examples, which can be related to one another through a central theory, such as the one described here. If the community of researchers can agree on such a theory, then simulation studies can also play an important role in understanding the behavior of parameter estimates, particularly for the short time series (10–20 yr), which are all the data available in many fisheries. Significantly, Box and Jenkins (1970) included several sample data sets in their book, and these have since been used by numerous researchers to compare various methods with one another. Fisheries, likewise, could use such a carefully compiled benchmark, complete with

a description of the known biology for each example. In short, much interesting work remains to be done. I believe progress will be most rapid if the various individual efforts can find a common focal point and thus be seen clearly as parts of a larger whole.

Acknowledgements

I am very grateful for reviews of an earlier draft by R. Deriso and C. Walters. Their thoughtful comments led to substantial improvements in the final version. Furthermore, as I hope this paper makes clear, I share Walters' (1980) enthusiasm for Deriso's (1980) theoretical work. I would not have undertaken the development here without it. Finally, my thanks to M. Stocker for assistance in locating the herring data used for Fig. 1.

References

- BARANOV, F. I. 1925. On a question of the dynamics of the fishing industry. *Byull. Rybn. Khoz.* 8: 7–11. (Transl. from Russian by W. E. Ricker, 1945)
- BEVERTON, R. J. H., AND S. J. HOLT. 1957. On the dynamics of exploited fish populations. U.K. Minist. Agric. Fish. Food Fish. Invest. (Ser. 2) 19: 533 p.
- BOX, G. E. P., AND G. M. JENKINS. 1970. Time series analysis forecasting and control. Holden-Day, San Francisco, CA. xix + 553 p.
- BURKENROAD, M. D. 1948. Fluctuation in abundance of Pacific halibut. *Bull. Bingham Oceanogr. Collect.* Yale Univ. 11: 81–129.
- CLARK, C. 1972. The theoretical side of calculus. Wadsworth Publishing Co., Belmont, CA. 240 p.
- DERISO, R. B. 1980. Harvesting strategies and parameter estimation for an age-structured model. *Can. J. Fish. Aquat. Sci.* 37: 268–282.
- FOURNIER, D., AND C. P. ARCHIBALD. 1982. A general theory for analyzing catch at age data. *Can. J. Fish. Aquat. Sci.* 39: 1195–1207.
- GULLAND, J. A. 1955. Estimation of growth and mortality in commercial fish populations. U.K. Minist. Agric. Fish. Food Fish. Invest. (Ser. 2) 18(9): 46 p.
- KIMURA, D. K., J. W. BALSIGER, AND D. H. ITO. 1984. Generalized stock reduction analysis. *Can. J. Fish. Aquat. Sci.* 41: 1325–1333.
- KIMURA, D. K., AND J. V. TAGART. 1982. Stock reduction analysis, another solution to the catch equations. *Can. J. Fish. Aquat. Sci.* 39: 1467–1472.
- LUDWIG, D., AND R. HILBORN. 1983. Adaptive probing strategies for age-structured fish stocks. *Can. J. Fish. Aquat. Sci.* 40: 559–569.
- LUDWIG, D., AND C. J. WALTERS. 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. *Can. J. Fish. Aquat. Sci.* 38: 711–720.
- PELLA, J. J., AND P. K. TOMLINSON. 1969. A generalized stock production model. *Inter-Am. Trop. Tuna Comm. Bull.* 13: 421–496.
- POPE, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. *Int. Comm. Northwest. Atl. Fish. Res. Bull.* 9: 65–74.
- RICKER, W. E. 1954. Stock and recruitment. *J. Fish. Res. Board Can.* 11: 559–623.
1958. Handbook of computations for biological statistics of fish populations. *Bull. Fish. Res. Board Can.* 119: 300 p.
1975. Computation and interpretation of biological statistics of fish populations. *Bull. Fish. Res. Board Can.* 191: xviii + 382 p.
- ROFF, D. A. 1983. Analysis of catch/effort data: a comparison of three methods. *Can. J. Fish. Aquat. Sci.* 40: 1496–1506.
- SCHAEFER, M. B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-Am. Trop. Tuna Comm. Bull.* 1: 25–56.
1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. *Inter-Am. Trop. Tuna Comm. Bull.* 2: 245–285.
- SCHNUTE, J. 1977. Improved estimates from the Schaefer production model: theoretical considerations. *J. Fish. Res. Board Can.* 34: 583–603.
1982. A manual for easy nonlinear parameter estimation in fishery research with interactive microcomputer programs. *Can. Tech. Rep. Fish. Aquat. Sci.* 1140: xvi + 115 p.
- SCHNUTE, J., AND D. FOURNIER. 1980. A new approach to length–frequency analysis: growth structure. *Can. J. Fish. Aquat. Sci.* 37: 1337–1351.
- SKUD, B. E. 1975. Revised estimates of halibut abundance and the Thompson–Burkenroad debate. *Int. Pac. Halibut. Comm. Sci. Rep.* 56.
- STOCKER, M., V. HAIST, AND L. ROSENFELD. 1982. Summaries of British Columbia herring sample data for the 1981–82 season. *Can. Data Rep. Fish. Aquat. Sci.* 356: v + 273 p.
- THOMPSON, W. F., AND F. H. BELL. 1934. Biological statistics of the Pacific halibut fishery. (2) Effects of changes in intensity upon total yield and yield per unit gear. *Int. Fish. Comm. Rep.* 8.
- TSAY, R. S., AND G. C. TIAO. 1984. Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationarity and nonstationary ARMA models. *J. Am. Stat. Assoc.* 79: 84–96.
- UHLER, R. 1980. Least squares regression estimates of the Schaefer production model: some Monte Carlo simulation results. *Can. J. Fish. Aquat. Sci.* 37: 1284–1294.
- WALTER, G. G. 1973. Delay-differential equation models for fisheries. *J. Fish. Res. Board Can.* 30: 939–945.
- WALTERS, C. J. 1980. Comment on Deriso's delay-difference population model. *Can. J. Fish. Aquat. Sci.* 37: 2365.
- WALTERS, C. J., AND D. LUDWIG. 1981. Effects of measurement errors on the assessment of stock–recruitment relationships. *Can. J. Fish. Aquat. Sci.* 38: 704–710.

Appendix A. Survival Calculations

To prove (1.24), notice that

$$N_{a+1, t+1} = \tau_t N_{at} \text{ from (1.6)}$$

$$= \tau_t (N_{at} - C_{at}) / \phi_t \text{ from (1.7)}$$

$$= \sigma_t (N_{at} - C_{at}) \text{ from (1.10).}$$

The remainder of this appendix derives versions (a), (b), and (c) of the survival equations (1.11)–(1.12). It is sufficient to consider constant M , q , and E during just one year with time t from 0 to 1. For this discussion only, let P_t represent the cohort population at time t , and let Q_t represent the total catch of the cohort from time 0 to time t . We are interested primarily in P_0 and P_1 , the initial and final populations, as well as the total catch Q_1 taken during year 0. Then, from (1.7) and (1.24)

$$(A.1) \quad \phi_0 = (P_0 - Q_1) / P_0$$

$$(A.2) \quad \sigma_0 = P_1 / (P_0 - Q_1).$$

When fishing and natural mortality are concurrent, assume that the population P_t and the catch Q_t are governed by the differential equations

$$(A.3) \quad dP/dt = -(M + qE)P$$

$$(A.4) \quad dQ/dt = qEP.$$

It follows from (A.3) that

$$(A.5) \quad P_t = e^{-(M+qE)t} P_0$$

and then from (A.4)–(A.5) that

$$(A.6) \quad Q_t = [qE/(M + qE)][1 - e^{-(M+qE)t}]P_0.$$

Expressions for P_1 and Q_1 follow from (A.5) and (A.6). Substituting these into (A.1)–(A.2) gives (1.11b)–(1.12b).

If fishing occurs before natural mortality, then from (A.6) with $M = 0$, the total catch

$$(A.7) \quad Q_1 = [1 - e^{-qE}]P_0$$

is removed first. This leaves a new initial population

$$P'_0 = e^{-qE}P_0$$

to which (A.5) applies with $E = 0$; consequently,

$$(A.8) \quad P_1 = e^{-M} e^{-qE} P_0.$$

Substituting (A.7)–(A.8) into (A.1)–(A.2) gives (1.11a)–(1.12a).

Finally, if natural mortality occurs first, then the number of surviving fish available to the fishery is

$$P_1'' = e^{-M}P_0$$

computed from (A.5) with $E = 0$. This gives the new initial population to which (A.6) applies with $M = 0$; consequently,

$$(A.9) \quad Q_1 = [1 - e^{-qE}] e^{-M}P_0.$$

Subtracting the catch Q_1 from the population P_1'' which survived natural mortality gives the same resulting value of P_1 as in (A.8) above. Substituting (A.8)–(A.9) into (A.1)–(A.2) gives (1.11c)–(1.12c).

Appendix B. Recruitment Parameters

The logarithmic derivative of the recruitment curve (1.30) is

$$(B.1) \quad \frac{1}{R} \frac{dR}{dS} = \frac{1 - (1 + \gamma)\beta S}{S(1 - \beta\gamma S)}.$$

Consequently, $dR/dS = 0$ when $S = [(1 + \gamma)\beta]^{-1}$; this is the expression for S_m in (1.31). Substituting that value of S into (1.30) also gives the expression for R_m in (1.31). Similarly, the value S_0 in (1.32) gives $R = 0$ in (1.30) for $\gamma > 0$.

The expressions (1.31) for S_m and R_m can be solved jointly for α and β with the results

$$(B.2) \quad \alpha = (R_m/S_m)(1 + \gamma)^{(1/\gamma)}$$

$$(B.3) \quad \beta = [S_m(1 + \gamma)]^{-1}.$$

Substituting (B.2)–(B.3) into (1.30) gives (1.33). Similarly, (1.36) follows from (1.31)–(1.32), and after substituting (1.36) into (1.33), one can obtain (1.34).

Appendix C. Equilibrium Calculations

The principal equation (4.2) follows from the catch equation (211) in Table 5 as described in Section 4. This expression for C^* vanishes when $\phi = 1$ or when

$$(C.1) \quad \phi(\alpha V - \rho\alpha v\sigma\phi) = (1 - \sigma\phi)(1 - \rho\sigma\phi).$$

Equation (C.1) is quadratic in ϕ ; also, on dividing by ϕ^2 , (C.1)

is quadratic in $1/\phi$, and the roots turn out to be

$$(C.2) \quad 1/\phi = [F \pm (G^2 + H)^{1/2}]/2$$

where

$$F = \sigma + \alpha V + \rho\sigma$$

$$G = \sigma + \alpha V - \rho\sigma$$

$$H = 4\rho\sigma\alpha(V - v).$$

Since $\alpha > 0$, $1 > \sigma > 0$, $1 > \rho > 0$, and $V > v$, it follows that F , G , and H are all positive; furthermore,

$$\begin{aligned} F + (G^2 + H)^{1/2} &> F + G \\ &= 2(\sigma + \alpha V) \\ &> 2 \end{aligned}$$

since $\sigma + \alpha V > 1$ by hypothesis. Thus, choosing the “+” sign in (C.2) gives a value $1/\phi > 1$, i.e. $\phi < 1$. Similarly,

$$\begin{aligned} F - (G^2 + H)^{1/2} &< F - G \\ &= 2\rho\sigma \\ &< 2 \end{aligned}$$

so that the “−” sign in (C.2) gives $\phi > 1$. This shows that (C.1) has a unique root in the range $0 < \phi < 1$ corresponding to “+” in (C.2) and, consequently, proves the validity of the expression (4.8) for ϕ_l .

By Rolle’s theorem (Clark 1972, p. 111), since $C^* = 0$ at ϕ_l and 1, there is a point ϕ_m between ϕ_l and 1 where

$$(C.3) \quad dC^*/d\phi = 0.$$

After considerable algebra, the condition (C.3) can be shown equivalent to (4.10). This calculation is facilitated by writing

$$(C.4) \quad C^* = K(1 - L^\gamma)$$

where the expressions K and L have definitions evident from (4.2). On taking the derivative of the logarithm of each side of (C.4), one can show that (C.3) is equivalent to

$$(C.5) \quad d(\log L)/d\phi = \gamma^{-1}(L^{-\gamma} - 1)d(\log K)/d\phi.$$

Condition (4.10) for ϕ_m follows from (C.5), after calculating the logarithmic derivatives of K and L .