Statistical Models for the Analysis of Ageing Error

Laura J. Richards, Jon T. Schnute, A. R. Kronlund, and Richard J. Beamish

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

Richards, L. J., J. T. Schnute, A. R. Kronlund, and R. J. Beamish. 1992. Statistical models for the analysis of ageing error. Can. J. Fish. Aquat. Sci. 49: 1801–1815.

We present statistical models for estimating the true age distribution of a population, based on multiple readings from individual fish. There are two steps to this process. The first involves estimating a classification matrix that defines the probability of assigning an age *a* to a fish when its true age is *b*. Since true age is unknown, we require an assumption related to ageing error bias; we assume that the true age is the most probable value for the observed age. True age proportions, or alternatively, true ages of fish in the sample are then estimated in the second step. Our methods allow us either to conduct both steps simultaneously or to estimate true age proportions from a previously estimated classification matrix. We illustrate our methods with data on walleye pollock (*Theragra chalcogramma*). We recommend that multiple independent readings be obtained for a subset of structures in future ageing studies and that ageing error be considered in subsequent analyses. Sample sizes must be increased with increasing ageing error to achieve a specified precision in estimates of true age proportions.

Nous présentons des modèles statistiques destinés à estimer la distribution de l'âge réel d'une population à partir de mesures multiples sur des poissons pris individuellement. La démarche comporte deux étapes. La première est le calcul d'une matrice de classification qui définit la probabilité d'assigner un âge *a* à un poisson tandis que son âge réel est *b*. Étant donné qu'on ne connaît pas l'âge réel, nous devons émettre une hypothèse concernant l'erreur systématique dans la détermination de l'âge; nous posons que l'âge réel est la valeur la plus probable de l'âge observé. Dans la seconde étape, on estime les proportions des âges réels, ou encore l'âge réel des poissons de l'échantillon. Nos méthodes nous permettent soit d'effectuer les deux étapes simultanément soit d'estimer les proportions des âges réels à partir d'une matrice de classification calculée antérieurement. Nous illustrons nos méthodes avec des données sur la goberge de l'Alaska (*Theragra chalcogramma*). Nous recommandons d'obtenir des mesures indépendantes multiples pour un sous-ensemble de structures dans les études futures sur la détermination de l'âge, et de considérer dans les analyses subséquentes l'erreur dans la détermination de l'âge. Il est nécessaire d'augmenter la taille des échantillons à mesure que l'erreur augmente pour arriver à une précision donnée dans l'estimation des proportions des âges réels.

Received November 4, 1991
Accepted March 24, 1992
(JB293)

Reçu le 4 novembre 1991
Accepté le 24 mars 1992

Catch–age data play a fundamental role in fish stock assessments (e.g. Megrey 1989) and in other studies of recruitment, growth, and mortality. Although many such analyses assume that fish ages are measured without error, this assumption is rarely met (Beamish and McFarlane 1983). Consequently, year-class strength can be incorrectly determined. Lai and Gunderson (1987) and Tyler et al. (1989) demonstrate how ageing error can lead to inappropriate yield projections and overfishing. Biased estimates of year-class strength can also mask important stock–recruit relationships and environmental correlates of year-class strength (Fournier and Archibald 1982; Hollowed et al. 1987; Myers and Drinkwater 1989; Bradford 1991).

In this paper, we present statistical models for estimating the true age distribution of a population, based on multiple age readings of individual fish. Although recent catch–age models can incorporate ageing error when it is known (Fournier and Archibald 1982; Methot 1989, 1990), general statistical methods have not been presented for assessing ageing error empirically. We describe several maximum likelihood estimation procedures; the one chosen for a particular situation depends on the input data and the specific question of interest.

Two main issues must be considered in any analysis of ageing error. The first deals with estimation of the classification matrix, which defines the probability of assigning an age a to a fish with true age b. The second issue involves estimation of true age proportions within a sample of fish, when individual ages may be assigned with error. Our analysis can address both issues simultaneously, based on a sample of structures for which two or more independent age readings are available. Other authors (e.g. Kimura and Lyons 1991) have advocated use of such data to assess reader bias and variability in order to maintain quality control.

The observed age a of a fish is a random variable that presumably bears some relationship to its true age b. Since true age is unknown, construction of the classification matrix forces us to make an assumption related to ageing error bias. In this paper, we assume that the observation a takes the value b with highest probability. We cannot assume that a is an unbiased estimate of b because observed ages must fall within the life span of the species. For example, due to truncation in the age distribution, mean assigned age will overestimate true age for the youngest fish and underestimate true age for the oldest fish.

The goal of most ageing error analyses is to estimate true age proportions. For example, if all fish in a sample have true age b, the observed ages typically take a range of values in the vicinity of b. In general, ageing error smooths differences in year-class strength, making strong year-classes appear weaker and weak year-classes appear stronger (Fournier and Archibald 1982; Rivard 1989; Kimura and Lyons 1991). The extent of smoothing is governed by the classification matrix. Mathematically, the expected value of the vector of observed age proportions is the product of the classification matrix and the vector of true proportions. This process leads to smoothing because each true proportion is distributed among adjacent ages. An error analysis seeks to reverse this process by allocating observed proportions back to the correct ages. As discussed by several authors (e.g. Pella and Robertson 1979), this is accomplished mathematically using the inverse classification matrix.

Estimation of true age proportions based on a known classification matrix represents one special case of analyses described in this paper. We also consider the need to estimate both the true proportions and the classification matrix simultaneously. Alternatively, estimates of true ages of individual fish within a sample may be required rather than the overall true age proportions. Such estimates could be obtained either from a known classification matrix or simultaneously with estimation of the classification matrix.

A naive method of estimating the true age proportions might be first to assign an age to each fish in a sample, based on multiple readings per fish, and then to estimate age proportions from these ages. This method is incorrect because the assigned age will sometimes deviate from the true age. Unless the age of each fish can be perfectly determined, the resulting age proportions still contain ageing error. The smoothing effect of ageing error is reduced, but not eliminated by such a process. We illustrate this feature in our analysis below.

The process of stock identification using mixture models (Pella and Robertson 1979; Fournier et al. 1984; Hoenig and Heisey 1987; Wood et al. 1987) has features in common with our analysis. In stock identification, a learning sample is used to relate a surrogate variable, such as a particular meristic, to the stock of origin. This relationship can then be used to estimate the proportions of fish from each stock in a sample for which only the surrogate variable is known. For the age data described here, the learning sample consists of multiple age readings; however, the true age of each fish is not known, unlike the true stock in a mixture learning sample. Consequently, our likelihood analysis differs in certain respects from that developed for stock mixtures.

No model can detect a systematic difference between observed and true ages. For example, if the ageing method fails to recognize the first annulus of each otolith, then the observed age could differ consistently by 1 yr from the true age. Our method cannot substitute for careful age validation. Rather, given reasonable assumptions about the classification matrix, it provides statistical corrections to age distributions estimated from multiple age readings.

Rigorous model formulation necessitates statistical detail. Models are presented in section 1, along with methods of statistical inference. Section 2 describes an application of our general methods to data on walleye pollock (*Theragra chalcogramma*). Reader effects are evident in the pollock data and a model tailored specifically to these data is illustrated in section 3. More general methods may be required for other data sets and we discuss several possible model extensions in section 4. To simplify the text as much as possible, some of the model equations are presented in tables. We use parentheses ( ) and brackets [ ] to denote equations in text and tables, respectively.

1. Statistical Model

1A. Concepts and Notation

Before we describe the statistical model, we must first discuss the raw data and develop model notation (Table 1). We assume

TABLE 1. Model notation in the context of samples for which one (1) or several (2+) readers examine each age structure.

Symbol	Readers	Description
		Age descriptors
a	2 +	Observed age-class ($a = 1, \ldots, A$)
b	2 +	True age-class ($b = 1, \ldots, A$)
c	1	Observed age-class ($c = 1, \ldots, A$)
d	1	True age-class ($d = 1, \ldots, A$)
		Indices
i	2 +	Fish index ($i = 1, \ldots, I$)
j	2 +	Reader index ($j = 1, \ldots, J$)
k	1	Fish index ($k = 1, \ldots, K$)
m	1, 2 +	Parameter index for $\mathbf{\Phi}$ ($m = 1, \ldots, M$)
n	1, 2 +	Parameter index for $\mathbf{\Theta}$ ($n = 1, \ldots, N$)
		Age vectors and matrices
a_{ij}	2 +	Age-class of fish i according to reader j
$\mathbf{A}_i$	2 +	Vector $(a_{ij})_{j=1,\ldots,J}$ of observed age-classes for fish i
$\mathbf{A}$	2 +	Matrix $[a_{ij}]_{i=1,\ldots,I;\, j=1,\ldots,J}$ of all observed age-classes
b_i	2 +	True age-class of fish i
$\mathbf{B}$	2 +	Vector $(b_i)_{i=1,\ldots,I}$ of true fish age-classes
c_k	1	Observed age-class of fish k ($k = 1, \ldots, K$)
$\mathbf{C}$	1	Vector $(c_k)_{k=1,\ldots,K}$ of observed fish age-classes
f_c	1	Number of observations on age-class c
d_k	1	True age-class of fish k ($k = 1, \ldots, K$)
$\boldsymbol{D}$	1	Vector $(d_k)_{k=1,\ldots,K}$ of true fish age-classes
		Parameters and probabilities
$\mathbf{\Phi}$	1, 2 +	Parameter vector that determines the classification matrix $\mathbf{Q}$
ϕ_m	1, 2 +	Component of $\mathbf{\Phi}$ ($m = 1, \ldots, M$)
$\mathbf{\Theta}$	1, 2 +	Complete parameter vector for a given model
θ_n	1, 2 +	Component of $\mathbf{\Theta}$ ($n = 1, \ldots, N$)
p^*_c	1, 2 +	Observed proportion of fish in age-class c
$\mathbf{P}^*$	1, 2 +	Vector $(p^*_c)_{c=1,\ldots,A}$ of observed age proportions
p_b	1, 2 +	Probability of true age b
$\mathbf{P}$	1, 2 +	Vector $(p_b)_{b=1,\ldots,A}$ of true age probabilities
$q(a \mid b, \mathbf{\Phi})$	1, 2 +	Probability of reading age-class a, given the true age-class b
$\mathbf{Q}(\mathbf{\Phi})$	1, 2 +	Classification matrix $[q(a \mid b, \mathbf{\Phi})]_{a=1,\ldots,A,\, b=1,\ldots,A}$

that a sample of age structures has been collected from a fish stock and, in general, that multiple independent age readings have been obtained for each structure in the sample. Repeated readings by a single reader could serve as independent observations, provided that the experimental design minimizes the possibility of recognizing previously read structures. Alternatively, readings could be conducted by different readers following standardized methods. At the end of this section, we consider a more complex case, where multiple readings are obtained for all age structures in a subsample and the remaining structures are assigned at random to one of the readers.

Our initial analysis requires a total of J readings for each structure, with $J > 1$. For each reading j ($j = 1, \ldots, J$), age

a_{ij} defines the age assigned to fish i ($i = 1, \ldots, I$), where a can take values from 1 to A. Although we use the term "age", our analysis treats a as an integer-valued age-class. Thus, the values $a = 1$ and $a = A$ may refer to specified minimum and maximum ages, respectively. The integer vector $\mathbf{A_i} = (a_{ij})_{j=1,\ldots,J}$ includes the ages assigned to fish i by the J readings, and the integer matrix $\mathbf{A} = [a_{ij}]_{i=1,\ldots,I;\, j=1,\ldots,J}$ represents the entire observed data set of multiple readings of all age structures.

Associated with each fish i is a true (but generally unknown) age b_i, and the integer vector $\mathbf{B} = (b_i)_{i=1,\ldots,I}$ represents the true ages of all fish in the sample. For each b, p_b denotes the probability that a fish in the sample has true age b; thus, the vector $\mathbf{P} = (p_b)_{b=1,\ldots,A}$ is subject to the constraints

$$(1.1) \qquad \sum_{b=1}^{A} p_b = 1; \quad p_b \geq 0.$$

The number A of ages defines the length of the real vector $\mathbf{P}$. Also, entries in the integer vector $\mathbf{B}$ range from 1 to A, while the length of $\mathbf{B}$ is the number I of fish in the sample.

The relationship between observed and true fish ages is described by a classification matrix $\mathbf{Q}(\mathbf{\Phi}) = [q(a|b,\mathbf{\Phi})]_{a=1,\ldots,A;\, b=1,\ldots,A}$, which may be defined by a generic parameter vector $\mathbf{\Phi} = (\phi_m)_{m=1,\ldots,M}$. For a specified $\mathbf{\Phi}$, each matrix entry describes the probability $q(a|b,\mathbf{\Phi})$ of assigning age a to a fish, given its true age b. These probabilities are subject to the constraints

$$(1.2) \qquad q(a|b,\mathbf{\Phi}) \geq 0$$

$$(1.3) \qquad \sum_{a=1}^{A} q(a|b,\mathbf{\Phi}) = 1 \text{ for each } b.$$

As discussed in the introduction, our analysis requires the additional assumption

$$(1.4) \qquad q(b|b,\mathbf{\Phi}) > q(a|b,\mathbf{\Phi});\ a \neq b.$$

We refer to (1.4) as the "modal" assumption, since we assume that the true age is assigned with the highest (modal) probability.

Although we have represented entries in $\mathbf{Q}(\mathbf{\Phi})$ by the notation $q(a|b,\mathbf{\Phi})$, we could have used indexed notation $q_{ab}(\mathbf{\Phi})$ to emphasize the matrix character of $\mathbf{Q}(\mathbf{\Phi})$. Thus, we assume that rows and columns of $\mathbf{Q} = [q_{ab}]$ correspond to observed and true ages, respectively. Assumption (1.3) requires that columns of $\mathbf{Q}$ sum to 1, while (1.4) states that the modal probability in each column occurs along the diagonal, where $a = b$.

In practice, the dependency of $\mathbf{Q}$ on $\mathbf{\Phi}$ must be expressed analytically. We consider two possible representations for $\mathbf{Q}(\mathbf{\Phi})$. In the first we let

$$(1.5) \qquad \mathbf{\Phi} = (\sigma_1, \sigma_A, \alpha)$$

and define $q(a|b,\mathbf{\Phi})$ by the sequence

$$(1.6) \qquad \sigma(b) = \begin{cases} \sigma_1 + (\sigma_A - \sigma_1)\dfrac{1 - e^{-\alpha(b-1)}}{1 - e^{-\alpha(A-1)}}; & \alpha \neq 0 \\[2ex] \sigma_1 + (\sigma_A - \sigma_1)\dfrac{b-1}{A-1}; & \alpha = 0 \end{cases}$$

$$(1.7) \qquad x_{ab}(\mathbf{\Phi}) = \frac{1}{\sqrt{2\pi}\sigma(b)} e^{-\frac{1}{2}\left[\frac{a-b}{\sigma(b)}\right]^2}$$

$$(1.8) \qquad q(a|b,\mathbf{\Phi}) = \frac{x_{ab}(\mathbf{\Phi})}{\sum_{a=1}^{A} x_{ab}(\mathbf{\Phi})}.$$

In this scheme, $\sigma(b)$ in (1.6) can be loosely interpreted as the standard deviation of the observation a in (1.7). The parameters σ_1 and σ_A provide lower and upper bounds to $\sigma(b)$ for true ages $b = 1$ and $b = A$, respectively, where $\sigma_1 > 0$ and $\sigma_A > 0$. Thus, $\sigma(1) = \sigma_1$ and $\sigma(A) = \sigma_A$. The parameter α determines nonlinearity of the function $\sigma(b)$ in (1.6) (Schnute 1981), where $\sigma(b)$ becomes linear in b as α tends to 0. For a given b, the right side of (1.7) is maximized when $a = b$; thus, the mode in each column of $\mathbf{X} = [x_{ab}]$ occurs on the diagonal, consistent with the assumption (1.4) for $\mathbf{Q}$. Furthermore, the condition (1.8) guarantees that $\mathbf{Q}(\mathbf{\Phi})$ satisfies (1.3). Consequently, starting from $\mathbf{\Phi}$ in (1.5), the sequence (1.6)–(1.8) defines a matrix $\mathbf{Q}(\mathbf{\Phi})$ that conforms to the necessary criteria (1.2)–(1.4). The parameter $\sigma(b)$ differs from a true normal standard deviation in this context because we have used the density function (1.7) directly, rather than integrals of this function. We refer to (1.6)–(1.8) as a "normal model" for $\mathbf{Q}$.

Our second representation of $\mathbf{Q}(\mathbf{\Phi})$ depends on the four parameters

$$(1.9) \qquad \mathbf{\Phi} = (\sigma_1, \sigma_A, \alpha, \beta).$$

In this scheme, which we call an "exponential model", $\mathbf{Q}$ is determined sequentially by (1.6),

$$(1.10) \qquad x_{ab}(\mathbf{\Phi}) = \sigma(b)^{|a-b|^\beta},$$

and (1.8), subject to the constraints

$$(1.11) \qquad 0 < \sigma_1 < 1, \quad 0 < \sigma_A < 1.$$

Assumptions (1.6) and (1.11) guarantee that $\sigma(b) < 1$; consequently, for each b, the right side of (1.10) is maximum when $a = b$. The extra parameter β controls the extent to which the classification matrix $\mathbf{Q}$ is dominated by its diagonal entries; as $\beta \to \infty$, $\mathbf{Q}$ becomes an identity matrix.

The concepts underlying these two representations of $\mathbf{Q}(\mathbf{\Phi})$ can readily be generalized. If $\mathbf{X} = [x_{ab}]$ is any matrix with $x_{ab} \geq 0$ satisfying the modal assumption, then the matrix $\mathbf{Q}$ defined by (1.8) is a classification matrix. The appropriate formulation of $\mathbf{Q}$ obviously depends on the data structure. Our parsimonious representations may be adequate for many (but not all) situations.

Our discussion so far assumes that the observed data consist of J age readings for each of I fish structures. This design is too narrow to encompass the range of practical problems associated with ageing error. We need to include the possibility that, for some structures, only individual readings are available. For example, with large samples it may be practical to obtain multiple readings for a subsample of fish and individual readings for the remainder. In this design, each of I structures is examined J times, and the remaining K structures are examined once. (If readings are performed by multiple readers, then the K structures would be assigned to readers at random.) We discuss the merits of this approach in our worked example below.

To extend our analysis to such experimental designs, we need to introduce additional notation for structures that have been examined once. Let c_k defne the age ($c = 1, \ldots, A$) assigned to fish k ($k = 1, \ldots, K$). The integer vector $\mathbf{C} = (c_k)_{k=1,\ldots,K}$ includes these observed ages. In the example above, the observed data consist of the matrix $\mathbf{A}$ of multiple age readings and the vector $\mathbf{C}$ of single age readings. We denote corresponding true ages by $\mathbf{B}$ and $\mathbf{D} = (d_k)_{k=1,\ldots,K}$, respectively, so that d_k is the true age of a fish with observed age c_k. In this context, the data $\mathbf{A}$ are analogous, but not identical, to a learning sample in a mixture problem. As we

TABLE 2. Six possible questions associated with parameters estimated from the data **A** and **C**.

1. *Age one fish*: Given the vector $\mathbf{A}_i$ for fish i and the classification matrix **Q**, what is an estimate of the true age b_i?
2. *Estimate the classification matrix and age all fish*: Given the matrix **A**, what are estimates of **Q** and **B**?
3. *Estimate the classification matrix and the true age probabilities*: Given the matrix **A**, what are estimates of **Q** and **P**?
4. *Age all additional fish*: Given the vector **C** and the classification matrix **Q**, what is an estimate of **D**?
5. *Estimate age probabilities from additional data*: Given the vector **C** and the classification matrix **Q**, what is an estimate of **P**?
6. *Estimate the classification matrix and true age probabilities from all available data*: Given the matrix **A** and the vector **C**, what are estimates of **Q** and **P**?

show below, the combination **A** and **C** can be used to estimate simultaneously the classification matrix $\mathbf{Q}(\mathbf{\Phi})$ and the probabilities **P**.

1B. Likelihood Functions

Fish age data can be used for many purposes, and appropriate analytical procedures depend on the question of interest. To focus our discussion and relate our results to the existing literature, we present a list of typical questions in Table 2. Table 3 distinguishes these questions mathematically, depending on the input data, the complete model parameter vector $\mathbf{\Theta} = (\theta_n)_{n=1,\ldots,N}$, and whether or not the classification matrix $\mathbf{Q}(\mathbf{\Phi})$ is presumed known (i.e. estimated from other data).

Question 1 asks for an estimate of the age of fish i, given **Q** and the vector $\mathbf{A}_i$ of observations. In this case, $\mathbf{\Theta} = (b_i)$, so that only one parameter is to be estimated. If ages of all fish are required, then there must be one parameter for each fish. Furthermore, if the classification matrix **Q** is also unknown, as in question 2, then an additional M parameters $\mathbf{\Phi} = (\phi_m)_{m=1,\ldots,M}$ are necessary to specify $\mathbf{Q}(\mathbf{\Phi})$. For stock assessment applications, an estimate of the vector **P** of true age probabilities is usually more important than an estimate of **B**. As clarified by questions 2 and 3, estimates of **B** and **P** involve different parameter vectors $\mathbf{\Theta}$ and correspondingly different numbers of parameters. Questions 4–6 relate to the information gained from single readings of additional structures. With **Q** given, the data **C** could be used to estimate the true age vector **D** (question 4) or the true age probabilities **P** (question 5). If all available data **A** and **C** are used, then both $\mathbf{\Phi}$ and **P** can be estimated simultaneously. Answers to questions 3 and 6 involve identical parameters, but the additional data **C** enable a larger sample size for question 6.

The estimation problems in Tables 2 and 3 can involve large numbers of parameters. In theory, $\mathbf{\Phi}$ could contain $M = A(A-1)$ parameters, since by (1.3), $A - 1$ parameters determine each column of $\mathbf{Q}(\mathbf{\Phi})$. Because **P** involves an additional $A - 1$ parameters, questions 3 and 6 could be formulated with as many as $N = A(A-1) + A - 1 = A^2 - 1$ parameters. In practice, the length of $\mathbf{\Phi}$ depends on the assumed structure of $\mathbf{Q}(\mathbf{\Phi})$. A parsimonious description, such as (1.5)–(1.8), may be more justifiable statistically.

Precise answers to the questions posed in Table 2 depend on likelihood functions (Table 4) that relate observed data to unknown parameters. The simplest likelihood [4.1] is defined by the probability that reading j results in age a_{ij} being assigned to fish i, given $\mathbf{\Phi}$ and the true age b_i. Since we assume age readings to be independent, the product of likelihoods [4.1] across readings gives the joint probability [4.2] of observing the data vector $\mathbf{A}_i$. Thus, [4.2] is the likelihood required to answer question 1. Similarly, question 2 is addressed by the product [4.3], which relates the data matrix **A** to the parameter vector $\mathbf{\Theta} = (\mathbf{\Phi},\mathbf{B})$.

Likelihoods [4.4]–[4.5] include a summation because they describe the compounded effect of two statistical processes. First, for any fish i, the true age b occurs with probability p_b. Second, given b, the vector $\mathbf{A}_i$ of observed ages occurs with probability $\prod_{j=1}^{J} q(a_{ij}|b,\mathbf{\Phi})$. The sum over all possible true ages b in [4.4] accounts for the total probability assigned to $\mathbf{A}_i$. An additional product across fish i in [4.5] addresses question 3.

Likelihoods [4.6]–[4.10] incorporate the additional data **C** from fish aged by a single reading of an age structure. The likelihood [4.6] is analogous to [4.1] and defines the probability that age c_k is assigned to fish k, given $\mathbf{\Phi}$ and its true age d_k. The product [4.7] of probabilities [4.6] across all fish k gives the likelihood required to answer question 4. The likelihood [4.8] involves a sum over d, weighted by the probability p_d, as in [4.4]. Thus, the answer to question 5, which requires an estimate of **P**, is addressed by the product [4.9]. In general, the additional data **C** are used in conjunction with the matrix **A** of multiple age readings. The likelihood [4.10] relates both **A** and **C** to the parameters $\mathbf{\Theta} = (\mathbf{\Phi},\mathbf{P})$, as required to address question 6. Since the data **A** and **C** are independent, the joint likelihood [4.10] is formed from the product of likelihoods [4.5] and [4.9].

1C. Statistical Inference

The likelihood functions $L(\mathbf{\Theta})$ cited in Table 3 can be used to obtain inferences on the parameters $\mathbf{\Theta}$. For most purposes, it is convenient to use twice the negative log likelihood

TABLE 3. Data, number of observations, role of **Q**, parameter vector $\mathbf{\Theta}$ to be estimated, dimension N of $\mathbf{\Theta}$, and likelihood function associated with each of the six questions in Table 2. The likelihood function is referenced by the corresponding equation number in Table 4.

Question	Data	Observations	**Q** known?	$\mathbf{\Theta}$	N	Likelihood
1	$\mathbf{A}_i$	J	Yes	b_i	1	[4.2]
2	**A**	$I \times J$	No	$\mathbf{\Phi},\mathbf{B}$	$M + I$	[4.3]
3	**A**	$I \times J$	No	$\mathbf{\Phi},\mathbf{P}$	$M + A - 1$	[4.5]
4	**C**	K	Yes	**D**	K	[4.7]
5	**C**	K	Yes	**P**	$A - 1$	[4.9]
6	**A**,**C**	$I \times J + K$	No	$\mathbf{\Phi},\mathbf{P}$	$M + A - 1$	[4.10]

TABLE 4. Likelihoods associated with various estimation problems. To simplify the notation, $\mathbf{Q}(\mathbf{\Phi})$ is expressed as $\mathbf{Q}$.

[4.1]	$L(a_{ij} \mid b_i, \mathbf{Q}) = q(a_{ij} \mid b_i, \mathbf{\Phi})$
[4.2]	$L(\mathbf{A}_i \mid b_i, \mathbf{Q}) = \prod_{j=1}^{J} q(a_{ij} \mid b_i, \mathbf{\Phi})$
[4.3]	$L(\mathbf{A} \mid \mathbf{B}, \mathbf{Q}) = \prod_{i=1}^{I} \prod_{j=1}^{J} q(a_{ij} \mid b_i, \mathbf{\Phi})$
[4.4]	$L(\mathbf{A}_i \mid \mathbf{P}, \mathbf{Q}) = \sum_{b=1}^{A} p_b \prod_{j=1}^{J} q(a_{ij} \mid b, \mathbf{\Phi})$
[4.5]	$L(\mathbf{A} \mid \mathbf{P}, \mathbf{Q}) = \prod_{i=1}^{I} \sum_{b=1}^{A} p_b \prod_{j=1}^{J} q(a_{ij} \mid b, \mathbf{\Phi})$
[4.6]	$L(c_k \mid d_k, \mathbf{Q}) = q(c_k \mid d_k, \mathbf{\Phi})$
[4.7]	$L(\mathbf{C} \mid \mathbf{D}, \mathbf{Q}) = \prod_{k=1}^{K} q(c_k \mid d_k, \mathbf{\Phi})$
[4.8]	$L(c_k \mid \mathbf{P}, \mathbf{Q}) = \sum_{d=1}^{A} q(c_k \mid d, \mathbf{\Phi})\, p_d$
[4.9]	$L(\mathbf{C} \mid \mathbf{P}, \mathbf{Q}) = \prod_{k=1}^{K} \sum_{d=1}^{A} q(c_k \mid d, \mathbf{\Phi})\, p_d$
[4.10]	$L(\mathbf{A}, \mathbf{C} \mid \mathbf{P}, \mathbf{Q}) = L(\mathbf{A} \mid \mathbf{P}, \mathbf{Q})\, L(\mathbf{C} \mid \mathbf{P}, \mathbf{Q})$

$$(1.12) \quad l(\mathbf{\Theta}) = -2 \log L(\mathbf{\Theta})$$

which we term the "inference function". Estimates $\hat{\mathbf{\Theta}}$ of the parameters $\mathbf{\Theta}$ are obtained by maximizing $L(\mathbf{\Theta})$, or equivalently, minimizing $l(\mathbf{\Theta})$ in (1.12). Thus, $\hat{\mathbf{\Theta}}$ is determined by

$$(1.13) \quad l(\hat{\mathbf{\Theta}}) = \min_{\mathbf{\Theta}} l(\mathbf{\Theta}).$$

Some parameter estimates may be conditional on other specified parameter values, and we indicate these with the symbol "~". For example, let $\mathbf{\Theta} = (\mathbf{\Theta}_1, \mathbf{\Theta}_2)$, where $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ have dimension N_1 and N_2, respectively, with $N_1 + N_2 = N$. Given $\mathbf{\Theta}_1$, the corresponding estimate $\tilde{\mathbf{\Theta}}_2$ and minimum value $\tilde{l}$ of the inference function are defined by

$$(1.14) \quad \tilde{l}(\mathbf{\Theta}_1) = l(\mathbf{\Theta}_1, \tilde{\mathbf{\Theta}}_2(\mathbf{\Theta}_1)) = \min_{\mathbf{\Theta}_2} l(\mathbf{\Theta}_1, \mathbf{\Theta}_2).$$

In general, we use $l_i(\mathbf{\Theta})$ or $\tilde{l}_i(\mathbf{\Theta}_1)$ to denote an inference function corresponding to question i. For example,

$$(1.15) \quad l_2(\mathbf{\Theta}) = -2 \log L(\mathbf{A} \mid \mathbf{B}, \mathbf{Q})$$

and

$$(1.16) \quad l_3(\mathbf{\Theta}) = -2 \log L(\mathbf{A} \mid \mathbf{P}, \mathbf{Q})$$

represent the inference functions addressing questions 2 and 3, respectively.

Contours of $\tilde{l}(\mathbf{\Theta}_1)$ determine confidence regions for $\mathbf{\Theta}_1$, with $\mathbf{\Theta}_2$ considered a vector of nuisance parameters. Since twice a difference in log likelihoods is asymptotically χ^2-distributed, an approximate probability level can be associated with the boundary of a confidence region. The true value of $\mathbf{\Theta}_1$ in (1.14) lies inside the contour

$$(1.17) \quad \tilde{l}(\mathbf{\Theta}_1) = \tilde{l}(\hat{\mathbf{\Theta}}_1) + \chi^2_\omega(N_1)$$

with approximate probability ω, where $\chi^2_\omega(N_1)$ is the ω-level for the χ^2-distribution with N_1 degrees of freedom. Thus, the difference $\tilde{l}(\mathbf{\Theta}_1) - \tilde{l}(\hat{\mathbf{\Theta}}_1)$ is scaled to probability units by the χ^2-distribution.

Our analysis often depends on the special choice $\mathbf{\Theta}_1 = \mathbf{\Phi}$ in (1.14). For example, in question 1, $\tilde{b}_i(\mathbf{\Phi})$ denotes the estimated true age of fish i conditioned on $\mathbf{\Phi}$. Thus, given $\mathbf{\Phi}$, $\tilde{l}_1(\mathbf{\Phi})$ denotes the inference function minimized with respect to the true age b_i.

Uncertainty in $\hat{\mathbf{\Theta}}$ can be assessed from the estimated asymptotic covariance matrix

$$(1.18) \quad \hat{V}[\hat{\mathbf{\Theta}}] = \left[\frac{1}{2} \left(\frac{\partial^2 l(\hat{\mathbf{\Theta}})}{\partial \theta_i\, \partial \theta_j} \right)_{i=1,\ldots,N;\, j=1,\ldots,N} \right]^{-1}$$

derived from asymptotic properties of $\hat{\mathbf{\Theta}}$ in (1.13) (Kendall and Stuart 1979, p. 59–60). Estimated standard errors $\hat{\mathrm{SE}}[\hat{\theta}_n]$ for the components of $\hat{\mathbf{\Theta}}$ are then given by $\sqrt{\hat{V}_{nn}}$, where $\hat{V}_{nn}$ is a diagonal component of (1.18).

Another statistical issue pertains to selecting an appropriate model structure for $\mathbf{Q}$. In the example below, we examine representations for $\mathbf{Q}(\mathbf{\Phi})$ obtained from both (1.7) and (1.10). We use the Akaike information criterion (Akaike 1974)

$$(1.19) \quad \mathrm{AIC} = l(\hat{\mathbf{\Theta}}) + 2N$$

to simplify the model identification process. A model with a low AIC value relative to other comparable models is considered to provide a good model fit. The term $2N$ in (1.19) acts as a penalty for the number of parameters; a model is favored if it can achieve a low value of $l(\hat{\mathbf{\Theta}})$ with few parameters. Akaike's (1974) method can be used for informal comparisons, where a model is a candidate for the best fit if its AIC value lies within 2 units of the minimum AIC value. More formal hypothesis tests can be based on the χ^2-distribution in (1.17).

1D. Special Cases

In the literature on stock composition models, various methods have been proposed for obtaining parameter estimates and corresponding variances. For example, Hoenig and Heisey (1987) and Kimura and Chikuni (1987) use the EM algorithm. Other authors (e.g. Pella and Robertson 1979) employ a classification-correction procedure. In their appendix, Wood et al. (1987) review some of the proposed methods and present arguments favoring the direct use of maximum likelihood. This approach requires no specialized techniques; a generic minimization algorithm is sufficient. We follow this approach in the analysis here (Appendix). Furthermore, we use straightforward numerical derivatives to evaluate $\hat{V}[\hat{\mathbf{\Theta}}]$ in (1.18).

The vectors $\mathbf{\Phi}$ and $\mathbf{P}$ are real and their estimates can be found with the aid of any search algorithm for function minimization across a real domain. By contrast, $\mathbf{B}$ and $\mathbf{D}$ are integer vectors. Their components can be estimated by evaluating the appropriate inference function l at all possible discrete values $(1, \ldots, A)$ and finding the minimum. For example, question 1 requires estimation of only one integer parameter (b_i), and this estimate can be obtained by examining the A possible values of the inference function.

Question 2 involves both real and integer parameters. To perform the necessary estimation from [4.3], first assume that $\mathbf{\Phi}$ and $\mathbf{Q} = \mathbf{Q}(\mathbf{\Phi})$ are given and use ordinary minimization to answer question 1 for each i; thus, find the estimate $\tilde{b}_i(\mathbf{Q})$ that minimizes the inference function obtained from [4.2]. Since [4.3] is a product of factors [4.2], the estimate

$$(1.20) \quad \tilde{\mathbf{B}}(\mathbf{Q}) = (\tilde{b}_i(\mathbf{Q}))_{i,\ldots,I}$$

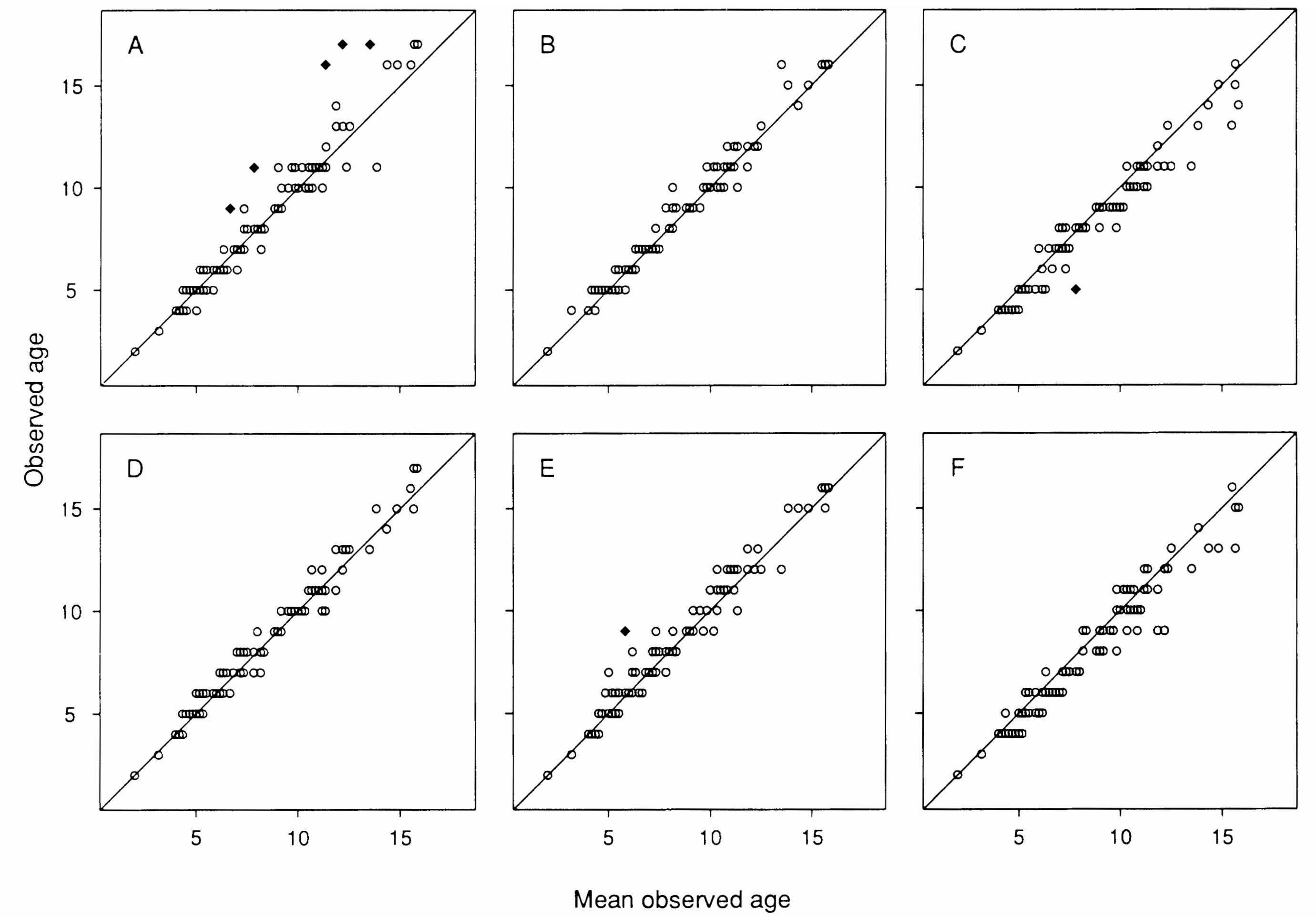

FIG. 1. Relationship between observed fish age a_{ij} and mean observed age $\frac{1}{J}\sum_{j=1}^{J} a_{ij}$ for readers A–F. A line of slope 1 is included for reference in each panel. The solid diamonds correspond to outlying points in Fig. 3.

minimizes the inference function corresponding to [4.3]. From (1.20), define the inference function

$$(1.21) \quad \tilde{l}(\mathbf{A}|\mathbf{Q}) = l(\mathbf{A}|\tilde{\mathbf{B}}(\mathbf{Q}),\mathbf{Q})$$

which depends only on $\mathbf{Q}$ and thus on the real parameter vector $\mathbf{\Phi}$. Finally, minimize $\tilde{l}(\mathbf{A}|\mathbf{Q}(\mathbf{\Phi}))$ as a function of $\mathbf{\Phi}$ to obtain the estimates $\hat{\mathbf{\Phi}}$, $\hat{\mathbf{Q}} = \mathbf{Q}(\hat{\mathbf{\Phi}})$, and $\hat{\mathbf{B}} = \tilde{\mathbf{B}}(\hat{\mathbf{Q}})$. These estimates resolve question 2.

The parameters $\mathbf{\Phi}$ and $\mathbf{P}$ of question 3 are real. However, estimation of $\hat{\mathbf{P}}$ can be difficult if one or more (unknown) components of $\mathbf{P}$ are 0. To ensure that the constraints (1.1) are met during minimization, we use surrogate parameters z_b, defined such that

$$(1.22) \quad p_b = \frac{1}{2}(1 - \cos \pi z_b).$$

Thus, values z_b in the range (0,1) or, more generally, in the range $(-\infty, \infty)$ correspond to values p_b in the range (0,1). We choose an age b_0 for which p_{b_0} appears relatively large, use (1.22) for $b \neq b_0$, and then compute

$$(1.23) \quad p_{b_0} = 1 - \sum_{b \neq b_0} p_b.$$

Thus, the parameter vector $\mathbf{P}$ is determined by $A - 1$ surrogate parameters z_b with $b \neq b_0$.

Question 4 requires estimating the integer vector $\mathbf{D}$, given $\mathbf{Q}$ and $\mathbf{C}$. In this case, there is precisely one parameter d_k for each observation c_k, and the estimate $\hat{d}_k$ maximizes $q(c_k|d)$ as a function of d. Thus, the estimate can be obtained by inspecting row c_k of the matrix $\mathbf{Q}$. Incidentally, the modal hypothesis (1.4) does not imply that $\hat{d}_k = c_k$, because (1.4) asserts only that the column maximum occurs on the diagonal of $\mathbf{Q}$.

For question 5, the estimates $\hat{\mathbf{P}}$ sometimes have a simple analytical representation. In general, the number K of additional fish is much greater than the number A of ages. If f_c represents the frequency of observing age c, then the observed proportion of age c fish is given by the ratio

$$(1.24) \quad p_c^* = \frac{f_c}{\sum_{c=1}^{A} f_c}.$$

The observed proportions (1.24) define the vector $\mathbf{P}^* = (p_c^*)_{c=1,\ldots,A}$. The likelihood

$$(1.25) \quad L(\mathbf{C}|\mathbf{P},\mathbf{Q}) = \prod_{c=1}^{A} \left[\sum_{d=1}^{A} q(c|d,\mathbf{\Phi})\, p_d \right]^{f_c}$$

is equivalent to [4.9] and accounts for the grouping of observations. A simple proof (Wood et al. 1987, appendix, proposition 1) based on (1.25) gives

TABLE 5. Estimates of the parameters $\Phi = (\sigma_1, \sigma_A, \alpha, \beta, \gamma)$ for questions 2 and 3, based on the normal (Norm), exponential (Exp), and reader effects (Rdr) models. The corresponding AIC and minimum values of the inference function are also given for comparison. Different cases represent the full models and models reduced by the constraints $\alpha = 0$, $\beta = 1$, or $\gamma = 1$. Values b are listed for which estimates $\hat{p}_b = 0$.

Case	Question	Model	N	$\hat{\sigma}_A$	$\hat{\sigma}_1$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	b with $\hat{p}_b = 0$	$l(\hat{\Theta})$	AIC
1	2	Norm	127	0.495	2.038	−0.166	—	—	—	1645.8	1899.8
2	2	Norm	126	0.369	1.411	0	—	—	—	1650.6	1902.6
3	2	Exp	128	0.080	0.534	−0.012	1.204	—	—	1600.7	1856.7
4	2	Exp	127	0.074	0.523	0	1.204	—	—	1600.7	1854.7
5	2	Exp	127	0.088	0.461	−0.046	1	—	—	1607.3	1861.3
6	3	Norm	18	0.458	1.858	−0.109	—	—	13 15 17	2189.8	2225.8
7	3	Norm	17	0.366	1.472	0	—	—	13 15 17	2191.7	2225.7
8	3	Exp	19	0.101	0.604	−0.061	1.187	—	17	2141.4	2179.4
9	3	Exp	18	0.072	0.537	0	1.184	—	17	2141.7	2177.7
10	3	Exp	18	0.095	0.477	−0.052	1	—	17	2147.0	2183.0
11	3	Rdr	25	0.046	0.352	−0.130	0.849	0.704	13 17	2014.3	2064.3
12	3	Rdr	24	0.031	0.252	0	0.851	0.682	13 17	2015.9	2063.9
13	3	Rdr	24	0.044	0.380	−0.131	0.869	1	13 17	2015.2	2063.2
14	3	Rdr	23	0.028	0.271	0	0.864	1	13 17	2017.1	2063.1

(1.26) $\hat{\mathbf{P}} = \mathbf{Q}^{-1}\mathbf{P}^*$

if the conditions (1.1) are satisfied by $\hat{\mathbf{P}}$. Equation (1.26) was discussed in general terms in the introduction. The inverse classification matrix $\mathbf{Q}^{-1}$ acts on the observed proportions $\mathbf{P}^*$ to reverse the smoothing effect of ageing error and thus to give the estimate $\hat{\mathbf{P}}$.

2. Pollock Example

2A. Data Description

The walleye pollock data were obtained from a reference sample prepared by the Alaska Fisheries Science Center (National Marine Fisheries Service, Seattle, WA). The sample contained 125 otolith pairs from the Gulf of Alaska and four areas of the Bering Sea, collected between October 1984 and February 1985. Each otolith pair was independently examined by six experienced readers from three countries using the break and burn method (Chilton and Beamish 1982).

Our analysis requires a value for the maximum age A. For the pollock data, assigned ages for one fish ranged from 21 to 24 yr and the next oldest assigned age was 17 yr. We excluded data on this fish to avoid extrapolation for missing ages. Thus, we set $A = 17$. Although data from old fish may be important for biological interpretation, our purpose here is illustration of the basic statistical methods. We seek estimates of the classification matrix $\mathbf{Q}(\Phi)$ within an age range where multiple readings are available.

We examined the remaining data by comparing ages assigned by each reader to the mean assigned age (Fig. 1). Readers A and E tend to overestimate and readers C and F underestimate ages relative to the mean. A more complex model, with extra parameters to accommodate reader effects, is presented in section 3. Here, we treat reader effects as part of the overall model variance. Since there are no large consistent deviations by any one reader, the data are adequate to address questions 2 and 3. Observed ages range from 2 to 17 yr, with the variance in assigned age tending to increase with the mean. There is complete agreement on the three fish assigned an age of 2 yr. Following the notation in Table 1, the data consist of $I = 124$ fish, with $J = 6$ readings of each age structure. The formal notation (Table 1) would use $a = 1, \ldots, 16$ to reference the 16 age-classes. For clarity, however, we let the symbol a denote an observed age $a = 2, \ldots, 17$.

2B. Question 2 Analysis

Technical details of model application are deferred to the Appendix. Parameter estimates

(2.1) $\hat{\Phi} = (\hat{\sigma}_1, \hat{\sigma}_A, \hat{\alpha}, \hat{\beta})$

for question 2 are provided in Table 5, cases 1–5, along with values of $l_2(\hat{\Theta})$ and the AIC. These five cases differ in the model used for $\mathbf{Q}$ and constraints imposed on the parameters Φ. For the normal model from (1.7), the full parameter vector (1.5) is included in case 1 whereas α is constrained to 0 in case 2. For the exponential model from (1.10), we consider the full parameter vector (1.9) (case 3) and parameter vectors reduced by the constraints $\alpha = 0$ (case 4) and $\beta = 1$ (case 5). Based on AIC values from (1.19), cases 3 and 4 are the most plausible candidates for the pollock data; AIC values for other cases are more than 6 units above the minimum. Cases 3 and 4 employ the exponential model and they can be compared by formal hypothesis tests using (1.17). It is not possible to reject the hypothesis that $\alpha = 0$, since minimum inference values $l_2(\hat{\Theta})$ for cases 3 and 4 agree to the reported number of decimal places. Thus, case 4 provides the best model fit for question 2.

The success of the exponential model is best understood by comparing the estimates $\mathbf{Q}(\hat{\Phi})$ for normal and exponential models. Estimated probability distributions $q(a|b, \hat{\Phi})$ for case 4 are more peaked than the corresponding distributions for case 1 (Fig. 2), declining rapidly from the mode at $a = b$. Thus, the exponential model leads to relatively leptokurtotic distributions. For most fish, however, the rounded mean observed age, the modal observed age, and the estimated true age $\hat{b}_i$ all agree. Estimates $\hat{\mathbf{B}}$ differ between cases 1 and 4 for only 8 (6%) of the 124 fish. For these fish, the exponential model tends to select the most frequent observation a whereas the normal model is more sensitive to outliers.

Because inference functions in this paper are not expressed as sums of squares, residual analysis in the usual sense is not possible. Nevertheless, the inference function (1.15) for question 2 is computed as a sum of positive terms, where each is the negative log of a probability. To investigate the weight attached to each observation, we define the inference contribution

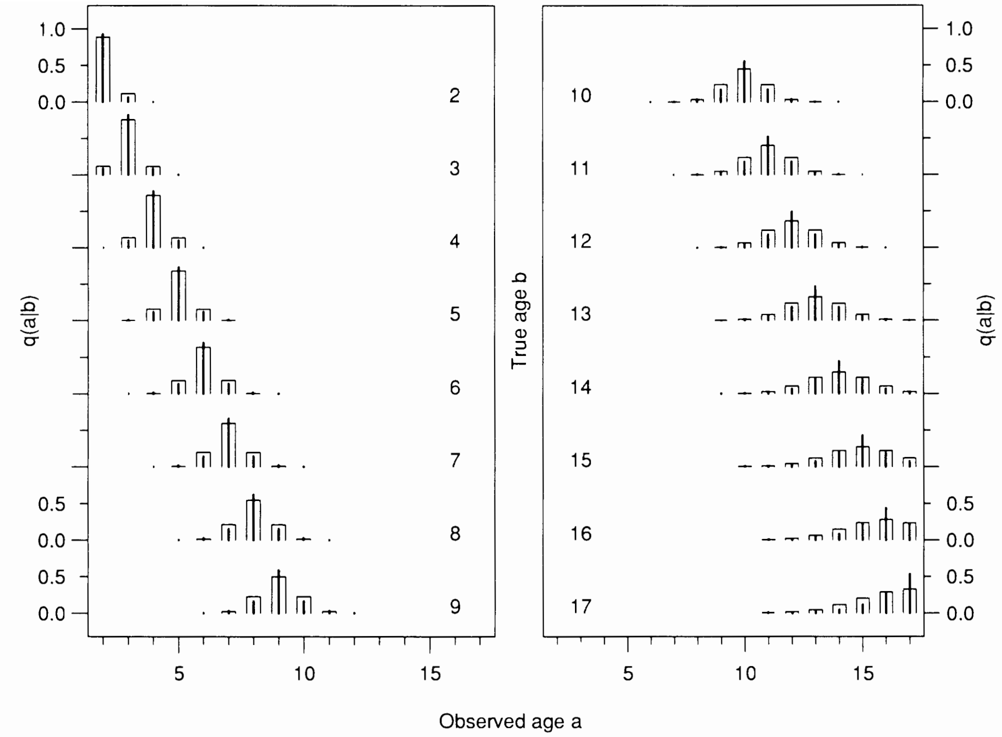

FIG. 2. Relationship between estimated probability $q(a|b,\hat{\Phi})$ and observed age a, given the specified true age b. Open bars represent the normal model (Table 5, case 1) for question 2 and centered vertical lines represent the exponential model (case 4).

(2.2) $\quad \epsilon_{ij} = -2 \log q(a_{ij}|\hat{b}_i, \hat{\Phi})$

from a single term of the inference sum. Figure 3 shows that the contributions ϵ_{ij} increase approximately linearly with the discrepancy

(2.3) $\quad \delta_{ij} = |a_{ij} - \hat{b}_i|$

between observed and estimated age-classes. For the pollock data, the discrepancy δ_{ij} takes integer values between 0 and 6; however, values δ_{ij} in Fig. 3 are jittered to improve the data display. Gaps in the frequency distribution of the contributions arise because ϵ_{ij} in (2.2) can admit only certain values. There are seven observations in Fig. 3 from six fish with contributions $\epsilon_{ij} > 10$. These observations are identified in Fig. 1; five are due to reader A, one to reader C, and one to reader E. Clearly, inconsistencies among readers inflate the inference function (1.15) and increase the uncertainty in $\hat{\mathbf{B}}$.

2C. Question 3 Analysis

As described in the introduction, most ageing error analyses aim to determine true age proportions **P**. Thus, question 3 has greater practical importance than question 2. For the pollock data, the results for question 3 (Table 5) are similar to those obtained for question 2. The exponential model leads to much lower AIC values than the normal model. Case 9 provides the best overall fit; again, we cannot reject the hypothesis that $\alpha = 0$.

Estimates $\hat{\mathbf{P}}$ vary little with final model choice among cases 6–10. Table 6 illustrates these results for cases 8–10, which are based on the exponential model. Differences among these cases relate to the assumed structure of **Q**. Table 7 provides a means of comparing estimates $\hat{\mathbf{Q}}$, where for each age b, the variance $V[a|b]$ can be calculated directly from the corresponding column of $\hat{\mathbf{Q}}$. Variances $V[a|b]$ for cases 8–10 are similar at young ages, but case 9 has a lower variance for older ages. Thus, flexibility in the parameter α allows for larger ageing error at these ages. Our inability to reject the hypothesis that $\alpha = 0$ reflects our small sample size at old ages. This observation has implications for sample design. Sample sizes must be large enough to estimate the distribution $q(a|b)$ for each age b. Typically, ageing error is larger and fewer age structures are available at older ages. Thus, it may be desirable to increase the number of readings of older age structures.

Small sample sizes at older ages also complicate the estimate $\hat{\mathbf{P}}$. In a typical fisheries data set, the exact dimension of **P** is poorly determined; the observed maximum age A is a random variable dependent on small values p_b for large b. For the normal model for **Q** (cases 6–7), we obtain the estimates $\hat{p}_{13} = \hat{p}_{15} = \hat{p}_{17} = 0$. Similarly, for the exponential model for **Q** (cases 8–10), we obtain the estimate $\hat{p}_{17} = 0$. Thus, maximum likelihood estimates for these cases lie on the parameter boundary defined by the constraint (1.1). Nevertheless, all values p_b ($b = 2, \ldots, 17$) are free parameters in the calculation, and we determine N and the AIC accordingly. In essence, our analysis suggests that the proportion of fish older than age 12 is small and poorly determined. A more rigorous analysis might use the new parameter

(2.4) $\quad p_+ = \sum_{b=13}^{17} p_b$

in place of p_{13}, and treat $p_{14}, \ldots, p_{17}$ as nuisance parameters. Although this would circumvent problems associated with the random behaviour of A, it would make the analysis much more complicated technically. We prefer the less formal approach here, with the understanding that information on proportions p_b with $b \geq 13$ is really limited to the composite parameter p_+ in (2.4). Estimates $\hat{p}_+$ are indeed similar among cases 8–10 (Table 6).

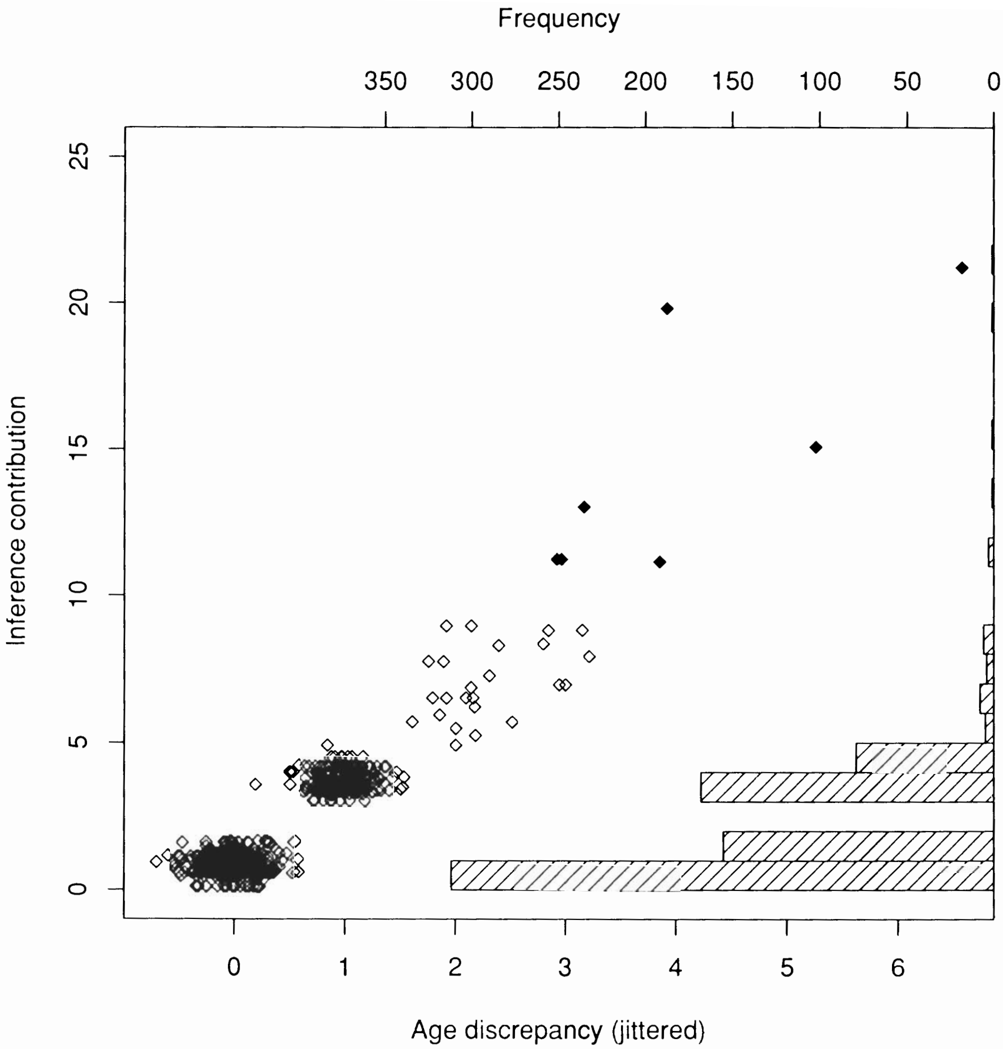

FIG. 3. Relationship between inference contributions ϵ_{ij} in (2.2) and the age discrepancy δ_{ij} in (2.3), based on the model in Table 5, case 4. The quantity δ_{ij} is integer-valued, but has been jittered to better display the data. Solid diamonds indicate values $\epsilon_{ij} > 10$. Horizontal bars reflect the frequency distribution of ϵ_{ij} values.

An informal residual analysis for question 3 can be conducted as for question 2. The positive inference contribution for one fish,

$$(2.5) \qquad \zeta_i = -2 \log \left[\sum_{b=1}^{A} \hat{p}_b \prod_{j=1}^{J} q(a_{ij}|b, \hat{\mathbf{\Phi}}) \right]$$

increases approximately linearly with

$$(2.6) \qquad \eta_i = \sum_{j \neq j'} |a_{ij} - a_{ij'}|$$

a measure of disparity among readings for fish i (Fig. 4). Since a_{ij} and b are integers, the quantity ζ_i in (2.5) admits only certain values, and gaps are apparent in the distribution of (2.5). A few fish contribute large values (2.5) and thus have high leverage in the inference function. The four fish with values $\zeta_i > 30$ also contribute to outlying points in Fig. 1 and 3.

Approximate confidence intervals for $\hat{p}_b$ can be obtained from $\hat{\mathrm{SE}}(\hat{p}_b)$, determined by the asymptotic covariance matrix in (1.18), or from the relationship (1.17) involving the χ^2-distribution. The two methods lead to 95% intervals of similar width for each value b (Table 8), except where the lower bound lies on the constraints $p_b = 0$. Likelihood intervals from (1.17) are located asymmetrically around the estimates $\hat{p}_b$; midpoints of the interval from (1.18) lie well above the corresponding estimates $\hat{p}_b$. Such asymmetries are intuitively appealing because symmetric intervals defined by $\hat{\mathrm{SE}}(\hat{p}_b)$ would have negative lower bounds in the cases $b = 3$ and $b = 14$. Estimates $\hat{p}_b$ are assumed to be unconstrained in (1.18) and we cannot obtain meaningful measures $\hat{\mathrm{SE}}(\hat{p}_b)$ at the boundary of the [0,1] range for p_b. For example, we cannot use (1.18) to obtain $\hat{\mathrm{SE}}(\hat{p}_{17})$. Nevertheless, because of computational intensity associated with (1.17) (Appendix), we resort to (1.18) as our measure of uncertainty.

If ages were known perfectly, the expected variance of $\hat{\mathbf{P}}$ could be estimated directly from the multinomial distribution, where the multinomial variance estimate $\hat{p}_b(1 - \hat{p}_b)/I$ depends on the number of fish I in the sample. Because ageing error introduces additional uncertainty, the estimated variance $\hat{V}[\hat{p}_b]$ from (1.18) is generally larger than the corresponding multinomial variance. Thus, a larger sample size is required to estimate $\hat{p}_b$ with a specified precision when ageing error is present than when exact ages are known. The relationship between $\hat{V}[\hat{p}_b]$ and $\hat{p}_b(1 - \hat{p}_b)$ is approximately linear, with a slope near

TABLE 6. Percentages corresponding to the estimates $\hat{\mathbf{P}}$ for cases 8–14 in Table 5. The estimate of p_+ in (2.4) is also given for comparison.

	$100 \times \hat{p}_b$						
b	Case 8	Case 9	Case 10	Case 11	Case 12	Case 13	Case 14
2	2.4	2.4	2.4	2.4	2.4	2.4	2.4
3	0.8	0.8	0.8	0.8	0.8	0.8	0.8
4	3.9	3.8	3.9	3.4	3.4	3.4	3.3
5	23.8	23.8	23.7	24.3	24.3	24.3	24.4
6	5.7	5.7	5.7	5.9	5.8	5.9	5.8
7	19.2	19.2	19.4	19.3	19.3	19.3	19.3
8	5.4	5.4	5.2	5.1	5.1	5.1	5.1
9	6.3	6.3	6.3	6.2	6.3	6.3	6.3
10	9.0	9.1	9.3	9.9	10.0	9.9	9.9
11	12.4	12.4	11.9	10.3	10.3	10.4	10.5
12	4.3	4.3	4.2	6.7	6.7	6.6	6.6
13	0.9	1.0	1.3	0.0	0.0	0.0	0.0
14	0.9	1.0	0.7	1.1	1.2	1.0	1.1
15	1.8	1.7	2.0	1.4	1.3	1.5	1.4
16	3.1	3.1	3.0	3.2	3.2	3.2	3.2
17	0.0	0.0	0.0	0.0	0.0	0.0	0.0
+	6.8	6.8	7.0	5.7	5.7	5.7	5.7

$1/I$ (Fig. 5). The estimated variance for ages 2 and 3 conforms very nearly to the multinomial prediction because these ages are measured with little error. Thus, when $b = 2$ or $b = 3$, $q(a|b)$ is nearly 1 when $a = b$ and 0 otherwise. For $b > 3$, $\hat{V}[\hat{p}_b]$ exceeds the multinomial prediction, where large deviations for ages 10–12 correspond to high levels of ageing error. The small deviation for age 16 is an artifact of our analysis, due to the fact that $\hat{\mathbf{P}}$ lies on the constraint $p_{17} = 0$. Thus, if age 17 is deemed impossible, then increased certainty is attached to a reading $a_{ij} = 16$.

In summary, this analysis of variance in $\hat{\mathbf{P}}$ shows that precision depends on the number I of fish age structures, rather than the total number of readings IJ. Multiple readings of individual structures are informative for $\mathbf{Q}$, but not for $\mathbf{P}$. Furthermore, as ageing error increases, data from individual age structures become less and less informative for $\mathbf{P}$. For this reason, we emphasize that sample sizes must be increased with increasing ageing error to yield a specified precision in $\hat{\mathbf{P}}$.

As discussed in the introduction, the true age proportions might be estimated naively from the frequency distribution of the estimates $\hat{\mathbf{B}}$ for question 2. Let $\mathbf{P}' = (p'_b)_{b=1,\ldots,A}$ represent age proportions determined from $\hat{\mathbf{B}}$. We have stated earlier that ageing error tends to smooth the observations $\mathbf{P}^*$ and that age proportions $\mathbf{P}'$ do not fully account for the smoothing effect. This is because of uncertainty in $\hat{\mathbf{B}}$. As J increases, $\hat{\mathbf{B}}$ asymptotically approaches its true value under the modal hypothesis (1.4). However, J will be small for most applications, leading to bias in the calculation of $\mathbf{P}'$. Evidence of smoothing in the pollock data is illustrated in Fig. 6. For the two dominant ages $b = 5$ and $b = 7$, $p^*_b < p'_b < \hat{p}_b$ as expected. Similarly, $p^*_b > p'_b > \hat{p}_b$ for ages 4, 6, and 8, adjacent to the dominant ages. Thus, the distributions defined by the observations $\mathbf{P}^*$, the naive estimates $\mathbf{P}'$, and the likelihood estimates $\hat{\mathbf{P}}$ become progressively more peaked.

For the pollock data, we are unable to distinguish statistically among $\mathbf{P}^*$, $\mathbf{P}'$, and $\hat{\mathbf{P}}$. For example, each observation p^*_b lies within $2 \times \hat{\mathrm{SE}}(\hat{p}_b)$, and the inference function $l_3(\hat{\mathbf{\Theta}})$ is relatively insensitive to small variations in $\hat{p}_b$. Differences among $\mathbf{P}^*$, $\mathbf{P}'$, and $\hat{\mathbf{P}}$ are still biologically meaningful, however; small deviations in $\hat{\mathbf{P}}$ can have a large impact on any subsequent analyses. Our inability to reject the observations $\mathbf{P}^*$ as estimates of $\hat{\mathbf{P}}$ simply reflects our small sample of 124 fish. The appropriate choice of sample size is an obvious subject for future work.

3. Reader Effects

In section 1, we described general analytical methods needed to estimate ageing error from a sample with multiple age readings. Here, we suggest how the methods might be tailored to a specific data set. We present a more complex model for the pollock data, acknowledging two features of these data not considered above. First, the J readings of each age structure were conducted by the same J readers. Thus, we can examine the data for possible reader effects. Second, the readers agreed completely on the age of the youngest fish. Our revised model acknowledges this by assuming that no ageing error is associated with ages 1 to A_1.

The extended pollock model depends on the $5 + J$ parameters

$$(3.1) \quad \mathbf{\Phi} = (\sigma_1, \sigma_A, \alpha, \beta, \gamma, r_1, \ldots, r_J)$$

subject to the constraints

$$(3.2) \quad \sum_{j=1}^{J} r_j = 0$$

where r_j represents the relative bias of reader j at age A and γ determines the change in bias with age. Based on the reader effects in (3.1), the revised classification matrix must be indexed by reader. We define the classification matrix $\mathbf{Q}_j$ by the sequence

$$(3.3) \quad \sigma(b) = \begin{cases} \sigma_1 + (\sigma_A - \sigma_1)\dfrac{1-e^{-\alpha(b-A_1-1)}}{1-e^{-\alpha(A-A_1-1)}}; & \alpha \neq 0, b > A_1 \\ \sigma_1 + (\sigma_A - \sigma_1)\dfrac{b-A_1-1}{A-A_1-1}; & \alpha = 0, b > A_1 \end{cases}$$

$$(3.4) \quad w(b) = \gamma + (1-\gamma)\frac{b - A_1 - 1}{A - A_1 - 1}$$

$$(3.5) \quad x_{abj}(\mathbf{\Phi}) = \sigma(b)^{|a-b-w(b)r_j|^{\beta}}; \quad \alpha > A_1, b > A_1$$

$$(3.6) \quad q_j(a|b, \mathbf{\Phi}) = \frac{x_{abj}(\mathbf{\Phi})}{\sum_{a=A_1+1}^{A} x_{abj}(\mathbf{\Phi})}; \quad \alpha > A_1, b > A_1$$

$$(3.7) \quad \mathbf{Q}_j = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}'_j \end{pmatrix}$$

where we assume that ages 1 to A_1 can be determined without error. Thus, the classification matrix $\mathbf{Q}_j$ in (3.7) is composed of four submatrices. The $A_1 \times A_1$ identity matrix $\mathbf{I}$ accounts for the youngest A_1 ages. Two rectangular null matrices $\mathbf{0}$ ensure that $q_j(a|b,\mathbf{\Phi}) = 0$ for $a \neq b$ when $b \leq A_1$ or $a \leq A_1$. The $(A - A_1) \times (A - A_1)$ matrix $\mathbf{Q}'_j$ describes ageing error associated with the oldest $A - A_1$ ages.

Equations (3.3)–(3.6), defining elements $q_j(a|b,\mathbf{\Phi})$ of $\mathbf{Q}'_j$, are analogous to (1.6)–(1.10) earlier. The modification (3.3) simply allows $\sigma_1 = \sigma(A_1 + 1)$ to represent the first age where ageing error is present. Reader effects enter in equation (3.5), where $w(b)r_j$ acts as an offset in the definition of $x_{abj}(\mathbf{\Phi})$ for a given b and j. We assume that reader effects vary linearly with age, dependent on the weight function $w(b)$ in (3.4). From (3.4), the offset $w(b)r_j$ takes the values γr_j and r_j when $b = A_1 + 1$ and $b = A$, respectively. Thus, the parameter r_j describes reader effects at the oldest age and the parameter γ describes the reduction in r_j at age $A_1 + 1$. Because $\sigma(b)$ and β are independent of j, reader effects result in only minor changes to the shape of the distribution $q_j(a|b,\mathbf{\Phi})$ for a given b. These changes are due to the mode shift in (3.5), combined with the normalization (3.6).

The likelihood function to address question 3 is given by

$$(3.8) \qquad L(\mathbf{A}|\mathbf{P}, \mathbf{Q}_j) = \prod_{i=1}^{I} \sum_{b=1}^{A} p_b \prod_{j=1}^{J} q_j(a_{ij}|b,\mathbf{\Phi}),$$

a modification of [4.5]. Parameter estimates obtained from the inference function corresponding to (3.8) are listed in Table 5 (cases 11–14). We considered the full model as well as models reduced by the constraints $\alpha = 0$ and $\gamma = 1$. Cases 11–14 agree as to the estimates $\hat{p}_{13} = 0$ and $\hat{p}_{17} = 0$ and have similar AIC values. There is a large decrease in AIC values between cases 8–10 and 11–14. Obviously, reader effects contribute significantly to model fit. The estimates $\hat{\sigma}_1$ and $\hat{\sigma}_A$ are also smaller for cases 11–14, although strict comparisons with cases 8–10 are complicated by a corresponding change in $\hat{\beta}$. Furthermore, σ_1 describes age 2 for cases 8–10 and age 3 for cases 11–14 because of the modification (3.3). The reader effects are small in magnitude (Table 9). With $\gamma = 1$ (cases 13–14), each estimate $\hat{r}_j$ represents a maximum offset in (3.5) of less than 0.5 yr. Estimates are slightly larger in cases 11–12, where the estimates $\hat{\gamma} < 1$ indicate that age-dependent effects $w(b)r_j$ increase with b.

An ageing error analysis is typically conducted to obtain estimates $\hat{\mathbf{P}}$. These estimates are relatively unaffected by the inclusion of reader effects (Table 6). Differences are minor between most corresponding values $\hat{p}_b$ for cases 8–14. The largest differences between the exponential and reader effects models occur for ages 11–12, the ages with the largest deviations in $\hat{V}[\hat{p}_b]$ from the expected multinomial variance in Fig. 5. Reader bias for age 12 is clearly evident from Fig. 1. Readers A and C each assigned age 12 to one structure whereas readers B and E assigned age 12 to 10 and 13 structures, respectively.

Variances $V[a|b]$ for cases 11–14 cannot be calculated directly as in cases 8–10. The exponential model in section 2 involves a single classification matrix $\mathbf{Q}$, while the model here depends on J distinct matrices $\mathbf{Q}_j$. To relate this model to the earlier one, suppose that (for a given b) the random variable a results from compounding two processes. First, reader j is selected at random and then a reading is obtained from the error distribution $q_j(a|b)$ for that reader. The resulting variance of $a|b$ can be expressed as the sum

$$(3.9) \qquad V[a|b] = E_j[V_a[a|b, j]] + V_j[E_a[a|b, j]]$$

where (3.9) is a general result for such compound distributions. Here, $V_a[a|b, j]$ measures the variability of a single reader and the expectation E_j averages this within-reader variabiilty. Similarly, $E_a[a|b, j] - b$ measures the bias of a single reader, and the variance of this quantity (the same as the variance of $E_a[a|b, j]$) measures bias variability among readers. Thus, (3.9) both provides a means of computing $V[a|b]$ for the reader effects model and partitions this variance into components within and among readers.

Variances $V[a|b]$ are similar for exponential and reader effects models, tending to increase with b (Table 7). Higher values $V[a|b]$ are associated with cases 8, 10, 11, and 13, for which α is unconstrained. The partitioning of $V[a|b]$ into components within and among readers is illustrated for case 11 only, but comparable results apply to other cases. The ability to partition this variance accounts for the drop in AIC values between cases 8–10 and cases 11–14. However, the magnitude of $V[a|b]$ can be attributed mainly to within-reader variability. The small among-reader variability tends to increase with b, consistent with our assumption of age-dependent reader effects in cases 11–12. Our general conclusion for this analysis is that reader variability can be partitioned, but not reduced, by extending the exponential model to include reader effects. Because corrections to age proportions stem from the composite ageing error, estimates $\hat{\mathbf{P}}$ (Table 6) are similar between exponential and reader effects models.

The model (3.1)–(3.7) is a simple implementation of reader effects, consistent with the modal hypothesis (1.4). We have

TABLE 7. Calculated variances $V[a|b]$ for cases 8–14 in Table 5. For case 11, the components $E_a[V_j[a|b, j]]$ and $V_j[E_a[a|b, j]]$ of (3.9) are also given, indicated by the columns V_jE_a and E_aV_j, respectively.

	Case 8	Case 9	Case 10	Case 11			Case 12	Case 13	Case 14
b	$V[a\|b]$	$V[a\|b]$	$V[a\|b]$	E_aV_j	V_jE_a	$V[a\|b]$	$V[a\|b]$	$V[a\|b]$	$V[a\|b]$
2	0.10	0.07	0.12	0.00	0.00	0.00	0.00	0.00	0.00
3	0.22	0.19	0.23	0.14	0.01	0.15	0.11	0.16	0.12
4	0.30	0.28	0.33	0.27	0.04	0.31	0.28	0.33	0.30
5	0.37	0.36	0.41	0.35	0.05	0.40	0.39	0.41	0.40
6	0.44	0.44	0.49	0.41	0.06	0.47	0.49	0.47	0.49
7	0.52	0.54	0.58	0.47	0.07	0.54	0.58	0.54	0.58
8	0.61	0.64	0.69	0.55	0.08	0.63	0.68	0.62	0.67
9	0.72	0.76	0.81	0.64	0.08	0.72	0.78	0.71	0.77
10	0.86	0.89	0.96	0.75	0.09	0.84	0.88	0.82	0.87
11	1.02	1.05	1.13	0.88	0.10	0.98	1.00	0.97	0.98
12	1.22	1.21	1.32	1.04	0.11	1.14	1.11	1.14	1.10
13	1.43	1.38	1.50	1.20	0.11	1.32	1.20	1.32	1.19
14	1.61	1.49	1.62	1.35	0.11	1.46	1.25	1.49	1.24
15	1.69	1.49	1.66	1.44	0.10	1.55	1.22	1.59	1.21
16	1.68	1.37	1.64	1.51	0.08	1.59	1.10	1.68	1.10
17	1.80	1.27	1.75	1.74	0.03	1.77	0.94	1.96	0.97

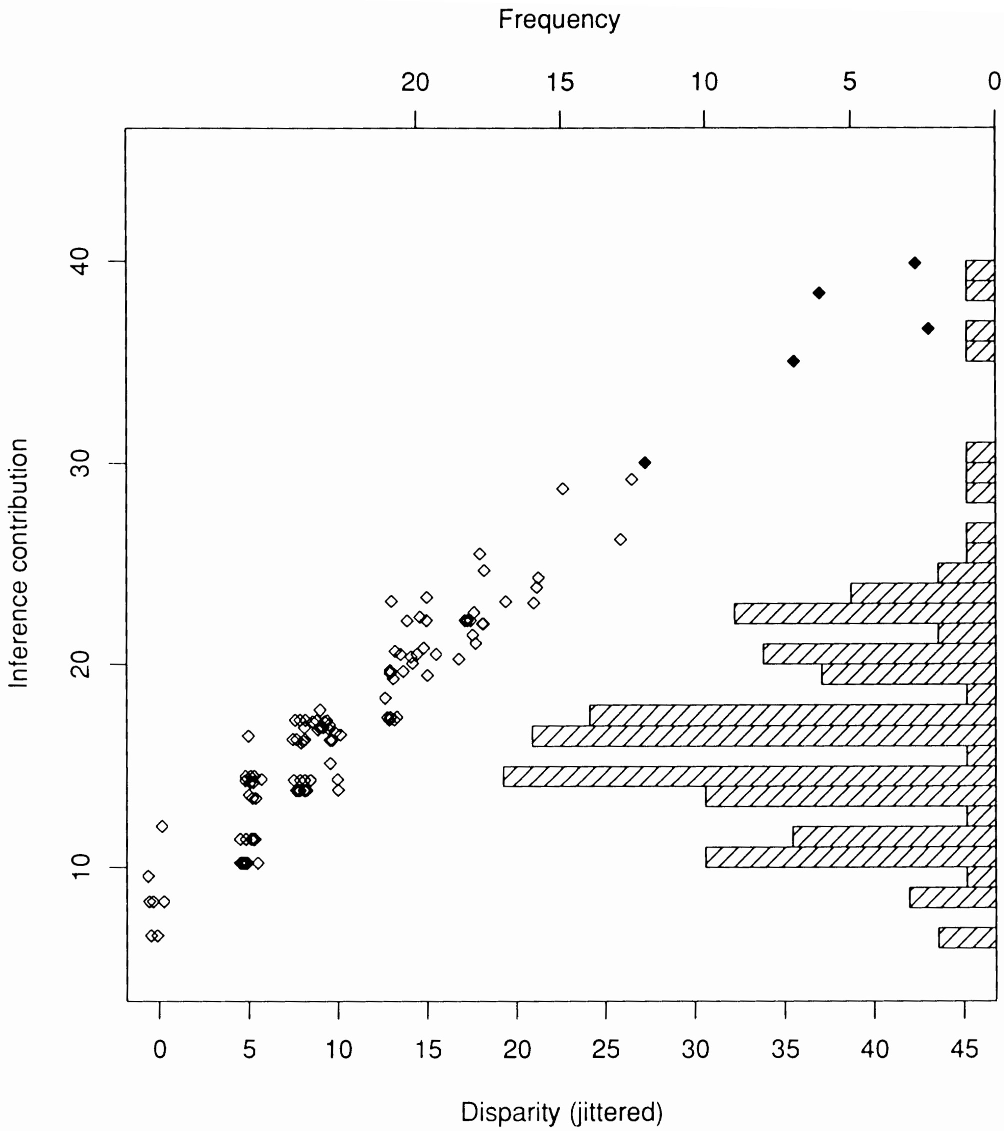

FIG. 4. Relationship between inference contributions ζ_i in (2.5) and the disparities η_i in (2.6), where η_i is calculated from the model in Table 5, case 9. The quantity η_i can take only certain integer values (0, 5, 6, 8, . . .), but has been jittered to better display the data. Solid diamonds represent the five fish identified as outliers in Fig. 1 and 3. Horizontal bars reflect the frequency distribution of ζ_i values.

TABLE 8. For selected ages b from case 9 in Table 5: estimates $\hat{p}_b$, widths $2 \times 1.96 \times \hat{\text{SE}}(\hat{p}_b)$ of 95% confidence intervals obtained from the asymptotic covariance matrix (1.18), 95% confidence intervals from the χ^2-distribution in (1.17), and corresponding midpoints and widths of the latter intervals. For $b = 14$, the left endpoint of the 95% interval is constrained at $p_b = 0$.

b	$\hat{p}_b$	$2 \times 1.96 \times \hat{\text{SE}}(\hat{p}_b)$	95% interval	Midpoint	Width
3	0.008	0.031	(0.000,0.035)	0.018	0.035
4	0.038	0.073	(0.012,0.085)	0.048	0.073
11	0.124	0.137	(0.065,0.200)	0.133	0.136
14	0.010	0.058	(0.000,0.050)	0.025	0.050

assumed that the shape of the distribution $q_j(a|b,\mathbf{\Phi})$ for a given b is similar for each reader, with the distribution mode shifted according to the relative bias $w(b)r_j$. More complex formulations could be implemented. For example, $w(b)$ could depend nonlinearly on b, and the parameters σ_1, σ_2, α, and β could also be indexed by reader.

4. Concluding Remarks

We have presented the principal analytical methods needed to estimate ageing error from a sample with multiple age readings. One possible model extension incorporating reader effects was discussed in section 3. Other modifications may also be required to tailor the general model to a particular data set. We recognize, for example, that our formulations of **Q** may be inappropriate. Bradford (1991) postulates a classification matrix **Q** with substantial under ageing of older fish. This leads to highly skewed probability distributions $q(a|b)$ for a given b. By contrast, the distributions in Fig. 2 are symmetric for true ages near the center of the age range.

We avoided the problem of small sample sizes at old ages by eliminating the oldest fish from our analysis. This may be impractical when old fish are important for biological interpretation. A common solution to this problem is to let age-class A denote all fish of that age and older. However, grouping older ages may necessitate a different representation for the final row

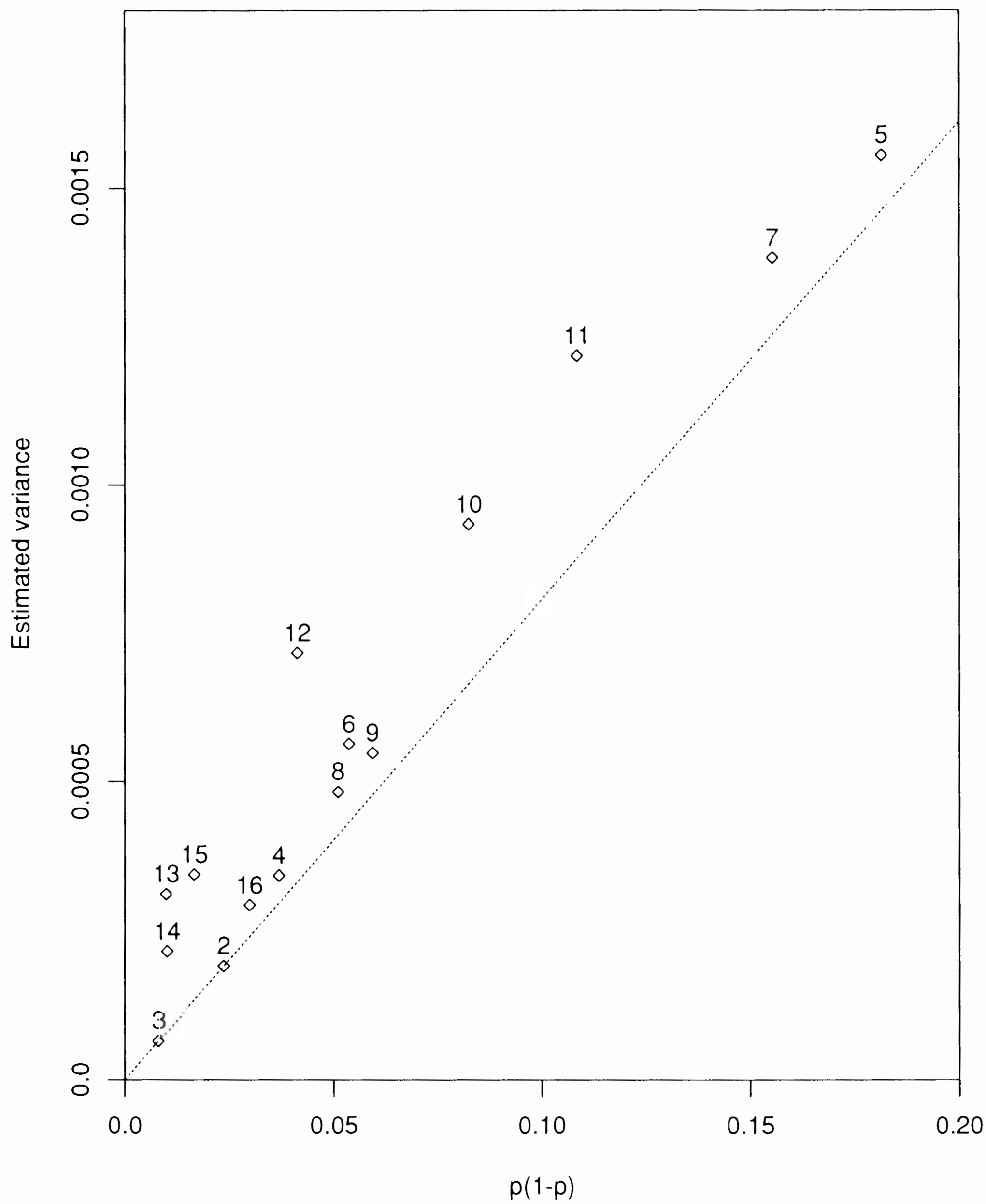

FIG. 5. Relationship between estimated variance $\hat{V}[\hat{p}_b]$ from (1.18) for the specified true age b and the product $\hat{p}_b(1 - \hat{p}_b)$. The dotted line has slope $1/I$, where $I = 124$, indicating the theoretical relationship for a multinomial distribution.

and column of **Q**. Grouping could also bias the estimates $\hat{p}_b$ in the vicinity of $b = A$, if the classification matrix estimated from one population is applied to another population with a different mix of older ages. In general, it is better to estimate $\hat{\mathbf{P}}$ from as many ages as possible. Ages can be grouped subsequently if required for other types of analyses. Models could also be designed using the concepts developed in (2.4), although technical implementation may be difficult.

Under most circumstances, it is not possible (or perhaps even desirable) to obtain precisely J readings of each age structure. However, the models can easily be generalized to account for unequal sample sizes. If J_i represents the number of readings for structure i, then from [4.5] the likelihood function to address question 3 is given by

$$(4.1) \qquad L(\mathbf{A}|\mathbf{P},\mathbf{Q}) = \prod_{i=1}^{I} \sum_{b=1}^{A} p_b \prod_{j=1}^{J_i} q(a_{ij}|b,\mathbf{\Phi}).$$

The likelihood [4.10] is a special case of (4.1), where $J_i = J$ for a subset of $I - K$ structures and $J_i = 1$ for the remaining K structures. Thus, the general likelihood function (4.1) can be used to address both questions 3 and 6.

In some situations, age increments from a specified date are known, but true ages are not. For example, distinguishing marks may naturally occur on ageing structures or may be induced with oxytetracycline injections (Thomas 1983; Beamish and McFarlane 1987; Casselman 1987). If fish are marked on a known date and later recovered, then two sets of readings on each age structure can be used to estimate a classification matrix in the context of question 2. For fish i and reader j, let a_{ij} and a'_{ij} denote estimated ages at recovery and mark formation, respectively. The likelihood for these data is given by

$$(4.2) \qquad L(\mathbf{A}, \mathbf{A}'|\mathbf{B},\mathbf{Q}) = \prod_{i=1}^{I} \prod_{j=1}^{J} q_j(a_{ij}|b_i,\mathbf{\Phi})\, q_j(a'_{ij}|b_i - t_i,\mathbf{\Phi})$$

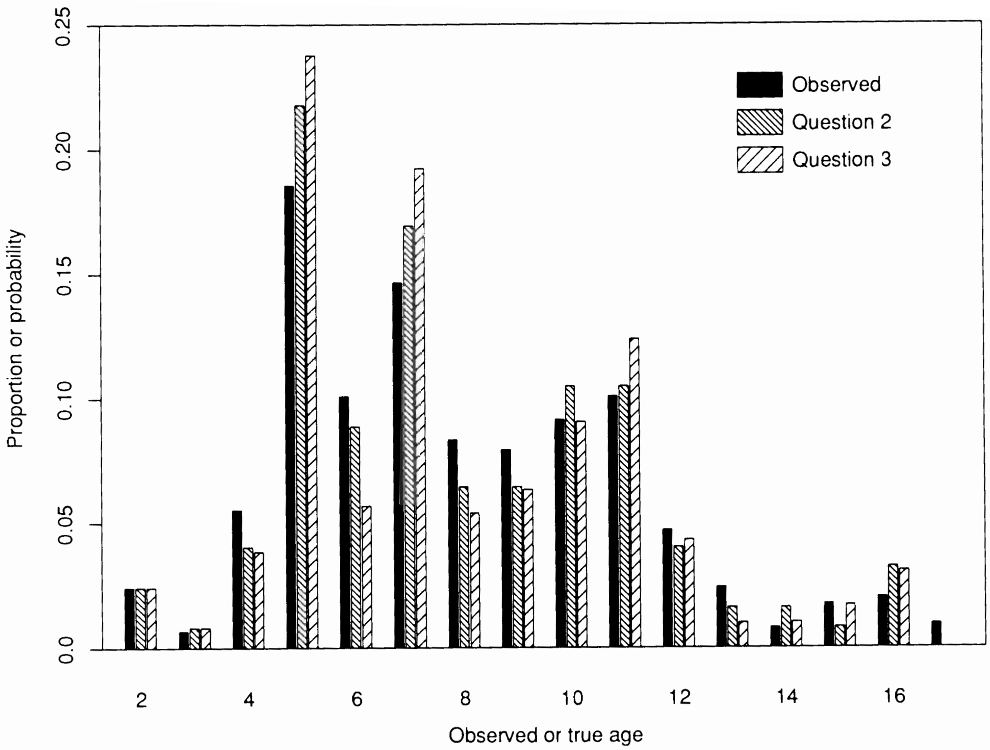

FIG. 6. Histograms relating ages a or b to observed proportions p_a^*, proportions p_b' estimated from $\hat{\mathbf{B}}$ in question 2 (Table 5, case 4), and estimated probabilities $\hat{p}_b$ in question 3 (Table 5, case 9).

TABLE 9. Estimates of reader effects ($r_1, . . ., r_6$) in units of years for cases 11–14 in Table 5.

Case	$\hat{r}_1$	$\hat{r}_2$	$\hat{r}_3$	$\hat{r}_4$	$\hat{r}_5$	$\hat{r}_6$
11	0.120	0.015	−0.459	0.527	0.316	−0.519
12	0.130	0.014	−0.469	0.536	0.318	−0.529
13	0.078	0.006	−0.356	0.424	0.252	−0.405
14	0.082	0.005	−0.357	0.424	0.250	−0.404

where the matrix $\mathbf{A}'$ has elements a'_{ij}, t_i is the known time period between mark formation and recovery of fish i, and $q_j(a|b,\mathbf{\Phi})$ corresponds to a reader-dependent distribution as in (3.6). This representation assumes that readings are independent and that the values t_i are not revealed to readers.

Our models could also be modified to incorporate the case in which true ages are actually known for a sample of fish. Multiple readings from structures of known true age would then constitute a learning sample, and likelihood analysis would reduce to that familiar in mixture models (Fournier et al. 1984; Hoenig and Heisey 1987; Wood et al. 1987). In this case, a modal assumption would not be needed, since an adequate learning sample would provide the information necessary to estimate the entire classification matrix $\mathbf{Q}$. A model for $\mathbf{Q}$, such as (1.6)–(1.8), may still be desirable, however, to reduce the required number of parameters.

Our analysis does not specifically address the issue of experimental design. In section 2, we noted that additional readings of age structures (i.e. larger J) may provide increased precision in $\hat{\mathbf{Q}}$. Because ageing error tends to increase with age, additional readings may be particularly important for older ages. Thus, it may be appropriate to conduct a greater number of readings on age structures taken from fish presumed to be older. Similarly, larger random samples of structures (i.e. larger I) are required to achieve a specified precision in $\hat{\mathbf{P}}$ when ageing error is present, relative to a comparable data set with sampling error but no ageing error. Question 6 suggests one possible experimental strategy, in which multiple readings are performed on age structures from I fish, and age structures from an additional K fish are each examined once. These additional readings give increased precision in $\hat{\mathbf{P}}$. The likelihood (4.1) could be used to analyze all such experimental designs.

Among-reader variability is a well-known problem in age determination. For example, Kimura and Lyons (1991) document how the percentage agreement between two readers of a structure tends to decrease with increasing fish age. Percentage agreement also depends on the species. Averaging across ages, Kimura and Lyons (1991) report an agreement of 64% for walleye pollock, compared with 41, 44, and 79% agreement for Pacific ocean perch (*Sebastes alutus*), sablefish (*Anoplopoma fimbria*), and Pacific hake (*Merluccius productus*), respectively. Thus, for some species and readers, any single reading has a high probability of being wrong.

Ageing error corrections are applied infrequently in fisheries science, despite the known complications of ignoring this error. For example, Hollowed et al. (1987) and Myers and Drinkwater (1989) recognized the potentially confounding effects of ageing error in their attempts to relate year-class strength to environmental factors. Their analyses depended on previously published data, however, and they were unable to determine the influence of ageing error on their conclusions. Clearly, accurate ages are required to estimate year-to-year differences in fish survival and the dependency of survival on environmental factors. Our results (Fig. 6) reinforce the observation by several authors that ageing error acts to smooth differences in year-class strength, thus masking any such relationships.

Any method of age determination must be validated across all age-classes (Beamish and McFarlane 1987) before estimates of age and ageing error can be meaningful. Even with validation, however, there will generally be variability among readings of a given structure. In some cases, this variability is addressed by consultation, discussion, and perhaps gentle coercion in order to resolve disagreements in interpretation. It is then assumed that resolved ages represent true ages. Such an approach does not correct for ageing error; it merely standardizes an interpretation, and it may even lead to a false sense of precision. By contrast, our method uses variability among readings to assess statistical error. We exploit this knowledge to obtain corrected estimates of the true age distribution.

Our analysis provides an explicit means of dealing with the long-standing problem of measurement error in fish age determination. We believe that similar analyses should form a routine component of studies dependent on fish age data. In particular, multiple readings should be obtained for a subset of age structures and used to estimate the classification matrix **Q**. Single readings of additional structures can then help to improve precision in the estimate of **P**. We caution that, as ageing error increases, sample sizes must also be increased to obtain a specified level of precision.

Acknowledgements

We wish to thank Dr. Daniel Kimura and the organizers and participants of the workshop on pollock ageing methodology, held from September 10 to 14, 1990, at the Sea Fisheries Institute, Gydnia, Poland. In particular, Dr. Jerzy Janusz provided one of us (R.J.B.) with the data used here. Thoughtful comments by an anonymous reviewer and Drs. Michael Bradford, John Hoenig, and Jake Rice helped improve the manuscript.

References

AKAIKE, H. 1974. A new look at the statistical identification model. IEEE Trans. Auto. Control 19: 716–723.

BEAMISH, R. J., AND G. A. MCFARLANE. 1983. The forgotten requirement for age validation in fisheries biology. Trans. Am. Fish. Soc. 112: 735–743.

1987. Current trends in age determination methodology, p. 15–42. *In* R. C. Summerfelt and G. E. Hall [ed.] Age and growth of fish. Iowa State University Press, Ames, IA.

BRADFORD, M. 1991. Effects of ageing errors on recruitment time series estimated from sequential population analysis. Can. J. Fish. Aquat. Sci. 48: 555–558.

CASSELMAN, J. M. 1987. Determination of age and growth, p. 209–242. *In* A.H. Weatherley and H.S. Gill [ed.] The biology of fish growth. Academic Press, London, UK.

CHILTON, D. E., AND R. J. BEAMISH. 1982. Age determination methods for fishes studied by the groundfish program at the Pacific Biological Station. Can. Spec. Publ. Fish. Aquat. Sci. 60: 102 p.

FOURNIER, D., AND C. P. ARCHIBALD. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39: 1195–1207.

FOURNIER, D., T. D. BEACHAM, B. E. RIDDELL, AND C. A. BUSACK. 1984. Estimating stock composition in mixed-stock fisheries using morphometric, meristic, and electrophoretic characteristics. Can. J. Fish. Aquat. Sci. 41: 400–408.

HOENIG, J. M., AND D. M. HEISEY. 1987. Use of a log-linear model with the EM algorithm to correct estimates of stock composition and to convert length to age. Trans. Am. Fish. Soc. 116: 232–243.

HOLLOWED, A. B., K. M. BAILEY, AND W. S. WOOSTER. 1987. Patterns of recruitment of marine fishes in the Northeast Pacific Ocean. Biol. Oceanogr. 5: 99–131.

KENDALL, M., AND A. STUART. 1979. The advanced theory of statistics. Vol. 2. 4th ed. MacMillan Publ. Co. Inc., New York, NY. 748 p.

KIMURA, D. K., AND S. CHIKUNI. 1987. Mixtures of empirical distributions: an iterative application of the age–length key. Biometrics 43: 23–35.

KIMURA, D. K., AND J. J. LYONS. 1991. Between-reader bias and variability in the age-determination process. U.S. Fish. Bull. 89: 53–60.

LAI, H. L., AND D. R. GUNDERSON. 1987. Effects of ageing error on estimates of growth, mortality and yield per recruit for walleye pollock (*Theragra chalcogramma*). Fish. Res. 5: 287–302.

MEGREY, B. A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Am. Fish. Soc. Symp. 6: 8–48.

METHOT, R. D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Am. Fish. Soc. Symp. 6: 66–82.

1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data, p. 259–277. *In* L.-L. Low [ed.] Proceedings of the symposium on application of stock assessment techniques to gadids. INPFC Bull. 50.

MITTERTREINER, A., AND J. SCHNUTE. 1985. Simplex: a manual and software package for easy nonlinear estimation and interpretation in fishery research. Can. Tech. Rep. Fish. Aquat. Sci. 1384: 90 p.

MYERS, R. A., AND K. DRINKWATER. 1989. The influence of warm core rings on recruitment of fish in the Northwest Atlantic. J. Mar. Res. 47: 635–656.

NELDER, J. A., AND R. MEAD. 1965. A simplex method for function minimization. Comput. J. 7: 308–313.

PELLA, J. J., AND T. L. ROBERTSON. 1979. Assessment of composition of stock mixtures. U.S. Fish. Bull. 77: 387–398.

PRESS, W. H., B. P. FLANNERY, S. A. TEUKOLSKY, AND V. T. VETTERLING. 1986. Numerical recipes, the art of scientific computing. Cambridge University Press, Cambridge, UK. 818 p.

RIVARD, D. 1989. Overview of the systematic, structural, and sampling errors in cohort analysis. Am. Fish. Soc. Symp. 6: 49–65.

SCHNUTE, J. 1981. A versitile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci. 38: 1128–1140.

THOMAS, R. M. 1983. Back-calculation and time of hyaline ring formation in the otoliths of the pilchard off southwest Africa. S. Afr. J. Mar. Sci. 1: 3–18.

TYLER, A. V., R. J. BEAMISH, AND G. A. MCFARLANE. 1989. Implications of age determination errors to yield estimates, p. 27–35. *In* R. J. Beamish and G. A. McFaralane [ed.] Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models. Can. Spec. Publ. Fish. Aquat. Sci. 108.

WOOD, C. C., S. MCKINNELL, T. J. MULLIGAN, AND D. A. FOURNIER. 1987. Stock identification with the maximum-likelihood mixture model: sensitivity analysis and application to complex problems. Can. J. Fish. Aquat. Sci. 44: 866–881.

Appendix

In this Appendix, we report technical details relevant to the estimates in Tables 5–8. Questions 2 and 3 are addressed by minimizing the inference functions (1.15) and (1.16), respectively. We applied the simplex search method (Nelder and Mead 1965; Press et al. 1986), as implemented by Mittertreiner and Schnute (1985). For question 2, the simplex method was used to obtain M of the parameter estimates, and the remaining I parameters were estimated by direct maximization. For question 3, all $M + A - 1$ parameters were obtained with the simplex method. In spite of the large number of parameters (up to 25), the simplex method performed adequately. For example, estimates for case 3 were achieved in 115 iterations from reasonable initial values; these required 3 min on a 486 33-mHz personal computer. Corresponding estimates for case 8 required approximately 4220 iterations (26 min). Individual iterations were more costly for question 2 due to the maximization (1.20)–(1.21). Total run time was much longer for question 3, however, because of the larger number of parameters involved.

The estimated asymptotic covariance matrix (1.18) was determined from numerical derivatives, where we follow the procedure described by Mittertreiner and Schnute (1985) to verify convergence. For these calculations, it is necessary to use the actual parameters p_b, rather than the surrogate parameters z_b in (1.22). Furthermore, the calculations must be performed twice, choosing a different age b_0 as the calculated parameter p_{b_0} in (1.23) for each run. Thus, $A - 1$ estimates $\hat{\mathrm{SE}}(\hat{p}_b)$ can be obtained from the first run, but a second run is required to obtain the final estimate. Estimates $\hat{\mathrm{SE}}(\hat{p}_b)$ are consistent between runs for the $A - 2$ unconstrained parameters p_b.