Should total landings be used to correct estimated catch in numbers or mean weight-at-age?

Peter Lewy and Hans Lassen


Many ICES fish stock assessment working groups have practised Sum Of Products, SOP, correction. This correction stems from a comparison of total weights of the known landings and the SOP over age of catch in number and mean weight-at-age, which ideally should be identical. In case of SOP discrepancies some countries correct catch in numbers while others correct mean weight-at-age by a common factor, the ratio between landing and SOP.

The paper shows that for three sampling schemes the SOP corrections are statistically incorrect and should not be made since the SOP is an unbiased estimate of the total landings. Calculation of the bias of estimated catch in numbers and mean weight-at-age shows that SOP corrections of either of these estimates may increase the bias.

Furthermore, for five demersal and one pelagic North Sea species it is shown that SOP discrepancies greater than 2% from the landings are very unlikely. Larger discrepancies probably are indications of problems with the sampling design. The proper action is to reexamine the sampling programme and to revise it where needed.

Key words: Sum Of Product (SOP) correction, sampling error, sampling, fish stock assessment.

Received 1 September 1995; accepted 29 July 1996.

P. Lewy and H. Lassen: Danish Institute for Fisheries Research, Department of Marine Fisheries, Charlottenlund Castle, DK-2920 Charlottenlund, Denmark.

Introduction

Fish stock assessments are often based on catches in numbers and mean individual weights by age and year, which allows for calculation of landed weight. The Sum Of Products (SOP) of mean individual weight and catch in numbers over age groups compared to the reported landings is discussed in this paper. It is postulated that correction for a difference between SOP and reported landing is neither necessary nor desirable.

The mean individual weights and the catch in numbers are obtained from sampling of the landings while the total landing may be obtained from a statistical programme or from sampling. The SOP over age-of-catch in number-at-age and mean weight-at-age should ideally match the estimated landing but because of variation and biases in the sampling programmes this is seldom the case. The SOP correction consists of applying a correction factor to either mean weights or numbers caught so that SOP and landings match exactly. The SOP correction is therefore related to a national sampling programme for estimation of catch in numbers and mean weight-at-age of a given fish stock.

It has been a common practice in some ICES fish stock assessment working groups to correct for the SOP differences in order to obtain agreement between the landings and SOP. For example, in ICES (1994b) the SOP corrections are not presented explicitly, but are included in standard computer programs used for calculating total international catch in numbers and mean weight-at-age. This practice has been supported by ICES’s Advisory Committee for Fishery Management (ACFM), as in Anon. (1995) where it is stated that “If for instance the SOP does not match the catch in tonnes data exactly the Working Group should identify and describe the causes and corrections should be made if possible. Whether it is most appropriate to correct catch in tonnes or the catch numbers and weight data should be considered by the working group.”

It is the thesis of this paper that SOP corrections are neither necessary nor desirable. A minor SOP deviation is to be expected and should not give rise to any corrections. Indeed for some sampling schemes a SOP correction introduces a bias in the estimated mean weights or catches. Major SOP discrepancies from the landings indicate incorrect sampling procedures.
and the proper action is a reinvestigation of the entire programme.

The procedures for obtaining catch-at-age sometimes include a correction for ice in the boxes, overweight in boxes, etc. These factors are corrections of the total landings and not SOP corrections as discussed in this paper.

The mean weight-at-age provided by ICES assessment working groups is often a long-term average value applicable for projections. A similar situation occurs if the mean weights are derived from a general length-weight relationship. SOP based on such mean weights-at-age would of course not be expected to match the landing exactly in any particular year. The catch in numbers or the mean weights do not require SOP corrections in these situations.

The present paper studies three often used sampling programmes to obtain mean weight-at-age and catch in numbers. These programmes include random sampling of the catches and two Age–Length Key (ALK) sampling strategies. For each of these sampling programmes the paper considers the possible bias of estimates of SOP, catch in numbers and of mean weight-at-age. The variance of SOP is also evaluated to get an impression of the magnitude of the SOP deviation from the actual landings.

Fish sampling problems in general have been considered in many contexts, such as workshops (Doubleday and Rivard, 1983; ICES, 1994a). Two-phase sampling schemes using the length for stratification have been considered by Tanaka (1953), Anderson (1964) and Lewy (1995). The SOP correction issue has only occasionally been dealt with (Lassen and Lewy, 1981; Lewy, 1982).

Materials and methods

Estimates of SOP, catch in numbers and of mean weight-at-age are investigated for three different sampling schemes. They all are two-phase stratified sampling programmes for a given fish stock, where the primary sample deals with the length distribution of the catch and the secondary sample is an Age–Length Key (ALK) sample. In the primary sample the length of each fish is measured and the sample divided into length groups. For each length group the mean weight per fish and the corresponding standard deviation are determined. The ALK is a random sample of fish drawn from each length group. Each of these fish are aged and weighed thus providing the age- and weight-distribution. The notation used is given in Appendix I. The estimators used for catch in numbers for length \( l \) and age \( a \), \( N_{la} \), catch in numbers for age \( a \), \( N_a \), and mean weight-at-age, \( T_a \), respectively are:

\[
\hat{N}_{la} = \frac{L}{\Omega} \frac{\sum m_{la}}{n} \frac{m_i}{m_i}
\]

\[
T_a = \frac{1}{\hat{N}_a} \sum \hat{N}_{la} \hat{\phi}_{la}
\]

where \( L \) denotes total weight landed, \( \Omega \) denotes the mean weight of the fish in the primary sample and \( m_{la} \) denotes the mean weight by length and age in the secondary sample.

Since \( \hat{\phi}_l = \frac{a}{m_i} \) and \( \Omega = \frac{1}{n} \sum n_i \Omega_l \), SOP becomes

\[
\text{SOP} = \sum_a \hat{N}_a T_a = \frac{1}{\Omega} \sum_l n_i \hat{\phi}_l
\]

Equation (4) shows that possible SOP deviations from the landings \( L \) are due to differences in mean weight by length group in the primary length and the secondary ALK samples.

All three sampling schemes assumes that a primary random sample of length frequency is drawn. The three sampling schemes considered are:

(i) Constant ALK sampling. The ALK samples \( m_{la} \) are random subsamples obtained from random length samples \( n_i \) where \( m_{la} \) denotes the number of fish of length “l” and age “a” and \( n_i \) denotes the number of fish in length group “l” in the primary sample. Furthermore, the number of samples \( m_i \) is the same for each length group “l”, that is \( m_i = k \), where \( k \) is a constant.

(ii) Proportional sampling. As in case (i), \( m_{la} \) are random subsamples from the primary samples, but in this case \( m_i/n_i = m/n \) for all length groups “l”.

(iii) Independent sampling. The random length samples, \( n_i \), and the random age/length samples, \( m_{la} \), are independent samples collected from separate sampling programmes.

Bias of estimated SOP, \( N_a \), and \( W_a \) and coefficient of variation of SOP

In order to evaluate the effect of a possible SOP correction the possible bias of the estimates (2), (3) and (4) has been estimated for the three sampling schemes considered. Coefficient of variation of estimated SOP has also been calculated. Only the bias of the order \( n^{-1} \) has been considered in the calculations. The assumptions made are given in Appendix II.

Formulae for the relative bias of estimated catch in number, mean weight-at-age and SOP and coefficient of variation of SOP have been derived as outlined in Appendix III and are summarized in Table 1.
The bias of estimated catch in number-at-age, $N_a$, has been evaluated using weight-at-age and age composition data for annual landings of six North Sea demersal species (ICES, 1994b) and North Sea herring (ICES, 1994c). The bias of the estimated mean weight-at-age, $T_a\bar{W}$, and the coefficient of variation of the estimated SOP have been evaluated using weight-at-length and length distribution data from the Danish commercial sampling programme in 1993 for six North Sea demersal and one pelagic species. The six demersal species are the four roundfish species, cod (Gadus morhua), haddock (Melanogrammus aeglefinus) and saithe (Pollachius virens) and the two flatfish species, plaice (Pleuronectes platessa) and sole (Solea solea). The pelagic species is herring (Clupea harengus).

The primary length samples of roundfish and herring were grouped into length intervals of 1 cm while 0.5 cm was used for the flatfish. The three sampling schemes may be combined with further stratification, for instance by port or commercial size category. In such cases the bias needs to be calculated for each strata and then combined to provide the required total.

### Results

Estimated SOP is unbiased or approximately unbiased (sampling scheme 3), see Table 1. Equation (4) shows that SOP depends only on the mean weight-by-length group in the primary length samples, $\bar{W}$, and the secondary ALK samples, $\Omega_a$. Since these mean weights can both be calculated without knowing the age composition it follows that SOP is independent of the age composition. In particular, possible misreadings of age do not influence the SOP. Next, the magnitude of the bias of catch-at-age and mean weight-at-age and the variance of the SOP is evaluated.

### Bias of catch-at-age

The bias of estimated catch in numbers at age depends on the number of length samples and the relation between mean weight-at-age and mean weight of all fish in the sample for all three sampling schemes. The relative bias multiplied by the number of length samples is calculated for the six species considered and shown in Table 2. The relative bias (multiplied by $n$) is negative for the small fish and positive for large fish (Table 2). For age groups 0–4 of the roundfish and for all age groups of sole, plaice and herring the bias is less than 1% when the number of fish measured is greater than 200. For cod and haddock older than 4 years, measurement of about 1000 fish is required for obtaining a bias less than 1%.

### Bias of estimated mean weight-at-age

Four general characteristics can be deduced from Table 1: (i) the bias approaches zero when the number of fish aged approaches the number of fish measured, i.e. when $m$ approaches $n$; (ii) the bias is inversely related to the catch in number at age; (iii) the bias is negative, and, (iv) for small $m$ compared to $n$ the relative bias is proportional to $-N_a/N_a$ and $1/m$.

Let us consider situations where large biases may occur: as the bias increases when $m/n$ decreases we only need to evaluate situations where $m/n$ is small. In this case we have:

### Table 1. Approximate relative bias of estimated SOP, catch in number at age and mean weight and coefficient of variation of SOP.

<table>
<thead>
<tr>
<th>Relative bias of estimates and CV of SOP</th>
<th>(a) Constant</th>
<th>(b) Proportional</th>
<th>(c) Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(SOP) - L$</td>
<td>$0$ (Exact)</td>
<td>$0$ (Exact)</td>
<td>$0$</td>
</tr>
<tr>
<td>$E(N_a) - N_a$</td>
<td>$1/n (1-W_a/W)$</td>
<td>$1/n (1-W_a/W)$</td>
<td>$1/n (1-W_a/W)$</td>
</tr>
<tr>
<td>$E(T_a\bar{W}_a)$</td>
<td>$N/\bar{W}_a (C_m D_n/m n)$</td>
<td>$N/\bar{W}_a (1/m 1/n)$</td>
<td>$N/\bar{W}_a (1/m 1/n)$</td>
</tr>
<tr>
<td>$CV^2 (SOP)$</td>
<td>$\mu^2 (A - B)/m n$</td>
<td>$(1 - 1/m) \mu^2 B$</td>
<td>$1/m \mu^2 B$</td>
</tr>
</tbody>
</table>
The constants $D_a$ and $C_a$ given in Appendix I have been evaluated using data from the Danish North Sea sampling programme 1993. It is found that $0.02 < C_a < 2$ while $0.2 < D_a < 0.98$ ($D_a$ is by definition less than or equal to 1). Normally $C_a$ is about 1. No systematic relationship between the constants and the age was found.

The magnitude of the bias has been illustrated using age composition data for North Sea cod 1994 (ICES, 1994b) and by assuming that $m=1000$ and $C_a$ and $D_a$ equal to 1. The results are shown in Table 3.

Table 3 indicates that the bias of estimates of mean weight-at-age may be considerable for age groups seldom caught. In the case of North Sea cod the relative bias is larger than 35% for fish of 7 years old and older. If, for example, the relative bias is required numerically less than 5% and if $N/N_a=100$ then in some cases $m$ must be greater than 4000 in programme (a) and 2000 in programmes (b) and (c). As this may not be possible, instead the number of fish aged, $m$, by length group should be as close as possible to the number of fish sampled, $n_a$, for the length groups containing these rare fish.

### Variation of SOP

Approximations of coefficient of variation of estimated SOP are given in Table 1. These approximations depend on the coefficients $A$, $B$, and $\mu$ given in Appendix I. The three coefficients are calculated for the six species considered by size category. The results are given in Table 4. It shows that the estimate of variation of the weight of individual fish, $\mu$, is about 0.1 for all species and size categories. It can be shown that the same applies to the unstratified case where the fish are not divided into size categories. Correspondingly, Table 4 shows that the upper limit of the coefficient $A$ is about 2 while $B$ lies between 1 and 1.6 and therefore $0.14/\sqrt{m}$ is an applicable upper limit for the coefficient of variation of estimated SOP for all three sampling schemes.

To illustrate the magnitude of possible deviations between estimated SOP and the landings, the probability that the relative deviation exceeds $h\%$ has been calculated using the common upper limit of $0.14/\sqrt{m}$ and by assuming that the estimated SOP is normally distributed. The upper limit of the probability given by the last term in
Equation (5) is shown in Figure 1 for $100 \leq m \leq 1000$, and for $h = 0.01, 0.02$ and $0.03$. Figure 1 shows that when the annual number of fish aged $m$, is larger than 300, SOP deviations larger than 2% have probabilities less than 1.4%, while deviations larger than 1% have probabilities less than 22%. In general SOP deviations are likely to be less than 2% even for small values of $m$.

Discussion

Possible discrepancies of SOP from the landing are due to differences between mean weight-at-length in the length sample and mean weight-at-length in the ALK sample. The results show that if the fish sampled in the ALK programme are random samples then large SOP deviations from the landing are very unlikely even for small numbers of fish sampled in the ALK. If the relative SOP deviation is larger than twice the coefficient of variation of estimated SOP, it may be caused by an incorrect sampling programme or errors in the calculations. These problems need to be thoroughly examined so that the errors can be identified and removed.

Problems with the sampling programme in case of the sampling schemes (a) and (b) may consist in ALK subsamples by length group which are not taken randomly among the length samples such that systematic sampling errors occur. Similarly this will happen for sampling scheme (c), where the length samples and the ALK samples are taken independently of each other. Most likely there are systematic errors in the selection procedure of samples in at least one of these two sampling schemes.

SOP corrections of catch in numbers or mean weight-at-age have probably been made in order to obtain catch predictions and estimates of stock size and fishing mortality which are not affected by the SOP deviation from the landings. It was concluded above that no SOP correction should be made. An obvious question is therefore how the SOP deviations should be considered. The assessment methods used up to now, conventional Virtual Population Analysis (Gulland, 1965), various tuning methods (Darby and Flatman, 1994) and catch predictions, are deterministic approaches where catches and mean weight at age are assumed to be known without errors. Using these methods the SOP deviation from landings may be considered as a coarse estimate of the sampling error generated on predicted catches. A better approach of including sampling errors into assessments is to apply the so-called integrated models (e.g. Doubleday, 1976; Fournier and Archibald, 1982; Deriso et al., 1985; Lewy, 1988; Sparholt, 1990), that is standard stochastic models incorporating expected value and variance of the stochastic variables. This enables calculation of the precision of estimated stock size and fishing mortality and thereby (for given level of predicted fishing mortality) the precision of the predicted catches.

Conclusions

1. No SOP correction should be made as a minor SOP deviation from total weight landed is to be expected in most cases. SOP corrections of catch in numbers and mean weight-at-age may lead to estimates with stronger bias than the original estimates. The reasons for this are: (i) If all fish sampled are aged, SOP is identical and equal to the landings and no SOP correction is possible (or desired); and, (ii) SOP is an unbiased estimate of the landing for the three sampling schemes considered.  

2. Estimated mean weight-at-age is negatively biased. For small numbers of fish aged $m$, compared to the numbers of fish sampled, the bias is proportional to

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Species & Size category & \(\hat{\mu}\) & A & B \\
\hline
Cod & 1 & 0.09 & 1.5 & 1.0 \\
& 2 & 0.07 & 1.4 & 1.0 \\
& 3 & 0.10 & 1.6 & 1.0 \\
& 4 & 0.08 & 1.5 & 1.0 \\
& 5 & 0.09 & 1.6 & 1.0 \\
Haddock & 1 & 0.09 & 1.6 & 1.6 \\
& 2 & 0.09 & 1.6 & 1.0 \\
& 3 & 0.10 & 2.0 & 1.1 \\
Saithe & 1 & 0.10 & 1.8 & 1.5 \\
& 2 & 0.13 & 2.0 & 1.1 \\
& 3 & 0.10 & 1.8 & 1.1 \\
Plaice & 1 & 0.10 & 1.9 & 1.2 \\
& 2 & 0.10 & 2.0 & 1.0 \\
Sole & 1 & 0.10 & 2.1 & 1.1 \\
& 2 & 0.10 & 2.1 & 1.1 \\
Herring & 1 & 0.09 & 1.6 & 1.1 \\
\hline
\end{tabular}
\caption{Estimated annual values of \(\hat{\mu}\), A and B by species and commercial size category.}
\end{table}

\[\hat{\mu} = \frac{1}{\sum(n_i - 1) \sum(n_i - 1) s_i \Omega_i}\]

\[A = Q \sum P_i \left(\frac{W_i}{W}\right)^2\]

\[B = \sum P_i \left(\frac{W_i}{W}\right)^2\]

Source: Danish commercial sampling programme for the North Sea 1993.
N/Na and 1/m. In cases of large values of N/Na (which is the case for old fish seldom caught), the bias may be considerable unless one ensures that the number of fish aged increases towards the number of fish sampled for the critical length groups. The bias may be removed using the formulas given in Table 1. SOP corrections using a factor less than one will increase the negative bias.

3. Estimated catch in numbers are biased by the order n⁻¹ (n is the number of fish sampled). Small fish are positively and large fish negatively biased. For the six species considered the bias is less than 1% when the number of fish sampled is greater than 1000. The bias may be removed using the relevant formula in Table 1. No common factor can correct for the (small) age dependent biasness.

4. Even though SOP is an unbiased estimate of the landings, SOP will vary randomly around the landing (except when all the fish sampled are aged). However, the deviation due to stochastic errors is small. Coefficient of variation of SOP is proportional to √1/m – 1/n, √1/m – 1/n, and √1/m, respectively, for the sampling schemes (a), (b) and (c) (c is a constant greater than one). For the species examined the coefficient of variation of SOP is less than 0.14/√m. In general this implies that SOP deviations due to stochastic errors will be less than 2% even for small numbers of fish aged. If SOP deviates more than 2% from the landing it is probably due to systematic sampling errors, e.g. that the length and age/length sampling programmes represent different fish populations. Such errors need to be examined and removed. On the other hand, a small SOP deviation does not guarantee the quality of the sampling programme.

It should be noted that ageing problems do not affect the SOP deviation.

References


Anon. 1995. Guidance to assessment working groups from ACFM and the secretariat. ICES.


Appendix I

Notation

L denotes total weight of the catch
N denotes total number of fish catched
Nf denotes the number of fish caught of length "l" and age "a"
 Nf denotes the number of fish caught of length "l" and age "a".
 n denotes total number of fish sampled in the length sample
 n denotes the number of fish in the length sample of length "l"
 n denotes the number of fish in the length sample of length "l" and age "a"
 m denotes the number of fish aged, m=∑m=∑m
 m denotes the number of fish aged of length "l".
 W, Wp, Wf and Wf denote mean weights per fish in the catch
 Ω, Ω, Ω and Ω denote mean weights in the length sample
 Ω, Ω, Ω and Ω denote mean weights in the ALK sample
 SOP, N, Nf, Nf and T denote the estimators of L, N, Nf, Nf, and Wf, respectively
 Q denotes the number of length groups

\[ \sigma_i = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (W_i - \bar{W})^2} \]

\[ s_i = \sqrt{\frac{1}{n_i} \sum_{i=1}^{n_i} (\Omega_i - \bar{\Omega})^2} \]

\[ \bar{\mu} = \frac{1}{(n_i - 1)} \sum_{i=1}^{n_i} \frac{s_i}{\Omega_i} \]

\[ P_1 = \frac{N_i}{N} \]

\[ A = Q \sum P_i \left( \frac{W_i}{W} \right)^2 \]

\[ B = \sum P_i \left( \frac{W_i}{W} \right)^2 \]

\[ C_s = Q \sum \frac{W_{ia}}{W_s} \frac{N_{ia}}{N} \left( \frac{N_i - N_{ia}}{N} \right) \]

\[ D_s = 1 - \sum \frac{W_{ia}}{W_s} \frac{N_{ia}}{N_{ia}} \frac{N_{ia}}{N_{ia}} \]

VAR denotes the variance
COV denotes the covariance
cV denotes coefficient of variation.

Appendix II

Assumptions

In order to estimate the potential bias of the estimates (1), (2) and (4) the following sample characteristics are used:

\( (m_{it}, m_{ia}, m_{ia}, m_{ia}) \) are multi-hypergeometrically distributed in case (a) and (b) and multinomially distributed in case (c).

\( (n_i, n_i, n_i, n_i, n_i, n_i) \) and \( (n_{ia}, n_{ia}, n_{ia}, n_{ia}) \) are multinomially distributed.

\( n_i, \Omega, \Omega_s, \Omega_s, \Omega_s \) and \( n_{ia} \) are independent stochastic variables.

With respect to the weight we have \( E_{\Omega_s}(\Omega_s) = W_{ia} \) and \( E_{\Omega_s}(\Omega_s) = W_i \) where \( W \) denotes the true mean weight.
It is known from the sampling theory (Cochran, 1977) that
\[ \text{VAR}(\tilde{\Omega}_i) = \frac{\sigma_i^2}{n_i} \left( 1 - \frac{n_i - 1}{N_i - 1} \right) \sim \frac{\sigma_i^2}{n_i} \]
for \( n_i \ll N_i \).

where
\[ \sigma_i^2 = \frac{2}{N_i} \sum_{i=1}^{N_i} (W_{ii} - \tilde{W}_i)^2 \]

In cases (a) and (b) we have that
\[ E_w(\tilde{\omega}_l) = \sqrt{\tilde{\omega}_l} \text{ and } E_w(\tilde{\omega}_l) = \sqrt{l} \text{ and } \]

where \( s_i^2 = \frac{1}{n_i} \sum_{i=1}^{n_i} (\Omega_{ii} - \tilde{\Omega}_i)^2 \). We also know that \( s_i \) is approximately an unbiased estimate of \( \sigma_i \).

In case (c) we have
\[ E(\tilde{\omega}_l) = \sqrt{\tilde{\omega}_l} \text{ and } E(\tilde{\omega}_l) = \sqrt{l} \text{ and } \]

Furthermore, the following approximations have been used in the calculations:
\[ \frac{1}{Y} \sim \frac{1}{E(Y)} \left( 2 - \frac{Y}{E(Y)} \right) \]

and
\[ E \left( \frac{X}{Y} \right) \sim \frac{1}{E(Y)} \left( E(X) - \text{COV}(X,Y) \right) \frac{E(Y)}{E(Y)} \]

where \( X \) and \( Y \) are stochastic variables.

The approximations may be used if \( 0 < Y < 2E(Y) \). In the paper the latter condition will be considered fulfilled if the probability of this event is greater than 95%.

### Appendix III

#### Derivations of formulas in Table 1

**Expected value of estimated SOP**

In the case of the sampling schemes (a) and (b) we have for a given primary sample that \( E(m_i) = \Omega_i \). Using Equation (4) one obtains \( E(SOP) = L \), showing that the estimated SOP is an unbiased estimator of \( L \). If all fish sampled are aged \( (n = m) \) the estimated SOP is identical to the landings, \( L \). In case of sampling scheme (c) one can show that the estimated SOP approximately is an unbiased estimator of \( L \). This approximation holds if

\[ n > 0.1 \sqrt{\frac{2.7}{n} \sum_{i=1}^{n} \left( \frac{W_i}{W} \right)^2} \]

For the six species considered one can show that this means that at least one fish from each length group should be sampled so that the number of fish measured should be equal to or greater than the number of length groups. As this always is fulfilled the approximation used is valid.

**Expected value of estimated catch in number-at-age**

In the case of the sampling schemes (a) and (b) we obtain

\[ E(N_{ia}) = \frac{L}{n} E \left( \frac{n_{ia}}{\Omega_i} \right) \]

\[ \sim \frac{L}{n \tilde{W}} \left( E(n_{ia}) - \frac{\text{COV}(n_{ia}, \sum_k n_{kb} W_{kb})}{n \tilde{W}} \right) \]

\[ = N_{ia} \left( 1 + \frac{1}{n} \left( 1 - \frac{\tilde{W}_i}{\tilde{W}} \right) \right) \]

In case of scheme (c) one can obtain the same formula when assuming that the mean weight-at-length is independent of the age. Equation (1) in this Appendix leads to the bias shown in Table 1. The approximations used requires that \( n \) is equal to or greater than the number of length groups, which is always the case. The approximation is therefore applicable.

**Expected value of mean weight-at-age**

Using the definition of \( T_a \) (Equation (3) in text) the expected value is

\[ E(T_a) = \sum_i W_{ia} E \left( \frac{\tilde{N}_{ia}}{N_a} \right) \]

(2)

We therefore have to find the expected value, \( E(N_{ia}/N_a) \). For the sampling schemes (a) and (b) one finds that

\[ E(T_a) \sim W_a \left( 1 - \frac{N_a C_a - D_a}{N_a m - n} \right) \]
which leads to the relative bias shown in Table 1. The coefficients $C_a$ and $D_a$ are given in Appendix I. For the sampling schemes (b) and (c) the formulas the relative bias given in Table 1 can be derived similarly. In all three cases the approximations are valid for the six species considered if

$$\frac{n_a m}{n} > 3.6$$

In practice this is equivalent to saying that the number of fish aged should be larger than four for all age groups considered, which is fulfilled for reasonable sampling schemes.

Coefficient of variance of estimated SOP

We have calculated previously the expected value of the estimated SOP. In order to find the coefficient of variance we therefore need to find $E(SOP^2)$. This requires that the variance of the sample mean weights are known.

As $SOP = \frac{\sum n_i \hat{w}_i}{\sum n_i \hat{\Omega}_i}$ we therefore have to find

$$E(SOP^2) = E\left(\left(\frac{\sum n_i \hat{w}_i}{\sum n_i \hat{\Omega}_i}\right)^2\right).$$

It can be shown that in case of sampling scheme (a) we get

$$E\left(\frac{SOP^2}{L^2}\right) \approx 1 + \frac{Q}{m} \sum \frac{P_i^2 \sigma_i^2}{W^2} - \frac{1}{n} \sum \frac{P_i \sigma_i^2}{W^2}$$

(3)

where $P_i = \frac{N_i}{N}$. The coefficient of variation of the weight of the individual fish, $\sigma_i/W_i$, has been tested to equal the common coefficient of variation, $\mu$, for all length groups and species considered. Equation (3) of this Appendix can therefore be written as

$$E\left(\frac{SOP^2}{L^2}\right) \sim 1 + \mu^2 \left(\frac{A}{m} - \frac{B}{n}\right)$$

where $\mu$, $A$, and $B$ are given in Appendix I. The expected value, $E(SOP^2/L^2)$, has similarly been calculated for the sampling schemes (b) and (c). The resulting coefficients of variation of estimated SOP are given in Table 4.