ICES Journal of Marine Science, 54: 399–411. 1997

Instantaneous separable VPA (ISVPA)

Z. I. Kizner* and D. A. Vasilyev

Kizner, Z. I., and Vasilyev, D. A. 1997. Instantaneous separable VPA (ISVPA). – ICES Journal of Marine Science, 54: 399–411.

The approach suggested is in the class of cohort methods; it is a new technique for processing catch-at-age data on species having short (within a year) periods of fishery. The method can also be regarded as an approximation to more general conditions when fishery varies continuously during the year. In many cases it enables a more complete extraction of information on exploited populations and fishery from the catch-at-age matrices, including the natural mortality coefficient and terminal fishing parameters, without using any auxiliary data (survey data, fishing effort series etc.). A number of numerical experiments using simulated data illustrate the methodology and demonstrate the merits of the suggested method.

© 1997 International Council for the Exploration of the Sea

Key words: VPA, cohort, instantaneous fishing, separability, convergence, minimization.

Received 1 June 1995; accepted 24 July 1996.

Z. I. Kizner: Department of Life Sciences, Bar-Ilan University, Ramat-Gan 52900, Israel. D. A. Vasilyev: Russian Federal Research Institute of Fisheries and Oceanography (VNIRO), 17 V. Krasnoselskaya St., Moscow 107140, Russia.

Introduction

One of the popular approximations of the Virtual Population Analysis (VPA), the so-called cohort analysis by Pope (1972), is based on two simple formulae:

$$C_{i,j}=\frac{F_{i,j}}{F_{i,j}+M}N_{i,j}[1-e^{-(F_{i,j}+M)}] \qquad (1)$$

$(i=1, \ldots, n;\ j=1, \ldots, m)$ and

$$N_{i,j}=(N_{i+1,j+1}e^{M/2}+C_{i,j})e^{M/2} \qquad (2)$$

$(i=1, \ldots, n-1,\ j=1, \ldots, m-1)$, where

- i – year index,
- n – total number of years,
- j – index of an age group ($j=1$ corresponds to the first age group present in the catch data),
- m – total number of age groups in the catches,
- $N_{i,j}$ – abundance (number of individuals) of the j-th age group at start of the i-th year,
- $C_{i,j}$ – catch from the j-th age group in the i-th year,
- $F_{i,j}$ – fishing mortality coefficient for the j-th age group in the i-th year,
- M – natural mortality coefficient.

Equation (1) expresses the total catch from the j-th age group, accumulated in the i-th year, if the dynamics of the group abundance N and the accumulated catch C (at time t) during the year are governed by the well known equations: $dN/dt=-(F+M)N$ and $dC/dt=FN$, where F and M do not depend on t (indices are omitted). Equation (2) is traditionally regarded as a discrete approximation of a continuous process; it becomes an exact one if the catch, $C_{i,j}$ is taken instantaneously in the middle of the i-th year.

However, there are many exploited stocks with such short periods of fishing that the latter may be regarded as momentary. In such a case if the period of fishing falls in the middle of a year, Equation (1) should be replaced by

$$C_{i,j}=\varphi_{i,j}N_{i,j}e^{-M/2}, \qquad (3)$$

where $\varphi_{i,j}$ plays the role similar to that of $F_{i,j}$ in Equation (1) but cannot be called a fishing mortality coefficient. Strictly speaking, it is the fraction of the abundance of the j-th age group, taken as catch in the middle of the i-th year. Later on we will study this model, which may be called Instantaneous Separable VPA, or ISVPA for short. The word ‘‘Instantaneous’’ means that the catch is assumed to be taken instantaneously once a year, while the word ‘‘Separable’’ shows that we accept the hypothesis of separability (i.e. of age selectivity of the fishery), similar to that by Pope and Shepherd (1982). The acronym ISVPA should not be confused

*Corresponding author.

1054–3139/97/040399+13 $25.00/0/jm960188 © 1997 International Council for the Exploration of the Sea

with that of Integrated Stochastic VPA by Lewy (1988).

It is clear that in reality the fishing season does not necessarily fall within the middle of the calendar year. For the model it means that instead of factors $e^{M/2}$ and $e^{-M/2}$ the Equations (2) and (3) must contain factors e^{TM}, $e^{(1-T)M}$, and e^{-TM}, where T is a given constant, $0<T<1$ (or even e^{T_iM}, $e^{(1-T_i)M}$, e^{-T_iM}). However, when T does not depend on i, these corrections do not make the model more general: the use of the original Equations (2) and (3) in such a case simply means that the time to which $N_{i,j}$ corresponds is shifted by $T-1/2$ with respect to start of the i-th year. That is why below we will assume $T=1/2$.

We would like to emphasize that our ISVPA, can be regarded as being an approximate method for assessment of continuously exploited age-structured populations. It should be noted that the assumption of a constant fishing mortality coefficient during a year, that underlies conventional VPA, is also only an approximation. These two hypotheses are in fact two opposite limit cases in the framework of cohort methods.

One should not be suprised by the propinquity of some of the ideas of our method to those of Separable VPA by Pope and Shepherd (1982): the hypothesis of separability causes it. But the simplicity of our approach and the different mathematical structure of the equations, allows the estimation of additional terms (e.g. the terminal fraction $\varphi_{n,m}$ and sometimes the natural mortality coefficient), removing the need for ''tuning'' (calibration) using auxiliary information (such as survey or fishing effort data). Testing the suggested approach demonstrates its applicability to different types of catch-at-age data.

Formulation of the problem

The hypothesis of separability means in this context that

$$\varphi_{i,j}=f_i \cdot s_j, \tag{4}$$

where f_i is proportional to the fishing effort (a year effect), while s_j is the selectivity of the fishery (an age effect). These two variables are analogues of fishing and selective patterns in conventional separable VPA. We will call f_i simply an effort, and suppose the selectivity to be normalized:

$$\sum_{j=1}^{m} s_j=1. \tag{5}$$

The catch-at-age matrix $\|C_{i,j}\|$ is given, while the unknown vectors $\{s_j\}$, $\{f_i\}$ as well as the matrix $\|N_{i,j}\|$ are to be found from Equations (2)–(5). As for the coefficient M, two variants are considered: M is a given function of age, or M is unknown.

Substituting Equation (4) into Equation (3) we obtain $2mn-(m+n)+2$ equations for $mn+m+n$ unknown values, if M is given. If M is to be estimated along with $\{s_j\}$, $\{f_i\}$ and $\|N_{i,j}\|$, we have $mn+m+n+1$ or $mn+2m+n$ unknowns (for constant M and for that dependent on age correspondingly). When m and n are sufficiently large, the number of equations exceeds the number of unknowns. This means that in general, such a system of equations has no exact solution; one can only speak of a ''solution'' that minimizes a certain loss function. As by Equations (3) and (4) we essentially attempt to present the catches in separable form, it would be reasonable to seek a ''solution'' that satisfies Equations (2) and (5) exactly and secures the best fit of the estimated catches, $f_is_jN_{i,j}e^{-M/2}$, to the actual data, $C_{i,j}$, according to the least squares principle (see section on Iterative procedure).

The number of unknowns is usually large while their orders of magnitude may be uncertain even if one deals with the logarithms of true unknowns. This makes the use of a standard nonlinear minimization method difficult. A special procedure for determining all the unknowns as functions of M and f_n enables us to reduce the problem to minimizing a function of two variables (see section on Method for solving the problem).

When Equation (2) is regarded as containing the ''true'' catches, $C_{i,j}$, the model defined by Equations (2)–(5) may be called catch controlled. For sufficiently smooth input data this approach is adequate. Highly variable data (or data with strong random errors) may require a certain modification of the method to help overcome such a variability. For this purpose an alternative, effort-controlled version of the model can be used. It is obtained by substitution of estimated catches, $f_is_jN_{i,j}e^{-M/2}$, into Equation (2) for the ''true'' ones. In other words, at the stage of evaluation of the sizes of the age groups, $N_{i,j}$, the following equation resulting from (2)–(4), should be used instead of (2):

$$N_{i,j}=\frac{N_{i+1,j+1}e^{M}}{1-f_is_j} \tag{6}$$

$(i=1, \ldots, n-1;\ j=1, \ldots, m-1)$.

Method for solving the problem

Main relationships

The general idea behind our approach is a conventional one for cohort methods: as soon as the coefficient of natural mortality, M, and the terminal elements, $\{N_{n,j}\}$ and $\{N_{i,m}\}$, of the matrix $\|N_{i,j}\|$ are known, all other elements, in the catch-controlled version of the model, can be successively determined from Equation (2). In such a case the terminal values can only be found from Equations (3)–(5) (if we cannot or do not want to use any auxiliary information). In the effort-controlled version, the efforts and selectivities are needed for this purpose in addition to M, and Equation (6) must be used instead of Equation (2).

Suppose there exists an exact solution of the problem, i.e. that the expansion of the fraction $\varphi_{i,j}$ into product of the effort and selectivity (4) is an exact one. Let us sum up both parts of the Equation (4). Taking into account the normalization relationship (5) one obtains:

$$f_i = \sum_{j=1}^{m} \varphi_{i,j} \tag{7}$$

and

$$s_j = \frac{\sum_{i=1}^{n} \varphi_{i,j}}{\sum_{i=1}^{n} f_i} \tag{8}$$

These will be the main formulae of the iterative procedure for determining the series $\{f_i\}$ and $\{s_j\}$.

Iterative procedure

Within the following procedure M is regarded as being a given constant or a function of age.

The calculations start with setting the initial distributions of the effort and selectivity, $\{f_i\}$ and $\{s_j\}$; the normalizing condition (5) must be kept. Then the iterative procedure itself works.

Every iteration consists of the following steps. First, $\{N_{n,j}\}$ and $\{N_{i,m}\}$ are evaluated from Equations (3) and (4), then all other $N_{i,j}$ are determined through Equations (2) or (6). After that the matrix $\|\varphi_{i,j}\|$ is evaluated according to formula (3), and $\{f_i\}$ and $\{s_j\}$ are determined by means of Equations (7) and (8). To avoid divergence of the procedure it is convenient to redetermine s_m and s_{m-1} replacing them by their arithmetic mean:

$$s_m = s_{m-1} = \frac{\sum_{i=1}^{n} (\varphi_{i,m} + \varphi_{i,m-1})}{2\sum_{i=1}^{n} f_i}.$$

Finally, to follow the process of convergence, the sum of squares of residuals for the catches,

$$SS(IT) = \sum_{i=1}^{n} \sum_{j=1}^{m} (C_{i,j} - \hat{C}_{i,j})^2 \tag{9}$$

(if the error in catch-at-age data may be treated as an additive one), or for the logarithms of catches (in the case of multiplicative error structure),

$$SS(IT) = \sum_{i=1}^{n} \sum_{j=1}^{m} (\ln C_{i,j} - \ln \hat{C}_{i,j})^2, \tag{10}$$

is calculated. Here IT is the iteration index, and $\hat{C}_{i,j} = f_i s_j N_{i,j} e^{-M/2}$.

The convergence is regarded to be a true one if SS (IT) becomes stable with the growth of IT, and the asymptotic SS value, SS*, is sufficiently close to min SS (IT) (the experiments carried out demonstrate that SS* is indeed close to min SS (IT) – see Results and discussion).

Minimization of SS*(M, f_n)

As the sizes of the age groups in the last year, $\{N_{n,j}\}$, are determined through Equations (3) and (4), and then the fractions, $\varphi_{i,j}$, are evaluated from Equation (3), the terminal effort, f_n, does not change (i.e. stays equal to the initial guess) during the iterative procedure described. By varying f_n within a certain interval and repeating the whole sequence of calculations for every new f_n value, one can find an estimate of f_n corresponding to the minimum SS*, obtaining the best description of the catch-at-age matrix by the model for the given M (within the region of convergence of the iterative procedure). This idea works out very often (see Results and discussion) and can easily be realized using any standard minimization routine (or even ‘‘by hand’’) as M is assumed known.

This approach can also be applied to solving the problem when M is unknown. In such a case, by taking into account the approximate character of all our estimates and their dependence on the quality of the input data, it is reasonable to suppose M to be constant (to make the model more robust) and minimize SS*(M, f_n) as a function of two variables M and f_n. For this purpose any standard two-variable minimization routine can be used, with our iterative procedure serving as a subroutine for determining the loss function SS*.

Treatment of zero catches

When zeros in the matrix $\|C_{i,j}\|$ indicate missing values of the catch-at-age data, formal use of the procedure, described above, may lead to considerable loss of accuracy of the estimates or even to incorrect results. However, the method described above can easily be generalized so as to enable ‘‘reconstruction’’ of missing data and increase of accuracy. For this purpose it is reasonable to use the effort-controlled version of the model, based on Equation (6) (though it is possible to use the catch-controlled version with Equation (2), referring to Equation (6) only when calculating abundances corresponding to these deceptive zero catches). Further, for all missing catches we formally set $\varphi_{i,j}$ to be zero, but then instead of Equations (7) and (8) we use the following equations for determination of f_i and s_j.

$$f_i = a_i \sum_{j=1}^{m} \varphi_{i,j}, \tag{11}$$

$$s_j = \frac{b_j \sum_{i=1}^{n} \varphi_{i,j}}{\sum_{k=1}^{m} \left(b_k \sum_{i=1}^{n} \varphi_{i,k} \right)}. \tag{12}$$

In Equations (11) and (12)

$$a_i = \frac{m}{m - K_i}, \qquad b_j = \frac{n}{n - L_j},$$

where K_i and L_j are correspondingly the numbers of missing data in the i-th row and j-th column of the matrix $\|C_{i,j}\|$. Naturally, as soon as the series $\{s_j\}$ is computed, its two last elements, s_{m-1} and s_m must be replaced by their average.

Both the original and the generalized procedures work when all the terminal catches, $\{C_{n,j}\}$ and $\{C_{i,m}\}$, are non-zero (see section on Iterative procedure). Otherwise one more simple correction of the iterative procedure should be made. If a certain terminal element of the catch-at-age matrix is zero, the last non-zero catch in the corresponding diagonal should be used in the calculations as a terminal catch for this cohort. For example, if $C_{q,m}=0$ and $C_{q-k,m-k}$ is the last non-zero catch from the corresponding cohort ($0<k<q\leqslant n$ and $k<m$), then $N_{q-k,m-k}$ is determined by Equations (3) and (4), while all earlier $N_{i,j}$ are found through Equation (6) (or (2), see above). Afterwards, the matrix $\|\varphi_{i,j}\|$ (except for the elements $\varphi_{q-k+1,m-k+1}, \ldots, \varphi_{k,m}$) is evaluated according to Equation (3) and $\{f_i\}$ and $\{s_j\}$ are evaluated by means of the pair of Equations (7) and (8) or (11) and (12). Finally, as soon as the series $\{f_i\}$ and $\{s_j\}$ are known, we are able to successively estimate the elements $N_{q-k+1,m-k+1}, \ldots, N_{k,m}$ of the matrix $\|N_{i,j}\|$ through Equation (6), if necessary (certainly, the corresponding estimated catches, $\hat{C}_{q-k+1,m-k+1}, \ldots, \hat{C}_{k,m}$, will always differ from zero).

Data for testing the model

For the purpose of testing the approach, seven simulated data sets, DS0–DS6, were utilized. The catch-at-age matrix DS0 was generated with the use of an operating model based on the equations:

$$N_{i+r,1} = R(P_i), \qquad P_i = \sum_{j=1}^{m} N_{i,j} \tag{13}$$

$(i \geq r+1)$,

$$N_{i+1,j+1} = N_{i,j} e^{-M}(1 - \varphi_{i,j}) \tag{14}$$

$(i \geq 1,\ j=1, \ldots, m-1)$, and (3)–(5), where

- r – recruitment age,
- P_i – parental (and exploited) stock in the i-th year which we will call simply a stock,
- $R(P_i)$ – stock–recruitment function (known).

The initial conditions for Equation (14), i.e. the age distribution of the abundance in the first year, $\{N_{1,j}\}$, as well as the first age group abundance for the next $r-1$ years, $N_{2,1}, \ldots, N_{r,1}$ (if $r>1$) were set.

This operating model can be regarded as a reversed analogue, with respect to time and age, of the initial effort-controlled model (3)–(6). Here $N_{i+1,j+1}$ is calculated as soon as $N_{i,j}$ has been calculated, and therefore we call this a perspective model. Now the vectors $\{f_i\}$, $\{s_j\}$, and natural mortality coefficient, M, are given, while the catch-at-age matrix is one of the results of simulation. For definiteness, the age of recruitment, r, is assumed to be equal to the age of maturation, and this is given the index 1 (though, in fact the model permits the age of maturation to exceed r). In the course of simulation, the Shepherd (1982) stock–recruitment relationship was used.

The simulation procedure consisted of two parts. At the first stage, a carrying capacity steady state of a non-exploited population containing eight age groups was simulated by setting zero fishing effort and continuing the computations until the changes in all the model variables became negligible. Note that the carrying capacity level in the model (3)–(5), (13), (14) depends only on M (taken as 0.2) and on the form and parameters of the recruitment relationship. At the second stage, the harvesting started and continued for 44 years yielding the matrices $\|C_{i,j}\|$ and $\|N_{i,j}\|$ ($i=1, \ldots, 44$; $j=1, \ldots, 8$). The obtained matrix $\|C_{i,j}\|$ served as the ''pure'' input data set, DS0, for ISVPA, while the given vectors $\{f_i\}$ and $\{s_j\}$, and the obtained matrix $\|N_{i,j}\|$ were regarded as standards (''true'' values) for the subsequent comparison with the results produced by ISVPA.

Data sets DS1–DS4 containing measurement errors were obtained from DS0 by adding normally distributed random errors (noise), $err(\sigma)C_{i,j}$, to the catches $C_{i,j}$. The relative error, $err(\sigma)$, had zero mathematical expectation and standard deviation $\sigma=0.1$, 0.2, 0.3 and 0.4 for DS1, DS2, DS3 and DS4 correspondingly.

In order to test the ability of ISVPA to reconstruct the parameters of the population and fishing in the case of continuous (throughout the year) fishing, two more simulated data sets, DS5 and DS6, have been utilized. These are the most noisy data among the six simulated data sets provided by the ICES Workshop on Methods of Fish Stock Assessments in Reykjavik in 1988 by means of a rather complicated operating model based on the conventional principles; the two data sets are referred to in Anon. (1993) as Data Sets 5 and 6. Fishing exploitation of a stock by two trawler and one longliner fleets, by one fleet with fixed nets, and by three research vessels, was simulated providing estimates of catch-at-age data for ages 3–12 for a period of 30 years for each of seven fleets. Catchability was assumed to contain trends in the two commercial fleets. Noise was introduced into the data in the form of lognormal and

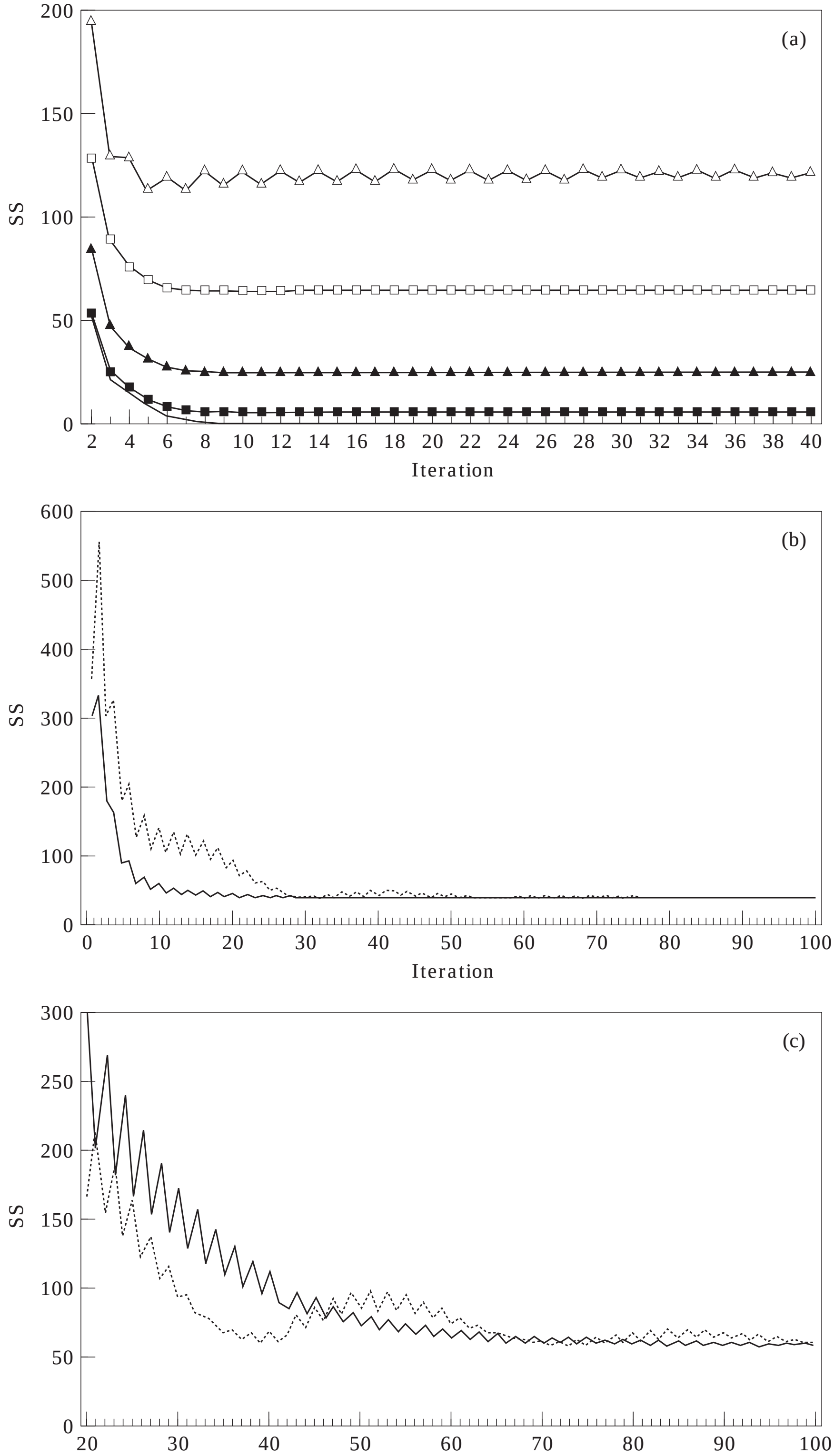

Figure 1. Convergence of the iterative procedure in terms of SS (IT) for data sets (a) DS0–DS4 (DS0 (——), DS1 (—■—), DS2 (—▲—), DS3 (—□—), DS4 (—△—), (b) DS5, and (c) DS6: 1 – given M = 0.2 (——), 2 – estimated M (· · ·).

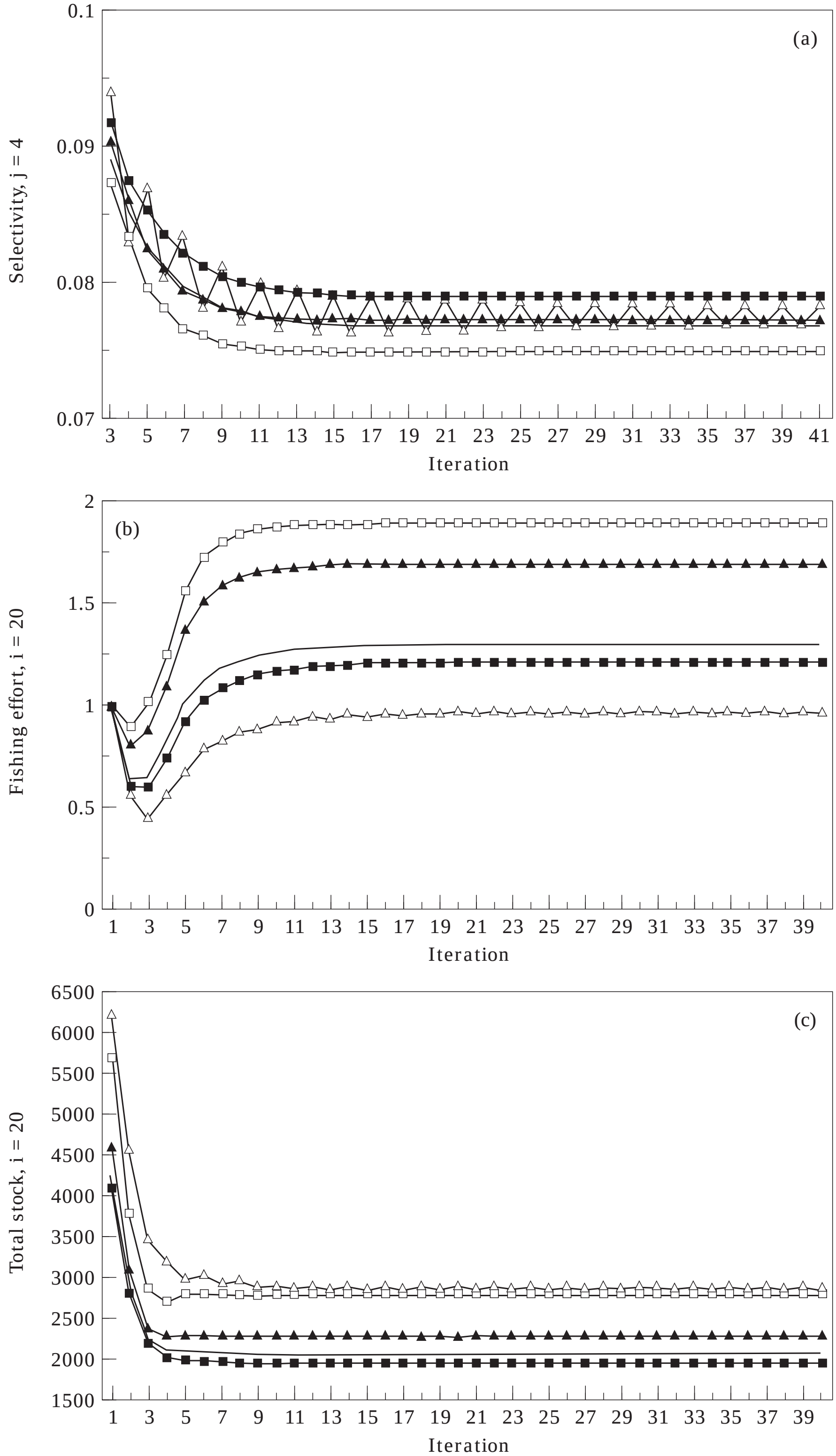

Figure 2. Convergence in terms of (a) selectivity at j = 4, (b) fishing effort at i = 20, and (c) total stock at i = 20 for data sets DS0–DS4: DS0 (——), DS1 (—■—), DS2 (—▲—), DS3 (—□—), DS4 (—△—).

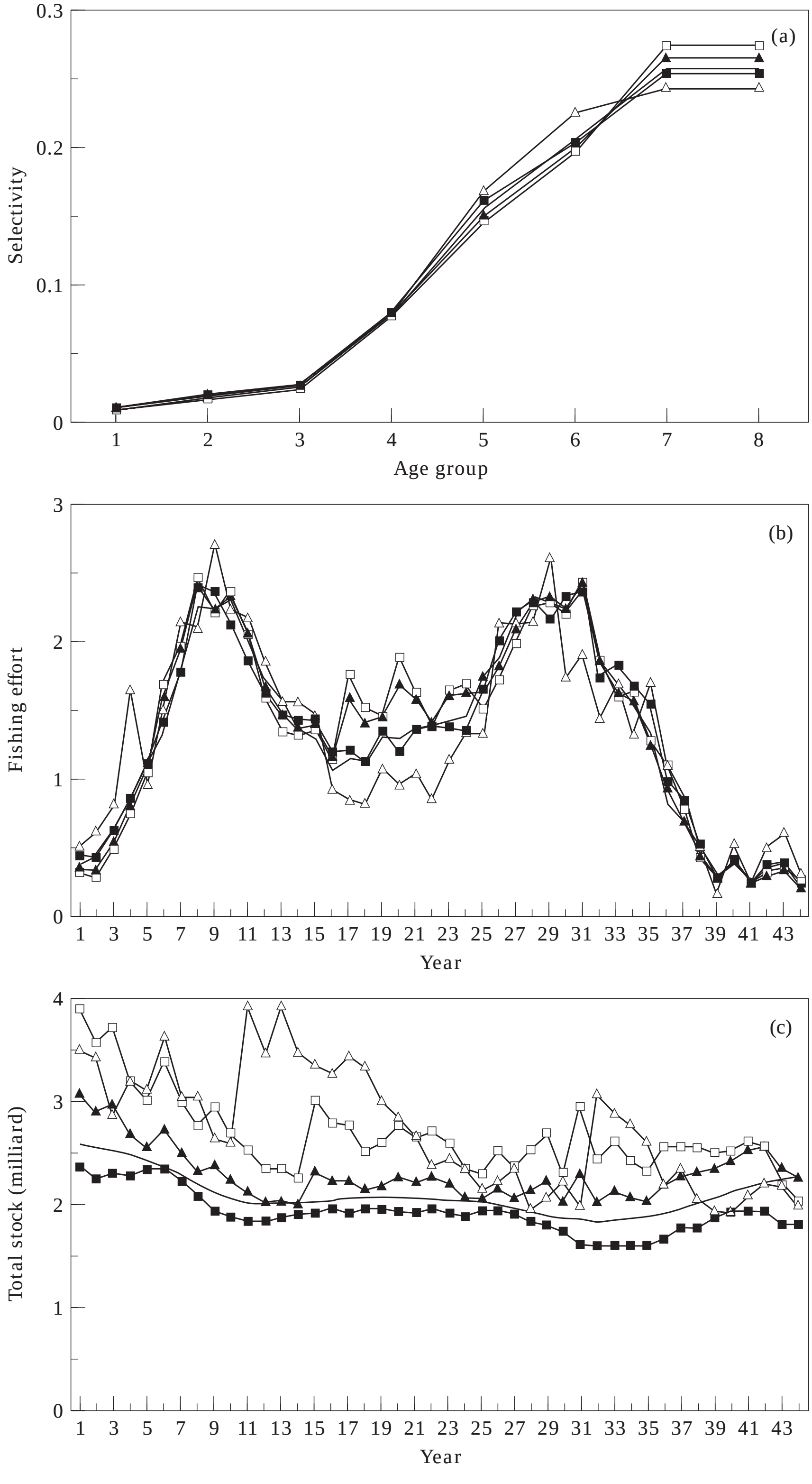

Figure 3. Estimated variables: (a) selectivity, (b) fishing effort, and (c) total stock for data sets DS0–DS4 (the estimates for DS0 exactly coincide with the ''true'' values): DS0 (——), DS1 (—■—), DS2 (—▲—), DS3 (—□—), DS4 (—△—).

Table 1. Frequency distribution of the percentage discrepancy between the estimated and true stock size, DS1.

Range of discrepancy	1	2	3	Stock-at-age 4	5	6	7	8	Total stock
70% ≤ discrepancy	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50% ≤ dis. < 70%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30% ≤ dis. < 50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.5	0.0
10% ≤ dis. < 30%	0.0	2.3	2.3	2.3	4.5	9.1	9.1	13.6	0.0
− 10% ≤ dis. < 10%	43.2	52.2	61.4	70.5	77.3	72.7	72.7	63.6	72.7
− 30% ≤ dis. < − 10%	54.5	45.5	36.3	27.2	18.2	18.2	18.2	15.9	27.3
− 50% ≤ dis. < − 30%	2.3	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.0
− 70% ≤ dis. < − 50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
discrepancy < − 70%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 2. Frequency distribution of the percentage discrepancy between the estimated and true stock size, DS2.

Range of discrepancy	1	2	3	Stock-at-age 4	5	6	7	8	Total stock
70% ≤ discrepancy	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50% ≤ dis. < 70%	6.8	9.1	6.8	6.8	6.8	4.5	2.3	0.0	0.0
30% ≤ dis. < 50%	13.6	13.6	11.4	11.4	6.8	6.8	9.1	11.4	0.0
10% ≤ dis. < 30%	25.0	22.7	27.3	22.7	25.0	25.0	25.0	22.7	0.0
− 10% ≤ dis. < 10%	45.5	43.2	43.2	47.7	40.9	36.4	29.5	18.2	87.5
− 30% ≤ dis. < − 10%	9.1	11.4	11.3	11.4	20.5	27.3	31.8	36.4	12.5
− 50% ≤ dis. < − 30%	0.0	0.0	0.0	0.0	0.0	0.0	2.3	11.3	0.0
− 70% ≤ dis. < − 50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
discrepancy < − 70%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

gamma-distributed process and measurement errors (different for different age groups and fleets), while natural mortality rate was assumed to be 0.2 for all ages and years. DS5 is characterized by separability of fishing mortality rate at age, F, for each fleet, while F in DS6 is not separable. All the details regarding the simulation of DS5 and DS6 can be found in Anon. (1993).

Results and discussion

Both the catch- and effort-controlled version of ISVPA were examined in the experiments by using the simulated data described in the section on Data for testing the model. The limit (asymptotic) SS (IT) value, SS*, was indeed rather close to min SS (IT) for both of the criteria (9) and (10), and the estimates were accurate enough for the data sets DS1 and DS2 with low level of noise (see, e.g. Figs 1–3). In the case of ‘‘pure’’ data, DS0, the estimates of all the parameters of stock and fishery dynamics completely coincide with the ‘‘true’’ parameters independently of the type of model (catch- or effort-controlled). Correspondingly, in this case $SS^*(M, f_n)=0$ for both of the criteria (9) and (10). However, in the cases of considerable noise in the data (DS3 and DS4), the minimum $SS^*(M, f_n)$ was better pronounced when the effort-controlled version of ISVPA along with the criterion (10) was used. That is why the main results of the numerical experiments corresponding to this case are presented below.

Convergence

The iterative procedure, described above converges within a rather wide range of M and f_n values independently of the initial distributions of f_i and s_j, whatever catch-at-age matrix is taken. The illustrations of convergence, presented in Figures 1 and 2, correspond to the ‘‘best’’ M and f_n, those providing the minimum of $SS^*(M, f_n)$, and to homogeneous initial distributions of the selectivity and fishing effort, $s_j=1/m$, $f_i=1$ (at $i \le n-1$). All the variables estimated within the iterative procedure are incorporated in the function SS(IT) given by Equation (10), and, as it can be seen from Figure 1a and Figure 2, they converge simultaneously with SS(IT). This proves that the stabilization of SS(IT) may indeed be used as a criterion for stopping the iterative process.

It is clear from Figure 1(a) that the higher level of noise in the catch-at-age data causes higher values of

Table 3. Frequency distribution of the percentage discrepancy between the estimated and true stock size, DS3.

Range of discrepancy	Stock-at-age 1	2	3	4	5	6	7	8	Total stock
70% ≤ discrepancy	22.7	18.2	13.6	9.1	11.4	11.4	9.1	4.5	0.0
50% ≤ dis. < 70%	6.8	11.4	13.6	11.4	4.5	2.3	4.5	4.5	4.5
30% ≤ dis. < 50%	15.9	9.1	6.8	11.4	15.9	11.4	9.1	4.5	40.9
10% ≤ dis. < 30%	40.9	38.6	22.7	18.2	6.9	13.6	15.9	22.7	50.0
− 10% ≤ dis. < 10%	2.3	13.6	34.2	40.8	43.2	34.1	22.8	13.9	2.3
− 30% ≤ dis. < − 10%	9.1	6.8	6.8	6.8	13.6	20.4	25.0	22.7	2.3
− 50% ≤ dis. < − 30%	2.3	2.3	2.3	2.3	4.5	6.8	13.6	22.7	0.0
− 70% ≤ dis. < − 50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
discrepancy < − 70%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 4. Frequency distribution of the percentage discrepancy between the noisy and true catch, and between the estimated and true stock size, DS4.

Range of discrepancy	Catch-at-age 1	2	3	4	5	6	7	8	Total catch
70% ≤ discrepancy	9.1	4.5	9.1	9.1	4.5	4.5	4.5	2.3	0.0
50% ≤ dis. < 70%	13.6	15.9	6.8	13.6	11.4	9.1	2.3	6.8	0.0
30% ≤ dis. < 50%	11.4	18.2	20.5	13.6	13.6	15.9	6.8	2.3	0.0
10% ≤ dis. < 30%	6.8	15.9	20.5	9.1	22.7	9.1	11.4	18.2	38.6
− 10% ≤ dis. < 10%	29.6	18.3	4.5	22.9	22.7	22.8	15.9	25.0	45.5
− 30% ≤ dis. < − 10%	11.4	6.8	13.6	15.9	11.4	25.0	29.5	18.2	13.6
− 50% ≤ dis. < − 30%	9.1	13.6	11.4	6.8	11.4	6.8	11.4	18.2	2.3
− 70% ≤ dis. < − 50%	4.5	4.5	9.1	4.5	2.3	4.5	15.9	4.5	0.0
discrepancy < − 70%	4.5	2.3	4.5	4.5	0.0	2.3	2.3	4.5	0.0
Range of discrepancy	**Stock-at-age 1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total stock**
70% ≤ discrepancy	22.7	18.2	13.6	13.6	9.1	9.1	9.1	11.4	9.1
50% ≤ dis. < 70%	15.9	15.9	13.6	9.1	4.5	2.3	2.3	2.3	15.8
30% ≤ dis. < 50%	13.6	13.6	18.2	18.2	22.7	15.9	9.1	11.4	18.2
10% ≤ dis. < 30%	18.2	25.0	18.2	15.9	15.9	15.9	11.4	2.3	34.1
− 10% ≤ dis. < 10%	11.4	9.1	13.6	15.9	18.2	27.3	29.5	38.6	20.5
− 30% ≤ dis. < − 10%	11.4	9.1	11.4	11.4	9.1	2.3	13.6	13.6	2.3
− 50% ≤ dis. < − 30%	6.8	9.1	11.4	15.9	18.2	22.7	11.4	4.5	0.0
− 70% ≤ dis. < − 50%	0.0	0.0	0.0	0.0	2.3	4.5	13.6	15.9	0.0
discrepancy < − 70%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

SS(IT) and, hence, of SS*(M, f_n). This is due to a worse description of the dynamics of the stock–fishery system by the model as compared with the ''true'' (i.e. simulated) dynamics (Fig. 3).

The convergence in the case of DS0–DS3 is very fast and actually takes 15 to 20 iterations. In the case of DS4, slowly decaying saw-tooth type oscillations with a 2-year periodicity can be observed. Although the trends in SS(IT) and other parameters stabilize rather fast, the above mentioned oscillations become significant at the stage of searching for the minimum of SS*(M, f_n), and therefore about 150 to 200 iterations are required. Convergence of the iterative procedure for data sets DS5 and DS6 was observed after about 250 iterations.

Quality of the estimation

All of the estimated values of the natural mortality coefficient, M = 0.18, 0.22, 0.26 and 0.25 for the data sets DS1–DS4, respectively, are close to the ''true'' M = 0.2, though the estimates for DS1 and DS2 are somewhat more accurate than those for DS3 and DS4. Similarly, the quality of the estimates of other parameters obtained by processing the data sets DS1–DS4, in general, falls with the growth of noise in the catch data (Fig. 3). Both the estimated efforts, $\{f_i\}$, and selectivities, $\{s_j\}$, closely correspond to the ''true'' ones. The highest deviations in the estimated efforts correspond to DS3 and DS4 and fall mainly in the period from i = 17 to i = 22 (where they

Table 5. Frequency distribution of the percentage discrepancy between the estimated and true stock size, DS5.

Range of discrepancy	Stock-at-age (M is given) 3	4	5	6	7	8	9	10	11	12	Total stock
$70\% \leq$ discrepancy	6.7	0.0	0.0	0.0	0.0	3.3	0.0	0.0	0.0	13.3	0.0
$50\% \leq$ dis. $< 70\%$	6.7	10.0	10.0	10.0	6.7	6.7	3.3	3.3	3.3	0.0	3.3
$30\% \leq$ dis. $< 50\%$	10.0	10.0	13.3	13.3	16.7	10.0	6.7	0.0	0.0	3.3	3.3
$10\% \leq$ dis. $< 30\%$	23.3	23.3	16.7	13.3	16.7	10.0	20.0	20.0	3.3	3.3	13.4
$-10\% \leq$ dis. $< 10\%$	13.3	16.7	20.0	26.7	16.7	20	13.3	16.7	23.4	10.0	63.3
$-30\% \leq$ dis. $< -10\%$	33.3	33.3	33.3	30.0	33.2	40.0	36.7	20.0	20.0	20.0	16.7
$-50\% \leq$ dis. $< -30\%$	6.7	6.7	6.7	6.7	10.0	10.0	16.7	33.3	40.0	23.4	0.0
$-70\% \leq$ dis. $< -50\%$	0.0	0.0	0.0	0.0	0.0	0.0	3.3	6.7	10.0	16.7	0.0
discrepancy $< -70\%$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	0.0
Range of discrepancy	**Stock-at-age (M is estimated) 3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total stock**
$70\% \leq$ discrepancy	3.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3	0.0
$50\% \leq$ dis. $< 70\%$	0.0	0.0	0.0	0.0	3.3	3.3	3.3	3.3	0.0	0.0	0.0
$30\% \leq$ dis. $< 50\%$	10.0	10.0	10.0	10.0	6.7	6.7	6.7	0.0	3.3	3.3	3.3
$10\% \leq$ dis. $< 30\%$	6.7	16.7	16.6	23.3	20.0	20.0	13.3	6.7	0.0	3.3	3.3
$-10\% \leq$ dis. $< 10\%$	26.7	26.7	26.7	20.0	13.3	13.3	10.0	23.3	26.6	10.0	23.4
$-30\% \leq$ dis. $< -10\%$	26.7	23.3	26.7	26.7	30.0	30.0	30.0	20.0	16.8	20.0	66.7
$-50\% \leq$ dis. $< -30\%$	23.3	23.3	20.0	20.0	26.7	20.0	30.0	36.7	33.3	23.4	3.3
$-70\% \leq$ dis. $< -50\%$	3.3	0.0	0.0	0.0	0.0	6.7	6.7	10.0	20.0	16.7	0.0
discrepancy $< -70\%$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	0.0

Table 6. Frequency distribution of the percentage discrepancy between the estimated and true stock size, DS6.

Range of discrepancy	Stock-at-age (M is given) 3	4	5	6	7	8	9	10	11	12	Total stock
$70\% \leq$ discrepancy	10.0	10.0	10.0	10.0	10.0	10.0	13.3	13.3	13.3	10.0	0.0
$50\% \leq$ dis. $< 70\%$	3.3	3.3	3.3	3.3	3.3	3.3	0.0	0.0	0.0	3.3	6.7
$30\% \leq$ dis. $< 50\%$	3.3	3.3	6.7	6.7	6.7	3.3	3.3	3.3	3.3	3.3	3.3
$10\% \leq$ dis. $< 30\%$	13.3	10.0	10.0	13.4	10.0	16.7	16.7	16.7	3.3	3.3	26.7
$-10\% \leq$ dis. $< 10\%$	33.4	33.4	26.7	23.3	20.0	10.0	13.3	6.7	20.0	6.7	53.3
$-30\% \leq$ dis. $< -10\%$	30.0	30.0	33.3	30.0	36.6	43.3	30.0	26.7	20.0	23.4	10.0
$-50\% \leq$ dis. $< -30\%$	6.7	10.0	6.7	10.0	6.7	6.7	13.4	20.0	26.8	16.7	0.0
$-70\% \leq$ dis. $< -50\%$	0.0	0.0	3.3	3.3	6.7	6.7	10.0	13.3	13.3	23.3	0.0
discrepancy $< -70\%$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	0.0
Range of discrepancy	**Stock-at-age (M is estimated) 3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total stock**
$70\% \leq$ discrepancy	6.7	6.7	6.7	6.7	6.7	6.7	3.3	3.3	0.0	10.0	0.0
$50\% \leq$ dis. $< 70\%$	0.0	0.0	3.3	3.3	0.0	0.0	10.0	6.7	13.3	3.3	0.0
$30\% \leq$ dis. $< 50\%$	3.3	6.7	3.3	3.3	6.7	6.7	0.0	3.3	0.0	3.3	3.3
$10\% \leq$ dis. $< 30\%$	3.3	0.0	6.7	6.7	6.7	6.7	6.7	3.3	3.3	3.3	6.7
$-10\% \leq$ dis. $< 10\%$	13.3	16.7	16.7	23.3	20.0	16.6	16.7	20.0	6.7	6.7	26.7
$-30\% \leq$ dis. $< -10\%$	43.3	43.3	36.7	33.3	33.3	33.3	36.6	13.3	20.0	23.4	56.6
$-50\% \leq$ dis. $< -30\%$	23.4	16.6	16.6	16.7	20.0	20.0	26.7	33.4	43.3	16.7	6.7
$-70\% \leq$ dis. $< -50\%$	6.7	10.0	10	6.7	6.6	10.0	10.0	16.7	10.0	23.3	0.0
discrepancy $< -70\%$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3	10.0	0.0

reach about 40%). The errors in the estimated selectivities are smaller and reach their maximum at the youngest age groups (25% at $j=1$ for DS3), while the highest error at $j=8$ is only about 8% (DS4).

The frequency distributions of the percentage discrepancy between the estimated and "true" stock-at-age (and catch-at-age for DS4) values are presented in Tables 1–6. The discrepancies are determined as

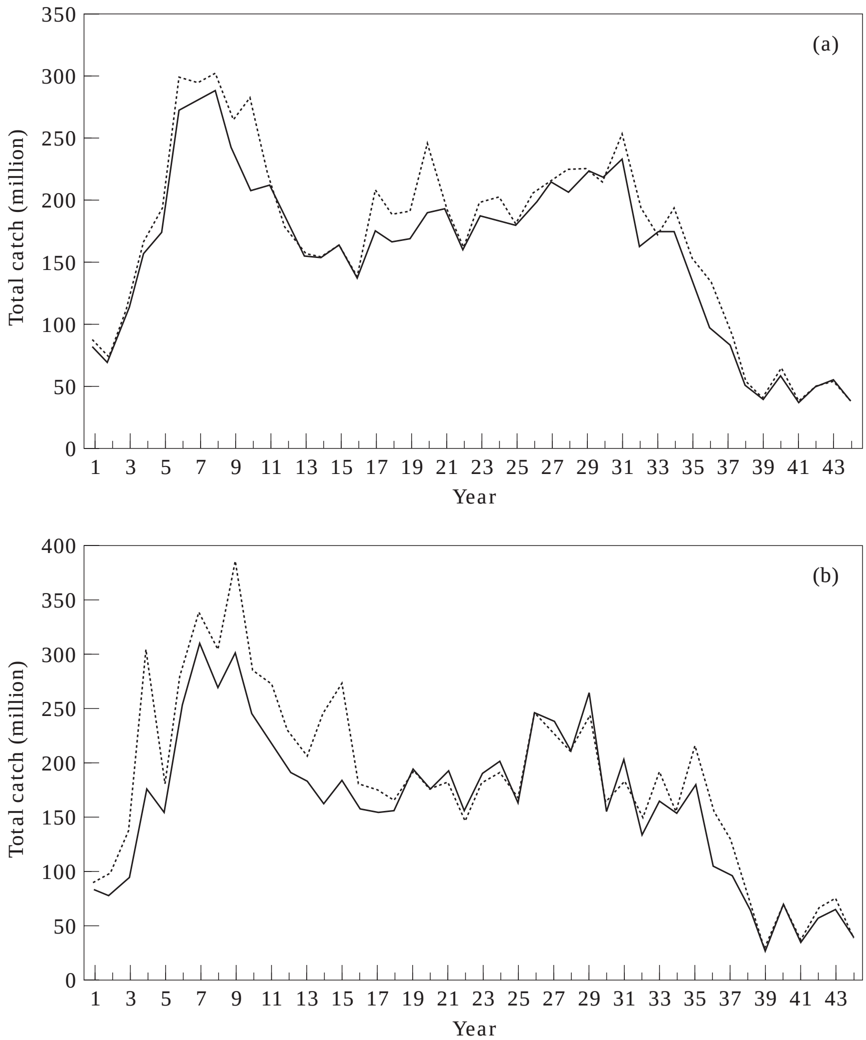

Figure 4. The approximation of noisy total catch by the model for (a) DS3 and (b) DS4: 1-noisy catch (——), 2-estimated catch (· · ·).

100[(Estimate/Truth) − 1], following the recommendations of Anon. (1993). The deviations of the estimated total stock size with respect to the ''true'' increase from DS1 to DS4 (Tables 1–4), and in the case of DS4 the maximal deviations are much higher (about 100% at i = 11 and 13) than for the estimates of fishing effort. Such high deviations can be explained by the occasional emergence of a number of cohorts with enormously systematic positive errors in the catch-at-age matrix (see Table 4). For example, the mean ratio between the noisy (DS4) and ''true'' (DS0) catches taken from the cohort, recruited in the year i = 11, is about 1.5 (while the maximal is 2). As the estimate of the natural mortality coefficient (M = 0.25) was rather close to the ''true'' value (M = 0.20), the accumulation of these systematic positive errors in the catches in the course of calculating the cohort size back from i = 18, j = 8 to i = 11, j = 1 became apparent in the extremely high cohort sizes in the early years (especially, at i = 11), i.e. in the youngest age groups. This accumulation resulted in the peak at i = 11 (DS4) on Figure 3. Similarly, a check up of the peaks at i = 1, 6, 13, 17 and 32 showed that they (and the rise of the graph within the interval from i = 11 to 21) are the results of accumulation of systematic positive errors in catches. This effect can also be observed in Table 4. So, the model actually reconstructs a virtual population dynamics that corresponds to the noisy catch data in the best way. The quality of such a reconstruction in the case of DS3 and DS4 can be judged by comparison of the estimated catches, $f_i s_j N_{i,j} e^{-M/2}$, with the noisy data (Fig. 4). From this point of view the results of processing DS3 and DS4 look much better.

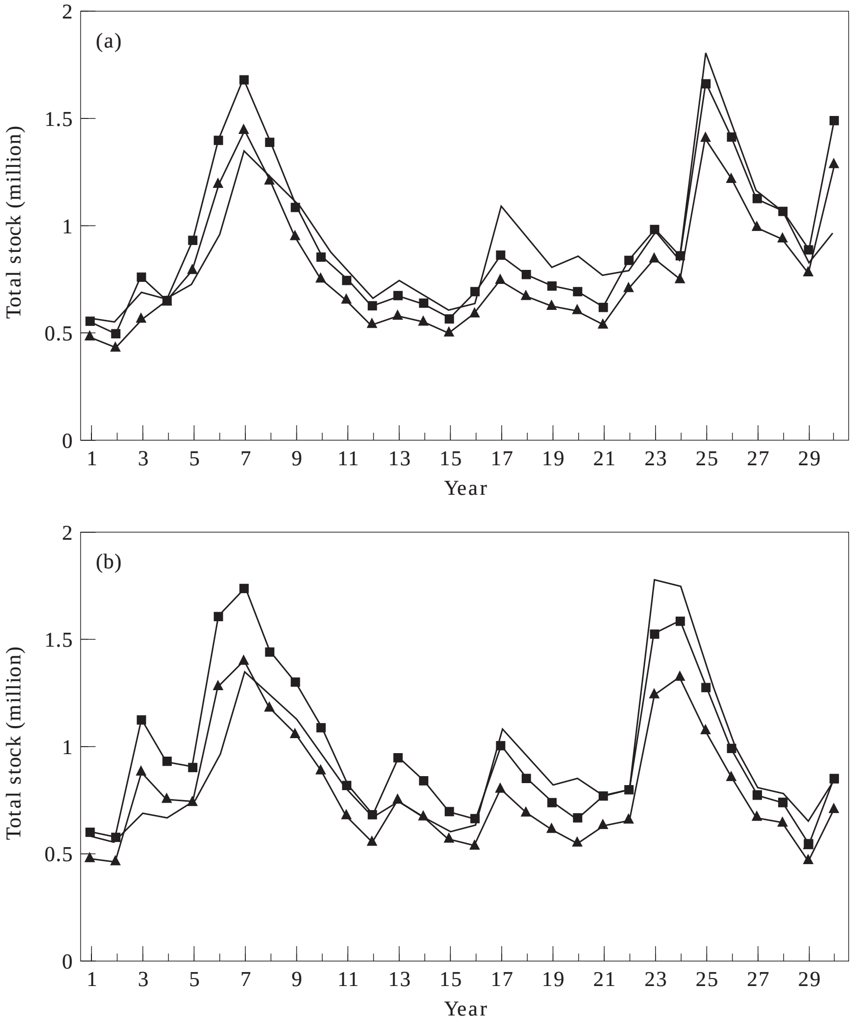

Figure 5. The ''true'' and estimated total stocks for (a) DS5 and (b) DS6: 1 – ''true'' stock (——), 2 – estimated at given M = 0.2 (f_n corresponds to min SS* (0.2, f_n)) (—■—); 3 – estimated at M and f_n providing min SS* (M, f_n) (—▲—).

In the case of DS5 and DS6 we analysed both the problem with the given natural mortality coefficient (M = 0.2, SS* depends only on f_n), and the full problem that includes estimation of M and f_n by minimization of the function SS* (M, f_n). The estimates of M in the latter case are M = 0.16 for DS5 and M = 0.15 for DS6. The results of the total stock assessment are presented in Figure 5. In each of the cases the stock dynamics was reconstructed rather well (though the results for DS6 in the first years look somewhat better when M is estimated equally with f_n, while in the last years they are better when M = 0.2 is set). The corresponding statistics of the deviations are presented in Tables 5 and 6.

Conclusion

Both hypotheses of a constant fishing mortality coefficient during the year (which is characteristic for conventional VPA and standard separable VPA) and of instantaneous harvesting (ISVPA) are approximations of the real situation when the fishing intensity varies during a year. The results of the numerical experiments demonstrate the ability of ISVPA to provide a sufficiently good description of the stock–fishery dynamics (without using any auxiliary information such as survey results etc.) not only in the case of instantaneous fishing, but also when the stock is continuously exploited.

Main difficulties in dealing with VPA techniques are usually related to uncertainty in natural mortality coefficient values and the terminal fishing rate. So, the capability of the method to determine f_n and M (under certain conditions), proved by the experiments, must be regarded as one of the merits of ISVPA.

Another problem which one may come across both at the stage of tuning VPA or interpretation of its results is connected with determination of a correspondence between the fishing mortality coefficient and the fishing effort. It is clear that in the case of ISVPA such a problem does not arise due to the direct proportionality between the parameter f_i and the fishing effort.

A by-product of our work, the perspective model (3)–(5), (13), and (14), being combined with the ISVPA model (2)–(5) or (3)–(6), can serve for constructing sustainable yield curves, but this deviates from the subject of the present paper, and the corresponding results will be published separately.

Acknowledgements

The authors use this opportunity to thank Dr Chris Darby and another, anonymous reviewer for their very helpful comments and recommendations that considerably promoted our work and the presentation of its results.

References

Anon. 1993. Reports of the Working Groups on Methods of Fish Stock Assessment. Copenhagen, 9–16 June 1987, Reykjavik, 6–12 July 1988, Nantes, 10–17 November 1989. ICES cooperative research report No. 191. International Council for the Exploration of the Sea. Copenhagen, Denmark.

Lewy, P. 1988. Integrated stochastic virtual population analysis: estimates and their precision of fishing mortalities and stock sizes for the North Sea whiting stock. Journal du Conseil International pour l'Exploration de la Mer, 44: 217–228.

Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. ICNAF Research Bulletin, 9: 65–74.

Pope, J. G. and Shepherd, J. G. 1982. A simple method for the consistent interpretation of catch-at-age data. Journal du Conseil International pour l'Exploration de la Mer, 40: 176–184.

Shepherd, J. G. 1982. A versatile new stock–recruitment relationship for fisheries, and the construction of sustainable yield curves. Journal du Conseil International pour l'Exploration de la Mer, 40: 67–75.