TISVPA overview

A characteristic feature of ISVPA (including its special "triple-separable" case - TISVPA) consists in intentional implementation of principles of robust statistics in procedures of estimation of the model parameters. Just this helps to diminish the influence of errors in the data on the results of analysis and to extract more information about stock-fishery system from available data. The theory of robust statistics was formed in the middle of the XX-th century, although its origin may be attributed to much earlier works considering various distributions and behavior of various statistics. Interest to robustness, which generally may be understood as a measure of independence of results on assumptions standing behind the procedures used to get these results, was dealt with the perception that in many cases common statistic assumptions, like normality of distribution and independence of observations, are only a very rough idealization of real situation. It turned out that statistics and methods of their estimation, ideal in "classic" conditions, for real data, including substantial noise component, often gives unacceptable result. Just this situation is common for fish stock assessment. The theory of robustness is a deeply developed area of statistics at present, but its practical implementation for estimation of parameters of fish stock assessment models is still rather rare and "ad hoc". The ISVPA model includes several features which help it to operate with "real" (that is strongly noisy) data. Among them: robust loss functions, possibility to ensure unbiased solution, independence of estimated selection pattern upon user's choice about its overall shape, implementation of different options concerning mutual validity of assumptions about quality of catch-at-age data and stability of selection pattern, possibility to exclude influence of year-to-year survey catchability variations caused by difference in survey conditions, etc.

The formulations of data source – specific components of the overall model loss function could be quite different depending on properties of the data set under consideration.

The model does not work with likelihoods, which are very popular because they helps to give ideologically clear mixing of evidences from different sources of data. The reason for that is the following: classic likelihood functions are known to be extremely non-robust. For example, Y. Chen and D. Fournier (1999) reminded that "in formulating likelihood functions, data have been analyzed as if they are normally, identically, and independently distributed. It has come to be believed that the first two of the assumptions are frequently inappropriate in fisheries studies. In fact, data distributions are likely to be leptokurtic and (or) contaminated by occasional bad values giving rise to outliers in many fisheries studies." In their studies Y. Chen and D. Fournier showed that "the existence of outliers may greatly bias the derived posterior distributions. The likelihood of having outliers in fisheries studies implies that posterior distributions may be unreliable. This may lead to erroneous results on the dynamics of fish stocks and subsequently the adaptation of an inappropriate strategy in managing fisheries resources." Noel G. Cadigan and Ransom A. Myers (2001) examined two maximum likelihood estimators of the sequential population model (SPA) parameters. These estimators were based on assumption that the stock-size indices were described by the lognormal or gamma distributions. Using simulations, the authors found that both types of estimators could have significant biases; however, their results indicated that it could be preferable to use the gamma model, because it tended to have lower bias and variability, even when the true distribution of the stock-size indices was lognormal.

Generally speaking, the following approaches are commonly used in order to overcome non-robust properties of classic likelihoods:
- implementation of classic distributions with heavy tails (in order to better accommodate outliers);
- implementation of mixed, or so called mixture- distributions;
- implementation of exotic and extremely flexible distributions.

A lot of robust distributions are summarized, for example, by K.Passarin (2004), including: power-series distributions; extended power distribution; the Student-t distribution; elliptical distribution, which is considered to be a family of symmetric distributions including *inter alia* the normal and the Student−t distributions.

This group of approaches is aimed at a better accommodation of large errors in observation. But the question whether by doing this we can sufficiently reduce the influence of outliers on the solution still remains. F. Hampel (2002) underlines that …"The most common way out in practice still seems to be the replacement of the original parametric model, such as normality, by another, more complicated ad hoc model. *These models are, strictly speaking, as unrealistic as the original model*; if (as is frequently the case) they are chosen with good intuition, they do work for a full neighborhood of the original model, *but this can only be proven by robustness theory*."

Another group of approaches comprises construction of quasi-likelihoods based on reduced influence of "bad points" (M-estimates, etc.) (e.g. see Hampel et al., 1986). But when the model simultaneously includes a number of different quasi-likelihoods with artificially reduced influence of some points, the problem of mutual weighting of information from different sources rises again.

F. Hampel (2002) also pointed that "...some Bayesians may want to cling to their original model and to an unmodified likelihood function, yet be somewhat robust. For them I offer the following tentative suggestion. All they have to do is to replace the most extreme observations by pseudo-observations, which behave like data from the ideal model and do not contain dangerous outliers". This seems to be a very useful generalized suggestion which in many cases could be easier to implement rather than to construct an exotic likelihood and then to prove that it really helps. There are a lot of ways to detect outliers and then to do something with them. One of the most simple ones is α-winsorization (Huber, 1981) and procedures built on its basis. Unfortunately, the procedure of classic α-winsorization is itself not robust since it is based on such non-robust measure of scale, as standard deviation. However, it is not difficult to develop a more robust and efficient winsorization procedure (Vasilyev, 2004) on the basis, for example, of so called "X-84"-rule introduced by P.Huber (Hampel et al., 1986). For the simulated catch-at-age data containing 5% of outliers it was shown (Vasilyev, 2004) that "robustified" winsorization procedure gave almost twice better improvement of results of stock assessment compared to classic procedure (see Vasilyev, 2005).

It is also necessary to mention that dealing with procedures of iterative detection and improvement of outliers it is important to have a sufficiently good initial estimate, i.e. the one obtained using all points. Otherwise the procedure can run in the wrong direction. Therefore, the model itself should be already sufficiently robust, when it is applied to all data including not yet detected outliers. Often, if the initial model is robust enough, it appears that further games with outliers are comparatively not so much important. That is why in the ISVPA (TISVPA) model parameters estimation procedures the choice is made for robust lost functions rather then to likelihoods which are known to have problems with robustness.

The ISVPA model was tested within the ICES procedures and is used for data exploration and stock assessment for several stocks within the ICES Working Groups.

The special case of the ISVPA model - TISVPA is intended to better reflect generation - dependent peculiarities in selection pattern which can be caused by many reasons, including natural and artificial ones (see Vasilyev, 2006).

The TISVPA (Triple Instantaneous Separable VPA) model in the version which was discussed at the ICES Working Group on Methods of Fish Stock Assessment (2006) is an extension of the ISVPA model (of its version 2004.3, first presented at the ICES Study Group on Assessment Methods Applicable to Assessment of Norwegian Spring-Spawning Herring and Blue Whiting Stocks (ICES, 2004)). In a few words, the model now can represent fishing mortality coefficients (more precisely – exploitation rates) as a product of three parameters: f(year)*s(age)*g(cohort), that is it gives possibility to estimate within the model an additional set of generation-dependent parameters in separable representation of exploitation rates. This set of parameters is intended to adapt traditional separable representation of fishing mortality (as a product of age-dependent and year-dependent factors) to situations when several year classes may have peculiarities in their interaction with fishing fleets caused by different spatial distribution, higher attractiveness of more abundant schools to fishermen, or by some other reasons. There also could be some artificial reasons which cause generation-dependent peculiarities in selection pattern, for example, errors in aging.

The above mentioned generation-dependent factors (g-factors) can be estimated and applied not to the whole interval of age groups used in the model, but to some age “window”. This helps (1) to be closer to real situations (when it is known that only some range of age groups may have generation-dependent peculiarities in their interactions with fishing fleets) and (2) to diminish the influence of age groups having data of lower quality (usually these are youngest and oldest ages). For age groups which are outside the chosen age interval, the g-factors are stated to be equal to unit, but as a result of global normalization of all g-factors to unit by average, used in the model to balance the parameter estimation procedure, they may get somewhat different values. The user can choose the age-window for application of g-factors by setting the first and the last age for estimation of g-factors. He also can not use them at all – in such a case the TISVPA model is reduced to the ISVPA model.

Two sub-models with respect to the generation-dependent peculiarities in selection pattern are reserved in the model:

1 - sub-model of “within-year effort redistribution by ages”;

2- sub-model of “gain (loss) in selection”.

The first sub-model assumes that in each year more fishing-attractive cohorts borrow some amount of fishing effort from other cohorts by increasing its selection at the expense of diminished selections for other age groups in this year. The second one assumes that some cohorts has increased (or reduced) selections, but this does not cause direct change in selections for others.

The same way, as in ISVPA, the TISVPA parameter estimation procedures are based on some principles of robust statistics what helps to diminish influence of errors (noise) in the data on the assessment results and contain options, similar to those of the ISVPA model, addressing statistical meaning of the catch-at-age - based solution and dealing with auxiliary data.

Brief description of the model is summarized in the table below; the detailed algorithm of the TISVPA model is given in the Appendix.

Model	**TISVPA**
Version	**2009.1**
Model type	A separable model is applied to one or two periods, determined by the user. The separable model covers the whole assessment period. **It is possible to include the third, generation-dependent, factor into separable representation.**
Selection	The selection at oldest age is equal to that of previous age; selections as function of age s(a) are normalized by their sum to 1. For the plus group the same mortality as for the oldest true age. **If generation dependent factors are included**, then s(a,y)=s(a)g(cohort). s(a,y) can be normalized for each year by their sum to 1 – sub-model of “within-year effort redistribution by ages, or not – sub-model of “gain (loss) in selection”. The matrix of g-factors is normalized to give global average = 1.
Estimated parameters	
Catchabilities	The catchabilities by ages and fleets can be estimated or assumed equal to 1. Catchabilities are derived analytically as exponents of the average logarithmic residuals between the catch-derived and the survey-derived estimates of abundance.
Plus group	The plus group is not modelled, but the abundance is derived from the catch assuming the same mortality as for the oldest true age.
SSB surveys	Considered as absolute or relative. If considered as relative, coefficient of proportionality is derived analytically as exponent of the average logarithmic residuals between the catch-derived and the survey estimates of SSB.
Surveys in year (terminal + 1)	Can be taken into account (in assumption that fishing pattern in the year (terminal+1) is equal to that of terminal year)
Objective function	The objective function is a weighted sum of terms (weights may be given by user). For the catch-at-age part of the model, the respective term is: • sum of squared residuals in logarithmic catches, or • median of distribution of squared residuals in logarithmic catches MDN(M, fn), or • absolute median deviation AMD(M, fn). For SSB surveys it is sum of squared residuals between logarithms of SSB from cohort part and from

	surveys. For age- structured surveys it is SS, or MDN, or AMD for logarithms of N(a,y) or for logarithms of proportions-at-age, or for logarithms of weighted (by abundance) proportions-at-age.
Variance estimates/ uncertainty	For estimation of uncertainty parametric conditional bootstrap with respect to catch-at-age, (assuming that errors in catch-at-age data are log-normally distributed, standard deviation is estimated in basic run), combined with adding noising to indexes (assuming that errors in indexes are log-normally distributed with specified values of standard deviation) is used.
Other issues	Three error models are available for the catch-at-age part of the model: • errors attributed to the catch-at-age data. This is a strictly separable model ("effort-controlled version") • errors attributed to the separable model of fishing mortality. This is effectively a VPA but uses the separable model to arrive at terminal fishing mortalities ("catch-controlled version") • errors attributed to both ("mixed version"). For each age and year, F is calculated from the separable model and from the VPA type approach (using Pope's approximation). The final estimate is an average between the two where the weighting is decided by the user or by the squared residual in that point. Four options are available for constraining the residuals on the catches: 1. Each row-sum and column-sum of the deviations between fishing mortalities derived from the separable model and derived from the VPA-type (effort controlled) model are forced to be zero. This is called "unbiased separabilization" 2. As option 1, but applied to logarithmic catch residuals. 3. As option 1, but the deviations are weighted by the selection-at-age. 4. No constraints on column-sums or row-sums of residuals. If "triple-separable" version is used, then option 2 also produces cohort-sum equal to zero. For options 1 and 2, as well as for option 3 if not the whole age range is chosen for application of g-factors, the listed above conditions are somewhat compromised, but they are still valid for generation-independent s(a).
Program language	Visual Basic

REFERENCES

Cadigan N.G., and R.A.Myers. 2001. A comparison of gamma and lognormal maximum likelihood estimators in a sequential population analysis. Can. J. Fish. Aquat. Sci. 58: p.560–567.
Chen, Y., and D. Fournier. 1999. Impacts of atypical data on Bayesian inference and robust Bayesian approach in fisheries. Can. J. Fish. Aquat. Sci. 56: p. 525–1533.
Constable, A.J.; de la Mare, W.K. (1996). A generalised model for evaluating yield and the long-term status of fish stocks under conditions of uncertainty. CCAMLR Science 3: 31□54.
Dunn A., D.J. Gilbert, S.M. Hanchet, 2005.Further development and progress towards evaluation of an Antarctic toothfish (*DISSOSTICHUS MAWSONI*) stock model fro the Ross SeaSEA. WG-FSA-SAM-05/12. 24 p.
Hampel, F. 2002. Some thoughts about classification. Research Report No. 102. January 2002. Seminar fur Statistik Eidgenossische Technische Hochschule (ETH) CH-8092 Zurich Switzerland. 19 p.
Hampel, F. R., Ronchetti, E.M., Rousseeuw, P. J., Stahel,W. A. 1986, Robust Statistics: The Approach Based on Influence Functions, Wiley, New York
Passarin, K. 2004. Robust Bayesian estimation. UNIVERSITÀ DELL'INSUBRIA, FACOLTÀ DI ECONOMIA, 11/2004.
Vasilyev, D. 2004. Winsorization: does it help in cohort models? ICES CM2004/K:45.
Vasilyev, D. 2005. Key aspects of robust fish stock assessment. Moscow, VNIRO Publishing, 2005, 103 p.
Vasilyev, D. 2006 Change in catchability caused by year class peculiarities: how stock assessment based on separable cohort models is able to take it into account ? (Some illustrations for triple-separable case of the ISVPA model - TISVPA). ICES CM 2006/O:18. 35 p.

APPENDIX

DESCRIPTION OF THE TISVPA (version 2009.1)

D.A.Vasilyev <dvasilyev@vniro.ru>

Introduction

The TISVPA (Triple Instantaneous Separable VPA) is an extension of the ISVPA model in its version 2004.3. The extension consists in possibility to estimates within the model an additional set of generation-dependent parameters in separable representation of exploitation rates. This set of parameters is intended to adapt traditional separable representation of fishing mortality (as a product of age-dependent and year-dependent factors) to situations when several year classes may have peculiarities in their interaction with fishing fleets caused by different spatial distribution, higher attractiveness of more abundant schools to fishermen, or by other reasons.

The above mentioned generation-dependent factors (g-factors) can be estimated and applied not to the whole interval of age groups used in the model, but to some age "window". The user can choose this window by setting the first and the last age for estimation of g-factors. He also can not use them at all – in such a case the TISVPA model is reduced to the "ordinary" ISVPA model.

Two sub-models with respect to these generation-dependent peculiarities are reserved in the model:

1 - model of "within-year effort redistribution by ages"

2- model of "gain (loss) in selection"

The first sub-model assumes that in each year more fishing-attractive cohorts borrow some amount of fishing effort from other cohorts by increasing its selection at the expense of diminished selections for other age groups in this year. The second one assumes that some cohorts has increased (or reduced) selections, but this does not cause direct change in selections for others.

The same way, as in ISVPA, the TISVPA parameter estimation procedures are based on some principles of robust statistics what helps to diminish the influence of error (noise) in catch-at-age data on the results if the assessment. Special parameterization of the model makes it unnecessary to use any preliminary assumptions about the age of unit selection and about the shape of selection pattern.

The model

The Instantaneous Separable VPA (the ISVPA) group of models is designed for stock assessment when catch-at-age data are noisy; auxiliary information may be incorporated, or not used at all (if it is not available or considered as unreliable). The term "instantaneous" means that similarly to the cohort analysis introduced by Pope (1972) the catch is assumed to be taken "instantaneously", that is within a comparatively short period during the year. Approximation of instantaneous catch is absolutely correct for short fishing seasons, but it also can be regarded as being an approximate method for assessment of continuously exploited age-structured populations. In should be noted that the assumption of the constant fishing mortality coefficient during a year, that underlines conventional VPA, is also only a approximation. These two hypotheses are, in fact, the two opposite marginal simplifications in the frame of cohort models. The ISVPA acronym should not be confused with that of the Integrated Stochastic VPA by Lewy (1988).

Let us recall that Pope's Cohort Analysis is based on the observation equation (Baranov's catch equation):

$$C_{a,y} = F_{a,y}/(F_{a,y} + M)*N_{a,y}\,[1 - e^{-(F_{a,y} + M)}] \tag{1}$$

(a = 1,..., m; y = 1,...,n),

and the dynamic state equation:

$$N_{a,y} = (N_{a+1,y+1}e^{M/2} + C_{a,y})e^{M/2} \tag{2}$$

(a = 1,..., m-1; y = 1,...n-1), where a is the age index, m is the total number of age groups, y is the year index, n is the total number of years, $N_{a,y}$ is the abundance of age group a in year y,$C_{a,y}$ is the catch from age group a in year y, M is the instantaneous natural mortality coefficient (may be constant or represent a function of age). For simplicity, a=1 and y=1 are, respectively, the first age group and the first year in the available data.

Equation (1) expresses the total catch from age group a, accumulated in the y-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 $C_{a,y}$ is taken instantaneously in the middle of the year y.

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 on the middle of the year, equation (1) may be replaced by

$$C_{a,y}= \varphi_{a,y}\, N_{a,y}e^{-M/2}, \tag{3}$$

where $\varphi_{a,y}$ plays the role similar to that of $F_{a,y}$ in equation (1) but cannot be called a fishing mortality coefficient. Strictly speaking, it is the fraction of the abundance of the a-th age group, taken as catch in the middle of the year y. The model (2)-(3) can be regarded as an "instantaneous" analogue of the VPA. The word "separable" shows that the hypothesis of separability (i.e. of age selectivity of the fishery) is accepted.

In terms of the TISVPA in its traditional separable case (ISVPA) it means that

$$\varphi_{a,y} = s_a \cdot f_y \tag{4}$$

where f_y is proportional to the fishing effort (a year effect), while s_a is the selectivity of the fishery (an age effect). Further we will call them an effort factor and a selectivity factor.

If it is assumed that the assumption about selection pattern can be violated by some cohort-dependent effect, then the following representation can be used ("triple-separable" version):

$$\varphi_{a,y} = s_a \cdot f_y \cdot g_{cohort} \tag{4.1}$$

Selectivity factors in the model are normalized:

$$\sum_{a=1}^{m} s_a = 1 \tag{5}$$

If triple-separable version is used, then g-factors are normalized to give global average for the whole matrix of g-factors equal to 1:

$$\sum_{y=1}^{n-1} \sum_{a=1}^{m-1} g_{a,y} = 1/[(n-1)(m-1)]$$

(n - number of years),

where

$$g_{a(j)1,\, y(j)1} = g_{a(j)1+1,\, y(j)1+1} = g_{a(j)1+2,\, y(j)1+2} = \ldots = g_{a(j)k, y(j)k} = g_j\,;$$

$a_{(j)1}$ - index of youngest age group, and $a_{(j)k}$ - index of oldest age group in the cohort j under consideration.

If triple-separable version is used, then an additional normalization allows to get sub-models of two kinds of "physical" process of changes in selection pattern (or two sub-versions with respect to g-factors):

1 – sub-model of "within-year effort redistribution by ages". Here the following *additional* normalization is used *for each year*:

$$\sum_{a=1}^{m} s_{a,y} = s_a g_{cohort} = 1 \tag{5.1}$$

2- sub-model of "gain (loss) in selection" - this additional normalization is not used.

It is clear that in reality the fishing season does not necessarily fall on the middle of the calendar year. For the model it means that instead of factors $e^{M/2}$ и $e^{-M/2}$ the Equations (2) and (3) must contain factors $e^{\beta M}$, $e^{(1-\beta)M}$ and $e^{-\beta M}$, where β is the given constant ($0 < \beta < 1$). For simplicity in further explanations we will use β=1/2.

As can be seen, calculation of abundances in Equation (2) is undertaken directly via catch values. Catch values in this case are treated as true, the same way as in deterministic cohort models. But separabilization of the model makes it possible to look for unique values of $N_{a,y}$. . By this reason the version of the model determined by Equations (2)-(5) can be called ***catch controlled***. In this version of the model the role of separabilization consists only in estimation of terminal populations and this version may be regarded simply as a method of tuning of ordinary cohort analysis, while the loss function of the model (e.g. the sum of squared residuals between logarithms of actual and theoretical catches) may be regarded as a measure of inseparability of the catch-at-age data (in logarithmic form).

The effort-controlled version of the ISVPA, which do not treat catch-at-age data as true, is based on another dynamic state equation, resulting from substitution of the expression for theoretical catch $\hat{C}_{a,y} = s_{a,y} f_y N_{a,y} e^{-M/2}$ instead of actual catch $C_{a,y}$ into Equation (2):

$$N_{a,y} = \frac{N_{a+1,y+1}e^{M}}{1 - s_{a,y} f_y} \qquad (2')$$

(Naturally, for ordinary separable representation (ISVPA) $s_{a,y}=s_a$ for every y.)
Thus, in the abundance estimation by this version of the model it is implied that separable representation of fishing mortality is true and residuals are attributed to errors in catch-at-age data. Here the value of loss function may be regarded as a measure of “precision” of catch-at-age data (if we assume that the fishery is fairly separable).

In practice in most cases both assumptions (that catch-at-age data are precise or fishery is well separable) are rather far from reality. If there are some ideas about their relative validity it is possible to use ***mixed*** version of the ISVPA in which the equation of the stock dynamics is a mixture (with the coefficient set up by the user) of equations (2) and (2’). In this version of the ISVPA the same weight (or the “level of relative confidence”) of the two assumptions is used for all points.

Since the user often has no preliminary perception about the relative validity of the above mentioned assumptions and since the relative weight of these assumptions could be highly different for different points (a,y), the 4-th version of the ISVPA called ***mixed with weighting by points*** is also available. In this version for *every point (a,y)* equations (2) and (2’) are weighted with reciprocal squared residuals between the given catch*(a,y)* value and its respective “theoretical” value $\hat{C}_{a,y} = s_{a,y} f_y N_{a,y} e^{-M/2}$ where $N_{a,y}$ is calculated by equation (2) or (2’). These weights are recalculated at every iteration within the iterative procedure of the model parameters estimation (see below).

Equation (2) or (2’) is treated as an exact one and serves for calculation of the matrix $\|N_{y,a}\|$ through M and $\|C_{y,a}\|$ (in the catch controlled version) or M and the vectors s_a , f_y and g_{cohort} (in the effort controlled version). Equations (3)-(4), postulating the separability, or age selectivity of fishing, are regarded as approximate ones, and the unknowns M, s_a , g_{cohort} and f_y are estimated so that to reduce the residual in Equation (3) to the minimum possible (as a rule, the squared logarithmic error is meant). Equation (5) is a normalizing condition and is treated as an exact one.

Estimated values of $\varphi_{a,y}$ can be recalculated into traditional instantaneous coefficients of fishing mortality $F_{a,y}$ by the formula: $F_{a,y} = -ln(1-\varphi_{a,y})$, which becomes obvious if we rewrite the equation (2') as

$$Ln\ (N_{a,y} / N_{a+1,y+1}) = \mathrm{M} - \ln(1-\varphi_{a,y})$$

and to compare it with traditional VPA equation:

$$Ln\ (N_{a,y} / N_{a+1,y+1}) = F_{a,y} + M.$$

Algorithm of the model

The algorithm of each version of the ISVPA generally consists of a '*core*', in which all the model parameters are evaluated from the iterative procedure with the given natural mortality coefficient, M, and terminal fishing effort, f_n, and an the outward '*shell*' (a loop in which the best M and f_n are fitted).

The 'core' is represented in the program ***by 4 iterative procedures***. The three procedures which are described in details below are designed to ensure "unbiasness" of the solution, each in its own sense.

The 4-th procedure is intended to produce the best fit to catch-at-age data, but the solution will be free from any restriction on bias. The 4-th procedure is a rather time consuming derivative-free procedure, but experiments with very noisy data showed that if parameters are strongly interdependent and the minimum is flat this procedure works better (gives a better fit) compared to some tested algorithms, including Marquardt-Levenberg and Simplex ones.

*Basic iterative procedure (procedure A) (*marked as *nonlog* in the menu*)*

Within any ISVPA iterative procedure the given M and f_n are not changed. The calculations start with setting up of the initial values of the fishing effort, f_y at y=1,..., n-1 and selectivity, s_a; at a=1,..., m (the normalizing condition (5) must be followed). Each iteration consists of the following steps.

First, the terminal vectors $\{N_{a,n}\}$ and $\{N_{m,y}\}$ are evaluated by (3), then all other $N_{a,y}$ are determined from (2) or (2'). After that the matrix of fractions $\|\varphi_{a,y}\|$ is evaluated by the Equation

$$\varphi_{a,y} = \frac{C_{a,y}}{N_{a,y}} e^{M/2}, \tag{6}$$

and $\{f_y\}$ and $\{s_a\}$ are determined as

$$f_y = \sum_{a=1}^{m} \varphi_{a,y} \tag{7}$$

and

$$s_a = \frac{\sum_{y=1}^{n} (\varphi_{a,y} / g_{a,y})}{\sum_{a=1}^{m} \sum_{y=1}^{n} (\varphi_{a,y} / g_{a,y})} \tag{8}$$

To improve the convergence, s_m and s_{m-1} are replaced with their arithmetic mean:

$$s_m = s_{m-1} = \frac{\sum_{y=1}^{n} (\varphi_{m,y} / g_{m,y} + \varphi_{m-1,y} / g_{m-1,y})}{2\sum_{a=1}^{m} \sum_{y=1}^{n} (\varphi_{a,y} / g_{a,y})}. \tag{8.1}$$

Note that the selectivity values remain normalized since the initial normalization.

Generation-dependent factors g_{cohort} = are estimated as follows:

$$g_j = (g_{a(j)1,\, y(j)1} + g_{a(j)1+1, y(j)1+1} + \ldots + g_{a(j)k, y(j)k}) / k, \tag{9}$$

where $g_{a(j),\, y(j)} = \dfrac{\varphi_{a,y}}{s_a f_y}$

Strictly speaking, the symbol $\varphi_{a,y}$ is allotted to the estimate of the fraction given by formula (6) at each iteration *IT*. To avoid confusion, its separable analog, which also can be evaluated at each iteration, will be designated as $\varphi^{sp}_{a,y} = s_{a,y} \cdot f_y$.

Assume that the convergence is already achieved, and $\varphi_{a,y}$ and $\varphi^{sp}_{a,y}$ are limits of the corresponding fractions at $IT \to \infty$. When we deal with the 'pure', completely separable data, the convergence means that $\varphi_{y,a} = \varphi^{sp}_{y,a}$. However, in the general case, when the catch-at-age data do not correspond to completely separable fishing (and contain errors), the two fraction estimates, $\varphi_{a,y}$ and $\varphi^{sp}_{a,y}$ must differ. This difference could serve as a measure of non-separability in the data, thus appearing in the role of a random error, $\varepsilon_{a,y}$, in the fraction $\varphi_{a,y}$ with respect to the separable fraction $\varphi^{sp}_{a,y}$:

$$\varphi_{a,y} = s_a \cdot f_y + \varepsilon_{a,y}. \tag{10}$$

Now let us clarify the question of whether our separable estimates of φ are unbiased or not. Such an analysis requires calculation of the mathematical expectation of the random values ε. It is reasonable to regard such errors within each age group at y=1,...,n-1 as being independent and equally distributed. When this is the case, the averaging of ε within the same age group furnishes the required estimation of the bias. At $IT \to \infty$ relationships (5), (7) and (10) yield:

$$f_y = \sum_{a=1}^{m} (s_a f_y + \varepsilon_{a,y}) = f_y + \sum_{a=1}^{m} \varepsilon_{a,y}$$

or

$$\sum_{a=1}^{m} \varepsilon_{a,y} = 0, \qquad (11)$$

for each year *y*. Similarly, at $IT \to \infty$, relationships (5), (8), (10) and (11) involve:

$$s_a = \frac{\sum_{y=1}^{n} (s_a f_y + \varepsilon_{a,y})}{\sum_{a=1}^{m} \sum_{y=1}^{n} (s_a f_y + \varepsilon_{a,y})} = s_a + \frac{\sum_{y=1}^{n} \varepsilon_{a,y}}{\sum_{y=1}^{n} f_y},$$

or

$$\sum_{y=1}^{n} \varepsilon_{a,y} = 0 \qquad (12)$$

for each age group *a* (certainly, transformation (9) does not break this result). Relationships (11) and (12) prove that the separable estimates of φ supplied by this iterative procedure are unbiased. This is valid for traditional “double-separable” representation. If $s_{a,y}$ is used instead of s_a , then this condition can be somewhat compromised.

*Weighted arithmetical mean procedure (procedure B) (*marked as *nonlog w-d* in the menu*)*

When the selectivity is strongly dependent on age, the errors corresponding to different age groups hardly can be regarded as equally distributed (although, relationship (10) shows that their mean over age also equals zero). In this case, a modified iterative procedure could be appropriate, in which inverse selectivity values serve as weights at the stage of calculating the efforts.

Within this, 'weighted' iterative procedure, relationship (7) is replaced with the following equation for calculating the efforts:

$$f_y = \frac{1}{m}\sum_{a=1}^{m}\frac{\varphi_{a,y}}{s_{a,y}}\,, \tag{13}$$

(which is also an algebraic consequence of the separability hypothesis), and the efforts are calculated from (13) using the selectivity values from the previous iteration. Thereupon the current selectivity values are computed from (8) and g-factors – from (9).

Analysis of statistical sense of the solution for this procedure is similar to the previous one. At $IT \to \infty$ relationships (5), (13) and (10) result in:

$$f_y = \sum_{a=1}^{m}(s_a f_y / s_a + \varepsilon_{a,y} / s_a) = f_y + \sum_{a=1}^{m}(\varepsilon_{a,y} / s_a)$$

or

$$\sum_{a=1}^{m}(\varepsilon_{a,y} / s_a) = 0, \tag{11'}$$

for each year y. Similarly, at $IT \to \infty$, relationships (5), (8), (10) and (11') will give:

$$s_a = \frac{\sum_{y=1}^{n}(s_a f_y / s_a + \varepsilon_{a,y} / s_a)}{\sum_{a=1}^{m}\sum_{y=1}^{n}(s_a f_y / s_a + \varepsilon_{a,y} / s_a)} = s_a + \frac{\sum_{y=1}^{n}(\varepsilon_{a,y} / s_a)}{\sum_{y=1}^{n} f_y}\,,$$

or

$$\sum_{y=1}^{n}(\varepsilon_{a,y} / s_a) = 0 \tag{12'}$$

for each age group a. Relationships (11') and (12') prove that the separable estimates of φ weighted by selectivity factor, supplied by this iterative procedure are unbiased. Again, this is valid for "double-separable" procedure. If $s_{a,y}$ is used instead of s_a , then this condition can be somewhat compromised.

"Logarithmic" (geometrical mean) procedure (procedure C)

Logarithmic transformation of the relationships (3) and (4) leads to the third iterative algorithm, similar to the basic and the weighed arithmetic mean ones but dealing with logarithms of C, φ, s, f, etc. Within this, logarithmic iterative procedure relationships (6) - (8), that are used at the IT-s iteration, must be replaced with:

$$\ln \varphi_{a,y} = \frac{M}{2}\ln\frac{C_{a,y}}{N_{a,y}}, \tag{14}$$

$$\ln f_y = \frac{1}{m}\sum_{a=1}^{m} \ln\left(\frac{\varphi_{a,y}}{s_a g_j}\right), \tag{15}$$

$$\ln s_a = \frac{1}{n}\sum_{y=1}^{n} \ln\left(\frac{\varphi_{a,y}}{f_y g_j}\right), \tag{16}$$

$$\ln s_m = \ln s_{m-1} = \frac{1}{2n}\sum_{y=1}^{n} \left(\ln\frac{\varphi_{m,y}}{g_j f_y} + \ln\frac{\varphi_{m-1,y}}{g_j f_y}\right)$$

and

$$\ln g_j = (\ln g_{a(j)1,\, y(j)1} + \ln g_{a(j)1+1, y(j)1+1} + \ldots + \ln g_{a(j)k, y(j)k}) / k\,, \tag{17}$$

where $g_{a(j),\, y(j)} = \dfrac{\varphi_{a,y}}{s_a f_y}$

It is necessary to mention that in this and in all other procedures for “short” generations (less than 2 points in the catch-at-age data matrix), values of g_j are not recalculated within iterations and remains the same (their discrepancy from initial guess (equal to 1) is due only to normalization.

When evaluating f_y by (15), selection-at-age and g-factors are taken from the previous iteration. At the end of each iteration, selectivities must be re-normalized so that to satisfy condition (5). This procedure can also be called "weighed geometrical mean procedure", as from (15) and (16) it immediately follows that f_y, s_a and g-factors equal to the geometrical means of $\varphi_{a,y}$ weighed by s_a or f_y or generation factor respectively.

It is easy to show, that this iterative procedure stops when “estimated” logarithmic catches are unbiased (residuals have zero mean) simultaneously within years, age groups **and generations** (this will be illustrated below). In order to understand the statistical meaning of the convergence of the procedure, it is convenient to use the notion of estimated catch,

$\hat{C}_{a,y} = s_a f_y g_j N_{a,y} e^{-M/2}$, and present $\varphi_{y,a}$ in the form:

$$\varphi_{a,y} = g_j s_a f_y \frac{C_{a,y}}{\hat{C}_{a,y}} \tag{18}$$

Let us consider the limits at $IT \to \infty$ of all the variables participating in the model. Therefore the fractions $\varphi_{a,y}$, which is determined by equation (4.25)-(4.28), can be replaced with that given by relationship (4.29), where $\hat{C}_{a,y}$ is substituted by $\hat{C}^*_{a,y}$, the catch estimates supplied by the iterative procedure at at $IT \to \infty$. This substitution implies:

$$\sum_{a=1}^{m}[\ln C_{a,y} - \ln \hat{C}^{*}_{a,y}] = 0 \quad , \tag{19}$$

$$\sum_{y=1}^{n}[\ln C_{a,y} - \ln C^{'*}_{a,y}] = 0 \ , \tag{20}$$

$$\sum_{j=1}^{k}[\ln C_{a,y}(j) - \ln C^{'*}_{a,y}(j)] = 0 , \tag{21}$$

Equation (21) is valid only for generations, participating in evaluation of g_j (see above). The meaning of (19)-(21) is that the log-transformed estimates of catches are unbiased for each age group, each year, and each generation. If not all available age groups, but some age window is chosen for application of g-factors, then equation (21) may be compromised, while (19) and (20) remain valid.

Loss functions

In accordance with the assumptions about the error structure in the data the solution of the model can be based on the standard minimization of sum of squared residuals or on the minimization of some more robust loss functions: the median of distribution of squared residuals or the absolute median deviation of residuals.

Minimization of the median, *MDN*, of squared residuals (that is, the use of the least median or the LMSQ principle) instead of their sum (the classical LSQ-principle) sometimes is referred to be more resistant to outliers, i.e. those elements of the data set which go far beyond the reasonable confidence limits and, hence, are suspicious of containing extremely high errors (O'Brien, 1997; Hampel et al., 1986).

According to this concept, an alternative ISVPA solution can be seeked as providing estimates of M and f_n, which secure the minimum of the median of the distribution of the squared logarithmic residuals,

$$SE_{a,y} = (\ln C_{a,y} - \ln \hat{C}^{*}_{a,y})^2$$

(a = 1,...,m; y=1,...,n). The corresponding loss function will be denoted as $MDN(M, f_n)$.

In practice, the median of a random series is estimated by rearranging its elements in a descending or increasing order and taking the central element of the new series or the mean of two central elements (depending on whether the total number of the elements is odd or even). However, when used within the framework of ISVPA, this estimate sometimes may cause a certain roughness of the surface $MDN(M, f_n)$. In order to make the loss function smoother, the median is estimated here as the mean of a number (for example, 10) of central elements of the ordered series of $SE_{a,y}$. So, in this version of ISVPA, the iterative procedures for estimating the vectors *f* and *s* remain the same as described above, the only difference being the use of the behavior of the median as an indicator of their convergence. Numerical experiments ascertain workability of the three versions of the ISVPA iterative procedures combined with the LMSQ principle.

As it was already noted, in order to smooth the median estimates, averaging over a number of central elements of the ordered series of squared residuals is suggested. Certainly, the number of central elements can vary from one or two to $m \cdot n$, the total length of the series. However, in the latter case, the averaging results in estimation of the mathematical expectation rather than the true median of the squared residuals. Thus, in fact, the suggested approach (when averaging over a number of central squared residuals is applied) can actually be regarded as a compromise between the true median minimization and the conventional least squares criterion. The advantage of this compromise is that, according to our experience, the use of the least squares approach leads to a sufficiently smooth loss function, while the minima of $MDN(M, f_n)$ are better pronounced.

One of the central issues in fitting a model to real data is the choice of the fitting criterion. Statistically, the use of the LSQ criterion is equivalent to accepting the hypothesis of normality of the distribution of the residuals (in the case when the sum of squared logarithmic residuals is minimized, the errors themselves are supposed to be logarithmically normal). What is the reason for using the median minimization approach? What kind of iterative procedure matches well the LMSQ criterion? To illuminate the nature of combining the LMSQ criterion with the ISVPA, let us consider the third, weighed logarithmic version of the iterative procedure.

It has been shown above that the logarithmically transformed theoretical estimates of catches are unbiased. Strictly speaking, it means only that the mathematical expectation of the corresponding residuals is zero. We, however, believe that in practice, the distributions of the logarithmic residuals are often almost symmetric. This is confirmed by our numerous computer tests with both simulated and real data. Clearly, if a random value ε is distributed symmetrically the median of its squares, ε^2, indicates the compactness of the distribution of ε: the higher the median of ε^2, the greater the variance of ε. Conversely, the lower the median of the distribution of ε^2, the more compact is the distribution of ε. Thus, by minimizing the median of the squared logarithms of the catches residuals resulting from estimation of catches by means of the weighed logarithmic iterative procedure, the maximal allowable compactness of the distribution of the errors themselves is reached, consequently, providing a reasonable fit of the model to the catch-at-age data in the sense of the conventional maximum likelihood concept.

Such a statistical justification cannot be given to the median minimization approach when the first (**A**) or the second (**B**) version of the TISVPA iterative procedure is used, as neither of them impose any reasonable condition on the errors in the logarithmically transformed catches. From this point of view for these versions the conventional least squares approach seems to be more appropriate.

On the other hand, the approach when the quality of fitting is measured by some “window” in the distribution of residuals which does not include the tails of the distribution, could be considered a means to suppress the influence of outliers on the solution (because the residuals corresponding to outliers are located near the margins of the distribution and will not affect the value of the median). From this point of view minimization of the median seems to be appropriate for procedures **A** and **B** also.

In addition to the two above mentioned TISVPA objective functions, the absolute median deviation $AMD(M, f_n)$, i.e. the median of the absolute deviations of model residuals from their median value, known as one of the most robust measures of scale (Huber, 1981), also may be used. According to my experience in some cases (for example, when distribution of residuals, still having zero mean, has nonzero median) *AMD* gives more pronounced minimum with respect to *MDN(SE)* - minimization. However, if the data are not informative (for example, if historical changes in catches and in stock are not pronounced) the *AMD* may be not sufficiently sensitive and it may be better to use the *MDN*.

Now let us say a few word about the procedure of estimation of the “best” (in the sense of the loss function chosen) values of (f_n, M). The choice of the procedure in the ISVPA is based on the following considerations:

- algorithmic simplicity, taking into account that in the *outer loop* only two (or even one, if *M* is considered as known) parameters are to be estimated;
- if the loss function surface has more than one minimum - to give possibility to start minimization in the vicinity of the required minimum and to arrive at it even if the surface is very flat (this implies that gradient methods may be ineffective).

Numerous simulation experiments have indicated that the method, which is not the fastest, but which allows us to reach precisely the minimum even if the error surface is very flat and the minimum is local, is the method of “lowering by coordinates” with successively diminishing steps. The step of the procedure (i.e. the increase in the tested parameter value) is fixed by the program and after the minimum is reached with this step, the latter is reduced by a factor of 10. When the minimum is attained again the step value is reduced by the same factor, and so on, till the minimum of the tested parameter value is reached with the required precision.

It is necessary to mention that while minimization of the sum of squared errors multiple minima are almost never encountered (here the problem is that for noisy data minimum of *SSE* is often reached at a marginally high or low value of the tasted parameter), for the median minimization the surface of the loss function (as a function of f_n and M) may have complex structure. That is why before the final run with precise estimation of the model parameters it is recommended to make preliminary point-by-point scanning of the (f_n ; M) area with sufficiently small step (e.g. 0.1 for f_n and 0.01 for M). The TISVPA program realization gives such a possibility.

Suppression of inter-iteration oscillations

When the level of noise in the initial data is high, the estimated effort and selectivity, as well as the sum of squared residuals, *SSE*, vs. the number of iteration, *IT*, contain a few explicit slowly decaying modes of oscillations superimposed on a certain rapidly stabilizing trends. These oscillations slow down the convergence of the *SSE* to its limit, *SSE**, or of the *MDN* to *MDN**, or the *AMD* to *AMD** at $IT \to \infty$, thus becoming significant at the stage of searching for the minimum of $SSE^*(M, f_n)$ or $MDN^*(M, f_n)$, as in practice, at every M and f_n the iterative process is stopped at a finite *IT*. The most notable in this context are the saw-tooth type oscillations with a 2-year periodicity, i.e., those with the highest frequency. Conventional method for filtering oscillations and extraction of trends from numerical series is a moving averaging. We, however, are dealing with an iterative process, where at any iteration *IT*, the current selectivity, $s_{IT}(a)$, or the effort, f_{IT}, estimate is calculated after the previous value, $s_{IT\text{-}1}(a)$ or $f_{IT-1}(y)$, was found. That is why, by defining the corrected selectivity and effort estimates at *IT*-th iteration, $s'_{IT}(a)$ and $f'_{IT}(y)$, as

$$s'_{IT}(a) = \alpha s_{IT-1}(a) + (1-\alpha) s_{IT}(a) \tag{22}$$

$$f'_{IT}(y) = \alpha f_{IT-1}(y) + (1-\alpha)\, f_{IT}(y) \tag{23}$$

and by a proper choice of the coefficient $0<\alpha<1$, the desired filtration, similar to the moving averaging, can be achieved. According to (22), all the selectivity estimates, which were computed at the previous iterations, participate in the correction for the current, *IT*-th iteration. The same is valid for the effort (see (23)). So, the size of the averaging interval in this filtration procedure increases with the growth of *IT*. Nevertheless, as the weights of the last, *IT*-th, iterations remain constant, while the weights of the early iterations decay, the suggested filtering procedure can be regarded as an analog of a conventional moving averaging. The effective averaging interval is determined by the choice of α: the smaller α, the narrower the effective averaging interval. Experiments showed that the choice of α do es not affect the result: they are almost identical for tested range of α from 0 to 0.95.

Treatment of zero catches.

Existence of zero values in the catch-at-age matrix is known to be a rather complicated (and may be logically controversial when dealing with the logarithmic residuals) problem which is solved differently in different methods. In the ISVPA the following algorithm is applied:

1. If $C_{a,y}$=0, then the value of $\varphi_{a,y}$ is taken equal to its "theoretical" value, that is

$\varphi_{a,y}= s_{a,y} f_y$.

2. Residuals for points of zero catches are taken equal zero.

3. Stock abundance is computed as follows:

3.1. If $N_{a+1,y+1}$ >0 and $C_{a,y}$=0, than $N_{a,y}$ is computed by (2).

3.2. If $N_{a+1,y+1}$ =0 and $C_{a,y}$=0, than $N_{a,y}$ = 0.

3.3. If $N_{a+1,y+1}$ =0 and $C_{a,y}$>0, than $N_{a,y}$ is computed by (3), similar to the terminal points.

3.4. If $N_{a+1,y+1}$ >0 and $C_{a,y}$>0, than $N_{a,y}$ is computed by equation (2) or (2') or their mixture, according to the version chosen.

Estimation of ISVPA parameters without limitation on bias

To test experimentally the role of limitation on bias, imposed by the above described ISVPA procedures, an additional, free of such limitations parameter estimation procedure was developed.

For "direct" fitting of multi-parameter models the Marquardt-Levenberg and Gauss-Newton method are traditionally used (Bard, 1974), as it was done, for example, in the CAGEAN (Deriso et al., 1985) and the ICA (Patterson, 1994). But in our case the use of these methods is complicated by normalization equation (2.5): parameters are becoming inter-dependent. Attempt to use the Simplex-method (Schnute, 1982) was also unsuccessful: for the case of many parameters the procedure is very time-consuming and also requires very qualified initial guess for parameters (the result is extremely sensitive to its choice).

Therefore, the procedure of "direct" search for the ISVPA parameters free of limitations on bias was finally arranged as follows. The same was, as it was done with "iterative" inner ISVPA procedures, the procedure was designed as two concentric loops. In outer loop optimization by (f_n , M) is made, while the parameters {s_a }and {f_y} (except of f_n) are estimated in the inner loop.

The inner loop is arranged as follows. Each parameter is optimized in succession, while the order of optimization appeared to be important. Starting with a set of initial guesses for all parameters $s_1,...,s_m$ и $f_1,...,f_{n-1}$, optimization begins from f_{n-1} ; after that the value of f_{n-2} is optimized, and so on till f_1 . Further, the best value (from point of view of the loss function) of s_1 is estimated, the other values of s_a being changed by means of normalization equation (5). The found value of s_1 is then "frozen up" and the "best" value of s_2 is searched for (here the normalization equation (5) is applied to the rest of selectivity factors: $s_3,...,s_m$,). Then the next, s_3 , selectivity factor is estimated, and so on till s_{m-2}. . The rest of selectivities, $s_m=s_{m-1}$, appears to be already estimated by the normalization equation. After that the procedure returns back to the estimation of f_1 , and the same sequence of calculations is repeated till convergence.

The above described procedure gives the solution free from restrictions on bias. For "clean" catch-at-age data (simulated data without noise) the procedure gives absolutely correct estimates of all parameters (as well as "iterative" procedures **A**, **B** and **C**). For noisy simulated data and for real data the solution based on this "unrestricted" fitting procedure as a rule is much worse, while the final value of the loss function may be lower than for "unbiased" solutions.

It should be noted that implementation of the above described procedure of "parameter-by-parameter" optimization for the *median* minimization could be problematic if one (or a group) of parameters $s_1,...,s_m$ and $f_1,...,f_{n-1}$ occasionally influences only those values of residuals which are located in tails of the distribution of residuals and, hence, do not influence the median value.

Dealing with auxiliary information

There is possibility to include up to three SSB indices and up to seven age structured stock abundance indices into the model. In such a case, the ISVPA loss function will include additional components representing measures of discrepancy:

- for each SSB index : between logarithms of the SSB from the cohort part of the model and from surveys;

-for each age-structured index: between logarithms of abundance (a,y) from the cohort part of the model and from surveys (whether corrected to the estimated age-dependent "fleet catchabilities" or not).

The model fitting could be done not only with the survey abundance-at-age data, but also with the survey age proportions and "weighted" survey age proportions (see below).

Thus, for each age-structured index the discrepancy may be measured as the traditional sum of squared residuals, or by the MDN, or the AMD. The measure can be stated independently for each of "fleet".

For the SSB indices the only available measure in the model is the sum of squared residuals (because, as a rule, available number of years of the SSB surveys is rather low).

The program

Current realization of the ISVPA is made in Visual Basic and can be run within any Windows environment. If Visual Basic is installed on your computer it will be enough to copy only executable file. If not, you should use the TISVPA set-up package.

Input files are blank-separated text files, including:

- "necessary" files: catch-at-age by years, weight-at-age by years in the stock and maturity-at-age by years;

-"optional" files (not obligatory): natural mortality by ages, up to three files with the SSB estimates by years and up to seven files with age-structured abundance indices by years.

All input files must be copied to the **C:\vbisvpa** directory or its subdirectories.

Output files include: the file with records of minimization (minim.out), the file with results (its name is given by the user) of initial ("basic") run, as well as bootstrap output files:

1) bootf.out - includes the effort factor estimates by years and bootstrap runs;

2) bootm.out - includes the natural mortality estimates by ages and bootstrap runs (if it were regarded as an unknown parameter);

3) boots1.out and boots2.out - include the estimates of selectivities (for the first and the second time intervals) by ages and bootstrap runs (the program permits fitting of two selectivity patterns for two different successive time intervals);

4) bootssb.out - includes the SSB estimates by years and runs;

5) boottsb.out - includes the estimates of total stock biomass by years and runs;

6) bootntrm.out - includes the terminal year abundance estimates by ages and runs.

The procedure of working with the program is the following.

Primary choice is to use "triple" or ordinary separability assumption. If triple version is chosen – choose the first and the last ages for estimation of g-factors and sub-model (first or second). If ordinary ("double") separabilization is chosen, the model will be reduced to ISVPA (in version 2004.3)

1. The first thigh to do while running the program is to enter the names of catch-at-age and weight-at-age files. If they are located directly in the C:\vbisvpa directory one should simply print their names (with extension). If they are stored in some sub-directory of C:\vbisvpa one should print the name of this subdirectory prior to the name of the file.

2. After that one will be asked about the **situation with natural mortality**: 1) to find *M* as an age-independent value; 2) to find it as a simple quadratic function of age; or 3) to use known values of *M(a)*. If you choose option 2, you will be asked to enter the age of the minimum *M* (as a rule it can be taken equal to the age of 'mass' maturity). If option 3 is chosen, you will be asked to enter the name of the file with known *M(a)* values.

3. Next you will have to choose the method of the parameter estimation. There are four options available. Option 1 will produce solution with “unbiased separabilization”; option 3 will lead to “unbiased weighted separabilization”; option 2 will ensure “unbiased” estimates of logarithms of all parameters; option 4 will produce solution corresponding to the best fit to logarithmic catches, not restricted by any condition on bias. While using option 4 one should be patient as it is time-consuming. In most cases option 1 or 2 is recommended. It is strongly recommended **not to use option 4** when you minimize the median as the error surface can be too "broken".

4. The next choice is what to minimize. It is possible to minimize the sum of squared residuals in the logarithmic catches, or the median of distribution of squared residuals in the logarithmic catches $MDN(M, f_n)$, or the absolute median deviation $AMD(M, f_n)$. For noisy data it is recommended to minimize the *MDN* or *AMD*.

5. Selection of the first and the last year of analysis and the last year of first selectivity pattern (**the program gives possibility to fit two different selectivity patterns for two different successive time intervals**). After that it is required to input the first and the last age groups. Naturally, they should be within the range of the input data. After that you will be asked whether the oldest age in the data is a “normal” age group, or a +-group?

6. Next question is about the “version” of the program (1. Catch-controlled, 2. Effort-controlled, 3. Mixed, 4. Mixed, weighted by points). Version 1 is preferable if fishery is known to be extremely non-separable. It also can be useful as a part of “mixed” versions 3 and 4. Version 2 is preferable if M is considered as an unknown parameter and/or the data are very noisy.

7. If version 3 is chosen you should input relative weight of the catch-controlled routine.

8. You could (1) scan the error surface or (2) look for a precise solution. If scanning is chosen, you will be asked about minimum and maximum values of the parameter (f_{term} or (M and f_{term})) and of the "step". It is recommended to make scanning first as there could be several local minima of the loss function. Option 2 allows to find a precise solution. If there are several local minima, you could look for a solution corresponding to the required minimum making a proper choice of an initial guess about the parameter and a sufficiently small initial step. Please note that if the “scan” mode was chosen, the output file will contain the result at the **rough** minimum of the loss function. To get the result at the precise minimum you have to start the program again and to choose the option called “precise solution”. If "precise solution” is looked for, you should input the value of initial guess for f_{term} or (M and f_{term}) as well as the value of the initial step in the searching procedure.

9. Next you will have to set the value of the “inter-iteration smoother”. In most cases any value within 0.5-0.9 will be OK. In case of very noisy data, to suppress possible oscillations you

could take a higher value - up to 0.9. Don’t worry about the “precise” value of this parameter: if the procedure converges - it is OK. Experiments proved that the final result will be the same even at 0.95.

10. If you have chosen the median minimization, you should input the number of central elements of the ordered series of squared residuals (or residuals) to use as its measure. In most cases 10 points is OK. If the error surface contains too many local minima it could be useful to increase the number of central elements; if minimum is too flat - you may diminish the number of central elements. It is noteworthy that this setting will be used for the MDN or the AMD measures everywhere (for indices also, if one of these measures will be used for some of them).

11. Enter the part of the year for the peak of catches (since the model is based on Pope’s approximation of “instantaneous” catch). If the fishing is uniform all over the year - enter the traditional value of 0.5.

12. Enter the name of the output file. It will be in C:\vbisvpa directory.

13. You can display the currents results on the screen. This will slow down the calculations, however, you would be able to watch the process.

14. Input the maturity-at-age file name.

15. You will be asked whether to include SSB surveys or not. If you want to do it, you will have to input names of the SSB survey files by years (up to 3).

16. If you have age-structured abundance indices, you can use up to seven different indices. If you want to include these indices, input their names.

17. If any auxiliary information is used, you will be asked to input weight for the catch-at-age- derived component in the overall loss function (any value is possible, including 0).

18. If SSB surveys are included: for each of them input weights for components of the overall loss function which represents the measures of their closeness to the cohort part -derived estimates of the SSB (for the SSB indexes only one sort of measure is available - the sum of squared residuals between their logarithmic values).

19. Input part from the beginning of the year till the period when the surveys have been made (the same should be done for all SSB indices).

20. If SSB surveys are included: for each of them, input values of the standard deviation from the lognormal distribution which will be used in the stochastic runs.

21. If SSB surveys are included: state whether to treat each of them as absolute or relative indices.

22. If age-structured indices are included, input part from the beginning of the year till the period when the age-structured survey has been made (for each kind of survey).

23. If age-structured indices are included, state the type of the index (e.g. the mature fish, the whole stock, or the immature fish).

24. If age-structured indices are included: for each of them, input weights for the components of the overall loss function which represent the measure of their closeness to the cohort part, derived estimates of abundance.

25. If age-structured indexes are included: for each of them answer whether: to estimate age-dependent catchabilities or not (if you choose not to do it it will be assumed that $q(a)$=1).

26. If age-structured indexes are included - choose for each of them what measure of closeness of fit will be used: the MDN, SSE, or AMD.

27. If age-structured indices are included – for each of them, choose the terms you want to compare at tuning : (1) logarithmic abundances(a,y) from the modeled cohort part or logarithmic abundances(a,y) from the survey; 2) a logarithmic abundance (a,y) (from the cohort part of the model) or a logarithmic age structure(a,y) from the surveys; 3) a logarithmic age structure of the stock (a,y) (from the cohort part) and a logarithmic age structure(a,y) from the surveys.

28. If age-structured indexes are included: for each of them, enter the values of the standard deviation of the lognormal distribution which will be used in stochastic runs.

29. When calculations are finished, you can make stochastic runs. Current version of the program gives possibility to run parametric conditional bootstrap with respect to catch-at-age, (assuming that errors in catch-at-age data are log-normally distributed, standard deviation is estimated in basic run), combined with adding noise to indexes (assuming that errors in indexes are log-normally distributed with specified values of standard deviation).

If something goes wrong or in an undesirable direction, it is always possible to **stop the program** by clicking the button “stop”. The program will return to the initial (input) screen and you can run it again. The only what is necessary to remember when using “stop by user” is that if the “direct search” option for inner parameters is used, you have to let the program to finish at least one inner cycle (that is to finish calculation of inner parameters for at least one f_{term}) and to stop it after that (otherwise interrupt will cause error and abortion of the program).

The current version of the program allows one to use surveys for the (terminal+1) year (that is for year without known catch-at-age data). Fishing pattern in this year is assumed equal to that of the “true” terminal year. In such a case all input files should be entailed to include data for this year which becomes terminal; the catch-at-age file should include zero values of the catch-at-age for this year.

References

1. Anon. 1999. Reports of the Working Groups on the Assessment of Mackerel, Horse Mackerel,Sardine and Anchovy (1999). ICES CM 2000/ACFM.
2. Bard Y. 1974. Non-linear parameter estimation. - NY, Academic Press, 1974. - 349 p.
3. Deriso R.B., T.J. Quinn II and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci., 42, N 4: 815-824.
4. Doubleday W.G. 1976. A least squares approach to analysing catch at age data. ICNAF Res. Bull,12: 69-81.
5. Fournier D. and Archibald C.P. 1982. A General Theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci.,39: 1195-1207.
6. Fournier , D. A, J. Hampton, and J.R. Sibert. 1998. MULTI FAN-CL: a length- based, age structured model for fisheries stock assessment, with applications to South Pacific albacore. Can. J. Fish. Aquat. Sc i. 55: 1-12.
7. Gudmundsoon, G. 1986. Statistical considerations in the analysis of catch-at-age observations. J. Cons. int. Explor. Mer, 43: 83-90.
8. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A.. 1986. Robust Statistics. The approach based on Influence Function. John Wiley & Sons, NY.
9. Hilborn, R., et al. 2000. COLERANE. A generalized age structured stock assessment model. Users manual. Version 1.
10. Huber, P. J. 1981. Robust statistics. John Wiley & Sons, NY.
11. Kimura, D.K. 1986. Variability, tuning and simulation for the Doubleday-Deriso Catch-at-Age model. Can. J. Fish. Aquat. Sci. 46: 941-949.
12. Kizner Z.I. and D.Vasilyev. 1993. Instantaneous Separable VPA (ISVPA) with determination of natural mortality coefficient. ICES Statutory Meeting (1993) C.M. 1993/D:2. 45p.
13. Kizner Z.I. and D.A.Vasilyev. 1997. Instantaneous Separable VPA (ISVPA). ICES Journal of Marine Science, 54 , N 3: 399-411.
14. . Lewy P. 1988. Integrated stochastic virtual population analysis: estimates and their precision of fishing mortalities and stock sizes for the North Sea whiting stock. J. Cons. int. Explor. Mer. 1988. Vol. 44. P. 217-228.
15. Patterson K.R. 1995. Technical reference for the Integrated Catch-at-Age Programmes, Version 1.2 . SOAFD Marine Laboratory. Aberdeen.
16. Pope J.G. 1972. An investigation of the accuracy of virtual population analysis. Int. Commn. Northwest Atl. Fish. Res. Bull. 1972. N9. P. 65-74.
17. Pope J.G. 1974. A possible alternative method to virtual population analysis for the calculation of fishing mortality from catch at age data. ICNAF Res. Doc. 74/20. 16 pp.
18. Pope J.G., Shepherd J.G. 1982. A simple method for consistent interpretation of catch-at-age data. Cons. int.Explor.Mer.,40: 146-184.
19. Vasilyev D.A. 1998. Separable Methods of Catch-at-age Analysis From Point of View of Precautionary Approach. ICES Study Group on the Precautionary Approach to Fishery Management (Copenhagen, 3-6 February 1998). Working Paper N 11. 7 pp.
20. Vasilyev D. 1998a. Separable cohort procedures with internal property of unbiasness of the solution. ICES C.M. 1998 / BB:3, 11 pp.
21. Vasilyev D. 2001. Cohort models and analysis of commercial bioresources at informational supply deficit. Moscow, VNIRO Publishing, 2001. 98 pp.
22. Vasilyev D. 2003. Is it possible to diminish the impact of unaccounted time trends in age structured surveys' catchability on the results of stock assessment by means of separable cohort models ? ICES CM 2003/X:03. 13 pp.
23. **Vasilyev, D. 2005. Key aspects of robust fish stock assessment. M: VNIRO Publishing, 2005. 105 p.**
24. **Vasilyev, D. 2006. Change in catchability caused by year class peculiarities: how stock assessment based on separable cohort models is able to take it into account? (Some illustrations for triple-separable case of the ISVPA model - TISVPA). ICES CM 2006/O:18. 35 p.**