NORTH-EAST ARCTIC COD STOCK ASSESSMENT BY MEANS OF THE TISVPA MODEL

D.A.Vasilyev, T.I.Bulgakova

INTRODUCTION

North-East Arctic cod stock assessment in frames of ICES is traditionally carried out by means of the XSA model (Shepherd, 1991). From the moment of its appearance in the arsenal of the ICES stock assessment methods, the XSA model became very popular because it gave possibility to input into the stock assessment procedure a set of available abundance indices simultaneously and to carry out a detailed diagnostics of the solution.

On the other hand, experience of implementation of the XSA model has revealed its substantial shortcomings. First, precise statistical meaning of the model solution is unknown, because the solution is attained not by direct maximization of the declared objective function, but as a result of convergence of a specific iterative procedure. This iterative procedure may not converge in some cases. Even if it converges to some solution, this solution can be far from true maximum (minimum) of the declared objective function. An additional statistical ambiguity in the solution comes from the implementation of the so called shrinkage procedures.

Second, the solution is totally defined by trends in the auxiliary information (surveys, catches per unit effort, etc.), while the signal from one of the most representative data source, catch-at-age, is not used directly (XSA is based on simple (not separable) cohort model).

Third, the XSA model does not deal with robustness and principles of robust statistics, what is important in dealing with practical (essentially nosy) data.

Let us mention also some other imperfections of the XSA model: it does not give possibility to use directly the available "integral" stock indices, such as total or spawning stock biomass estimates; it is impossible to get rid of the influence of year-to-year variations in survey conditions, for example, by tuning not on abundance indices, but on age proportions, as it can be done in some more up-to-date models.

Understanding of the above mentioned in the last decade has resulted in more rare use of the XSA model for serious stock assessment in the practice of ICES working groups, and in increasing implementation of more advanced (more often – separable) models, which are able to take into account all the available information about the stock, including the information about its current size directly contained in catch-at-age data. These models are well known in the ICES and include such models as ICA, AMCI, SAD, SMS, etc. The ISVPA model, including its extended version TISVPA, may be attributed to this group. The main characteristic feature of this model consists in intentional implementation of principles of robust statistics in order to diminish influence of errors in the data on the results of the assessment (Vasilyev, 2005).

THE MODEL AND THE DATA.

In most of contemporary stock assessment models the so called separable representation of fishing mortality coefficients is used, at least - to some extent, that is fishing mortality coefficients can be represented as a product of two vectors: one, depending on age, and second – depending on the year. Such a parametrization helps to diminish the number of unknowns and to extract the information about current stock size directly from catch-at-age. But often it is believed that separable models are excessively structurally restrictive to ascribe possible irregularities in interactions between specific cohorts and fishery. This question is considered in the version of the ISVPA model named TISVPA (Triple Instantaneous Separable VPA), which is used here. 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 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 generations 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. Besides, such an approach helps to diminish to some extent the influence of some systematic errors in the data.

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 traditional 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 have increased (or reduced) selections, but this does not cause direct change in selections for others. In our assessment the first, the "within-year effort redistribution by ages" – sub-model was used.

An important feature of the TISVPA model consists in implementation of principles of robust statistics in procedures of estimation of 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 model options include: 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.

Brief description of the model is summarized in the table below.

Model	**TISVPA**
Version	**2006.1**
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

In our assessment we used the same data, as the Arctic Fisheries Working Group used in the North-East Arctic cod stock assessment by means of XSA (ICES, 2007): catch-at-age, mean weight-at-age, and proportions of mature fishes by ages and by years. The value of instantaneous natural mortality coefficient was taken equal to 0.2 for all ages (cannibalism was taken into account by means of incorporation of its numerical estimates into catch-at-age matrix). Catch-at-age data also included *Russian* estimates of unreported catches. The same way, as in the assessment by means of XSA, 4 age-structured stock abundance indices were used: Russian trawl *cpue* ("fleet 1"); joint bottom trawl surveys ("fleet 2"); joint acoustic surveys (Barents Sea and Lofoten) – "fleet 3", and Russian bottom trawl surveys ("fleet 4"). In our assessment all available data were taken, while in XSA runs only data from 1997 were used in tuning.

In the TISVPA model the generation-dependent factors (g-factors) were estimated for age groups from 3 till 11, because in younger ages the impact of cannibalism-derived "artificial" catches is high, and in older age groups including the plus-group the data are more noisy, what may make the results less stable.

The version of the model used for assessment assumed the equal probabilities of existence of errors in catch-at-age and in separable representation of fishing mortality (the so-called "mixed" version). An additional restriction on the solution was the guaranteed unbiased model description of log-transformed catch-at-age data. For catch-at-age data a clear minimum of the respective component of the model loss function was found already by traditional minimization of sum of squared residuals in logarithmic catch-at-age (see figure 1). For fleet 1 the median of distribution of squared residuals in abundance-at-age was minimized. For the fleets 2, 3 and 4 the only option which gave minima of the respective components of the model loss function (that is which allowed to avoid the deterioration of the solution caused by substantial noise) was found to be the minimization of sums of squared residuals in logarithmic age proportions, weighted by estimated abundances (on figures this measure of closeness of fit is presented as Pw(a,y)). This may be attributed to substantial year-to-year changes in survey conditions, which cause differences in effective survey catchability and, hence, in absolute abundance estimates from surveys, while the data on relative age composition of the stock could remain more reliable.

RESULTS

As it can be seen form figure 1, all the sources of information (catch-at-age and 4 kinds of auxiliary data) contain mutually non-contradicting information about the stock size (positions of respective minima are very close to each other). It supports the choice of the model options.

Figure 2 represents the residuals for all the data used in the assessment. As it can be seen, the residuals for catch-at-age contain no apparent cohort pattern, that is incorporation of generation-dependent factors into the model helped to take into account possible peculiarities in interaction of different generation with the fishery. The estimates of generation-dependent factors (g-factors) are presented on figure 3. Implementation of generation-dependent factors in the separable model of fishing mortalities allows substitute uniform for all years selection pattern by probably more realistic adaptive selection matrix (figure 4).

The estimated by means of TISVPA values of cod fishing stock (includes fish of age 3 and older) are shown on figure 5. Figure 6 compares the results of TISVPA to the results of the XSA model (ICES 2007). The estimated trends in spawning stock biomass (SSB), fishing stock biomass (B3+) and in fishing mortality estimated by the two models are generally similar, except for 2005 and 2006. The XSA model shows some decline of the stock, beginning from 2004, while the TISVPA-derived estimates show continuation of its tendency to increase with the SSB estimate for 2006 approaching historically highest value of 1992 on the time interval considered.

Both models show decrease in fishing mortality in 2006, while in the results of the TISVPA model this decrease is more pronounced.

Retrospective analysis shows sufficient stability of the results: figure 7 represents the TISVPA-derived estimates of SSB, recruits at age 1, and the fishing mortality estimates when the data for 2006, 2005, 2004 and 2003 were successively taken away from the consideration.

Analysis of uncertainty made by means of conditional parametric bootstrap shows sufficient reliability of the results with the trends in the bootstrap-median values of the stock biomass being similar to trends in the key run of the model. Figure 8 shows the bootstrap-derived estimates of uncertainty in the estimates of SSB, in age-dependent selection pattern, as well as for stock abundance by ages in 2006. Let us mention that the estimates of age-dependent selection pattern have rather narrow confidence intervals and are rather stable, what can be considered as a support for possibility of application of separable models for assessment of this stock.

CONCLUSION

The results of the our assessment show that the attention to peculiarities of interactions of generations with the fishery, as well as the attention to robustness of analysis, pertinent to the TISVPA model, results in much higher estimates of North-East Arctic cod stock spawning stock biomass for recent years in comparison to the results of the Arctic Fisheries Working Group (AFWG), obtained by means of the XSA model using the same sources of data (878 000 tonnes against 535 000 tonnes for 2006). The estimate of total stock is 1.6 times higher: 2 072 000 tonnes against 1 298 000 tonnes.).

Short term forecast shows that for reference fishing mortality value equal to $F_{5\text{-}10} = F_{pa} = 0.4$ the recommended TAC value in 2007г. is equal to 678 000 tonnes, and in 2008 – 672 000 tonnes, the value of SSB being increased till 1 193 000 tonnes in 2008.

The XSA model traditionally used by AFWG for the North-East Arctic cod stock assessment is to be considered as only one of great variety of age-structured stock assessment models used for this purpose in the present world, and the model which is not without known important shortcomings. Our results obtained using the TISVPA model, the model aimed at robustness and with attention to generation-dependent peculiarities, may indicate that the XSA-derived results may lead to erroneous perception of the stock and hardly can be regarded as the only and sufficiently reliable basis for further management decisions.

REFERENCES

ICES 2007. ICES AFWG Report. 2007, p. 128-265

Kizner Z.I. and D.A.Vasilyev. 1997. Instantaneous Separable VPA (ISVPA). ICES Journal of Marine Science, 54 , N 3: p.399-411.

Shepherd J.G. 1991. Extended Survivors' Analysis: an improved method for analysis of catch-at-age data. //ICES Working Group on Methods. June 1991. Working paper . -16 p.

Vasilyev, D. 2004. Winsorization: does it help in cohort models? ICES CM2004/K:45.

Vasilyev D. 2005. Key aspects of robust fish stock assessment. M: VNIRO Publishing, 2005. 105 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.

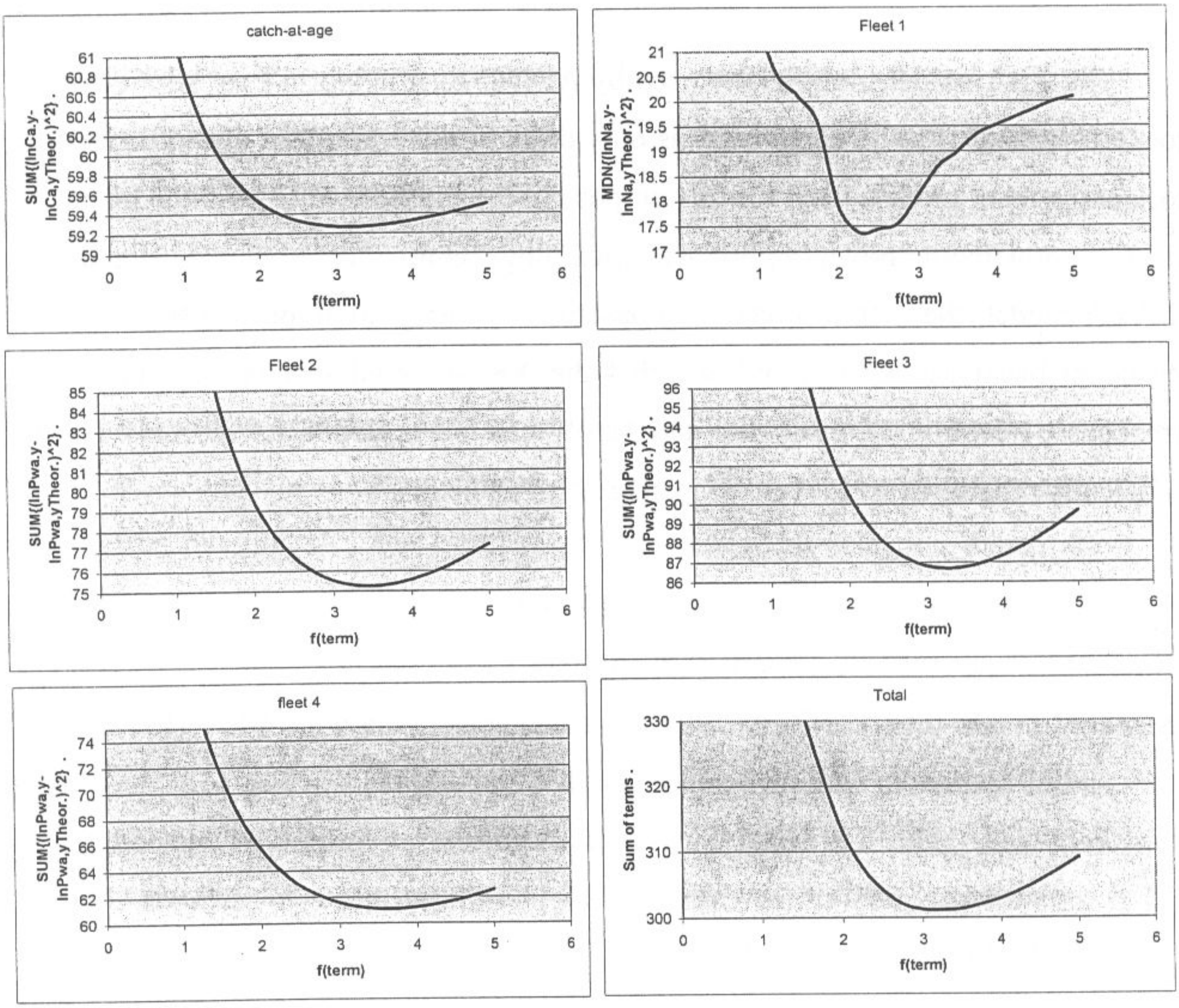

Figure.1 Profiles of the components of the model loss function.

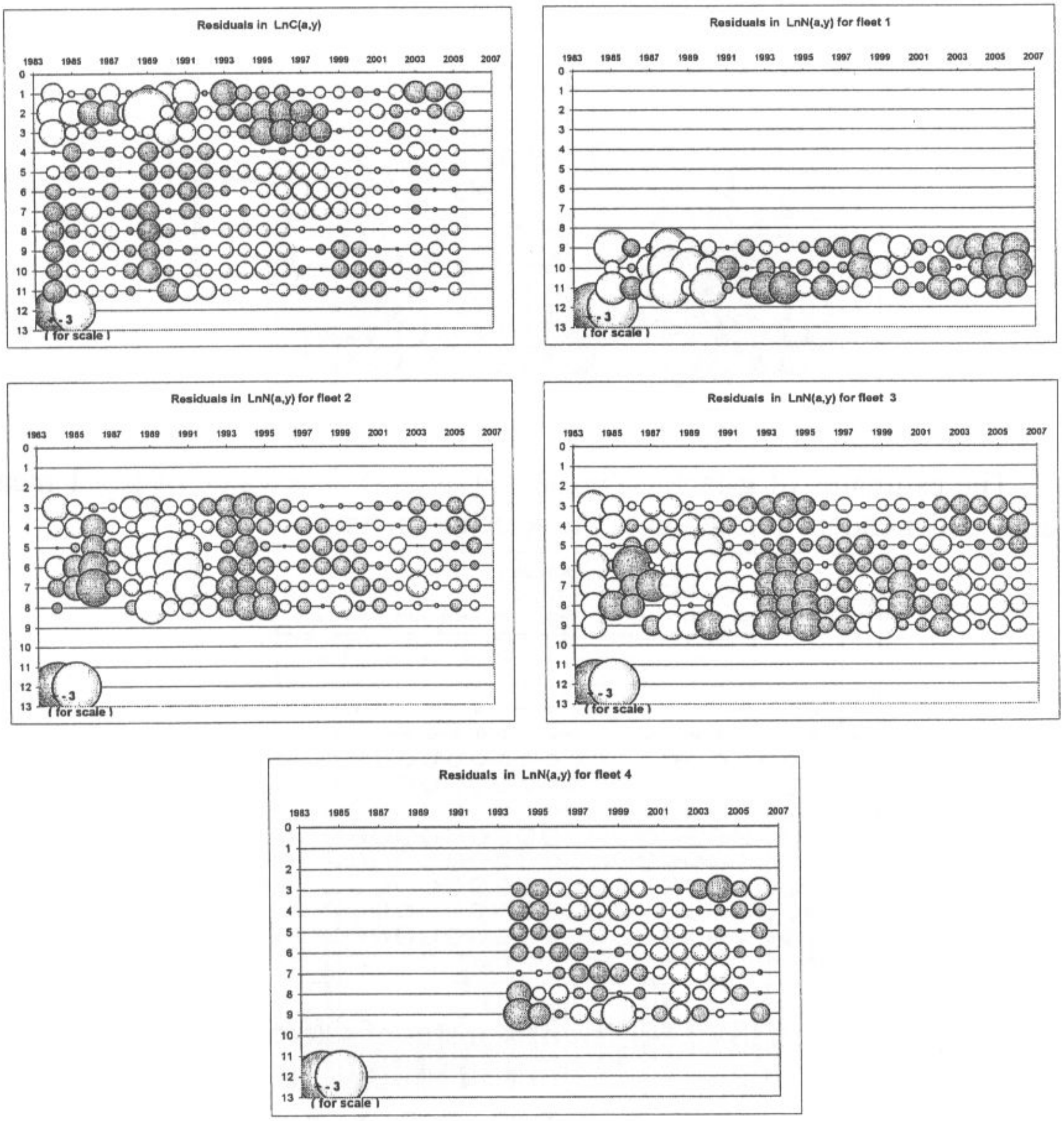

Figure 2. Bubble-plots of the model residuals for each source of input data.

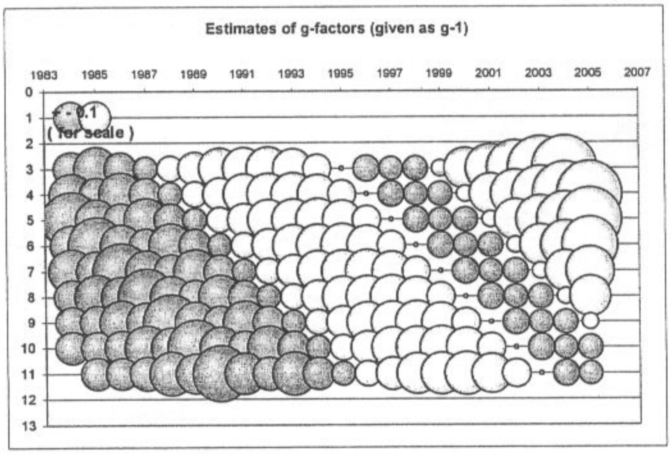

Figure 3. The estimates of g-factors (presented as g-1). The value of g = 1 means no generation-dependent correction of the selection pattern.

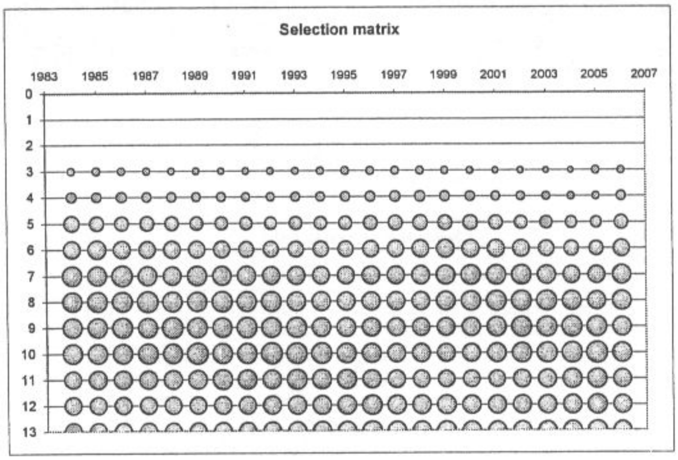

Figure 4. The TISVPA-derived estimates of selection matrix.

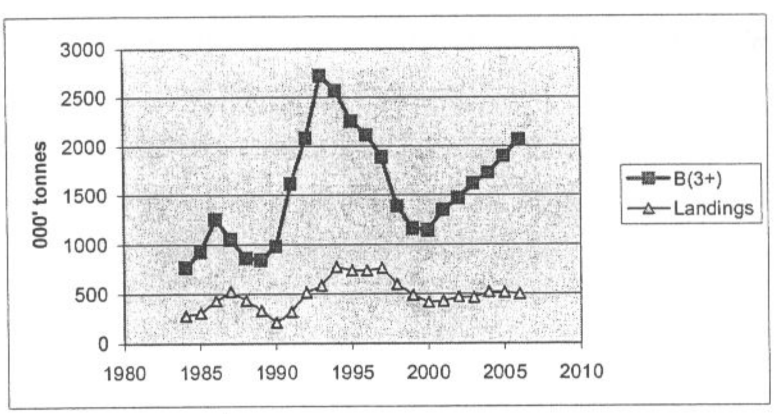

Figure 5. Fishing stock biomass estimates (TISVPA).

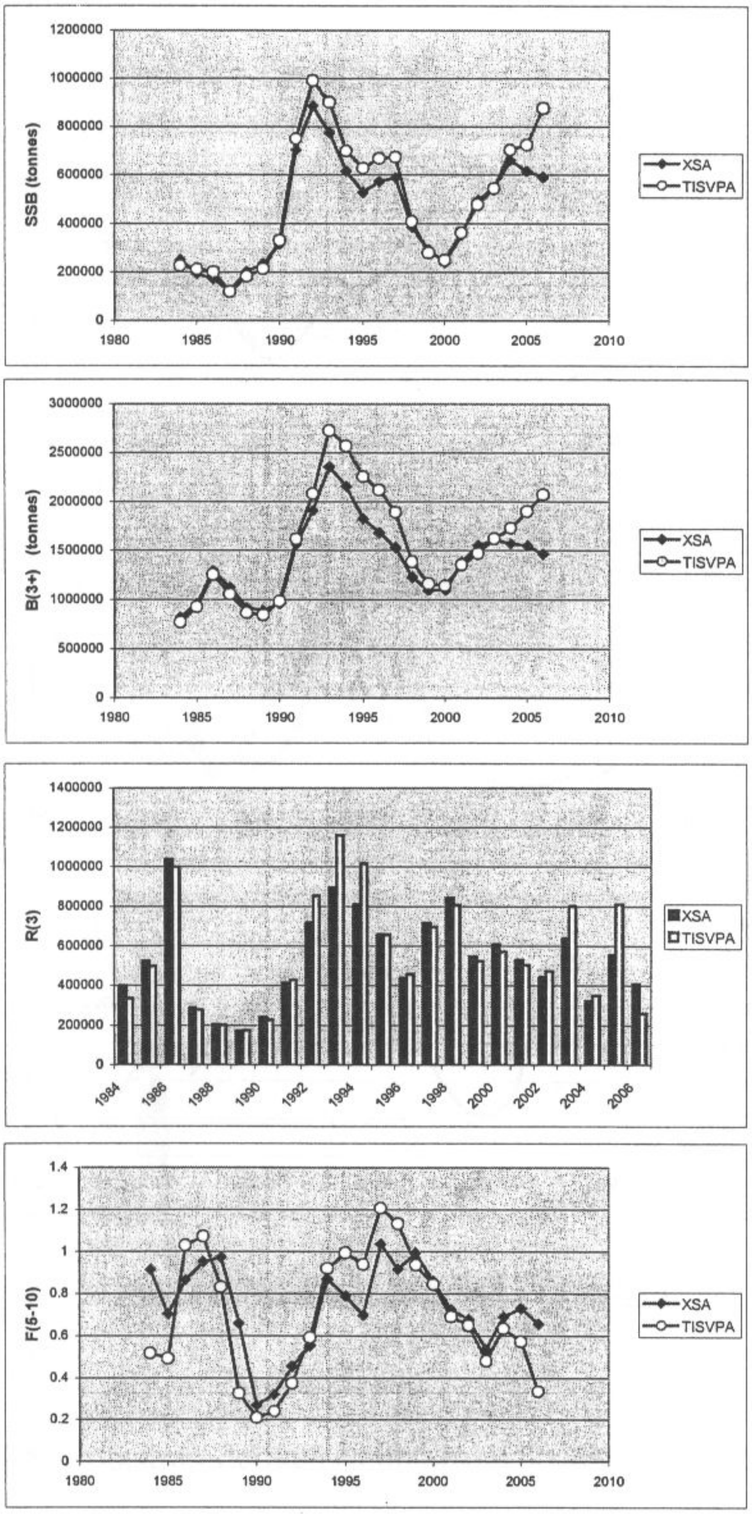

Figure 6. Comparison of the results obtained by means of XSA and TISVPA.

For comparison the results of the run SVPASA15/V15 are taken (ICES 2007). This run was considered by the AFWG as final. Russian estimates of unreported catches were used.

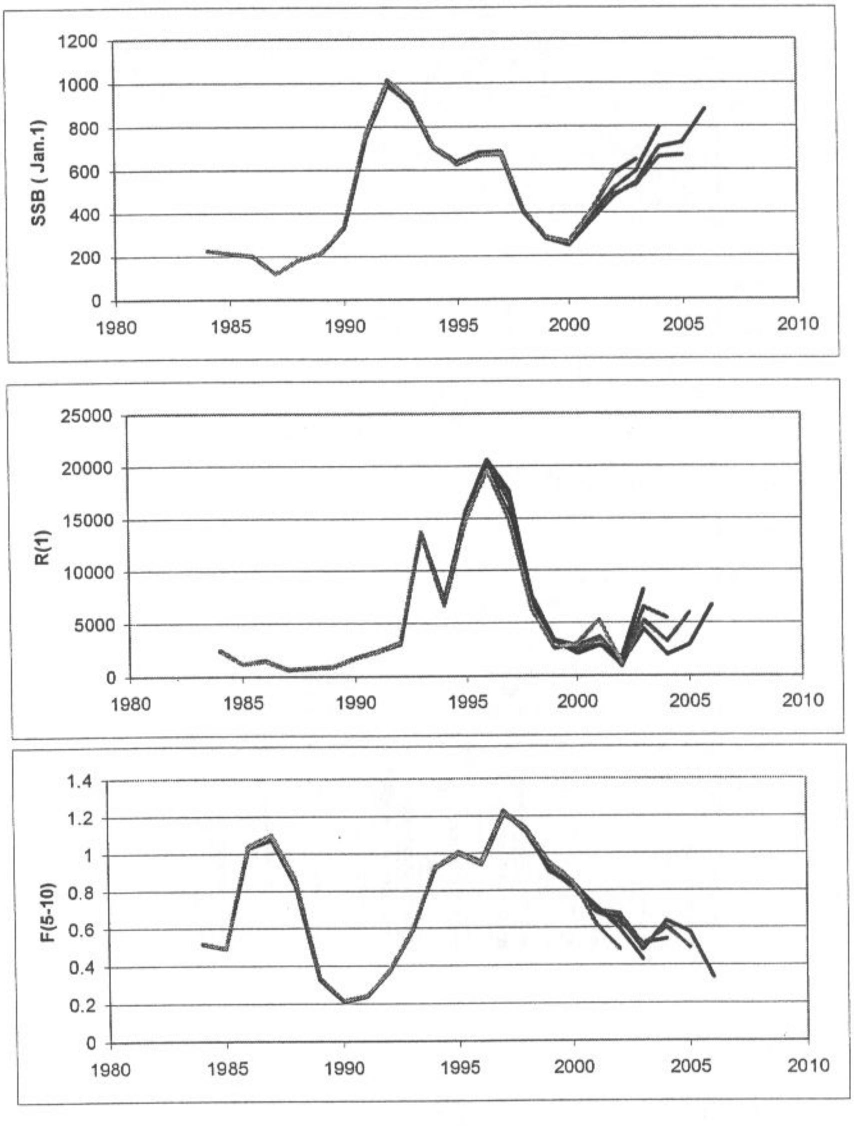

Figure 7. TISVPA retrospective runs.

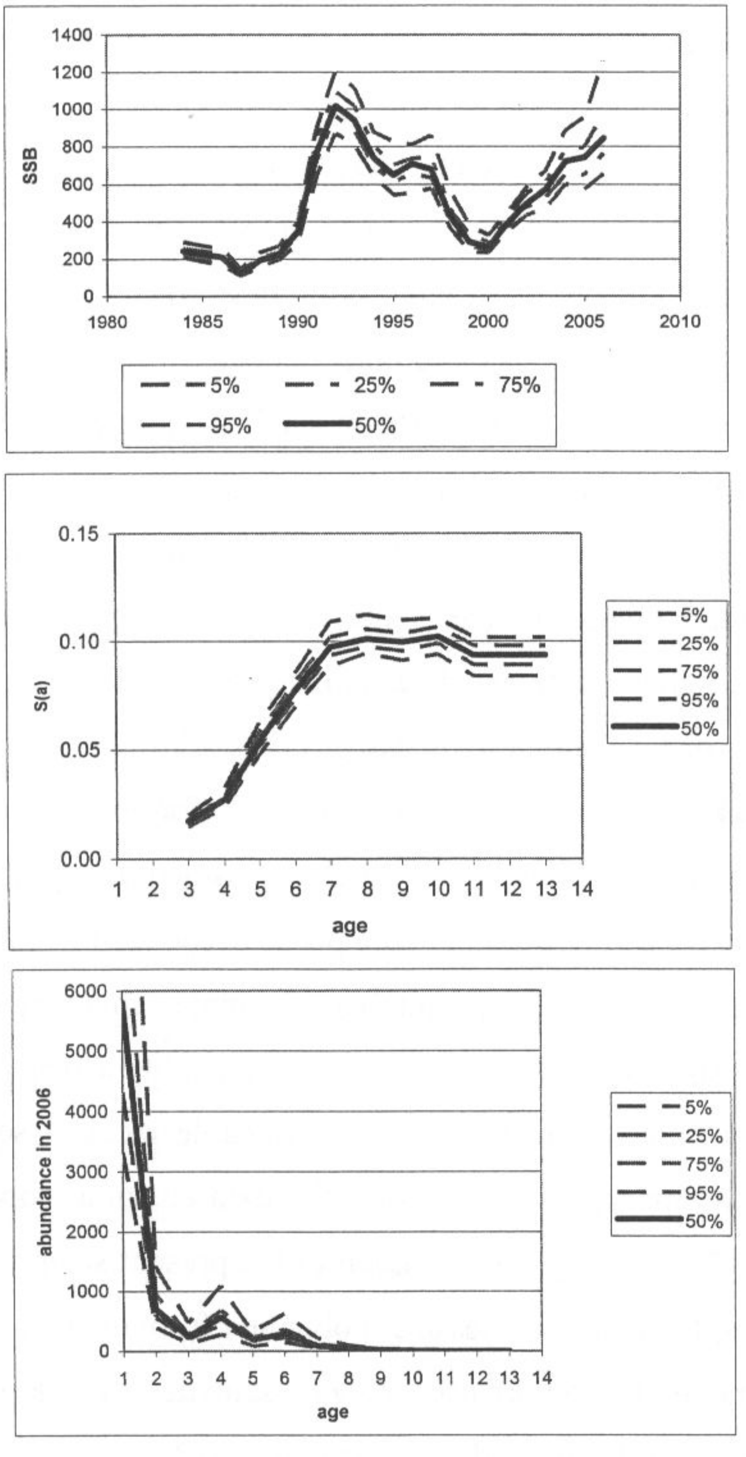

Figure 8. TISVPA. Bootstrap-derived estimates of uncertainty in the results.