

Northwest Atlantic



Fisheries Organization

Serial No. N2180

NAFO SCR Doc. 93/5

SCIENTIFIC COUNCIL MEETING - JUNE 1993

On Stabilization of Commercial Fishes Stock Estimates,
Obtained by Means of the Adaptive Approach

by

P. S. Gasiukov

Atlantic Scientific Research Institute of Marine Fisheries and Oceanography (AtlantNIRO)
5 Dmitry Donskoy Street, Kaliningrad 236000, Russia

1. Introduction

In 1991 the article was published by Canadian scientists. (Sinclair A. et al, 1991), with analysis of total allowable catch (TAC) estimates reliability for major commercial species in NAFO areas during several years. Conclusions, made by the authors, conformed the fact, that stock estimates may vary considerably from year to year. Scientists of the working group utilizing new IKES methods made the similar conclusions (Annon, 1991).

This phenomenon may be called as estimate instability and is typical to many methods of virtual population analysis (VPA) tuning, utilized in international fishery organizations work.

This problem should be studied in details, as the commercial fishes stock estimates instability may result in erroneous decisions on management.

2. Adaptive framework

The adaptive framework was chosen from numerous methods of VPA tuning (Gavaris S., 1983). The main reasons of this are following. At present the method is in common use. It was utilized to assess stock of the major commercial fishes in the North-West Atlantic and in work of NAFO, ICCAT, CCAMLR.

As compared to other methods of VPA tuning, the adaptive framework method has several advantages. First, the method is based on the clear criterion of unknown parameters estimation. Second, the method permits to utilize the data of observations, obtained from various sources.

Let us confine to the adaptive framework formulation, assuming that the data of a commercial species observation include the data on catch age composition, direct observations of abundance indices based on the strawl, in accounting surveys and data on catch-per-effort by age groups.

In such case a task is determined as following: it is necessary to estimate abundance of the object researched by age groups, in the terminal year, based on the following criterion:

$$SS = \lambda_{RV} \cdot SS_{RV} + \lambda_{cpue} \cdot SS_{cpue} \quad (I)$$

$$SS_{RV} = \sum_{a=1}^{a_k} \sum_{y=1}^{y_k} \left(\text{obs} \ln I_{ay}^{RV} - \text{calc} \ln I_{ay}^{RV} \right)^2 \quad (2)$$

$$SS_{cpue} = \sum_{a=1}^{a_k} \sum_{y=1}^{y_k} \left(\text{obs} \ln I_{ay}^{cpue} - \text{calc} \ln I_{ay}^{cpue} \right)^2 \quad (3)$$

where *obs* - values observed;

calc - values calculated;

λ_{RV} , λ_{cpue} - weight factors;

I_{ay}^{RV} - abundance index based on trawl accounting surveys;

I_{ay}^{cpue} - catch per effort by ages;

a_k - the oldest age group;

y_k - terminal year.

The calculated values in the objective function are obtained by means of creating appropriate regression equations at the assumption of multiplicative pattern of error:

$$I_{ay}^{RV} = q_a^{RV} \cdot N_{ay} \cdot \xi^{RV} \quad (4)$$

$$I_{ay}^{cpue} = q_a^{cpue} \cdot N_{ay} \cdot \xi^{cpue} \quad (5)$$

where q_a^R and q_a^{cpue} - appropriate proportionality factors;

ξ^{RV} and ξ^{cpue} - random errors;

N_{ay} - abundance estimate for *a* age group in *y* year, based on VPA.

The first component of the objective function is a mean quadratic error of approximation of the data observed in trawl accounting surveys, and the second component is a mean quadratic error of approximation of the catch rate data. Since the initial data of various scale the model requires to include the weight factors for each component of the objective function.

Thus, the adaptive framework task is to search values of age groups abundance in the last fishing year which provides the minimum mean quadratic error sum of the objective function. (1).

3. Retrospective analysis

The task of the model (1)-(5) stability is solved by means of retrospective analysis. This analysis required the artificially created situation when the model is dynamically run starting from a definite year and finishing with terminal year.

For example there are data on silver hake in the Scotian area, utilized to estimate the stock size in NAFO, starting from 1977 till the current moment (Waldron D.J. et al., 1992, Gasiukov F.S., 1993). So the task of hake stock estimation may be solved based on restricted data from 1977 through 1985 and the same calculation method, utilized at present. Then the next calculation is carried out using the data, including 1986, further - 1987 and so on till the last fishing year. Thus the ensemble of estimates is obtained, which determines the method behaviour at the final time period.

During such dynamical running the estimates of the current terminal year are compared to estimates, obtained by means of the total information available. Both those estimates have errors, caused by random and regular errors of catch age composition assessment, average fish weight by ages, natural mortality and also by the effect of uneven fishery distribution within a year, migrations and other factors (Ultang O., 1977). At the same time the estimate for the current terminal year contains random and regular errors, caused by the method of VPA tuning. In the case of small errors those estimates may be considered as stable, otherwise the estimates are of unstable pattern.

Tables 1-4 and figure 1 show the results of the retrospective analysis for data on hake, starting from 1984 through 1990, obtained at $\lambda_{RV}=I$ and $\lambda_{cpue}=I$. Total biomass, biomass from the age of 2+, summed abundance from the age of 2+ and average weighted coefficient of fishing mortality are presented.

Those results revealed that retrospective estimates for particular years may considerably deviate from the estimates, obtained on total information. For example if the year of 1985 is used as a terminal one, the method available should result in biomass estimate of 413.8 thousand tons. At the same time, the same estimate, obtained by means of VPA, based on the information for 1977-1992, will amount to 282.1 thousand tons. Similarly when the year of 1987 is used the as a terminal one, the total biomass estimate will amount to 431 thousand tons, and retrospective estimate based on data for 1977-1992 will be 197 thousand tons. It is evident, that total biomass estimate may deviate from the estimate, based on the total information, up to 70%.

Similar significant deviations are observed in estimates of biomass at age 2+, total abundance, average weighted coefficient of fishing mortality.

Taking in account the fact, that abundance and biomass estimates for each terminal year are used as a base to predict the prospect stock dynamics and to calculate TAC, it is evident, that unstable estimates could result in wrong management decisions with serious consequences.

Thus a special feature of parameter assessment method, using the adaptive framework, is a considerable variation of those parameters at transfer from year to year. In other words abundance estimates and parameters do not correlate when a new batch of information is added after the current fishing year is over.

4. Methods to obtain stable estimates.

The estimate instability revealed is common for the tasks, named as "ill-posed tasks". The term of "ill-posed tasks" (Tikhonov A.N., Arsenin V.A., 1986) implies significant variation of the final result at some change of initial information. The idea of ill-posed task solving methods was used as a base to improve the adaptive framework and to provide stable estimates.

The task of functional (1) minimization is equivalent to solving of a system of equations when each of them could be written as following:

$$\ln q_a^{RV} + \ln N_{ay}(M, N_a, C) = \ln I_{ay}^{RV} \quad (6)$$

$$\ln q_a^{opus} + \ln N_{ay}(M, N_a, C) = \ln I_{ay}^{opus} \quad (7)$$

The term $N_{ay}(M, N_a, C)$ shows that the N_{ay} estimate, obtained by means of VFA, is a function of natural mortality M , actual catches by years and ages C , unknown abundance estimates in a terminal year and several other factors. That equation system could be written in the form of operator equation:

$$Az = u \quad (8)$$

where z - unknown values,

A - operator,

u - vector of known right terms of the equation system.

Its solution (z^*) is of the form:

$$z^* = \arg \inf_z \rho(Az, u) \quad (9)$$

where ρ - is a distance in the vector space U ,

\inf - precise lower limit of function ρ . At adaptive approach (1), SS is an equivalent of function ρ , vector of unknown abundance estimates N_a is an equivalent of the vector z and logarithms of appropriate indices are equivalents of the right terms of the equation system.

This equation is difficult to solve as the operator A and right term vector u are not exactly known due to occurrence of some elements, contained observation errors. Besides, model

relations assumed in VPA may not correspond to the actual processes. As a result, approximated solution obtained in accordance to traditional methods (8) may significantly deviate from a precise solution. Such tasks are classified as ill-posed ones.

To solve the latter, special methods have been developed, and one of them is based on the concept of regularization algorithm (Tikhonov A.N., Arsenin V.A., 1986). The task of minimum searching (9) is substituted by a new task, obtained by means of function ρ - replacement by smoothing functional by Tikhonov:

$$\rho_{\alpha}(z) = \rho(Az, u) + \alpha \cdot \Omega(z) \quad (10)$$

where α - so called regularization parameter,

Ω - additional functional, referred to as "stabilizer".

The norm of vector in finite dimensional space is an example of stabilizer.

Such modification alters the task characteristics, and in definite conditions, related to selection of parameter α , the task modified has improved characteristics as compared to the initial one. In particular, A.N.Tikhonov showed that the selection of regularization parameters sequence, agreed with the initial data error could be provided in such a way, and sequence z_{α} would approach precise solution of equation (8) at $\alpha \rightarrow 0$. Besides, the solution would obtain the stability required.

The task of searching the minimum functional (1), discussed from point of view of ill-posed tasks solution (Morosov, N.A., 1987), resulted in the similar expression in the form of functional by Tikhonov.

The major problem in ill-posed tasks solution by means of regularization algorithm is the estimation of parameter α . The above mentioned result by Tikhonov is formulated in terms of sequence of z_{α} values at $\alpha \rightarrow 0$. That law of α decreasing corresponds to errors of observations simultaneously approach 0. However in practical tasks solution no situation is created when errors of observations approach 0. Those errors are of fixed value so regularizing parameter is to be correlated to the former values.

Several methods are proposed to make it. The generalized principle of deviation is the most widely practised one (Tikhonov A.N., Arsenin V.A., 1986) (Morosov V.A., 1987). That method requires an a priori estimation of accuracy both right terms of equation system and the operator. There are also approximal methods for regularization parameter in cases, when researcher obtains information on unknown z values pattern. Thus the paper by Tikhonov A.N. et al (1983) considers such methods providing compact isolation in the domain of z definition or in such cases, when the unknown function is monotonic or convex one. In other words the task of regularization parameter estimation which depends on the initial information accuracy, may be approximately solved based on a priori data on unknown values behaviour for some sets.

That method might be more simple as applied to the task of parameter estimation in the model of adaptive framework. The latter is based on the feature of VPA estimates approaching the "true" values within a zone, which may be called the VPA convergence interval. The latter term implies a time interval, separated from a terminal year by a distance, appropriate to exclude actually the effect of VPA tuning upon abundance estimates and fishing mortality by age groups.

We shall illustrate that idea taking silver hake in Scotian area as an example and utilizing the approach firstly used by Butterworth D.S. et al (1990). Figure 2 shows plots of temporal trend in silver hake summed biomass, calculated by means of VPA method at terminal fishing mortality factors of actually unattainable maximum and minimum values. It could be with reasonable confidence stated that fishing mortality rates, typical for that fishing object, are within that interval of values. The values 10 and 0.03 were selected as those extreme values.

As the results presented evidence, VPA tuning effect is followed for no more than 3-4 last years of fishery. When the difference between maximum and minimum values of terminal fishing mortality are made comparable with the scatter of fishing mortality estimates, obtained by means of VPA method withing convergence interval, then tuning effect will be already insignificant in two years. That example shows, that regardless of VPA tuning method, total biomass values could be apriori considered known within the convergence interval, and could be used to determine the regularization parameter.

Let us assume, that the unknown vector \mathbf{z} is estimated by means of minimization of smoothing functional by Tikhonov (10). Let us solve that task for each point of time moments sequence $t_i \in [t_0, t_1, \dots, t_k, \dots, T]$, using a dynamical approach, characteristic of a retrospective analysis. Besides, let us assume, that within the interval $[t_0, t_k]$ (analogue of VPA convergence interval), estimated of certain function $\mathbf{T}(\mathbf{z}_k)$ from unknown parameters \mathbf{z}_k are known. Then we may estimate regularization parameter for time moment T of the term of the best approximation function $\mathbf{T}(\mathbf{z}_k)$ by means of analogue function of argument \mathbf{z}_k in the interval $[t_0, t_k]$, obtained by means of extreme task (10) solving in that interval. If $\mathbf{T}(\mathbf{z}_k)$ is the known value of function \mathbf{T} in the point t_k , $\mathbf{T}^{calc}(\mathbf{z}_k)$ is the calculated estimate of that function at $\mathbf{z}_k^* = \arg \min \rho(\mathbf{u}_k)$ in the interval $[t_0, t_k]$, then regularization parameter is defined as following:

$$\alpha_{min} = \arg \min \left\{ \left(\mathbf{T}(\mathbf{u}_k) - \mathbf{T}^{calc}(\mathbf{u}_k) \right)^2 \right\} \quad (II)$$

The calculation sequence is as following:

1. Initial value of regularization parameter $\alpha=0$ is determined.

2. The task (6) is solved in the interval $[t_0, T_1]$.
3. Function T is determined at point t_k .
4. The task (6) is solved in the interval $[t_0, t_k]$ for a set of parameter α estimates, and only the estimate provided (7) is selected.
5. The initial α value is substituted with estimate, obtained at step 4, and steps 2-5 are repeated until α changes significantly.

If there are several points where function T is known, equation (11) may be written as following:

$$\alpha_{min} = \underset{k=y}{\operatorname{argmin}} \sum_{k=y}^{y_2} \left\{ (T(u_k) - T^{calc}(u_k)) \right\}^2 \quad (I2)$$

5. Adaptive framework definition, using regularization algorithm.

The objective function of the adaptive framework, taking in account stabilization, could be written as following:

$$SS = \lambda_{RV} \cdot SS_{RV} + \lambda_{cpue} \cdot SS_{cpue} + \alpha \cdot \|N^* - N_\alpha\|^2 \quad (I3)$$

where N^* is a test element,

- $\| \|$ - a sign of vector N_α norm,
- α - regularization parameter.

As is evident from (13), apart from traditional adaptive approach, an additional term of stabilizing functional appears in the objective function record, and the task of estimate unknown N_α values in the terminal year is complicated due to the requirement to determine the regularization parameter and to set the sample element value.

Let us note the apparent characteristic of the objective function (13): when the test element value coincides with the vector of age groups abundances in the terminal year, the stabilizing functional becomes 0, and expression of the objective function coincides with (1).

Let us assume, that weight factors are known, or the algorithm for their calculation is set. Then, the calculation sequence for unknown values estimation by means of (13), is as following:

1. The test element value is assumed 0.
2. Stock estimation is carried out by means of the adaptive approach at the objective function (13). Therefore factor $\alpha = 0$, which coincides with a traditional version of the adaptive framework. Thus biomass B_y estimates are obtained for all y , y_1, y_2, \dots, y_k .
3. Time interval Y is selected from VPA convergence zone $y=[y_1, y_2]$ together with appropriate biomass estimates B_{y_1}, \dots, B_{y_2} .
4. VPA tuning is carried out by means of the adaptive framework on some set of regularization parameter α values by each year of fishery from the y interval selected.

5. From the set of regularization parameter values, mentioned in item 4, such $\alpha_{min}^{(1)}$ is selected, which provides the minimum mean quadratic error between biomass estimates, obtained at step 2 for $y \in Y$, and estimates, obtained by means of the adaptive approach (13), when years y_1, \dots, y_2 are used as the terminal one.

6. The stock is estimated at $\alpha_{min}^{(1)}$ by means of the adaptive approach in the $[y_0, y_k]$ interval, vector of average long-term abundance by age groups is calculated, which is used further as a test element N^* .

7. Steps 3,4,5 are carried out, using a new sample element. 8. A new value of $\alpha_{min}^{(2)}$ is used for the final VPA tuning by means of the adaptive approach (13).

In general, a cyclic repetition of item 7 of the algorithm should be continued until fulfilment of the formal convergence criterion. However, practical calculations show that already after the second step the improvement is so insignificant that the subsequent calculations become unreasonable.

The question may arise on the regularization parameter α values, which results in the need to carry out the calculations under item 4. That term of algorithm is more strictly written as following:

4. The regularization parameter $\alpha_{min}^{(t)}$ value is determined, which provides the minimum value of mean standard deviation between biomass estimates in years y_1, \dots, y_2 , obtained at step 2, and biomass estimates, obtained by means of the adaptive framework, when the years y_1, \dots, y_2 are used as the terminal one.

Meanwhile the estimation of factors λ_{RV} and λ_{cpue} have not been mentioned. Those values could be calculated by means of one of the following methods: 1. If statistical characteristics of observation data are known a priori, the weight factors could be calculated by means of the method, proposed by D. Kimura (1987).

2. When complicated calculations are fulfilled, the weight factors could be determined by means of "moving check" procedure (Gasiukov P.S., 1990).

3. According to Shepherd J.S. (1992) the weight factors could be determined by means of including estimates of random component variance of (4) and (5) into the procedure of calculations.

4. Weight factors could be determined as inverse values of random component variance of (4) and (5) in the VPA divergence interval, carrying out a preliminary calculation over the entire information.

5. The factors may be calculated at the step of regularization parameter estimation according to the criterion (12).

It should be noted that the robust procedures are to be used in any method of weight factors calculations based on data observed or estimated.

The algorithm of VPA tuning presented is based on the principle of regularization algorithm by Tikhonov. Thus the regularization parameter is determined according to the criterion of minimum deviations of retrospective biomass estimates from the estimates, obtained by means of VPA in the interval of the convergence method. Note, that difference between those estimates is a base to judge on their stability.

Thus, the adaptive approach with a regularization is a two-step procedure of optimization where the first step corresponds to the extreme task of adaptive framework solution at set values of the regularization parameter, and the second step corresponds to the extreme task solution for estimation of those parameters optimal values. In other words, the regularization parameter determination is carried out by means of estimating the objective function minimum, and each objective function calculation is the solution of the adaptive framework extreme task.

6. Retrospective analysis utilizing method of stable estimates

Calculations utilizing the regularization algorithm were carried out on the same information on silver hake, which was used in the retrospective analysis.

The regularization parameter is determined in VPA convergence zone in the interval of the years 1984-1990. In addition the total hake biomass at age 2+, obtained by means of VPA method for 1977-1992 is assumed to be known.

Estimation of weight factors λ_{RV} and λ_{cpue} was carried out by means of the method, mentioned above at number 4. A median of absolute values residues, multiplied by a constant, was used as the estimate of random values variance in (4) and (5). The parameters became (after λ_{RV} set to 1) equal to:

$$\alpha = (9.0E-8)^2$$

$$\lambda_{RV} = 1.0$$

$$\lambda_{cpue} = 1.16.$$

The results of calculation of total biomass, biomass at age 2+, total abundances and average weighted fishery mortality rates are presented in tables 5-8 and in the figure 3. Comparison with similar results, obtained by means of traditional algorithm, shows that the behaviour of curves significantly alters. Instead of characteristics "fan", the estimates were obtained, in good agreement to the VPA retrospective estimates. Thus, biomass at age 2+ deviates from the retrospective estimate only in one case and approach 25 thous. tons (1986). On other cases such deviation never exceeds 20 thous. tons, which is less than 14% of the absolute value. Similar results are obtained for other abundance indices and parameters of the hake population. That calculation shows, that the adaptive framework in definition the regularization

algorithm to solve ill-posed tasks provides stable estimates when the regularization parameter is estimated by the criterion of being close to the retrospective estimates.

Conclusion

At least for some commercial fish species the estimates of stock abundance and biomass, fishing mortality rate, obtained by means of the adaptive framework are characterized by high instability, which appears as significant (up to 100%) deviations of values by fishing years from the retrospective estimates, calculated by means of VPA method for the entire time period.

To stabilize stock estimates it is suggested to utilize the methods of ill-posed task solution. Utilization of those methods is carried out at assumption of known biomass values for the period, equal to VPA convergence interval.

Such approach provides stable stock estimates. Average deviation of such estimates from the retrospective ones is less than 10%.

Some problems are worth to be studied further:

1. For some objects of fishery, where estimates of observed data accuracy are available, it is useful to obtain the regularization parameter value, using the generalized "principle of deviation". Some predictable difficulties are related to the problems of estimation of abundance indices variance for trawl surveys and catch age composition proportions variance (Gasiukov P.S., 1977), (Dorovskikh R.G., Gasiukov P.S., 1989). So the values, estimated in the model, not always agree to the estimated based on observed data, which leads to absence of respective equation solution.

That problem solution (e.g. for silver hake) could require repeated analysis of the entire information available at new principles, utilizing robust statistics.

2. Bias of estimations, obtained using the regularization algorithms is to be seriously studied. Theory of that problem is well developed (Efron B., 1982), however its practical fulfillment requires a lot of calculations, and is to be carried out on efficient computer.

REFERENCES

1. Sinclair A., Gascon D., O'BOYLE R., Rivard D. and Gavaris S. Consistency of Some Northwest Atlantic Groundfish Stock Assessments. Scientific Council Studies No 18. Special Session on Management Under Uncertainties 5-7 September 1990.
2. Anon. 1991. Report of the Working Group on Methods of Fish Stock Assessments. St. John's, Newfoundland 20-27 June 1991. ICES C.M. 1991/Assess:25.
3. Gavaris S. 1988. An adaptive framework for the estimation of population size. Canadian Atl. Fish. Sci. Adv. Comm. (CAFAC) Res. Doc. 88/29. 12p.

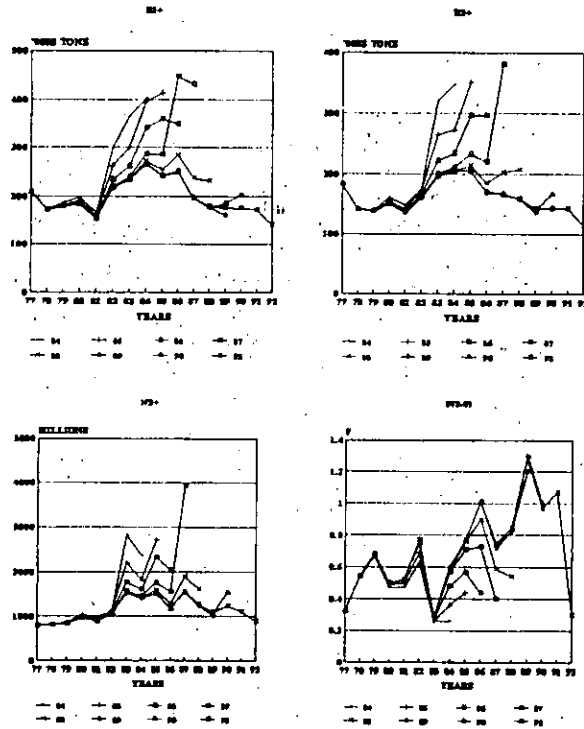


Figure 1. Results of the retrospective analysis of silver hake for 1984-1990

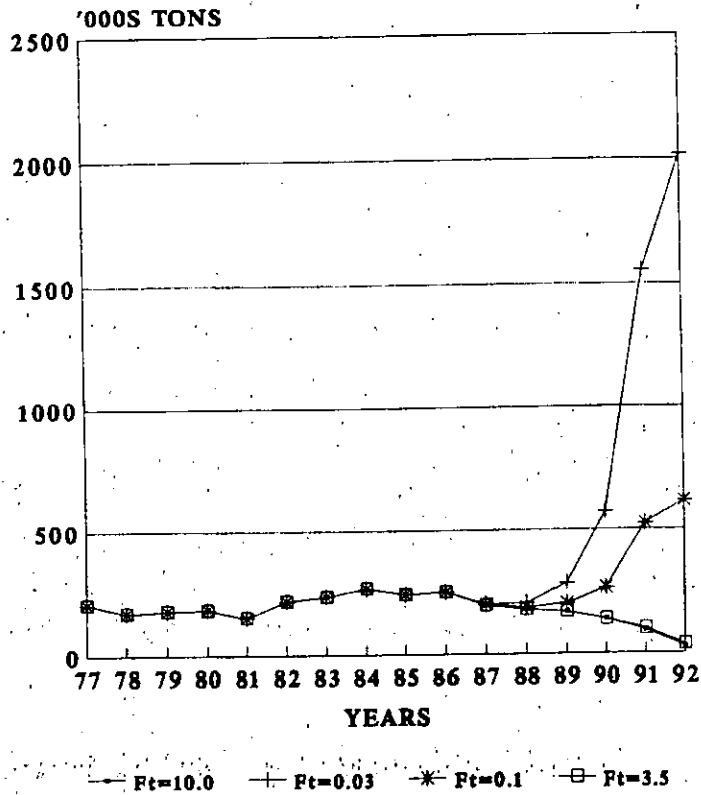


Figure 2. Range of hake biomass estimates

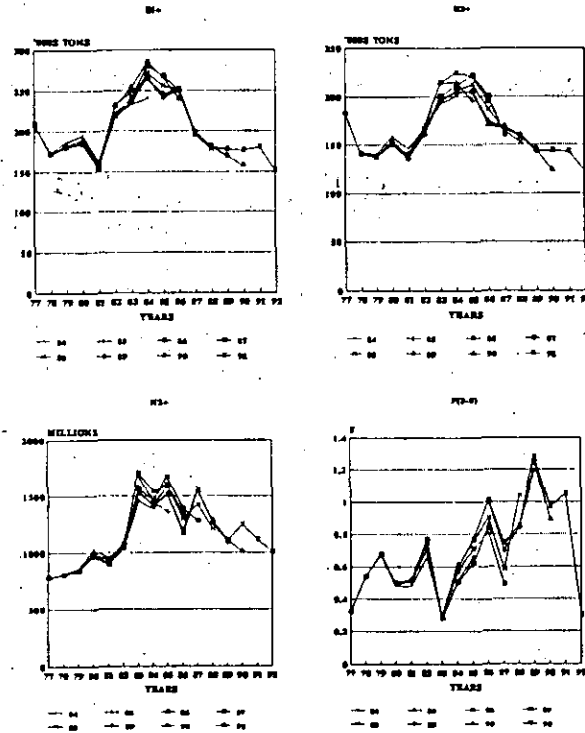


Figure 3. Results of the retrospective analysis of silver hake, obtained by means of the regularization algorithm