NAFO Sci. Coun. Studies, 16: 149-152

Bootstrap Estimation of the Confidence Intervals of Stock and TAC Assessments with the use of Dynamic Surplus Production Models

Z. I. Kizner
All-Union Research Institute of Marine Fisheries and Oceanography (VNIRO)
17 V. Krasnoselskaya, Moscow B-140, 107140, USSR

Abstract

Using a dynamic production model along with commercial fishery data, it is possible to evaluate the estimates of stock abundance for every year of the period of an intensive fishery (including the last one) and the TACs for a few years ahead. But being dependent upon the initial fishery data, these estimates are to be regarded as random ones, and therefore, it is necessary to provide them with confidence intervals. Taking into account non-linearity of the models under consideration, and the fact that usually data series are sufficiently short, the only real way to get the confidence intervals is the residual (conditional) bootstrap. The corresponding procedure is described and discussed. Two hypotheses (the error is an additive or a multiplicative one) and two kinds of bootstrap techniques (parametric and non-parametric) are compared.

Introduction

When dealing with the problem of stock and total allowable catch (TAC) assessment, one has to sometimes operate with comercial fishery data which do not reflect the age structure of the exploitable population. In such a case, a surplus production model can serve as a mathematical instrument for the investigation.

The initial data which are used for model fitting, especially, catch-per-unit-effort (CPUE) series contain errors of a different nature. Those errors can be regarded as random ones. This implies that any dynamic production model for TAC forecasting should be constructed as an **observation error model** (using the terminology of Walters, 1986). Therefore the fitting procedure for a non-linear model cannot be a simple regression, but must contain a certain iterative procedure providing gradual tuning of the model to best describe real stock dynamics. In such a case, any direct analytical estimation of confidence intervals of the model parameters, the stock size and TAC estimates (which are random values too) can be carried out very rarely, and the bootstrap technique offers the only real way to get the corresponding confidence intervals.

As an example, the dynamic production model with the control through fishing effort suggested by Kizner (MS 1989; MS 1990), is considered below. The model may be shown to be stable (Kizner, MS 1990), therefore, its confidence intervals are rather narrow compared to other versions of the model (e.g. the model with the control through catch). The model itself is described in the next section, followed by the fitting procedure. The bootstrap procedure is then described, where two approaches are compared.

The Model

Two equations expressing balance of the stock biomass and the proportion between the biomass and CPUE:

$$B_{i+1} = B_i + G(B_i) - C_i \qquad \text{... (1)}$$

$$V_i = qB_i \qquad \text{... (2)}$$

will serve as a basis of the following constructions

where

- B_i — biomass at start of the year i
- V_i — CPUE at start of the year i
- C_i — catch in the year i
- G( ) — production function: $G(B_i) = rB_i(1-B_i/K)$ and $G(B_i) = rB_i(1-1nB_i/1nK)$ according to Schaefer and Fox, respectively
- q,r,K – positive constants: q – catchability coefficient, K – carrying capacity.

Here and below in this section we operate only with 'model' (estimated) variables (except C_i and f_i).

Substitution of equation (2) into (1) and replacing C_i in (1) by f_i $(V_i + V_{i+1})/2$ reduces the system to one equation with respect to CPUE:

$$V_{i+1} = V_i + qG(V_i/q) - qf_i(V_i + V_{i+1})/2$$

which gives:

$$V_{i+1} = \frac{(1 - qf_i/2)V_i + qG(V_i/q)}{1 + qf_i/2} \qquad \text{... (3)}$$

Here f_i is the fishing effort in the year i. For example, when G is the Schaefer function, the governing equation is:

$$V_{i+1} = V_i \frac{1 - qf_i/2 + r(1 - V_i/qK)}{1 + qf_i/2} \quad \text{... (4)}$$

The TAC forecasts are calculated in this model as

$$TAC_{n+1} = f_{0.1}(V_{n+1} + V_{n+1+1})/2$$

where every new CPUE value is related to the previous one through the relationship analogous to equation (3). For example, for the Schaefer-type model

$$V_{n+k+1} = V_{n+k} \frac{1 - qf_{0.1}/2 + r(1 - V_{n+k}/qK)}{1 + qf_{0.1}/2}$$

for k = 1, 000., l (V_{n+1} is determined by equation (4), $f_{0.1}$ is a given control action and is described by the model parameters).

The stock biomass estimates can be obtained using equation (2).

Fitting Procedure

First, the initial (start) 'model' CPUE must be evaluated as

$$V_2 = (CPUE_1^{obs} + CPUE_2^{obs})/2$$

where the actual (observed) CPUEs are denoted as 'obs'.

Then, the first approximations of the model parameters q, r, K must be given (the values of the parameters of the corresponding 'process error' models can be taken) and the first approximations of the estimated V_i(i = 3, ..., n+1) must be evaluated through equation (3) (from equation (4), for the case of the Schaefer surplus production function).

Every next approximation of the estimates of the series (V_i) and of the set of the model parameters must be found in the course of the iterative procedure of minimizing the functional

$$\sum_{i=2}^{n} [(V_i + V_{i+1})/2 - CPUE_i^{obs}]^2 \quad \text{... (5)}$$

if the error is supposed to be additive or

$$\sum_{i=2}^{n} [\ln((V_i + V_{i+1})/2) - \ln CPUE_i^{obs}]^2 \quad \text{... (6)}$$

if the error is supposed to be multiplicative.

On the output of the procedure described, one has the final estimates of q, r, K, as well as V_i for i = 3, ..., n+1.

Bootstrap Estimation of the Confidence Intervals

The initial data series ($CPUE_i^{obs}$) is in fact only a sample (and usually a rather short one) from any set of possible CPUE values (parent population). If a number of such samples were available, we could repeat the whole computational procedure over and over again to obtain a lot of estimates of the model parameters, of TACs and biomass values, and then using conventional statistical methods to evaluate corresponding confidence intervals. The residual (conditional) bootstrap (Efron, 1982), which is based on this very idea, is actually a kind of Monte-Carlo approach to evaluating the statistical characteristics of the above-mentioned estimates by means of simulation of artificial input data statistically similar to the initial data series.

Starting from the maximum likelihood principle, the residual bootstrap regards the residuals $\varepsilon_i = CPUE_i^{obs} - V_i$ as being a sample representing the random component in the input data, when the error is an additive one. Supposing all ε_i are equally distributed, we can take any permutation of the residuals and add every term of the new sequence to corresponding 'model' (estimated) CPUE values to obtain a new data series (a replication) similar to but different from the initial one. In the case of the multiplicative error, the ratios $\delta_i = CPUE_i^{obs}/V_i$ should be rearranged, and then every 'model' CPUE value must be multiplied by corresponding δ_i to get a new artificial data series. This described approach is called non-parametric bootstrap.

A modification of this method, called parametric bootstrap, can be obtained by the use of a generator of pseudo-random numbers distributed just as the residuals are. Since the method of least squares is used when fitting the model, it is only natural to use a generator of normal (if the error in the initial data is supposed to be additive) or lognormal (in the case of the multiplicative error) numbers; minimization of the functionals (5) and (6) should be performed in the first and the second case correspondingly.

Results and Discussion

A comparison of different variants of the described bootstrap technique was carried out with the use of a computer program made by the author in co-operation with V. Babajan and M. Matushansky (Babajan *et al.*, MS 1989). It was found that for an approximately 25-year series it is sufficient to produce about 200–250 replications to get more or less accurate estimates of the confidence intervals. Another result is that for the present model, the hypothesis of the multiplicative

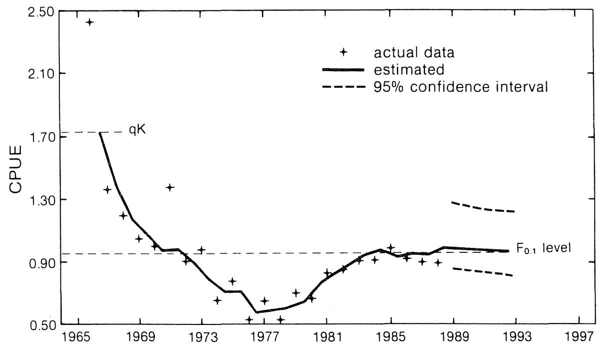

Fig. 1. Actual data, estimated CPUE and 95% confidence intervals for forecasted TACs for Cape hake in ICSEAF Division 1.5.

TABLE 1. Main results of the computations (numbers in parentheses are the coefficients of variation, i.e. the ratios of standard errors to corresponding estimates).

Estimates and coefficients of variation			
Model parameters			
q =	0.001412(0.2691)		
K =	1221.294312(0.1758)		
r =	0.554281(0.2098)		
Parameters of the equilibrium CPUE *vs* effort relationship			
a =	1.724625(0.1095)	b =	0.004394(0.1876)
F_{msy} strategy			
MSY	169.234924 (0.0494)		
E_{msy}	196.257065 (0.0906)		
B_{msy}	610.647156 (0.1758)		
$CPUE_{msy}$	0.862313(0.1095)	At start of the current year	
$F_{0.1}$ strategy		B_t	665.309737 (0.2314)
		B_t/K	0.544758 (0.0945)
$E_{0.1}$	176.631348 (0.0906)	B_t/B_{msy}	1.089516 (0.0945)
$B_{0.1}$	671.711853 (0.1758)		
$CPUE_{0.1}$	0.948544 (0.1095)	RY=G(B_t)	165.804106 (0.0521)
Fitting statistics			
SS = 0.330712	Residual mean = 0.005056	S.D.E. = 0.125385	

nature of the error and, consequently, minimization of the functional (6) and utilization of the generator of lognormal numbers are preferable.

It goes without saying, that the increase in number of replications increases the accuracy of the confidence intervals estimate (say 1,000, which seems to be enough in the present case). But the upper limit of this number depends also on the capacity of the computer.

The results of application of the described approach to analysis of the Cape hake fishery (in the ICSEAF Div. 1.5) can serve as an illustration of the parametric bootstrap estimation of confidence intervals of the model parameters, TAC forecasts and current stock size (see Fig. 1 and 2, and Table 1).

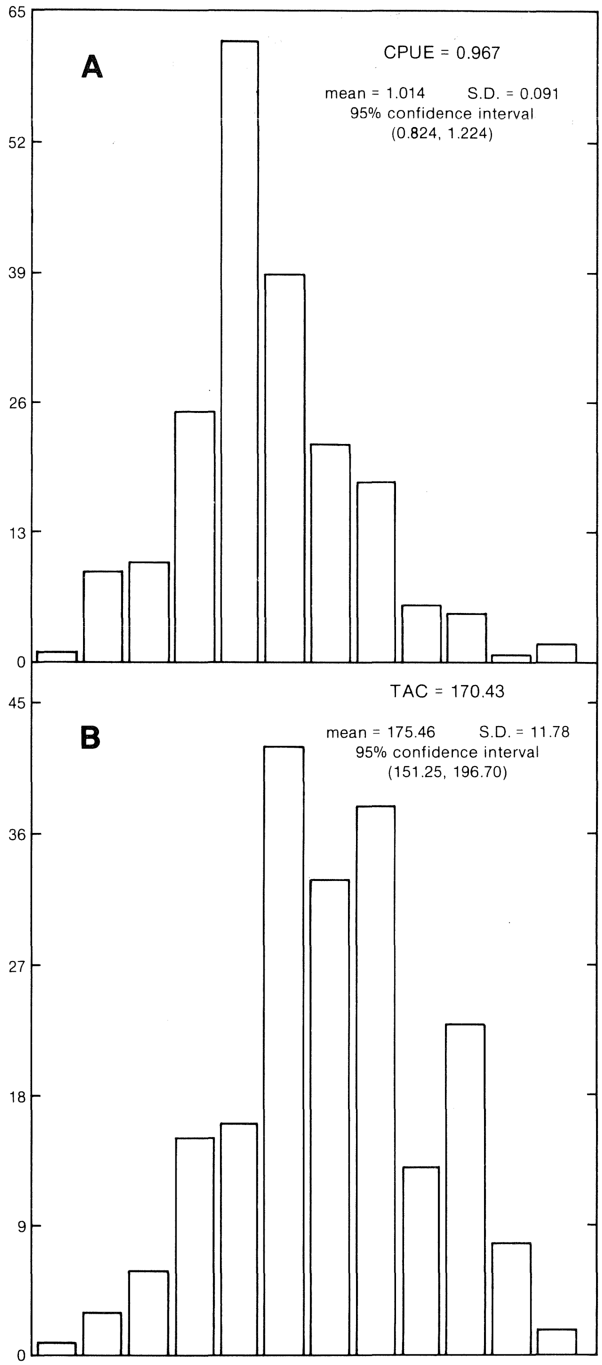

Fig. 2. Histograms of forecasted (**A**) CPUEs and (**B**) TACs 3 years ahead (1991) for 200 bootstrap replications (the data are shown in Fig. 1).

References

BABAJAN, V. K., Z. I. KIZNER, and M. V. MATUSHANSKY. MS 1989. Three versions of a dynamic production model (brief comments of BKM computer program). ICSEAF/

SAC/89/S.P./44, 1989.

EFRON, B. 1982. The Jackknife, the Bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, 1982.

KIZNER, Z. I. MS 1989. On the stability of dynamic surplus production models used in ICSEAF. Resume. ICSEAF/89/M.P./12, 1989.

MS 1990. On the stability of dynamic surplus production models used in ICSEAF. ICSEAF/SAC/90/S.P./3, 1990.

WALTERS, C. J. 1986. Adaptive management of renewable resources. McMillan, New York, 1986.