

ESTIMATES OF GROWTH AND ABUNDANCE FROM SIMULATED SIZE COMPOSITION DATA USING THE STOCK SYNTHESIS MODEL

Methot, R.D.

*National Marine Fisheries Service, Alaska Fisheries Science Center,
Bin C15700, Bldg. 4, 7600 Sand Point Way NE, Seattle, Washington 98115-0070, U.S.A.*

SUMMARY

The Stock Synthesis Model (Methot, 1990) was used to analyze four sets of simulated size composition data. Each 15-year data series included annual total catch, effort, and catch size composition. The model estimated von Bertalanffy growth parameters, dispersion of size at age, annual recruitment, fishery selectivity, and fishing mortality. The model was able to accurately fit the data in circumstances of high and low contrast in the actual recruitment, and of high and low dispersing of size at age.

RESUME

Le modèle de synthèse du stock (Methot, 1990) a été utilisé pour analyser quatre jeux de données simulées sur la composition de taille. Chaque série de quinze années comprenait la capture annuelle totale, l'effort et la composition de taille de la prise. Le modèle a estimé les paramètres de croissance de von Bertalanffy, la dispersion de la taille avec l'âge, le recrutement annuel, la sélectivité de la pêche et la mortalité par pêche. Le modèle a pu ajuster les données de façon précise dans des conditions de fort contraste et de contraste peu accusé du recrutement réel, et de dispersion élevée et faible de la taille avec l'âge.

RESUMEN

Se usó el modelo Síntesis de Stock (Methot, 1990) para analizar cuatro conjuntos de datos simulados de composición por talla. Cada una de las series de datos de 15 años incluía la captura anual total, esfuerzo y composición por talla de la captura. El modelo estimó parámetros de crecimiento de von Bertalanffy, dispersión de la talla por edad, reclutamiento anual, selectividad de la pesquería y mortalidad por pesca. El modelo pudo ajustar con precisión los datos en circunstancias de contraste alto y bajo en el reclutamiento real, y de dispersión alta y baja de la talla por edad.

Synthesis Model

The Synthesis model simulates the dynamics of an age-structured exploited population and simulates the processes by which the fisheries and surveys affect and observe this population (Methot, 1990). The model adjusts parameters (such as annual recruitment, fishing mortality, selectivity patterns) to maximize the quality of the model's fit to various kinds of data (survey biomass, fishery effort or CPUE, fishery and survey size and age composition). The result should be a description of the population that is consistent with each type of data, but smoothes out the measurement variability associated with the observations. In practice the various types of data may indicate patterns that are inconsistent with each other. Here the model is perhaps even more useful by providing a framework for identifying and exploring these discrepancies. The goodness-of-fit to the data is quantified in terms log-likelihood, and the best parameters are estimated by an inverse Hessian method that utilizes numerical calculation of the gradient and curvature of the log-likelihood surface (Methot, 1990).

The age-structured version of Synthesis basically assumes that all fish of a given age can be adequately described by a mean body weight and a mean selectivity. Its simulation can be briefly described as:

N_{ya} = population numbers in year y and age a (note that multiple periods within year can be defined)

M = instantaneous rate of natural mortality (may be age-specific)

W_{ytc} = body weight at age observed for fishery type t

S_{atc} = gear and age-specific selectivity (note that selectivity patterns can be changed over time)

E_{ytc} = fishing mortality rate in year y , fishery t for fully available ages (e.g. ages for which $S_{atc} = 1.0$)

$F_{ytc} = E_{ytc} S_{atc}$ = fishing mortality rate by year, fishery, and age

$Z_{ya} = M + \sum_t F_{ytc}$ = total mortality rate

$C_{ytc} = N_{ya} \frac{F_{ytc}}{Z_{ya}} (1 - e^{-Z_{ya}})$ = catch at age

$C_{yt} = \sum_a C_{ytc} W_{ytc}$ = catch biomass

$N_{y+1,a+1} = N_{ya} e^{-Z_{ya}}$ (with accumulation in the terminal age).

The age-structured synthesis model is primarily designed to examine age composition data, but it is able to calculate expected length compositions by applying a length at age distribution, $L@A_{a1}$, to the estimate of the catch at age, then summing over all ages. However, the age-structured model does not explicitly model growth, nor does it allow fishery selectivity to vary by size, rather than age.

The size-structured synthesis model maintains most characteristics of the age-structured model, but adds a length dimension to the existing time and age dimensions of the simulated population. The population age composition is expanded into the population age/length composition through use of the same length at age matrix, $L@A_{a1}$, used in the age-structured model. However, the size model is designed to estimate the growth parameters that describe this matrix. Selectivities are defined as a function of size and/or age and applied to the size- and age-structured population to generate estimates of the size and age composition of the catch.

$N'_{yalt} = L@A_{a1} N_{ya}$ = numbers at age/length in year y

where:

$L@A_{a1}$ is the proportion of the age a fish that are in the l^{th} size bin; assumes that mean size at age in the true population can be described by a von Bertalanffy growth function and that the distribution of size at age is normal

N_{ya} is the mean numbers at age during time period y

The age- and size-structured population is acted on by fisheries and surveys described by size- and/or age-specific selectivity, S_{latc} , patterns. The selectivity patterns are generally described as the product of two logistic curves (one ascending and one descending). The catch at age/length is then simply:

$$C_{yalt} = S_{latc} N'_{yalt}$$

Row and column sums of the estimated catch at age/length then indicate the expected proportion at length and proportion at age:

$$P_{lt} = \frac{\sum_a C_{yalt}}{\sum_{a1} C_{yalt}}$$

is proportion at length;

$$p_{at} = \frac{\sum_1 C_{y,at}}{\sum_{a1} C_{y,at}}$$

is proportion at true age.

Generally, the p_{at} vector is further processed by an ageing error matrix to simulate the effect of variability in the age determination process. In other words, the model's expected age composition is blurred to the extent that we believe the observed age compositions have been blurred.

The calculated mean body weight at age for each fishery depends on the size-specific selectivity and on the distribution of size at age:

$$W_{at} = \frac{\sum_1 S_{it} L@A_{a1} W_i}{\sum_1 S_{it} L@A_{a1}} = \text{mean weight of age } a \text{ fish in fishery } t$$

where:

W_i is the mean weight of fish in the i^{th} size bin.

Method (1990) also indicates how the $C_{y,at}$ matrix can be manipulated to produce estimates of mean body length at age that take into account size-selectivity and variability in the ageing process.

Application to Simulated Data

Four simulated data sets were analyzed. They differed by the degree of contrast in recruitment and the degree of dispersion in the size at age:

HC-BM: high contrast, blurred modes
 HC-CM: high contrast, clear modes
 LC-BM: low contrast, blurred modes
 LC-CM: low contrast, clear modes.

The following information from each data set were prepared for analysis with Synthesis:

1. Catch - annual total catch in numbers. Normally synthesis uses catch in weight, but by defining all fish to weigh one kg, numbers and weight are equivalent.
2. Effort - The F values in the supplied data tables were input as estimates of relative effort. Synthesis estimates a constant of proportionality between these values and estimated F .

3. Size composition - annual catch at size was condensed from the supplied 50 bins into 25 size bins (nominally labelled 10 cm through 34 cm).

Model Assumptions

1. The size composition data are from ages 1-10. Because Synthesis always uses a terminal accumulator age it was necessary to extend the model to age 11 and to give the age 11 fish a very high level of natural mortality to remove them from the modeled population.
2. The natural mortality is 0.2 for ages 1-10.
3. Selectivity to the fishery is a logistic function of size and large fish are fully available. This size-selectivity pattern is assumed to be constant during the time series.
4. Growth can be described by a von Bertalanffy function with the t_0 parameter replaced by a parameter that describes mean size at age 1.5. The dispersion of size at age follows a normal distribution with the standard deviation of size at age increasing linearly with mean size. This trend in the standard deviation is described by two parameters which are the standard deviation at age 1 and the standard deviation at age 11. Inspection of the size composition data (Figure 1) indicates that the dispersion of size at age increases during the first few years of the time series. This unusual pattern could be replicated in Synthesis only by allowing the model to estimate separate parameters for the dispersion of size at age during the first few years of the simulation (Table 1).
5. Weighting factors: The effort observations were assumed to have a CV of 5% and the size compositions were assumed to be as precise as a random sample of 200 fish (which implies a CV of about 20% for a size bin that contributes 10% of the catch).
6. The abundance at ages 1-5 in the first year are estimated as individual parameters. The abundance at age for ages 6-11 at the beginning of the first year is in equilibrium with an estimated, constant recruitment level and a catch equal to the catch in the first year of the time series. This condensation of 6 older ages into one parameter is reasonable given the blurring of the size modes for the older fish.
7. The total catch is known perfectly and the model can estimate a fishing mortality factor for each year such that the observed and estimated total catch match perfectly. The observed effort data are assumed to be proportional to the fishing mortality factor.

PARAMETERS

The 34 parameters estimated by the model are listed in Table 1.

FIT TO DATA

The model achieves an essentially perfect fit to the effort time series (Figure 2) and the size composition data (Figures 3-6). The effective sample sizes for these fits to the size composition are several thousand measured fish. The only persistent residual in these fits is an underestimation of the catch at a size of about 22 cm. The difference between the Beta function used to generate the data and the normal distribution used to fit the data does not seem to have adversely affected the model's performance. At the beginning of the time series the estimated dispersion of size at age is essentially constant for all ages. As expected, the model estimates that the standard deviation of size at age increases during the first few years of the time series (Table 1). This empirical process seems to adequately replicate the algorithm used to generate the data. Note that the algorithm used to generate the size composition data could not generate the first year's size composition. The data generator should have been run for a few years before the start of the time series in order to remove the transition from dispersion defined only by the Beta function, to dispersion defined by the growth algorithm.

COMPARISON TO KNOWN VALUES

The estimated logistic selectivity functions in synthesis could not be quite as steep as the actual exponential function used to generate the data (Figure 7). This difference explains the persistent residual at 22 cm in the fit to the size composition data.

The model's estimates of the initial age composition are close to the actual values (Figure 8), although the blurred mode runs had residuals at ages 3 and 5.

The model's estimates of the recruitment time series are essentially perfect (Figure 9).

DISCUSSION

A time series of size composition data with a consistent modal pattern and a time series of effort were sufficient to describe the characteristics of these simulated fished populations. In another application to actual data (Methot, unpublished manuscript) the distinct size modes of young fish in an annual trawl survey were analyzed simultaneously with the blurred size modes of older fish in the fishery. I believe it is generally true that size composition data which contain a consistent modal pattern and some auxiliary data (fishery effort or CPUE, survey CPUE, etc.) are sufficient to develop an essentially age-structured assessment. It is not necessary to parse the blurred size mode of older fish into ages before analysis; this information has been lost. This is especially true if the assumption of asymptotic selectivity is correct. Synthesis acknowledges that entropy wins; rather than attempt to create several age data elements out of a blurred size

mode, synthesis bases its comparison to the actual observation of a blurred mode composed of several age groups. Of course, this approach only works well if there are contrast in the data. In some initial trials with these data sets it was apparent that the likelihood surface for the Low Contract, Blurred Mode data set was rather flat. The fact that the final model convergence occurred so close to the true population values is somewhat fortuitous in this case.

In this application, synthesis was able to estimate ascending size selectivity, growth parameters, and the recruitment time series because it linked the information from 15 years of data. In principle, the model could be applied to a single size composition observation. This would require an assumption of constant recruitment and the model probably could estimate no more than 2-3 of the 5 defined growth parameters and, perhaps, the ascending size selectivity. As the user profiled the model on total mortality, Z , the model would essentially do a classical Z/K analysis. As more years of data are added, the model will gain precision in its estimates of the growth parameters, and gain the ability to estimate variation in recruitment. It may be instructive to re-analyze these simulated data sets first using only one year of data, then sequentially adding more years.

There are a other circumstances in which more information is needed in order to provide unambiguous results. If a large fraction of the age groups are accumulated in the right hand size mode, then direct age determinations will be necessary to provide reasonably precise estimates of the mortality rate within these age groups. Some phenomenon can introduce bias in simple models. For example, large fish may be able to avoid the fishery or survey gear. In this case the size-selectivity pattern is dome-shaped. Estimation of the level of decline in selectivity for larger fish is highly confounded with estimates of the maximum size, dispersion of size at older ages, and mortality at older ages. Some assumptions or independent estimates of some parameters will be required. In other real situations, older fish may migrate into deeper water away from the primary fishing grounds. In this case, there is declining vulnerability with increasing age. The author's experience with modelling sablefish on the U.S. west coast indicates that age data are necessary to detect and estimate this phenomenon.

Synthesis provides the capability to describe selectivity as a function of size and age, and to simultaneously explore goodness-of-fit to size composition, age composition, mean size at age, and age composition within a specified size range. This flexibility allows it to span the full range from simple two-parameter biomass based models, to size-based models as described here, to complex multiple fishery, age-structured models.

LITERATURE CITED

Methot, R.D. 1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. INPFC Bull. 50: 259-277.

Table 1. Parameters in synthesis model. All listed parameters were estimated by the model, except NATMORT for ages 1-10 which was fixed at the indicated value, and EFFORT SCALER which is a nuisance parameter that is automatically adjusted by the model.

PARAMETER	HC-BM	HC-CM	LC-BM	LC-CM
NATMORT (FIXED)	0.2	0.2	0.2	0.2
NATMORT@11	1.42	1.28	4.38	1.43
Logistic Selectivity Function				
INFLECTION SIZE	17.3	18.2	16.9	18.3
SLOPE	0.355	0.334	0.379	0.331
EFFORT SCALER (AUTO)	0.970	1.137	0.933	1.200
Growth parameters				
SIZE @ AGE 1.5	13.76	13.78	13.71	13.78
MEAN ASYMPTOTIC SIZE	29.08	29.67	29.35	29.68
K	0.445	0.410	0.448	0.410
STDV @ AGE 1	1.24	0.57	1.32	0.57
STDV @ AGE 11 IN 1980	1.15	0.65	0.97	0.66
STDV @ AGE 11 IN 1981	1.41	1.01	1.42	1.04
STDV @ AGE 11 IN 1982	1.67	1.24	1.57	1.25
STDV @ AGE 11 IN 1983	2.21	1.41	1.83	1.32
STDV @ AGE 11 IN 1984	2.10	1.36	2.18	1.44
STDV @ AGE 11 IN 1985	2.56	1.55	2.37	1.55
STDV @ AGE 11 IN 86-94	2.77	1.60	2.73	1.64
Recruitment and Initial Age Composition				
RECRUIT FOR AGES 6-11	0.846	0.687	0.757	0.665
1980 AGE 5	0.156	0.315	0.416	0.323
1980 AGE 4	0.563	0.481	0.508	0.459
1980 AGE 3	0.813	0.659	0.827	0.632
1980 AGE 2	0.734	0.771	0.747	0.745
1980 RECRUITMENT	0.970	0.933	0.974	0.907
1981 RECRUITMENT	1.050	0.963	1.037	0.942
1982 RECRUITMENT	5.041	4.887	1.036	0.936
1983 RECRUITMENT	1.363	1.072	1.221	1.089
1984 RECRUITMENT	0.833	0.784	0.876	0.776
1985 RECRUITMENT	1.046	1.059	1.116	1.017
1986 RECRUITMENT	0.759	0.752	0.813	0.728
1987 RECRUITMENT	1.215	1.190	1.280	1.178
1988 RECRUITMENT	2.984	2.968	1.058	0.936
1989 RECRUITMENT	0.912	0.761	0.854	0.764
1990 RECRUITMENT	0.926	0.922	0.974	0.896
1991 RECRUITMENT	0.779	0.757	0.804	0.731
1992 RECRUITMENT	0.859	0.838	0.878	0.809
1993 RECRUITMENT	0.767	0.721	0.774	0.704
1994 RECRUITMENT	0.989	0.981	1.011	0.977

LC-CM SIZE COMPOSITION

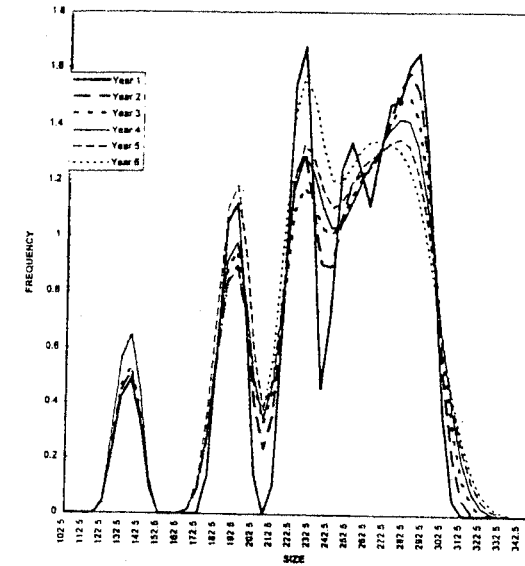


Figure 1. Size compositions for first six years in the LC-CM data set. Note the lower dispersion in size-at-age in the early years.

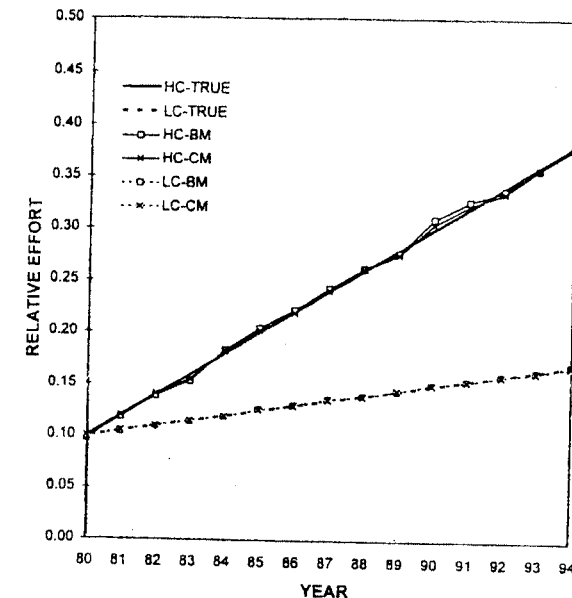


Figure 2. Synthesis model fits to the effort data.

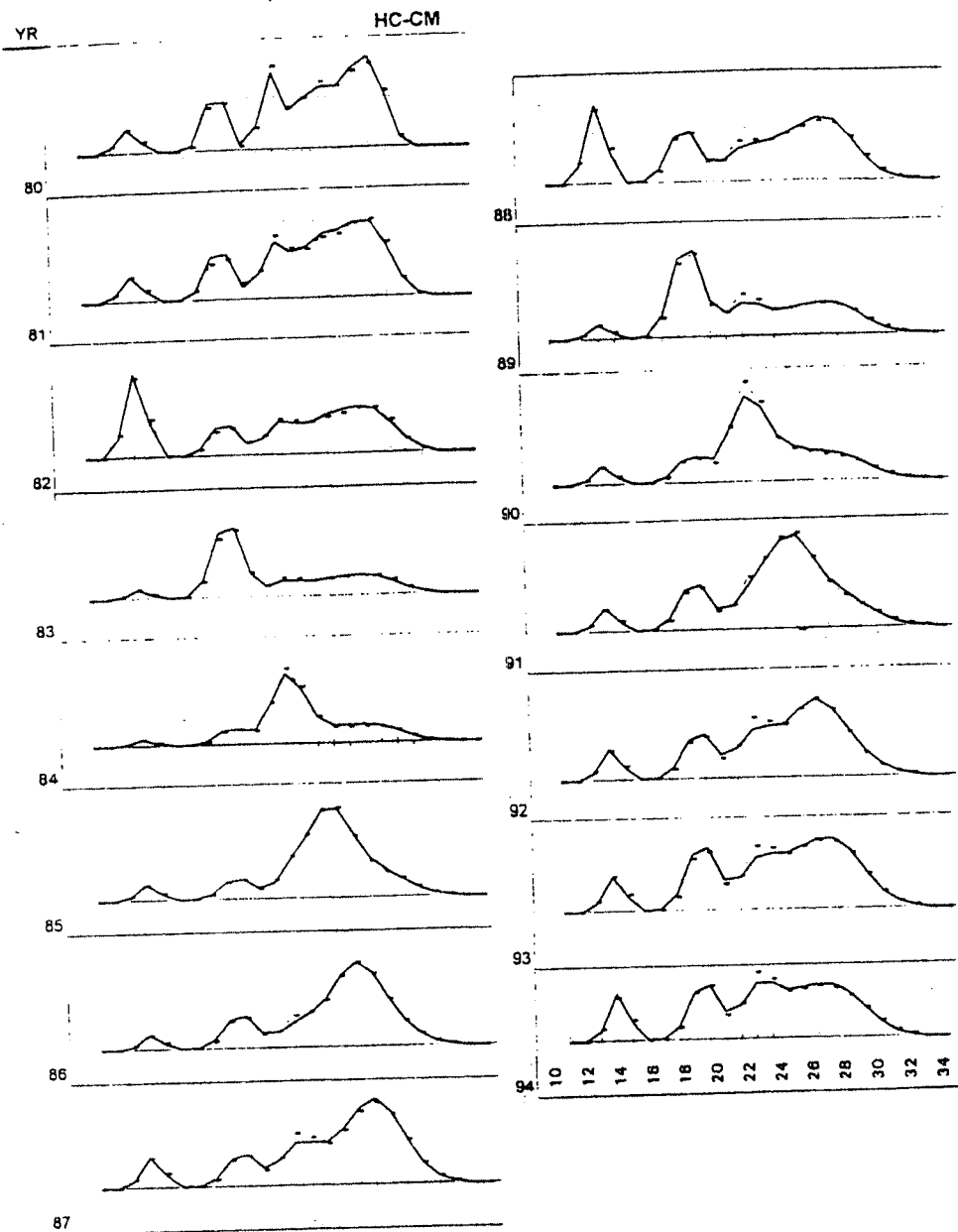


Figure 3. Synthesis model fits to the time series of size compositions in the HC-CM data set. The bold line is the model estimate of proportion at size, the dashed line with tic marks is the observed proportion.

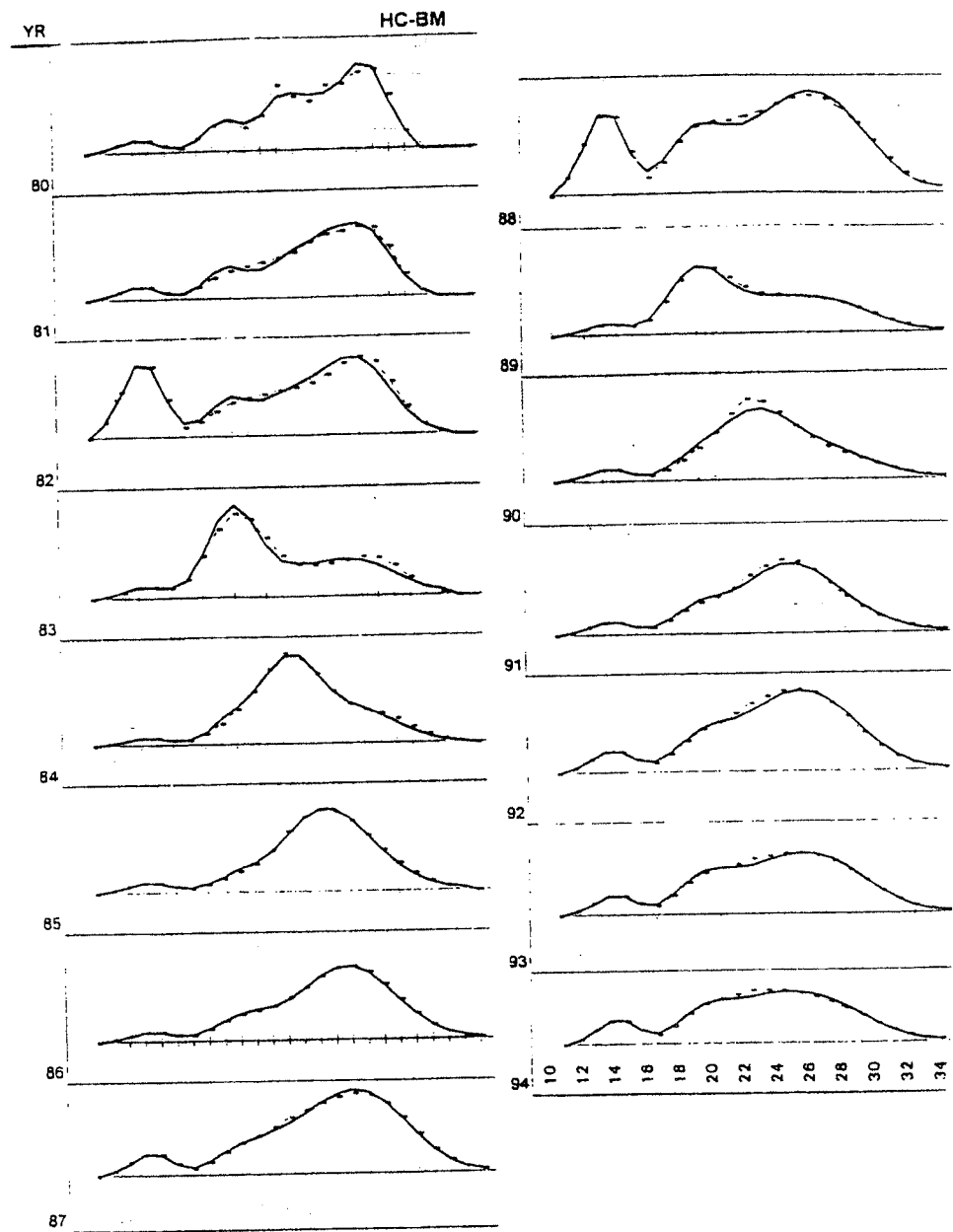


Figure 4. Synthesis model fits to the time series of size compositions in the HC-BM data set.

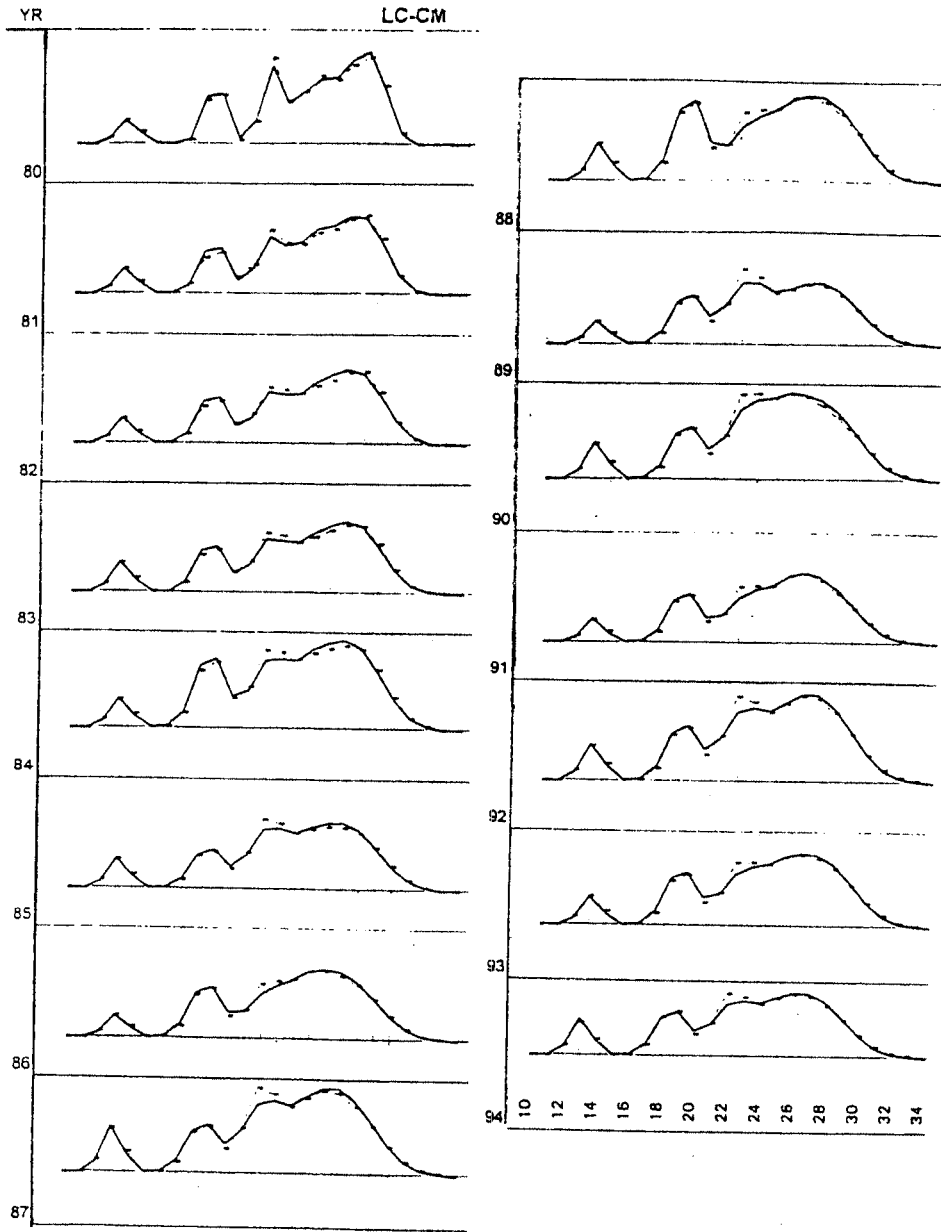


Figure 5. Synthesis model fits to the time series of size compositions in the LC-CM data set.

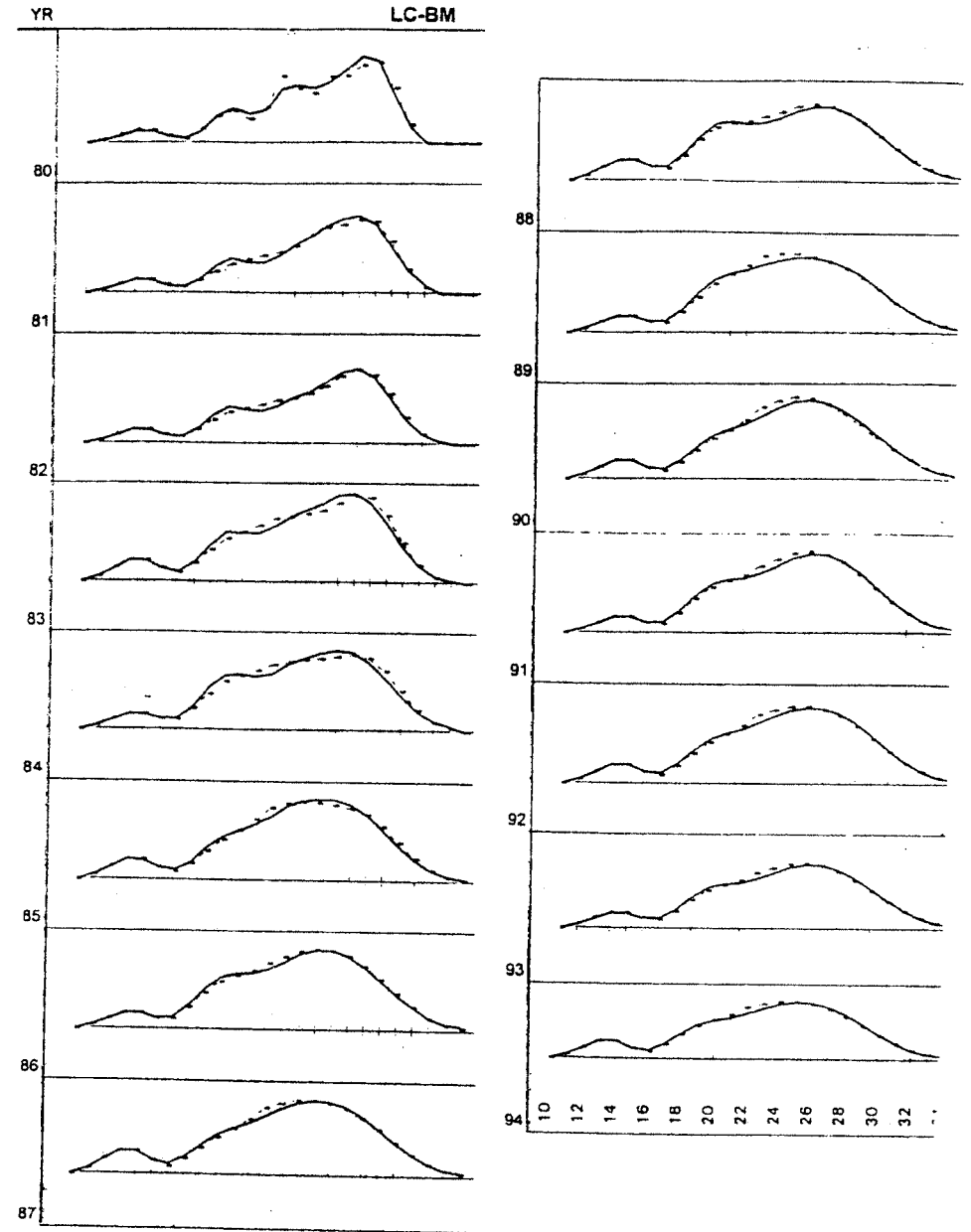


Figure 6. Synthesis model fits to the time series of size compositions in the LC-BM data set.

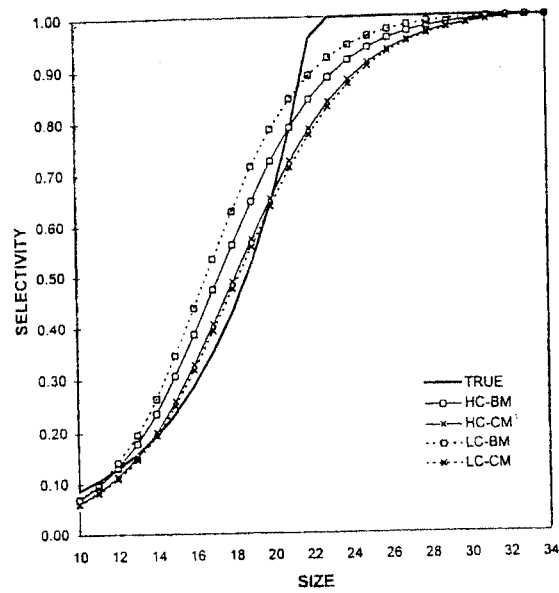


Figure 7. True and estimated size-selectivity patterns.

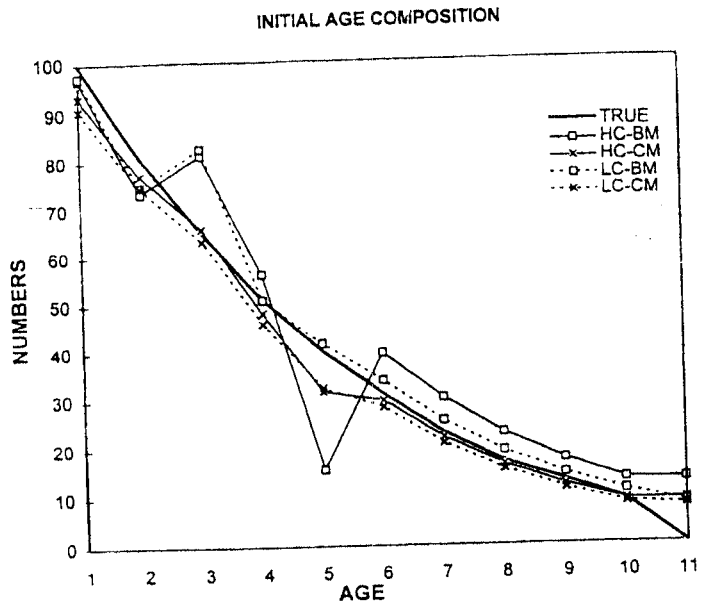


Figure 8. Comparison of true and estimated numbers-at-age in the first year of the time series.

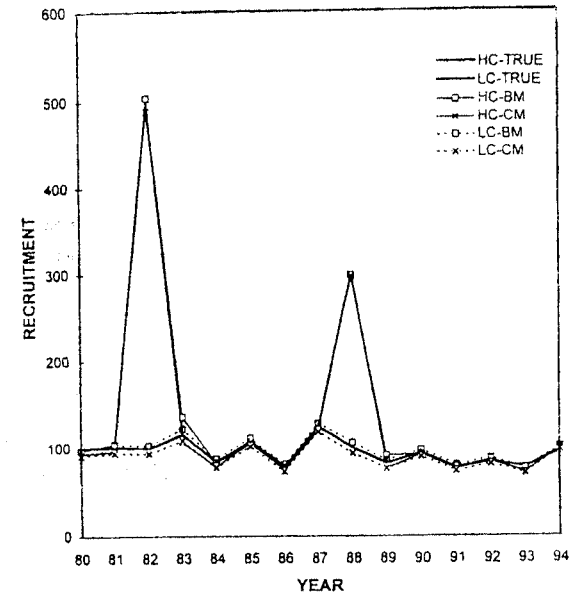


Figure 9. Comparison of true and estimated time series of recruitment at age 1.