

# ANALYSIS OF ST. ANDREWS SIMULATED DATASETS USING LENGTH-BASED SEPARABLE SEQUENTIAL POPULATION ANALYSIS

SCRS/1994/063

Col.Vol.Sci.Pap. ICCAT, 44 (2) : 61-66 (1995)

*Kimura, D.K.*

*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

Analysis of simulated data sets generated at the ICCAT Workshop on the Technical Aspects of Methodologies which Account for Individual Growth Variability by Age, held at St. Andrews Biological Station in 1993 (SCRS/93/17), indicated that the Length-based Separable Sequential Population Analysis model (SCRS/93/51) gave biased results. The apparent reason for these biases is given in SCRS/94/118 titled, "Addendum to SCRS/93/51". A re-analysis of the simulated data sets provided in this addendum shows much improved results.

## RESUME

L'analyse des jeux de données simulées élaborés lors de la Consultation ICCAT sur les Aspects techniques des Méthodologies pour intégrer la Variabilité individuelle de la Croissance par Age (Station biologique de St. Andrews, NB, Canada, 1993 - document SCRS/93/17) a signalé que le modèle d'analyse des populations séquentielle basée sur la longueur (SCRS/93/51) donnait des résultats biaisés. La cause apparente de ces biais est indiquée dans le document SCRS/94/118, intitulé "Addendum to SCRS/93/51". Une nouvelle analyse des jeux de données simulées présentée dans cet addendum montre de bien meilleurs résultats.

## RESUMEN

El análisis de conjuntos de datos simulados, generados en las Jornadas de Trabajo ICCAT sobre Aspectos Técnicos de las Metodologías que explican la Variabilidad del Crecimiento Individual por Edad, que tuvo lugar en la Estación Biológica de St. Andrews, Canadá (SCRS/93/17), indicaba que el modelo "Análisis secuencial de Población Separable basado en la talla" (SCRS/93/51) daba resultados sesgados. La razón aparente de estos sesgos se da en el documento SCRS/94/118 "Addendum al SCRS/93/51". Un nuevo análisis de los conjuntos de datos simulados, en dicho Addendum, presenta mucho mejores resultados.

## Introduction

At the ICCAT Workshop on the Technical Aspects of Methodologies which Account for Individual Growth Variability by Age, held at the St. Andrews Biological Station, New Brunswick, Canada, we offered a Length-based Separable Sequential Population Analysis (LSSPA) model for the analysis of swordfish (*Xiphias gladius*). Also at that meeting it was decided that simulated datasets with "tuna-like" characteristics would be useful for evaluating the many possible methods of analyzing length-based fishery data where catches and catch-per-unit-effort (cpue) data are available by length category with very little if any age information. In this paper we describe briefly the results of analyzing these simulated datasets using the LSSPA model.

## Materials and Methods

The LSSPA model we use is described by Kimura and Scott (1994, SCRS 93/51). Briefly, this model consists of the "usual" core separable model which includes the numbers in each year-class at the youngest age that they first enter the model, fishing mortality as a product of "full" annual mortality and age selectivity, and a size selectivity curve which together with a von Bertalanffy growth curve is used to generate the age selectivities. The von Bertalanffy growth model included a variance parameter which was used to describe variability of size-at-age. An effort was made to keep the model as sparse as possible so that model output would not depend on model nuances that might not be fully understood.

The simulated datasets are described in COM-SCRS/93/17 and consisted of four datasets generated by varying two variables at two levels. The first of these variables was described as high and low contrast, meaning that the high contrast time series contained occasional large year-classes; while the low contrast time series contained only modest variability in year-class strength. The second of these variables was described as blurred or clear modes, referring to the modal length-at-age distributions. Blurred modes means the length-at-age distributions at different ages overlapped significantly; while clear modes referred to non-overlapping length-at-age distributions. The four simulation datasets were as follows:

1. High contrast, blurred modes.
2. High contrast, clear modes.
3. Low contrast, blurred modes.
4. Low contrast, clear modes.

Input parameters for the LSSPA model were as follows:

1. Natural mortality  $M=0.2$ .
2. From COM-SCRS/93/17  
von Bertalanffy growth parameters:  $t=135.4$ ,  $L=295.4$ ,  $kbrody=0.670$ .  
a.  $cvl=0.045$  clear modes.  
b.  $cvl=0.08$  blurred modes.

Because these parameters are based on "observations" from the high contrast simulations, they may have been an important source of bias when fitting the low contrast simulations.

3. Total biomass was estimated by converting the estimated numbers at age using weight-at-age estimated from COM-SCRS/93/17 relationships:  $L=300[1-\exp(-.4t)]$  and  $W=0.00001L^3$ .

4. Selectivity and fishing effort assumed to be known and correct.

Besides these modeling parameters, LSSPA used only catch-at-length data.

### Results

The LSSPA model was fit to the four simulated datasets using least squares (Fig. 1) and least absolute deviations (Fig. 2). Least absolute deviations is generally thought to be more robust than least squares, though less efficient under the usual normality assumptions.

For least squares fits, histograms of residuals are given on the log-scale for cpue in Fig. 3 and catch in Fig. 4. For least absolute deviation fits, histograms of residuals are given on the log-scale for cpue in Fig. 5 and catch in Fig. 6.

Numeric modeling results are shown in Table 1 for least squares fits, and Table 2 for least absolute deviation fits. Values shown in these tables are year index, instantaneous full-selection annual fishing mortality rate, recruitment in numbers (1-yr-olds), age 1-10yr in numbers, age 1-10yr in biomass.

In examining the results of the LSSPA fits to the simulations, one should recall that our growth data was taken from the COM-SCRS/93/17 average length-at-age data, which was derived from the high contrast data simulations. Therefore these growth parameters may have been a significant source of bias for fitting the low contrast datasets. Keeping this in mind, possible conclusions from these model fits are:

1. Fitting the high contrast datasets presented no problems.
2. Blurred modes, when the degree of blurring is known, does not appear to present additional difficulties in the model fits.
3. Using the high contrast growth data may have caused significant bias in the model fits to the low contrast data.
4. Using least absolute deviations appeared to recover the trends in recruitment and abundance, but the bias in recruitment and abundance trends remained.

If the mean length-at-age and variabilities differ between the high contrast and low contrast data series, it might be useful to refit the low contrast data series with more correct growth parameters to see if this is the source of bias. An alternative explanation for the apparent poorer fits to the low contrast data is that the process noise purposely introduced in the simulations represented a greater noise to signal ratio for the low contrast simulations than the high contrast simulations.

Table 1. Least squares fit to the four test data sets using the LSSPA model.

A. High contrast, blurred modes.				
year	fiest	num1	num1-10	bnum1-10
1	0.175	83	297	22430
2	0.158	95	312	25247
3	0.150	500	731	31501
4	0.161	117	683	44408
5	0.186	83	592	54046
6	0.206	106	525	55107
7	0.223	77	444	51229
8	0.258	117	417	43648
9	0.280	278	563	39888
10	0.302	78	477	39793
11	0.328	87	408	38122
12	0.360	70	332	33189
13	0.399	75	279	26005
14	0.448	69	243	21437
15	0.515	84	234	17859
B. High contrast, clear modes.				
year	fiest	num1	num1-10	bnum1-10
1	0.084	103	425	37819
2	0.101	99	427	42112
3	0.100	586	912	48797
4	0.106	120	836	62952
5	0.112	90	731	73906
6	0.122	125	669	75999
7	0.148	72	557	70035
8	0.171	105	504	62635
9	0.201	235	590	55227
10	0.223	64	489	51139
11	0.247	75	412	46116
12	0.275	57	332	39220
13	0.341	57	261	28398
14	0.445	44	209	22386
15	0.723	43	167	16540
C. Low contrast, blurred modes.				
year	fiest	num1	num1-10	bnum1-10
1	0.171	86	303	22755
2	0.133	100	322	25792
3	0.115	101	344	28907
4	0.112	117	379	32189
5	0.120	83	372	35169
6	0.127	107	388	37070
7	0.131	77	368	38046
8	0.147	117	381	35561
9	0.154	93	373	35433
10	0.160	76	348	34903
11	0.168	86	337	33606
12	0.177	70	312	31765
13	0.187	76	298	29721
14	0.199	70	281	27544
15	0.210	88	287	25999
D. Low contrast, clear modes.				
year	fiest	num1	num1-10	bnum1-10
1	0.080	113	446	39275
2	0.082	120	465	44145
3	0.074	118	478	47669
4	0.072	141	513	50881
5	0.073	95	494	53647
6	0.077	135	516	54591
7	0.091	72	455	51438
8	0.103	105	445	49458
9	0.120	79	407	46073
10	0.132	63	362	42928
11	0.146	71	332	38760
12	0.164	54	292	34327
13	0.190	54	260	29929
14	0.235	43	221	25286
15	0.323	43	192	21144

Table 2. Least absolute deviation fit to the four test data sets using the LSS\ A model.

A. High contrast, blurred modes.				
year	first	num1	num1-10	bmass1-10
1	0.114	84	373	38629
2	0.142	89	331	29066
3	0.161	452	699	34536
4	0.185	114	649	45014
5	0.204	80	557	52043
6	0.219	100	489	51636
7	0.243	74	414	47320
8	0.273	112	386	39338
9	0.297	267	528	36012
10	0.322	79	452	36189
11	0.346	86	387	34886
12	0.380	68	315	30508
13	0.417	72	264	24135
14	0.467	65	229	19865
15	0.538	80	220	16486

B. High contrast, clear modes.				
year	first	num1	num1-10	bmass1-10
1	0.111	88	385	38387
2	0.134	90	362	34719
3	0.155	464	734	39379
4	0.174	105	669	49162
5	0.191	77	570	55572
6	0.210	100	502	54656
7	0.238	72	411	46958
8	0.263	111	389	40744
9	0.288	265	529	36892
10	0.307	72	446	36723
11	0.333	83	382	35275
12	0.365	66	310	30773
13	0.409	69	260	24200
14	0.459	61	222	19827
15	0.546	68	203	16219

C. Low contrast, blurred modes.				
year	first	num1	num1-10	bmass1-10
1	0.114	84	408	47384
2	0.124	89	329	28133
3	0.127	90	338	30812
4	0.129	103	357	32735
5	0.136	79	347	34185
6	0.141	96	355	34848
7	0.144	72	336	34939
8	0.156	109	347	32051
9	0.162	88	341	31509
10	0.169	74	322	31140
11	0.177	82	314	30238
12	0.184	69	294	28940
13	0.192	74	283	27447
14	0.200	69	269	25800
15	0.209	90	279	24653

D. Low contrast, clear modes.				
year	first	num1	num1-10	bmass1-10
1	0.111	88	378	37846
2	0.119	91	358	34160
3	0.123	89	359	35668
4	0.130	107	377	36607
5	0.130	76	360	37160
6	0.137	98	367	37294
7	0.145	71	331	33442
8	0.151	110	351	32641
9	0.158	89	345	32129
10	0.165	72	323	31580
11	0.173	82	314	30556
12	0.179	67	293	29181
13	0.187	72	280	27617
14	0.198	65	263	25799
15	0.211	77	263	24391

Figures

Fig. 1. Least squares fits to the four simulation datasets. Comparisons are made using known 1 yr-old recruitment, and known total 1-10 yr-old population numbers.

Fig. 2. Least absolute deviation fits to the four simulation datasets. Comparisons are made using known 1 yr-old recruitment, and known total 1-10 yr-old population numbers.

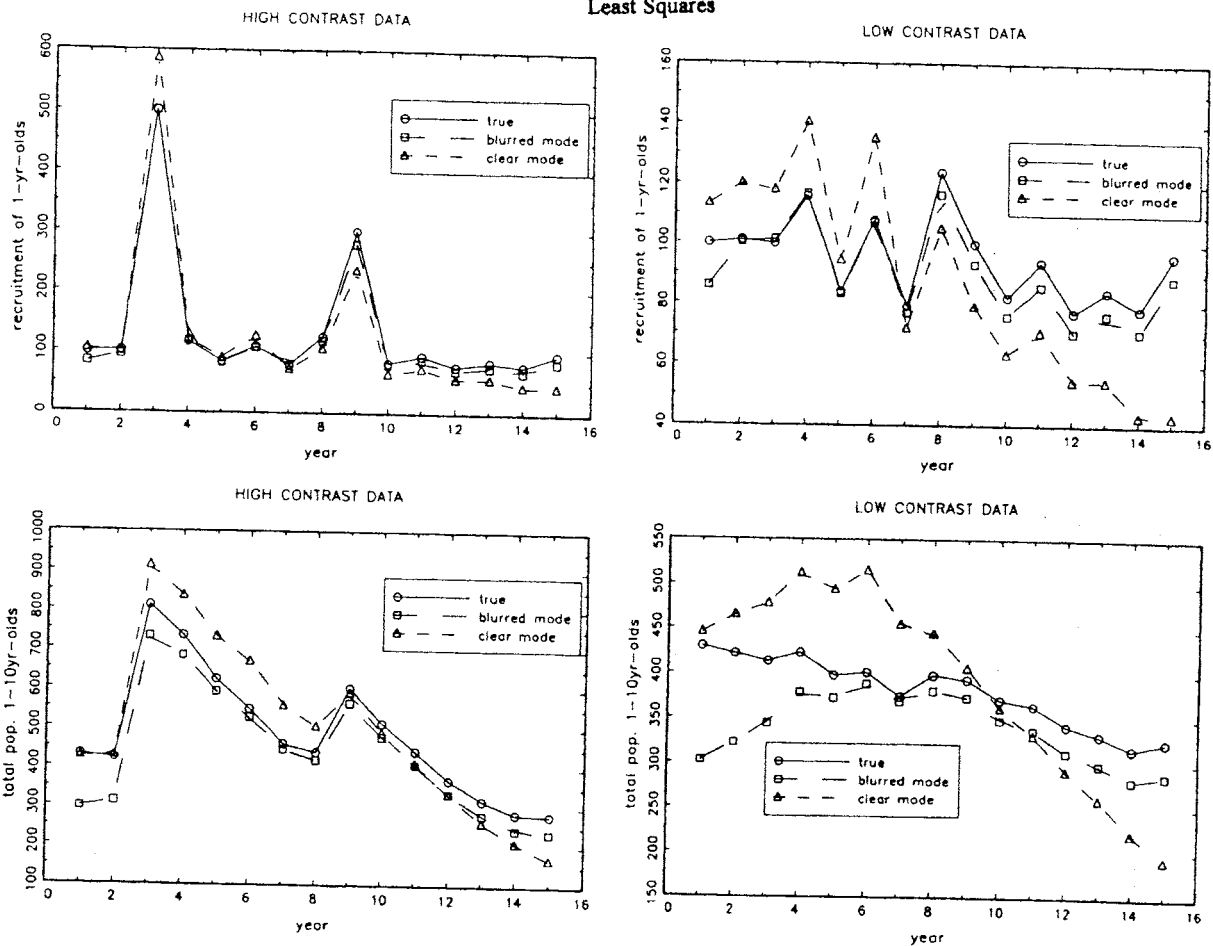
Fig. 3. Histogram of  $\log(\text{cpue})$  residuals from least squares fits of the LSSPA model to the four test datasets.

Fig. 4. Histogram of  $\log(\text{catch})$  residuals from least squares fits of the LSSPA model to the four test datasets.

Fig. 5. Histogram of  $\log(\text{cpue})$  residuals from least absolute deviation fits of the LSSPA model to the four test datasets.

Fig. 6. Histogram of  $\log(\text{catch})$  residuals from least absolute deviation fits of the LSSPA model to the four test datasets.

### Least Squares



### Least Absolute Deviations

