Application of high-breakdown robust regression to tuned stock assessment models

Victor R. Restrepo
Rosenstiel School of Marine and Atmospheric Science, Cooperative Unit for Fisheries Education and Research
University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149
E-mail address: vrestrepo@rsmas.miami.edu

Joseph E. Powers
National Marine Fisheries Service, Southeast Fisheries Science Center
75 Virginia Beach Drive, Miami, Florida 33149

Abstract.–Many tuned assessment models, such as sequential population analysis and nonequilibrium production models, are cast in the form of least-squares minimization routines. It is well known that outliers can substantially alter the results of least-squares methods. Indeed, in the process of conducting stock assessments, much time and effort are often spent in discussing the merits of individual data points and in evaluating the impact that including or excluding them has on the perceived stock status. Unfortunately, straight-forward statistical tests for detecting outliers have been developed only for univariate statistics or for the simplest of linear models and are generally useful to test for a single outlier only. In this paper, we apply a high-breakdown robust regression technique, least trimmed squares, to two assessment models using North Atlantic swordfish and West Atlantic bluefin tuna as examples. We illustrate how robust regression can be used as an initial step in statistically detecting outliers before the more efficient least-squares minimization can be used.

Manuscript accepted 30 July 1996.
Fishery Bulletin 95:149–160 (1997).

Tuned stock assessment models are statistical methods that analyze time series of fishery catch data in conjunction with auxiliary information (indices of relative abundance, fishing effort, etc.) to yield estimates of stock abundance and exploitation rates over time. Such methods are widely used today by stock assessment working groups throughout the world because they provide an objective and statistically defensible way to assess the status of stocks and to derive management advice. The two primary methods are sequential population analysis (SPA: Fournier and Archibald, 1982; Deriso et al., 1985; Pope and Shepherd, 1985; Kimura, 1989; Methot, 1990; Powers and Restrepo, 1992; Gavaris[1]) and nonequilibrium production models (Pella and Tomlinson, 1969; Hilborn, 1990; Hilborn and Walters, 1992; Prager, 1994). SPA's are typically age structured and production models are not, although there are exceptions to this generalization in the references just cited. Both types of methods, however, share the commonality of often being cast as nonlinear least-squares minimization problems.

Despite efforts to standardize all steps involved in a stock assessment (from data collection, preparation of model inputs, to running the models), stock assessments are rarely automated and, more often than not, generate controversy. In our experience with different fora, a common cause for controversy is as follows: various data sets are presented to a working group and then the group collectively decides on the sets of data and model assumptions to be used. The consensus selection is typically termed the "base case." Individual data points are then scrutinized for exclusion from further analyses to determine the robustness of the overall assessment to the sensitivity changes. This partial "sensitivity analysis" can, in practice, be undesirable because perceptions of what results ought to be like may influence which data or data points are scrutinized and thus generate controversy; not every working group participant has the same perception. The lack of an a priori objective selection process could lead working groups astray (Restrepo and Powers, 1995). A so-

[1] Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29, 12 p. Biological Station, Department of Fisheries and Oceans, St. Andrews, New Brunswick, Canada EOG 2X0.

lution to this problem lies in a method that would objectively identify—and deal with—"outliers."

Statistical tests have been developed for identifying outliers (see Barnett and Lewis, 1994), but most of the straight-forward approaches can only deal with a few outliers in univariate analyses or in linear regression. High-breakdown robust regression methods (Rousseeuw, 1984; Rousseeuw and Leroy, 1987) hold promise for addressing the issue, as suggested by several recent papers in fisheries literature (e.g. Chen et al., 1994; Chen and Paloheimo, 1994). The goal of high-breakdown robust regression is to provide model estimates that are insensitive to contamination (up to 50%) by outliers and thus will serve to identify outlying observations. However, most robust regression applications in fisheries science (Chen et al., 1994; Chen and Paloheimo, 1994) and in statistics literature have been developed for linear problems (but, see Stromberg, 1993). In this study we seek to illustrate the application and usefulness of this tool by using two nonlinear examples: a nonequilibrium production model for North Atlantic swordfish, *Xiphias gladius,* and a sequential population analysis for West Atlantic bluefin tuna, *Thunnus thynnus.* Both stocks are assessed by the Standing Committee on Research and Statistics (SCRS) of the International Commission for the Conservation of Atlantic Tunas (ICCAT). The analyses presented here are illustrative and are not intended to replace those of the SCRS.

Methods

Assessment models

Assuming a normal (Gaussian) error structure, the typical tuned assessment method minimizes the squared deviations (residuals, r) between observed and predicted indices of abundance:

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n_i} (I_{ij} - \hat{I}_{ij})^2 = \min \sum_{i=1}^{m} \sum_{j=1}^{n_i} (r_{ij}^2), \qquad (1)$$

for m indices, each with n_i observations. The prediction of each index, $\hat{I}_{ij}$, comes from a population model, such as a surplus production model or a sequential population analysis. Alternatively, the minimization can be made in terms of observed and predicted catches or in terms of observed and predicted fishing effort. Note that some maximum-likelihood approaches do not make the normal error assumption (e.g. Fournier and Archibald, 1982); we focus on those approaches that are in a least-squares framework or that can be transformed to one, which include iteratively reweighted least squares and some forms of maximum likelihood.

In this paper, we give robust regression examples using two population models. A detailed explanation of these methods is beyond the scope of this paper and readers are referred to the citations given below. The surplus production model corresponds to a Schaeffer (logistic) form, fitted as nonequilibrium time series by using the continuous time method presented by Prager (1994). This method estimates parameters describing the carrying capacity, rate of intrinsic population growth, initial biomass, and catchability coefficients that best explain observed time series of relative abundance according to the criterion in Equation 1. The sequential population analysis corresponds to a tuned virtual population analysis method known as ADAPT, an age-structured assessment framework popular in the east coast of North America. Details on ADAPT can be found in Powers and Restrepo (1992, 1993), Punt (1994), and Gavaris.[1] ADAPT estimates age-specific fishing mortality rates in the last year of data and catchability coefficients that satisfy Equation 1, while forcing cohorts to conform to exponential survival through time:

$$N_{a+1,y+1} = N_{a,y} e^{-z_{a,y}},$$

where N denotes stock size in numbers, Z denotes instantaneous total mortality, and a and y are subscripts for age and year.

Data sets

The data set used with the nonequilibrium production model is for North Atlantic swordfish as employed by ICCAT in its 1994 assessment (ICCAT, 1995). This data set consists of total landings (in weight) for the period 1950–93 and of a single standardized longline series of catch per unit of effort (CPUE, used as a measure of relative abundance), spanning the period 1963–93 (Table 1). After a series of sensitivity tests, ICCAT assumed in its "base case" analysis that the initial biomass in 1950 was a known quantity, equal to 0.875 times the stock's carrying capacity. Thus, 3 parameters were estimated: carrying capacity, intrinsic rate of growth, and a constant of proportionality (q) relating the series of relative abundance (X) to absolute biomass units (B). The minimization of Equation 1 was done in log scale, i.e. $I_{ij} = \ln(X_{ij})$ and $\hat{I}_{ij} = \ln(qB_j)$.

The data for the SPA is for West Atlantic bluefin tuna, also as employed by ICCAT in its 1994 base case assessment (ICCAT, 1995). It consisted of catch

Table 1

North Atlantic swordfish, *Xiphias gladius*, data used for the nonequilibrium production model (from ICCAT, 1995). Relative abundance is in Kg/1,000 standard hooks, standardized from Canadian, Japanese, Spanish, and U.S. longliners. t = metric tons.

Year	Landings (t)	Relative abundance	Year	Landings (t)	Relative abundance
1950	3,646	—	1973	6,001	—
1951	2,581	—	1974	6,301	—
1952	2,993	—	1975	8,776	421.69
1953	3,303	—	1976	6,587	353.66
1954	3,034	—	1977	6,352	393.92
1955	3,502	—	1978	11,797	649.61
1956	3,358	—	1979	11,859	338.57
1957	4,578	—	1980	13,527	430.69
1958	4,904	—	1981	11,138	310.18
1959	6,232	—	1982	13,155	356.96
1960	3,828	—	1983	14,464	287.88
1961	4,381	—	1984	12,753	286.12
1962	5,342	—	1985	14,348	265.94
1963	10,189	1,258.10	1986	18,447	255.54
1964	11,258	467.29	1987	20,234	217.30
1965	8,652	294.86	1988	19,614	207.62
1966	9,338	273.50	1989	17,299	196.90
1967	9,084	320.22	1990	15,865	199.20
1968	9,137	269.55	1991	15,224	194.02
1969	9,138	233.95	1992	15,593	182.55
1970	9,425	274.25	1993	16,977	172.27
1971	5,198	—			
1972	4,727	—			

at age from 1970 to 1993 for ages 1 to 10+ (Table 2), and of 7 indices of relative abundance assumed to track different segments of the population (Table 3; see Fig. 4). A number of assumptions were made and these can be found in Appendix BFTW-2 of ICCAT (1995). The parameters estimated were 7 constants of proportionality relating each index of relative abundance to absolute biomass or numbers and 4 fishing mortalities in 1993 (for ages 2, 4, 6, and 8).

We reiterate that we chose the same data sets and model structures as those in ICCAT (1995) for illustrative purposes. It may be worthwhile to investigate the results of robust regression techniques applied to alternative data (e.g. indices obtained with a different standardization procedure) or to formulations (e.g. different assumptions about known quantities and other constraints).

Robust regression

Several robust minimization criteria discussed in Rousseeuw and Leroy (1987) have been applied to fisheries data (see Chen et al., 1994). In contrast with the method of least squares, the goal of these techniques is to moderate the influence of outliers in the parameter fitting process (Eq. 1). Of particular interest to us are the so-called "high-breakdown" methods that are insensitive to up to 50% contamination by outliers, because they can effectively be used as an objective method to identify outliers.

Two high-breakdown robust regression methods are least median squares (LMS) and least trimmed squares (LTS). LMS minimizes the median of the squared residuals and LTS minimizes the sum of the lowest xn squared residuals, where x is a fraction (less than 1.0 to 0.5) defined by the user. The results of an LMS regression and an LTS regression with a 50% trim are essentially very similar, although the LTS one is statistically more efficient (Rousseeuw and Leroy, 1987). In our initial experimentation with fisheries assessment models, we found that the LTS minimum was somewhat easier to find (the LMS could sometimes not converge, indicating that a large number of restarts may be required). Therefore, we limited our investigation to the LTS minimization criteria discussed below. This can be either

$$LTS_1 = \sum_{i=1}^{m} \min \sum_{j=1}^{n_i/2+1} (r^2)_{j:n_i} \tag{2}$$

or

$$LTS_2 = \min \sum_{i=1}^{m} \sum_{j=1}^{n_i/2+1} (r^2)_{j:n_i}, \tag{3}$$

Table 2

West Atlantic bluefin tuna, *Thunnus thynnus*, catch at age data (in numbers) used for the sequential population analysis (from ICCAT, 1995).

	Age									
Year	1	2	3	4	5	6	7	8	9	10^+
1970	64,886	105,064	127,518	21,455	3,677	914	176	172	535	3,726
1971	62,998	153,364	38,360	46,074	672	1,673	2,109	1,350	1,133	5,957
1972	45,402	98,578	33,762	3,730	3,857	118	569	576	261	5,519
1973	5,105	74,311	30,482	7,161	2,132	1,451	953	1,544	555	4,444
1974	55,958	20,056	21,094	6,506	3,170	683	916	913	1,081	12,508
1975	43,556	148,027	8,328	11,963	821	547	317	671	1,651	9,472
1976	5,412	19,781	72,393	2,910	2,899	344	206	1,168	558	14,033
1977	1,274	22,419	9,717	32,139	4,946	3,633	957	513	1,109	13,532
1978	5,133	10,863	20,015	6,315	10,530	4,061	655	472	341	11,982
1979	2,745	10,552	16,288	14,916	3,448	3,494	2,612	599	557	12,283
1980	3,160	16,183	11,068	8,881	2,866	2,982	5,533	3,454	1,061	12,213
1981	6,087	9,616	16,541	5,244	6,023	3,721	2,884	3,211	2,764	10,621
1982	3,528	3,729	1,654	498	342	751	477	519	896	3,077
1983	4,173	2,438	3,268	894	866	911	1,402	1,353	1,039	5,628
1984	868	7,504	1,848	2,072	2,077	1,671	594	759	1,091	4,574
1985	568	5,523	12,310	2,814	4,329	4,019	1,024	612	698	5,603
1986	563	5,939	7,135	3,442	1,128	1,726	931	520	345	5,335
1987	1,513	13,340	9,137	5,491	4,385	2,318	1,566	1,251	1,014	3,856
1988	4,850	9,149	11,745	3,933	4,144	4,220	2,258	1,631	1,600	4,555
1989	787	12,877	1,679	3,815	1,713	2,082	2,677	1,864	1,461	5,356
1990	2,368	4,238	17,958	1,947	2,747	1,825	1,629	2,388	1,522	4,253
1991	3,327	14,533	10,761	2,924	1,650	2,166	2,347	1,946	1,915	4,485
1992	420	5,985	1,997	711	1,425	737	1,916	1,870	1,323	4,383
1993	329	1,130	5,215	3,689	2,089	1,883	1,598	2,456	1,479	2,922

where the notation $j{:}n_i$ indicates that the squared residuals are sorted in ascending order from $j=1$ to n_i; note that $n_i/2+1$ is actually an integer value equal to $n_i/2$ when n_i is even and equal to $n_i/2+1$ when n_i is odd. Equations 2 and 3 are two different minimization objectives that differ in the way they treat multiple series of relative abundance data. In Equation 2, the trimmed sums of squared residuals are computed separately for each index and then added to the objective function being minimized. Thus, the individual indices are de facto given equal weighting. In Equation 3, the trimming is done over all available data points, regardless of which relative abundance series they belong to. Thus, the LTS_1 formulation forces each available series to contribute to the objective function, whereas the LTS_2 formulation could plausibly eliminate indices that fit very poorly in comparison with the others. An analogous distinction can be made for the LS fit by giving either equal weight to all available data series (as in Eq. 1) or by assigning weights to each series in proportion to their mean squared errors. The latter has often been accomplished by means of iterative reweighting (Powers and Restrepo, 1992) or maximum likelihood (Punt, 1994).

Algorithms for high-breakdown robust regression are notoriously computation-intensive, even in the simplest univariate linear regression case (Rousseeuw, 1984; Rousseeuw and Leroy, 1987; Steele and Steiger, 1986). A typical algorithm for a linear robust regression with p parameters goes like this: For a large number of times, s, select p data points, do a least-squares regression (LS) and compute the corresponding robust objective function (e.g. sum of trimmed squares) for the complete data set. The LTS solution is given by the parameter estimates and results in the lowest robust objective function value. In the linear case, the value of s is chosen such that, for a given fraction of data contamination and a given p, at least one of the s subsamples is not contaminated (Rousseeuw and Leroy, 1987). The choice of s in the nonlinear case is not clearcut. However, in the linear case the values of s grow very rapidly with p and percent contamination; therefore many available algorithms set $s=3{,}000$ for $p>9$ (Rousseeuw and Leroy, 1987). Similar values were used here for the nonlinear case.

Algorithms for nonlinear robust regression are rare, owing partly to the increased computational

Table 3

West Atlantic bluefin tuna, *Thunnus thynnus*, relative abundance indices (from ICCAT, 1995). The larval index is in relative biomass units. while all others are in relative numbers. The numbers below each index label are the ages or range of ages that each index is assumed to represent. TL = tended line, LL = longline, RR = rod and reel, GOM = Gulf of Mexico, NWA = Northwest Atlantic.

Year	Canada TL 10^+	Japan LLGOM 10^+	Japan LLNWA 1–9	Larval GOM 8^+	US LLGOM 8^+	US RR 8^+	US RR 1–5
1974	—	1.4670	—	—	—	—	—
1975	—	1.0200	—	—	—	—	—
1976	—	0.8960	0.8134	—	—	—	—
1977	—	0.6700	1.7822	1.7704	—	—	—
1978	—	0.9350	1.4621	4.2341	—	—	—
1979	—	0.9380	0.5476	—	—	—	—
1980	—	1.5130	1.0327	—	—	—	1.2109
1981	2.3489	0.5610	1.4812	0.9575	—	—	0.1274
1982	2.1095	—	0.7121	1.1008	—	—	1.3417
1983	1.5621	—	0.5022	0.8977	—	2.4703	0.7816
1984	1.0718	—	0.8527	0.4750	—	1.0949	—
1985	0.5131	—	0.9967	—	—	1.0483	0.5366
1986	0.6157	—	0.5725	0.1897	—	0.7324	0.9995
1987	0.3991	—	1.1490	0.3236	1.7544	0.6933	1.2138
1988	0.6271	—	0.8773	1.4146	0.6842	1.3195	1.6059
1989	0.4561	—	0.7417	0.5803	1.0526	0.6808	1.3339
1990	0.2965	—	0.7754	0.3446	1.1404	0.6204	0.7331
1991	—	—	0.7523	0.2652	1.5614	0.7694	1.3277
1992	—	—	1.8813	0.4464	0.5263	0.8727	0.7968
1993	—	—	1.0675	—	0.2807	0.6981	0.9912

requirements. Although the LS solutions for the linear case (as described in the previous paragraph) can be accomplished with simple matrix manipulations, nonlinear LS solutions require iterative computations. Stromberg (1993) presented a multistage algorithm for nonlinear regression that is similar to the one outlined above, succeeded by a direct minimization of the robust objective function by using the simplex search of Nelder and Mead (1965). Building upon Stromberg's ideas, we reviewed algorithms for an LTS_1 solution to the bluefin tuna SPA. On the basis of these results and the work of Stromberg (1993), we adopted the algorithm below but acknowledge that there are many other possible fruitful options to be explored, such as "simulated annealing" (Corana et al., 1987). Our algorithm uses the fact that the simplex search of Nelder and Mead (1965) requires p+1 starting guesses, denoted by v vertices, for each of the p parameters being estimated.

1 Find the LS estimate for the entire data set. The estimates $(p)_{LS}$ are used as starting guesses for step 2.

2 Repeat s times:
  a) Set initial parameter guesses at random from within 10 times the $(p)_{LS}$ estimates from step 1.
  b) Find the LTS estimates for the complete data set by using the starting values from step 2a.
  c) Restart step 2b until the objective function (either Eq. 2 or Eq. 3) does not change appreciably.
  d) Save the parameter estimates corresponding to the $(p+1)_{LTS}$ parameter sets with the lowest objective function value.

3 Initialize the v vertices for the simplex search with the best $(p+1)$ parameter sets from the s solutions from step 2 and find the LTS estimate for the entire data set. As in step 2, carry out restarts as needed.

This algorithm is a direct robust minimization search that is initialized s times from a Monte Carlo grid centered around the LS solution. It is computationally intensive, but this seems necessary given the multi-modal nature often encountered in the LTS or LMS objective function. For this study we used s = 500. For both the swordfish nonequilibrium production model and the bluefin tuna SPA analyses, step 2 involved 5 restarts on average and thus made the total number of minimizations greater than 2,500. It should be noted that this search algorithm does not guarantee that a global minimum LTS solution is going to be found. Therefore, we favor multiple replicates and restarts so that there is some confidence that the solution is globally minimal. At this point we have no firm guidance about the tradeoffs between the number of replicates (s) and the number of restarts other than to say that replicates are probably more important than restarts. For example 500 replicates with 5 restarts seems preferable to 25 replicates with 100 restarts.

Dealing with outliers

Aside from biological or fishery considerations, statistical outliers are data points whose residuals, scaled by the dispersion of errors,

$$\frac{r_{ij}}{\sigma}$$

are far from the mean scaled residual. For the simple LS minimization (Eq. 1), the overall dispersion of the residuals is the mean squared error (MSE),

$$\sigma = \sqrt{\frac{1}{\sum_{i=1}^{m} n_i} \sum_{i=1}^{m} \sum_{j=1}^{n_i} (r_{ij}^2)}.$$

For the LTS regression, the dispersion is similarly computed as a robust measure of average dispersion (σ_i for index i in Eq. 2 or σ for all data points in Eq. 3):

$$\sigma_i = 3.7444 \sqrt{\frac{1}{n_i} \sum_{j=1}^{n_i/2+1} (r^2)_{j:n_i}}$$

for LTS_1, Eq. 2, or

$$\sigma = 3.7444 \sqrt{\frac{1}{\sum_{i=1}^{m} n_i} \sum_{i=1}^{m} \sum_{j=1}^{n_i/2+1} (r^2)_{j:n_i}}$$

for LTS_2, Eq. 3.

The constant 3.7444 is a correction factor used to achieve consistency with normal error distributions (Rousseeuw and Leroy, 1987). As a rule of thumb, Rousseeuw and Leroy (1987) suggest that absolute values of scaled residuals larger than 2.5 can be treated as statistical outliers. Owing to the small number of observations in some of our data series, we use a threshold based on the t-distribution with $\alpha = 0.01$ and $n-1$ degrees of freedom. After obtaining the LTS estimates, we carried out a new least-squares minimization excluding from the analyses any absolute scaled residuals greater than the corresponding critical value. We refer to this final result as the "trimmed LS" solution.

Results

Swordfish nonequilibrium production model

The swordfish data represent a simple example with a single index. Nevertheless, there are several very large deviations between the observed and predicted index values, when the traditional least-squares (LS) solution to the nonequilibrium production model fit is computed (Fig. 1). Indeed, these deviations have generated considerable debate (ICCAT, 1995). Therefore, we applied the robust regression techniques of the LTS algorithm and the trimmed LS method of outlier detection to this example. The LTS solution (with a 50% trim) was computed 500 times with 5 restarts each. There did not appear to be problems of multiple minima with this example because vir-

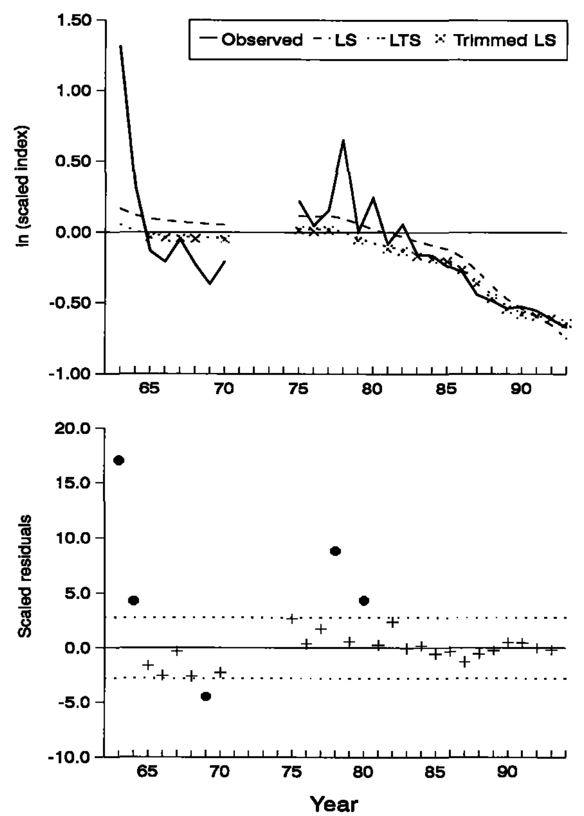

Figure 1

Biomass index values for North Atlantic swordfish using least squares (LS), least trimmed squares (LTS), and trimmed least squares (Trimmed LS). The top panel shows observed and predicted index values. The bottom panel shows the scaled residuals from the LTS fit compared with the critical value for the outlier detection criterion (dashed lines); solid circles are those data points considered to be outliers and not included in the final trimmed LS solution.

tually all of the 500 solutions converged to the same value. The outlier detection criteria identified five of the original 27 data points (19%), all of which occurred before 1981 (Fig. 1). Predicted index values with both the LTS and trimmed LS solutions were higher than the LS predictions prior to the late 1980's and lower than LS predictions in recent years (Fig. 1).

Predicted relative biomass values with either the LTS or trimmed LS solutions are higher than the initial LS solution (ICCAT, 1995), particularly in the 1990's (Fig. 2) and suggest less of a decline in the population. Absolute biomass predictions with the LTS method were generally higher than those from the initial LS solution, whereas trimmed LS solutions were lower (Fig. 2). The trimmed LS solution results in biomass levels that are lower than those in the other two methods; however, the decline over the time series is less. Biomass projections were made under two strategies: 1) a recovery strategy in which future fishing mortality rate was fixed at the value that would produce maximum sustainable yield and 2) a status quo strategy in which the fishing mortality would be fixed at the 1993 level. The LTS and trimmed LS projections indicate that both recovery and decline is not as rapid as that predicted from the initial LS solution (Fig. 2).

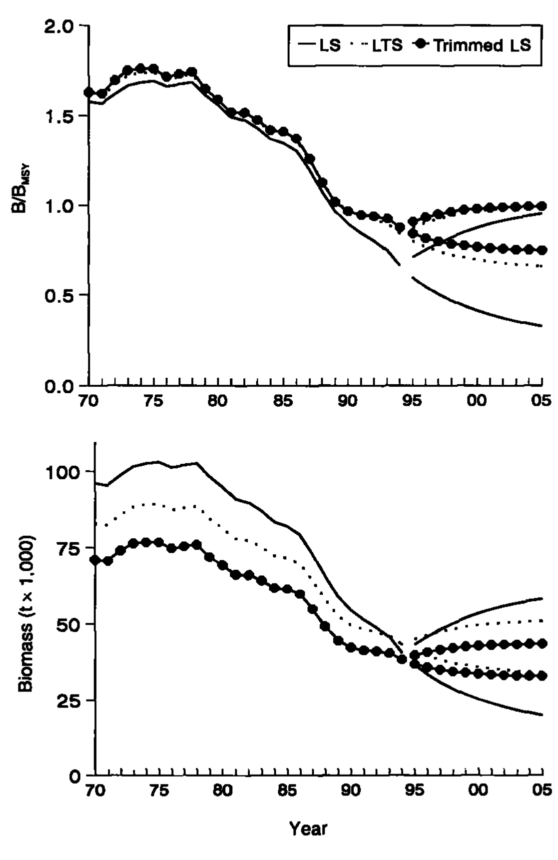

Figure 2

Predicted biomass relative to biomass at maximum sustainable yield (B/B_{MSY}, top panel) and absolute biomass (bottom panel) resulting from LS, LTS, and trimmed LS solutions. The left side of the graphs show the production model estimates. The right sides of the graphs are projections made with the fishing mortality rate at maximum sustainable yield (F_{MSY}) and with the fishing mortality rate in 1993 (F 93). Ascending limbs were projected by using F_{MSY}, descending limbs by using F 93.

The robust regression techniques applied here tend to provide a better fit to the index data points in recent years at the expense of the data points in the earlier years of the series. Indeed, several of the points identified through the outlier detection process were those data points for which there was much debate regarding variability and bias (ICCAT, 1995). However, some of the data points identified here were not identified by ICCAT (1995); therefore, we reemphasize the point that the selection of outliers should be based on objective criteria.

Bluefin tuna SPA

As mentioned before, a high-breakdown robust regression objective function can possess multiple minima. Figure 3 illustrates this point with the LTS_1 objective function plotted around ± 50% of the final estimate for one of the parameters, while all other parameter values were fixed at their solution. The figure highlights the need for an exhaustive search owing to the multimodal nature of the response surface.

Figure 4 shows the observed indices of relative abundance in the first column, the scaled residuals

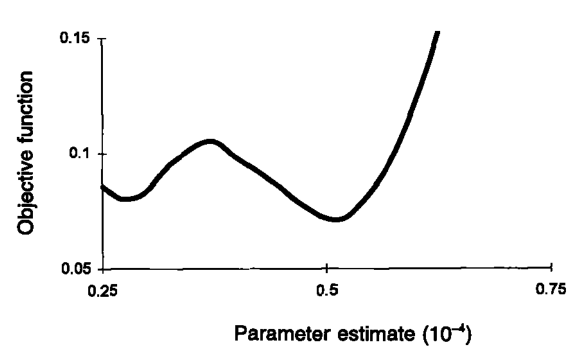

Figure 3

Trimmed squares objective function (Eq. 2) plotted around the solution for one of the parameters estimated in the bluefin tuna sequential population assessment (catchability for the U.S. rod and reel large fish index, Table 3). The plot shows that multiple local minima can occur in robust regression problems.

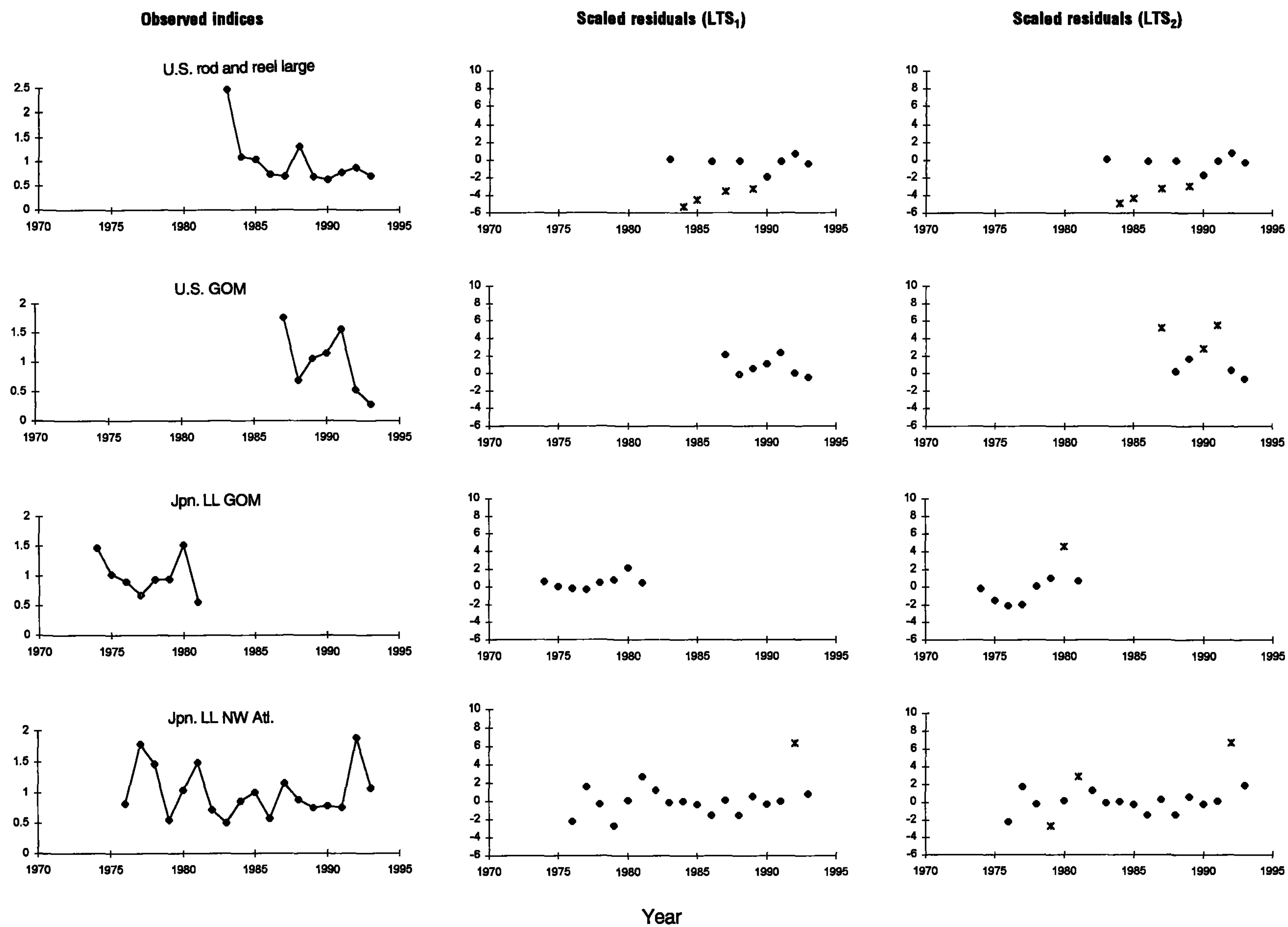

Figure 4

Results from the bluefin tuna sequential population analyses using least trimmed squares (LTS) regression. The left-hand column shows the seven available indices of relative abundance (see Table 3). The middle and right-hand side columns show the scaled residuals resulting from the minimizations with Equations 2 and 3, respectively. Crossed symbols identify statistical outliers at the 1% significance level.

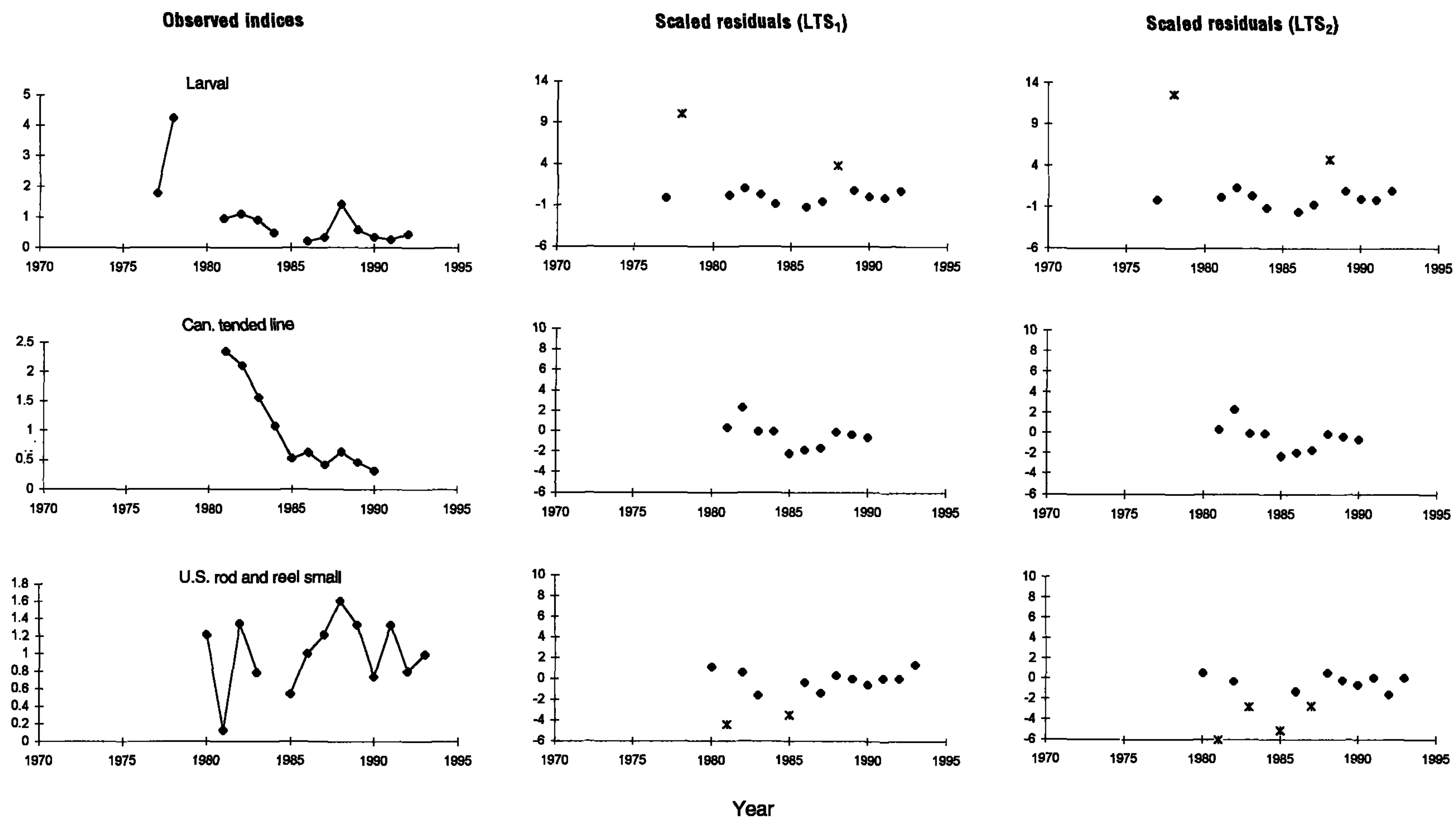

Figure 4 (continued)

from the LTS_1 fit (second column), and the LTS_2 fit (last column). The open symbols indicate which data points were identified as outliers according to the *t*-test criterion mentioned previously. The LTS_1 regression, which gives equal consideration to all index series, identified 9 outliers (11% of the total index data points). The LTS_2 approach, which gives more weight to the better-fitting series, identified the same 9 observations as outliers, and an additional 8 (21% of the total number of data). The 1978 estimate from the larval index stands out as a particularly large outlier (Fig. 4). But perhaps more importantly in terms of the effect on the SPA results, the 1992 data point for the Japanese Northwest Atlantic longline index, is also identified as a large outlier. That is, because of the convergence properties of the ADAPT approach, the more recent data tend to have a larger impact on the estimates of current stock status.

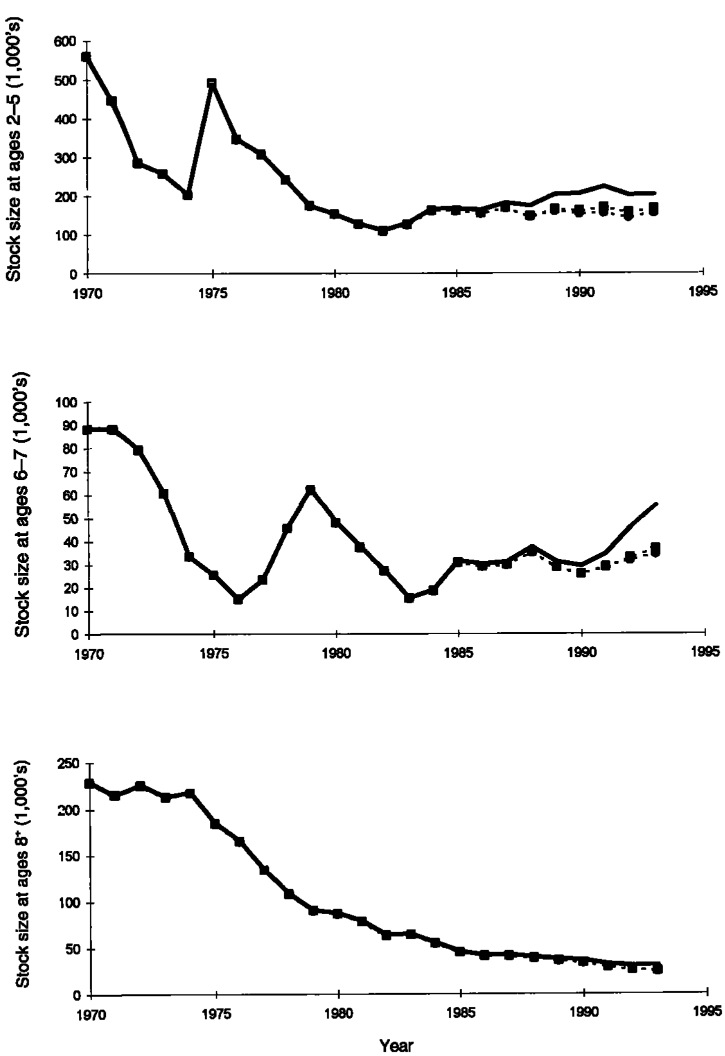

Figure 5

Bluefin tuna stock size estimates for 3 groups of ages. The solid line represents the estimates from the least-squares solution with all the available data, as in the ICCAT assessment. The dashed lines show the least squares estimates after removal of the data points identified as outliers in Figure 4. Squares = after minimization with Equation 2; circles = after minimization with Equation 3. Squares and circles overlap.

Figure 5 shows the estimated stock size trajectories for 3 age groupings that ICCAT assessments focus on: small fish (ages 2 to 5), medium fish (ages 6 and 7), and spawners (ages 8 and older). The solid line without symbols represents the initial LS solution (Eq. 1), as in the 1994 ICCAT assessment. The 2 dashed lines with symbols (virtually indistinguishable from each other) represent the final trimmed LS solutions, i.e. after removal of the outliers identified in Figure 4. Note that all the stock size estimates are identical in the first half of the time series, owing to the convergence properties of the SPA. Differences in 1990's stock size estimates before and after trimming are most notable for small and medium bluefin tuna (Fig. 5). For this example, the final trimmed LS solutions estimate lower current stock sizes (Fig. 5) and correspondingly higher current exploitation rates (not shown).

The impact that these differences in the estimates have on management recommendations can be appreciated in Figure 6, which shows a 10-year projection of the stock's spawning biomass at two levels of constant landings considered by ICCAT. These projections were made by using the same assumptions as those in the assessment (Appendix BFTW-2 in ICCAT, 1995): essentially, that recruitment is constant after a certain parental biomass level and that the 3 most recent recruitment values from the SPA are poorly estimated and are replaced by the geometric mean recruitment from past years. The top panel in Figure 6 is a projection made by assuming 2,000 metric tons (t) landings after 1993: the lower panel assumes 2,660 t landings after 1993. The solid lines represent the LS solution as in the ICCAT assessment, and the dashed lines represent the LS solutions after trimming (squares for results from the LTS_1 solution and circles for results from the LTS_2 solution). The projections made without removing outliers are optimistic and suggest a continued increase in parental biomass even at the higher level of landings. The projections made after trimming, on the other hand, are less optimistic. These suggest a more modest increase in spawning biomass at the 2,000 t level of landings, or a decline in spawning biomass after 7 years of 2,660 t landings (Fig. 6).

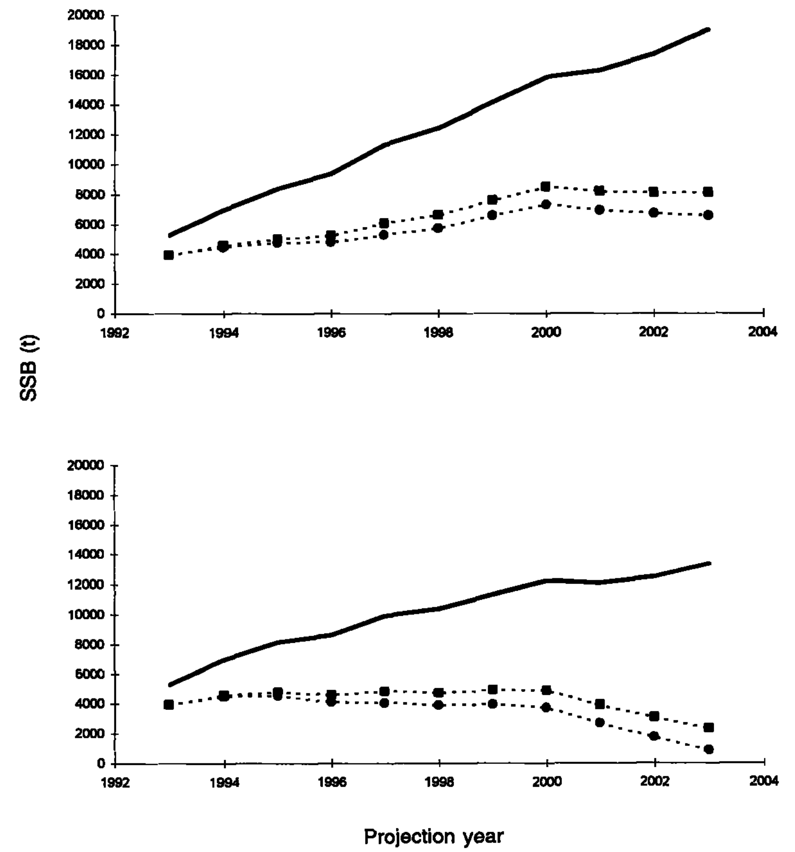

Figure 6

Bluefin tuna spawning stock size projections at two levels of constant landings: 2,000 metric tons (t) (top panel), and 2,660 t (bottom panel). Lines and symbols are the same as those in Figure 5. SSB = spawning stock biomass.

Discussion

The robust regression methods as applied to tuned population assessment models may be helpful in several ways. The methods can be used as an alternative minimization criterion to obtain estimates of the population parameters. They can also be used to identify outliers for elimination from subsequent fitting. In either case, much of the subjectivity that can enter discussions about individual data points during working group meetings would be eliminated. The latter aspect (identification and elimination of outliers) is especially useful because, after elimination of the outliers, one can then go on and conduct the normal bootstrap (Punt, 1994) or Monte Carlo (Restrepo et al., 1992) analyses used to evaluate uncertainty in the estimates. The robust regression methods could be used to screen the outliers, and then the other methods could be used to estimate variability and to project the population status under different management scenarios. Presently, computation time would preclude incorporating bootstrap or Monte Carlo techniques directly into the LTS search. Removing outliers should, also, have a moderating effect on the so-called retrospective patterns (Sinclair et al., 1990), some of which are caused by outliers in the indices (ICES, 1995).

It is important to keep in mind a point of caution when removing statistical outliers from an assessment. Observations that appear to be outliers are so in the overall context of data-model. That is, it is possible that a data point is considered as either an outlier or not, depending on the model formulation, constraints, etc. For example, if the bluefin tuna indices of abundance had been considered to be lognormally distributed instead of normally-distributed, the LTS regression may have identified more or fewer observations as outliers. A related point is that we do not advocate rushing to eliminate outliers automatically from stock assessments. Instead, a first step should be to look into reasons why such observations may seem like outliers, e.g. undetected transcription errors or environmental influences that were not accounted for in the analysis. Additionally, the outlier detection would identify candidates for sensitivity analysis in an objective manner. Instead of determining data points that are influential on the results and trying to determine if those points could be considered outliers, we are advocating the converse.

The outlier detection procedures outlined here inherently assume symmetry in the response surface. Thus, it is expected that the trimmed LS technique will provide results similar to those coming from bias correction procedures used in bootstrapping methods (e.g. Prager, 1994). Both methods assume that the underlying distributions are symmetrical and adjust the results in order to maintain that symmetry. However, if model constraints or other features of the model or data force the response surface to have an underlying (but unknown) skewed distribution, then the outlier selection process outlined here might falsely identify some data points as outliers. Conversely, the least trimmed squares (LTS) solutions make no assumptions about the shape of the response surface. Therefore, we expect that the LTS method could be robust to those situations where the distribution is skewed. Nevertheless, with judicious application, robust regression is expected to be a useful tool for evaluating and selecting data appropriate for tuning stock assessment models.

Acknowledgments

We are grateful to two anonymous reviewers for their critical review of this manuscript. Support for this study was provided through the Cooperative Unit for Fisheries Education and Research (CUFER) by National Oceanic and Atmospheric Administration Cooperative Agreement NA90-RAH-0075.

Literature cited

Barnett, V., and T. Lewis.
1994. Outliers in statistical data, 3rd ed. John Wiley and Sons, New York, NY, 584 p.

Chen, Y., D. A. Jackson, and J. E. Paloheimo.
1994. Robust regression approach to analyzing fisheries data. Can. J. Fish. Aquat. Sci. 51:1420–1429.

Chen, Y., and J. E. Paloheimo.
1994. Robust regression approach to estimating fish mortality rates with a cohort-based model. Trans. Am. Fish. Soc. 123:508–518.

Corana, A., M. Marchesi, C. Martini, and S. Ridella.
1987. Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm. ACM Trans. Math. Software 13:262–280.

Deriso, R. B., T. J. Quinn II, and P. R. Neal.
1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815–824.

Fournier, D., and C. P. Archibald.
1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195–1207.

Hilborn, R.
1990. Estimating the parameters of full age-structured models from catch and abundance data. Bull. Int. N. Pac. Fish. Comm. 50:207–213.

Hilborn, R., and C. J. Walters.
1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York, NY, 570 p.

ICCAT (International Commission for the Conservation of Atlantic Tunas).
1995. Report for the biennial period, 1994–1995, Part I (1994), vol. 2. Int. Comm. Cons. Atl. Tunas, Madrid, Spain, 283 p.

ICES (International Council for the Exploration of the Sea).
1995. Report of the working group on methods of fish stock assessments. Int. Coun. Explor. Sea Coop. Res. Rep. 199, 147 p.

Kimura, D. K.
1989. Variability, tuning, and simulation for the Doubleday-Deriso catch-at-age model. Can. J. Fish. Aquat. Sci. 46:941–949.

Methot, R. D.
1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. Int. N. Pac. Fish. Comm. Bull. 50:259–277.

Nelder, J. E., and R. Mead.
1965. A simplex method for function minimization. The Computer J. 7:308–313.

Pella, J. J., and P. K. Tomlinson.
1969. A generalized stock production model. Inter-Am. Trop. Tuna Comm. (IATTC) Bull. 13:420–496.

Pope, J. G., and J. G. Shepherd.
1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Int. Explor. Mer 42:129–151.

Powers, J. E., and V. R. Restrepo.
1992. Additional options for age-sequenced analysis. ICCAT Coll. Vol. Sci. Pap. 39: 540–553.
1993. Evaluation of stock assessment research for Gulf of Mexico king mackerel: benefits and costs to management. N. Am. J. Fish. Manage. 13:15–26.

Prager, M. H.
1994. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. 92: 374–389.

Punt, A. E.
1994. Assessments of the stocks of Cape hakes *Merluccius* spp. off South Africa. S. Afr. J. Mar. Sci. 14:159–186.

Restrepo, V. R., and J. E. Powers.
1995. Useful methods for dealing with outliers in stock assessments. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. Vol. XLIV:151–154.

Restrepo, V. R., J. M. Hoenig, J. E. Powers, J.W. Baird, and S. C. Turner,
1992. A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. Fish. Bull. 90:736–748.

Rousseeuw, P. J.
1984. Least median of squares regression. J. Am. Stat. Assoc. 79:871–880.

Rousseeuw, P. J., and A. M. Leroy
1987. Robust regression and outlier detection. John Wiley and Sons, New York, NY, 329 p.

Sinclair, A., D. Gascon, R. O'Boyle, D. Rivard, and S. Gavaris
1990. Consistency of some Northwest Atlantic groundfish stock assessments. North Atlantic Fisheries Organization (NAFO) Sci. Council Study 16:59–77.

Steele, J. M., and W. L. Steiger
1986. Algorithms and complexity for least median of squares regression. Discr. Appl. Math. 14:93–100.

Stromberg, A. J.
1993. Computation of high breakdown nonlinear regression parameters. J. Am. Stat. Assoc. 88:237–244.