Standardizing Swordfish, *Xiphias gladius*, Longline Catch per Unit of Effort Using General Additive Models

PIERRE KLEIBER and NORMAN BARTOO

Southwest Fisheries Science Center
National Marine Fisheries Service, NOAA
P.O. Box 271
La Jolla, California 92038-0271

ABSTRACT

Catch per unit of effort in the Japanese longline fishery for swordfish in the Pacific Ocean was analyzed using General Additive Models (GAM's). Twenty years of catch and effort data were standardized using various GAM's which incorporated variables of time, proportion of swordfish in the catch, and month of catch. Separate analyses were conducted for various areas.

A modeling exercise was conducted to discover whether it is effective to use the proportion of swordfish in the catch as a proxy for targeting of swordfish by the fishery. Results showed that under some systematic trends in abundance of either swordfish or bycatch species, real catch-per-unit-effort trajectories for swordfish were mis-estimated; without independent external data, this method of analysis applied to this data set has a high risk of misinterpretation.

Introduction

The Japanese longline fishery for swordfish in the Pacific Ocean has the broadest combination of spatial and time coverage of all the fisheries catching swordfish. Swordfish in the Japanese longline catch is and has been both a target species and a valuable incidental catch. The Japanese longline catch-and-effort data series is considered to be the most complete single data set covering the range of tunas and billfish in the Pacific and as such has been used to draw inferences on tuna and swordfish stocks (Bartoo and Coan, 1989).

An important feature of the Japanese longline fleet operations is that the target species varies and with it the fishing strategy, such as time of day, time of year, geographic location, depth of set, type of bait, and use of lightsticks. Such factors are highly likely to affect the power of the gear to attract and capture swordfish. However, not all of these factors are reported in the data set. This paper examines the use of proportion of swordfish and other species in the catch as an indicator of targeting, as an approach to standardizing catch per unit of effort (CPUE).

Data Selection

The data used were catch, in number of fish, and effort, in number of hooks, as reported by the Fisheries Agency of Japan (1963–82) for the fishing years 1960–80. The data were summed by $5° \times 5°$ square by month, yielding approximately 60,000 records. Two spatial stratifications of the data were used: a Pacific-wide examination including all areas of catch, and four regions selected because they had fisheries targeting swordfish and were relatively high-catch areas (Bartoo and Coan, 1989; Fig. 1).

CPUE Standardization

The reason for standardizing CPUE is to remove from the data any variation due to effects other than fish abundance. This is usually accomplished by some sort of multivariate statistical technique with CPUE as a dependent variate explained by a suite of independent variates, including time (e.g. Punsly and Deriso, 1991). To the extent that the independent variates other than time account for all variation in CPUE due to things other than variation in abundance and random noise, the time effect estimates the trajectory of abundance in time.

In our analysis we used a general additive model (GAM) technique (Chambers and Hastie, 1992). This technique allows numerical independent variables to

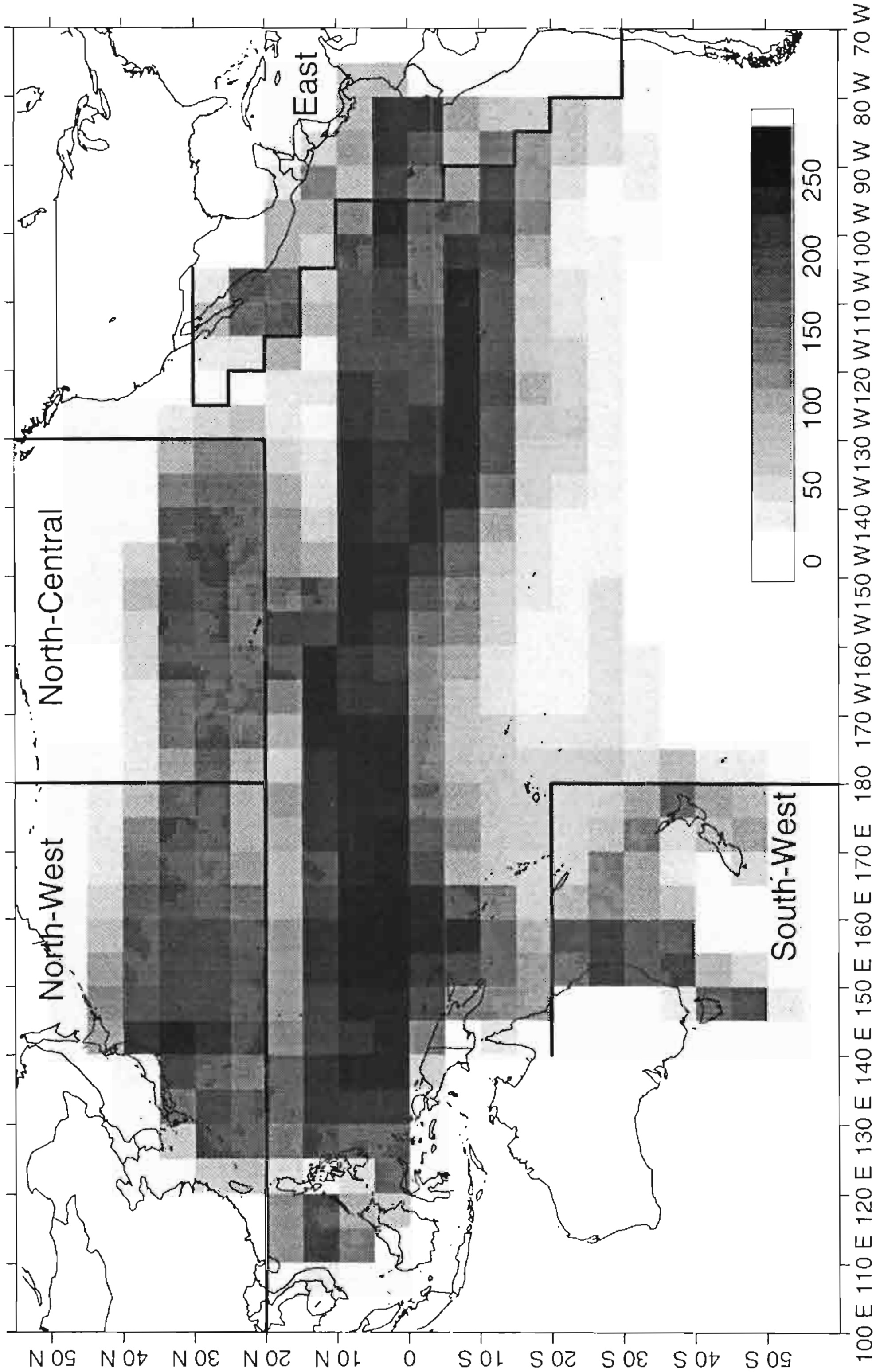


Figure 1
Four regions from which swordfish CPUE data from the Japanese longline fleet were selected for independent GAM analyses. Shading indicates density of longline sets cumulated from 1960 to 1980.

have nonlinear effects on the dependent variable as determined by a smoothing algorithm (Cleveland, 1979). Therefore the effect of an independent variable is not tied to a particular mathematical function but instead is constrained only by the smoothing algorithm.

Our simplest model, in which we entered time as a non-linear variable, is given by

Model 1: $\log(\text{CPUE} + \varepsilon) = a + \text{lo}(t)$

where ε is a small value (0.1) to handle the case where there is zero catch, and $\text{lo}(t)$ is a local regression (LOESS; Chambers and Hastie, 1992) smoothing function of time given as the year and decimal fraction thereof (e.g. 85.5 for 1 July 1985). In this and other cases the LOESS functions were of degree 1 and span parameter 0.5.

Because the data do not contain direct information on species targeted, we used a proxy variable consisting of the arcsine transform of the swordfish catch as a proportion of the total catch of tunas and swordfish. The arcsine transform of a variate is the inverse sine of the square root of the variate. It is applied to proportions to spread out the values approaching zero and one, that is, values near the ends of the range of values that a proportion can take (Sokal and Rohlf, 1969). Our second model is given by

Model 2: $\log(\text{CPUE} + \varepsilon) = a + \text{lo}(t) + \text{lo}(P)$

where P is the arcsine transform of the proportion of swordfish in the catch.

To investigate seasonal variability within years, we used a third model given by

Model 3: $\log(\text{CPUE} + \varepsilon) = a + \text{lo}(t) + \text{lo}(P) + \text{factor}(M)$

where factor(M) is the month entered as a categorical (factor) variable. To enter M as a continuous, numerical variable, we would like it to be a circular variable (i.e. with January and December adjacent to each other in the same way as all other pairs of adjacent months). However, we could not find a way to do that with the software we had to work with.

We did not progress to further models with additional independent variables, such as environmental variables, because specific, set-by-set data other than catch were not available. Broader-scale indices could have been used, such as average temperature or the southern oscillation index, but there is a danger that such variables affect swordfish abundance rather than catchability. They would therefore obscure rather than reveal the abundance signal that we wished to elucidate from the CPUE data.

Table 1

Deviance (%) in swordfish CPUE in different areas explained by the models. All tests were significant at $P<0.000001$.

		Deviance explained				
	Test	Northwest	North-central	East	Southwest	Total
Model 1	add lo(t)	2	12	6	4	0.3
Model 2	add lo(P)	66	91	82	84	80
Model 3	add factor(M)	72	93	83	88	81

Standardization Results

We first tried to detect a Pacific-wide abundance trend. Because the whole data set was too large for the software to accommodate in a single analysis, we chose a random subsample of approximately 10% of the records. Using Model 1, which had time as the only independent variate, a gentle, approximately 15-yr cycle was evident (Fig. 2A). The model explained only a small percent of the raw deviance, yet its statistical effect was still significant (Table 1). With Model 2, which added the proxy targeting variable, the temporal effect was smoothed into a steady decline from 1960 to 1980 (Figure 2B). The targeting variable P had a positive effect which leveled off at high values of P (Figure 2C). Thus, as would be expected, CPUE tends to increase with increasing proportion of swordfish in the catch. Repeated subsamplings did not change these results materially.

To allow for the possibility that swordfish abundance in the different regions would vary independently, we conducted separate analyses of 4 regions within the Pacific (Fig. 1). A large number (65%) of the longline records from the tropical regions, where swordfish is not targeted but is retained, were eliminated (Bartoo and Coan, 1989). In this case the software was able to handle all the data in each region. Using Model 1, different patterns of temporal variation were indeed apparent (Fig. 3): an overall declining trend in the northwest and north-central Pacific, a rising trend in the southwest Pacific, and a rise followed by a fall in the northeast Pacific. Under Model 2, the addition of our proxy variable for targeting changed the picture somewhat (Fig. 4). The temporal effect was flattened and the span (maximum – minimum) was reduced, particularly in the north-central and the southwest regions. Model 2 showed a more or less decreasing time trend in all

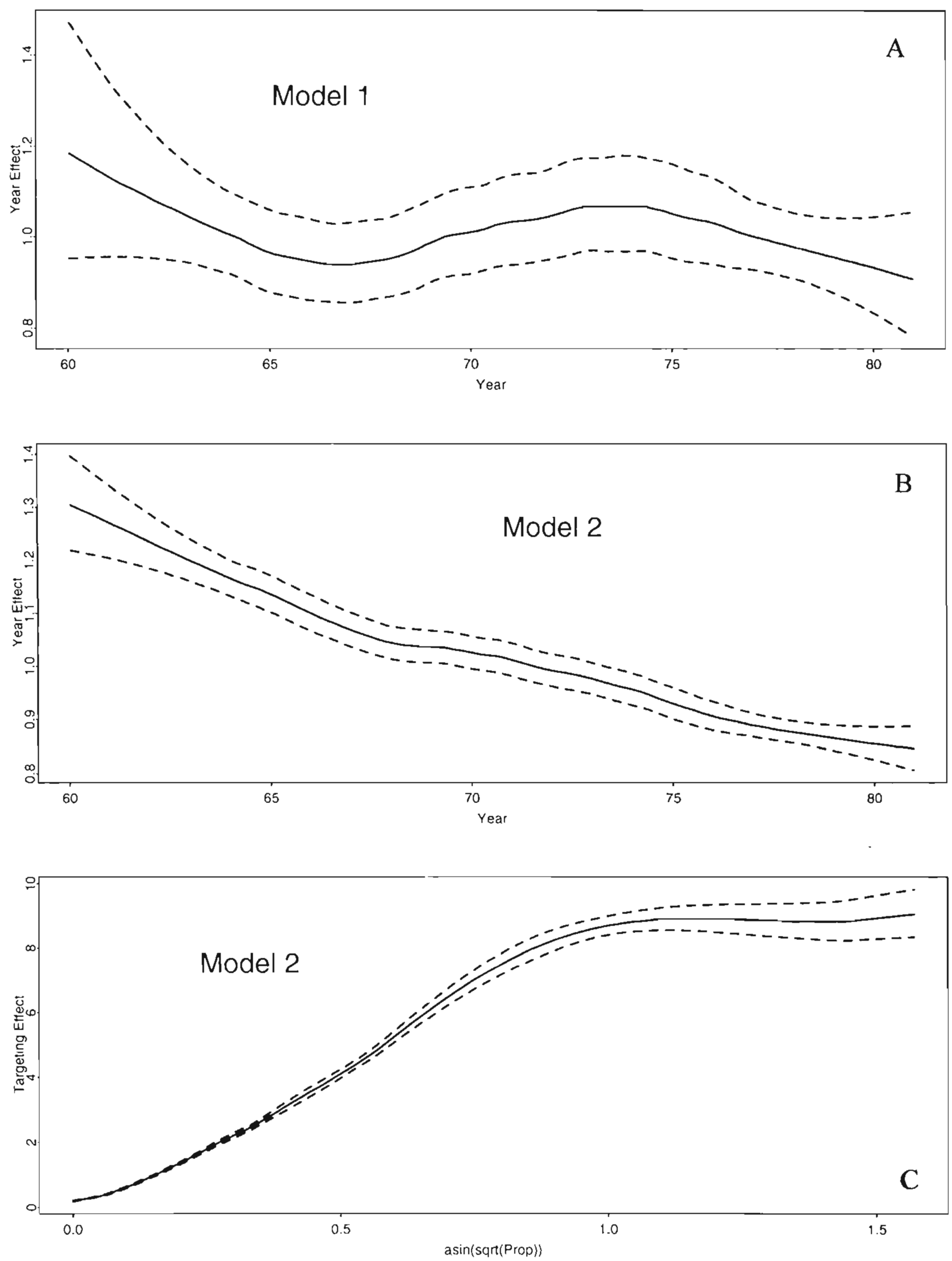


Figure 2

Results from GAM analysis of a random subsample of swordfish CPUE data from the Japanese longline fleet, 1960–80. (A) year effect with Model 1; (B) year effect with Model 2; (C) effect of proxy targeting variable with Model 2. Dashed lines indicate ± 2 standard errors.

regions from 1960 to 1980, the trend in the southwest region being the reverse of that shown in Model 1.

The functional effect of the proxy targeting variable P was monotone increasing in all regions except in the

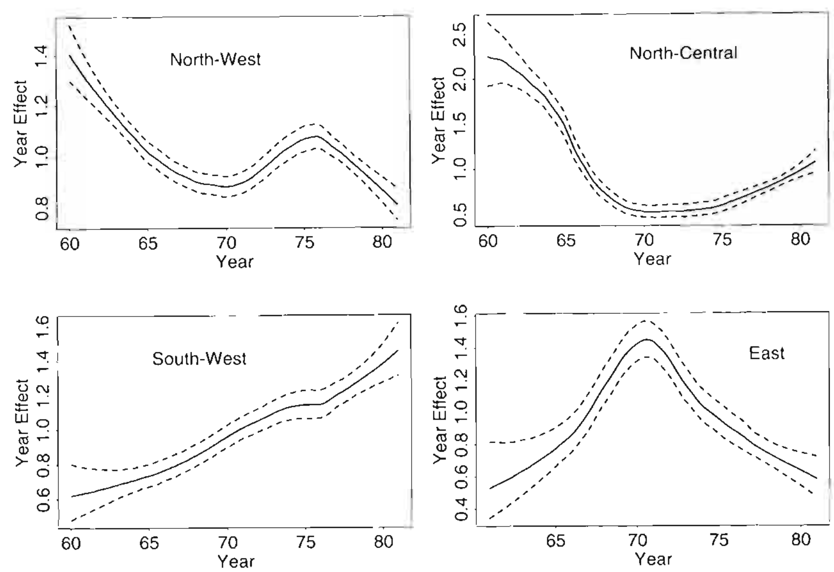


Figure 3
Year effects in four regions analyzed independently under Model 1. Dashed lines indicate ± 2 standard errors.

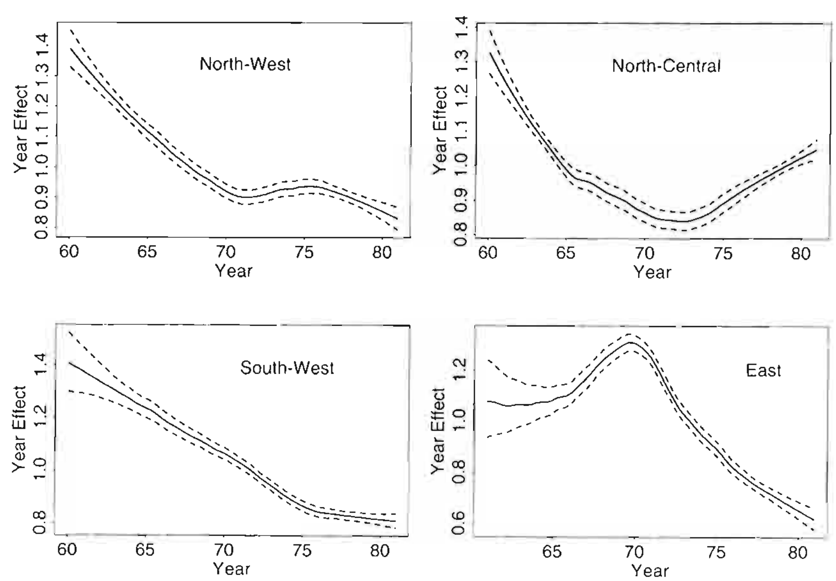


Figure 4
Year effects in four regions analyzed independently under Model 2. Dashed lines indicate ± 2 standard errors.

northwest, which showed a downward trend for high values of P (Fig. 5). In all cases the targeting effect was more pronounced over the range examined than the effect of time over 20 yr.

When residuals were plotted around the time function for Model 1 in the northwest region (Fig. 6), it was obvious that very little of the deviance in the data is explained by the model, as is true for all the regions

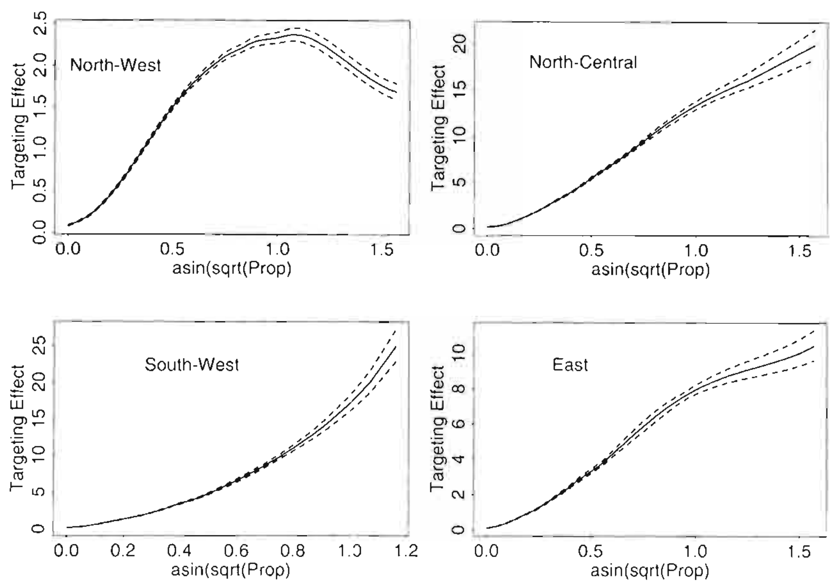


Figure 5

Effects of proxy targeting variable in four regions analyzed independently under Model 2. Dashed lines indicate ± 2 standard errors.

(Table 1). The model fit was still significant statistically, due to the large number of data points, for all regions but the southwest. With addition of the targeting variable (Model 2), much more deviance was explained by the model, between 66% and 91% (Table 1), and the span of the residuals was reduced by 3 orders of magnitude in the northwest (Fig. 6). Seasonality in the residuals was also revealed, which was not picked up by the time-effect function, lo(t). We could have attempted to capture this variability by setting the LOESS span parameter to a much smaller value than 0.5. However, we chose to let lo(t) focus on interannual variability, and added a month variable (Model 3) which accounted for most of the seasonality and reduced the span of residuals by another order of magnitude (Fig. 6). The month effect showed a minimum in the summer and a maximum in the winter in the northwest region; the basic shapes of the time-effect and targeting-effect functions in Model 3 were not appreciably different from Model 2 (Fig. 7; cf. Fig. 4, 5).

For the north-central and southwest regions, similar results to those in Figures 6 and 7 were observed in the reduction of the span of residuals when progressing from Model 1 to 2 to 3, and in the seasonality revealed by Model 2 residuals and captured by the month effect in Model 3. As would be expected, seasonality in the southwest was 6 mo. out of phase with the two northern regions. As expected for a region that spans the equator, the eastern region revealed no seasonality. Therefore in this case there was no reduction in span of residuals by Model 3 beyond that achieved by Model 2.

Simulation Modeling

An important question about our analyses is whether the proxy targeting function is really doing its job. Ideally this function would be sensitive only to the relative emphasis that fishermen are putting on catching swordfish rather than other target species. While the proportion of swordfish in the catch would seem to reflect the targeting of swordfish relative to other species, it would also be expected to reflect the abundance of swordfish relative to other species.

With a simple simulation model, we investigated this and other possible problems inherent in a proxy targeting variable that uses catch of other species. The model produces synthetic CPUE data for swordfish and for an alternate, non-target species which is a composite of all tuna species caught. In the model, 2,500 fishing sets are distributed throughout 20 yr of simulated time, the date of each set being randomly chosen from a uniform distribution. For each set a random choice is made, according to a probability level, *sprob*, which is the probability that a set targets swordfish. The value of *sprob* can either be held constant or can be varied with the date of the set. The mixture of swordfish and non-swordfish sets, therefore, is either more or less constant (being

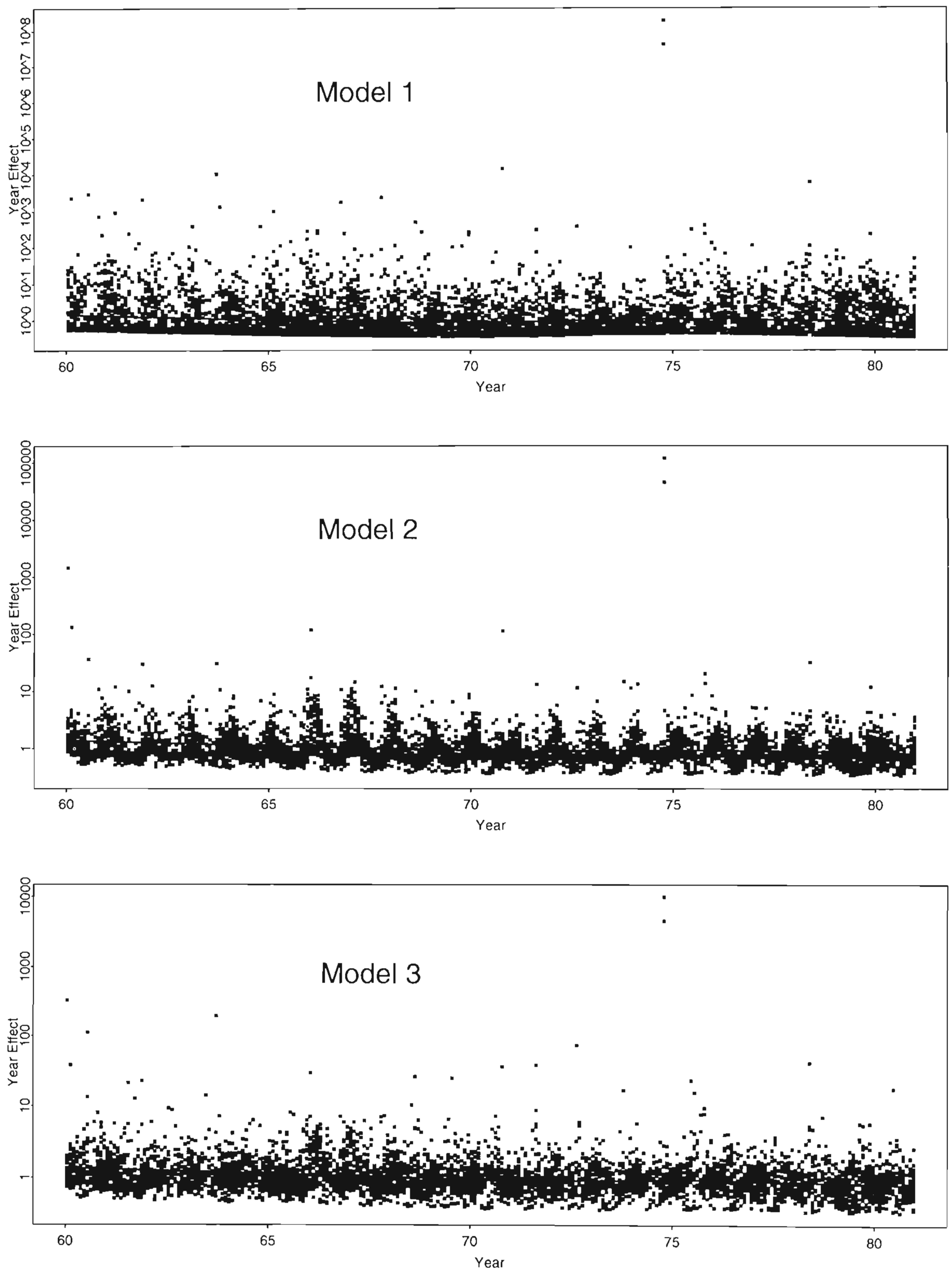


Figure 6
Time effect shown by residuals in northwest region under (from top to bottom) Models 1, 2, and 3.

subject to stochastic variation) or it has additional variation with time governed by the variation in *sprob*. The catchability of swordfish is assumed to be 100 times greater for swordfish sets than for non-swordfish sets, whereas the composite catchability of other species is assumed to be twice as great for non-swordfish sets than

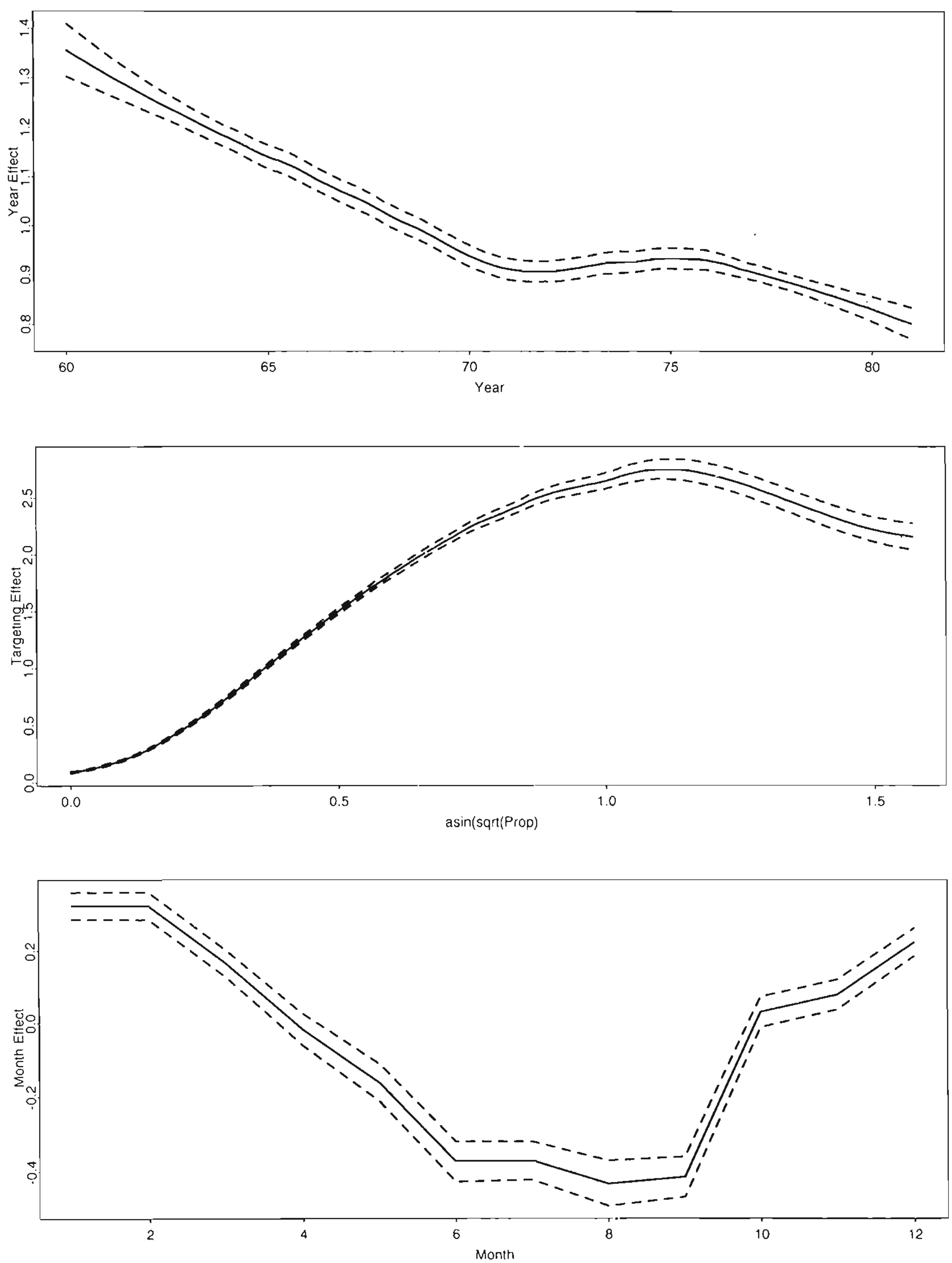


Figure 7
Year, targeting, and month effects in northwest region under Model 3. Dashed lines indicate ± 2 standard errors.

for swordfish sets. The milder response of the composite (other) species catchability to targeting reflects the assumption that some other species are likely to be more vulnerable to swordfish sets and some to non-

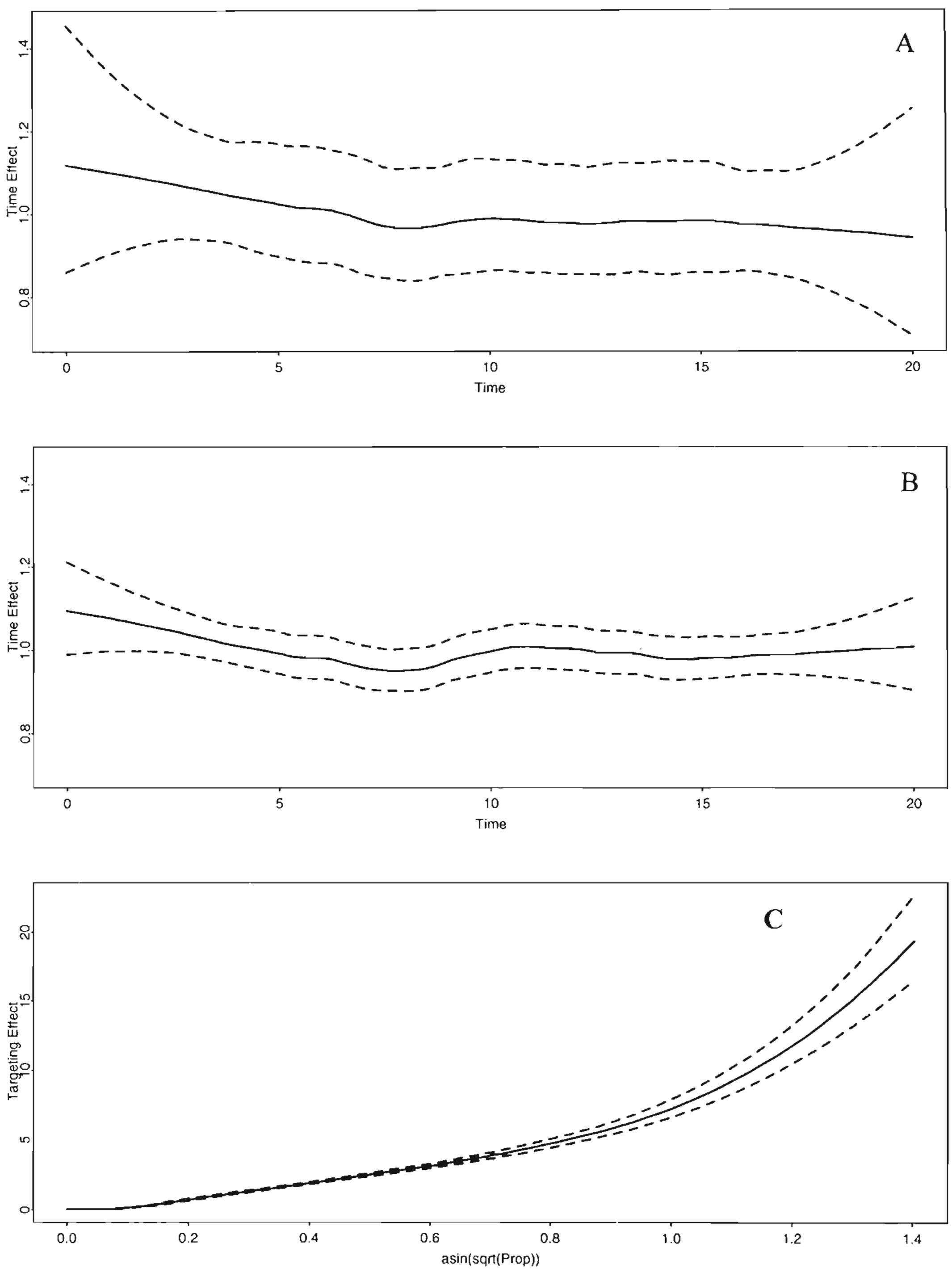


Figure 8

GAM analysis of simulated data with constant abundance of swordfish, constant abundance of other species, and constant targeting probability. (A) Year effect under Model 1; (B) year effect under Model 2; (C) effect of proxy targeting variable under Model 2. Dashed lines indicate ± 2 standard errors.

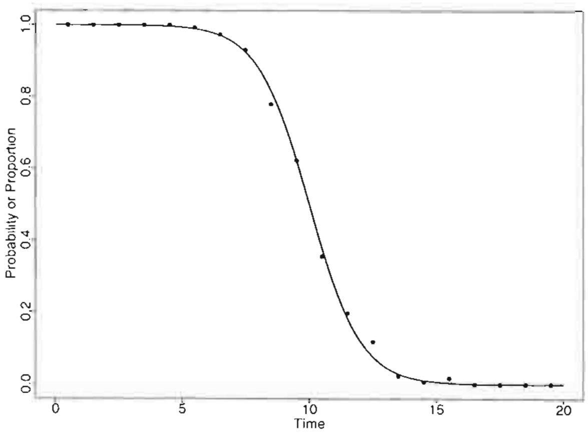


Figure 9
Targeting probability, *sprob,* as a function of time in a GAM analysis of simulated data. Points are actual proportions of sets chosen to target swordfish by 1-yr intervals in one of the simulations.

swordfish sets. Both catchabilities are subjected to lognormal stochastic error. The abundance of either swordfish or the composite (other) species can be held constant or forced to follow any arbitrary trajectory in time. The synthetic data consist of catches (catchability times abundance) of swordfish and other species for each set.

We produced synthetic data sets for a variety of combinations of variability in swordfish abundance, other fish abundance, and swordfish targeting probability, *sprob.* The data sets were then subjected to Model 1 and Model 2 GAM analyses. We did not trouble with Model 3 because we were interested in interannual rather than seasonal variability.

Figure 8 shows GAM results for a situation in which swordfish abundance, other fish abundance, and *sprob* are all held constant. Though the time effects in both Model 1 and Model 2 are not completely flat, attention to the confidence regions shows little indication of a time trend in swordfish abundance in either case. However, the width of the confidence region is reduced in Model 2, reflecting an innate correlation between the dependent variable (swordfish CPUE) and the targeting variable due to the fact that swordfish catch is used in calculating both variables.

To investigate the effect of a change in targeting we forced the targeting probability, *sprob,* to vary with time as in Figure 9. The proportion of sets chosen to target swordfish is also shown for each year. The GAM results for this situation (Fig. 10) show a dramatic drop in apparent swordfish abundance under Model 1, as would be expected. The addition of the targeting variable in Model 2 only partially accounts for the change in targeting. Thus even in this ideal situation, that is, no changes in the abundance of other species, the proxy targeting variable is only partially successful in correcting a false drop in apparent swordfish abundance.

When the abundance of other fish change over time, the problems are even worse. Figure 11 shows GAM results from a scenario in which swordfish abundance is constant, and targeting is constant, but the abundance of other species drops by a factor of 10 over the 20 yr of simulation. The Model 1 results show some variation, but within the confidence region. However, the time effect in Model 2 indicates a dramatic drop in apparent swordfish abundance. In this case the proxy variable creates a false drop in apparent swordfish abundance that did not exist without the proxy variable.

From these results and the results of many other simulations, it is evident that is it possible to simulate almost any scenario of misleading indications of trends in fish abundance from CPUE data, with or without standardization using a proxy targeting variable. Therefore it is clear that there is very limited information about targeting contained in data on the proportion of different species in the catch.

Discussion

In order to standardize swordfish CPUE, we wanted to account for variation in fishing strategy that targets fishing effort either toward swordfish or toward other species. We used the proportion of swordfish in the catch as a proxy variable for targeting, but our modeling results show that it cannot be relied upon to correct for targeting effects in the CPUE time series. Therefore the declining overall swordfish CPUE (Fig. 2) cannot be unequivocally interpreted as indicating a decline in swordfish abundance. Our analyses of separate regions eliminated much of the non-swordfish longline effort in equatorial regions, but the fisheries in those regions remained multi-species fisheries. Some measure of confidence might be taken in results for three of the regions from the fact that the patterns of variation in time are essentially identical with and without the targeting variable (Fig. 3, 4). However, given the simulation results, caution is called for, and the reversal in the time trend for the southeast region depending on the model used must be viewed with extreme caution. The fact that the targeting variable was able to greatly reduce the residual deviances (Fig. 6) could in large part be due to the inherent correlation between swordfish CPUE and the targeting variable.

To adequately analyze catch and effort data from multi-species fisheries such as the Japanese longline fleet, there is a great need for information that is directly relevant to targeting, e.g. depth of set, time of day, and use of light sticks. Provision for such data on

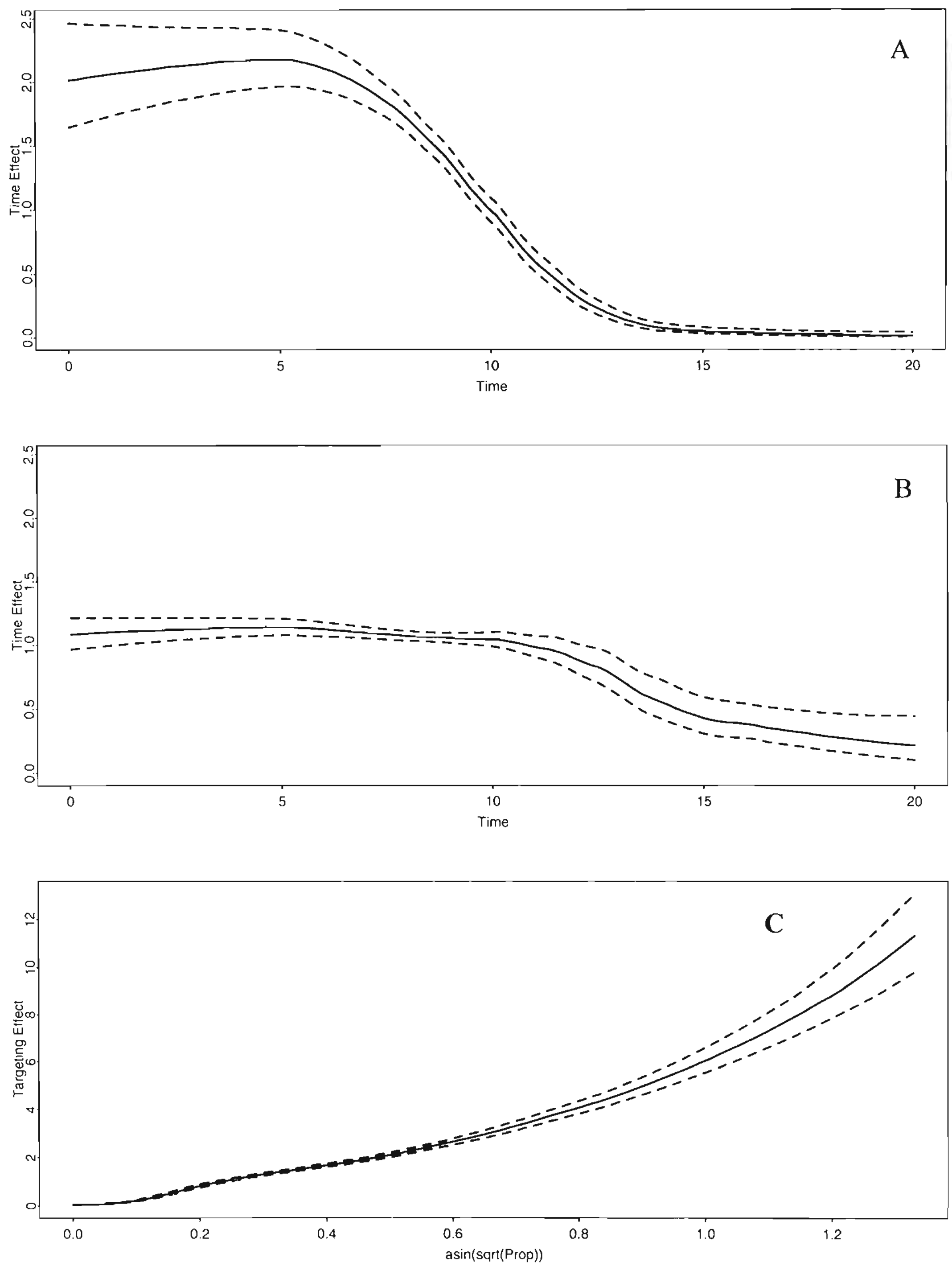


Figure 10

GAM analysis of simulated data with constant abundance of swordfish, constant abundance of other species, and declining targeting probability. (A) Year effect under Model 1; (B) year effect under Model 2; (C) effect of proxy targeting variable under Model 2. Dashed lines indicate ± 2 standard errors.

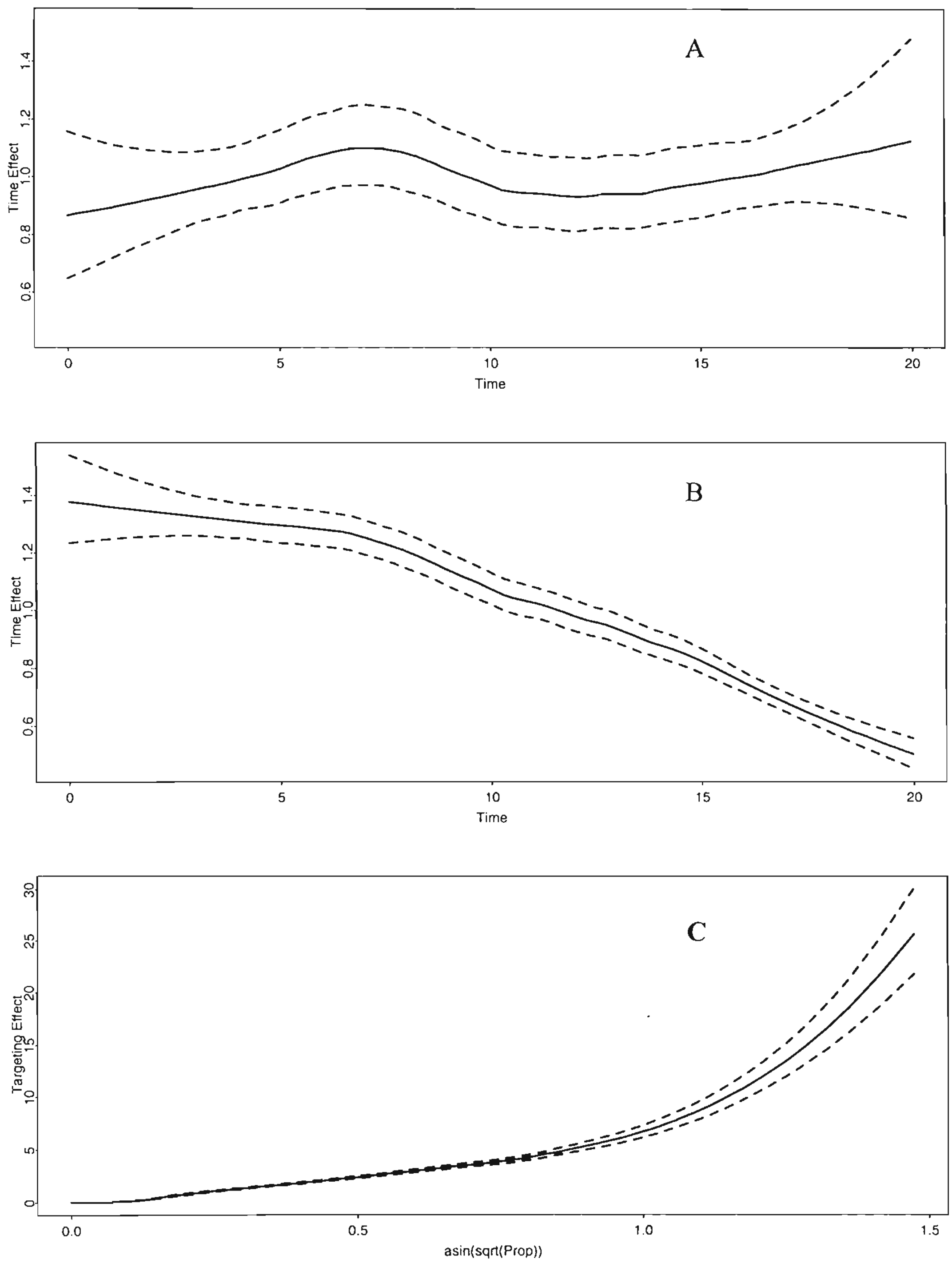


Figure 11

GAM analysis of simulated data with constant abundance of swordfish, declining abundance of other species, and constant targeting probability. (A) Year effect under Model 1; (B) year effect under Model 2; (C) effect of proxy targeting variable under Model 2. Dashed lines indicate ± 2 standard errors.

log sheets will improve the future situation, but will not fix the existing published time series. Because these time series cover extended time periods (20–30 yr) they are potentially of great value, but they would be immensely more valuable if accompanied by targeting data. It may be the case that unpublished (and not computer-encoded) data relevant to targeting is stored away in attics or basements somewhere. Though difficult and time consuming, a data "rescue" effort among fusty shoe boxes and the like might be highly rewarding.

Literature Cited

Bartoo, N. W., and A. L. Coan Jr.
1989. An assessment of the Pacific swordfish resource. *In* R. H. Stroud (ed.), Planning the future of billfishes: research and management in the 90's and beyond. Part I, Fishery and stock synopses, data needs and management, p. 137–151. Natl. Coalition Mar. Conserv., Savannah, GA.

Chambers, J. M., and T. J. Hastie.
1992. Statistical models in S. Wadsworth and Brooks, Pacific Grove, CA, 608 p.

Cleveland, W. S.
1979. Robust local weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. 74:829–836.

Fisheries Agency of Japan.
1963–82. Annual reports of effort and catch statistics by area on Japanese tuna longline fishery, 1960–1980. Dep. Res. Develop., Fish. Agency of Japan.

Punsly, R. G., and R. B. Deriso.
1991. Estimation of the abundance of yellowfin tuna, *Thunnus albacares*, by age groups and regions within the eastern Pacific ocean. Inter-Am. Trop. Tuna Comm. Bull. 20:99–131.

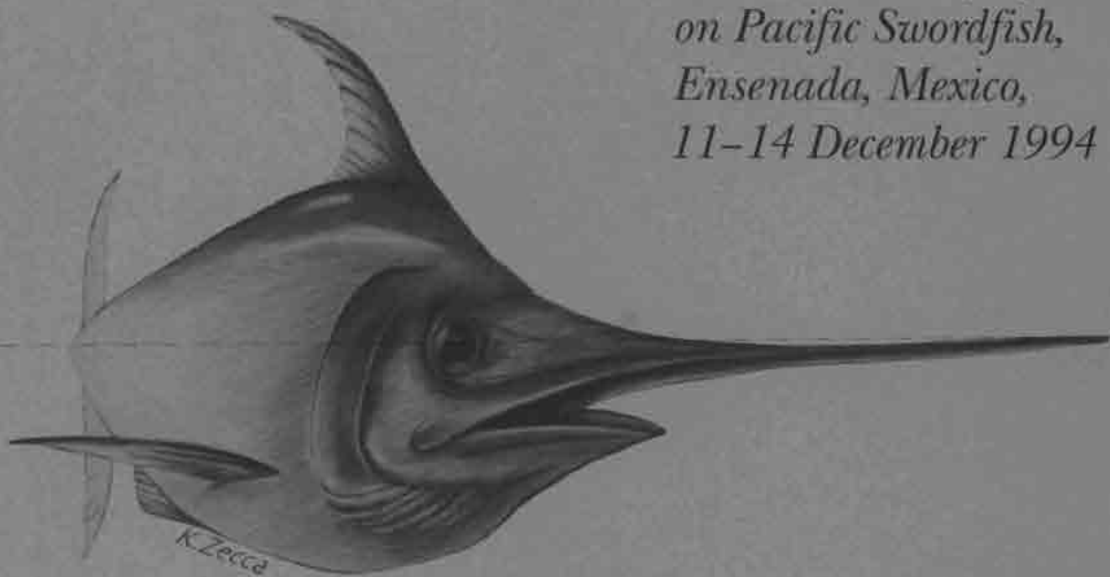
Sokal, R. R., and F. J. Rohlf.
1969. Biometry. Freeman, San Francisco, CA, 776 p.

NOAA Technical Report NMFS 142

December 1998

Biology and Fisheries of Swordfish, *Xiphias gladius*

Papers from the International Symposium on Pacific Swordfish, Ensenada, Mexico, 11–14 December 1994



Edited by

Izadore Barrett
Oscar Sosa-Nishizaki
Norman Bartoo

U.S. Department of Commerce

NOAA Technical Report NMFS 142

A Technical Report of the *Fishery Bulletin*

Biology and Fisheries of Swordfish, *Xiphias gladius*

Papers from the International Symposium on Pacific Swordfish, Ensenada, Mexico, 11–14 December 1994

Izadore Barrett
Oscar Sosa-Nishizaki
Norman Bartoo (editors)

Cover illustration by Katherine Zecca

December 1998

U.S. Department of Commerce
Seattle, Washington

Suggested reference

Barrett, Isadore, Oscar Sosa-Nishizaki, and Norman Bartoo (eds.). 1998. Biology and fisheries of swordfish, *Xiphias gladius*. Papers from the International Symposium on Pacific Swordfish, Ensenada, Mexico, 11-14 December 1994. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, 276 p.

Purchasing additional copies

Additional copies of this report are available for purchase in paper copy or microfiche from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161; 1-800-553-NTIS; http://www.ntis.gov.

Proprietary products

The National Marine Fisheries Service (NMFS) does not approve, recommend, or endorse any proprietary product or proprietary material mentioned in this publication. No reference shall be made to NMFS, or to this publication furnished by NMFS, in any advertising or sales promotion which would indicate or imply that NMFS approves, recommends, or endorses any proprietary product or proprietary material mentioned herein, or which has as its purpose an intent to cause directly or indirectly the advertised product to be used or purchased because of this NMFS publication.