STANDARDIZATION OF U.S. RECREATIONAL FISHING SUCCESS FOR SAILFISH *(ISTIOPHORUS PLATYPTERUS)* 1973-92, USING GENERAL LINEAR MODEL TECHNIQUES

SCRS/1993/100

Col.Vol.Sci.Pap. ICCAT, 42 (2) : 346-352 (1994)

Farber, M.I.

Southeast Fisheries Science Center, National Marine, Fisheries Service,
75 Virginia Beach Drive, Miami, Florida 33149, USA

SUMMARY

This report presents the results of an analysis of U.S. recreational fishing success for sailfish over the period 1973-1992. The data base included only tournaments that were directed at sailfish. A General Linear Model (GLM) was applied to generate indices of abundance using numbers of sailfish caught per unit of fishing time (CPUE). The original boat-trip records were aggregated into day-location records for the purpose of this analysis. These records were then weighted according to the number of boat-trips they included. Factors considered in the analysis were year, region, bi-month (wave) period, and appropriate two-way interactions. Area-specific and combined regions r^2 values for CPUE were very similar at approximately 0.3.

RESUME

Ce rapport présente les résultats d'une analyse de l'essor de la pêcherie sportive des Etats-Unis durant la période de 1973-92. La base de données comprenait uniquement les championnats visant le voilier. Un modèle linéaire généralisé (GLM) a été appliqué pour élaborer des indices d'abondance en utilisant un nombre de voiliers capturés par unité de temps de pêche (CPUE). Les registres d'origine de navires-sorties ont été ajoutés aux registres de jour-lieu pour mener à bien cette analyse. Ces registres ont ensuite été pondérés suivant le nombre de navires-sorties. Les facteurs pris en considération dans l'analyse étaient, l'année, le secteur, la période bi-mensuelle (vague) et les intéractions appropriées dans deux sens. Les valeurs r_2 spécifiques de zone et secteurs combinés pour la CPUE étaient autour de 0.3.

RESUMEN

Este informe presenta los resultados de un análisis del éxito obtenido por la pesquería deportiva de pez vela de Estados Unidos durante el período 1973-1992. La base de datos incluía sólo concursos dirigidos al pez vela. Se aplicó un modelo lineal generalizado (GLM) para generar índices de abundancia utilizando números de pez vela capturados por unidad de tiempo de pesca (CPUE). Para realizar este análisis, se añadieron los registros originales de barco/marea a los de día/situación geográfica. Después, estos registros fueron ponderados de acuerdo con el número de barcos/mareas que incluían. Los factores que se tuvieron en cuenta en el análisis fueron: año, región, período (onda) bimensual, y las adecuadas interacciones en dos sentidos. Los valores r^2 para CPUE específicos de las áreas, y regiones combinadas fueron muy similares, de aproximadamente 0.3.

INTRODUCTION

Indices of abundance for sailfish (*Istiophorus platypterus*) were first established from U.S. recreational fishing data from the western North Atlantic by Beardsley and Conser (1981) for the period 1972-1978. Only nominal rates of recreational fishing success have been reported since that investigation, the most recent being Prince et al. (1990). Catch-rates must be standardized to be considered as a potentially reliable indicator of changes in relative abundance. In this study, a general linear model (Farber et al. 1992; Browder and Prince 1990, and others) is used to standardize annual sailfish recreational catch-rates for two main factors -- region and time-period.

Standardized CPUEs are used as indices of abundance in stock assessment analyses using virtual population analysis (VPA) and surplus production modeling. Standardized estimates from the present study are being applied to non-equilibrium stock-production models by Jones and Farber (1993).

Twenty years of data were analyzed from two regions within the study area, as well as the combined-regions. The separate regional models provide additional perspective because, although separate regional estimates can be obtained from a combined-regions model, they will all display the same pattern over time. By modeling the regions separately, one can determine how the regional trends differ from each other and from the combined-regions trend. This information is useful in evaluating the trends from the viewpoint of stock assessment.

The recreational billfish survey of the Southeast Fisheries Science Center, National Marine Fisheries Service (NMFS), was the source of the data. Beardsley and Conser (1981) described the survey and discussed the potential for obtaining indices of abundance from survey data. A comprehensive review of this survey (for 1972-1986) was presented by Prince et al. (1990) and not repeated here. The scope and direction of the survey has remained essentially the same throughout the 20-yr period, however, certain aspects of data collection methods were changed in 1986. Budgetary constraints reduced the number of tournaments that could be attended by NMFS personnel. In place of this coverage, the survey team began to actively solicit tournament data on a volunteer basis from cooperative tournament personnel. This proved to be a cost-effective way of obtaining information from tournaments that NMFS representatives were not able to attend and even greatly expanded coverage in certain areas. The data collected were not as comprehensive as that obtained by NMFS representatives but are considered to be sufficiently detailed for this analysis.

DATA AND METHODS

Much of the information involved in defining the source of the data, the preparation of the data, and the analysis techniques are detailed in Browder and Prince (1990), and generally are not repeated in this report unless necessary for continuity, clarity, or comparison. The methods used in the present paper are identical in many respects to those of Browder and Prince (1990) and Farber et al. (1992) in their analyses of blue marlin and white marlin. The primary difference is that in this analysis only data from tournaments targeting sailfish are included, thereby limiting regions of interest and main effects defined in the general linear model (GLM).

Preparation of the Data

Individual records for each fish were first combined into one record for each trip. These records were further combined into a single record for each day and location. The rationale for combining the data into day-location records was presented in Browder and Prince (1990). Two geographic regions were defined that contained virtually all[1] tournaments that could be designated as sailfish tournaments: (1) Florida East Coast and (2) Florida Keys. The Florida East Coast (FEC) refers to the area of Fort Pierce south through Miami. The Keys are a series of connected islands located south of Miami, from Key Largo to Key West. There is a relatively high concentration of sailfish in this area, with sailfish often found close to shore (Carter and Farber 1993). To reduce the problem of small sample sizes within a time-area stratum, two-month periods ("waves") were defined to take into account the seasonal variation in the data (as done by previous investigators), with wave 1 defined as January-February.

Data Analysis

Exploratory Data Analyses:

The nominal data, consisting of the sampled number of sailfish caught with their associated sampled fishing effort (Figures 1a, 2a, and 3a), and the mean annual values of CPUE in number of fish per 100 hrs (Figures 1b, 2b, and 3b), were plotted to provide a view of the time-series before any manipulation of the data. Mean values from nominal data can be biased by the distribution of both fishing effort and sampling effort each year. Annual variation in data distribution can produce a spurious pattern of change in mean values over time. The GLMs are designed to produce estimates that are not affected by changes in data distribution and therefore might provide a more reliable picture of change over time. On the other hand, standardized annual estimates from a GLM are influenced by the fit of the model to the data. It is therefore useful to compare the pattern of change in standardized estimates from GLMs to that in the original data.

An understanding of hours of fishing per individual trip and hours of fishing per day-location record also was obtained in the exploratory phase of analysis. Note that the fishing technique was not recorded until 1979. The fishing technique recorded on the original individual fish record was either trolling or live-baiting because the tournaments included in this study were specifically directing effort at sailfish. Sailfish is considered a bycatch of the effort for tournaments targeting marlin and therefore data from those events are not considered in this analysis. Live-baiting is very popular among anglers targeting sailfish and represented the majority (over 62%) of the individual records. However, because the individual records were combined into trip records, and then combined into day-location records, the technique could not be included in the data compilation. For this reason, annual CPUE trends for each technique were evaluated to give insight into the consequences of combining the data. It was found that in some years the trolling CPUEs were higher and in some years the live-bait CPUEs were higher. Consequently, combining the data across technique only had an averaging affect on the nominal CPUEs.

General Linear Models:

The GLMs were developed to obtain standardized estimates of annual CPUE, which is generally thought to be indicative of changes in relative annual abundance. These annual estimates were plotted to examine the pattern of change over time. Separate models were constructed for the combined-regions and for the two individual regions. The purpose of constructing separate regional models was to compare the pattern of temporal change for

[1] The number of records analyzed, by region, does not include 304 trips (29 records) from South Carolina or Puerto Rico tournaments. Each of these events were sampled in only two years, 1989-1990, and 1977 and 1983, respectively.

the two regions to that in the combined data. The objective was to determine whether any one region strongly influenced the pattern of change in the combined data and whether the regions differed in their pattern of change.

"Region" and "wave" were the effecting factors included in the combined-regions models. The main effect defined in the regional models was "wave". Each region and wave formed a stratum, or cell, of the analysis in the combined-regions models. The region-wave interaction was included in the models to account for regional differences in seasonal patterns of fishing success. One constraint in the selection of data and "effects" to include in the model was to minimize the number of sparse or empty cells that can compromise an analysis (Searle 1987).

The GLMs were formulated following the methodology of Farber et al. (1992). The approach was to define dummy variables and explicitly include all levels of each effect in the GLM. The independent variables were loaded explicitly, rather than declaring cases or "effects". A dummy variable was defined for each alternative (or "level") of each effect and each alternative main effect and interaction term was placed in the model as a separate independent variable. This approach allowed for the exclusion of the terms and interactions that resulted in empty cells in each model. For instance, certain wave-region interactions were excluded because there were no records for the particular wave in that region. There is no way to exclude empty cells when classes are declared in formulating GLMs. The "proc glm" module in SAS was used (SAS 1991). It was not possible to run a class-oriented model in SAS on the data. Each record was "weighted" in the GLM analysis by the number of trips it comprised (as explained in Farber et al. (1992)) to take into account the original number of trips making up a record.

The dependent variables of the analysis were of the general form:

$$\ln \{[(M \times N) + c] / E\}$$

where N = number of fish hooked or caught, M is a constant multiplying factor, c is a constant necessary in a logarithmic model of this form when zero values are present in the data, and E is hours of fishing. It was necessary to add a constant to the data before taking the log because there were records in which there was effort but no fish were caught. Following Farber et al. (1992), an M of 300 was used with $c=1$.

Fighting time was not subtracted from fishing time because doing so would necessitate eliminating records that did not include information on fighting time (the length of time a fish was on the hook). Although records without fighting time make up only a small portion of the database, they may be unevenly distributed among cells. Furthermore, cells with small sample sizes might be disproportionately affected by excluding them. Also, fighting time is generally short for sailfish and often estimated post facto.

The annual least-squares means produced by the models were back-transformed and divided by three to be expressed in terms of fish caught per 100 hrs of fishing effort. A correction for bias inherent in the back-transformation process was applied (Duan 1983). For all the graphic presentations (other than nominal values), the standardized annual CPUEs estimated by the models were adjusted to a scale based on 1.00 in 1973 before being plotted. Therefore, patterns of change rather than the absolute magnitude of change are revealed by these plots. The adjustment facilitates plotting on the same scale and making comparisons. The 1973-adjusted standardized values are dimensionless.

RESULTS

Nominal Data

The nominal values for effort and numbers caught do not reflect the change in total catch or fishing effort. Rather, they represent the changes in tournament sampling effort over time. Nominal catch and effort for the FEC and Keys combined generally increased throughout the years of the survey, with a sharp increase noted from the early-1980s through the mid-1980s, until 1987 when sampled effort began to decrease yearly (Figure 1a). Nominal fishing success for sailfish was generally stable over the two decades, with fluctuations, but was exceptionally high in 1980 (Figure 1b). Trends in all nominal values for the FEC region (Figures 2a and b) were nearly identical to the combined data.

Nominal catch has been generally stable for the Keys region, while effort fluctuated to a low in 1981, then increased to a high in 1987, and then decreased annually through 1992 (Figure 3a). Nominal CPUE fluctuated from 1973 to a low in 1985, then increased through 1992 to the 1973 level (Figure 3b).

Standardized Annual CPUE Rates

Standardized annual fishing success rates resulting from the GLMs, developed for each of the regions and for the combined-regions were highly significant (F-prob < 0.001) and explained from 28% to 31% of the variation in the data (Table 1). The residuals were generally normally distributed. Nevertheless, the estimated 90% confidence intervals should be viewed as approximations. Graphical results from the GLMs are presented in terms of annual 1973-adjusted standardized rates in dimensionless units (Figures 4-6). For the combined-regions, the standardized CPUEs fluctuated over the period 1973-1981, with a relative high of 1.04 in 1977 (Figure 4). The series then decreased annually to a low of 0.37 in 1985, and then exhibited a general increasing trend through 1992 to a level of 0.85.

The results of the FEC model present a generally flat relative CPUE over the period 1974 through 1992 with values in the 0.3 - 0.5 range, with a high (other than the initial year of 1.00) of 0.69 in 1977 (Figure 5). The Keys model exhibits the most variability with relative CPUE values above 1.0 occurring in 1975, 1981, 1982, and 1990 (Figure 6). There is a general increasing trend with fluctuations from the low of 0.38 in 1985 to 0.93 in 1992.

DISCUSSION

Standardized indices of fishing success rates were generated because nominal fishing success rates may not reflect true changes in relative abundance. Uneven sampling coverage can create spurious change in mean success rates if locations with high success rates (or, conversely, low success rates) contribute more records some years than others. This problem can be expected to affect the data from the billfish survey, because coverage is not uniform across areas and tournaments and is not consistent from one year to another. Standardized estimates produced by general linear models estimate least square means by weighing all strata equally, regardless of number of records.

For both the combined model and the Keys model there were periods of declining (1981-1985) and increasing (1985-1992) CPUEs. However, there is apparently no significant change in abundance since 1973. For the FEC model, relative abundance appears stable since 1974. In comparing r^2 values among billfish species, the combined-regions CPUE model for sailfish ($r^2=31\%$) was greater than the results for blue marlin (26%) and less than for white marlin (37%) found by Farber et al. (1992). This is evidence of some consistency in GLM analysis of billfish data from the recreational billfish survey.

The combined model results exhibit greater variability than the FEC model (which comprised the majority of the data) over the 20-yr period. For both the combined and Keys models, the apparent decrease in abundance over the period 1981-1985 has reversed itself with increasing CPUEs through 1992. Point estimates are near 1.0, with the upper confidence limit including 1.0. The CPUEs for the FEC model have been very consistent since 1974 though never again reaching the initial 1973 value of 1.0.

LITERATURE CITED

Beardsley, G.L. and R.J. Conser. 1981. An analysis of catch and effort data from the U.S. recreational fishery for billfishes (Istiophoridae) in the western North Atlantic Ocean and Gulf of Mexico, 1971-78. U.S. Fish. Bull. 79(1):49-68.

Browder, J.A. and E.D. Prince. 1988. Exploration of use of tournament and dock catch and effort data to obtain indices of annual relative abundance for blue and white marlin, 1972 through 1986. Inter. Comm. Conserv. Atl. Tunas, Coll. Vol. Sci. Pap. 28:287-299.

Browder, J.A. and E.D. Prince. 1990. Standardized estimates of recreational fishing success for blue marlin and white marlin in the western North Atlantic Ocean, 1972-1986. In: R.H. Stroud (ed.), Planning the Future of Billfishes, Research and Management in the 90's and Beyond. Proceedings of the Second International Billfish Symposium, Kailua-Kona, HA. August 1-5, 1988. National Coalition for Marine Conservation, Inc. Savannah, GA. p. 215-229.

Carter, R.L. and M.I. Farber. 1993. Summary of a 1992-1993 volunteer survey of billfish landings from the Florida East Coast and Keys. ICCAT Working Document SCRS/93/101.

Duan, N. 1983. Smearing estimate: A nonparametric retransformation method. J. Am. Stat. Assoc. 78:605-610.

Farber, M.I. 1990. Evaluating statistical bias in using catch-rate indices from the U.S. recreational billfish fishery for estimating abundance by the use of a simulation model. Ph.D. Dissertation, University of Miami, Coral Gables, FL. 231p.

Farber, M.I. and R.J. Conser. 1983. An update on the status of stocks of blue marlin and white marlin in the Atlantic Ocean. Inter. Comm. Conserv. Atl. Tunas, Coll. Vol. Sci. Pap. 18(3):674-692.

Farber, M.I., J.A. Browder, and J.P. Contillo. 1992. Standardization of recreational fishing success for marlin in the western North Atlantic Ocean, 1973-1991, using generalized linear model techniques. ICCAT Working Document SCRS/92/62.

Jones, C.D. and M.I. Farber. 1993. An exploratory surplus-production model analysis of sailfish (*Istiophorus platypterus*) in the western Atlantic Ocean. ICCAT Working Document SCRS/93/99.

Prince, E.D., A.R. Bertolino, and A.M. Lopez. 1990. A comparison of fishing success and average weights of blue marlin and white marlin landed by the recreational fishery in the western Atlantic Ocean, Gulf of Mexico, and Caribbean Sea, 1972-1986. In: R.H. Stroud (ed.), Planning the Future of Billfishes, Research and Management in the 90's and Beyond. Proceedings of the Second International Billfish Symposium, Kailua-Kona, HA. August 1-5, 1988. National Coalition for Marine Conservation, Inc. Savannah, GA. p. 159-178.

SAS. 1991. SAS® System for Linear Models, Third Edition, SAS Institute, Inc., Cary, NC. 329p.

Searle, S. R. 1987. Linear Models for Unbalanced Data. John Wiley. New York. 536p.

Table 1. General linear models developed from the 1973-1992 U.S. billfish survey database for sailfish CPUE. Only tournaments that targeted sailfish were included in the analysis. Regions are the Florida East Coast (FEC) and the Florida Keys (KEYS). The combined-region model includes FEC and KEYS. The number of records represents the daily sum of the individual boat-trips; and the r^2 is the percent of the total variance explained by each model.

REGION	NUMBER OF RECORDS	NUMBER OF TRIPS	r^2
FEC	579	22,355	0.31
KEYS	335	9,973	0.28
Combined	914	32,328	0.31

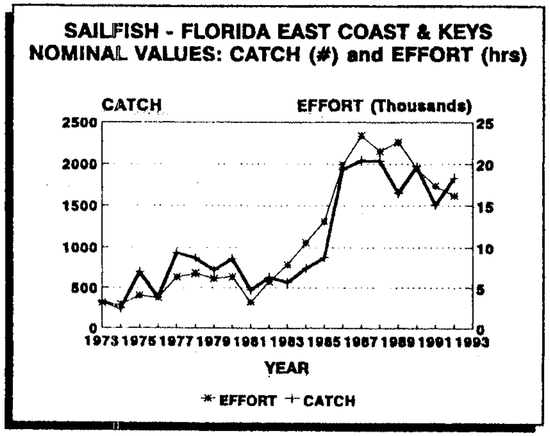

Figure 1a. Nominal values for number of sailfish caught, with associated sampled effort (hrs), over the period 1973-1992 from the U.S. recreational billfish survey, from the combined Florida East Coast and Keys database.

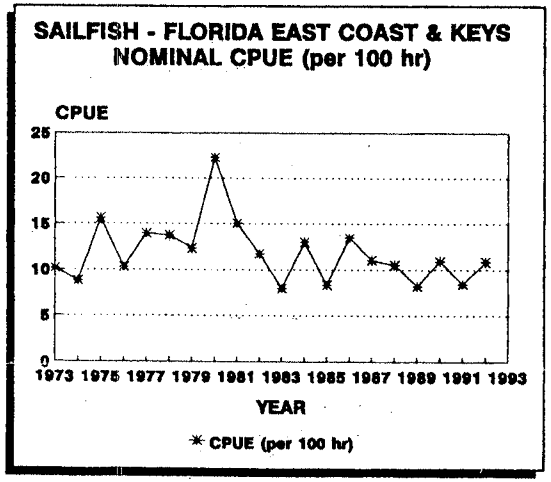

Figure 1b. Nominal value for sailfish CPUE (number of fish per 100 hrs) over the period 1973-1992 from the U.S. recreational billfish survey, from the combined Florida East Coast and Keys database.

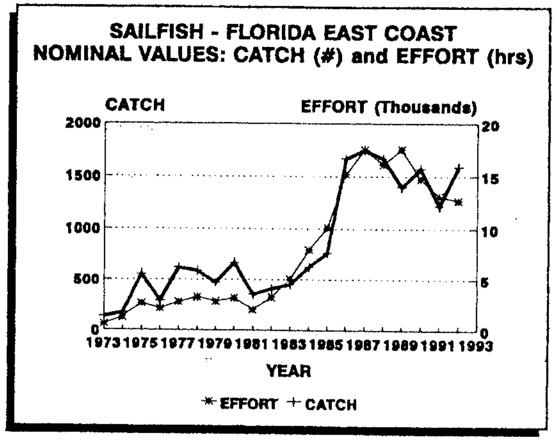

Figure 2a. Nominal values for number of sailfish caught, with associated sampled effort (hrs), over the period 1973-1992 from the U.S. recreational billfish survey, from the Florida East Coast database.

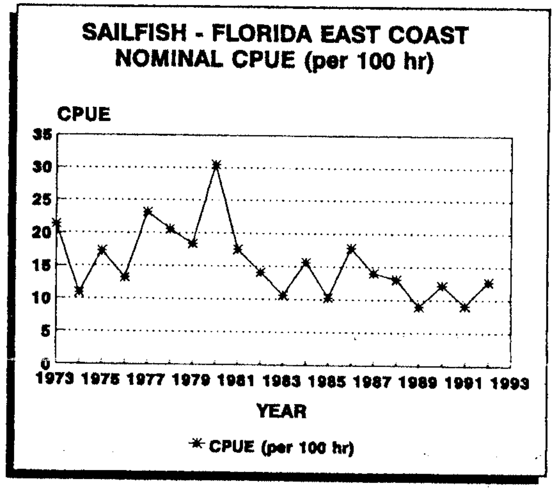

Figure 2b. Nominal value for sailfish CPUE (number of fish per 100 hrs) over the period 1973-1992 from the U.S. recreational billfish survey, from the Florida East Coast database.

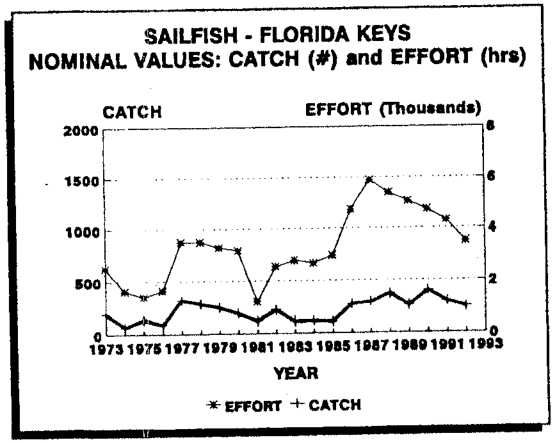

Figure 3a. Nominal values for number of sailfish caught, with associated sampled effort (hrs), over the period 1973-1992 from the U.S. recreational billfish survey, from the Florida Keys database.

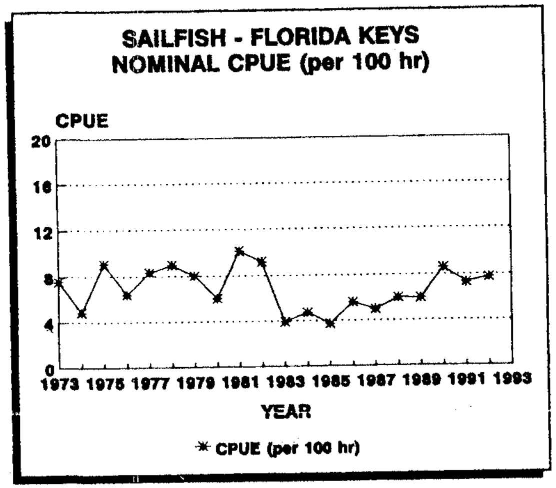

Figure 3b. Nominal value for sailfish CPUE (number of fish per 100 hrs) over the period 1973-1992 from the U.S. recreational billfish survey, from the Florida Keys database.

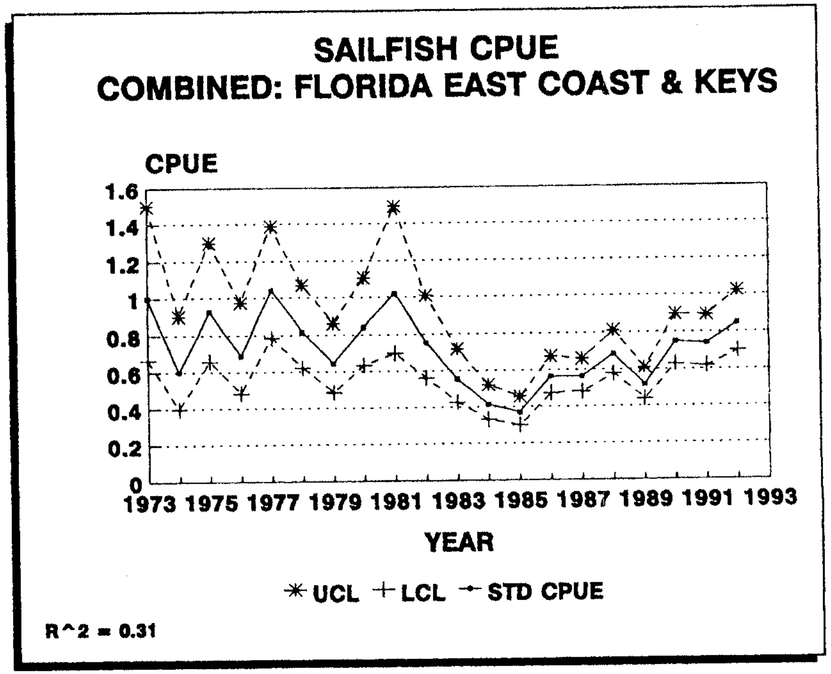

Figure 4. Annual 1973-adjusted standardized mean CPUE (STD CPUE) values for sailfish from the U.S. recreational billfish survey, from the combined Florida East Coast and Keys GLM. The mean rates are plotted with associated approximate 90% upper (UCL) and lower (LCL) confidence limits. Units are dimensionless.

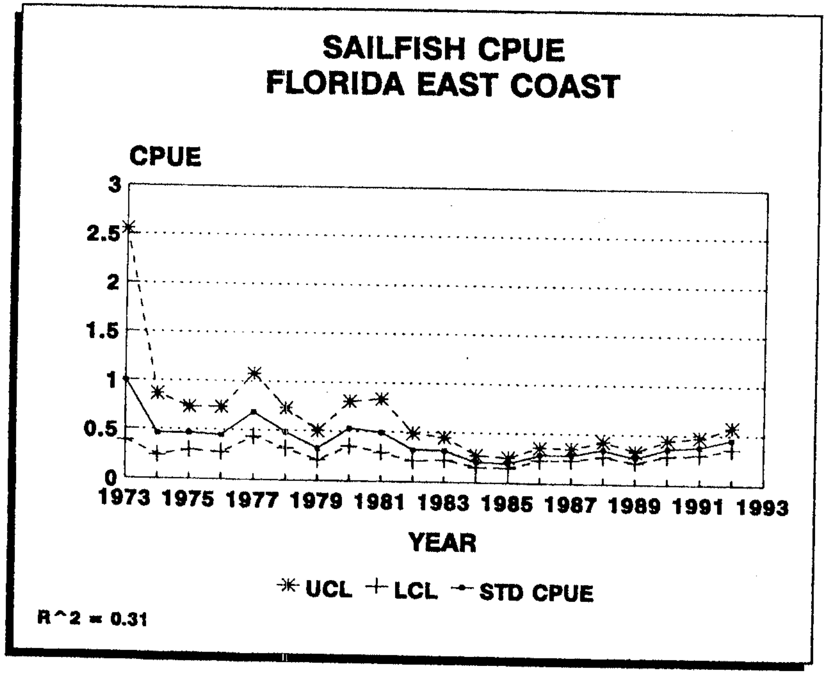

Figure 5. Annual 1973-adjusted standardized mean CPUE (STD CPUE) values for sailfish from the U.S. recreational billfish survey, from the Florida East Coast GLM. The mean rates are plotted with associated approximate 90% upper (UCL) and lower (LCL) confidence limits. Units are dimensionless.

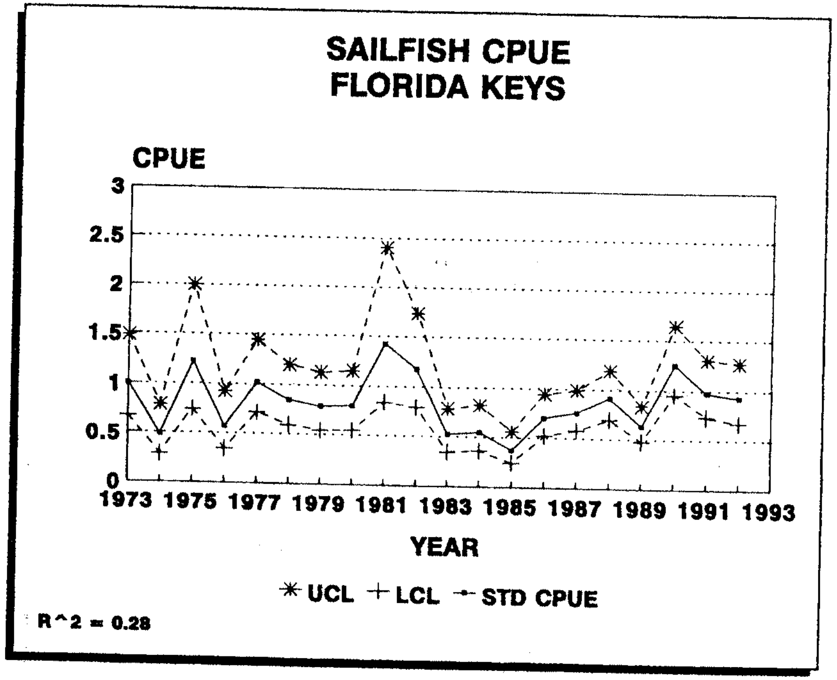

Figure 6. Annual 1973-adjusted standardized mean CPUE (STD CPUE) values for sailfish from the U.S. recreational billfish survey, from the Florida Keys GLM. The mean rates are plotted with associated approximate 90% upper (UCL) and lower (LCL) confidence limits. Units are dimensionless.