

THE EFFECTS OF CHANGING METHODS AND FISHERY REGULATIONS ON CPUE SERIES USED FOR TUNING POPULATION ASSESSMENT MODELS

SCRS/1995/082

Col.Vol.Sci.Pap. ICCAT, 45 (2) : 139-143 (1996)

Hester, F.J.

East Coast Tuna Association

SUMMARY

There is a recognized problem of dealing with possible statistical outliers in CPUE indices used in tuning the VPA for the western Atlantic bluefin assessment. This paper argues that the first step in addressing this type problem is to re-examine the basic data composing the index. Such an approach is used by most stock assessment groups. Several examples are given of potentially biased abundance indices used for stock assessment and how the assessment group decided on the likely source of the bias and dealt with the problem. In the case of the 1993 assessment of the western Atlantic bluefin stock the assessment group departed from usual practices.

RESUMÉ

Un problème a été constaté lorsque l'on traite d'éventuels points statistiques dispersés dans les indices de CPUE utilisés pour ajuster la VPA dans l'évaluation du thon rouge ouest-atlantique. Le présent document avance que le premier pas pour aborder ce genre de problème est d'effectuer un nouvel examen des données de base qui composent cet indice. Cette approche est utilisée par la plupart des groupes d'évaluation de stock. Ci-après plusieurs exemples d'indices d'abondance potentiellement biaisés utilisés pour l'évaluation des stocks, et la façon dont le groupe s'est prononcé quant à la source probable des biais et a résolu la question. Dans le cas de l'évaluation de 1993 du stock ouest-atlantique de thon rouge, le groupe s'est éloigné des processus usuels.

RESUMEN

Existe un problema reconocido al tratar con posibles puntos estadísticos dispersos en los índices de CPUE usados en el ajuste del VPA para la evaluación del atún rojo del Atlántico oeste. En este documento se arguye que el primer paso para tratar este tipo de problema es examinar de nuevo los datos básicos que componen el índice. Este enfoque es aplicado por la mayor parte de los grupos de evaluación de stocks. Se dan varios ejemplos de índices de abundancia con sesgos potenciales, que se usan en la evaluación de stocks y de cómo el grupo de evaluación decidió cual era la posible fuente del sesgo y trató el problema. En el caso de la evaluación del stock de atún rojo del Atlántico oeste efectuada en 1993, el grupo de evaluación abandonó las prácticas habituales.

INTRODUCTION

NRC 1994 raised the issue of different modeling assumption and error structure assumptions on abundance indices used in the 1993 Atlantic bluefin tuna assessment for the west. Depending on the model and error assumptions used the VPA stock biomass estimates could be quite different. For example the base case 8+ spawning stock estimate changed by over 40% depending on the model and error assumption used. As a result, considerable discussion occurred on the matter of which error structure assumption are best suited for developing these and similar indices. However, inspection of the indices reveals the reason for the different results lies in the change of a single point on each CPUE series that departs markedly from the trends: These are the 1983 data point for the US R&R 8+ index and the 1992 data point for the US Gulf of Mexico directed longline index. Since the VPA tuning is sensitive to trends in indices, the occurrence of a point considerably above or below the average in a series may have a large impact on how the index affects the estimates if the point is at the beginning or end of a series.

In the case of these two indices, the degree a point departs from the general average trend of a series changes under different modeling assumptions. The SCRS is aware of this problem and has sought some objective method of dealing with what appear to be statistical outliers. One approach is the present interest in error structure models. Another has been to look for statistical means to identify and reject outliers. This paper argues that concern over the choice of which type of error structure to use, or how to develop mathematical criteria to objectively reject outliers is misplaced until the data producing the aberrant point have been carefully examined. When the resulting assessment is found to be unduly influenced by one or two points out of many, the integrity of these points and the data from which they were derived is open to question. It simply may be that some change in the fishery or in the way the data were collected is a likely cause of the marked departure from the general trend of the series.

THE PROBLEM

US Rod and Reel 8+

This index has undergone considerable modification since it was first proposed to the SCRS in 1983. A chronology of these changes is given in Appendix 1. The index as used in the 1993 western Atlantic bluefin assessment is repeated here in Table 1.

Table 1. US Rod and Reel Large Fish index used in the 1993 assessment.

Year	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Index	15.16	5.80	2.50	0.94	1.87	2.45	2.12	1.83	2.71	0.83

The question of the 1983 data point has been a contentious issue for several years. In 1991 the SCRS reject including it in the index because it was aberrant, and was developed from a different data series than subsequent years. Further, data for 1982 collected under the same protocol were no longer available leaving the high 1983 point to begin the abundance index series (Fig 1). The index was revised for the 1993 SCRS. After considerable discussion the 1983 data point was used in tuning even though the 1983 point had a large impact on the stock size estimates and despite the fact that the residual plot indicated 1983 was somewhat unusual (Fig 2).

US Gulf of Mexico Longline By-Catch.

This index was first considered by the SCRS in 1991, but not used for tuning the VPA for the western Atlantic assessment because it began in 1987 and comprised only four years data. The index uses information from mandatory log books from a selected group of US longline vessels fishing in the Gulf of Mexico and the Florida Strait that consistently catch bluefin tuna as by-catch. At SCRS 1993 the index was again considered and used in the assessment. The most recent year of data, 1992, was a year when a regulatory change occurred. From 1987 through 1991 the vessels were allowed to retain two bluefin per trip, all others being discarded. Beginning in 1992 this allowance was changed to one bluefin per trip, with an additional requirement that the vessel have at least 1134 kg of other species on board before landing a bluefin. Following the change in regulations in 1992, the index showed a sharp decline with that data point being about eight percent of that for 1991, which is unlikely to be solely related to a change in abundance. At the same time catch of other species increased sharply (Fig 3).

The analysis (Cramer and Scott, 1994) recognized the potential problem and included both bluefin reported kept and discarded. Their GLM considered as variables: year, area, season, vessel, swordfish catch rate, yellowfin catch rate, depth and hooks per mile. No consistent relationship between bluefin catch rate and yellowfin catch rates (by area and season) were observed, however, bluefin catch rate varied inversely with swordfish catch rate. Yellowfin catch rate was not used further in the analysis. The exclusion of yellowfin catch rates as a variable, in retrospect, appears problematic as there is a significant negative relation between bluefin CPUE and the catch of other species. (Fig 3).

The index used in tuning the VPA is repeated here as Table 2.

Table 2. Standardized index of abundance for US Gulf of Mexico longline by-catch of bluefin tuna used in the 1993 assessment.

Year	1987	1988	1989	1990	1991	1992
Index	0.87	0.40	0.65	0.68	0.92	0.06

The 1992 data point, being an order of magnitude less than the other years, seemed suspect. Therefore the index was presented both with and without 1992, however, the VPA was tuned using the entire series.

DISCUSSION.

The assessment was sensitive to these data points. Table 3 presents the results of several sensitivity runs made at the 1993 ICCAT.

Table 3. Sensitivity of VPA results to alternate tuning procedures. (After Broadhead, 1993)

Age Group	Base Case	GOM minus 1992	R&R minus 1983	Minus 1992 & 1983
2-5	84.1	86.8 Delta%3.2	92.9 Delta%10.5	96.9 Delta%15.2
6-7	28.4	29.4 Delta%3.5	31.6 Delta%11.3	33.1 Delta%16.5
8-10+	32.8	36.1 Delta%10.1	43.4 Delta%32.3	48.1 Delta%46.6
SBB 8+ 1993/1975 Ratio	32.8/222.5 0.147	36.1/224.4 0.161 Delta%9.5	43.4/228.5 0.190 Delta%29.2	48.1/231.2 0.208 Delta%41.5

Developing indices such as these is a basic approach for stock assessments. The usual practice in developing an index is to determine available information or to design of a system of monitoring catch and effort. This is followed by the collection of these data to obtain a time series, standardization of the indices over time, and evaluation of the index in terms of reliability of the information and its performance in the assessment model. In general evaluation takes the form of examination of the index for unusual changes in trend, or aberrant points that appear not caused by corresponding changes in abundance. This may include examination of goodness of fit between the index and other information such as estimated catches or catch-at-age matrices. In the following section I give some examples of how this evaluation has been carried out in other assessments, and how the assessment groups dealt with the problems that they encountered. In the examples given the problems stem from three causes. These are:

1. There was some change in the conduct of the fleet that increased or decreased fishing efficiency,
2. There was a change in the way the fishery was regulated,
3. A single aberrant point appeared in the index.

EXAMPLES.

1. Changing Fishing Efficiency.
 - a. Spanish Bait Boats.

This is a nominal CPUE index based on catch rates by Spanish bait boats operating in the Bay of Biscay from 1970 to the present. The index was evaluated by the SCRS in 1992 for use in tuning a VPA for bluefin in the eastern Atlantic and Mediterranean.

A non-random pattern in the residuals of the index fit to the catch-at-age data suggested a systematic difference between the VPA and the actual CPUE data. The SCRS decided the use of sonar by the fleet after 1977 could have led to a change in the effective effort of these vessels that may not be sufficiently controlled in calculating the annual CPUE values. The Committee therefore decided to separate the CPUE into two series (1970-1977 and 1978-1991) to account for the hypothesized change in effective effort.

b. Spanish Trap.

At the same 1992 meeting the SCRS observed that the residuals in initial VPA fits to the trap index for large fish were consistently negative in the period 1971-1981 and consistently positive thereafter. Examination of the catch-at-age data showed a consistently higher proportion of fish of ages 7, 8 and 9 compared to ages 10 and older in the trap catches since 1982. The Committee decided this apparent change was related to market factors or some other change not related to stock abundance and therefore separated the index into two series for use in tuning the VPA: one series from 1971 through 1981, the other from 1982 to present.

c. French Purse Seine.

These indices represent the CPUE at age 2 and 3 from French purse seiners operating the Mediterranean over the period 1978 to present. The French scientist observed that the use of spotter aircraft become general in 1982. Therefore it was decided to divide the series at 1981 and use only the latter portion (1982 to present). The earlier period was believed too short (only four years) to be used in the VPA.

2. Changing Fishery Regulations.

a. US Surf Clams.

The National Marine Fisheries Service developed an index of landings-per-unit-effort (LPUE) in the US EEZ based on logbook records of hours fished (by Class 3 clam dredgers), and reported landings. This fishery has been regulated by quotas since 1978. In addition, size limits were put in place in 1982 and a six-hour limit on fishing time in 1985. These two regulations were removed beginning with 1991; quotas remain.

After GLM standardization of LPUE in the Northern New Jersey area (the main fishing area) an unexpected sharp discontinuity appeared in 1990 (SARC 1994), with LPUE decreasing about 50% in 1991 and then maintaining a similar trend to that of the earlier period. The Committee decided that with removal of time limits on fishing effort was more accurately reported in terms of actual hours fished during the 1991 to present period than during 1982-1990. The change in reporting was considered to be a potentially important bias that would affect the utility of LPUE as a relative abundance index in the important Northern New Jersey region. This was examined in an initial run of a modified DeLury model for Northern New Jersey that used standardized LPUE as a relative abundance index. Residuals for the estimated catchability coefficients during 1991-1994 indicated a systematic bias. As a result, an analysis to adjust the standardized LPUE series was performed. The standardized and corrected series are shown in Figure 4.

3. Changing Data Collection Methods.

a. US Surf Clams.

Research vessel survey catch was used to provide fishery independent estimates of abundance for surf clams of the US northeast coast. Catches are number of clams per standard tow. The numbers are reproduced here in Table 4.

Table 4. Survey abundance estimates of surf clams standardized to dredge tow time.

Year	1982	1983	1984	1986	1982	1992	1994
Number of Clams	96.2	86.3	71.5	58.1	61.1	59.1	176.5

Survey results for 1994 showed about a three-fold increase in numbers per tow (Fig 5). This was attributed to an increase in sampling efficiency. The sensitivity of various model results to inclusion of the 1994 data point was tested by projecting the survey indices and determining if the observed survey points fell within the confidence intervals of the projected indices. Because the 1994 survey values were so large, relative to the time series, several runs were made using the 1994 recruit index, replacing it by the recent average recruit index. Replacing the index with the average resulted in improved model fits. The average was used for the final assessment.

CONCLUSIONS .

A few examples of screening procedures and resulting decisions have been given above. In each instance indices were judged to be in need of closer examination because a trend over time occurred with the residuals, or if a single point stood out as an apparent outlier. In each case the assessment group took some action that resulted in rejection of data, division into two indices, or the development and application of a correction term. Standing in contrast are the two large bluefin indices for the western Atlantic, which were accepted and used despite similar, but more severe, problems. Clearly, a standard approach needs to be agreed to for dealing with data problems in the future in the assessment group for the western Atlantic. In particular, a consistent approach needs to be agreed to on when to consider dropping certain data from CPUE series prior to GLM analysis, or when to remove certain suspect data point from the indices after GLM analysis.

RECOMMENDATIONS

1. In cases where the series are used to tune an assessment model such as a VPA, due consideration should be given to the increased impact such points have when they occur at the beginning or end of a series, and such point should be evaluated carefully.
2. It is worthwhile to seek objective methods to identify what may be statistical outliers.
3. The original data producing single data points and indices that have an inordinate effect on VPA estimates should be re-examined critically.
4. When the data the outlier represents can be associated with a particular departure from the basic collection protocol and assumptions, the problem should be dealt with as a data problem and the common practice of deleting data prior to the final GLM analysis, or breaking the index series should be followed.
5. In the case of the assessment of Atlantic bluefin tuna the same standards should be applied when evaluating indices of abundance for any fishery.

REFERENCES

Cramer, J. and G.P.Scott. 1994. Indices of Abundance for Large Bluefin Tuna, *Thunnus thynnus*, from the U.S. Mandatory Pelagic Longline Fishery in the Gulf of Mexico and off the Florida East Coast. ICCAT Col. Vol. Sci. Pap. 42(1): 164-171.

National Research Council, 1994. An Assessment of Atlantic Bluefin Tuna. Nat'l Acad. Press, Wash. D.C. 148 pp.

Northeast Fisheries Science Center, In Press. Report of the Fall 1994 NEFSC Stock Assessment Workshop (19th SAW).

Note: References to the U.S. Large Fish Rod and Reel Index will be found in the Appendix.

APPENDIX 1.

US ROD AND REEL - HANDLINE CPUE INDEX

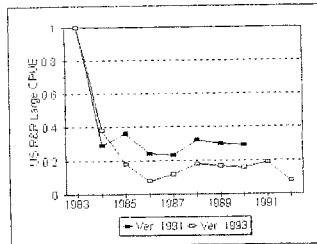


Fig 1. US Rod and Reel CPUE indices presented to the SCRS. The 1983 data point was rejected in 1991 for use in tuning the VPA.

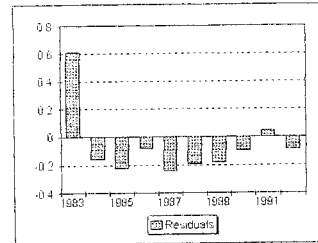


Fig 2. Standardized residuals for the 1993 US Rod and Reel index used in tuning the 1993 VPA Base Case.

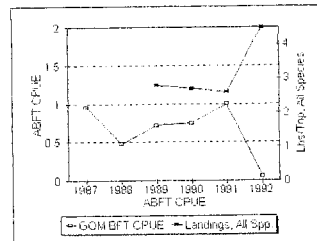


Fig 3. The relation between ABFT catch/effort in the Gulf of Mexico U.S. longline by-catch fishery and the landings per trip of all other species.

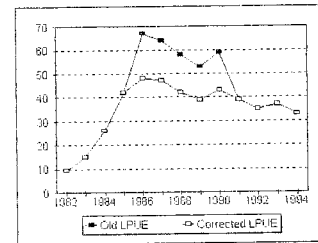


Fig 4. Surf clam landings per Unit Effort corrected for under reporting the hours fished.

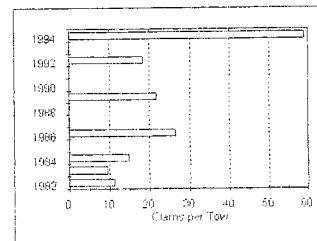


Fig 5. Standardized surf clam research survey catches of fully recruited clams (>140mm).

The following is a chronological summary of major modification made to the index since it was first used by the SCRS to tune the VPA. The quotes extracted from the respective SCRS documents presenting the index. References to these papers are provided at the end of this summary for those wishing to consult the complete text.

Brown and Turner (1989)

"Preliminary investigation indicated that interviews conducted in 1987 were biased towards successful trips, ... data from 1987 had to be excluded from all final analyses and results.

"Each trip interview record included data on: date, gear (with the exception of 1982)..."
 "...the GLM analysis for 1983-1986 ... indicated that gear did not contribute significantly to the model. [therefore] 1982 data were added."

Brown and Turner (1990)

"Data for 1982 were excluded from the analysis as there were too few observations to insure a balanced design."

"Interviewer reports from 1987 suggested that a bias existed towards interviews of successful trips. Data for 1987 were therefore excluded from the analysis.

Cramer and Brown (1991)

Data from 1982 were excluded as in Brown and Turner (1990).

"...the relative scarcity of trips... [excluded using] vessels fishing out of western Long Island (NY), New Jersey and states further south."

Turner, Cramer and Brown (1992)

"The existence of two survey systems in 1994-1984 complicates the use of some data from those years (specially 1984) in developing large bluefin indices. For 1982, basic data from the large fish survey is not currently accessible, though summary information is. In 1983 data for the two surveys is readily distinguished. The 1984 data from the surveys are not as readily distinguished, so that careful selection of data to use is needed."

"There was sufficient information to test gear only in GOMA in the early years when fishing with handlines was more frequently recorded. Gear accounted for about 2% of the total SS in the analysis on disaggregated data and about 4% and 13% on the analyses on data aggregated over year and gear, and year and gear and SST respectively. Inclusion of the effect in both GOMA and STHN would have resulted in substantial unbalance in the design and its effect would have been confounded with area, therefore handlines were excluded from the analysis."

"...1982 excluded..."(supra)

1987 was included in this and the following analyses.

Cramer and Turner (1994)

"As noted in the earlier paper, data from 1982 is not currently available."

Data from 1987 was included (supra).

REFERENCES

Brown, C.A. and S.C. Turner. 1989. Standardized Catch Rates of Large Bluefin Tuna in the New England (U.S.) Rod and Reel/Handline Fishery. ICCAT Col. Vol. Sci. Pap. 30(2): 302-310.

Brown, C.A. and S.C. Turner. 1990. Standardized Catch Rates of Large Bluefin Tuna in the New England (U.S.) Rod and Reel/Handline Fishery. ICCAT Col. Vol. Sci. Pap. 32(2): 248-254.

Cramer, J. and C. A. Brown. 1991. Standardized Catch Rates of Large Bluefin Tuna in the New England (U.S.) Rod and Reel/Handline Fishery. ICCAT Col. Vol. Sci. Pap. 35(2): 294-301.

Turner, S. C., J. Cramer and C. A. Brown. 1992. Standardized Catch Rates of Large Bluefin Tuna, *Thunnus thynnus* in the New England (U.S.) Rod and Reel/Handline Fishery. ICCAT Col. Vol. Sci. Pap. 39(3): 758-771.

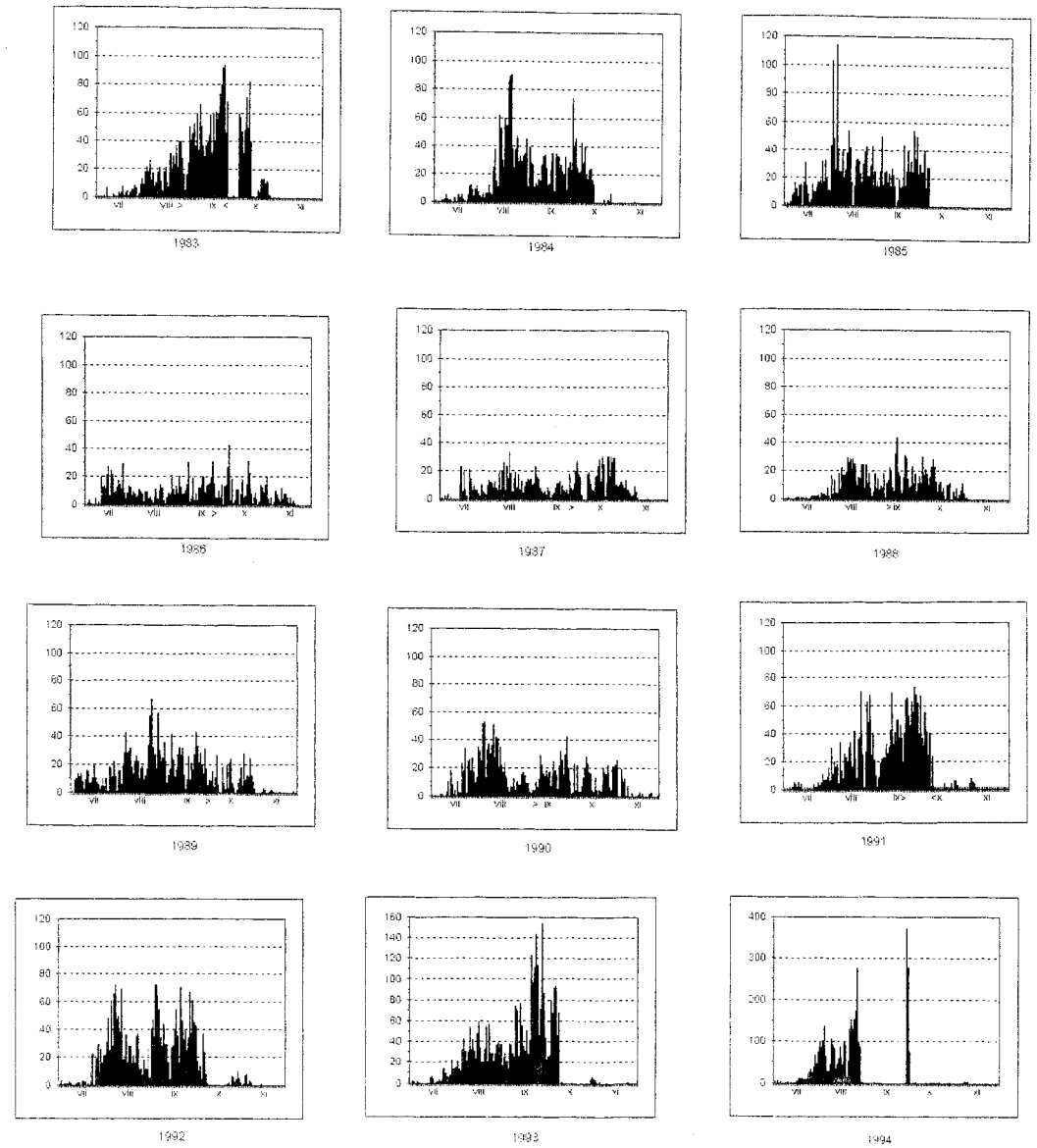
Cramer, J., S.C. Turner. 1994. Large Bluefin Tuna, *Thunnus thynnus*, indices of abundance from the Rod and Reel and Handline Fishery off the Northeast United States. ICCAT Col. Vol. Sci. Pap. 42(1): 156-168.

And a Note. Last year there was discussion about the potential bias changing the daily bag limit from one to two fish might have had on this index. In response the NMFS provided the following information regarding dates when the two fish per day regulation was in force:

Year	Date Reg. Began	Date Reg. Ended	Percent Total Catch @ 2/day
1983	8/10	9/10	56
1984	Not implemented		NA
1985	Not implemented		NA
1986	9/10	12/31	36
1987	9/11	12/31	30
1988	8/27	12/31	52
1989	9/15	12/31	17
1990	8/24	12/31	46
1991	9/6	9/27	38

Because of the way data for the index were collected (few observation were made after mid-September), an effect, if any is unlikely to be detected in 1986, 1987 or 1989. The quota was not reached in 1988 even though the bag limit was raised to two fish per day in late August. Because few cells contain two fish per day catch data this effect would be difficult to assess using GLM.

One might *a priori* suspect that CPUE was higher during the two-fish-per-day period in 1983 and 1991 since the index in part is based on catch per line hour and part of that time would be time to finding the first fish. Arguably, less time to find the second fish would be likely if for no other reason than the running time involved from when the lines went into the water until the boat reached the fishing area. Appendix Figure 1 shows the daily reported landings from 1983 through 1994. Time when the daily bag limit was two fish per day are denoted by > <. The first of each month is shown by the corresponding Roman numeral. In both 1983 and 1991, catch rates during the two-fish-per-day period were used in developing the index. How great an effect this has on the GLM estimates is unclear. For 1983, during the approximately four month fishery, 58% of the catch occurred during the 30 day period when two fish were allowed per day.



Appendix 1, Fig. 1. Daily Landings of Giant Bluefin by General Category Boats. Note the scaling change for 1993 and 1994.