SCRS/1987/068

Col.Vol.Sci.Pap. ICCAT, 28 : 287-299 (1988)

EXPLORATION OF THE USE OF TOURNAMENT AND DOCK CATCH AND EFFORT DATA TO OBTAIN INDICES OF ANNUAL RELATIVE ABUNDANCE FOR BLUE AND WHITE MARLIN 1972 THROUGH 1986

J. A. Browder, E. D. Prince
Southeast Fisheries Center
75 Virginia Beach Drive, Miami, Florida, U.S.A.

SUMMARY

By applying a general linear model to billfish tournament and dock catch and effort data, indices of abundance for the years 1972 through 1986 were obtained from the following measures of fishing success: numbers of blue marlin hooked and caught and white marlin hooked and caught per unit of fishing time. Compressing the data from a record-per-trip to a record-per-day basis improved the fit of the residuals to a normal distribution and increased the percent of variation in fishing success that could be explained by the model. For example, the R^2 for number of blue marlin hooked per unit of fishing time was increased from 4 percent to 22 percent. In the analysis, the compressed records were weighted according to number of trips that they covered. The R^2 for number of white marlin hooked or caught was nearly twice that for blue marlin hooked or caught. The influencing factors considered in the analysis were sample type (tournament or dock), area, month, and year. Both main effects and certain two-way interactions were examined. In addition, the effect of certain local environmental variables was tested. Sea state and cloud cover were significant explaining variables in models for blue marlin; sea state, but not cloud cover, was significant in explaining white marlin fishing success.

RESUME

En appliquant un modèle linéaire général aux données de prise et effort d'istiophoridés provenant des concours de pêche et des prises à quai, les indices de l'abondance pour la période 1972-1986 ont été obtenus à partir des mesures de la production de pêche suivantes: nombre de makaires bleus et blancs pêchés à l'hameçon et capturés --par unité de période de pêche. En comprimant les données d'un registre-par-sortie à un registre-par-jour, l'ajustement des résidus s'améliore à une distribution normale et accroît le pourcentage de variation de la production de la pêche qui pourrait être expliqué par le modèle. Par exemple, le R^2 pour le nombre de makaires bleus pêchés à l'hameçon par unité de période de pêche s'est accru, allant de 4 à 22 %. Dans l'analyse, les registres comprimés ont été pondérés au nombre de sorties. Le R^2 pour le nombre de makaire blanc pêché à l'hameçon ou capturé représentait près du double que celui de makaire bleu. Les facteurs influents étudiés dans l'analyse ont été classifiés par type d'échantillonnage (concours de pêche et prises à quai) par zone, mois et année. Les effets principaux et certaines interactons en double ont fait l'objet d'un examen. De plus, l'effet de certaines variables locales de milieu a été testé. L'état de la mer et des nuages étaient importants pour expliquer les variables des modèles du makaire bleu; l'état de la mer et non les nuages, a été significatif pour expliquer l'effectivité de la pêche du makaire blanc.

RESUMEN

Aplicando un modelo lineal general a los datos de concursos de pesca, sobre muelle, y de esfuerzo, de marlines, se obtuvieron índices de abundancia para los años 1972 a 1986 a partir de las siguientes mediciones de captura productiva: número de ejemplares de aguja azul enganchados con anzuelo y capturados, y de aguja blanca enganchados con anzuelo y capturados, por unidad de período de pesca. Condensar los datos de un registro por salida a un registro por día mejoró el ajuste de los residuales a una distribución normal y aumentó el porcentaje de la variación de la producción de pesca que podría explicarse mediante el modelo. Por ejemplo, el R^2 para el número de agujas azules capturadas con anzuelo por unidad de período de pesca aumentó de 4% a 22%. En los análisis, los registros resumidos se ponderaron de acuerdo con el número de viajes. El R^2 para el número de agujas blancas enganchadas con anzuelo o capturadas fue casi el doble que el de aguja azul enganchada con anzuelo o capturada. Los factores de influencia considerados en los análisis fueron del tipo de muestra (concurso o muelle), área, mes y año. Se examinaron los dos efectos principales y ciertas interacciones de doble sentido. Además, se comprobó el efecto de ciertas variables ambientales locales. El estado de la mar y la cobertura de nubes constituyeron variables de importancia significativa en los modelos para la aguja azul; el estado de la mar - y no la cobertura de nubes - fue un factor de peso a la hora de explicar el éxito en la pesca de aguja blanca.

INTRODUCTION

Development of reliable indices of abundance are important in assessing the status of stocks and determining trends. To provide a data base for stock assessment, a long-term survey of recreational catch and effort for billfishes off the U.S. east coast, the Gulf of Mexico, the Bahamas, and the Caribbean Sea was initiated in 1972 by the Southeast Fisheries Center's Miami Laboratory. Data have been collected annually since that time. Details of this survey are described in Beardsley and Conser (1981). In the present study, we use data for all of the main areas of the survey to develop separate indices of annual abundance for blue marlin (*Makaira nigricans*) and white marlin (*Tetrapturus albidus*) from 1972 through 1986. By using data from all areas surveyed, we attempted to minimize the possibility that variation in some oceanographic feature specific to an area (for instance, the Loop Current in the Gulf of Mexico) could distort the index by affecting local accessibility of the stock to fishing.

METHODS

The data set contained the following information relevant to this analysis: (1) number of vessels fishing each day, including those not catching any fish; (2) number of fish of each species hooked by each vessel each day; (3) number of fish of each species caught by each vessel each day; (4) number of hours fished by each vessel on each day; (5) number of hours of fighting time; (5) average wave crest height; and (6) cloud cover. The bases for estimating an index of abundance were, alternatively, number hooked per unit effort (HPUE) or number caught per unit effort (CPUE). The effort unit was "effective fishing time" in hours, defined as the number of hours fished minus fighting time. According to Beardsley and Conser (1981), HPUE may be a more reliable indicator of abundance than CPUE in recreational fisheries for billfish, because the weight of the fishing line, which varies among tournaments and among individuals, affects the proportion of billfish hooked that actually are boated. It is possible to identify hooked fish, because almost all billfish jump out of the water when they are hooked.

The raw data were contained in 15 data files, one for each year, and were in the form of a separate record for each fish. The first working data set was constructed from the raw file by combining the original records (one for each fish) into a record for each trip. After preliminary analysis, a second working data set was constructed from the first by combining all the trips for the same sampling-event day (tournament day or dock-sampling day) into a single record containing total numbers of fish and hours from all the trips. Each of the working files contained the data for all 15 years of the survey. The data sets were refined by eliminating (1) tournaments and locations where sailfish was the principal target, (2) locations with very few records (less than approximately 60), and (3) months with very few records. Records from the lower Florida east coast were excluded by these criteria, as were records from St. Petersburg, Florida; Mississippi; and Cozumel. Blue marlin catches in New England (areas north of New York) were so low that records from this region were eliminated from the data set used for the blue marlin analysis, although they were retained for the white marlin analysis. The months excluded from the analysis were November, December, January, and February, which had 373, 51, 23, and 122 records, respectively. The data set for white marlin consisted of 94,547 records, while that for blue marlin contained 94,294 records.

Six areas were defined as follows: (1) New England (white marlin) or New York (blue marlin) to Beaufort, North Carolina; (2) South Carolina to Daytona, Florida; (3) Florida Keys; (4) Bahamas; (5) Caribbean; and (6) Gulf of Mexico. Eight months were included in the data set: March through October. Two types of samples were defined: (1) dock sampling and (2) tournament sampling. Beardsley and Conser (1981) found significant differences in these two types of samples in the Gulf of Mexico. Table 1 shows the number of records in each type-area-month strata. In an approach similar to that of Robson (1966), Farber (1986), Turner (1986), and others, a general linear model (GLM) was developed to distinguish variation in catch rates among years by holding three other main sources of variation -- type, area, and month -- constant. A regression model was constructed in which a dummy variable represented each alternative of the four parameters -- type, area, month, and year. Additional dummy variables were assigned to represent each alternative of two-way interactions between area and type and area and month. One standard for each type, area, month, and year was defined, and dummies for the main effects of these standards and their two-way interactions were purposely omitted from the model, a necessity of this procedure. Interactions of year with other parameters were purposely excluded from the model in order that the year effect could be used as an index of annual abundance. In the analysis, each dummy variable was set equal to one or zero, depending on whether or not a record was from the type, area, month, or year that dummy represented. The regression analysis produced a regression coefficient for each dummy variable. The generalized predictor equation, which assumed a log-linear relationship between the dependent variable and the dummy variables, was as follows:

$$\ln (M\, Y_{hat} + 1) = B_0 + \text{sum}\,(B_t\, D_t) + \text{sum}\,(B_a\, D_a) + \text{sum}\,(B_m\, D_m) + \text{sum}\,(B_y\, D_y) + \text{sum}\,(B_{at}\, D_{at}) + \text{sum}\,(B_{am}\, D_{am}) \qquad (1)$$

where Y_{hat} = prediction of relative abundance

D_t = dummy variables for type $\quad t = 1,2$

D_a = dummy variables for area $\quad a = 1,\ldots,6$

D_m = dummy variables for month $\quad m = 3,\ldots,10$

D_{at} = dummies for area-type

D_{am} = dummies for area-month

B_0 is the constant and the B vectors are the regression coefficients for the variants of each factor. Only one condition of each factor can be met by each situation; therefore, for each prediction of the dependent variable (Y_{hat}), only one regression coefficient for each factor parameter is "turned on" (i.e., multiplied by D = 1 rather than D = 0). A standardized index of annual abundance (Y_{hat-y}) is calculated from the regression constant and the regression coefficient of the dummy variable for each year, as follows:

$$Y_{hat-y} = \{[e^{(B_0 + \text{sum}\,(B_y\, D_y))}] - 1\} / M \qquad (2)$$

As this equation indicates, the effect of each year on the dependent variable, the standardized annual index of abundance, is the sum of the constant and the regression coefficient for that year. Y_{haty} is a "standardized" estimate of annual abundance because factors other than year are held constant. A time series of predictions for each year indicates the annual variation in the data.

Ninety percent confidence limits around Y_{haty} were calculated on the basis of the variance of B_0 and each of the B_y's and the covariance of B_0 and B_y. According to Farber (NMFS/SEFC, Miami, pers. comm.), the covariance of B_0 and each B_y, when not directly available from statistics output, can be calculated as

$$\text{Covar}\,(B_0\, B_y) = -\,\text{Mean}\,(D_{yn})\ \text{Var}\,(B_y) \qquad (3)$$

where D_{yn} is the vector of values of a specific year dummy variable.

The original model consisted of the following main effects and interactions: type, area, month, year, area x type, and area x month. The standards selected for the model were: type = 2 (tournament), area = 6 (Gulf of Mexico), month = 8 (August), and year = 1983. The final model included the following: type, area-1 x type-1, area-5 x type-1, and area x month for all area x month interactions in the model. The same model design was used for all four dependent variables (blue marlin HPUE and CPUE and white marlin HPUE and CPUE).

This model, when applied to the first working data set, which consisted of a record for each trip, explained only 3.7% of the variation in the number of blue marlin hooked per unit of effort. An analysis of resulting residuals indicated that their distribution departed greatly from a normal distribution (Figure 1a).

A second working data set, in which the data were compressed into a record for each sampling-event day, was created in an attempt to improve the fit of the residuals to the normal distribution. To avoid the bias that might be caused by using records representing different levels of effort, we weighted records in our analysis according to the number of trips covered by each. Because the records were weighted in the analysis, compressing the data did not change the degrees of freedom of the analysis. The number of trips covered by each record varied from one to over 100 and averaged about 10. A large number of records with only one trip occurred in the data set due to dock-sampling visits in which only one boat was interviewed on a given day. Compressing the data reduced the total amount of variation in the data set because some of the individual variation among fishing units was eliminated. Not all of this type of variation was eliminated, however, since we still had a considerable number of records for only one trip. The amount of variation in the dependent variable that was explained by the model increased from the original 3.7% to approximately 22% when we applied it to the compressed data set, and the distribution of residuals more nearly approximated the normal distribution (Fig. 1b).

Once we had made the initial improvement in the distribution of the residuals, we tested the effect on this distribution of the multiplier (M) used in the log transform of the dependent variable (see equations 1 and 2). If zero values of the dependent variable occur in a data set, the addition of a constant to the value is necessary before making the log transformation, because the log of zero is not defined. Applying a multiplier to the variable before adding one is essentially the same as changing the constant from one. Traditionally, M = 100 is used with recreational fishing data, and results are expressed in terms of number caught per 100 hrs. We tested M = 100, 200, 300, 400, and 500. Assuming that the basic unit of effort was six hours, the most common length of one trip, using these multipliers was analagous to changing the constant from one to: 0.06 (M = 100), 0.03 (M = 200), 0.02 (M = 300), 0.015 (M = 400), and 0.012 (M = 500). An M was selected to minimize the sum of the absolute values of the coefficients of skewness (g_1) and kurtosis (g_2), as suggested by Berry (1987). A minimum was reached at $M = 300$, when $g_1 = -0.015$ and $g_2 = 0.162$ (Table 2). This is equivalent to adding a constant of 0.02 to the data after multiplying by six.

According to a Kolmogorov Smirnoff test, the distribution of weighted residuals still departed significantly ($p < 0.001$) from a normal distribution; therefore, we conducted an analysis of the residuals to determine whether the fit of the residuals to a normal distribution could be improved by eliminating groups of data. Residuals were grouped according to type, area, or month, and the distribution of weighted residuals within each group was examined (Tables 3, 4, and 5). Then residuals were grouped, first by the number of trips representing by the record (Table 6) or by the number of blue marlin hooked per record (Table 7). Certain groups displaying obviously aberrant distributions relative to the normal distribution expected for the complete set of residuals were excluded from the data set, one group at a time, and a separate regression analysis performed. The following eliminations were made: (1) type 1 (dock samples); (2) area 1 (New York to Beaufort, N.C.); (3) records representing less than 10 trips; (4) records representing less than 20 trips; and (5) records in which the number of blue marlin hooked was zero. As indicated in Table 8, none of these eliminations improved the fit of the residuals to the normal distribution; in fact, the opposite was true. The complete data base seemed preferable to any of these alternatives. Regression models and predictions are not affected by the lack of normality in the distribution of residuals, but, since we did not quite satisfy this requirement, calculated confidence limits for our predictions may not be accurate.

In addition to dummy variables representing alternative states of our parameters and their interactions, we also tested the effect of two environmental variables -- sea state (average wave-crest height) and cloud cover -- on the percent of variation in fishing success that could be explained by the model. Both of these variables were collected at the time catch and effort data were collected and had been included on the data records. Wind data, also included on the records, was not tested in the model because it was highly correlated with sea state. Sea state and cloud cover were poorly correlated with each other in the data.

RESULTS

Table 9 indicates the adjusted R^2's and other statistics of various alternative models for each of the four dependent variables. Generally speaking, the models for white marlin explained about twice as much of the variation in their dependent variables as the models for blue marlin. Although sea state and cloud cover were indicated to be highly significant ($p < 0.0001$) independent variables in blue marlin models, the models that included these variables explained little more of the variation in fishing success than models without them. For instance, the model with sea state explained 0.2% more of the variation in blue marlin hooked per unit effort than the model without this variable; and the model with both sea state and cloud cover explained 0.5% more of the variation in blue marlin hooked per unit effort than models without the these two variables. Although sea state was a highly significant ($p < 0.0001$) variable when added to the white marlin models, cloud cover was not.

The abundance indices resulting from the model are plotted by year in Figures. 2 and 3. Indices are reported as base numbers rather than logs; retransformation was accomplished by taking the exponent of the sum of the regression coefficient and the constant, then subtracting one, and then dividing by three. Division was by three rather than 300 so that the index could be reported in terms of 100 hrs of fishing. Within both species, the pattern of variation in HPUE and CPUE was similar. A feature common to three of the plots -- white marlin HPUE and CPUE and blue marlin HPUE -- is a minimum in 1978; blue marlin CPUE is lower in 1976 than in 1978. A maximum occurs in 1980 in white marlin HPUE and CPUE, but in 1981 in blue marlin HPUE and CPUE. A persistent decline after 1980 is evident in white marlin HPUE and CPUE; whereas a decline from 1981 to 1982 was followed by years of stability in blue marlin HPUE and a slight increase in blue marlin CPUE, followed by a decline from 1985 to 1986. Estimated 90% confidence limits around the indices are shown but may not be accurate because our residuals were not quite distributed normally.

Predictions of blue and white marlin HPUE based on area-month interactions holding "year" effects constant were used to examine seasonality within areas and to compare HPUE in the various areas within months. In Table 10a, the months in which blue marlin HPUE was lower than the overall mean estimated based on all the area-month interactions are listed by area in the first column and months in which it was greater than the mean are listed in the second column. Area-months excluded from the program because of lack of data (empty cells) or for certain other reasons are listed in the final column. In Table 10b, the areas in which blue marlin HPUE was lower than the overall mean are listed by month in the first column. Areas in which blue marlin HPUE was higher than the overall mean are listed in the second column. Areas excluded from the program are listed by month in the final column. The same information is given for white marlin HPUE in Table 11.

In most cases, the seasonality within areas suggested by model predictions is consistent with what we know about the fishery and the distribution of the fish (Table 10a). For instançe, the period of higher than average blue marlin HPUE is limited to one month in the area from New York to Beaufort. Early summer and fall (May, June, and September) are the months of greater than average HPUE in the area from South Carolina to Daytona. Higher than average catches occur in the Keys only in October. Higher than average blue marlin HPUE occurs from March through August in the Bahamas, June through October in the Caribbean, and May through October in the Gulf of Mexico. Clearly, the Caribbean, the Bahamas, and the Gulf of Mexico are the principal areas of the blue marlin fishery, as was reported previously by Prince and Bertolino (1986). Our model results indicate that HPUE is higher than average in these three areas during most months sampled.

Model predictions suggest that the seasonality and centers of interest of the white marlin fishery are somewhat different from that of blue marlin (Table 11a). The northeast, which includes New England, is an area of higher than average HPUE from June through October. White marlin HPUE is below average in the area from South Carolina to Daytona in all months sampled. Above average HPUE is found in May and July in the Florida Keys and in March and April in the Bahamas. May is the month of above average white marlin HPUE in the Caribbean. June through October are months of above average white marlin HPUE in the Gulf of Mexico. The northeast, including New England, and the Gulf of Mexico are indicated as the only two areas of consistently high white marlin HPUE for several months (Table 11).

DISCUSSION

The first question raised by the results is how valid are the indices of annual abundance produced by the models? The models' indications of seasonality in the various areas and the relative importance of the various areas within any given month are consistent with what is known about the fisheries for both blue and white marlin, suggesting that the models' perceptions of year-to-year variation in fishing success also are accurate. In addition, aspects of our results are supported by a previous analysis by Beardsley and Conser (1981). Using a portion of the same data set used in our analysis -- Gulf of Mexico data for the period 1972-1978, they compared a recreational HPUE from pooled dock and tournament data to the CPUE of the Japanese longline fishery operating in the same area at the same time. They concluded that white marlin HPUE was a reliable indicator of stock abundance because, within all three of the subareas examined, the measures of fishing success in the two fisheries were correlated over time. Their data showed a decline in white marlin HPUE in the Gulf of Mexico from a peak in 1975 through 1978, the last year covered by their analysis. Our results for white marlin using data from all areas are consistent with their results for the Gulf of Mexico in that a low occurs in 1978. In our analysis, however, the decline starts after a peak in 1974, rather than 1975. Beardsley and Conser (1981) were not able to find a relationship between a pooled recreational HPUE for blue marlin and the blue marlin CPUE of the Japanese longline fishery in the Gulf of Mexico.

A second question raised by our results is how can the models be improved? Explaining 20 to 40% of variation in the data leaves plenty of room for improvement. Some of the unexplained variation may be due to variation in fishing skills among individual units. We did not completely eliminate this source of variation by combining the data into event-day records, because, 2,672 records in the blue marlin data base, or approximately 10% of the dock-sampling records, still represent individual trips; and a total of 16,938 records represent between two and nine trips. If the sampling schedule were revised so that the number of trips sampled in each day at the dock is increased, this would reduce the individual variation in our data base consisting of a record for each sampling-event day. Reducing individual variation is one approach to explaining a greater percentage of variation in the data with a model.

Variation among locations within the same area is another possible source of variation. Preliminary analysis of data from the Gulf of Mexico (using the original record-per-trip data base) indicated a significant difference in blue marlin HPUE among locations. For instance, over the 15 yr period, HPUE at Galveston was significantly higher than anywhere else in the Gulf ($p < 0.05$). Some of the variation in the data might be explained by separating the data into more finely-graded areas. Our preliminary analysis suggested that, although differences among locations were significant, the percent of variation in blue marlin HPUE that could be explained by location differences was fairly small.

Beardsley and Conser's (1981) analysis indicated differences in the relative HPUE for blue marlin from dock samples and tournament samples among different regions within the gulf. This kind of inconsistency might have reduced the amount of variation that could be explained by our model. Creating more finely-graded areas might allow us to resolve this problem.

The extent to which we can use more finely defined areas in our model is limited by the distribution of the data. Increasing the number of areas would increase the number of empty area-month cells. Logistic problems also limit the extent to which areas can be subdivided. The size of the matrices that must be manipulated in regression analysis increase as the square of the number of variables. As we define the GLM, each two-way interaction is a variable; thus as the number of variants of any one factor is increased, the matrices can quickly become unmanageably large.

Long-term and short-term effects of variable oceanographic and meteorologic factors may be sources of variation in the data. We tested two variables that might affect either the efficiency of fishing units or the propensity of marlin to strike bait. Although both variables were highly significant in the blue marlin equation and one was highly significant in the white marlin equation, neither explained much of the data's variation -- possibly because the dummy variables for area-month interactions explained most of the spatial-temporal variation in the data that was due to sea state and cloud cover. The addition of these variables potentially could reduce variation only within areas and months or, possibly, among months of different years.

To get a better idea of how environmental factors affect day-to-day fishing success, we plan to analyze a special set of tournament data in which we have a daily record of fishing success at the same location for over a month -- data from the 1985 Cayman Island Tournament. As possible factors influencing fishing success, we will examine not only sea state and cloud cover but also moon phase, tidal cycle, wind speed, wind direction, barometric pressure, water temperature, and several other aspects of local conditions. The fact that some tournaments -- for instance, the San Juan and St. Thomas tournaments -- are scheduled solely on the basis of moon phase suggests that this might be an important factor affecting fishing success. Information gained from analysis of the special data set could provide knowledge to improve the present models. For instance, if moon phase is shown to be a highly significant factor explaining fishing success, it might be feasible to retrospectively add moon-phase information to the 15-yr recreational tournament and dock-sampling data base and reanalyze the data, incorporating this information into the model. Certainly it is feasible to add this information to records for future years.

Oceanographic factors such as ocean temperature and current patterns could affect the catchability of marlin stocks by affecting the date of their arrival in a fishing area or their position in the area relative to bases from which fishing vessels depart. Fishing success may be low some years because the tournament dates did not coincide with the time fish were in the area or because conditions kept the fish too far offshore to be accessible to fishing. Environmental data required to examine effects of this type would, of course, have to be area-specific. While incorporating ocean temperature in a data base covering several areas would be easy, incorporating current effects probably would be not be effective, even if possible, since specific oceanographic features differ by area. Considerable thought should be given to whether and how such data should be added to a model of the type we have developed here. Simply adding local ocean temperature to the model might not be very effective in reducing unexplained variation in HPUE and CPUE, because the area-month interactions already included in the model might account for much of the variation that local ocean temperature might otherwise explain. It might be more profitable to examine this type of environmental effect in models specific to each region than in a model covering all the regions.

CONCLUSIONS

The following conclusions can be drawn from this analysis:

1. The seasonality of fishing success and relative fishing success in the various areas predicted by the models for blue and white marlin are consistent with what is known about the fisheries for the two species and thus support the model's standardized annual estimates of HPUE and CPUE as indices of abundance.

2 The pattern of annual variation in white marlin abundance indicated by this analysis is consistent with the pattern for the first eight years of the record indicated by Beardsley and Conser (1981) from an analysis of a portion of the same data set. The variation in Beardsley and Conser's recreational HPUE for white marlin in the Gulf of Mexico was correlated with the CPUE of the Japanese longline fishery operating in the same location at approximately the same time.

2. Results of a previous analysis (Beardsley and Conser 1981) support the use of white marlin annual estimates of HPUE as an index of abundance.

3. Only about 20% of blue marlin and 40% of white marlin fishing success were explained by the models, which raises the question of how the

models might be improved by either (a) reducing the variation in the data or (b) adding other factors to explain additional variation.

4. Part of the unexplained variation probably is due to differences in anglers, equipment, and fishing site. This general category of variability might be reduced if more samples per day were taken at each sampling location. A low number of samples is more common in dock sampling than in tournament sampling and might be improved by sampling docks on weekends, rather than weekdays, when more samples can be obtained.

5. A greater percentage of the variation in the data might be explained by the model if environmental variables associated with marlin fishing success were added to the model. An analysis of data from the 1985 Cayman Island Tournament will be used to evaluate the potential for various environmental factors such as moon phase to explain variation in fishing success.

REFERENCES

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. Fish Bull 79: 49-68.

Berry, D. A. 1987. Logarithmic transformations in ANOVA. Biometrics 43: 439-456.

Farber, M. I. 1986. Swordfish catch per unit effort (CPUE) trends in the Northwest Atlantic based on Japanese longline data. Unpublished working document 86/1, Swordfish ASsessment Workshop, NMFS, SEFC, Miami, FL, April 1986. 20 p.

Prince, E. D., and A. R. Bertolino. 1986. Recreational CPUE for Atlantic blue marlin along the U. S. east coast, in the Bahamas, Caribbean, and Gulf of Mexico, 1972-1984. ICCAT Working Document SCRS/86/73. Miami Laboratory, Southeast Fisheries Center, National Marine Fisheries Service, Miami, FL. 3 p.

Robson, D. S. 1966. Estimation of relative fishing power of individual ships. Res. Bull. Int. Comm. N. W. Atl. Fish. 3:5-14.

Turner, S. C. 1986. Catch rates of bluefin tuna in the Japanese longline fishery recorded by United States observers. Int. Comm. Conser. Atl. Tunas, Col. Vol. Sci. Pap. 26:323-338.

Table 1. Distribution (total number) of raw samples in blue marlin file by type, area, and month, 1972-1986.

Month	Type	Area 1 NY-NC	 2 SC-Day	 3 FL Keys	 4 Bahamas	 5 Carib	 6 GOM
All	1. Dock	6,342	0	0	0	927	17,986
	2. Tourn	9,773	2,545	2,406	18,776	7,606	27,922
3. March	1	0	0	0	0	58	38
	2	0	0	0	3,591	0	0
4. April	1	0	0	0	0	30	245
	2	0	4	0	5,489	0	25
5. May	1	164	0	0	0	45	1,736
	2	27	2,032	291	3,229	223	1,709
6. June	1	783	0	0	0	100	3,117
	2	3,497	177	0	3,221	2,023	6,315
7. July	1	2,302	0	0	0	222	4,362
	2	2,304	0	72	3,226	688	9,436
8. Aug	1	2,055	0	0	0	318	4,113
	2	2,422	0	0	20	1,312	6,341
9. Sept	1	932	0	0	0	103	2,696
	2	1,523	332	0	0	2,148	3,289
10. Oct	1	106	0	0	0	51	1,679
	2	0	0	2,043	0	1,212	807

Table 2. Statistics of analysis of the effect on the distribution of residuals of changing the multiplier in the dependent variable.

Mult	Equiv. Const.	Coef. of Kurtosis	Coef. of Skewness	Sum of 1 & 2	K-S Z	R^2	F-stat	Mean SS
100	0.06	0.889	0.415	1.304	9.71	0.2137	502	0.4134
200	0.03	0.307	0.127	0.434	10.85	0.2180	516	0.6762
300	0.02	0.162	-0.015	0.177	15.39	0.2203	522	0.8698
400	0.015	-0.111	-0.105	0.216	18.79	0.2218	527	1.0259
500	0.012	0.092	-0.169	0.261	21.42	0.2228	530	1.1582

Note: weighted number of records = 94,294.

TABLE 3

DISTRIBUTION OF WEIGHTED RECORDS WITHIN STANDARDIZED RESIDUAL INTERVALS, BY TYPE, FROM REGRESSION OF BLUE MARLIN HPUE ON DUMMY VARIABLES.

		TYPE	
		1	2
9.99	-3.00	60	20
-3.00	-2.67	12	0
-2.67	-2.33	0	216
-2.33	-2.00	237	1646
-2.00	-1.67	800	1726
-1.67	-1.33	5798	250
-1.33	-1.00	3482	3791
-1.00	-0.67	1014	4454
-0.67	-0.33	255	6452
-0.33	0.00	595	12903
0.00	0.33	1233	13528
0.33	0.67	1758	11828
0.67	1.00	2139	6547
1.00	1.33	2616	3719
1.33	1.67	1922	1318
1.67	2.00	1469	429
2.00	2.33	903	107
2.33	2.67	481	54
2.67	3.00	275	33
3.00	9.99	212	12

TABLE 4

DISTRIBUTION OF WEIGHTED RECORDS WITHIN STANDARDIZED RESIDUAL INTERVALS, BY AREA, FROM REGRESSION OF BLUE MARLIN HPUE ON DUMMY VARIABLES.

		AREA					
		1	2	3	4	5	6
9.99	-3.00	0	0	0	0	80	0
-3.00	-2.67	0	0	0	0	12	0
-2.67	-2.33	0	43	0	105	0	68
-2.33	-2.00	432	25	0	469	0	957
-2.00	-1.67	16	23	9	1005	40	1433
-1.67	-1.33	0	95	0	128	67	5758
-1.33	-1.00	3886	103	232	390	400	2262
-1.00	-0.67	1719	228	64	1175	484	1798
-0.67	-0.33	683	351	24	1897	878	2874
-0.33	0.00	1063	357	830	3230	2337	5681
0.00	0.33	1493	241	555	3215	2164	7093
0.33	0.67	2053	524	507	3576	897	6029
0.67	1.00	1577	269	162	1775	586	4317
1.00	1.33	1565	232	0	1174	307	3057
1.33	1.67	834	42	23	422	121	1798
1.67	2.00	385	12	0	188	97	1216
2.00	2.33	227	0	0	27	30	726
2.33	2.67	105	0	0	0	28	402
2.67	3.00	38	0	0	2	2	266
3.00	9.99	44	0	0	0	3	177

TABLE 5

DISTRIBUTION OF WEIGHTED RECORDS WITHIN STANDARDIZED RESIDUAL INTERVALS, BY MONTH, FROM REGRESSION OF BLUE MARLIN HPUE ON DUMMY VARIABLES.

		MONTH							
		3	4	5	6	7	8	9	10
9.99	-3.00	0	0	0	3	8	44	9	16
-3.00	-2.67	0	0	0	12	0	0	0	0
-2.67	-2.33	0	0	43	68	105	0	0	0
-2.33	-2.00	0	126	234	629	563	134	159	38
-2.00	-1.67	352	343	292	870	296	232	92	49
-1.67	-1.33	0	79	833	605	1594	1080	1104	753
-1.33	-1.00	73	222	461	189	1981	2839	1119	389
-1.00	-0.67	396	367	469	1317	1868	801	202	48
-0.67	-0.33	450	351	609	2061	1218	1136	616	266
-0.33	0.00	516	1199	1390	2898	2530	1685	1948	1332
0.00	0.33	487	859	1172	3675	3683	2252	1760	873
0.33	0.67	668	1107	1564	3052	3051	2029	1266	849
0.67	1.00	219	614	1041	1595	2177	1374	1086	580
1.00	1.33	258	302	779	1003	1759	1148	842	244
1.33	1.67	166	120	189	648	730	765	419	203
1.67	2.00	81	39	157	300	502	509	179	131
2.00	2.33	13	29	93	172	241	313	99	50
2.33	2.67	0	19	47	77	171	141	53	27
2.67	3.00	2	9	46	44	61	60	47	39
3.00	9.99	7	11	35	22	74	41	23	11

TABLE 6

DISTRIBUTION OF WEIGHTED RECORDS WITHIN STANDARDIZED RESIDUAL INTERVALS, BY NUMBER OF TRIPS IN RECORD, FROM REGRESSION OF BLUE MARLIN HPUE ON DUMMY VARIABLES.

		NUMBER OF TRIPS							
		1	1-9	10-19	20-29	30-39	40-49	50-59	>=60
9.99	-3.00	18	32	10	20	0	0	0	0
-3.00	-2.67	5	7	0	0	0	0	0	0
-2.67	-2.33	2	29	104	43	38	0	0	0
-2.33	-2.00	41	592	684	374	138	0	54	0
-2.00	-1.67	217	951	832	358	168	0	0	0
-1.67	-1.33	1283	3974	513	97	30	41	110	0
-1.33	-1.00	447	2414	1463	931	539	434	506	539
-1.00	-0.67	77	535	848	1158	736	809	323	982
-0.67	-0.33	5	105	1114	1206	1392	1189	390	1306
-0.33	0.00	4	271	1453	1948	2335	1963	1069	4455
0.00	0.33	4	578	1982	2606	2690	1679	1642	3580
0.33	0.67	4	933	2534	2619	2685	2122	530	2159
0.67	1.00	17	1297	2111	1383	1396	785	265	1432
1.00	1.33	8	1481	1797	1048	769	576	326	330
1.33	1.67	18	1460	1032	363	134	128	105	0
1.67	2.00	29	1051	542	89	104	83	0	0
2.00	2.33	96	675	153	86	0	0	0	0
2.33	2.67	141	343	25	26	0	0	0	0
2.67	3.00	125	183	0	0	0	0	0	0
3.00	9.99	131	77	16	0	0	0	0	0

TABLE 7
DISTRIBUTION OF WEIGHTED RECORDS WITHIN STANDARDIZED RESIDUAL INTERVALS, BY NUMBER OF BLUE MARLIN BOOKED, FROM REGRESSION OF BLUE MARLIN HPUE ON DUMMY VARIABLES.

		NUMBER BOOKED							
		0	1	2	3	4	5	6	>=7
9.99	-3.00	80	0	0	0	0	0	0	0
-3.00	-2.67	12	0	0	0	0	0	0	0
-2.67	-2.33	216	0	0	0	0	0	0	0
-2.33	-2.00	1883	0	0	0	0	0	0	0
-2.00	-1.67	2486	10	30	0	0	0	0	0
-1.67	-1.33	5790	186	19	53	0	0	0	0
-1.33	-1.00	5538	827	556	20	61	43	0	228
-1.00	-0.67	1232	1369	1113	613	420	180	171	370
-0.67	-0.33	76	1292	1289	973	906	178	460	1533
-0.33	0.00	57	1156	1463	1425	941	1124	1003	6329
0.00	0.33	34	1394	942	1070	1070	741	1017	8493
0.33	0.67	0	1459	966	961	920	1065	688	7527
0.67	1.00	0	1077	949	301	637	502	406	4814
1.00	1.33	0	872	827	473	313	565	299	2986
1.33	1.67	0	548	539	490	308	200	123	1032
1.67	2.00	0	409	267	226	217	185	61	533
2.00	2.33	0	248	171	155	86	63	76	211
2.33	2.67	0	178	78	69	50	53	17	90
2.67	3.00	0	121	82	37	24	13	30	1
3.00	9.99	0	88	52	28	19	7	5	25

Table 8. Statistics of analysis of the effect on the distribution of residuals of excluding subsets of data for which the distribution of residuals was highly non-normal.

Group Excluded	Weighted Number Records	Coef. of Kurtosis	Coef. of Skewness	Sum of Abs. Val. 1 & 2	K-S Z	R^2	F-stat	Mean SS
None	94,294	0.162	-0.015	0.177	15.39	0.2203	522	0.8698
<10 trips	74,634	0.653	-0.582	1.235	17.11	0.2554	544	0.5467
<20 trips	57,421	0.973	-0.556	1.529	10.62	0.2914	536	0.4001
# hooked=0	76,890	0.760	0.246	1.006	6.58	0.2543	535	0.3516
Type 1	69,033	0.859	-0.579	1.438	17.51	0.2651	566	0.5466
Area 1	78,174	0.375	-0.079	0.454	19.37	0.1836	399	0.8475

Table 9. Statistics of the four basic models and alternative models containing environmental variables.

Excluded	Weighted Number Records	Coef. of Kurtosis	Coef. of Skewness	Sum of Abs. Val. 1 & 2	K-S Z	Adj. R^2	F-stat	Mean SS
Blue Marlin Hooked								
Dummies only	94,294	0.162	-0.015	0.177	15.39	0.2199	522	0.8698
Dummies & sea		0.169	-0.010	0.179	14.96	0.2219	518	0.8675
Dummies, sea, and sky		0.185	-0.005	0.190	15.31	0.2246	513	0.8650
Blue Marlin Caught								
Dummies only		0.503	0.511	1.014	21.28	0.2255	539	0.8097
White Marlin Hooked								
Dummies only	94,547	0.360	-0.157	0.517	14.76	0.4342	1,423	1.1593
Dummies & sea		0.363	-0.158	0.521	15.28	0.4354	1,402	1.1573
White Marlin Caught								
Dummies only	94,547	0.440	0.109	0.549	11.80	0.4061	1,269	1.1354

Table 10. Distribution of blue marlin HPUE by area and month relative to overall mean.

a. Months of lower and higher HPUE, by area.

	Months < mean	Months ≥ mean	Area-months not in Model
1. New York-Beaufort, N.C.	5,7,8,9,10	6	3,4
2. So. Carolina-Daytona, Fla.	4	5,6,9	3,7,8,10
3. Florida Keys	5,7	10	3,4,6,8,9
4. Bahamas		3,4,5,6,7,8	9,10
5. Caribbean	3,4,5	6,7,8,9,10	
6. Gulf of Mexico	3,4	5,6,7,8,9,10	

b. Areas of lower and higher HPUE, by month.

	Areas < mean	Areas ≥ mean	Area-months not in model
3. March	5,6	4	1,2,3
4. April	2,5,6	4	1,3
5. May	1,3,5	2,4,6	
6. June		1,2,4,5,6	3
7. July	1,3	4,5,6	2
8. Aug.	1	4,5,6	2,3
9. Sept.	1	2,5,6	3,4
10. Oct.	1	3,5,6	2,4

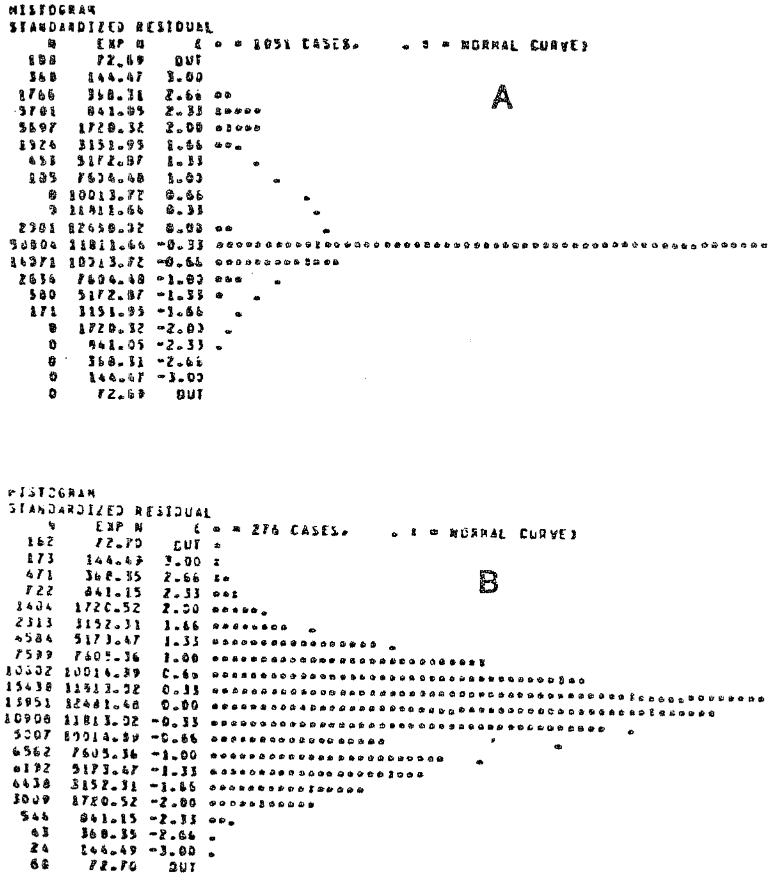
Table 11. Distribution of predicted white marlin HPUE by area and month relative to overall mean.

a. Months of lower and higher HPUE, by area.

	Months < mean	Months ≥ mean	Area-months not in model
1. New England-Beaufort, N.C.	5	6,7,8,9,10	3,4
2. So. Carolina-Daytona, Fla.	4,5,6,9		3,7,8,10
3. Florida Keys	10	5,7	3,4,6,8,9
4. Bahamas	5,6,7,8	3,4	9,10
5. Caribbean	3,4,6,7,8,9,10	5	
6. Gulf of Mexico	3,4,5	6,7,8,9,10	

b. Areas of lower and higher HPUE, by month.

	Areas < mean	Areas ≥ mean	Area-months not in model
3. March	5,6	4	1,2,3
4. April	2,5,6	4	1,3
5. May	1,2,4,6	3,5	
6. June	2,4,5	1,6	3
7. July	4,5	1,3,6	2
8. Aug.	4,5	1,6	2,3
9. Sept.	2,5	1,6	3,4
10. Oct.	3,5	1,6	2,4


```
HISTOGRAM
STANDARDIZED RESIDUAL
    N       EXP N       ( * = 1051 CASES,     . = NORMAL CURVE)
   108      72.69    OUT
   360     144.47   3.00
  1766     360.31   2.66
  3701     841.05   2.33
  5697    1720.32   2.00
  1926    3151.95   1.66
   653    5172.87   1.33
   185    7634.48   1.00
     0   10013.72   0.66
     9   11811.66   0.33
  2381   12650.32   0.00
 50804   11811.66  -0.33
 16371   10013.72  -0.66
  2636    7634.48  -1.00
   580    5172.87  -1.33
   171    3151.95  -1.66
     0    1720.32  -2.00
     0     841.05  -2.33
     0     360.31  -2.66
     0     144.47  -3.00
     0      72.69    OUT
```

A

```
HISTOGRAM
STANDARDIZED RESIDUAL
    N       EXP N       ( * = 276 CASES,     . = NORMAL CURVE)
   162      72.70    OUT
   173     144.49   3.00
   471     360.35   2.66
   722     841.15   2.33
  1404    1720.52   2.00
  2313    3152.31   1.66
  4584    5173.47   1.33
  7599    7605.36   1.00
 10302   10016.39   0.66
 15438   11813.02   0.33
 13951   12661.48   0.00
 10908   11813.02  -0.33
  5307   10016.39  -0.66
  6562    7605.36  -1.00
  6192    5173.47  -1.33
  6438    3152.31  -1.66
  3009    1720.52  -2.00
   546     841.15  -2.33
    63     360.35  -2.66
    24     144.49  -3.00
    68      72.70    OUT
```

B

Figure 1. Distribution of standardized residuals from regression of blue marlin HPUE on dummy variables, using "trip" records (A) and "day" records (B).

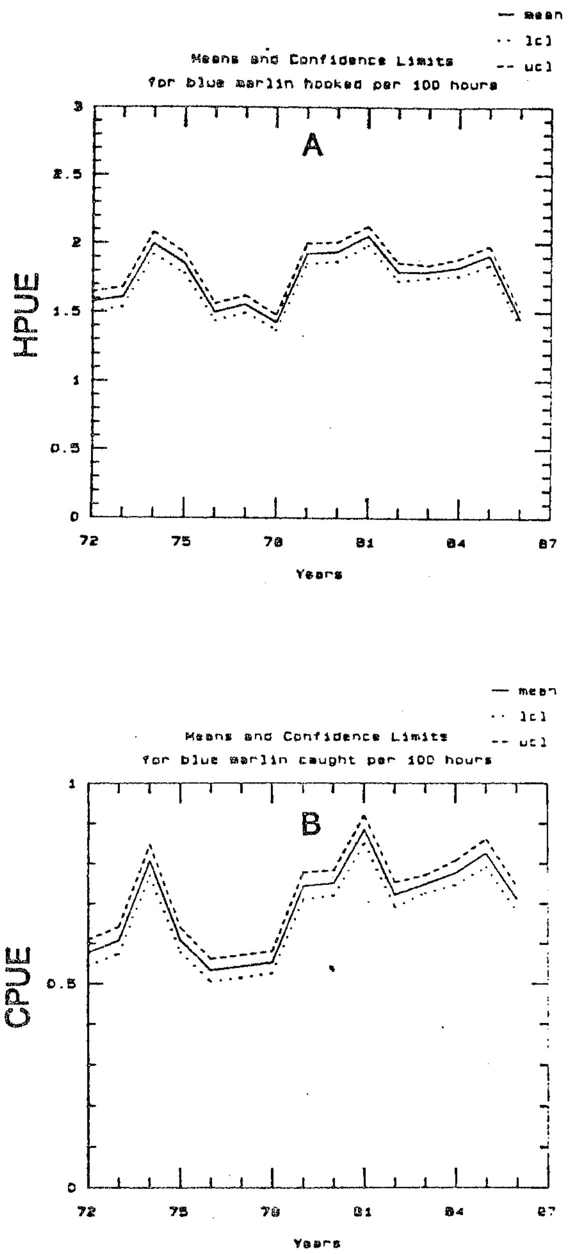


Figure 2. Standardized number of blue marlin hooked (A) and caught (B) per 100 hrs of fishing, by year, based on data from all six areas. Estimated 90% confidence limits are included.

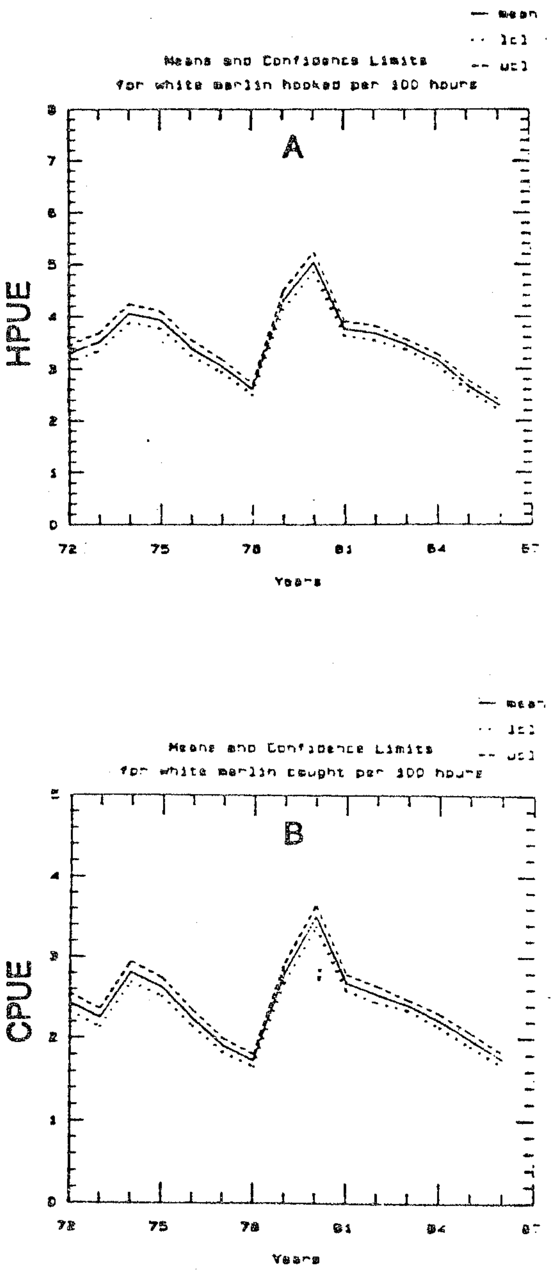


Figure 3. Standardized number of white marlin hooked (A) and caught (B) per 100 hrs of fishing, by year, based on data from all six areas. Estimated 90% confidence limits are included.