1999 ANNUAL REPORT

Estimation of Bycatch and Discards of Sharks and Other Species by the Hawaii Longline Fishery

William Walsh & Pierre Kleiber

May 1999

SUMMARY

The principal activities supported by this grant since June 1998 have been a series of statistical analyses that compared blue shark catch rates reported by National Marine Fisheries Service (NMFS) observers to those provided in logbook records from commercial vessels operating within the Hawaii-based commercial longline fishery. The analyses evaluated data gathered from March 1994 through December 1997, which represented the first 46 months of the fishery observer program in Hawaii. The objectives were to identify and characterize the effects of environmental, operational, and temporal factors on blue shark catch rates, and then to simulate these effects on a fishery-wide basis for use as a comparative standard on the fraction of trips (currently about 95%) that do not carry a fishery observer. This was considered an appropriate area of inquiry because blue shark has been the most numerous species in the catch of this fishery during this decade, and it remains unclear as to whether the seemingly large incidental catch (approximately 100,000 per year) has affected its population status. The first objective was attained by developing two different statistical models, known as a generalized additive model (GAM) and a regression tree, respectively, from the fishery observer data. The second objective was attained by using the GAM to generate a prediction corresponding to each reported catch, and representing a 'surrogate' observer. The major findings of the analyses were that blue shark catch rates exhibited no long-term trend of decrease, that a tractable suite of readily measured variables can serve as the basis of a meaningful simulation of blue shark catch rates, and that statistical simulation can be usefully employed to evaluate logbook reporting practices.

Progress to Date

This project has included three major categories of activities relevant to the Hawaii-based commercial longline fishery: preparation of logbook and observer data, a series of modern statistical analyses, and presentation of results. As of this writing (May 1999), the first two areas have been completed, and substantial progress has been achieved in the third.

The first requisite task, data processing, was performed periodically, to make the logbook data set as complete, correct, and amenable to manipulation as possible. At present, logbook data for the period November 1990-December 1998 (approximately 96000 longline sets) are properly identified to individual longline sets and readily manipulated with the statistical software package S-PLUS (MathSoft, Inc. 1996). The observer data from March 1994 to January 1999 are also available (approximately 2800 longline sets), can be manipulated with S-PLUS (MathSoft Inc. 1996), and have been properly linked to the logbook data, so that all sets witnessed by observers can be directly compared to the corresponding logbook reports.

Two major series of statistical analyses were conducted in the past year. The first consisted of the development of two models of blue shark CPUE from observer data: a generalized additive model (GAM) and a regression tree. Both models identified a similar suite of extrinsic factors related to blue shark CPUE, and comparison of the two demonstrated that they were mutually complementary in facilitating interpretation of the results. For example, both models identified temporal effects on blue

shark CPUE as significant, but the regression tree, which considered years and months as categorical rather than continuous variables, indicated that months were the principal source of variation. As such, temporal effects on blue shark CPUE were interpreted as reflections of seasonality, rather than any long-term changes in population status. The second series of analyses consisted of a fishery-wide "expansion"of the observer-based GAM to the logbook data. This entailed application of its coefficients to values of predictor variables provided in logbook reports. The result was an expanded data set that included a predicted catch corresponding to each reported catch. Comparisons of these predictions with the logbooks permitted identification of questionable data and yielded considerably improved insight into reporting (and under-reporting) of blue shark catches. These analyses, in addition to the cluster analysis described in the 1998 Annual Report, represent the output required under the second task area of this project.

The final task required by the project, presentation of results, is nearing completion. Three draft manuscripts have been completed, and are currently in the Honolulu Laboratory internal review process. The intention is to submit these to *Fisheries Research* (peer-reviewed) for publication. Abstracts from these manuscripts are attached.

Planned Activities

The only remaining activities for this project will consist of any manuscript revisions that may be required by the NMFS or journal review processes.

Budget Status

The budget from this project has been exhausted. There are no plans to seek additional funding.

Literature Cited

S-PLUS Version 3.4. 1996. MathSoft Inc. Seattle, WA.

Cluster Analysis of Logbook Data from the Hawaii-based Commercial Longline Fishery, with Emphasis on Blue Shark (Prionace glauca)

William A. Walsh

Pierre Kleiber

Abstract

A seven-year set (November 1990-October 1997) of logbook data from the Hawaii-based commercial longline fishery was subjected to cluster analysis by Ward's method. Data consisted of the enumerated taxonomic composition of the catch from individual longline gear sets. The objectives were to identify distinct taxonomic assemblages within the catch, to detect associations between such assemblages and fishing practices, and to estimate catch per unit effort (CPUE) for blue shark.

Three clusters were defined on the basis of catches of 12 species, each of which comprised at least 1% of the total catch of the fishery. These species in combination comprised 92.5-98.0% of within-cluster catches. The first cluster, and the most consistent in its catch composition, was dominated by blue shark and swordfish. These species within this cluster comprised higher percentages of the catch and had higher mean catch per unit effort than all other species in all clusters. The second cluster was dominated by bigeye tuna and albacore, and the third was dominated by swordfish and bigeye tuna. Both the blue shark / swordfish and bigeye tuna / albacore clusters were divisible into minor clusters. The two blue shark / swordfish minor clusters differed primarily in terms of blue shark catches. Bigeye tuna and albacore each dominated a minor cluster.

Catch composition was associated with specific categories of fishing trips. Longline sets from the blue shark / swordfish cluster were derived primarily from swordfish-directed and "mixed-species"-directed fishing. Sets from the bigeye tuna / albacore cluster were derived primarily from tunas-directed and "mixed-species"-directed fishing.

Blue shark catches and CPUE exhibited pronounced intra-annual variation. CPUE plots within the blue shark / swordfish minor clusters tracked roughly in parallel to one another.

We conclude on the basis of the consistency of the composition of the blue shark / swordfish cluster that distinct and readily identifiable assemblages exist within the catch. We conclude that catch composition is meaningfully associated with fishing practices. Finally, we conclude that the time series of CPUE within the various clusters, particularly the blue shark / swordfish cluster and its constituent minor clusters, show no evidence of long-term changes concerning blue shark, which was the most abundant species within the fishery during the seven-year study period.

Generalized Additive Model and Regression Tree Analyses of Blue Shark (Prionace glauca) Catch Rates by the Hawaii-based Commercial Longline Fishery

William A. Walsh

Pierre Kleiber

Abstract

Generalized additive model (GAM) and regression tree analyses were conducted with blue shark, Prionace glauca, catch per unit effort (CPUE) data collected aboard Hawaii-based commercial longline vessels by National Marine Fisheries Service observers from March 1994 through December 1997 (N = 2010 longline sets). The objective was to improve predictive capability by relating blue shark CPUE to a tractable suite of readily measured or computed variables. The GAM included nine spatio-temporal, environmental, and operational variables and explained 72.1% of the deviance of blue shark CPUE. Geographic location was the predominant influence on blue shark CPUE. Latitude exerted the strongest effects of any individual variable; longitude was the most influential variable when adjusted for the effects of all other factors. Temperature was the most important environmental factor, with relatively cold temperature associated with high CPUE. An adjusted time series plot revealed no apparent trends in blue shark CPUE since 1994. The initial regression tree included 68 terminal nodes and 11 predictors. It was refined to a final tree with 42 terminal nodes, which reduced the root mean deviance by 65.3%. The tree was partitioned first on latitude 26.6°N, and then branched out to reach terminal nodes after 2-8 additional partitionings. Sets south of this latitude were characterized by lower CPUE and partitionings on a greater number and variety of predictors. Northerly sets were characterized by higher and more variable blue shark CPUE. Predictions from the two analyses were highly correlated ($r=0.903$, $P << 0.001$). Moreover, use of these methods in combination aided greatly in the interpretation of results. We conclude that GAM and regression tree analyses can be usefully employed in assessment of blue shark CPUE in this fishery. We suggest that either or both of these models could serve as comparison standards for commercial logbooks.

Comparison of Logbook Reports of Incidental Blue Shark Catch Rates by Hawaii-based Longline Vessels to Fishery Observer Data by Application of a Generalized Additive Model

William A. Walsh

Pierre Kleiber

Abstract

A Generalized Additive Model (GAM) of catch-per-unit-effort (CPUE) for blue shark, Prionace glauca, was fitted to data gathered from March 1994 through December 1997 by National Marine Fisheries Service (NMFS) observers stationed aboard Hawaii-based commercial longline vessels (N = 2010 longline sets). Its coefficients were then applied to the values of predictor variables contained in logbook records from the remainder of fishery-wide effort during the study period (N = 41319 longline sets). The objective was to determine whether predictions generated by such a model could serve in lieu of observers on the much larger fraction of longline trips that do not carry an observer (approximately 95%). A series of data evaluation procedures led to rejection of 2.6% of the logbook data as apparently false or inaccurate. After deleting these sets, the relationship between logbook CPUE data and GAM predictions was expressed by

$$\log_e(Y+1) = 0.7952 \log_e(X+1) - 0.0586,$$

where Y = blue shark CPUE from logbooks and X = GAM predictions of CPUE ($r^2 = 0.307$; N = 40243). This regression indicated that logbook CPUE data varied directly with, but were generally less than, GAM predictions. A time series of monthly mean GAM predictions agreed much more closely with the trends in monthly mean CPUE values from logbooks than the corresponding plot of uncorrected monthly mean values from observers. Patterns of correspondence between logbook trends and GAM predictions were further refined by plotting the time series according to the type of fishing effort (e.g. tuna-directed, swordfish-directed). The highest logbook mean CPUE values, the highest mean GAM predictions, and the greatest differences between the two occurred consistently in mid-year on swordfish trips. In contrast, logbook mean CPUE values and mean GAM predictions were closest for tuna effort, but this reflected an order of magnitude reduction in the response scale rather than closely similar trends. We conclude that prediction with a GAM fitted to fishery observer data is a useful monitoring technique for the Hawaii-based commercial longline fishery. It allowed us to gain insight into fleet-wide and individual logbook reporting practices, to estimate the relationship between logbook data and predicted values, to characterize the bias in this relationship, and to identify patterns specific to each major sector of the fishery.