

# 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<sup>a,\*</sup>, Pierre Kleiber<sup>b</sup>, Marti McCracken<sup>b</sup>

<sup>a</sup>Joint Institute for Marine and Atmospheric Research, University of Hawaii, 2570 Dole Street, Honolulu, HI 96822, USA

<sup>b</sup>Honolulu Laboratory, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2570 Dole Street, Honolulu, HI 96822-2396, USA

Received 7 November 2000; received in revised form 14 May 2001; accepted 4 July 2001

---

## Abstract

A generalized additive model (GAM) of blue shark, *Prionace glauca*, catch rates (catch per set) was fitted to data gathered by National Marine Fisheries Service (NMFS) observers stationed aboard Hawaii-based commercial longline vessels ( $N = 2010$  longline sets) from March 1994 to December 1997. Its coefficients were then applied to the values of predictor variables, which were also contained in logbook records that described the remainder of fishery-wide effort during the study period ( $N = 41\,319$  longline sets). The objective was to determine whether predictions generated by such a GAM could serve in lieu of observers on the large fraction of longline trips that do not carry an observer (approximately 95%). After deleting data considered false or inaccurate, much of which was associated with a small number of vessels, the relationship between catch rates as reported in logbooks and GAM predictions was expressed by

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

where  $Y$  is the catch rate (i.e., the number of blue shark caught per set) and  $X$  the GAM predictions ( $R^2 = 0.307$ ,  $N = 40\,243$ ). Patterns of correspondence between logbook trends and GAM predictions were further refined by plotting the trends according to the type of fishing effort (e.g., tuna- or swordfish-directed). The highest mean catch rates reported in logbooks, the highest mean GAM predictions, and the greatest differences between the two occurred consistently in mid-year on swordfish trips. In contrast, mean values from logbooks and mean GAM predictions were closest for tuna-directed effort, but this reflected an order of magnitude reduction in the scale of catch rates rather than closely similar trends. A bootstrapping algorithm developed for the GAM yielded an estimate of 23.9% under-reporting for the study period, with approximate 95% prediction limits of 15.4–28.9%. 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 bias in this relationship, and to identify patterns specific to each major sector of the fishery.

© 2002 Elsevier Science B.V. All rights reserved.

**Keywords:** Blue shark; Generalized additive model; Incidental catch rates; Fishery observer data; Logbook reliability

---

## 1. Introduction

A previous paper (Walsh and Kleiber, 2001) presents a generalized additive model (GAM) of catch

---

\* Corresponding author. Tel.: +1-808-983-5346;

fax: +1-808-983-2902.

E-mail address: william.walsh@noaa.gov (W.A. Walsh).

rates (catch per longline set) for blue shark, *Prionace glauca*, in the Hawaii-based longline fishery. The model was fitted to data gathered by National Marine Fisheries Service (NMFS) observers stationed aboard Hawaii-based commercial longline vessels from March 1994 to December 1997. It was recommended for evaluation as a management tool because the underlying data were considered reliable; it explained a high proportion of the deviance of blue shark catch rates, and its terms, spatio-temporal, environmental, and operational factors, were amenable to meaningful interpretation. These authors suggested that such a model could serve as a standard against which to compare incidental blue shark catches reported in the logbooks of commercial vessels.

This paper presents a fishery-wide expansion of a GAM of blue shark catch rates (see Section 2.2) to the entire Hawaii-based longline fishery. This application is predicated upon the fact that the logbooks submitted by commercial vessels to NMFS contain the same types of data that are gathered by the fishery observers and required as independent variables in the GAM. The objective of this paper was to determine whether model predictions could effectively represent “surrogate” observers on the large fraction of longline trips that do not carry an observer (approximately 95%). Therefore, we predicted blue shark catch rates with a GAM fitted to observer data and then compared these values to the corresponding logbook data. Records from longline sets with large differences between these values were examined to assess whether data were false or erroneous. We then determined the relationships between logbook data and GAM predictions, overall and for each of the major types of fishing effort. The latter evaluation was performed because differences in fishing practices associated with specific target species are known to affect incidental catches of non-target fishes (He et al., 1997).

A study of this sort, with its critical assessment of data quality, was appropriate because the logbooks represent the principal source of information regarding incidental catch of blue sharks in this fishery. In fact, although incidental, blue shark has been the most numerous species in the entire catch of this fishery during this decade (He et al., 1997; Ito and Machado, 1999; Walsh and Kleiber, 2001). The results presented herein are timely since determination of the magnitudes of bycatch and bycatch mortality was recently

named as the foremost national objective for fisheries bycatch management (Schmitt, 1998; ‘bycatch’ referred to both fishery discards and retained incidental catches). The results are expected to prove useful in monitoring logbook reporting practices and in blue shark stock assessments founded upon catch data provided by observers, in logbooks, or both.

## 2. Materials and methods

Full details of the development and interpretation of a GAM of blue shark catch rates are presented in Walsh and Kleiber (2001). These authors employed methods similar to those of Bigelow et al. (1999). Specifics concerning this paper follow.

### 2.1. Comparison of blue shark catch rates from logbooks and observers

The correspondence between blue shark catch rates as reported in logbooks and by observers serving on the reporting vessels from March 1994 to December 1997 ( $N = 2010$  longline sets) was assessed by computing the linear regression of the former on the latter. Close agreement was expected because the same fish were being enumerated, the presence of an observer represented a form of oversight, and captains might confer with observers (or vice versa) before making log entries. This computation was intended to verify the comparability of logbook and observer data, at least under conditions favoring agreement.

### 2.2. GAM development and predictions

Evaluation of the correspondence between predicted and reported incidental blue shark catches was initiated by fitting a GAM to data gathered by fishery observers ( $N = 2010$  longline sets), and then generating a set of predicted catches that corresponded to the total fishery-wide effort for the study period. The latter entailed application of the coefficients from the GAM to the values of predictor variables contained in logbook records by use of the ‘predict.gam’ function in S-PLUS (MathSoft, 1996; Hastie, 1992). Longline sets conducted in the presence of observers and those with missing values for any predictor variable(s) were excluded from the analyses. The result was an

expanded data set comprised of a suite of predictor variables, the reported blue shark catch, and the corresponding predicted catch for each gear set ( $N = 41\,319$  longline sets). Because the GAM described catch rates in the presence of an observer, its predictions could conceivably represent surrogate observers for the greater number of longline sets conducted in their absence (approximately 95%).

The GAM used for prediction differed from that described by Walsh and Kleiber (2001) in the following respects. This GAM included eight (latitude, longitude, sea surface temperature, date, number of hooks, vessel length, visible proportion of the moon, and angle of the sun to the horizon) rather than nine predictor variables because 60.7% of the logbook records, derived largely from tuna-directed sets, did not include the ninth variable (numbers of light sticks per hook), and it was considered appropriate to delete this variable since it would have been impossible to distinguish missing data from zeroes. In addition, satellite sea surface temperature (SST) data obtained from the Climate Modeling Branch of the National Oceanic and Atmospheric Administration (NOAA) were substituted for direct temperature measurements because 30.2% of the logbook entries lacked this information. The linear regression of predictions from a GAM that used satellite SST data on those from an equivalent GAM with direct SST measurements was

$$Y = 1.001X \quad (1)$$

where  $Y$  is the predictions obtained with satellite data and  $X$  the predictions obtained with direct SST measurements ( $R^2 = 0.949$ ,  $N = 2010$ ). Finally, this GAM was based upon cubic splines rather than loess fitting procedures; splines exhibit superior properties regarding testing for goodness-of-fit relative to the latter (Venables and Ripley, 1999). The nonlinear degrees of freedom were determined iteratively to approximate the residual mean deviance from the loess model as closely as possible. The linear regression of predictions from the spline-based GAM on those from one fitted by loess methods was

$$Y = 0.995X \quad (2)$$

where  $Y$  is the predictions obtained with a GAM fitted by splines and  $X$  the predictions obtained by the loess method ( $R^2 = 0.991$ ,  $N = 2010$ ). See Hastie (1992) for details regarding these fitting procedures.

### 2.3. Comparison of GAM predictions with logbook catch reports

The detailed analysis of logbook data was initiated by transforming reported and predicted blue shark catch values to  $\log_e(X + 1)$ , plotting the former on the latter, and then assessing their correspondence in a six-step process, which in every case was based upon a careful examination of the archived original logbook forms. First, all longline sets with a reported catch of zero but a predicted catch of 25 or more blue sharks were sorted by vessel permit numbers. All trip records from vessels with 10 or more of these sets during the entire study period were then examined for evidence of under-reporting. Typical examples of circumstantial evidence included a markedly altered reporting pattern after a vessel changed its captain, or a series of trips under a certain captain that reported zero blue shark catches despite fishing by methods and in areas where high catch rates would be expected. The second step entailed examination of the records of all trips that included two or more sets with a reported catch of zero but a predicted catch of 25 or more blue sharks. The third step consisted of examination of the records of all trips with any sets that had a reported catch of zero but a predicted catch of 50 or more blue sharks. The fourth stage consisted of examination of the records of all trips north of  $30^\circ\text{N}$  that reported zero blue sharks for the trip, regardless of the GAM predictions. This check was included because Nakano et al. (1985) and Pearcy (1991) demonstrated that blue sharks are abundant at this latitude and northward. The fifth stage was an examination of all sets that yielded 100 or more blue sharks to determine whether abnormal conditions (e.g., gear malfunction or lengthy soak time) contributed to the high catch. The final stage consisted of examinations of all catch records submitted by captains whose reports were deemed questionable on the basis of any of the procedures described above and all records from the vessels involved if more than one captain had been in command during the study period. The intention was to try to determine whether any inappropriate logbook reporting was attributable primarily to the vessel or a specific captain(s). Trips regarded as examples of non-reporting or other forms of inaccuracies were deleted from the data set. Such judgements were performed conservatively; data were deleted only

when there appeared to be strong indications of inaccuracies and after verification against the archived original reports. The linear regression of reported blue shark catch rates on predicted blue shark catch rates was computed after completing this data evaluation. Several standard residuals plots (Draper and Smith, 1981) were examined to detect sources of variation in this regression.

#### 2.4. Effects of special factors on correspondence between predicted and logbook catch rates

Logbook data were also examined in relation to specific fishing practices and other factors that may affect catch rates. The NMFS Honolulu Laboratory categorizes longline trips on the basis of target species as swordfish-, tuna-, or mixed-species-directed (He et al., 1997; Ito and Machado, 1999). Monthly mean catch rates from logbooks and GAM predictions were plotted to assess the applicability of the GAM to each trip type. Logbook reports of blue shark catches were also examined to assess whether, and if so to what extent, the reporting practices of individual vessels changed in relation to the presence of observers.

#### 2.5. Bootstrapping algorithm for prediction intervals

The comparisons of GAM results and logbook reports permitted us to approximate total under-reporting of blue shark for the study period, overall and by trip type. We used bootstrap methods to estimate prediction intervals. This involved resampling residuals from an under-smoothed curve and adding these to the fitted values from an over-smoothed curve, as described in Davison and Hinkley (1997). The over-dispersion in the data was modeled by use of standardized Pearson's residuals. The Pearson's  $\chi^2$ -statistic divided by its degrees of freedom was used as the estimate of the dispersion parameter  $\kappa$ .

The principal inherent complexity of this analysis was that blue shark catches obtained on different sets within a trip were likely correlated. Initial attempts to resample the Pearson's residuals by set did not succeed in reconstructing the error structure. Therefore, we calculated the mean standardized Pearson's residual for each trip, and then obtained an adjusted error from the error minus the mean error, in order to reconstruct the original error structure.

Given the preliminary findings, the estimation was computed as follows, with the observed blue shark catch on set  $h$  of trip  $t$  represented by  $y_{th}$ , with  $h = 1, \dots, H_t$  and  $t = 1, \dots, T$ , and the corresponding fitted value and standardized Pearson residual represented by  $\hat{\mu}_{th}$  and  $\varepsilon_{th}$ , respectively. First, we fitted an under-smoothed GAM and calculated the standardized Pearson's residuals for observed sets. Secondly, we calculated the mean of the Pearson's residuals  $\bar{\varepsilon}_t$  and the adjusted residuals  $\varepsilon'_{th}$  for each trip. We then fitted an over-smoothed GAM to the observer data, used simple random sampling to generate a sample of mean residuals for the total number of trips and adjusted residuals for the total number of sets, and created a bootstrap sampled response, defined as  $y^*_{th} = \hat{\mu}_{th} + \varepsilon^*_{th} \sqrt{\kappa \hat{\mu}_{th}}$  where  $\varepsilon^*_{th} = \varepsilon_t^* + \varepsilon'_{th}$ . We fitted the originally specified GAM to the generated responses, and for each sum of interest (e.g., the study period total and totals by trip types) calculated the sum of the generated observed  $y^*_+$  and fitted values  $\mu^*_+$  to obtain the statistic

$$d^*_+ = \frac{y^*_+ - \hat{\mu}^*_+}{\sqrt{\kappa \hat{\mu}^*_+}} \quad (3)$$

We repeated the random sampling and subsequent steps 999 times, ordered the values of  $d^*$ , and calculated  $1 - \alpha$  prediction intervals according to the Bonferroni principle from

$$\begin{aligned} \hat{\mu}_+ + d^*_{+, ((R+1)\alpha)} \sqrt{\hat{\mu}_+ + \kappa}, \\ \hat{\mu}_+ + d^*_{+, ((R+1)(1-\alpha))} \sqrt{\hat{\mu}_+ + \kappa} \end{aligned} \quad (4)$$

where  $R$  is the number of bootstrap replicates (i.e., 999). Diagnostics indicated that this algorithm effectively reconstructed the original error structure.

### 3. Results

#### 3.1. GAM development

The GAM (Table 1) was fitted to the full observer data set ( $N = 2010$  longline sets) to conform as closely as possible to that in Walsh and Kleiber (2001). There were no attempts to optimize the order of entry of predictors in terms of the Akaike information criterion (AIC), or to revise GAM structure other than as described previously. As such, vessel length yielded

Table 1

Analysis of deviance of an eight-variable GAM of blue shark catch rates (catch per set)<sup>a</sup>

Predictor variable	$\Delta$ AIC	$\Delta$ Residual deviance	d.f.	$F_{\text{enter}}$	$P$
Latitude	13508.17	13523.970	6.90	158.010	0
Longitude	5265.92	5280.152	7.12	93.799	0
Temperature	2610.15	2623.811	6.83	59.286	0
Date	947.60	961.633	7.02	23.448	0
Hooks	374.232	388.301	7.03	9.802	$4.2 \times 10^{-12}$
Vessel length	417.998	431.927	6.96	11.718	$1.3 \times 10^{-14}$
Solar index	104.018	117.924	6.95	3.226	0.002
Lunar index	58.346	72.356	7.00	1.974	0.055

<sup>a</sup> The reductions in the AIC and residual deviance, degrees of freedom, and the  $F$ -test and associated significance, are presented for each term. Null deviance: 32 473.45, d.f.: 2009. Residual deviance: 9073.373, d.f.: 1953.179. Pseudo- $R^2$ : 0.721.

a greater reduction in the residual deviance than the number of hooks per set despite later entry into the model, and the lunar index was only marginally significant as a predictor ( $P = 0.055$ ). The predictions generated by this GAM were highly correlated ( $r = 0.938$ ,  $P \ll 0.001$ ) with those from the model in Walsh and Kleiber (2001), and its pseudo-coefficient of determination (pseudo- $R^2 = 0.721$ ) was identical to that in Walsh and Kleiber (2001).

### 3.2. Correspondence between logbook and observer data—observed trips

The paired blue shark catch rate data from logbooks and fishery observers (Fig. 1) were significantly correlated ( $r = 0.818$ ,  $N = 2010$  sets,  $P \ll 0.001$ ), but at least three factors contributed to departure from full agreement. The first involved under-reporting on three trips (7–13 sets per trip) when logbooks reported zeroes but observers reported 4–169 blue sharks per set. The second involved over-reporting on one trip (15 sets) that was revealed by points clustered above the expected  $45^\circ$  angle. These resulted from logbook entries of blue sharks as both finned and released, rather than finned or released. After deleting these four trips, the regression of logbook on observer data was

$$Y = 0.949X + 0.735 \quad (5)$$

where  $Y$  is the logbook catch per set data and  $X$  the corresponding observer data ( $R^2 = 0.813$ ,  $N = 1963$ ). The 95% confidence interval about the regression coefficient (0.929–0.969) indicated that blue shark catches were slightly under-reported in logbooks. The third factor was that the observer and logbook data

sets did not agree closely for sets reported as zero blue shark catches. Logbooks and observers reported 205 (10.4%) and 99 sets (5.0%) with zero blue sharks, respectively. Observers reported 10 or fewer blue sharks on 88% of these, but there were three sets when observers reported 27–230 blue sharks while logbooks reported zero. These differed from the cases of non-reporting because the logbooks reported positive blue shark catches on the preceding and subsequent sets of these trips. Finally, there were 32 sets when observers reported zero blue sharks but logbooks reported 1–12 blue shark(s).

### 3.3. Comparison of GAM predictions with logbook catch reports—unobserved trips

Blue shark comprised nearly one-fourth of the total reported longline catches on trips without observers throughout the study period (Table 2). The uncorrected mean was 7.85 blue shark per set. There were 7754 sets (18.8%) with logbook entries of zero blue sharks.

The evaluation procedures (Table 3) led to rejection of 2.6% of the logbook data as apparently false or inaccurate. The preponderance (94%) were sets reported as zero blue shark catches. Most of these questionable data were associated with eight particular vessels. Among the vessels in question, one changed captain during 1997, after which its average reported catch rate increased 18-fold. This was the only instance that the conduct of a captain appeared to be distinguishable from that of a vessel. Zeroes still comprised 16.8% of the logbook catch data remaining after these procedures.

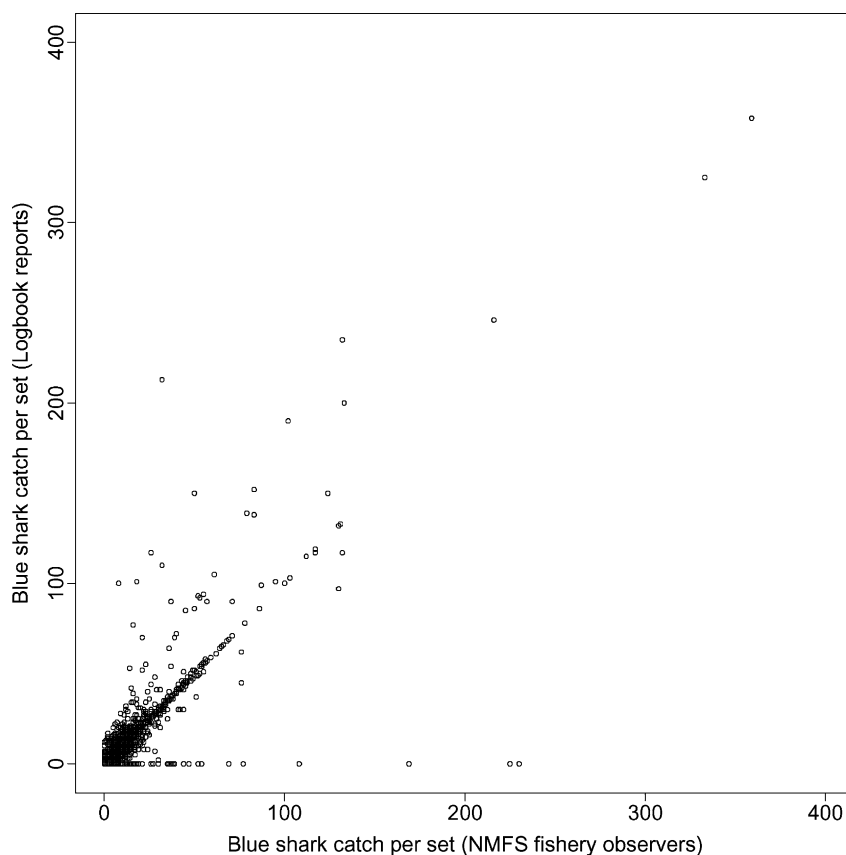


Fig. 1. Relationship between blue shark catch per set as reported in logbooks and by NMFS fishery observers.

The scatter plot of logbook data on the GAM predictions (Fig. 2) revealed a broad band of increase. The regression of the latter on the former (Table 4) was expressed by

$$\log_e(Y + 1) = 0.7952 \log_e(X + 1) - 0.0586 \quad (6)$$

where  $Y$  is the blue shark catch per set from logbooks and  $X$  the GAM predictions ( $R^2 = 0.307$ ,  $N = 40\,244$ ). The regression coefficient indicated that logbook data varied directly with, but were generally less than, GAM predictions. The data used to fit this regression included 130 sets with predicted catches

Table 2

Summary of effort and catches on trips of Hawaii-based commercial longline vessels without NMFS fishery observers from March 1994 to December 1997<sup>a</sup>

Year	Trips	Vessels	Gear sets	Total catch	Blue shark total catch	Percentage	Catch per set
1994–1997	4060	122	41319	1364462	324528	23.8	7.85
1994	819	122	7953	239522	74667	31.2	9.39
1995	1122 <sup>b</sup>	110	11180	387374	89078	23.0	7.97
1996	1085 <sup>c</sup>	104	11003	339765	89990	26.5	8.18
1997	1114	105	11183	397801	70793	17.8	6.33

<sup>a</sup> Total catch refers to all species of fishes listed in the logbook; other entries refer to blue shark.

<sup>b</sup> 45 trips begun in 1995 continued into 1996.

<sup>c</sup> 35 trips begun in 1996 continued into 1997.

Table 3  
Summary of a six-stage logbook data evaluation process

Criterion <sup>a</sup>	Data deleted		Data retained		Comments on deletions
	Trips	Sets	Trips	Sets	
Initial data set			4060	41319	
$BS_{GAM} \geq 25$ and $BS_{LOG} = 0$ , $\geq 10X$ per vessel	99	920	3961	40399	Logbooks from 17 vessels presented questionable data; 94% came from eight vessels
$BS_{GAM} \geq 25$ and $BS_{LOG} = 0$ , $\geq 2X$ per trip	8	85	3953	40314	One vessel deployed 56 sets on seven swordfish trips; all reported as $BS_{LOG} = 0$
$BS_{GAM} \geq 50$	2	29	3951	40285	One trip with 15 (of 19) sets as $BS_{LOG} = 0$ ; $BS_{GAM}$ range: 57.3–70.9
Latitude $\geq 30^\circ N$	3	40	3948	40245	Three trips from 34 to 36°N; all 40 sets as $BS_{LOG} = 0$
$BS_{LOG} \geq 100$	0	0	3948	40245	High catches did not result from abnormal conditions
Histories of vessels and captains	0	1	3948	40244	One logbook recording error

<sup>a</sup>  $BS_{GAM}$  and  $BS_{LOG}$  denote blue sharks per set as predicted by the GAM and reported in logbooks, respectively.

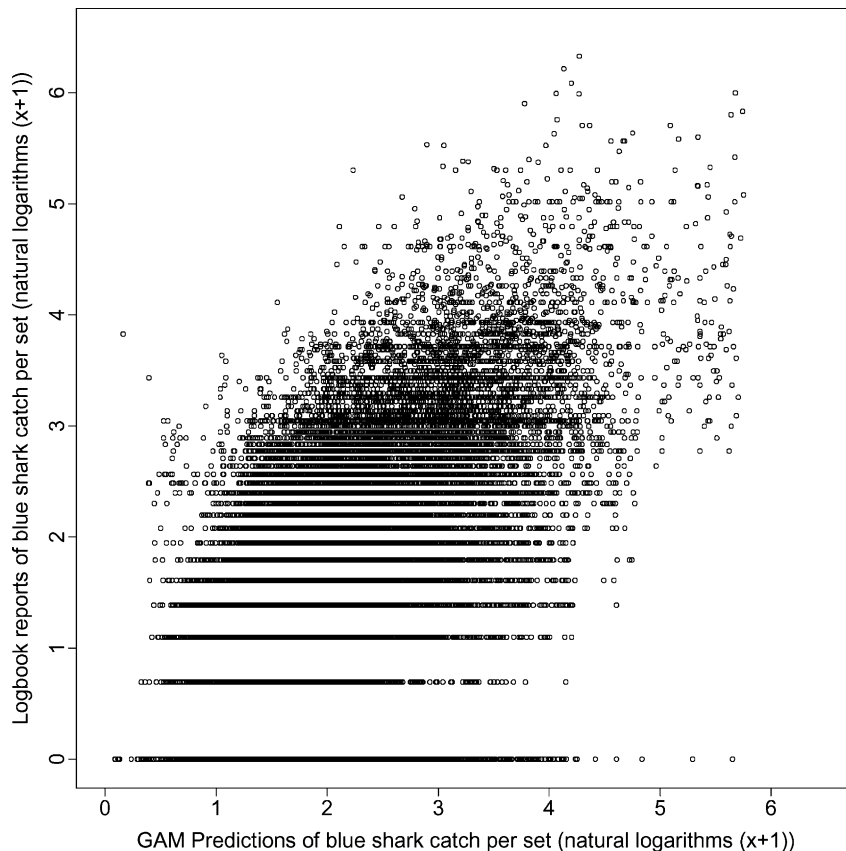


Fig. 2. Relationship between blue shark catch per set as reported in logbooks and as predicted by a GAM after transforming both variables to  $\log_e(X + 1)$ .

Table 4

Regression of logbook reports of blue shark catches per set on GAM predictions after transforming both variables to  $\log_e(X + 1)^a$ 

	d.f.	Value	Sum of squares	Standard error	Mean square	<i>t</i>	<i>F</i>	<i>P</i>	<i>R</i> <sup>2</sup>
Source of variation									
Regression	1		13556.51		13556.51		17837.62	0	
Residuals	40242		30583.74		0.76				
Parameter									
Intercept		−0.0586		0.0132		−4.4564		0.000	0.307
Coefficient		0.7952		0.0060		133.5576		0.000	

<sup>a</sup> 95% CI (regression coefficient): 0.7834–0.8070.

of 25 or more blue sharks but logbook reports of zero. These sets were retained because the original records provided insufficient evidence to warrant deletion. The much lower coefficient of determination of this regression compared to that which related logbook catch rates to those from observers reflected the

effects of both questionable but retained logbook values and the 20.5-fold increase in sample size ( $40\,244/1963 = 20.5$ ).

The residuals from the preceding regression (Fig. 3) exhibited a consistent seasonal trend. Positive monthly mean residuals indicated that logbook reports of catch

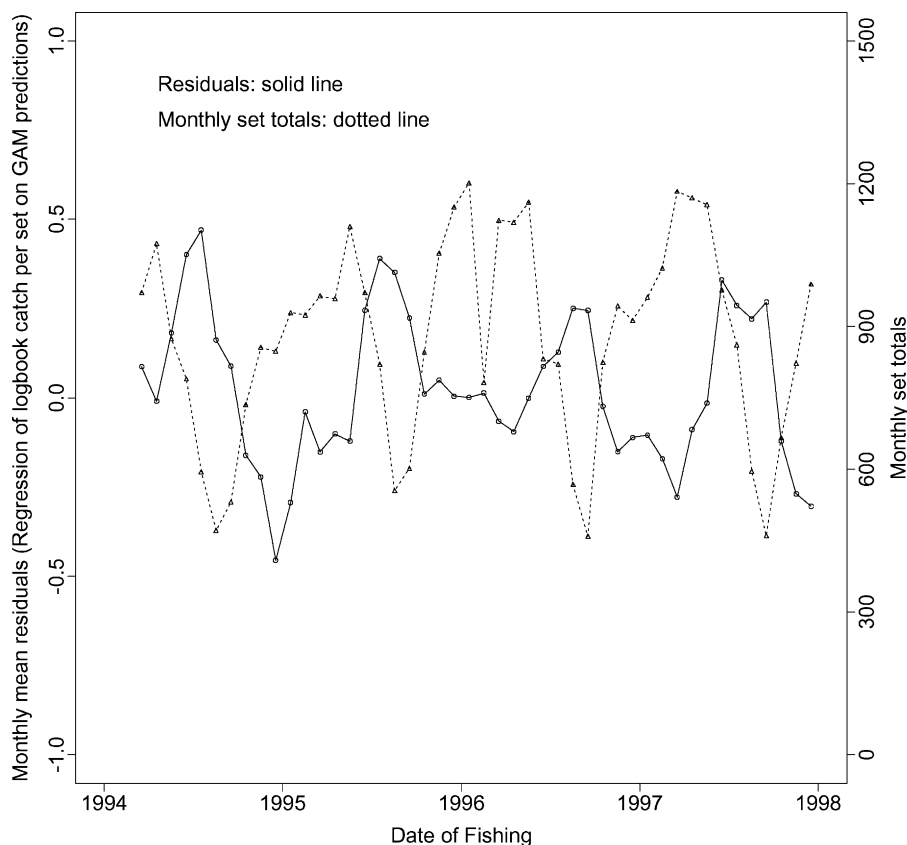


Fig. 3. Monthly mean residuals from the linear regression of blue shark catch per set as reported in logbooks on the predictions from a GAM (solid line with circles) along with monthly fishery-wide set totals (dotted line with triangles) from March 1994 to December 1997.



per set were high relative to predictions in mid-year, and vice versa. For example, the ratio of the sums of the reported to the predicted blue shark catches for June–August of all years, when the residuals were positive, was 0.99, whereas that for October–December of 1994, 1996, and 1997, when these were negative, was 0.65. This pattern reflected an inverse relationship with effort; monthly mean residuals were negatively correlated with set totals ( $r = -0.483$ ,  $P \ll 0.001$ ).

The monthly mean blue shark catch rates reported by observers and in logbooks were plotted with the mean GAM predictions (Fig. 4) to further characterize the temporal variation revealed by the residuals. Monthly mean catch rates from the logbooks and observers were significantly correlated ( $r = 0.305$ ,  $P = 0.042$ ), with similar patterns that included peaks in 1994 and 1995 and smaller increases in 1996. However, peaks did not always coincide, those in the observer data were much

more pronounced, and the similarities did not persist throughout the study period. There was lack of coincidence in 1994, when the main peaks in the two sets were offset by 1 month. The monthly mean catch rate patterns also differed in 1997, although the prominent observer data peak in January 1997 primarily reflected the activity of one vessel that averaged 94 blue sharks per set on an eight-set trip. In contrast, predictions from the GAM approximated the monthly logbook trends more closely than the uncorrected observer data ( $r = 0.410$ ,  $P = 0.005$ ), which were sometimes affected by small sample sizes. The predictions tended to track closely above the logbook monthly means. The highest peak in 1994 coincided with the peak prediction from the GAM, without the 1-month offset apparent in the uncorrected observer data. The heights of the predicted peaks in 1996 and 1997 were also less than those of the mean observer catch rates. The GAM did not reproduce

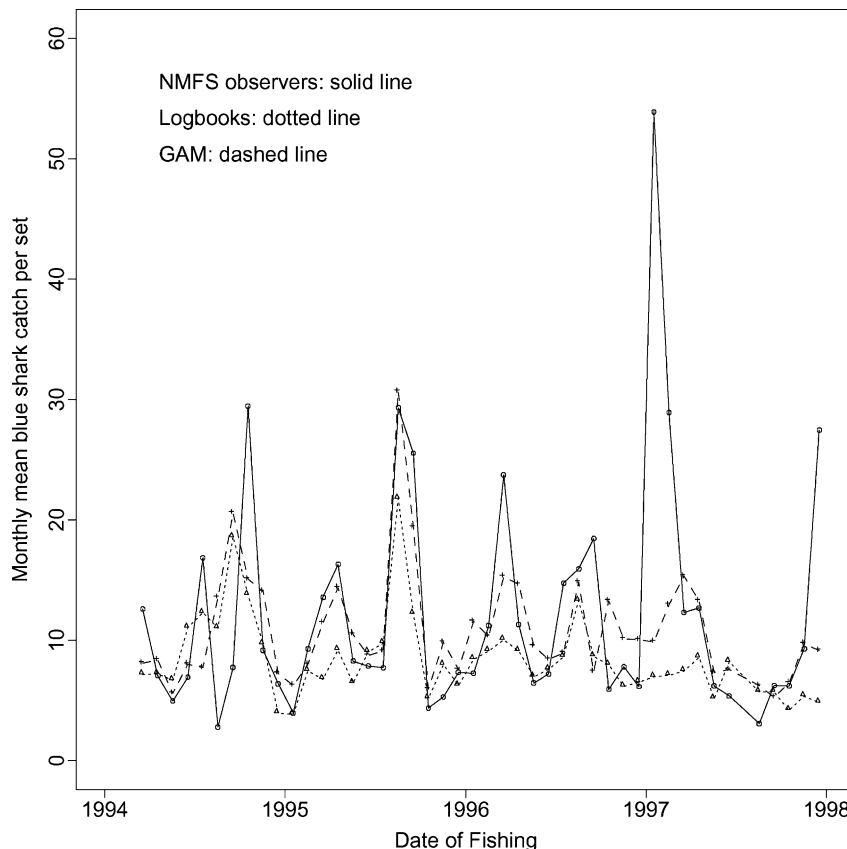


Fig. 4. Monthly mean blue shark catch per set as reported by NMFS fishery observers (solid line with circles), in logbooks (dotted line with triangles), and as predicted by a GAM (dashed line with crosses) from March 1994 to December 1997.

the two 1997 peaks in the observer data; in fact, the December 1997 peak in the observer data coincided with a slight downturn in the GAM predictions that was also apparent in the logbook trend.

### 3.4. Effects of fishing trip types

GAM predictions for each specific trip type (Figs. 5–7) exhibited many similarities to the mean logbook catch rates but on disparate scales and with characteristic patterns of error. The highest monthly mean logbook values were observed each summer on swordfish sets (Fig. 5); the GAM reproduced these features but at higher levels ( $r = 0.856$ ,  $P \ll 0.001$ ). Both logbook catch rates and GAM predictions for tuna sets (Fig. 6) exhibited an upward trend throughout the study period, but the closeness of the two sets

of values primarily reflected the order of magnitude reduction in the response scale relative to swordfish ( $r = 0.518$ ,  $P < 0.001$ ). The GAM also reproduced many of the features of logbook catch rates for mixed-species-directed sets ( $r = 0.708$ ,  $P \ll 0.001$ ), especially during 1996 and 1997 (Fig. 7). Its error consisted primarily of higher predicted than reported values in winter and early spring.

The composition and magnitude of fishing effort exhibited consistent seasonal patterns throughout the study period (Fig. 8). Tuna-directed sets comprised the largest fraction of fishing effort in all but 4 months and 58.9% of total effort throughout the study period. Swordfish-directed sets comprised the largest fraction during the first 3 months but declined to considerably lower levels by mid-1995, and then remained approximately stable as the smallest fraction of monthly effort

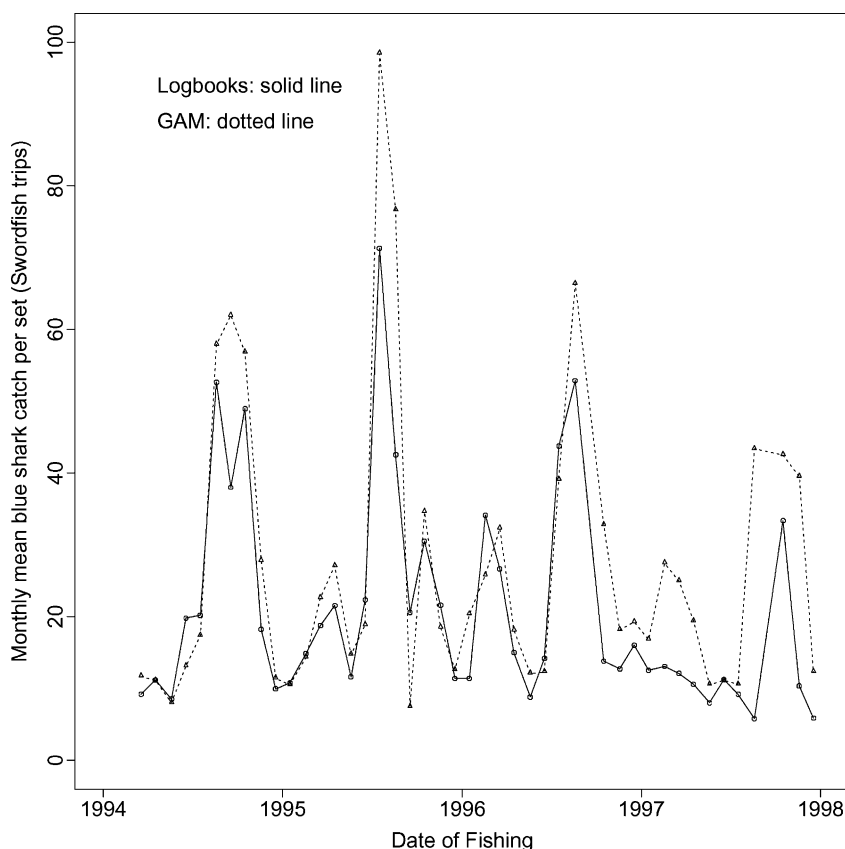


Fig. 5. Monthly mean blue shark catch per set on swordfish trips as reported in logbooks (solid line with circles) and as predicted by a GAM (dotted line with triangles) from March 1994 to December 1997. Note response scale.

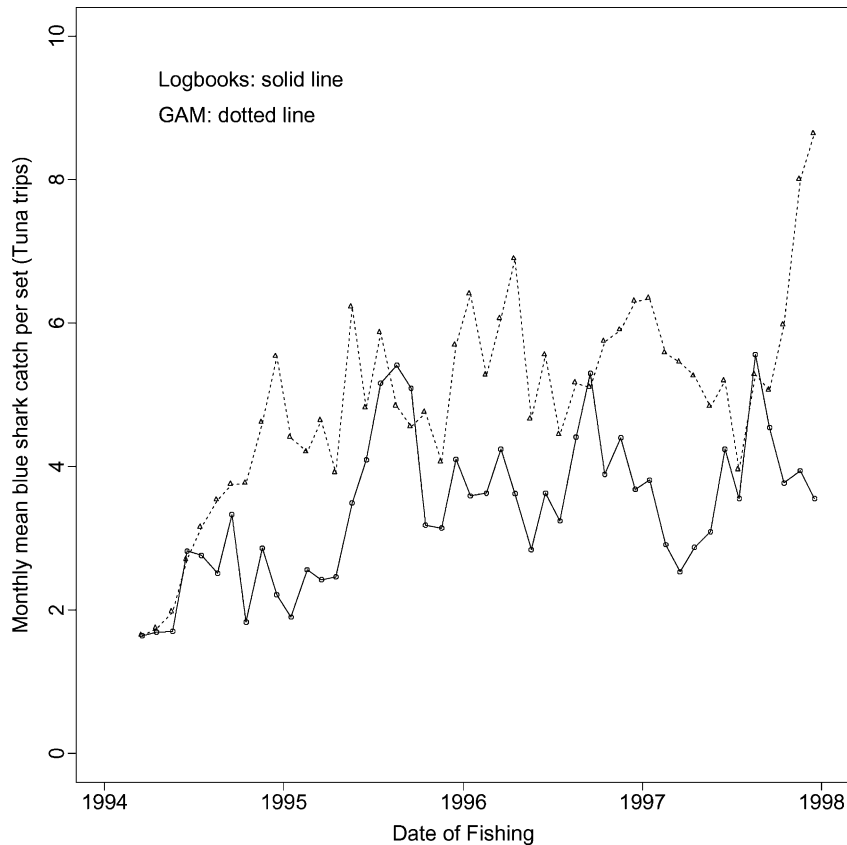


Fig. 6. Monthly mean blue shark catch per set on tuna trips as reported in logbooks (solid line with circles) and as predicted by a GAM (dotted line with triangles) from March 1994 to December 1997. Note response scale.

(17% of total effort). Both tuna and swordfish activity remained near their maximum levels from autumn to spring months. Mixed-species-directed sets (24.1% of total effort) were intermediate in number between tuna and swordfish sets each month after May 1995 and tended to peak in mid-year as tuna and swordfish activity declined towards their annual minima. It was noteworthy that the greatest differences between predicted and observed values tended to coincide with periods of low activity for swordfish and mixed-species trips, whereas large differences tended to coincide with high activity in the tuna sector. This was important because the actual discrepancy between model predictions and logbook reports for any given period is the product of the difference per set and the total number of sets.

Logbook reports listed greater numbers of zero blue shark catches on tuna sets than would have been

expected on the basis of trips with observers (see above). The overall 10.4% rate reported by logbooks in the presence of observers represented a weighted mean of 13.3% zeroes for tuna sets, 6.3% for swordfish sets, and 10.4% for mixed-species sets. The logbook data set included 21.1% zeroes for tuna sets, 8.4% for swordfish sets, and 12.1% for mixed-species sets; the difference in the proportion of zeroes was significant for tuna sets ( $\chi^2 = 40.15$ ,  $P \ll 0.001$ ).

### 3.5. Estimation of under-reporting on unobserved trips

The bootstrapping algorithm yielded an annual average under-reporting estimate of 26 507 blue shark on unobserved trips from March 1994 to December 1997 (Table 5). This would correspond to 23.9%, with approximate 95% prediction limits of

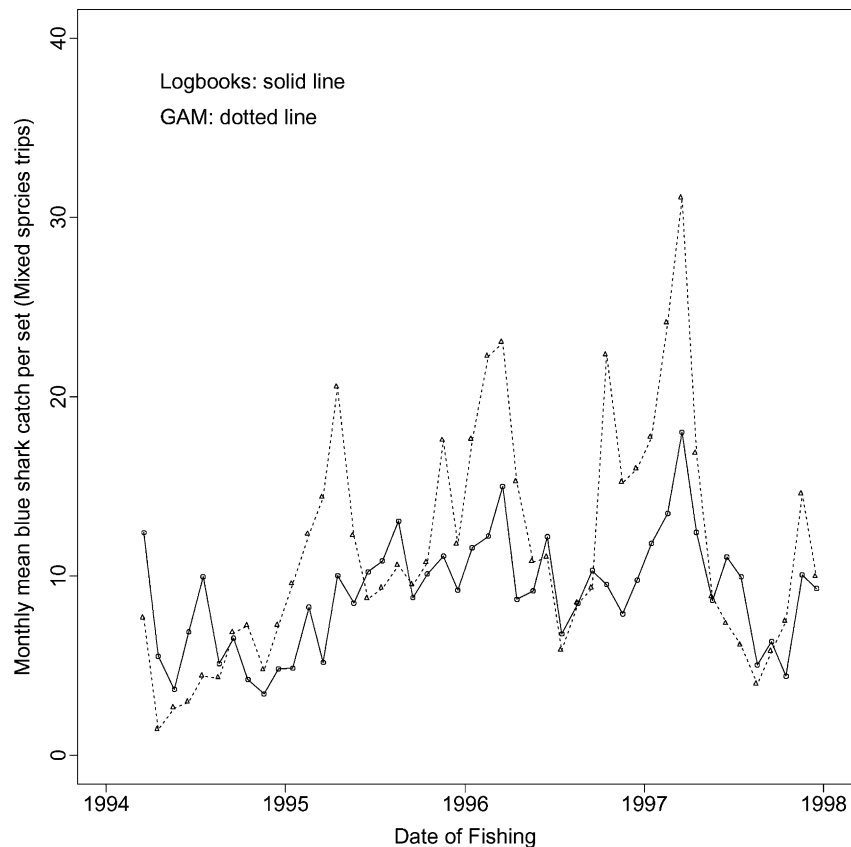


Fig. 7. Monthly mean blue shark catch per set on mixed-species trips as reported in logbooks (solid line with circles) and as predicted by a GAM (dotted line with triangles) from March 1994 to December 1997. Note response scale.

Table 5  
Summary of under-reporting of blue shark on unobserved trips in the Hawaii-based longline fishery<sup>a</sup>

Year	Trip type	GAM	Logbook	$\Delta$	95% Prediction interval	Estimated under-reporting (%)
1994–1997	All	424460	322850	101610	381990–454206	23.9 (15.4–28.9)
1994	Swordfish	60714	57551	3163	48334–80218	5.2 (–19.1–28.6)
	Tuna	13133	8952	4181	9869–18427	31.8 (9.2–51.4)
	Mixed-species	5334	6805	–1471	5021–8220	–27.6 (–35.5–17.2)
1995	Swordfish	49193	37767	11426	36690–58165	23.2 (–3.0–35.1)
	Tuna	32155	23011	9144	21042–38187	28.4 (–9.4–39.7)
	Mixed-species	36892	28224	8668	34247–53400	23.5 (6.7–40.5)
1996	Swordfish	27419	23207	4212	21932–34181	15.4 (–5.8–32.1)
	Tuna	37212	24903	12309	23824–43725	33.1 (–4.5–43.1)
	Mixed-species	57575	41842	15733	46919–69185	27.3 (10.8–39.5)
1997	Swordfish	20511	10496	10015	12323–23470	48.8 (14.8–55.7)
	Tuna	43070	26503	16567	24378–51481	38.5 (–8.7–48.5)
	Mixed-species	41252	33589	7663	34247–53400	18.6 (1.9–37.1)

<sup>a</sup> Total catch estimated by the GAM, the reported total catch, the difference between GAM and logbook totals, an approximate 95% prediction interval for the total catch, and point and interval (in parentheses) estimates for percent under-reporting are presented for each year and trip type.

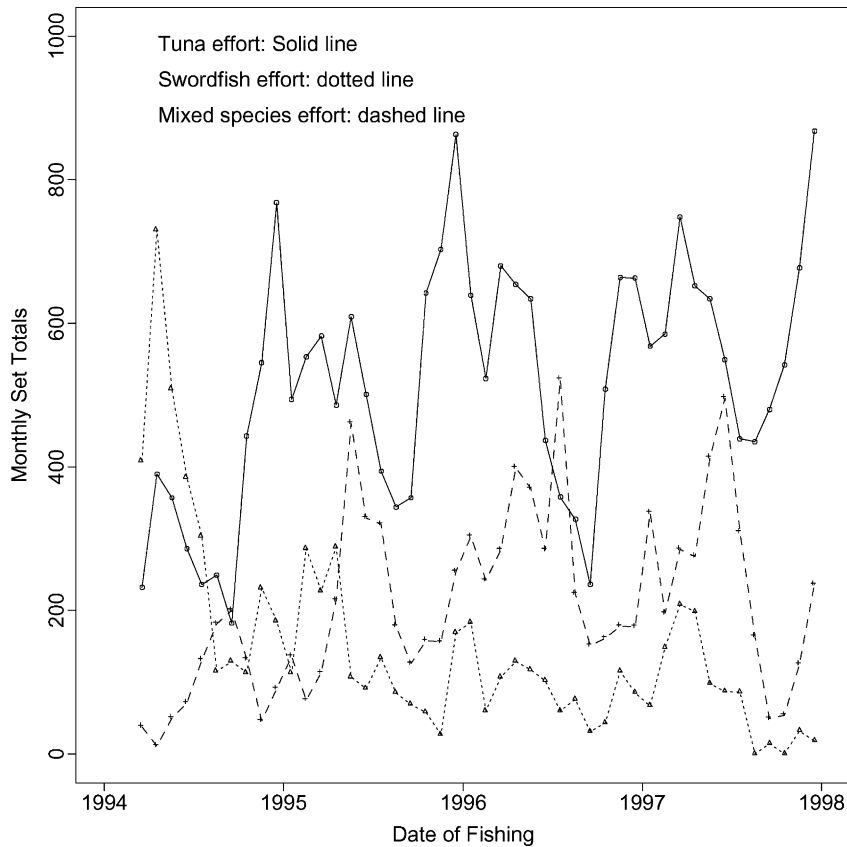


Fig. 8. Monthly set totals from March 1994 to December 1997 for tuna (solid line with circles), swordfish (dotted with triangles), and mixed-species (dashed line with crosses).

15.4–28.9%. Tuna-directed effort was the only sector of the fishery that exhibited indications of a trend, with increases each year in the difference between the GAM and logbook totals and the highest estimate for percent under-reporting each year except 1997, when the highest percentage was observed with swordfish trips.

The GAM predictions and logbooks were also compared for the deleted sets ( $N = 1075$ ) to assess their contribution to under-reporting. The GAM predicted a mean catch rate of 15.3 blue shark per set while the logbooks reported 1.6 blue shark per set. The estimate of total under-reporting on these sets was 14 766 blue shark. Thus, this relatively small number of sets (2.6% of the total effort) was the source of apparent under-reporting equivalent to 14.5% of the estimate of total under-reporting for the other 97.4% of the longline sets.

### 3.6. Effects of fishery observers on subsequent logbook reporting practices

Three cases exemplified improved reporting by previously non-compliant vessels after carrying a fishery observer. One vessel that had reported zero blue shark catches on eight consecutive trips before carrying an observer reported positive catches on 14 of its next 15 trips, and the one trip with zero blue sharks appeared reasonable for the area fished. A second vessel reported a 6.3-fold increase in blue shark catches after carrying an observer (computed as the average of nine subsequent trips). This vessel then changed captain and reported zero blue shark catches on eight of the next nine trips. A third vessel increased its reported blue shark catches 5.2-fold (computed as the average of eight subsequent trips) after carrying an observer. In contrast, however, one vessel that had

carried an observer reported zero blue shark catches on four trips later in the same year, one of which was a 13-set swordfish-targeted trip with GAM predictions of 45.9–56.4 blue sharks per set. These reports of zero or very low blue shark catches were among the deleted sets.

#### 4. Discussion

The results presented herein demonstrate that prediction of incidental blue shark catch rates with a GAM can serve at least four purposes. The method proved useful in identifying false or questionable logbook data, in clarifying the correspondence between logbook and observer data sets, in describing incidental catch patterns relative to types of fishing effort, and in improving understanding of logbook reporting behavior.

The direct comparison of observer and logbook data and the data evaluation procedures based on the initial GAM predictions revealed a low rate of non-reporting for blue shark, although certain problems were evident. The observation that incorrect records were submitted for three trips despite the presence of observers clearly demonstrated that some captains disregarded reporting responsibilities. However, the overall percentage of logbook data rejected as false or inaccurate (2.6%) suggested that most vessels met their responsibilities even without observers, although this value is probably an excessively conservative estimate of non-reporting. The first-stage evaluation criterion used to identify questionable data (i.e.,  $BS_{GAM} \geq 25$  and  $BS_{LOG} = 0$ ) required a catch slightly greater than three times the uncorrected mean catch rate, and other criteria were even more stringent. Still, it appears that reasonable compliance with regulatory requirements is the norm for incidental blue shark catches in this fishery.

The direct comparison of paired logbook and observer reports revealed three important points regarding data quality. The linear regression of logbook on observer blue shark catch rates demonstrated that the two were closely comparable. This result suggested that a GAM fitted to the observer data might accurately reproduce logbook catch trends. A countervailing issue was that the differences in numbers of zeroes reported by observers and in logbooks

suggested that observer data, logbook records, or both, contain some as yet inestimable degree of inaccuracy, some of which may simply reflect inherent difficulties of data gathering and record keeping at sea. Finally, the observation of three sets when observers reported 27–230 blue sharks but logbooks reported zeroes suggested that some logbook data were highly inaccurate, although this did not represent non-reporting per se because the other sets on these trips reported positive blue shark catches in their logbooks. Hence, it appears that even very high blue shark catches can be overlooked or misrecorded by human error.

The plot of logbook catch rates data on GAM predictions and the associated regression analysis revealed a strong relationship, but with a regression coefficient significantly less than a value of 1.0 that would have represented full agreement. This reflected the effects of an unknown remaining number of false zeroes, as well as other inaccuracies (e.g., logbook values that were under-reported, but not as zeroes) that remained undetected because the data evaluation procedures concentrated on identification of false zeroes. Moreover, it was not possible to partition the variance about the regression into pure error and lack of fit sums of squares (Draper and Smith, 1981) because the predictor variables formed unique combinations. Nevertheless, the seasonal pattern in the residuals suggested that some of its bias was comprehensible. It seems likely that a major underlying cause of inaccuracy was that captains and crews may have been engaged in higher priority activities than counting the incidental catch during particularly busy or productive times of the year. For example, the greatest differences between the logbook reports and GAM predictions for tuna-directed trips occurred during the first and fourth quarters (Fig. 6); the highest catch per unit effort (CPUE) for bigeye tuna is consistently attained during this period (R. Ito, personal communication). Hence, it does not seem necessary to invoke willful misconduct as an explanation for logbook inaccuracies under these circumstances.

The observation that the plots of uncorrected monthly mean catch rates from observers and logbooks exhibited both similarities and marked dissimilarities was not surprising. The allocation of observers to the different types of fishing effort was not closely similar to the overall percentages in the logbook data, and small observer sample sizes introduced distortions

in some months. Hence, monthly mean catch rate values from observers appear sufficient as rough indicators of general trends but insufficient for analytical or management purposes regarding incidental blue shark catch.

The plots of GAM predictions and mean logbook catch rates for the three trip types were important because these revealed both the agreement, representing the predictive utility of the GAM, and the timing and magnitude of error for each sector of the fishery. As such, sector-specific effects were delineated even in the context of a fishery-wide analysis based upon the continuous nature of the predictors. This argued in favor of the breadth of applicability of the GAM. These plots in combination with the time series of effort types also revealed fundamental differences among the fishery sectors in the correspondence between GAM predictions and mean logbook values; i.e., the greatest differences between predicted and reported values were observed during periods of high and low effort in the tuna and swordfish sectors, respectively.

The estimates of under-reporting generated by bootstrapping were important for three reasons. First, the estimate for total under-reporting was in reasonable agreement with a previous analysis of this fishery. He and Laurs (1998) evaluated blue shark reporting practices in this fishery from 1991 to 1995 and estimated that the overall average under-reporting rate was about 16%. Although the point estimate for under-reporting (23.9%) was greater, the 95% prediction interval presented herein (15.4–28.9%) encompasses this value. Agreement was not surprising given the 2-year data overlap. Secondly, He and Laurs (1998) inferred that a significant fraction of the under-reporting occurred in the tuna sector of this fishery. This inference is not intuitively obvious since blue shark CPUE on tuna sets is lower than in the other sectors of this fishery (Ito and Machado, 1999), but it is important because tuna sets comprise the largest sector of this fishery. Their inference was supported by the relatively high estimates generated by bootstrapping (28.4–38.5%). The observation that the percentage of zeroes in the logbook data for tuna sets (21.1%) was significantly greater than that reported in the presence of observers (13.3%), which presumably represents “optimal” logbook reporting, further supports the view that non-reporting occurred in this

sector of the fishery and requires further scrutiny. Finally, the comparison of the initial GAM predictions and logbook reports for the 2.6% of the longline sets that were deleted demonstrated that a small number of aberrant trips or individuals who habitually report the incidental catch inaccurately can be responsible for a relatively large proportion of the problems in estimation of the incidental catch.

The final aspect of the analysis also yielded useful results. The observations of behavior subsequent to the presence of observers demonstrated that some individuals did improve reporting practices, whereas others were apparently impervious to their responsibilities and exhibited no lasting effect. It would appear that GAM predictions could be used to assess whether additional observation might be needed on vessels that had submitted questionable logbook records.

By generating predictions of incidental blue shark catches with a GAM and conducting several statistical analyses, we gained insight into fleet-wide and individual logbook reporting practices, estimated the relationship between logbook data and predicted values, characterized bias in this relationship and suggested an underlying cause, and identified patterns specific to each major sector of the fishery. We conclude, therefore, that a GAM used as a predictive tool can effectively complement the activities of observers in the Hawaii-based longline fishery.

## Acknowledgements

The authors thank Keith Bigelow, Michael Musyl, and Samuel Pooley for reading and commenting on an earlier version of this manuscript, and Judith Kendig of the editorial staff of the NMFS Honolulu Laboratory for help in preparing the final manuscript. This paper is funded by Cooperative Agreement No. NA 67RJ0154\* from the National Oceanic and Atmospheric Administration.

## References

- Bigelow, K.A., Boggs, C.H., He, X., 1999. Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. *Fish. Oceanogr.* 3, 178–198.
- Davison, A.C., Hinkley, D.V., 1997. *Bootstrap Methods and their Application*. Cambridge University Press, New York.

- Draper, N.R., Smith, H., 1981. Applied Regression Analysis, 2nd Edition. Wiley, New York, 709 pp.
- Hastie, T., 1992. Generalized additive models. In: Chambers, J.M., Hastie, T.J. (Eds.), Statistical Models in Wadsworth and Brooks/Cole Advanced Books and Software. Wadsworth, Pacific Grove, CA, pp. 249–306.
- He, X., Laurs, M., 1998. Bycatch, finning, and economic value of blue shark in the Hawaii-based longline fishery. In: Gibble, N.A., McPherson, G., Lane, B. (Eds.), Shark Management and Conservation. Proceedings from the Sharks and Man Workshop of the Second World Fisheries Congress (Synopsis). Department of Primary Industries, Queensland, pp. 88–94.
- He, X., Bigelow, K.A., Boggs, C.H., 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fish. Res. 31, 147–158.
- Ito, R.Y., Machado, W.A., 1999. Annual Report of the Hawaii-based Longline Fishery for 1998. Administrative Report H-99-06. Honolulu Laboratory, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Honolulu, HI.
- MathSoft, 1996. S-PLUS Version 3.4. MathSoft, Inc., Seattle, WA.
- Nakano, H., Makiyara, M., Shimazaki, K., 1985. Distribution and biological characteristics of the blue shark in the central North Pacific. Bull. Fac. Fish. Hokkaido Univ. 36, 99–113 (in Japanese with English abstract).
- Pearcy, W.G., 1991. Biology of the transition region. In: Wetherall, J. (Ed.), Biology, Oceanography, and Fisheries of the North Pacific Transition Zone and Sub-Arctic Frontal Zone. NOAA Technical Report NMFS 105.
- Schmitt, R.A., 1998. An agenda for bycatch. Fisheries 23 (6), 6.
- Venables, W.N., Ripley, B.D., 1999. Modern Applied Statistics with S-Plus, 3rd Edition. Springer, New York.
- Walsh, W.A., Kleiber, P., 2001. Generalized additive model and regression tree analyses of blue shark (*Prionace glauca*) catch rates by the Hawaii-based commercial longline fishery. Fish. Res. 53, 115–131.