



# Environmental and spatial effects on the distribution of blue marlin (*Makaira nigricans*) as inferred from data for longline fisheries in the Pacific Ocean

NAN-JAY SU,<sup>1</sup> CHI-LU SUN,<sup>1,\*</sup> ANDRE E. PUNT<sup>2</sup> AND SU-ZAN YE<sup>1</sup>

<sup>1</sup>Institute of Oceanography, National Taiwan University, Taipei 10617, Taiwan

<sup>2</sup>School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195, USA

## ABSTRACT

Blue marlin is distributed throughout tropical and temperate waters in the Pacific Ocean. However, the preference of this species for particular habitats may impact its vulnerability to being caught. The relationship between spatio-temporal patterns of blue marlin abundance and environmental factors is examined using generalized additive models fitted to catch and effort data from longline fisheries. The presence of blue marlin, and the catch rate given presence, are modeled separately. Latitude, longitude, and sea-surface temperature explain the greatest proportion of the deviance. Spatial distributions of relative density of blue marlin, based on combining the probability of presence and relative density given presence, indicate that there is seasonal variation in the distribution of blue marlin, and that the highest densities occur in the tropics. Seasonal patterns in the relative density of blue marlin appear to be related to shifts in SST. The distribution and relative abundance of blue marlin are sufficiently heterogeneous in space and time that the results of analyses of catch and effort data to identify 'hotspots' could be used as the basis for time-area management to reduce the amount of blue marlin bycaught in longline fisheries.

**Key words:** blue marlin, distribution, environmental and spatial effects, generalized additive model

## INTRODUCTION

Blue marlin, *Makaira nigricans*, is a cosmopolitan species distributed throughout tropical, sub-tropical and temperate waters between 50°N and 50°S (Shaklee *et al.*, 1983). Most blue marlin are caught by longline vessels targeting tunas and billfishes, although small catches of blue marlin are also taken using surface gears, such as gillnets and harpoons, as well as by recreational, purse seine, and baitboat fisheries (Hinton, 2001). Analyses of genetic divergence based on allozymes, mtDNA and nuclear DNA (Graves and McDowell, 2003) and fishery catch rates suggest that blue marlin constitute a single stock in the Pacific Ocean (Kleiber *et al.*, 2003). The assumption of a single stock of blue marlin in the Pacific Ocean is also supported by the results of tagging experiments which have shown that blue marlin migrate long distances (>7000 miles in some instances) (Hinton, 2001).

Blue marlin is one of the most important bycatch species in the high seas fisheries of the Pacific Ocean (Molony, 2005). Bycatch species can be defined as species of minor economic value or which constitute a small (incidental) proportion of the catch. Such species do not determine the movement patterns of fishing vessels. However, the preferences of such species for particular habitat-related factors may impact their vulnerability to being caught. These factors include water temperature, current speed, and bottom depth (Goodyear, 2003). Unlike bottom depth, currents and temperature vary seasonally, and, if these factors influence distribution, may lead to seasonal changes in availability.

Several studies have highlighted the importance of temperature as a determinant of blue marlin distribution. For example, the results from temperature recordings on tags indicate that blue marlin spend the majority of their time at temperatures above 26°C (Block *et al.*, 1992). Seasonal latitudinal shifts in the distribution of blue marlin have been documented in the past (Anraku and Yabuta, 1959; Howard and Ueyanagi, 1965; Hinton and Nakano, 1996; Goodyear, 2003) and these may be partly a consequence of

\*Correspondence. e-mail: chilu@ntu.edu.tw

Received 2 December 2007

Revised version accepted 12 August 2008

seasonal changes in sea-surface temperature, SST (Goodyear, 2003).

Blue marlin are unable to sustain a swimming muscle temperature significantly above the ambient water temperature because they lack counter-current heat exchangers in their vascular systems (Brill *et al.*, 1999). Therefore, limitations on cardiac function may be responsible for the limiting effects of temperature on the distribution of blue marlin (Brill and Lutcavage, 2001). The vertical movements of marlins, including blue marlin, are also limited by temperature. Specifically, blue marlin prefer to be above the thermocline and appear to swim at the oceanic surface layer (Holland *et al.*, 1990). In addition, blue marlin apparently rarely range through the thermocline, even though they are capable of doing so (Block *et al.*, 1992), exhibiting a preference for the warm surface mixed layer above the thermocline (Hinton and Nakano, 1996; Seki *et al.*, 2002).

Hinton and Nakano (1996) noted that 75% of a small number of acoustically tagged animals were found within the mixed layer. However, blue marlin might respond in an indirect way to variation in sea level. For example, small skipjack (*Katsuwonus pelamis*) and yellowfin (*Thunnus albacares*) tunas are a particularly favored food for blue marlin (Brock, 1984) and movement and distribution of juveniles of tuna-like species may be strongly influenced by El Niño-Southern Oscillation events (Fiedler and Bernard, 1987), which could lead to changes to oceanographic features such as sea-level height and the depth of the mixed layer.

Layers of high zooplankton abundance often become feeding sites for predators when the vertical distribution of prey is limited by shoaling of the thermal structure (Seki *et al.*, 2002). Blue marlin, like other billfishes, do not form schools, but aggregate along oceanographic features, such as temperature fronts, that may be areas of increased productivity and relatively high prey abundance (Olson *et al.*, 1994). Therefore, it might be anticipated that areas of high chlorophyll concentration may also be areas of high

blue marlin abundance due to a response to prey availability.

The hydrography of a region is closely related to its bottom topography (Maravelias, 1999). For example, nutrients are transported to the surface by turbulence when the current speed increases over the shelf. Blue marlin, rarely found in coastal waters or near islands, prefer surface waters over deep and oceanic areas (Molony, 2005). Therefore, not only are blue marlin a pelagic species, but complex interactions between prey and predators may mean that bathymetry may determine the spatial distribution of this species at finer spatial scales.

The objectives of this study are to identify relationships between environmental factors and habitat preference for blue marlin, as such relationships can be elucidated using fishery catch and effort data, and then to use them to locate areas of highest relative abundance, which could provide a basis to manage the fisheries that catch this species. We explicitly consider the impact on inferences about the spatial and temporal distribution of a species such as blue marlin caused by the time scales for which the different environmental data types are available.

## MATERIALS AND METHODS

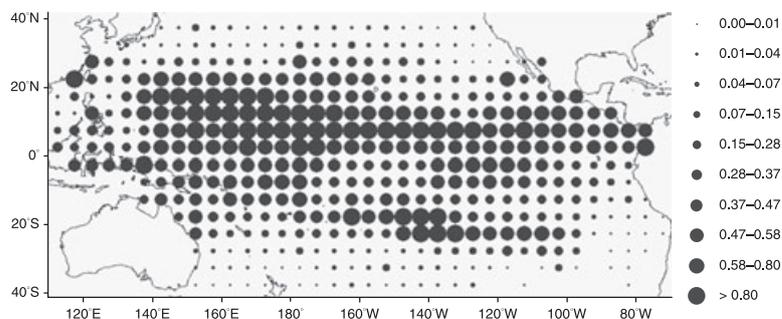
### Fishery data

Longline catch and effort data grouped by year (1950–2004), month, and 5° grid cell (Fig. 1) were obtained from the Oceanic Fisheries Programme of the Secretariat of the Pacific Community (<http://www.spc.int/oceanfish/Html/Statistics/Index.htm>). These data are aggregated over fleets (e.g. Japanese, Taiwanese and Hawaiian) due to data confidentiality issues. In addition to year and month, each catch-effort record contains information on location (latitude and longitude), fishing effort (in hooks), and catch (in number).

### Environmental data

The environmental variables considered in this study are: sea-surface temperature (SST), mixed layer depth

**Figure 1.** Nominal CPUE (number of fish caught/1000 hooks) distribution for blue marlin caught in the Pacific longline fisheries (1950–2004).



(MLD), sea-height anomaly (SHA), chlorophyll-*a* concentration (CHL), and bathymetry. Catch per unit effort (CPUE) is expressed as the number of fish caught by 5° cell divided by the corresponding total number of hooks, so the environmental variables were also averaged to 5° grid cell. The environmental data were sourced as follows: (1) monthly average SST (2° × 2°) for 1950–2004 from the NOAA/ERSST/V2 dataset provided by the NOAA/OAR/ESRL PSD (<http://www.cdc.noaa.gov/>); (2) monthly MLD (10 min resolution) for 1998–2004 (<http://web.science.oregon-state.edu/ocean.productivity/inputMldData.php>); (3) monthly SHA for 1998–2004 (<http://ferret.pmel.noaa.gov/NVODS/servlets/dataset/>); (4) monthly CHL (1° resolution) for 1998–2004 from the SeaWiFS dataset (<http://reason.gsfc.nasa.gov/OPS/>); and (5) the depth information (5 min resolution) from the Navy/DBDBV dataset (<http://www.ferret.noaa.gov/cgi-bin/dods/nph-dods/data/PMEL/>).

### Modeling

A variety of approaches can be used to model the relationship between fishery catch rates and environmental, temporal, spatial, and operational factors. The two most common approaches are generalized linear models (GLMs; McCullagh and Nelder, 1989) and generalized additive models (GAMs; Hastie and Tibshirani, 1990) (Maunder and Punt, 2004; Venables and Dichmont, 2004a). We chose to base the analyses on GAMs rather than the GLMs because GAMs can deal with non-linear relationships between covariates and response variables in a semi-parametric manner (Maunder and Punt, 2004), albeit at the expense of additional complexity when obtaining confidence intervals and applying hypothesis tests (Venables and Dichmont, 2004a). In addition, most of the previous studies which have explored relationships between spatial and environmental factors for pelagic fishes and fishery catch rates have been based on GAMs (e.g. Bigelow *et al.*, 1999; Rodríguez-Marín *et al.*, 2003; Walsh *et al.*, 2006).

The underlying assumption of a GAM model is that functions of the predictors are additive and smooth, so that the conditional mean can be expressed as:

$$g(\mu_i) = \mu + \sum_{j=1}^p f_j(X_i), \quad (1)$$

where  $g$  is the differentiable and monotonic link function,  $\mu = E(Y)$ , and  $f_j$  is a smooth function, such as a spline or a loess smoother.

A two-stage GAM approach (the delta approach; Lo *et al.*, 1992; Stefánsson, 1996; Ortiz and Arocha,

2004 – also referred to a hurdle model, Cragg, 1971) was used to deal with the fact that some 5° cells have no catch of blue marlin, but some effort. First, a GAM with a binomial response was used to analyze the data on presence or absence (1 or 0) of blue marlin (presence/absence, P/A, model) to determine the probability of catching blue marlin as a function of various covariates. A variety of models (CPUE models) based on the lognormal, gamma and Poisson error models were then used to model the catch rate given that it was positive (Ortiz and Arocha, 2004; Wise and Bromhead, 2004). The dependent variable was taken to be the number of fish caught per record, and fishing effort (i.e. number of hooks per record) was treated as an offset when using the Poisson distribution (Cooke and Lankester, 1996).

A forward stepwise technique was used to identify the appropriate set of explanatory variables for each model (P/A and CPUE). The Akaike Information Criterion (AIC), chi-square ( $\chi^2$ )  $P$ -values, and deviance were used to evaluate the significance of each predictor variable. All predictor variables were treated as continuous, and the effective degrees of freedom were estimated for each main factor. Latitude, longitude, and year were treated as interaction terms to account for the impact on catch rates of changes over time in how the gear is set owing to changes in targeting practices (Bigelow *et al.*, 2002). The full P/A and CPUE models can be written as:

$$\begin{aligned} \text{P/A} \sim & s(\text{Year}) + s(\text{Month}) + s(\text{Depth}) + s(\text{SST}) + \\ & s(\text{MLD}) + s(\text{SHA}) + s(\text{CHL}) + \\ & s(\text{Year, Latitude}) + s(\text{Year, Longitude}) + \\ & s(\text{Latitude, Longitude}); \end{aligned} \quad (2)$$

$$\begin{aligned} \text{CPUE} \sim & s(\text{Year}) + s(\text{Month}) + s(\text{Depth}) + s(\text{SST}) + \\ & s(\text{MLD}) + s(\text{SHA}) + s(\text{CHL}) + \\ & s(\text{Year, Latitude}) + s(\text{Year, Longitude}) + \\ & s(\text{Latitude, Longitude}), \end{aligned} \quad (3)$$

where P/A is zero if no blue marlin were caught (one otherwise), CPUE is the catch rate of blue marlin if some blue marlin were caught, and  $s(X)$  denotes a tensor spline smoother function of the covariate  $X$  or the interaction between two covariates. Tensor splines were selected over ordinary splines because the covariates are measured in different units (Wood, 2006).

Diagnostic plots, i.e. the distribution of residuals and quantile-quantile (Q-Q) plots, were used to assess the model fits and to select the most appropriate error distribution for the non-zero catch rates. In addition, the standardized deviance residuals were plotted

against the fitted values to check the assumed link function. The computational and presentational tasks were conducted using R version 2.4.1 (R Development Core Team, 2006).

Environmental data are not available for all of the years for which fishery data are available. Therefore, the P/A and CPUE models were fitted for two time periods (1950–2004 and 1998–2004). The models for each time period included spatial (latitude and longitude), temporal (year and month), and environmental (depth and sea-surface temperatures) effects. The model for 1998–2004 also included mixed layer depth, sea-height anomaly, and chlorophyll-*a* data as these data are available from 1998. Two models were fitted to the 1998–2004 data, one based on all of the environmental data and one based on the data used for the 1950–2004 model (the ‘reduced 1998–2004 model’). These two models allow an evaluation of the impact of the availability of different covariates and dataset lengths.

#### Predicting blue marlin densities

The selected P/A and CPUE models were used to predict the relative density of blue marlin over the Pacific Ocean, i.e. the P/A model was used to predict the probability of the presence of blue marlin in each 5° cell and the CPUE model was used to predict relative density given presence (Sacau *et al.*, 2005). The relative densities can be considered to be measures of the habitat preference for blue marlin (Maravelias, 1999).

Relative densities were predicted by month for every 5° cell in the Pacific Ocean for which environmental data were available to identify the spatial variation of blue marlin and how it changes over time. The spatial distributions of blue marlin were plotted by year and month for 1998–2004 (the years included in both models) using the model outputs from the 1950–2004 and 1998–2004 models to assess the impact of different predictor choices. The 25th, 50th, and 75th percentiles of the monthly predicted values of relative blue marlin density over the area for which data were available, were used to compare the results from the three models in a relatively straightforward manner.

The set of cells that may be ‘hotspots’ for blue marlin abundance were defined following Goodyear (2003) as those for which relative density is in the top 5%. The latitude and longitude for the set of cells that were in the ‘hotspot’ areas were averaged for each month during 1998–2004 for the models based on the 1950–2004 and 1998–2004 data to examine spatial variation and seasonal movement of ‘hotspots’. Frequency distributions of relative density by covariate level for the cells in the ‘hotspot’ areas were then used to explore which environmental conditions correspond to the (apparently) preferred habitats for blue marlin.

## RESULTS

The total number of observations used in the analyses is 140 019, of which 110 461 indicate a catch of at least one blue marlin. The highest catch rates of blue marlin occur in the tropics (Fig. 1).

#### Selection of an error model for the non-zero catch rates

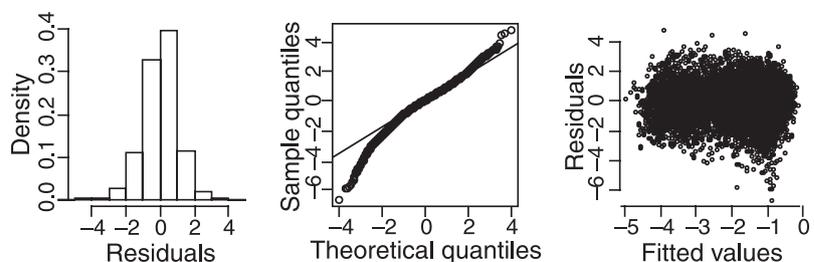
The residuals from the lognormal distribution conform to the greatest extent with its assumptions according to Q-Q plots, and because the residuals (in log-space) for the lognormal error distribution appear normal and there is no evidence for heteroscedasticity from the plot of the residuals against the fitted values (Fig. 2). The remaining analyses of the positive catch rates are consequently based on the lognormal distribution.

#### Selection of factors in models

Deviance tables (Table 1) are used to summarize the model selection process for the P/A and CPUE models for the two time periods (1950–2004 and 1998–2004). Table 1 lists the total deviance explained, the change in deviance due to each additional factor, the percentage of the total deviance explained by each additional factor, and the *P*-value derived from a chi-square test between models that differ in terms of one additional factor for each model.

All of the main effects are statistically significant at  $\alpha = 0.01$  in the P/A and CPUE models (Table 1). SST

**Figure 2.** Diagnostic plots for the lognormal error distribution when fitting the non-zero catch rates.



**Table 1.** Analysis of deviance tables for the 1950–2004 and 1998–2004 models.

	1950–2004			1998–2004		
	Residual deviance	% of total deviance explained	$P(\chi^2)$	Residual deviance	% of total deviance explained	$P(\chi^2)$
<i>P/A model</i>						
<NULL>	144 334			24 447		
+s(Year)	143 509	2.3	<0.01	24 429	0.2	<0.01
+s(Month)	143 484	0.1	<0.01	24 383	0.5	<0.01
+s(Depth)	141 367	5.8	<0.01	23 987	4.7	<0.01
+s(SST)	118 462	62.8	<0.01	18 641	63.1	<0.01
+s(MLD)	–	–	–	18 239	4.7	<0.01
+s(SHA)	–	–	–	18 213	0.3	<0.01
+s(CHL)	–	–	–	17 956	3.0	<0.01
+s(Year, Latitude)	113 813	12.7	<0.01	17 191	9.0	<0.01
+s(Year, Longitude)	109 519	11.8	<0.01	16 575	7.3	<0.01
+s(Latitude, Longitude)	107 844	4.6	<0.01	15 969	7.1	<0.01
Total deviance explained	25.3%			34.7%		
<i>CPUE model</i>						
<NULL>	233 136			27 455		
+s(Year)	212 975	16.4	<0.01	27 227	1.8	<0.01
+s(Month)	211 565	1.1	<0.01	27 169	0.5	<0.01
+s(Depth)	206 332	4.3	<0.01	26 696	3.8	<0.01
+s(SST)	128 124	63.6	<0.01	17 872	70.3	<0.01
+s(MLD)	–	–	–	17 645	1.8	<0.01
+s(SHA)	–	–	–	17 551	0.7	<0.01
+s(CHL)	–	–	–	17 413	1.1	<0.01
+s(Year, Latitude)	123 112	4.1	<0.01	16 800	4.9	<0.01
+s(Year, Longitude)	116 370	5.5	<0.01	16 131	5.3	<0.01
+s(Latitude, Longitude)	110 256	5.0	<0.01	14 897	9.8	<0.01
Total deviance explained	52.7%			45.7%		

$P$ , level of significance.

accounted for more than 60% of the explained deviance in the four models and leaving SST out of the model led to the greatest change in deviance and AIC (Table 2). Other factors that explained a substantial amount of the deviance (>5% in at least one model) were the three two-way interaction terms, depth (*P/A* models only), and year (*CPUE* model applied to data for 1950–2004). The total deviances explained by the models for the non-zero catches were 52.7 and 45.7%, whereas the deviances explained by the *P/A* models were 25.3 and 34.7% for analyses based on the 1950–2004 and the 1998–2004 datasets, respectively.

#### *Spatial distribution*

Figure 3 shows the nominal catch rates and the predicted relative density of blue marlin by 5° cell for 1998 and 2004 (years which all models cover), obtained by averaging the relative densities over months. The relative densities for January and July averaged over the years 1998–2004 from the 1950–2004 and 1998–2004

models and the nominal catch rates are shown in Fig. 4 (January and July, because these months represent the extremes of the spatial distributions). For ease of presentation, relative density in Figs 3 and 4 is categorized into five equal-sized groups.

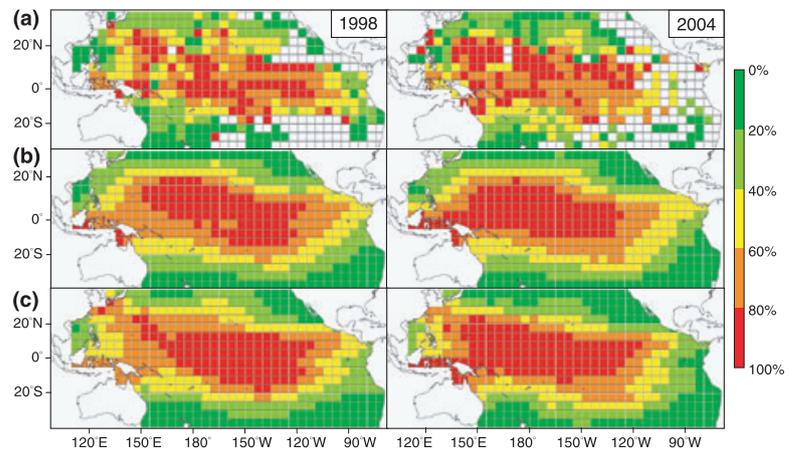
Both the probability of catching a blue marlin and the catch rate given that the catch is non-zero vary spatially and temporally. There is a broad range of latitudes and longitudes at which both the probability of catching a blue marlin and the catch rate given the catch is non-zero are high, with a notable increase in both of these quantities in the equatorial Pacific (results not shown). Qualitatively, the spatial patterns of relative density predicted by the two models are similar, i.e. both show seasonal variation in the areas of highest relative density of blue marlin and that the highest relative densities occur in the tropics between 20°N and 20°S, especially in the central Pacific Ocean. The spatial patterns of relative density for the 1998–2004 and reduced 1998–2004 models are essentially identical

**Table 2.** The residual deviance and the percentage relative changes in the explained deviance and AIC due to omitting each of the factors (i.e. omitting both the main effect and its interaction) in turn. Results are shown for two time periods.

	1950–2004			1998–2004		
	Residual deviance	Dev (%)	AIC (%)	Residual deviance	Dev (%)	AIC (%)
<i>P/A</i> model						
<NONE>	107 844			15 969		
–s(Year)	121 016	12.2	12.1	18 451	15.5	14.4
–s(Month)	120 096	11.4	11.2	18 467	15.6	14.5
–s(Depth)	120 749	12.0	11.8	18 584	16.4	15.2
–s(SST)	143 414	33.0	32.8	23 797	49.0	47.6
–s(MLD)	–	–	–	18 597	16.5	15.3
–s(SHA)	–	–	–	18 458	15.6	14.5
–s(CHL)	–	–	–	18 454	15.6	14.4
–s(Latitude)	120 119	11.4	11.2	18 453	15.6	14.4
–s(Longitude)	120 646	11.9	11.7	18 653	16.8	15.7
<i>CPUE</i> model						
<NONE>	110 256			14 897		
–s(Year)	164 188	48.9	14.0	18 149	21.8	6.7
–s(Month)	145 132	31.6	9.6	18 017	20.9	6.5
–s(Depth)	147 928	34.2	10.3	18 187	22.1	6.8
–s(SST)	216 170	96.1	23.7	26 668	79.0	20.5
–s(MLD)	–	–	–	18 025	21.0	6.5
–s(SHA)	–	–	–	18 012	20.9	6.5
–s(CHL)	–	–	–	18 030	21.0	6.5
–s(Latitude)	145 098	31.6	9.6	18 147	21.8	6.7
–s(Longitude)	149 371	35.5	10.7	18 565	24.6	7.5

Dev, the deviance.

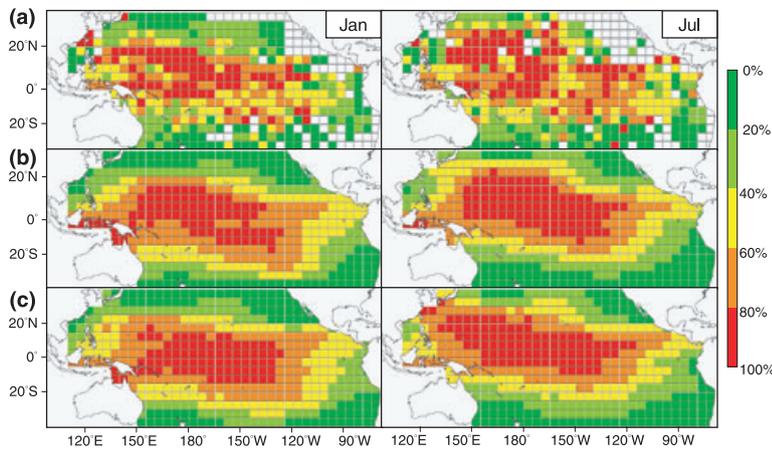
**Figure 3.** Relative density of blue marlin for 1998 and 2004 (averaged over months) based on (a) the nominal data, (b) the 1950–2004 model, and (c) the 1998–2004 model.



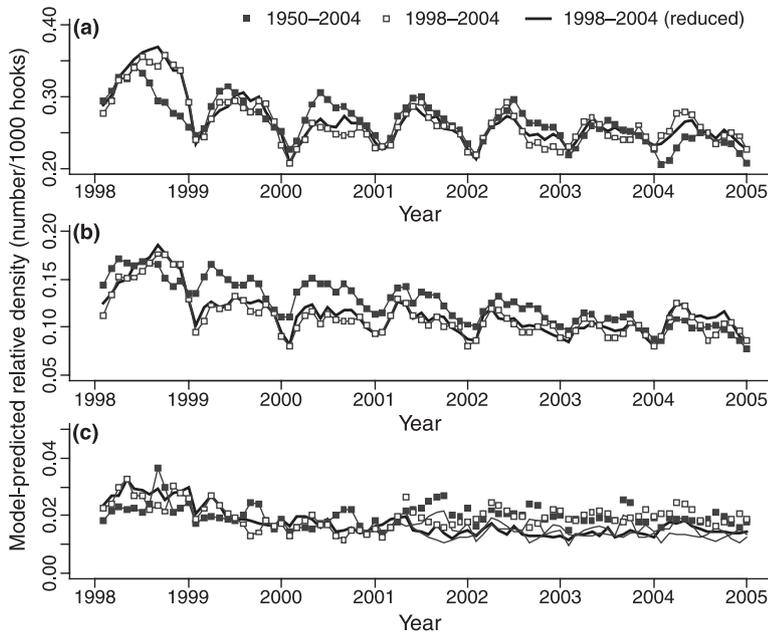
both qualitatively and quantitatively (results not shown). The model predictions in Figs 3 and 4 are smoother than the raw data. The nominal catch rates are high in some  $5^\circ \times 5^\circ$  cells which the model indicates should have low relative density on average (e.g. south of  $20^\circ\text{S}$ ). Such catch rates are caused by high catch with low effort and probably reflect sampling error rather than real effects.

*Differences among models*

Figure 5 plots the quartiles of model-predicted relative density across the cells with data for each month during 1998–2004 based on the 1950–2004 and 1998–2004 models as well as on the reduced 1998–2004 model. The time-series of quartiles from the 1950–2004 and 1998–2004 models exhibit similar patterns



**Figure 4.** Relative density of blue marlin for January and July (averaged over 1998–2004) based on (a) the nominal data, (b) the 1950–2004 model, and (c) the 1998–2004 model.



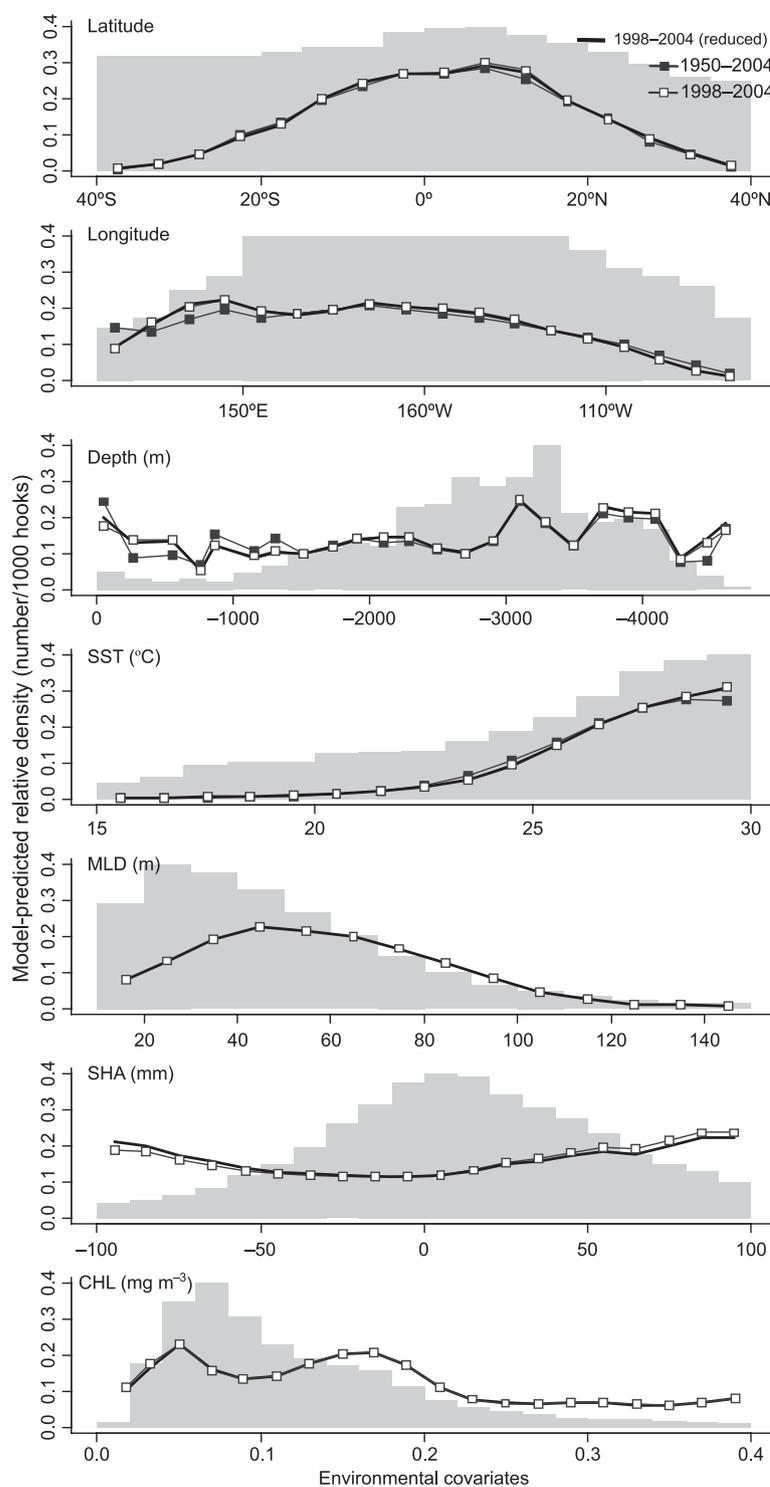
**Figure 5.** The (a) third, (b) second, and (c) first quantiles of monthly model-predicted relative density (1998–2004) over the spatial cells considered in the analyses from the three models.

of seasonality with little between-model differences, whereas the original and reduced 1998–2004 models lead to almost identical outcomes. However, the median levels of relative density differ between the model based on data from 1950–2004 and those based on data for 1998–2004.

The predicted relative densities of blue marlin during 1998–2004 over the cells with data between 40°N and 40°S are plotted against each level of the environmental and spatial covariates in Fig. 6 for the three models. As expected from previous studies, the highest relative densities are predicted to occur in the areas in which SST is higher than 26°C and between 10°S and 15°N. The impact of the other factors is not as clear as for SST, but the impact of

MLD on relative density is also fairly apparent. Most of the fishing effort is directed at areas for which the relative density of blue marlin (and the tropical tunas, which are the targets of the fisheries concerned) is high. The most notable exceptions to this are latitude and, to a lesser extent, sea-surface temperature.

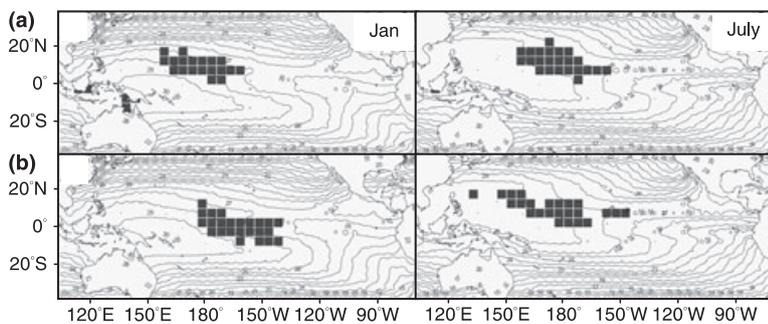
There are differences in the spatial distribution of ‘hotspots’ of blue marlin among months (averaging over 1998–2004) and there are again between-model differences (Fig. 7). Specifically, the areas of highest relative density predicted by the 1998–2004 model (and the reduced 1998–2004 model, results not shown) are somewhat broader in latitude and longitude and occur further towards the east (Fig. 7b). All



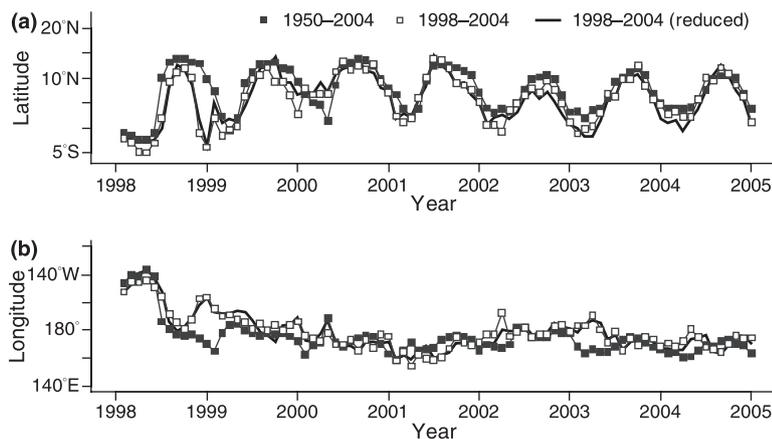
**Figure 6.** Average model-predicted relative density (number/1000 hooks) for each level of each covariate from the three models. The shaded histograms indicate the relative number of data points for each level of each covariate.

three models lead to clear indications of seasonal latitudinal movement of ‘hotspots’ (Fig. 8a). In particular, there is a southward shift in the areas of highest density during the first half of the year. The latitudes in which ‘hotspots’ occur seem consistent over time

and between models (Fig. 8a) and there is a trend over time towards ‘hotspots’ in the west from all three models (Fig. 8b). The two models based on the data for 1998–2004 provide nearly identical patterns in Fig. 8, whereas the 1950–2004 model leads to mark-



**Figure 7.** 'Hotspot' areas for blue marlin (the cells corresponding to the top 5% of relative density) based on (a) the 1950–2004 and (b) the 1998–2004 models, along with the spatial distribution of sea-surface temperature for January and July (averaged over 1998–2004).



**Figure 8.** The average (a) latitude and (b) longitude for the spatial cells in the 'hotspot' areas for each month during 1998–2004 based on the three models.

edly different estimates of the mean longitude of 'hotspots' before 2001.

Figure 9 explores the 'hotspot' areas further using frequency distributions by covariate level for the cells that correspond to the top 5% of the relative densities during 1998–2004 for the 1950–2004 and 1998–2004 models. The covariate levels leading to 'hotspots' are generally consistent between the 1950–2004 and 1998–2004 models, except for longitude and depth, with the 1950–2004 model indicating a somewhat higher frequency of 'hotspots' west of 150°W and in shallower water than the 1998–2004 model. As expected, 'hotspots' are predicted to occur when sea-surface temperature is above 27°C (Fig. 9).

## DISCUSSION

### Modeling approach

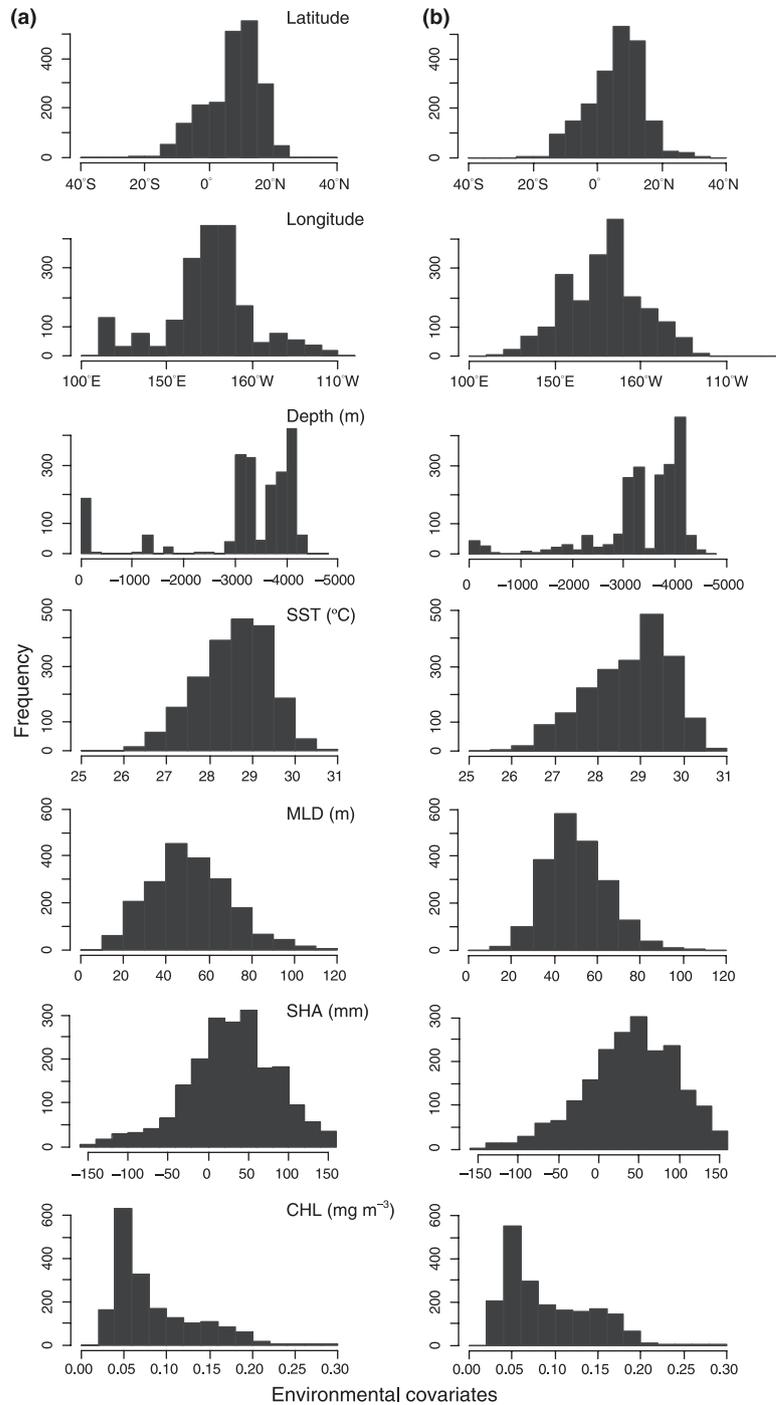
Catch-effort datasets, particularly those for relatively rare bycatch species, often contain many observations with zero catch. Zero catches can lead to computational problems if log-linear models are used to standardize the data because the logarithm of zero is not defined. Ignoring the zero records or replacing them with a small positive constant are two common ways to address this

problem (Hinton and Maunder, 2003). However, these methods may not deal adequately with the zero catches because the observations without catch may suggest environmental conditions or locations that are unsuitable for the species under investigation.

Considering the issue of presence separately from that of density given presence leads to a better understanding of the relationships between a fish species and the marine environment and ecosystem that maintains it (Maravelias, 1999). An advantage of the 'delta approach' used in this study is that it allows for explicit consideration of two key aspects of the ecology of blue marlin, i.e. distribution (presence/absence) and relative abundance (density).

Ortiz and Arocha (2004) examined alternative error distributions for the standardization of catch rates of non-target species, i.e. billfishes, in a pelagic tuna longline fishery. Their results indicated, consistent with those of this study, that a binomial error distribution for the probability of catching fish and a lognormal error model for the positive catches were most consistent with the available data.

Most of the relationships between the environmental variables and the catch rates/presence-absence for blue marlin were not linear, with the effective



**Figure 9.** Frequency distributions by covariate level for the spatial cells that are in the ‘hotspot’ areas (1998–2004) based on (a) the 1950–2004 and (b) the 1998–2004 models.

degrees of freedom being greater than one (results not shown). It was possible to capture these relationships adequately through the use of GAM. Although GAMs were used in this paper to capture spatial and environmental factors determining catch rates, similar results would have been obtained had other approaches such as GLMs and regression trees been applied. GLMs could have been used to model the

impact of spatial and environmental covariates, e.g. by treating longitude and latitude as cubic splines when fitting the GLM, but the choice of where to place the knots would have had to have been made through a trial and error process (see, for example, Venables and Ripley, 2002). Regression trees can deal with non-linear relationships among variables, and could be a viable alternative approach for identifying ‘hotspots’.

This approach would, however, not be ideal for creating maps of relative density, such as those in Figs 3 and 4.

A key reason for using models to analyze the data is that the nominal catch rates reflect where the fishery was operating and hence the average nominal catch rate in a given month in a grid cell implicitly weights levels of environmental variables by the amount of effort in that grid cell for each level. In contrast, GAMs (and other modeling techniques) remove the impact of how the data were collected (the non-randomness of the sampling process). For example, fishers appear to select areas based on SST. However, sufficient observations are available at a range of SST levels (see Fig. 6) to enable the impact of SST on catch rate to be distinguished. Naturally, no modeling framework should be used to infer outcomes from outside of the range covered by the data without appropriate care (in this study, for example, the model could be used to infer relative density south of Indonesia, but no such densities are reported as they would simply be extrapolation).

#### *Impact of environmental effects*

The results of this study identify relationships between catch rates of blue marlin and spatial, temporal, and environmental variables, providing insight into the underlying mechanisms determining the habitat preference and distribution of blue marlin in the Pacific Ocean. Although all the covariates have statistically significant impacts on both the presence of blue marlin, and the catch rate of this species given presence, SST is the most important (Table 1). The results indicate that blue marlin can be found over a large proportion of the Pacific Ocean, but prefer warm tropical areas, spend most time in the surface layer, and appear in deeper waters away from the continental shelf (Figs 6 and 9).

Latitude and longitude explain a large amount of the variance in the P/A and CPUE models (Table 1). However, of the various covariates considered in the GAM modeling, SST led to the largest reduction in the explained deviance and AIC when it was left out of the model (Table 2). The model results are therefore consistent with those of several previous studies (e.g. Brill and Lutcavage, 2001; Goodyear, 2003) and, specifically, those of tagging experiments which suggest that blue marlin are found in temperatures from 17 to 31°C, with most studies suggesting a preferred temperature of 24°C and higher (Holland *et al.*, 1990; Block *et al.*, 1992; Graves *et al.*, 2002).

It is clear that oceanic features, perhaps related to physiological limitations directly, or prey abundance

indirectly, play important roles in determining the preferred habitats of blue marlin (Seki *et al.*, 2002). The results in Figs 6 and 9 suggest that the highest catch rates of blue marlin are found when the thermocline is 30–60 m, which is consistent with the results of tagging experiments that suggest that blue marlin spend the majority of their time within 40 m of the surface with vertical movements confined to the surface mixed layer (Holland *et al.*, 1990; Block *et al.*, 1992; Graves *et al.*, 2002).

Brill and Lutcavage (2001) contend that chlorophyll-*a* concentration is an indirect surrogate measure of forage abundance for large pelagic fishes. In contrast, Fig. 9 shows that the highest relative densities of blue marlin occur in the cells where chlorophyll-*a* is relatively low (<0.2 mg m<sup>-3</sup>). The pelagic, tropical oceans are an oligotrophic ecosystem with high trophic efficiency for fish at the top of the trophic pyramid. Therefore, surface chlorophyll concentration does not completely explain the variance in primary production (Campbell *et al.*, 2002). Furthermore, the water column may be clear enough for blue marlin to be effective visual predators (Brill and Lutcavage, 2001). However, the power to identify an effect of chlorophyll-*a* on relative abundance may be low due to the resolution of the available catch and effort data (5° × 5°). For example, it is conceivable that blue marlin react to factors such as chlorophyll-*a* concentration and hence forage abundance at much finer spatial scales than could be examined in this study.

As a surface-dwelling species, blue marlin likely extend their vertical range with the deepening of the thermocline during El Niño events and may become more catchable to longline sets at this time (Ward and Myers, 2005). However, the results in Figs 6 and 9 do not confirm an increase in catch rates of blue marlin with deepening mixed layer depth and sea-level anomalies.

#### *Spatial distribution*

Blue marlin have the most tropical distribution of all billfishes (Molony, 2005) and are the most common billfish species in the catches of the longline fisheries in the tropical Pacific Ocean which target bigeye (*Thunnus obesus*) and yellowfin tuna (Hinton, 2001). The catch-effort dataset on which this study was based included data for several of the longline fisheries in the Pacific Ocean and hence covered the spatial distribution of the population (which any one dataset would not) and also pertained to a long time period (since the 1950s). This large dataset therefore provides an almost unique way to explore the distribution of a cosmopolitan species such as blue marlin.

The blue marlin is a highly migratory species and sometimes migrates to 40°N, but not regularly. In Hawaiian waters, Walsh *et al.* (2005) noted that large numbers of blue marlin occur in the autumn and early winter months with high landings reported. Although Hopper (1990) indicated that the summer influx of blue marlin in Hawaiian waters appears to be a spawning migration, blue marlin have been found highly fecund and spawning year-round in tropical waters (Molony, 2005). Therefore, seasonal shifts in the relative density of blue marlin may not relate to the spawning migration and hence reproductive strategy.

It is hard to identify any fish that is randomly distributed spatially (Brill and Lutcavage, 2001). In this study, latitude and longitude were significant in both the P/A and CPUE models (Table 1). It is evident that catch rates of blue marlin vary spatially and temporally (Figs 3 and 4). In particular, although the bulk of the population is found in warm tropical waters, there are also indications of changes in the preferred latitude seasonally (Fig. 8), but not in the fishing effort (results not shown).

#### *Sensitivity to model specifications*

The results from the model are not very sensitive to whether sea height, mixed layer depth, and chlorophyll-*a* are included in the analyses. However, the results, even for 1998–2004, are sensitive to using data for 1950–1997. Large changes in the fishing practices of the pelagic longline fisheries in the Pacific Ocean occurred during these years (Bigelow *et al.*, 2002). Specifically, there was a movement to gear designed to target the deeper-swimming bigeye tuna. This change makes it difficult to argue that catchability for surface-dwelling species such as blue marlin has remained constant over time and hence that relationships between covariates such as SST, latitude and longitude, and catch rates have been stationary. This likely lack of stationarity means that the results of the 1950–2004 model cannot be given as much weight as those from the 1998–2004 model even though the results from the various models are generally in agreement. In principle, more detailed information on gear (e.g. the number of hooks between adjacent mainline floats), could help evaluate changes in catchability (see, for example, Hinton and Nakano, 1996), but these data were not available owing to confidentiality issues. The lack of consistency in targeting practices over time is one reason why this paper has not focused on the predicted trends over time in relative abundance.

The results in Fig. 8 are suggestive of seasonal movement, but could equally well reflect changes in

availability. This source of uncertainty could be resolved through additional tagging studies.

#### *Implications for the conservation of blue marlin*

Effective management and conservation in the open oceans depends on determining the spatial distribution of species, fisheries, and the ecological system. For species such as billfish that have pelagic larvae and highly migratory adults, time-area closures may be the best way to minimize bycatches yet keep target catches high (Hobday and Hartmann, 2006). For example, Goodyear (1999) suggested that time-area restrictions on fishing could significantly reduce the bycatch of billfishes in pelagic longline fisheries without equivalent reductions to the catch of target species.

In principle, time-area restrictions on fishing should be based on prey distribution, environmental conditions, and habitat characteristics. The relationships among spatial, temporal, and environmental variables and blue marlin abundance derived in this paper could be used to predict the areas of highest density for this species (see Figs 3, 4 and 7) and hence form the basis for time-area management if there was a wish to reduce catches. Care needs to be taken, however, to ensure that 'hotspot' areas are relatively consistent over years. However, plots such as Fig. 7 indicate that whereas there is seasonal variability in the relative density of blue marlin, some areas have relatively consistent high catch rates of blue marlin among years. It is clear that there are differences between where 'hotspot' areas for blue marlin are predicted to be depending on the number of years (and covariates) included in the analyses. The results of the analyses of this paper suggest that relatively robust predictions of 'hotspot' areas can be made if only recent data are used, and that leaving some environmental covariates out of the analyses (there are undoubtedly other covariates which impact the spatial distribution of blue marlin) does not impact the qualitative predictions of where 'hotspots' occur. Finally, although the development of time-area restrictions should logically start from plots such as Fig. 7, the regional fisheries management organizations in the Pacific Ocean (IATTC and WCPFC) could refine the 'hotspots' using the confidential finer scale data.

#### ACKNOWLEDGEMENTS

We sincerely express our appreciation to the editor and the two anonymous reviewers for their valuable comments on an earlier version of this manuscript. This study was in part supported financially by the National Science Council and the

Fisheries Agency of Council of Agriculture, Taiwan, through the grants NSC93-2313-B-002-053, NSC94-2313-B-002-067, and 98-AS-15.1.2-FA-F2(2) to C.-L. Sun.

## REFERENCES

- Anraku, N. and Yabuta, Y. (1959) Seasonal migration of blue marlin. *Nankai Reg. Fish. Res. Lab. Rep.* **10**:63–71.
- Bigelow, K.A., Boggs, C.H. and He, X. (1999) Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. *Fish. Oceanogr.* **8**:178–198.
- Bigelow, K.A., Hampton, J. and Miyabe, N. (2002) Application of a habitat-based model to estimate effective longline fishing effort and relative abundance of Pacific bigeye tuna (*Thunnus obesus*). *Fish. Oceanogr.* **11**:143–155.
- Block, B.A., Booth, D.T. and Carey, F.G. (1992) Depth and temperature of the blue marlin, *Makaira nigricans*, observed by acoustic telemetry. *Mar. Biol.* **114**:175–183.
- Brill, R.W. and Lutcavage, M.E. (2001) Understanding environmental influences on movements and depth distributions of tunas and billfishes can significantly improve population assessments. In: *Islands in the Stream: Oceanography and Fisheries of the Charleston Bump*. G. Sedberry (ed.) Bethesda, MD: American Fisheries Society Symposium 25, pp. 179–198.
- Brill, R.W., Block, B.A., Boggs, C.H., Bigelow, K.A., Freund, E.V. and Marcinek, D.J. (1999) Horizontal movements and depth distribution of large adult yellowfin tuna (*Thunnus albacares*) near the Hawaiian Islands, recorded using ultrasonic telemetry: implications for the physiological ecology of pelagic fishes. *Mar. Biol.* **133**:395–408.
- Brock, R.E. (1984) A contribution to the trophic biology of the blue marlin (*Makaira nigricans* Lacépède, 1802) in Hawaii. *Pacific Sci.* **38**:141–149.
- Campbell, J., Antoine, D., Armstrong, R. et al. (2002) Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance. *Global Biogeochem. Cycles* **16**:1035.
- Cooke, J.G. and Lankester, K. (1996) Consideration of statistical models for catch-effort indices for use in tuning VPAs. *ICCAT Col. Vol. Sci. Pap.* **45**:125–131.
- Cragg, J.G. (1971) Some statistical models for limited dependent variables with applications to the demand for durable goods. *Econometrica* **39**:829–844.
- Fiedler, P.C. and Bernard, H.J. (1987) Tuna aggregation and feeding near fronts observed in satellite imagery. *Continental Shelf. Res.* **7**:871–881.
- Goodyear, C.P. (1999) An analysis of the possible utility of time-area closures to minimize billfish bycatch by U.S. pelagic longlines. *Fish. Bull. US* **97**:243–255.
- Goodyear, C.P. (2003) Spatio-temporal distribution of longline catch per unit effort, sea surface temperature and Atlantic marlin. *Mar. Freshw. Res.* **54**:409–417.
- Graves, J.E. and McDowell, J.R. (2003) Stock structure of the world's istiophorid billfishes: a genetic perspective. *Mar. Freshw. Res.* **54**:287–298.
- Graves, J.E., Luckhurst, B.E. and Prince, E.D. (2002) An evaluation of pop-up satellite tags for estimating postrelease survival of blue marlin (*Makaira nigricans*) from a recreational fishery. *Fish. Bull. US.* **100**:134–142.
- Hastie, T.J. and Tibshirani, R.J. (1990) *Generalized Additive Models*. London: Chapman and Hall.
- Hinton, M.G. (2001) Status of blue marlin in the Pacific Ocean. In: *Status of Tuna and Billfish Stocks in 1999*. Inter-Amer. Trop. Tuna Comm., La Jolla, CA: Stock Assessment Report, 1, pp. 284–319.
- Hinton, M.G. and Maunder, M.N. (2003) Methods for standardizing CPUE and how to select among them. *ICCAT Col. Vol. Sci. Pap.* **56**:169–177.
- Hinton, M.G. and Nakano, H. (1996) Standardizing catch and effort statistics using physiological, ecological, or behavioral constraints and environmental data, with an application to blue marlin (*Makaira nigricans*) catch and effort data from the Japanese longline fisheries in the Pacific. *Bull. Int. Am. Trop. Tuna Commun.* **21**:171–200.
- Hobday, A.J. and Hartmann, K. (2006) Near real-time spatial management based on habitat predictions for a longline bycatch species. *Fish. Management Ecol.* **13**:365–380.
- Holland, K.N., Brill, R.W. and Chang, R.K.C. (1990) Horizontal and vertical movements of Pacific blue marlin captured and released using sportfishing gear. *Fish. Bull. US.* **88**:397–402.
- Hopper, C.N. (1990) Patterns of Pacific blue marlin reproduction in Hawaiian waters. In: *Planning the Future of Billfishes*. R. H. Stroud (ed.) Savannah, GA: National Coalition for Marine Conservation, pp. 123–136.
- Howard, J.K. and Ueyanagi, S. (1965) Distribution and relative abundance of billfishes (Istiophoridae) of the Pacific Ocean. *Univ. Miami Inst. Mar. Sci. Stud. Trop. Oceanogr.* **2**:1–134.
- Kleiber, P., Hinton, M.G. and Uozumi, Y. (2003) Stock assessment of blue marlin (*Makaira nigricans*) in the Pacific using MUNTIFAN-CL. *Mar. Freshw. Res.* **54**:349–360.
- Lo, N.C., Jacobson, L.D. and Squire, J.L. (1992) Indices of relative abundance from fish spotter data based on delta-lognormal models. *Can. J. Fish. Aquat. Sci.* **49**:2515–2526.
- Maravelias, C.D. (1999) Habitat selection and clustering of a pelagic fish: effects of topography and bathymetry on species dynamics. *Can. J. Fish. Aquat. Sci.* **56**:437–450.
- Maunder, M.N. and Punt, A.E. (2004) Standardizing catch and effort data: a review of recent approaches. *Fish. Res.* **70**:141–159.
- McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*, 2nd edn. London: Chapman and Hall.
- Molony, B. (2005) *Summary of the Biology, Ecology and Stock Status of Billfishes in the WCPFC, with a Review of Major Variables Influencing Longline Fishery Performance*. Working paper EB-WP-2, 1st Meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission (WCPFC-SC1), New Caledonia: Noumea, 8–19 August 2005.
- Olson, D.B., Hitchcock, G.L., Mariano, A.J. et al. (1994) Life on the edge: marine life and fronts. *Oceanography* **7**:52–60.
- Ortiz, M. and Arocha, F. (2004) Alternative error distribution models for standardization of catch rates of non-target species from a pelagic longline fishery: billfish species in the Venezuelan tuna longline fishery. *Fish. Res.* **70**:275–297.
- R Development Core Team (2006) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, ISBN 3-900051-07-0, <http://www.R-project.org> [accessed 18 December 2006].
- Rodriguez-Marín, E., Arrizabalaga, H., Ortiz, M., Rodriguez-Cabello, C., Moreno, G. and Kell, L.T. (2003) Standardization of bluefin tuna, *Thunnus thynnus*, catch per unit effort in the baitboat fishery of the Bay of Biscay (Eastern Atlantic). *ICES J. Mar. Sci.* **60**:1216–1231.

- Sacau, M., Pierce, G.J., Wang, J. *et al.* (2005) The spatio-temporal pattern of Argentine shortfin squid *Illex argentinus* abundance in the southwest Atlantic. *Aquat. Living Resour.* **18**:361–372.
- Seki, M.P., Lumpkin, R. and Pierre, F. (2002) Hawaii cyclonic eddies and blue marlin catches: the case study of the 1995 Hawaiian International Billfish Tournament. *J. Oceanogr.* **58**:739–745.
- Shaklee, J.B., Richard, W.B. and Acerra, R. (1983) Biochemical genetics of Pacific blue marlin, *Makaira nigricans*, from Hawaiian waters. *Fish. Bull. US.* **81**:85–90.
- Stéfansson, G. (1996) Analysis of groundfish survey abundance data: combining the GLM and delta approaches. *ICES J. Mar. Sci.* **53**:577–588.
- Venables, W.N. and Dichmont, C.M. (2004a) GLMs, GAMs and GLMMs: an overview of theory and applications in fisheries research. *Fish. Res.* **70**:319–337.
- Venables, W.N. and Dichmont, C.M. (2004b) A generalized linear model for catch allocation: an example from Australia's Northern Prawn Fishery. *Fish. Res.* **70**:409–426.
- Walsh, W.A., Ito, R.Y., Kawamoto, K.E. and McCracken, M. (2005) Analysis of logbook accuracy for blue marlin (*Makaira nigricans*) in the Hawaii-based longline fishery with a generalized additive model and commercial sales data. *Fish. Res.* **75**:175–192.
- Walsh, W.A., Howell, E.A., Bigelow, K.A. and McCracken, M.L. (2006) Analyses of observed longline catches of blue marlin, *Makaira nigricans*, using generalized additive models with operational and environmental predictors. *Bull. Mar. Sci.* **79**:607–622.
- Ward, P. and Myers, R.A. (2005) Inferring the depth distribution of catchability for pelagic fishes and correcting for variations in the depth of longline fishing gear. *Can. J. Fish. Aquat. Sci.* **62**:1130–1142.
- Wise, B. and Bromhead, D. (2004) Striped marlin abundance: standardisation of CPUE. In: *Striped Marlin: Biology and Fishery*. D. Bromhead, J. Pepperell, B. Wise & J. Findlay (eds) Canberra: Bureau of Rural Sciences, pp. 171–194.
- Wood, S.N. (2006) Low rank scale invariant tensor product smooths for generalized additive mixed models. *Biometrics* **62**:1025–1036.