



Performance of methods used to estimate indices of abundance for highly migratory species

Patrick D. Lynch^{a,*}, Kyle W. Shertzer^b, Robert J. Latour^a

^a Virginia Institute of Marine Science, College of William & Mary, P.O. Box 1346, Gloucester Point, VA 23062, United States

^b National Oceanic and Atmospheric Administration, Southeast Fisheries Science Center, Center for Coastal Fisheries and Habitat Research, 101 Pivers Island Road, Beaufort, NC 28516, United States

ARTICLE INFO

Article history:

Received 17 October 2011

Received in revised form 3 February 2012

Accepted 4 February 2012

Keywords:

CPUE standardization

Habitat

Tuna

Billfish

Population dynamics

ABSTRACT

Estimating indices of abundance from fishery-dependent data requires that catch-per-unit-effort (CPUE) be standardized to account for factors that may have affected CPUE but are not related to changes in abundance. Such standardization is particularly important for highly migratory species (e.g., tunas, pelagic sharks, and billfishes), because of time-varying mismatches between distributions of abundance and the distribution of fishing effort. Two commonly applied methods for standardizing CPUE are generalized linear models (GLMs), which can account for changes in fishing practices in a straightforward linear fashion, and habitat-based standardizations (e.g., statHBS), which use nonlinear analysis to relate the distribution of fishing effort to the species distribution. We evaluated the accuracy of these methods over three patterns in vertical catchability as related to ocean temperature profiles, and 50 possible biomass trajectories using a simulation framework that followed the general effort dynamics of the Japanese longline fishery in the Atlantic Ocean from 1956 to 2009. Additionally, we propose a method for directly incorporating vertical habitat information into the linear models. Overall, we found the most accurate approach to be a delta-lognormal GLM with our unique habitat factor. The statHBS approach was the most accurate when catchability was simulated to peak in surface waters. However, statHBS was much more sensitive to errors in estimates of longline hook depths (i.e., habitats exploited). Based on these results, we recommend that relative abundance be estimated for highly migratory species following a delta-GLM approach that considers vertical habitats fished.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Management decisions in fisheries are often guided, at least in part, by the results of stock assessments. Therefore, in the interest of scientific integrity and sound management, it is important that scientists provide managers with accurate characterizations of stock dynamics and stock status. Assessment models often use fishery catch data and other inputs to estimate biomass trajectories and stock parameters by fitting predicted biomass to externally derived indices of relative abundance (Maunder and Starr, 2003). Such indices of abundance serve as ‘observed’ abundance trends and thus have an influence over the assessment results.

When derived from fishery-independent surveys that monitor the stock(s) being evaluated, indices of abundance may reliably depict abundance trends. However, many exploited species are either not monitored or not monitored comprehensively. For these

organisms, information on distribution and abundance is primarily obtained through catch and effort data from fisheries that either target or incidentally catch these species. To estimate relative abundance from these data, it is common practice to adjust the catch by the corresponding amount of effort and assume a proportional relationship between catch-per-unit-effort (CPUE) and abundance. However, because fisheries are not designed to collect random unbiased samples of the harvested populations, fishery-dependent CPUE must be standardized to account for factors (e.g., changes in fishing practices) that may cause the proportionality constant (catchability) to be time-varying (Wilberg et al., 2010; Ye and Dennis, 2009), thereby violating the assumed relationship between CPUE and abundance. Numerous approaches are available for standardizing CPUE to estimate relative abundance (Maunder and Punt, 2004); thus, to promote confidence in stock assessment results it is important to evaluate proposed methods under various assumptions and real-world conditions.

Highly migratory species (HMS) represent valuable global resources, but the high cost associated with a large scale pelagic survey has prevented the development of comprehensive monitoring programs for HMS (Bishop, 2006). Therefore, estimated trends

* Corresponding author. Tel.: +1 804 684 7885; fax: +1 804 684 7327.

E-mail addresses: pdlynch@vims.edu (P.D. Lynch), Kyle.Shertzer@noaa.gov (K.W. Shertzer), latour@vims.edu (R.J. Latour).

in abundance for HMS are typically derived from fishery-dependent data. Given the extensive spatial and temporal coverage of many HMS fisheries, such as the Japanese longline fishery (JLL), these time series, if appropriately standardized, may be capable of capturing true patterns in abundance. Longline data represent important sources for estimating indices of abundance for many HMS, but these data require careful consideration.

Changes in fishing practices within the JLL (Ward and Hindmarsh, 2007; Yokawa and Uozumi, 2001) have resulted in changes in vertical habitats exploited over time. Perhaps the most notable modification was a shift to deeper target habitats in response to a change in target species from yellowfin tuna, *Thunnus albacares*, and albacore, *Thunnus alalunga*, to bigeye tuna, *Thunnus obesus*, and bluefin tuna, *Thunnus thynnus*. While many HMS are known to exhibit vertical migrations, they have been shown to spend a large percent of their time in a relatively small depth or temperature (relative to the surface) range (Goodyear et al., 2008; Graves et al., 2002; Hoolihan et al., 2011; Horodysky et al., 2007; Kerstetter et al., 2003; Prince et al., 2010; among others). Considering the shift in JLL target habitat, the vertical distributions of HMS suggest that the proportion of their stocks removed by a given unit of effort (catchability) has not been homogeneous over time with respect to the vertical habitats exploited. This emphasizes the importance of including vertical fishing habitat in the CPUE standardization process when estimating relative abundance.

Historically, there have been two general classes of approaches to including vertical habitat information in CPUE standardization for HMS: generalized linear models (GLMs) and habitat-based standardization (HBS) (Goodyear, 2003; Hinton and Nakano, 1996; Maunder et al., 2006). In a GLM, environmental data are typically considered indirectly by including variables related to longline gear configurations as fixed effects to serve as proxies for habitats fished. Habitats are directly included with HBS, and this approach has been used to estimate relative abundance for several HMS in the Pacific Ocean (Bigelow et al., 2002; Bigelow and Maunder, 2007; Hinton and Nakano, 1996; Langley et al., 2005; Maunder et al., 2006). The HBS approach was originally described as a deterministic model (Hinton and Nakano, 1996); however, the preferred method is cast in a statistical framework (statHBS; Maunder et al., 2006). In the statistical approach, the relative catchability from a given pre-specified vertical habitat is estimated by relating the total catch for a longline set to the amount of effort estimated to occur in each habitat for each set (see Section 2.2 for more details). The GLM approach is commonly used for Atlantic HMS; for example, blue marlin, *Makaira nigricans*, CPUE from the United States longline fishery was standardized using a delta-lognormal GLM with fixed effects for area, fishing characteristics, and gear characteristics (Ortiz and Hoolihan, 2010). The GLM and HBS approaches may provide different trends in abundance, which have been shown to affect assessment results (Uozumi, 2003). Accordingly, comparing and evaluating the accuracy of GLMs and HBS has been identified as an important research priority (ICCAT, 2004).

In addition to changes in exploited habitats, another common consideration is the proportion of records with zero catch (Maunder and Punt, 2004). For pelagic longline fisheries, this proportion can be relatively high, particularly for bycatch species. A high proportion of zero catches may violate the assumptions of the statistical analysis, and when the data are assumed to follow a lognormal probability distribution (a common assumption), computational issues arise because the natural logarithm of zero is undefined. These concerns are relevant for GLMs and HBS methods cast in a statistical framework (Section 2.2). One common approach to account for zeros is to add a small constant to all catch records before analysis (Maunder and Punt, 2004). Other approaches do not require the analyst to alter the data, such as using an assumed probability distribution that can include zero observations (e.g., Poisson or

negative binomial), or modeling the proportion of zero observations and the observations with positive catches separately (i.e., the delta-GLM approach; Aitchison, 1955; Lo et al., 1992; Maunder and Punt, 2004; Stefánsson, 1996; among others).

In this study, we simulated catch and effort data over a range of specified biomass trajectories and patterns in vertical catchability. We compared the indices estimated by statHBS and several formulations of GLMs, including delta-GLMs, when fit to the simulated data. The simulated catch data incorporated a trend in fishing effort that is similar to that of the JLL in the Atlantic Ocean from 1956 to 2009. Previous studies have compared similar approaches to standardizing indices of abundance (Bigelow and Maunder, 2007; Goodyear, 2003; Maunder et al., 2006); however none of these directly compared the commonly used models evaluated in this analysis.

2. Methods

2.1. Data simulation

A simulation analysis is well suited for evaluating methods used to estimate relative abundance, because the true pattern in abundance is known. To simulate realistic catch data in this study, we specified fishing effort to follow the temporal dynamics of the JLL in the Atlantic Ocean. The change in target species and target fishing depth exhibited by this fishery highlights the importance of considering vertical habitats fished (e.g., depth) when estimating relative abundance from these data. Therefore, the data simulation propagated effort over a range of vertical habitats and specified catchability to vary by habitat.

It is common practice in fisheries to assume that CPUE is proportional to abundance (N) using the following general relationship (Maunder and Punt, 2004):

$$CPUE = qN \quad (1)$$

where q represents catchability. This fundamental relationship served as the basis for data simulation, and by expanding it to incorporate a habitat-specific q , we simulated catch data per longline set following:

$$C_{y,s} = \left[\sum_h q_h E_{y,s,h} B_y \right] e^{\varepsilon_{y,s}}, \quad \varepsilon_{y,s} \sim N(\mu = 0, \sigma = 2) \quad (2)$$

where $C_{y,s}$ is the catch in biomass for longline set s in year y , q_h refers to the catchability associated with vertical habitat h , $E_{y,s,h}$ is the total effort associated with habitat h in set s in year y , B_y is the exploitable biomass in year y , and $\varepsilon_{y,s}$ is a random deviation in catch for set s in year y . When generating random deviates, we selected a value for σ that enabled $C_{y,s}$ to cover a realistic range as compared to many HMS caught by the JLL in the Atlantic. Overall, this simulation was relatively simple in that neither spatial structure nor population structure (size, age, sex, etc.) were considered; however, these simplifying assumptions are consistent with recent assessments of many HMS, such as Atlantic marlins, that used a single-stock approach without population structure (ICCAT, 2006).

We simulated catch data over a period of 54 years (1956–2009), a time-span during which the JLL operated continuously in the Atlantic Ocean (ICCAT, 2006; Uozumi and Nakano, 1994). Biomass trajectories were specified for this period by declaring an initial biomass in the first year ($B_{1956} = 500,000$ t), with biomass in the following years determined as a random deviation from the previous year. The random deviates were derived from a normal distribution with a mean specified in each simulation as a random uniform number between $-10,000$ and $10,000$ and a standard deviation of $10,000$. Thus, the biomass trend followed a correlated random walk that increased when the mean of the normal distribution was large

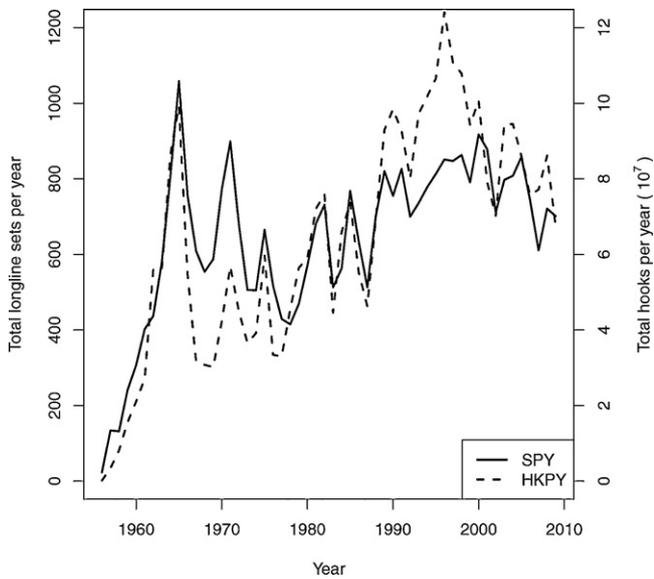


Fig. 1. Annual effort for the Atlantic Japanese longline fishery specified in the catch data simulation as total hooks and total sets per year (HKPY and SPY, respectively). The number of sets per year is an underestimate of the true number of sets for this fishery.

and positive, decreased when large and negative, and was stable when near zero. Also, a lower threshold of 100 t was imposed on B_y to prevent complete extirpation of the stock.

Fishing effort was specified as numbers of hooks, with one hook representing one unit of effort. Annual effort (total hooks fished per year) for the JLL in the Atlantic (Fig. 1) was determined using publically available catch and effort data from the International Commission for the Conservation of Atlantic Tunas (ICCAT).¹ The number of longline sets per year (Fig. 1) was determined from these catch and effort data as the total number of records per year, and the number of hooks fished per longline set was then fixed for each year at the number of hooks per year divided by the number of sets per year. The total number of records in the ICCAT database may not be an accurate representation of total sets per year, because each record may represent more than one set, but we feel this approach characterized the general pattern in sets per year for the JLL and improved efficiency in our simulation by reducing the total number of records (sets) generated.

To calculate the number of hooks from each set in each vertical habitat category, we specified a fishing depth for each hook in a set following catenary geometry (Bigelow et al., 2006; Yoshihara, 1951, 1954). These calculations required details pertaining to the configuration of longline gear over time, including the number of hooks per basket (HPB; i.e., hooks between floats), lengths of the branchline (b ; connects the hook to the mainline; also called gangion), floatline (f ; connects the float to the mainline), mainline between floats (ℓ), and the angle (φ) between horizontal and tangential of the mainline where it attaches to the floatline. The following equation was used to specify hook depths:

$$d_j = b + f + 0.5\ell \left\{ (1 + \cot^2 \varphi)^{1/2} - \left[\left(1 - 2 \frac{j}{(HPB + 1)} \right)^2 + \cot^2 \varphi \right]^{1/2} \right\} \quad (3)$$

where d_j is the depth of hook j and $j = 1$ to HPB for a given longline set. The number of HPB is often considered representative of longline target fishing depth with smaller numbers (3–6) used in shallower sets and larger numbers (10–20) used in deeper sets. Corresponding to a shift in target species, the proportion of sets with a large number of HPB increased throughout the 1980s (Serafy et al., 2004; Uozumi, 2003; Ward and Hindmarsh, 2007; Yokawa and Uozumi, 2001). This simulation followed the proportions outlined by Uozumi (2003) for HPB from 1975–1998. For years prior to 1975, we used the proportion reported in 1975, and for years after 1998, we used the proportion reported for 1998 (Fig. 2). The angle φ , although variable in practice, was set to 72° , a conventional assumption in previous studies (Ward and Myers, 2006), and the lengths of the longline components b and f increased over time, following the historical trend in gear configurations reported by Yokawa and Takuchi (2003) (Fig. 2). The value for ℓ was determined by assuming a fixed distance between longline hooks (45 m) throughout the simulation (i.e., $\ell = 45[HPB + 1]$; Ward and Hindmarsh, 2007). Finally, we reduced calculated hook depths by 25% to account for shoaling of the gear (Ward and Myers, 2006). Actual hook depth is likely influenced by several dynamic processes, and it has been shown that catenary algorithms do not accurately estimate longline hook depth (Rice et al., 2007; Ward and Myers, 2006). Thus, hook depths from our simulation were reflective of a general pattern for the JLL over time, but are not necessarily an accurate characterization of each set.

While it is important to have an understanding of hook depth, the vertical distributions of HMS are likely governed by a physiological response to the thermal properties of the water column (Brill and Lutcavage, 2001). Because these properties are highly variable over time and space, fishing depth does not necessarily reflect habitat fished. Therefore, we declared 16 vertical habitats (h) in which fishing effort could occur, each representing one degree deviations from sea surface temperature (i.e., 0 to -15°C). To simulate variability in the temperature at given hook depths, we randomly assigned each simulated longline set one of three possible temperature profile scenarios (i.e., relative temperature-at-depth): shallow, intermediate, or deep thermocline depth (Fig. 3). Relative temperature-at-depth was specified deterministically in each scenario to cover a range of possible temperature profiles. This effectively assigned each hook within each set to one of the 16 habitat categories and incorporated random variability between sets. Following the changes in longline gear configurations specified in the simulation, hook depths increased over time, which corresponded with deeper (cooler) habitats being fished (Fig. 2).

To relate habitat-specific fishing effort to vulnerability for the species being fished, we specified one of three potential scenarios for vertical habitat catchability (Fig. 3) in each simulation. The catchability in each habitat was influenced by relative temperature, and these scenarios were meant to cover the conceivable range for large pelagic fishes. While these organisms may spend the majority of their time in a preferred thermal regime, their vertical movement is likely a foraging strategy; therefore, the likelihood of taking a baited hook may not be constant with depth, meaning the vertical trend in catchability may not be directly related to the vertical distribution of the species (Goodyear et al., 2003). To account for uncertainty in vertical catchability, the three scenarios evaluated included peaks in different thermal habitats, relative to sea surface temperature (SST): surface, intermediate, and deep (Fig. 3). Each scenario was specified to have the same total catchability over the vertical habitats ($q_{total} = 2.5 \times 10^{-10}$), and q_h was determined as a proportion of total catchability allocated to each habitat using:

$$q_h = q_{total} P_h \quad h = 0, \dots, -15 \quad (4)$$

where in each scenario the proportion of total catchability for each habitat (P_h) was specified using the normal probability density

¹ <http://www.iccat.int>; data accessed September 2011.

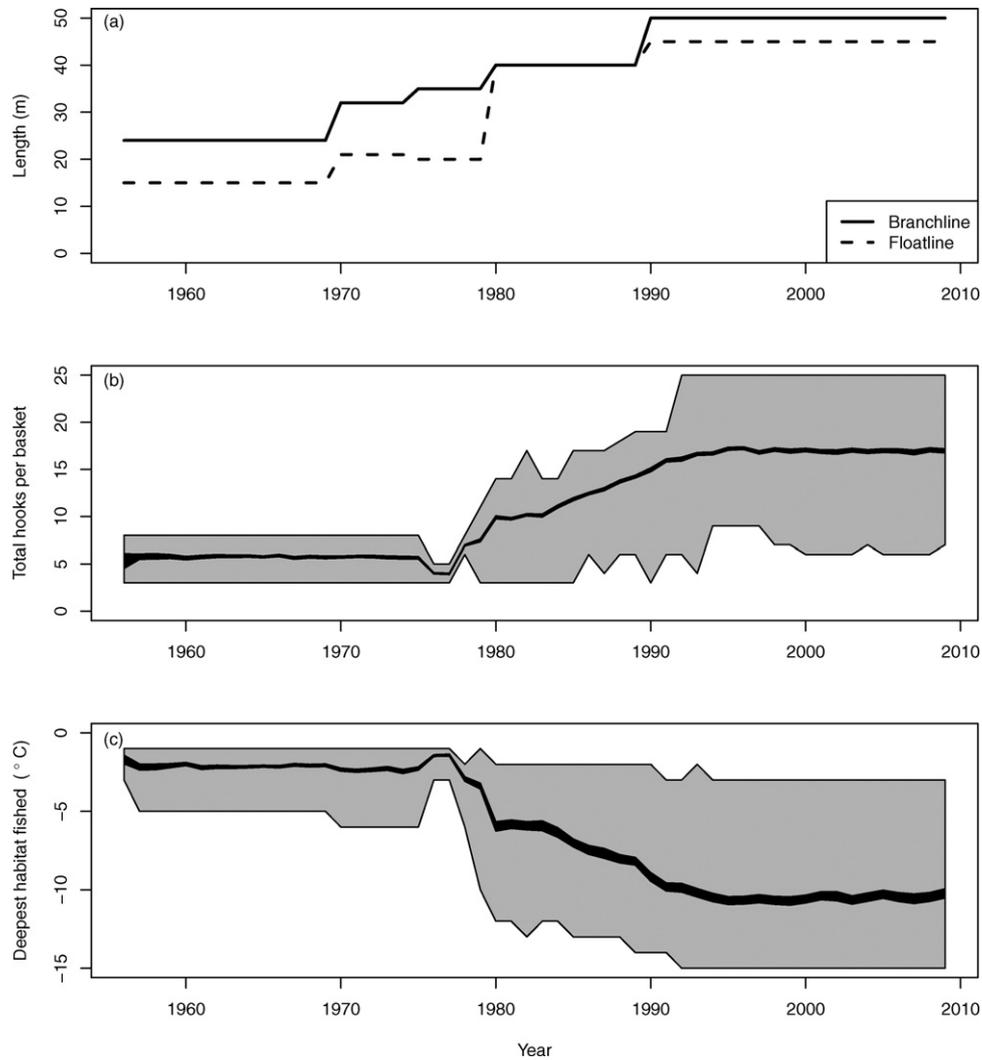


Fig. 2. Gear dynamics for the Atlantic Japanese longline fishery specified in the catch data simulation, including branchline and floatline lengths (a), and total hooks per basket (b) over time. Also, estimates of deepest habitats fished over time are presented (c). The total ranges (gray) and 95% confidence intervals (black) are presented for hooks per basket and deepest habitats fished.

function (PDF). The parameters of the normal PDF (μ , σ) varied between vertical catchability scenarios ($\mu = 0, 3, 5$ and $\sigma = 1.2, 1.5, 2.5$ for vertical catchability peaks at surface, intermediate, and deep habitats, respectively) and P_h was constrained to sum to one by dividing the probability densities in each habitat by the sum of the densities across h . We are unaware of any previously estimated patterns in vertical catchability for the JLL. The trends specified here allowed for the simulation of realistic catch data when following the effort dynamics of the JLL.

We repeated this simulation iteratively over 50 randomly determined biomass trajectories and the three possible scenarios for vertical catchability (Fig. 3) for a total of 150 simulated data sets. This captured the effects of a broad range of possible relationships between the abundance and vertical distribution of a large pelagic fish as related to the dynamics of the JLL. Also, to incorporate zeros in the catch data, sets with total catches less than 0.04 t were set to zero.

2.2. Approaches to estimating relative abundance

The 150 simulated datasets were used to compare the accuracy of indices of abundance estimated using three general approaches: the nonlinear statHBS model, GLMs, including delta-lognormal

GLMs (delta-GLMs), and non-standardized ‘nominal’ CPUE (calculated as average annual CPUE). The statHBS model, as described by [Maunder et al. \(2006\)](#), models catch data (C) from longline set i in a nonlinear framework as:

$$C_i = \hat{q}_{base} \hat{l}_t \sum_h \left(\hat{H}_h \sum_{i,j \in h} E_{i,j} \right) \quad (5)$$

where \hat{q}_{base} is overall catchability, \hat{l}_t represents the index of abundance over time t , \hat{H}_h is the difference in catchability for habitat h , and E_{ij} reflects the effort associated with hook j in set i (for this study each hook represented one unit of effort, so $E_{ij} = 1$). By multiplying the total number of hooks in habitat h by the change in catchability associated with that habitat (\hat{H}_h), effective effort is calculated for each longline set, thereby directly accounting for changes in vertical habitats exploited over time. The number of hooks per habitat category was calculated using catenary algorithms and longline configurations to determine hook depth, and the thermal properties of the water column to relate fishing depth to temperature relative to the surface (categorized as one degree deviations from SST). Estimates of hook depth and temperature-at-depth relied on the values specified for gear configurations and temperature profiles in the simulation. The estimable parameters of this model are

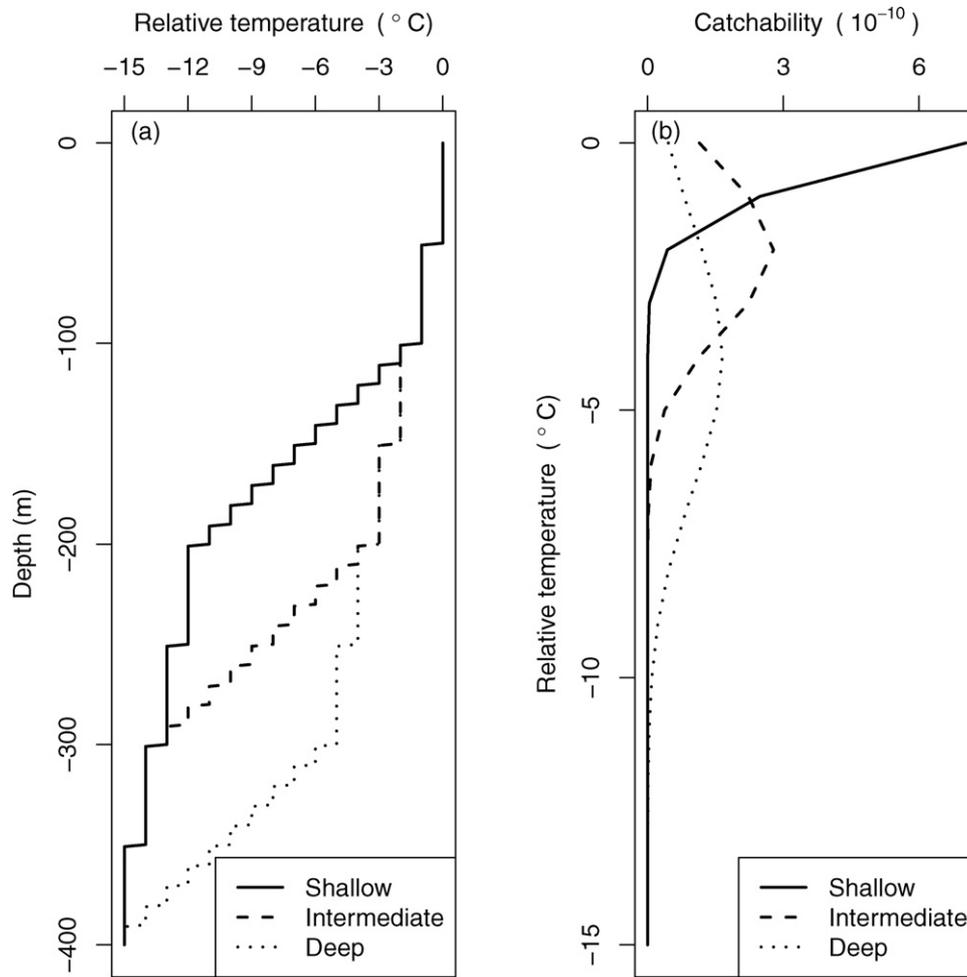


Fig. 3. Simulated ocean temperature profiles (a) and scenarios of vertical catchability (b). The temperature profiles were randomly assigned to simulated longline sets and were meant to reflect shallow, intermediate, and deep thermocline depths. The catchability scenarios exhibited peaks in surface (shallow), intermediate, and deep waters. The top of each panel represents the surface of the water with depth (a) and temperature (b) changing vertically.

\hat{q}_{base} and elements of the vectors \hat{l}_t and \hat{H}_h . Overall catchability is confounded with \hat{l}_t and \hat{H}_h , so l in the first year is set to one, and \hat{H}_h is constrained to sum to one. These parameters were estimated by fitting this model to simulated C_i by minimizing the following negative log-likelihood function:

$$-\ln L(\theta|C_i) = \sum_i \ln(\sigma) + \left[\frac{(\ln(C_i + \delta) - \ln(\hat{C}_i + \delta))^2}{2\sigma^2} \right] + \sum_h \left(\ln(\hat{H}_h) - \ln\left(\frac{1}{16}\right) \right)^2 \quad (6)$$

where \hat{C}_i is the predicted catch for longline set i , θ is a vector of parameters, δ is a small constant (1) used to prevent taking the log of zero, and σ is the estimated standard deviation (in log space). The last term was included to improve convergence by penalizing deviation from a uniform distribution under the constraint that \hat{H}_h must sum to one.

The GLM-based approach to standardizing CPUE data has been used in stock assessments for many years (Maunder and Punt, 2004). A GLM can accommodate non-normal error structure by relating the expected i th value of a response variable to a set of predictor variables through a link function (McCullagh and Nelder, 1989; Nelder and Wedderburn, 1972). Since GLMs are linear

models, a linear response is assumed between the function of the expected response and the explanatory variables as:

$$g(\mu_i) = x_i^T \beta \quad (7)$$

where g represents the link function, μ_i is the predicted response, x_i^T is a transposed vector of explanatory variables including an intercept and a categorical variable representing the time step (at a minimum), and β is a vector of parameters. For the GLM, CPUE adjusted by a small number (1×10^{-3}) rather than catch (as with statHBS) was the response variable and lognormal error structure was modeled by taking the natural logarithm of CPUE (the identity link function was used). Annual estimates of relative abundance and their respective standard errors were obtained as back-transformed year means with an infinite series lognormal bias correction following Lo et al. (1992).

The GLM is expanded in a delta-GLM, which models the probability of observing a zero catch as a function of the explanatory variables, and then separately fits a GLM (with the same explanatory variables) to the non-zero catches (Aitchison, 1955; Lo et al., 1992; Maunder and Punt, 2004; Stefánsson, 1996; among others). This approach is represented by:

$$Pr(Y = y) = \begin{cases} w, & y = 0 \\ (1 - w)f(y) & \text{otherwise} \end{cases} \quad (8)$$

In Eq. (8), w is the probability of observing a zero for the response variable (CPUE) and $f(y)$ is a model of the mean of the non-zero

CPUEs. For our analyses, the proportion of sets with positive CPUE ($1 - w$) was modeled using a binomial GLM with a logit link function, and the positive CPUEs were modeled with a lognormal GLM as previously described (though, without adjusting CPUE). Annual probabilities of positive catches from the binomial GLM were the back-transformed mean values for each year predicted when all additional factors were set to the level representing the observed mode (Maunder and Punt, 2004). Annual estimates of abundance were then obtained by multiplying the probability of positive observations for year y and the back-transformed bias corrected year means from the lognormal GLM of positive catches. To capture the precision of the abundance estimates, standard errors were calculated using the delta method (Lo et al., 1992; Seber, 1982).

It is important when evaluating the accuracy of GLMs and delta-GLMs to consider which explanatory variables should be included in the models. The year in which the catch occurred must be one of the variables, because the output of interest is relative abundance over time. The number of HPB is commonly used for longline catch and effort data as a proxy for fishing depth, so this factor was also considered in the analysis. Typically, when linear models are used to estimate relative abundance from longline data, habitats exploited are not directly considered, because information cannot be summed over multiple habitats as with statHBS (Maunder et al., 2006). However, catenary geometry and water column temperature profiles can be used to combine fishing and environmental data into a single variable to be incorporated directly into the GLMs. For this type of variable, it is important to consider the degree of detail since a factor with numerous levels can result in an over-parameterized model. Thus, using the catenary algorithm (Eq. (3)) to estimate hook depths, and the simulated temperature profiles to relate depth to temperature, we created a factor that represents the largest change in temperature relative to SST ($\text{Max}\Delta T$) for each longline set (i.e., deepest/coldest habitat fished per set).

To determine which predictor variables to include in the GLMs, a series of GLMs were evaluated using a range of configurations for the main effects of year, HPB, $\text{Max}\Delta T$, and each possible first-order interaction (following convention, interactions with year were modeled as random effects). In practice, model selection techniques, such as Akaike's Information Criterion (AIC) or analyses of residual deviance, are often employed when selecting a model for generating a standardized index of abundance (Maunder and Punt, 2004). However, in a simulation analysis, estimated abundance trends can also be compared with the simulated 'true' pattern in abundance to determine model performance. Thus, we fit each proposed GLM configuration to all simulated datasets and recorded AIC, residual deviance, and a metric for model accuracy (Section 2.3). Selection via AIC was based on AIC for each model minus the minimum observed AIC (ΔAIC). The model with $\Delta\text{AIC}=0$ represented the 'best' model of those evaluated (Burnham and Anderson, 2002). For our deviance analyses, we calculated the percent of deviance explained by the stepwise addition of each factor (Ortiz and Arocha, 2004), and if the percent explained was less than the somewhat arbitrary cutoff value of 1%, then the model was rejected (Maunder and Punt, 2004). For comparisons made between different random effects, models were fitted using restricted maximum likelihood estimation, while standard maximum likelihood estimation was used when comparing different configurations of fixed effects (Zuur et al., 2009). We conducted the analyses of model configurations using $\ln(\text{CPUE} + 1 \times 10^{-3})$ as a response variable, then repeated the analyses using only the records with positive CPUEs (i.e., $\ln(\text{CPUE}) > 0$) was the response). We considered four configurations of fixed effects ($\text{FE1} = \text{Year}$, $\text{FE2} = \text{Year} + \text{Max}\Delta T$, $\text{FE3} = \text{Year} + \text{Max}\Delta T + \text{HPB}$, and $\text{FE4} = \text{Year} + \text{Max}\Delta T + \text{HPB} + \text{Max}\Delta T \times \text{HPB}$), and fit the candidate models to the 150 simulated datasets. To summarize the selection metrics across all scenarios, we recorded their median values

and calculated the percent of times each model structure would have been selected following the three selection approaches. The summary statistics for model fit and model accuracy were then synthesized to select a single GLM structure for all simulated datasets. We did not allow the model structure to change during the simulations, because we did not want to introduce model selection as an additional level in our evaluations. The overall comparisons were then confined to five approaches: statHBS, a GLM and delta-GLM with the selected model structure, a delta-GLM without the $\text{Max}\Delta T$ factor (delta-GLM2) for evaluation of this unique variable, and nominal CPUE, because many studies (particularly meta-analyses) have used this approach to describe trends in abundance (e.g., Myers and Worm, 2003).

2.3. Model evaluation and comparison

The GLMs and statHBS model were fit to different response variables (CPUE and catch, respectively); therefore, common statistical comparisons of model fit (e.g., AIC) could not be used to compare these models. However, this does not represent a substantial drawback, because these statistical metrics are typically concerned with balancing model fit and parsimony for a given dataset, and the most important aspect in this evaluation is the accuracy of the estimated trend in annual biomass. Because estimated trends were treated as relative indices, the estimates from the various models were scaled for comparison with the associated 'true' biomass. Model comparisons were based on the annual percent difference in simulation scenario x for model m in year y ($\% D_{x,m,y}$) between true biomass and estimated biomass calculated as:

$$\% D_{x,m,y} = \frac{100|B_{x,y} - (\hat{I}_{x,m,y}/\bar{I}_{x,m,y})\bar{B}_{x,y}|}{B_{x,y}} \quad (9)$$

where $\hat{I}_{x,m,y}$ refers to the index of abundance estimated in simulation x for model m in year y . To facilitate the comparison of methods, the median of the annual percent differences (MPD) was calculated for each model m to serve as a single metric of accuracy for each simulated time series x and the standard deviation (SDPD) was calculated to characterize the variability of $\% D_{x,m}$ across years. Our performance metrics compared the absolute value (magnitude) of the relative errors. To evaluate the pattern (directionality) in the errors, we analyzed overlaid plots of 'true' and estimated biomass across all scenarios.

2.4. Sensitivity

The habitat variables specified in the GLMs and the statHBS model relied on estimates of hook depth for each longline set. By using the conditions described in the data simulation, these dynamics were known without error in the analyses. However, hook depth is notoriously difficult to estimate with accuracy (Rice et al., 2007; Ward and Myers, 2006), and uncertainty surrounding inputs for HBS methods have been shown to affect the accuracy of these approaches (Goodyear, 2003). Thus, we conducted a sensitivity analysis with a single simulated data set, assuming declining biomass over time and a peak in vertical catchability in surface waters. This scenario potentially reflects the vertical distribution and biomass trajectory of several HMS in the Atlantic Ocean. The statHBS model and a delta-GLM including the $\text{Max}\Delta T$ factor were then fit to the data assuming different degrees of uncertainty regarding estimated hook depth. Because estimates of hook depth may be biased, three scenarios were considered in the sensitivity analysis (hook depths always underestimated, always overestimated, or randomly over/underestimated in each longline set), and hook depth errors were specified over a range of 5–50% (incremented by 5%) for each scenario. The model evaluation metrics

(Section 2.3) were then compared across the range of errors specified.

2.5. Implementation

With the exception of the statHBS model, the simulation and analyses were executed using the statistical programming language R (R Development Core Team, 2011). Due to increased model complexity, statHBS was implemented using AD Model Builder (ADMB Project, 2011), which was called from R.

3. Results

The simulation model generated 150 separate catch and effort time series from 50 random biomass trends, each repeated three times to evaluate different profiles of vertical catchability. Of the 150 simulations, we highlighted nine as representative of the range of possibilities for patterns in relative abundance and vertical catchability (Fig. 4). When catchability was highest in surface waters (Fig. 4a–c), simulated catch peaked early in the time series, but

then decreased as the fishery began targeting deeper habitats. When catchability was highest in intermediate and deeper habitats (Fig. 4d–i), simulated catches tracked the true trend in abundance more closely. Also, because one component of this study is the treatment of records with a total catch equal to zero (i.e., either add a constant or use the delta approach), we summarized the proportion of zeros across all simulation scenarios (minimum, mean, and maximum proportions were 0.02, 0.11, and 0.41, respectively). The highest proportions of zero catches occurred when catchability was simulated to peak in surface waters or when biomass was simulated to decrease over time.

Selection of a model structure for the GLMs and delta-GLMs was not straightforward, because the three selection metrics did not always lead to the same conclusion (Table 1). The AIC approach clearly favored FE3, while the deviance analysis selected FE2 and FE4 most frequently. However, the accuracy of the models is their most important attribute, and MPD indicated that FE3 and FE4 exhibited the greatest accuracy, with FE4 selected slightly more often, and a lower overall MPD for FE3. In fact, the AIC approach selected the most accurate configuration in only 19.3% of the

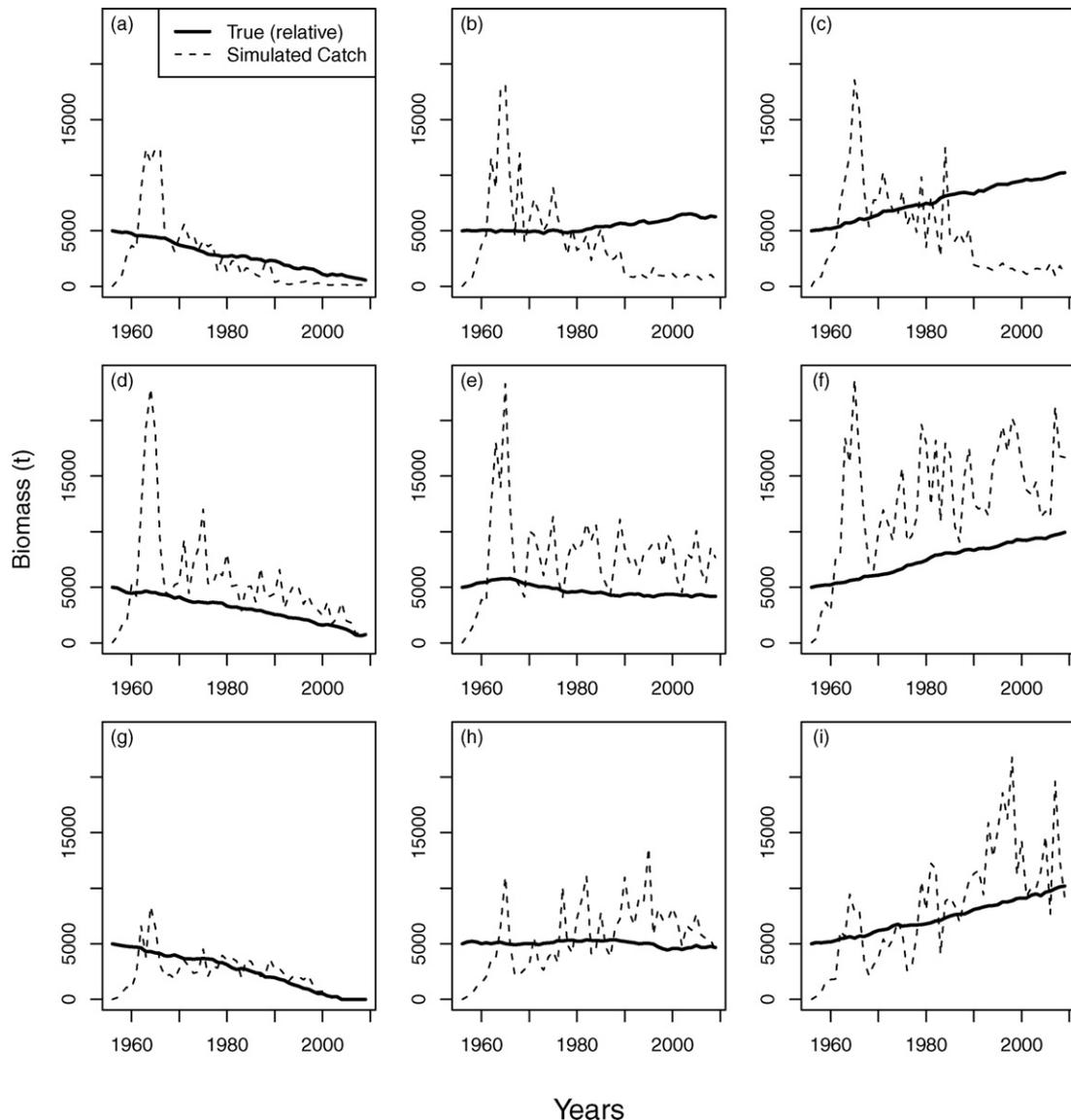


Fig. 4. Simulated catch data generated from a known underlying biomass trajectory for nine selected simulation scenarios where vertical catchability was assumed to peak in surface (a–c), intermediate (d–f), or deep waters (g–i), and the simulated ‘true’ biomass trajectory (scaled to total catch) was either decreasing (a, d, and g), stable (b, e, and h), or increasing (c, f, and i) over time.

Table 1
Summary of linear model selection metrics for different fixed effect structures. The four hypothesized model structures were evaluated when fit to all catch data adjusted by a constant, and when fit to the positive catch records only. Median values across all 150 simulated datasets were presented for the change in Akaike's Information Criterion from the minimum value (ΔAIC), the percent of the total residual deviation explained by the stepwise addition of each factor (% total dev.), and the median percent difference between estimated and true biomass (MPD). Also presented are the percent of times each model structure was selected by each corresponding selection metric (% S).

Model structure	ΔAIC	% total dev.	MPD	% S_{AIC}	% S_{Dev}	% S_{MPD}
<i>Response: $\log(CPUE + 1 \times 10^{-3})$</i>						
Year	1055.3	73.2	33.9	0.0	0.0	12.7
Year + Max ΔT	21.7	24.6	12.2	19.3	47.3	10.7
Year + Max ΔT + HPB	0.0	1.0	9.7	62.0	5.3	27.3
Year + Max ΔT + HPB + Max $\Delta T \times$ HPB	29.1	1.5	11.3	18.7	47.3	32.0
<i>Response: $\log(CPUE > 0)$</i>						
Year	762.1	75.6	35.0	0.0	0.0	8.7
Year + Max ΔT	10.8	21.2	9.1	33.3	42.7	25.3
Year + Max ΔT + HPB	0.0	1.2	10.1	58.0	8.0	29.3
Year + Max ΔT + HPB + Max $\Delta T \times$ HPB	37.5	1.9	11.1	8.7	49.3	30.0

scenarios, and the deviance analysis led to the most accurate model in only 10.7% of the scenarios. We decided to use FE3 rather than FE4 throughout the simulation runs, because it is a more parsimonious model, and also because the interaction term caused convergence issues with the delta-GLM. We did not observe improved accuracy when the model contained random effects to describe interactions with the Year coefficient; thus, these interactions were not included in the analyses.

Indices of abundance were estimated for each simulated data set following five approaches (delta-GLM, delta-GLM2, GLM, nominal, and statHBS) and evaluation metrics were calculated for each approach. The overall comparison of MPD for all simulations (Fig. 5a) suggests that the delta-GLM that included detailed vertical habitat information provided the most accurate estimates of relative abundance (i.e., the lowest overall median of MPD), though the variability of this metric spanned a slightly larger range than that for GLM. When accounting for different patterns in vertical catchability across all biomass trends (Fig. 6a–c), delta-GLM was most accurate, except when vertical catchability peaked in surface waters (Fig. 6a). Comparison of MPD over different biomass trajectories across all catchability scenarios (Fig. 6d–f) suggested that

delta-GLM was most accurate except when biomass increased over time (Fig. 6f). Overall variability in the accuracy of the estimates was lowest for delta-GLM2 (Fig. 5b), which exhibited the lowest SDPD when vertical catchability peaked in deep waters (Fig. 6i) and when biomass increased over time (Fig. 6l). However, in general, SDPD was relatively consistent across all approaches except when using nominal CPUE.

The nine highlighted simulation scenarios were used to visually evaluate estimated trends in abundance as compared to 'true' abundance trends (Fig. 7). The patterns in these plots were consistent across all simulated scenarios, and suggest that when vertical catchability peaked at the surface (Fig. 7a–c) all approaches overestimated abundance early in the time series, but underestimated abundance when catches declined with changes in fishing strategy. These patterns were exacerbated for less accurate approaches, particularly nominal CPUE. For scenarios where vertical catchability peaked in intermediate and deep waters, there were no consistent patterns in the errors, irrespective of biomass trajectory (Fig. 7d–i).

Finally, the sensitivity analysis indicated that for the scenario evaluated, the statHBS model was sensitive to hook depth uncertainty, yet the delta-GLM with detailed habitat information was not particularly sensitive to this error (Fig. 8). Model evaluation metrics for statHBS and delta-GLM were compared with baseline metrics calculated assuming no error in estimated hook depth as well as delta-GLM2 which did not include the Max ΔT factor. The accuracy of statHBS was compromised when hook depths were always underestimated or randomly over/underestimated. When hook depths were always overestimated, accuracy did not decline until hook depth estimates were at least 30% different from actual fishing depths (Fig. 8a). For the delta-GLM, accuracy was not substantially reduced as hook depth uncertainty increased; though when hook depths were always overestimated, a relatively small amount of error caused delta-GLM to be less accurate than delta-GLM2, but larger errors in estimated hook depths actually improved accuracy (Fig. 8b). The statHBS model also appeared more sensitive than delta-GLM to hook depth uncertainty in terms of SDPD (Fig. 8c and d). For the delta-GLM, SDPD was relatively consistent across the range of hook depth uncertainty for all scenarios.

4. Discussion

There are numerous challenges surrounding the assessment and management of HMS (Lynch et al., 2011). Among these include a lack of fishery-independent scientific research, resulting in a reliance on fishery catch and effort data for making inferences about population dynamics. In this study, we evaluated common approaches to estimating indices of abundance for HMS from longline fishery data when the fishery exhibited a change in target fishing depth. This shift in the vertical distribution of effort is a common feature of many pelagic longline operations to which HMS

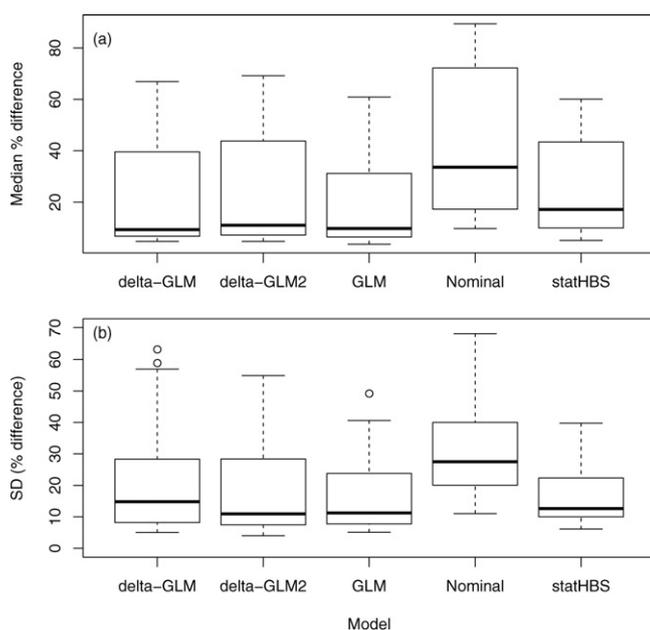


Fig. 5. Box and whisker plots of performance metrics, including median percent difference between estimated biomass and true biomass for a simulated dataset (a), and standard deviation (SD) of the percent differences (b) for five approaches to estimating an index of abundance. The solid line reflects the median, the box encompasses the interquartile range, the whiskers extend to the extreme values, and circles reflect potential outliers.

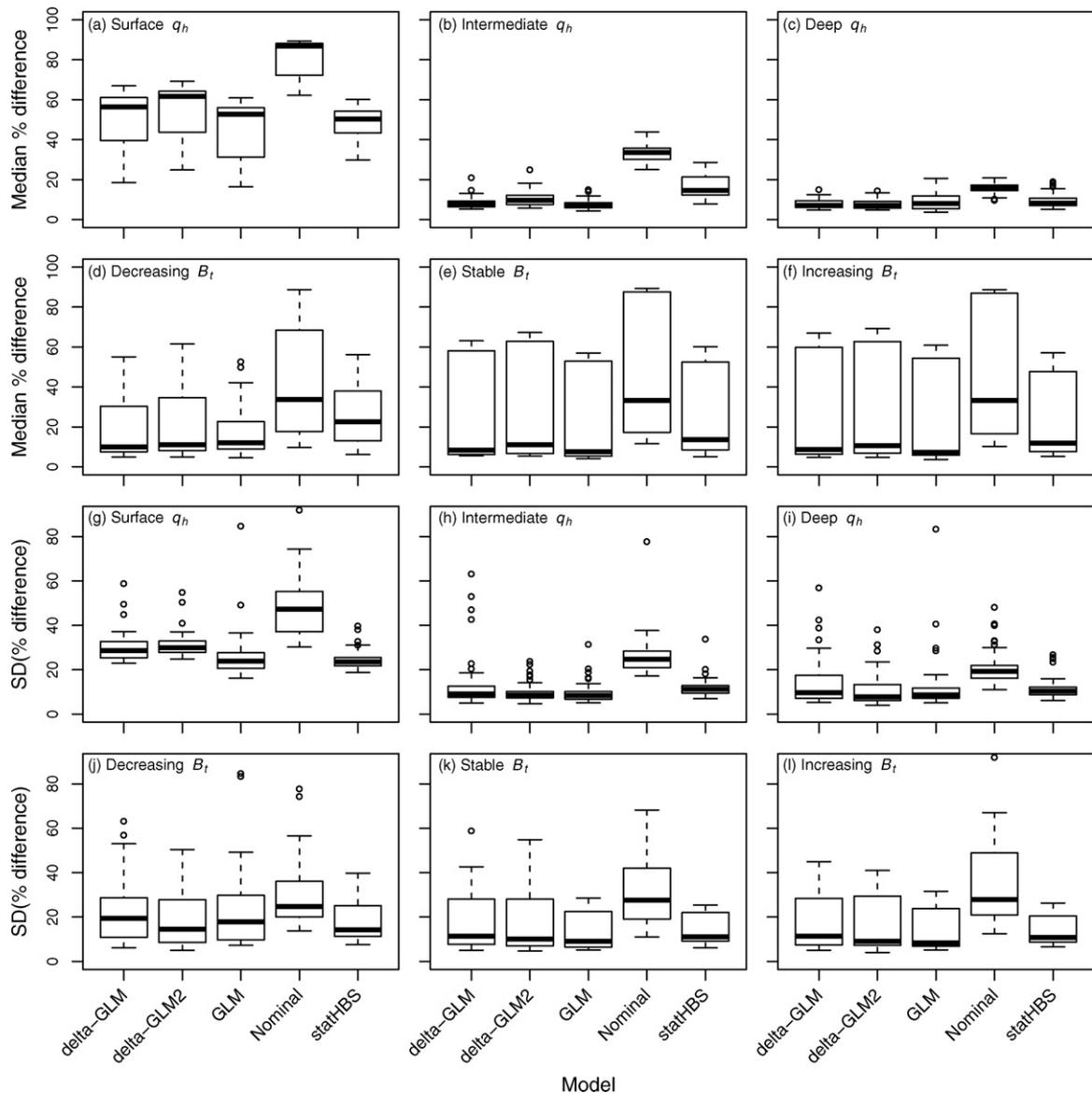


Fig. 6. Box and whisker plots of performance metrics, including median percent difference between estimated biomass and true biomass for a simulated dataset, and standard deviation (SD) of the percent differences for five approaches to estimating an index of abundance. The performance metrics were summarized for scenarios where vertical catchability (q_h) was simulated to peak in surface (a and g), intermediate (b and h), or deep waters (c and i), and where biomass (B_t) was simulated to decrease (d and j), remain stable (e and k), or increase (f and l) over time. The solid line reflects the median, the box encompasses the interquartile range, the whiskers extend to the extreme values, and circles reflect potential outliers.

are susceptible (Majkowski, 2007). Thus, we evaluated models that accounted for habitats exploited, such as statHBS and GLMs that included a unique habitat factor (Max ΔT). By including this habitat variable, we essentially created a bridge between statHBS and GLM approaches, because under both methods, estimates of vertical habitats exploited are derived by relating estimated longline hook depth to surrounding environmental conditions. Our overall conclusion from this study is that the most accurate approach was a delta-GLM that included the Max ΔT habitat factor. Including this information improved the performance of the linear models tested, and to our knowledge, this type of detailed habitat factor has not previously been used for obtaining annual estimates of relative abundance from longline data. However, delta-GLM did not exhibit the lowest variability in the errors of the approaches evaluated. This suggests that while delta-GLM may provide more accurate estimates of abundance across a time series, the magnitude of the error is less consistent than that for other methods. Still, given that SDPD does not differ substantially across approaches, we

maintain that delta-GLM is the preferred approach on the basis of MPD. It should be noted that we evaluated the base form of the statHBS model as described by Maunder et al. (2006). We did not consider alternative formulations; however, Maunder et al. (2006) suggested several that may improve the accuracy of the model.

While delta-GLM provided the most accurate estimates of abundance overall, other approaches provided the lowest MPD in several scenarios. For instance, when vertical catchability peaked at the surface, GLM and statHBS were generally more accurate than delta-GLM. However, it is important to recognize that under this catchability scenario all models provided relatively inaccurate estimates of abundance (Fig. 6a), even though the fishery and environmental data included in the analyses were known without error. A potential explanation for this phenomenon is that the shift by the fishery to deeper target habitats caused catches to decline over time to a level where there was no longer enough information to estimate annual biomasses with accuracy. This emphasizes the point that fisheries do not always sample populations effectively; thus,

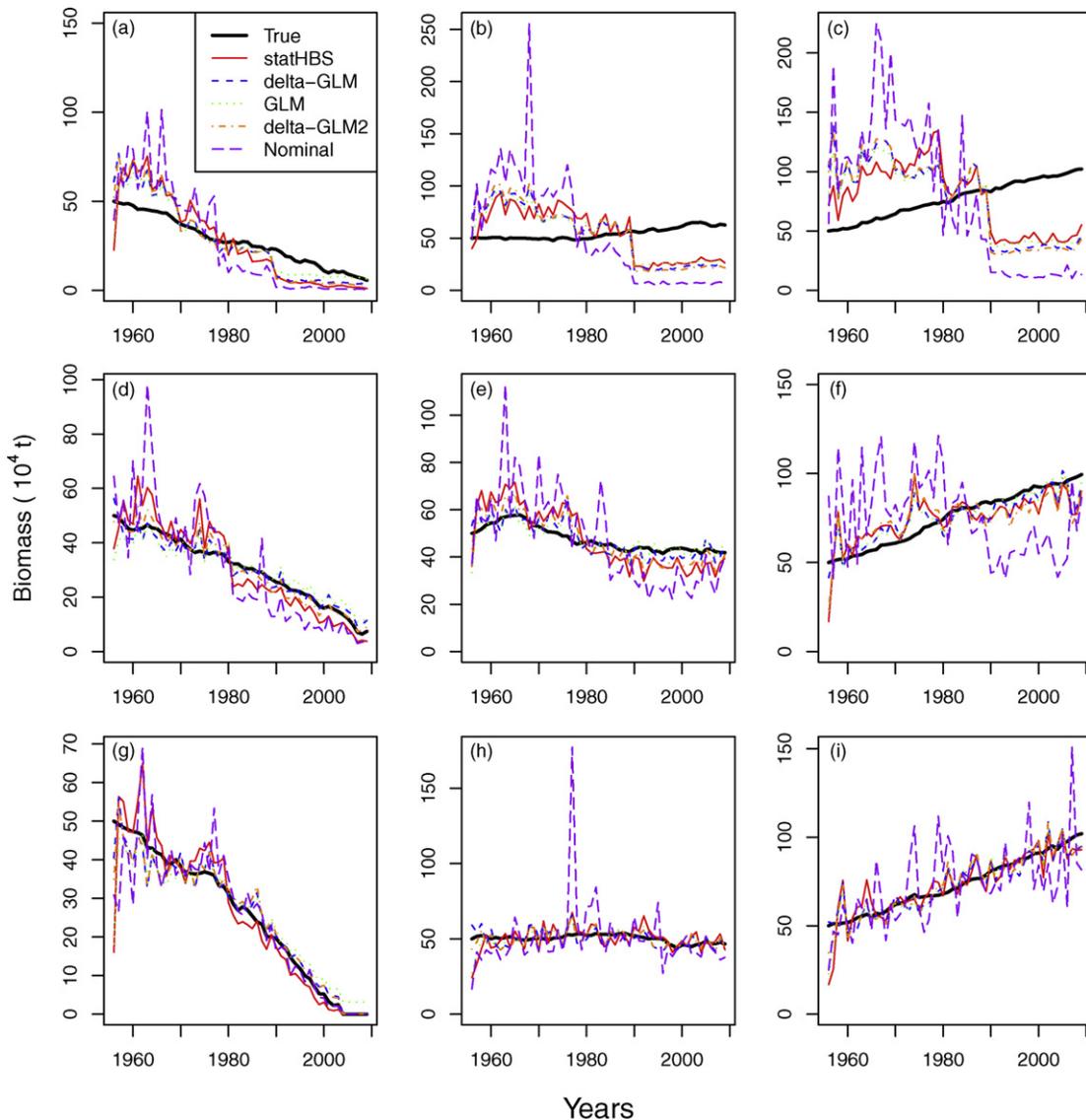


Fig. 7. Simulated 'true' biomass plotted against annual biomass from five approaches to estimating an index of abundance. Results are highlighted for nine selected simulation scenarios where vertical catchability was assumed to peak in surface (a–c), intermediate (d–e), or deep waters (g–i), and the known biomass trajectory was either decreasing (a, d, and g), stable (b, e, and h), or increasing (c, f, and i) over time.

it is risky to rely on fishery-dependent data for making inferences about population dynamics, even when changes in fishing strategy are properly considered.

When indices of abundance fail to capture stock dynamics, it is important to determine if the estimated trends exhibit hyperstability (i.e., abundance decreases more rapidly than the index) or hyperdepletion (the index decreases more rapidly than abundance) (Hilborn and Walters, 1992). For scenarios with a peak in vertical catchability at the surface, we observed evidence of hyperdepletion in estimated trends in abundance (Fig. 7a–c). If these scenarios captured true patterns in vertical catchability for bycatch species, then our results support the assertion by Uozumi (2003) that assessments of bycatch HMS in the Atlantic have been overly pessimistic.

In general, the delta approach did not result in substantial improvements over the traditional GLM. This suggests that the treatment of zeros (i.e., longline sets with total catch equal to zero) constitutes another important aspect of our evaluations. The delta approach directly accounts for zero observations, and for statHBS and GLM the data are adjusted to accommodate zeros.

Therefore, when the proportion of zero observations is relatively high, the delta approach would be expected to outperform methods that alter the data; however, statHBS and GLM performed better than delta-GLM in the scenario that resulted in the highest proportions of zeros (surface peak in catchability). Because the treatment of zeros was not the primary focus of our analyses, this unexpected result warrants further evaluation of the influence of zero observations on methods used to estimate indices of abundance. Additional studies should also consider the choice of assumed error distribution, since discrete distributions (e.g., Poisson, negative binomial) would not require the data to be altered.

Another aspect of our study that could benefit from additional research is the application of model selection criteria when generating indices of abundance. For our simulated scenarios and model configurations, neither AIC nor deviance analysis were reliable for predicting the model that provided the most accurate index of abundance. Further studies are necessary, because these approaches are commonly used when selecting models to standardize CPUE. Alternative selection approaches to consider include

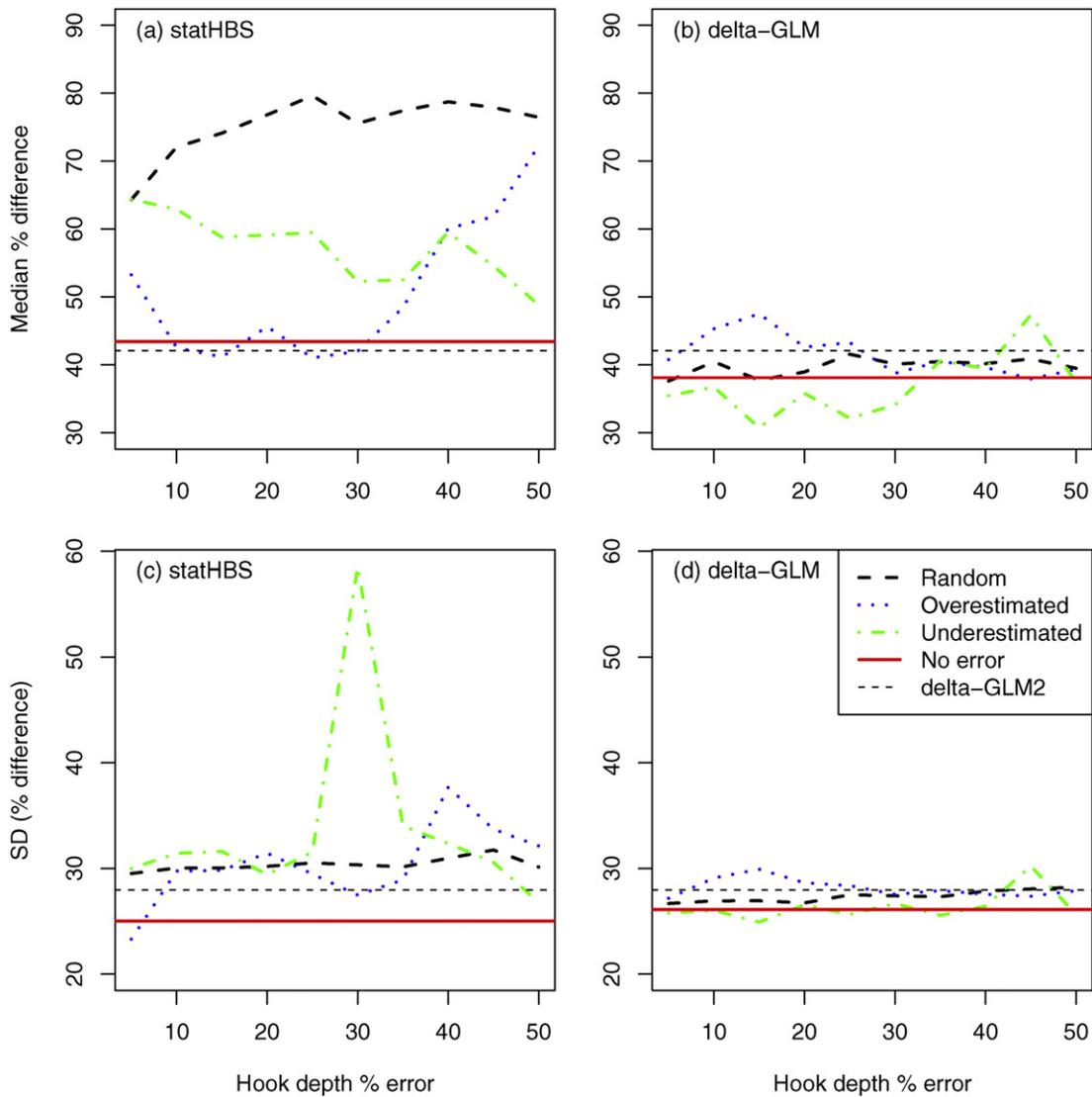


Fig. 8. Sensitivity plots for statHBS and a delta-GLM with an effect for habitats fished. These plots characterize the effect of hook depth uncertainty on the median percent difference between estimated and simulated biomass (a and b), and the standard deviation (SD) of the percent differences (c and d). The percent of error in estimated hook depth was applied as either overestimated (blue), underestimated (green), or with random directionality (black). The solid red line represents the result with no error in estimated hook depth, and the dashed line represents the result for a delta-GLM without an effect for habitats fished.

cross-validation (Maunder and Punt, 2004) and consistent information criteria (Shono, 2005).

Our catch and effort simulation was based on the dynamics of the Japanese Atlantic longline fishery, partly because this fishery has been recommended in this context (ICCAT, 2004), but also because it captures a scenario when changes in the vertical distribution of fishing effort may have affected catchability. Because fisheries continually evolve in response to advances in technology and shifts in the global demand for resources, dynamic catchability is likely a common feature underlying the catch data of many species, especially HMS caught by longline fisheries that have changed targeting practices. Thus, the data simulated in this study incorporated a shift in the distribution of fishing effort across vertical habitats to reflect a change in target species. While this important feature was incorporated, the simulation did not consider other potentially important dynamics, such as the geographical distribution of fishing effort over time, or variability due to size, age, or sex in the fish population. In some cases, including such detail may be essential. For instance, Prince et al. (2010) demonstrated that catchability of bycatch HMS may be higher

inside than outside the oxygen minimum zone of the eastern tropical Atlantic. Standardization of CPUE with GLMs could account for this by including an appropriately defined area factor in the linear model, whereas statHBS could be expanded to include a GLM component with this area factor (Maunder et al., 2006).

While the incorporation of vertical habitats fished improved estimates of relative abundance, its main practical limitation is the ability to estimate hook depths with accuracy. Catenary algorithms are typically used to generate static estimates of hook depth for an entire longline set, yet numerous factors (wind, hydrodynamics, behavior of hooked organisms, etc.) can cause substantial deviations from predicted depth for a given hook position both within and between sets (Bigelow et al., 2006; Rice et al., 2007; Ward and Myers, 2006). By conducting a sensitivity analysis over a range of errors in estimated hook depths, we quantified the effects of this uncertainty on the performance of statHBS and a delta-GLM. The high sensitivity of statHBS is perhaps due to the reliance of this method on a detailed characterization of hook depth, whereas the less sensitive delta-GLM used a simple categorical variable to describe vertical habitats fished. Although the Max ΔT variable

provided a simplistic depiction of habitats fished, its inclusion substantially improved the accuracy of the model (Table 1). Therefore, it is interesting that this influential factor was not particularly sensitive to error. In general, the sensitivity analysis provided a simple characterization of the influence of hook depth uncertainty. To provide a baseline evaluation, we selected a single simulated scenario that potentially reflects the vertical distribution and biomass trajectory of many HMS in the Atlantic Ocean (i.e., surface peak in catchability and declining biomass over time). For instance, stock assessments of Atlantic marlins indicated that their biomasses are in decline (ICCAT, 2006), and their vertical distributions have been characterized as surface oriented with occasional deep dives (Goodyear et al., 2008; Graves et al., 2002; Horodysky et al., 2007; Kerstetter et al., 2003; Prince et al., 2010). More comprehensive characterizations of hook depth uncertainty are warranted, and we encourage additional analyses that consider a variety of simulated scenarios. Additionally, our evaluations were conducted assuming no error in the catch data, gear dynamics, or oceanographic conditions. In practice there may be uncertainties surrounding each of these inputs, and a characterization of their effects would be useful. For example, temperature profiles for each longline set would ordinarily be obtained from a global ocean database, which may require interpolation, and therefore may not reflect the true temperature profile for each set.

In addition to expanding the evaluation of the effects of hook depth uncertainty, the results of this study highlight the importance of several areas of research. For instance, the development of a more sophisticated approach to modeling longline gear behavior as a function of environmental conditions could be useful in these applications. Accordingly, increasing the detail of fishery data reporting to include catch by hook position and corresponding environmental conditions could foster a better understanding of catches by habitat. Additionally, we defined vertical habitat categories as 1°C deviations from SST, but decisions about how to partition habitat may influence estimates of abundance that include this information. This emphasizes the importance of continued research on behavior, physiology, and habitat use of fishes, and in future analyses, model performance may be further improved using model selection (e.g., AIC) to determine an appropriate degree of detail in the habitat factor. However, based on the results of our selection of a fixed structure for the linear models, caution should be exercised when using traditional model selection metrics for CPUE standardization. Furthermore, the methods evaluated are best suited for estimating historical trends in abundance, yet predictions about future patterns are also important for fisheries management. In reality, complex dynamics likely govern true abundance trends, and most approaches for estimating relative abundance do not incorporate these relationships. However, nonlinear forecasting may be a promising approach for predicting future abundance when the governing equations are unknown (Glaser et al., 2011).

The focus of our study was on estimating relative abundance in the presence of changes in the vertical distribution of fishing effort, but a change in the geographical distribution of effort is an equally important consideration (Walters, 2003). Fishery-dependent data provide information from areas fished, which typically represent areas of high profitability. Without accounting for abundance trends in areas that were not fished, there is an implicit assumption that CPUE trends in areas fished are reflective of trends in areas that were not fished. In many cases this assumption may not be valid, and it has been shown to bias estimates of relative abundance (Campbell, 2004; Carruthers et al., 2010; Walters, 2003). To account for this source of bias in a CPUE standardization context, Carruthers et al. (2011) described a GLM approach to CPUE standardization that includes data imputation in unfished spatial cells. Their approach could easily be adapted to account for vertical habitats as described

herein, thereby directly addressing effects on catchability due to changes in the distribution of effort in three dimensions.

Since the description of HBS by Hinton and Nakano (1996), the choice between using HBS or a GLM approach to estimating the relative abundance of HMS has been controversial (Bigelow and Maunder, 2007; Goodyear, 2003; Goodyear et al., 2003; Maunder et al., 2006; Prince and Goodyear, 2006; Ward and Myers, 2005). Debate over this choice could be expected given the potential influence on assessment results (e.g., Uozumi, 2003) and the fact that regulations aimed at conserving HMS affect highly valued international fisheries. Nevertheless, the controversy emphasizes the importance of comparing and evaluating these methods, and the results of our study should be interpreted in the context of previous research (i.e., Bigelow and Maunder, 2007; Goodyear, 2003). In a simulation study, Goodyear (2003) compared GLM with the original deterministic formulation of HBS, and concluded that both methods can be accurate, but HBS can be strongly biased when input assumptions were erroneous. Bigelow and Maunder (2007) however, modeled catch rates using statHBS and GLM when applied to real fisheries data and found that statHBS fit best to the data, though the focus of their conclusions was on the importance of considering vertical habitat, rather than depth, in CPUE standardization for HMS. Recognizing the significance of this conclusion, we decided to evaluate not only statHBS, but also GLMs that consider habitat. Furthermore, given that delta-GLMs are a popular approach to CPUE standardization, we thought it was important to evaluate this method of dealing with zeros in the catch data. Thus, this study is the first to compare delta-GLMs and statHBS in a simulation context. In agreement with previous work, we conclude that the incorporation of vertical habitats exploited is important when estimating relative abundance of HMS from fishery-dependent data, specifically when there has been contrast in the habitats exploited over time. We further recommend that this information be included via delta-GLM rather than statHBS, unless there is substantial confidence in estimates of habitats exploited.

Acknowledgments

We thank J. Graves, K. Bigelow, P. Goodyear, A. Punt, T. Miller, J. Brubaker, C. Brown, J. Walter, E. Prince, E. Williams, K. Andrews, A. Buchheister, A. Horodysky, C. Wor, and two anonymous reviewers for providing thoughtful suggestions that vastly improved the quality of this research. Funding was provided by National Oceanic and Atmospheric Administration (NA09OAR4170119) and the Plumeri Award for Faculty Excellence at the College of William & Mary. This is Virginia Institute of Marine Science contribution number 3218. Views expressed are those of the authors, and do not necessarily represent findings or policy of any government agency.

References

- ADMB Project, 2011. AD Model Builder: automatic differentiation model builder. Developed by David Fournier and freely available from www.admb-project.org.
- Aitchison, J., 1955. On the distribution of a positive random variable having a discrete probability mass at the origin. *J. Am. Stat. Assoc.* 50, 901–908.
- Bigelow, K.A., Hampton, J., 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.
- Bigelow, K.A., Maunder, M.N., 2007. Does habitat or depth influence catch rates of pelagic species? *Can. J. Fish. Aquat. Sci.* 64, 1581–1594.
- Bigelow, K., Musyl, M.K., Poisson, F., Kleiber, P., 2006. Pelagic longline gear depth and shoaling. *Fish. Res.* 77, 173–183.
- Bishop, J., 2006. Standardizing fishery-dependent catch and effort data in complex fisheries with technology change. *Rev. Fish Biol. Fish.* 16, 21–38.
- Brill, R., Lutcavage, M., 2001. Understanding environmental influences on movements and depth distributions of tunas and billfishes can significantly improve stock assessments. In: Sedberry, G.R. (Ed.), *Island in the Stream: Oceanography and Fisheries of the Charleston Bump*. American Fisheries Society, Symposium 25, Bethesda, MD, pp. 179–198.

- Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag, New York, NY.
- Campbell, R.A., 2004. CPUE standardization and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fish. Res.* 70, 209–227.
- Carruthers, T.R., Ahrens, R.N.M., McAllister, M.K., Walters, C.J., 2011. Integrating imputation and standardization of catch rate data in the calculation of relative abundance indices. *Fish. Res.* 109, 157–167.
- Carruthers, T.R., McAllister, M.K., Ahrens, R., 2010. Simulating spatial dynamics to evaluate methods of deriving abundance indices for tropical tunas. *Can. J. Fish. Aquat. Sci.* 67, 1409–1427.
- Glaser, S.M., Ye, H., Maunder, M.N., MacCall, A.D., Fogarty, M., Sugihara, G., 2011. Detecting and forecasting complex nonlinear dynamics in spatially structured catch-per-unit-effort time series for North Pacific albacore (*Thunnus alalunga*). *Can. J. Fish. Aquat. Sci.* 68, 400–412.
- Goodyear, C.P., 2003. Tests of the robustness of habitat-standardized abundance indices using simulated blue marlin catch-effort data. *Mar. Freshwater Res.* 54, 369–381.
- Goodyear, C.P., Die, D., Kerstetter, D.W., Olson, D.B., Prince, E.D., Scott, G.P., 2003. Habitat standardization of CPUE indices: research needs. *Col. Vol. Sci. Pap. ICCAT* 55, 613–623.
- Goodyear, C.P., Luo, J., Prince, E.D., Hoolihan, J.P., Snodgrass, D., Orbesen, E.S., Serafy, J.E., 2008. Vertical habitat use of Atlantic blue marlin *Makaira nigricans*: interaction with pelagic longline gear. *Mar. Ecol. Prog. Ser.* 365, 233–245.
- Graves, J.E., Luckhurst, B.E., 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.* 100, 134–142.
- Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics & Uncertainty. Chapman and Hall, London.
- Hinton, M.G., 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 Japanese longline fisheries in the Pacific. *Inter-Am. Trop. Tuna Comm. Bull.* 21, 171–200.
- Hoolihan, J.P., Luo, J., Goodyear, C.P., Orbesen, E.S., Prince, E.D., 2011. Vertical habitat use of sailfin (*Istiophorus platypterus*) in the Atlantic and eastern Pacific, derived from pop-up satellite archival tag data. *Fish. Oceanogr.* 20, 192–205.
- Horodysky, A.Z., Kerstetter, D.W., Latour, R.J., Graves, J.E., 2007. Habitat utilization and vertical movements of white marlin (*Tetrapturus albidus*) released from commercial and recreational fishing gears in the western North Atlantic Ocean: inferences from short duration pop-up archival satellite tags. *Fish. Oceanogr.* 16, 240–256.
- International Commission for the Conservation of Atlantic Tunas (ICCAT), 2004. Report of the 2003 meeting of the ICCAT working group on assessment methods. *Col. Vol. Sci. Pap. ICCAT* 56, 75–105.
- International Commission for the Conservation of Atlantic Tunas (ICCAT), 2006. Report of the 2006 ICCAT billfish stock assessment (Madrid May 15–19, 2006). *Col. Vol. Sci. Pap. ICCAT* 60, 1431–1546.
- Kerstetter, D.W., Luckhurst, B.E., Prince, E.D., Graves, J.E., 2003. Use of pop-up satellite archival tags to demonstrate survival of blue marlin (*Makaira nigricans*) released from pelagic longline gear. *Fish. Bull.* 101, 939–948.
- Langley, A., Bigelow, K., Maunder, M.N., Miyabe, N., 2005. Longline CPUE indices for bigeye and yellowfin in the Pacific Ocean using GLM and statistical habitat standardization methods. Western and Central Pacific Fisheries Commission Science Committee 1 Stock Assessment Working Paper 8, 40 pp.
- Lo, N.C.H., Jacobson, L.D., 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.
- Lynch, P.D., Graves, J.E., Latour, R.J., 2011. Challenges in the assessment and management of highly migratory bycatch species: a case study of the Atlantic marlins. In: Taylor, W.W., Lynch, A.J., Schechter, M.G. (Eds.), Sustainable Fisheries: Multi-Level Approaches to a Global Problem. American Fisheries Society, Bethesda, MD, pp. 197–225.
- Majkowski, J., 2007. Global fishery resources of tuna and tuna-like species. FAO Fisheries Technical Paper 483, Food and Agriculture Organization of the United Nations, Rome.
- Maunder, M.N., Hinton, M.G., Bigelow, K.A., Langley, A.D., 2006. Developing indices of abundance using habitat data in a statistical framework. *Bull. Mar. Sci.* 79, 545–559.
- Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data. A review of recent approaches. *Fish. Res.* 70, 141–159.
- Maunder, M.N., Starr, P.J., 2003. Fitting fisheries models to standardized CPUE abundance indices. *Fish. Res.* 63, 43–50.
- McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models, 2nd ed. Chapman and Hall, London.
- Myers, R.A., Worm, B., 2003. Rapid worldwide depletion of predatory fish communities. *Nature* 423, 280–283.
- Nelder, J.A., Wedderburn, R.W.M., 1972. Generalized linear models. *J. R. Stat. Soc.* 135, 370–384.
- Ortiz, M., 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.
- Ortiz, M., Hoolihan, J.P., 2010. Updated blue marlin (*Makaira nigricans*) standardized catch rates from the U.S. pelagic longline in the northwest Atlantic and Gulf of Mexico 1986–2009. *Col. Vol. Sci. Pap. ICCAT* 66, 1692–1701.
- Prince, E.D., Goodyear, C.P., 2006. Hypoxia-based habitat compression of tropical pelagic fishes. *Fish. Oceanogr.* 15, 451–464.
- Prince, E.D., Luo, J., Goodyear, C.P., Hoolihan, J.P., Snodgrass, D., Orbesen, E.S., Serafy, J.E., Ortiz, M., Schirripa, M.J., 2010. Ocean scale hypoxia-based habitat compression of Atlantic istiophorid billfishes. *Fish. Oceanogr.* 19, 448–462.
- R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/> (accessed August 2011).
- Rice, P.H., Goodyear, C.P., Prince, E.D., Snodgrass, D., Serafy, J.E., 2007. Use of catenary geometry to estimate hook depth during near-surface pelagic longline fishing: theory versus practice. *N. Am. J. Fish. Manage.* 27, 1148–1161.
- Seber, G.A.F., 1982. The Estimation of Animal Abundance and Related Parameters. Oxford University Press, New York.
- Serafy, J.E., Diaz, G.A., Prince, E.D., Orbesen, E.S., Legault, C.M., 2004. Atlantic blue marlin, *Makaira nigricans*, and white marlin, *Tetrapturus albidus*, bycatch of the Japanese longline fishery, 1960–2000. *Mar. Fish. Rev.* 66 (2), 9–20.
- Shono, H., 2005. Is model selection using Akaike's information criterion appropriate for catch per unit effort standardization in large samples? *Fish. Sci.* 71, 978–986.
- Stefánsson, G., 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. *ICES J. Mar. Sci.* 53, 577–588.
- Uozumi, Y., 2003. Historical perspective of global billfish stock assessment. *Mar. Freshwater Res.* 54, 555–565.
- Uozumi, Y., Nakano, H., 1994. A historical review of Japanese longline fishery and billfish catches in the Atlantic Ocean. *Col. Vol. Sci. Pap. ICCAT* 41, 233–243.
- Walters, C., 2003. Folly and fantasy in the analysis of spatial catch rate data. *Can. J. Fish. Aquat. Sci.* 60, 1433–1436.
- Ward, P., Hindmarsh, S., 2007. An overview of historical changes in the fishing gear and practices of pelagic longliners, with particular reference to Japan's Pacific fleet. *Rev. Fish. Biol. Fish.* 17, 501–516.
- Ward, P., Myers, R.A., 2005. A method for 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.
- Ward, P.J., Myers, R.A., 2006. Do habitat models accurately predict the depth distribution of pelagic fishes? *Fish. Oceanogr.* 15, 60–66.
- Wilberg, M.J., Thorson, J.T., Linton, B.C., Berkson, J., 2010. Incorporating time-varying catchability into population dynamic stock assessment models. *Rev. Fish. Sci.* 18, 7–24.
- Ye, Y., Dennis, D., 2009. How reliable are the abundance indices derived from commercial catch-effort standardization? *Can. J. Fish. Aquat. Sci.* 66, 1169–1178.
- Yokawa, K., Takuchi, Y., 2003. Estimation of abundance index of white marlin caught by Japanese longliners in the Atlantic Ocean. *Col. Vol. Sci. Pap. ICCAT* 55, 484–501.
- Yokawa, K., Uozumi, Y., 2001. Analysis of operation pattern of Japanese longliners in the tropical Atlantic and their blue marlin catch. *Col. Vol. Sci. Pap. ICCAT* 53, 318–336.
- Yoshihara, T., 1951. Distribution of fishes caught by the long line. II. Vertical distribution. *Bull. Jpn. Soc. Sci. Fish.* 16, 370–374 (in Japanese).
- Yoshihara, T., 1954. Distribution of catch of tuna longline. IV. On the relation between K and ϕ_0 with a table and diagram. *Bull. Jpn. Soc. Sci. Fish.* 19, 1012–1014 (in Japanese).
- Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, New York.