IOTC-2011-WPB09-INF03

Fisheries Research 106 (2010) 249–260

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Using adaptive area stratification to standardize catch rates with application to North Pacific swordfish (*Xiphias gladius*)

Momoko Ichinokawa[a,*], Jon Brodziak[b]

[a] *National Research Institute of Far Seas Fisheries, Fisheries Research Agency, 5-7-1, Orido, Shimizu-ku, Shizuoka 424-8633, Japan*
[b] *Pacific Islands Fisheries Science Center, NOAA/NMFS, 2570 Dole Street, Honolulu, HI 96822-2326, USA*

ARTICLE INFO

Article history:
Received 4 January 2010
Received in revised form 10 August 2010
Accepted 10 August 2010

Keywords:
Standardizing catch-per-unit effort
Spatial stratification
Highly migratory species
Generalized linear model

ABSTRACT

This paper develops a new method to objectively construct an area stratification for standardizing catch-per-unit effort (CPUE) with generalized linear models (GLMs). This algorithm incorporates the advantages of binary recursion as used in regression trees to minimize a chosen objective function, and extends the concept of stepwise model selection to minimize an appropriate goodness-of-fit criterion for a chosen statistical model, such as GLM. The algorithm can adaptively search for area stratifications that achieved better GLM fits to the CPUE data. The new algorithm, which we call 'GLM-tree', is applied to swordfish CPUE data from Japanese longline vessels in the North Pacific as a case study. The GLM-tree algorithm was conducted with the fishery CPUE data under alternative assumptions about the structural complexity of the GLMs and alternative choices of goodness-of-fit criteria, e.g., Akaike or Bayesian information criteria. Results show that the GLM-tree algorithm created area stratifications more effectively than area stratification determined in an ad hoc manner, and made area stratifications with better fits to swordfish CPUE data until a goodness-of-fit criteria achieved minimum. The algorithm produced many alternative models under different model complexity and area stratifications, which could explain the swordfish CPUE data equally well, because the structural complexity of the GLMs can be compensated by increasing the number of areas. Effects of area stratifications on the estimates of standardized CPUE are also shown to indicate that estimates of the abundance indices tend to converge after a sufficient number of areas have been added.

1. Introduction

To achieve sustainable management of marine fish stocks, it is important to know how the resources have increased, decreased or fluctuated through time in relation to the fishery exploitation pattern. Because the absolute abundance of most fish stocks in oceanic ecosystems cannot be observed directly, relative abundance indices are often used to estimate the current status and historical trends of such stocks (Hilborn and Walters, 1992). Relative abundance indices are typically derived from research survey data or estimated from commercial or recreational fishery catch-per-unit effort (CPUE) data. In general, relative abundance indices for highly migratory tuna and billfish species, which inhabit wide range of ocean, are usually derived from commercial fishery CPUE (Maunder et al., 2006) with a few exceptions, e.g., Southern bluefin tuna (Eveson et al., 1999).

Nominal fishery CPUE is influenced by annual changes of stock abundance and other factors that change catchability (cf. Hilborn and Walters, 1992; Maunder and Punt, 2004; Bishop, 2006). Catchability of a highly migratory predator may vary by geographic region due to oceanographic conditions and the associated densities of prey. For example, consider the situation where the area of a traditional fishing ground is shrinking year by year from a wide range of ocean, which is the main habitat of a highly migratory species, to a narrow coastal area, which is a marginal habitat for the species. In this case, nominal fishery CPUE may decrease year by year even though the actual abundance of the stock has not changed. This situation can easily occur when fishing practices change, and may be especially important for measuring the relative abundance of non-targeted, by-catch species. The possible biases caused by changes in the spatial distribution of fishing effort and associated catchability need to be accounted for when estimating relative abundance indices using fishery CPUE data (Walters, 2003; Campbell, 2004; Ward, 2006).

Many of the impacts of spatiotemporal changes in fishing effort on CPUE, or other factors that affect catchability besides abundance, can be removed by applying statistical methods to standardize CPUE. Methods to standardize CPUE are well developed and are an important component of stock assessments (Hilborn and Walters, 1992; Maunder and Punt, 2004; Bishop, 2006). To standardize

* Corresponding author. Tel.: +81 54 336 6039; fax: +81 543 35 6035.
E-mail address: ichimomo@fra.affrc.go.jp (M. Ichinokawa).

0165-7836/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2010.08.001

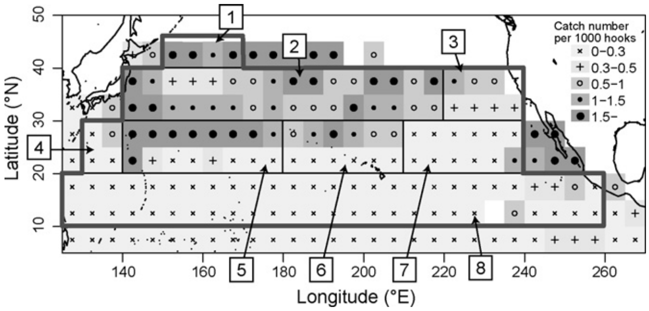

Fig. 1. Average nominal catch rates (numbers per 1000 hooks) of swordfish caught by Japanese longliners from 1975 to 2007 in the North Pacific, and area stratification defined by Nakano (1998). The fishery data in the region surrounded by thick gray lines were used in Nakano (1998) and this study. Several high CPUE areas (>1.5) were not included because of an insufficient time-series of fishing effort. Black solid lines partitioning the region indicate boundaries of the area stratification by Nakano (1998) with serial area labels from 1 to 8 in squares.

CPUE, the response variable (i.e., CPUE) is typically modeled using explanatory variables such as fishing region and season, gear settings, and year in a statistical model. Standardized abundance indices can be derived from the estimated year effects after the impacts of other predictors have been statistically removed to the extent possible. Generalized linear models (GLMs) are commonly used for standardizing CPUE (Gavaris, 1980; Kimura, 1981). Other statistical approaches for standardizing CPUE have been developed, including generalized additive models (GAMs, cf. Bigelow et al., 1999; Venables and Dichmont, 2004) and generalized linear mixed models (GLMMs; cf. Rodríuez-Marín et al., 2003; Venables and Dichmont, 2004). In this context, the definition of factor levels that affect CPUE, which can be created by dividing continuous explanatory variables into categorical ones, is an important step for applying statistical models such as GLMs, GAMs or GLMMs to standardize CPUE (Bishop, 2006).

Spatial stratification of the fishing area into spatial strata is particularly important for standardizing CPUE. Ideally, an appropriate spatial stratum is a region in which fish density is homogeneous and CPUE is influenced by explanatory variables in a similar manner (Bishop, 2006). However, many studies to standardize CPUE appear to determine spatial strata in an ad hoc manner. That is, spatial strata are often defined based on either the spatial distribution of fishing effort and nominal CPUE or oceanographic conditions, with the exception of a few studies using regression trees (i.e., Watters and Deriso, 2000; Walsh and Kleiber, 2001). Selecting appropriate area strata is especially important for standardizing CPUE of highly migratory species such as tunas and billfishes because fishery independent abundance indices are rarely available for these species. Further, the broad geographic ranges of tuna and billfish stocks suggest that spatial heterogeneity of catchability is likely to be a key factor for CPUE standardization.

Swordfish (*Xiphias gladius*) is a large pelagic billfish found in tropical, temperate, and subarctic waters (Nakamura, 1985). While swordfish is a cosmopolitan species, the distribution of swordfish CPUE caught by Japanese longliners in the North Pacific exhibits substantial spatial heterogeneity. The highest swordfish CPUE occurs in the northwest Pacific between 25°N and 45°N and in coastal waters off California, while the lowest CPUE has been observed in tropical waters (Uozumi and Uosaki, 1998; Nakano, 1998; Fig. 1). Nakano (1998) defined a total of eight oceanic strata for standardizing CPUE of swordfish in the North Pacific based on the heterogeneity of the distribution of average nominal CPUE (Fig. 1). Nakano found that there was a significant interaction between year and area effects when standardizing swordfish CPUE with a GLM. This implied that swordfish abundance trends differed among the 8 areas. Kleiber and Bartoo (1998) also found significant differences between swordfish abundance trends in the northwest Pacific (west of the International Date Line) and the north-central Pacific (International Date Line to 130°W, equator to 20°N). While both of these studies found that spatial heterogeneity was a predominant feature of the abundance trends of the swordfish population in the North Pacific, their results were derived from different assumptions about the appropriate spatial stratification used for standardizing CPUE. In particular, both studies selected a single spatial stratification a priori, and did not consider other possible scenarios of area stratification. Consequently, it was not possible to evaluate whether their choices of area strata were appropriate and sufficient to standardize swordfish CPUE.

This study develops a new method to objectively construct an area stratification for standardizing CPUE with GLM. This algorithm adaptively searches for the area stratification that produces the best GLM fit to the CPUE data. We apply the new algorithm to swordfish CPUE data from Japanese longline vessels in the North Pacific as a case study. The case study illustrates the relative effectiveness of the algorithm and shows how alternative assumptions about the structural complexity of the GLMs and alternative choices of goodness-of-fit criteria can affect the results of CPUE standardization. Empirical relationships among the number of areas, model complexity, and goodness-of-fit criteria are investigated to determine their relative influence on estimates of standardized CPUE.

2. Methods

2.1. Algorithm for area partitioning

To construct an area stratification that provided the best fit to the observed CPUE data, we used a binary recursive partitioning algorithm similar to what is used to construct regression trees (Breiman et al., 1984). Regression trees have been applied to select area strata for standardizing CPUE of bigeye tuna (Watters and Deriso, 2000) and blue sharks (Walsh and Kleiber, 2001). In these analyses, the objective function to be minimized was the total sum of squared differences between average and observed CPUE over strata created by the regression tree algorithm (Breiman et al., 1984). However, the objective function to be minimized in this study was a goodness-of-fit measure for the fit of the statistical model, such as a GLM, that was used to standardize CPUE.

The area-partitioning algorithm, which we call 'GLM-tree', was applied to partition the fishery area using the following three steps. First, the algorithm divided the current spatial domain into all possible pairs of strata, assuming a fixed spatial resolution defined by a set of regularly spaced dividing lines. This was the binary partition step. Second, the chosen statistical model was applied to fit the fishery CPUE data under each of the possible stratifications. This was the model fitting step. Third, a goodness-of-fit criterion measuring the fit of the statistical model to the observed CPUE data was used to select the stratification that produced the best fit over the set of possible stratifications. This was the optimization step for the current spatial domain. This three-step procedure was repeated recursively until the goodness-of-fit criterion could not be improved by an increase in the number of areas.

To illustrate the GLM-tree algorithm, consider a hypothetical case where the fishery area is the rectangular region extending horizontally from 140°E to 160°E and extending vertically from 10°N to 20°N in the North Pacific with a spatial resolution of 5° × 5° (Fig. 2). This spatial domain can be denoted as the rectangle R[140:160, 10:20] where the first coordinate indicates longitudinal extent and the second coordinate indicates latitudinal extent. In the first step of the GLM-tree algorithm, there are four possible pairs of strata that can be created by adding one new boundary with a 5° mesh; these are: (I) R[140:145, 10:20] and R[145:160, 10:20], (II) R[140:150, 10:20] and R[150:160, 10:20], (III) R[140:155, 10:20]

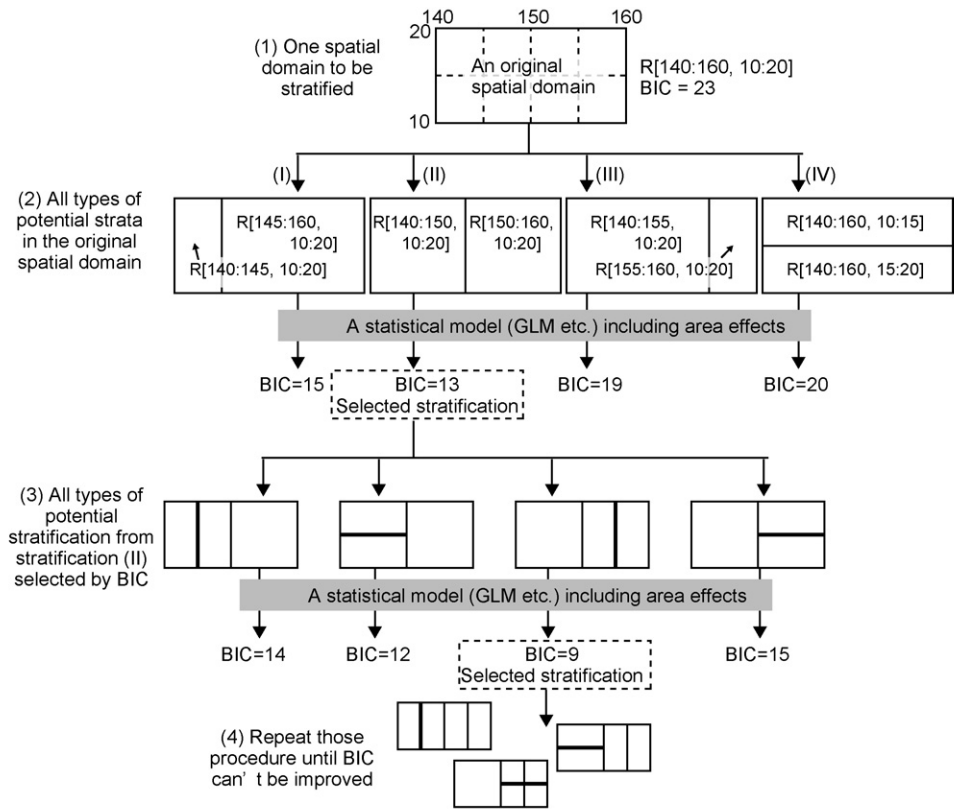

Fig. 2. Schematic diagram to show how the GLM-tree algorithm recursively constructs an area stratification for a given spatial domain. The notation R[140:160, 10:20] represents the rectangular region that extends from 140°E to 160°E and from 10°N to 20°N.

and R[155:160, 10:20], and (IV) R[140:160, 10:15] and R[140:160, 15:20]. In the second step, a GLM that includes the variable area was applied to standardize the CPUE data in the spatial domain using the four possible area stratifications. In the third step, a goodness-of-fit measure, e.g., Bayesian Information Criterion (BIC; Schwarz, 1978), was applied to select the stratification that produced the best fit to the CPUE data. For example, if the BIC values for the four stratifications of (I), (II), (III) and (IV), respectively, were 15, 13, 19 and 20, and the BIC value without area stratification was 23, then stratification (II) with BIC = 13 would be selected as the optimal stratification. In the next recursion, the stratification (II) would be partitioned using same three steps to construct a set of four possible stratifications, each consisting of 3 areas. Of these stratifications, the one with minimum BIC for the fitted GLM (e.g., BIC = 9 in Fig. 2) would be selected for the next step of the algorithm.

The GLM-tree algorithm requires a goodness-of-fit criterion to measure how well the statistical model can explain the observed CPUE. In this study, Akaike's Information Criteria (AIC; Akaike, 1973) and BIC were used as the goodness-of-fit criteria to be minimized. The AIC is calculated from model deviance (D) evaluated at the maximum likelihood estimate and a parameter penalty which depends on the number of model parameters (p) (Eq. (1)):

$$\mathrm{AIC} = D + 2p \tag{1}$$

For BIC, the parameter penalty increases with both the number of parameters and the sample size (n) of data points used to fit the model (Eq. (2)):

$$\mathrm{BIC} = D + \log(n)p \tag{2}$$

Because these two goodness-of-fit criteria are based on different underlying concepts of consistency (Burnham and Anderson, 2002), this study compared the performance of both AIC and BIC for creating area stratifications to standardize CPUE.

Another important difference between the GLM-tree and the regression tree algorithms was the recursive procedure to create partitions. For the regression tree algorithm, the recursive procedure was applied to each regional subgroup in order to 'grow' the tree for each subgroup independently. In contrast, for the GLM-tree algorithm, the recursive procedure was applied to all subgroups to simultaneously select a single boundary from the set of candidate stratifications produced by adding one boundary. By definition, the objective function for the regression tree algorithm was the total sum of squares: this function can be minimized by separately minimizing its individual components calculated for each subgroup. On the other hand, the objective function for the GLM-tree algorithm was a model-based goodness-of-fit criterion that cannot be separately calculated for each sub-group by definition. As a result, an increase in the number of areas within a subgroup can influence both the estimated model parameters and the goodness-of-fit of the model to the fishery data in the GLM-tree algorithm.

In summary, the GLM-tree algorithm was designed to increase the number of areas one at a time from the set of all candidate partitions. Computer code to implement the GLM-tree algorithm was written in the R language (R Development Core Team, 2008), and is available from the lead author by request. Estimation of model parameters for the GLMs was carried out using the biglm module (Lumley, 2006).

2.2. Case study: Japanese longline CPUE of swordfish in the North Pacific

The GLM-tree algorithm was applied to Japanese longline CPUE of swordfish in the North Pacific as a case study. Swordfish CPUE indices have typically been derived from analyses conducted by the Billfish Working Group of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. In one of these analyses, Nakano (1998) subjectively defined eight area strata for swordfish CPUE standardization (Fig. 1) based on visual inspection of the spatial distribution of average nominal CPUE of swordfish and fishing effort. In this study, we used the same overall spatial domain as Nakano (1998).

The data sets used in this study were also similar with those used in Nakano (1998), but the fishery data were updated until 2007, and sorted out to eliminate outliers. The fishery catch and effort data were collected from Japanese longline vessels and compiled by National Research Institute of Far Seas Fisheries (NRIFSF), Fisheries Agency of Japan. These data included species-specific catches of tunas, billfishes, and other bycatch species such as sharks, as well as details of fishing operations (e.g., number of hooks, gear configurations, and locations). Although Japanese longline data are reported by fishing captains in logbooks by each operation, NRIFSF typically aggregates the catch and effort data by $5^\circ \times 5^\circ$ area, month and hooks per basket (HPB) in order to average-out the influence of highly variable logbook reports and to keep the size of data sets manageable. As a result, both this study and Nakano (1998) used the aggregated longline data. The time period was from 1975 to 2007 in this study. Each aggregated area-month-HPB cell was required to have at least 10 fishing operations in order to provide a representative sample. Therefore, the cells with fewer than 10 operations were excluded from the analyses in this study although Nakano (1998) used all data.

A listing of the number of longline operations used in this study, by region, HPB category and year, is given in Table 1. Shallow longline sets consisting of 3–4 HPB occurred in the northwest region of the North Pacific (areas 1, 2 and 4–6) during the entire time period of 1975–2007, while shallow sets were rarely observed in the northeast (areas 3 and 7) or subtropical (area 8) regions. Moderate-depth longline sets consisting of 5–9 HPB during the 1970s and early 1980s were generally replaced by deeper sets consisting of >9 HPB after the late 1980s in all regions. This historical shift in the depth of fishing effort occurred worldwide (cf. Yokawa and Uozumi, 2001) due to a shift of targeting to bigeye tuna in tropical waters as well as other changes in the evolution of gear configurations used by Japanese longliners (Yokawa and Uozumi, 2001; Ward and Hindmarsh, 2007). Because the effects of year and HPB were confounded by area, GLMs that included both interaction terms of year and area, and year and HPB were not considered in the same model when applying stepwise model selection for choosing the form of the GLM for standardizing CPUE.

Standardization of swordfish CPUE (CPUE_{yijk}, where catch had units of number of swordfish per 1000 hooks, "y" indexed year, "i" indexed gear configuration, "j" indexed quarter, and "k" indexed area) was carried out using GLMs with a lognormal error distribution. Explanatory variables for the GLMs included year (Y), gear configuration of HPB (G), quarter (Q) and area (A) along with some interaction terms (Eq. (3)):

$$\log\left(\mathrm{CPUE}_{yijk} + \frac{\mu}{10}\right) = Y_y + G_i + Q_j + A_k + (\text{interaction terms}) \quad (3)$$

A constant term $\mu/10$ was added to the observed CPUE_{yijk} in order to rescale the value of zero catches (Campbell et al., 1996), where the value of μ was the overall average of nominal swordfish CPUE throughout the time series ($\mu = 0.78$). Although this approach for treating zero catch data may lead to biased predictions of CPUE in some circumstances (cf. Maunder and Punt, 2004), it did not have an important influence in this study because the percentages of zero catches were relatively low in the swordfish data sets (overall average of 8% zero catches with an average of 2–16% by area). CPUE predictors other than area were categorized a priori. The effect of gear configurations (G) was categorized into five classes of 3–4, 5–6, 7–9, 10–14 and 15–20 HPB. These categories were determined from GAM analyses using gam packages in R (Hastie, 2008) to investigate the nonlinear relationship between HPB and CPUE. The seasonal effect on CPUE (Q) was also categorized into four quarters: January–March (1st quarter), April–June (2nd quarter), July–September (3rd quarter) and October–December (4th quarter).

Potential interaction terms to be included in the GLMs were selected in a preliminary stepwise analysis using the area stratification from Nakano (1998). In this stepwise model selection analysis, BIC was used to eliminate 2nd order interaction terms one by one from the most complex model including all 2-order interaction terms except for both year and area (Y^*A) or gear and area (G^*A) due to confounding effects described above. Model selection was also evaluated using AIC for comparison, although those results were not emphasized because AIC tended to select the most complex model due to the large number of CPUE observations (Shono, 2005).

For GLMs that included the interaction term of Y^*A, the abundance index in year y (SCPUE_y) was calculated from the area weighting factors for each area indexed by a (f_a, $a = 1, 2, \ldots, A$) and from the abundance indices in year y and area a (SCPUE_{ya}) using the following equation (Campbell, 2004):

$$\mathrm{SCPUE}_y = \sum_{a=1}^{A} f_a \cdot \mathrm{SCPUE}_{ya}, \quad \text{where} \quad \sum_{a} f_a = 1 \quad (4)$$

The parameter SCPUE_{ya} was estimated from the least squares mean (a.k.a., population marginal mean) of CPUE in year y in area a (Searle et al., 1980) in the nominal scale and the area weighting factor (f_a). The factor was defined as the ratio of the number of $5^\circ \times 5^\circ$ blocks in area a to the total number of $5^\circ \times 5^\circ$ blocks considered. SCPUE_y was also estimated for models without the Y^*A term as the least squares mean in year y. The least squares means were originally estimated in the logarithmic scale, and converted to SCPUE_y or SCPUE_{ya} in the nominal scale by exponentiating the estimated least squares mean and adding the constant term. For example, SCPUE_y is calculated with $\exp(\mathrm{LCPUE}_y + \sigma_y^2/2) - \mu/10$, where the parameter LCPUE_y was least squares means in year y in the logarithmic scale and σ_y was the standard deviation of the estimated LCPUE_y. The value of SCPUE_y or SCPUE_{ya} was not evaluated when the corresponding least squares mean could not be calculated due to missing data. Least squares means were calculated using SAS (ver. 9.1 for SunOS 5.9 platform).

Swordfish abundance indices were also calculated from standardized CPUE using the area stratification from Nakano (1998). However, unlike Nakano (1998), which did not exclude strata having fewer than 10 operations, the least square means for the year effect could not be calculated for the model including the Q^*A term due to missing effort in the Area 3 and third quarter cell. In this case, the abundance indices were approximated by assuming the coefficients in those strata with missing data was the average of the other coefficients in the interaction term. The resulting approximate index was found to differ from the original index by less than 1 percent across years.

In order to assess the robustness of the calculated area stratifications, the GLM-tree algorithm was applied to 50 bootstrap data sets, which were constructed from the original CPUE data set by randomly re-sampling CPUE observations with replacement. In particular, the bootstrap re-sampling of observed CPUE was conducted for each year, quarter and HPB cell.

3. Results

Step-wise model selection using BIC and the area stratification defined by Nakano (1998; Fig. 1) selected the interaction terms of A^*Q, Q^*G and A^*G (Model III-A) when starting from the full model excluding Y^*A (Table 2). Similarly the interactions of A^*Q, Q^*G and Y^*A were selected (Model III-B) when starting from the full model

Table 1

Number of longline operations recorded in the data used for the swordfish CPUE standardization by year, region and gear configuration of hooks per basket (HPB). The region numbers follow the definition used in Nakano (1998, Fig. 1). Data for which the number of operations per each aggregated area, month and HPB cell were less than 10 were excluded from this analysis and this table.

	Northwest subarctic (Areas 1and 2)					Northwest temperate (Areas 4 to 6)					Northeast (Areas 3 and 7)					Subtropical (Area 8)				
	HPB 3–4	HPB 5–6	HPB 7–9	HPB 10–14	HPB 15–20	HPB 3–4	HPB 5–6	HPB 7–9	HPB 10–14	HPB 15–20	HPB 3–4	HPB 5–6	HPB 7–9	HPB 10–14	HPB 15–20	HPB 3–4	HPB 5–6	HPB 7–9	HPB 10–14	HPB 15–20
1975	779	4562	2097	177	0	862	2955	996	285	0	0	1338	150	0	0	527	8095	1223	1260	0
1976	858	8804	4830	719	0	791	5074	1749	835	0	0	810	111	32	0	807	8048	2201	1937	0
1977	991	7089	3686	1093	0	833	6495	3149	2123	0	0	611	677	46	0	264	7634	845	2954	12
1978	618	6962	2022	2217	0	744	5614	1931	1694	0	0	892	1559	795	0	0	6585	1796	5394	0
1979	651	7625	2464	3648	0	693	4586	1539	2381	20	0	795	1231	671	0	124	5197	861	4370	0
1980	357	6644	1993	5063	28	883	4835	1276	2586	34	0	345	384	397	0	206	3841	734	5114	0
1981	566	6751	1985	9013	0	1086	5096	1334	4646	23	0	1027	710	316	0	72	2703	817	7156	0
1982	681	5397	1739	8421	0	715	3967	896	4606	14	0	1141	730	289	0	218	1811	257	6217	0
1983	2333	4650	1966	9048	0	1682	1945	698	3801	0	0	376	1210	608	0	0	923	246	5977	15
1984	1331	4318	3700	11222	25	2405	1743	563	2673	55	0	122	1707	661	39	0	646	188	7214	19
1985	3120	3023	3412	9442	76	2624	1896	315	1862	10	0	0	713	267	0	0	239	144	5575	11
1986	5096	1901	1542	8475	328	3175	968	302	3448	172	0	0	1387	1465	13	0	282	132	8259	831
1987	4543	2009	2313	7021	297	4227	502	216	1560	328	0	0	1089	554	32	0	384	78	8459	1583
1988	4443	1399	2923	7635	466	2232	408	265	1170	283	0	11	781	992	0	0	391	95	12524	1368
1989	5641	505	3021	5990	604	2653	166	55	1868	211	0	19	2651	1584	0	0	536	154	9974	984
1990	4602	359	2513	4029	862	2498	23	289	943	304	0	0	928	879	0	0	219	0	5901	435
1991	4404	397	3053	5262	1282	2270	53	81	1091	598	0	16	1474	833	38	76	139	265	7479	2409
1992	4154	24	2656	2842	625	2858	15	283	962	526	0	31	1300	2591	0	0	13	15	3304	1842
1993	4536	29	2816	3293	1119	2995	0	342	1397	539	0	0	1093	1861	93	0	0	17	5357	2067
1994	3771	118	980	3120	1866	2277	42	90	970	587	0	0	840	2345	32	17	0	0	5140	1379
1995	3980	81	644	2285	1785	2353	14	36	697	718	0	41	721	1791	263	0	77	22	5171	2931
1996	4006	123	571	1931	925	2038	0	57	719	836	0	0	979	765	13	0	0	10	3425	1998
1997	3859	133	674	1552	1265	1625	0	95	586	1090	0	0	246	661	95	0	19	10	1899	1337
1998	3986	0	272	1150	1341	1642	0	78	442	818	0	0	468	860	110	0	0	49	1326	2640
1999	4671	0	220	1240	1688	1150	0	0	187	1465	0	0	369	1458	1707	0	0	0	1707	2898
2000	5841	0	62	394	1188	602	0	0	765	982	0	0	118	630	536	0	0	23	1411	2673
2001	5427	69	132	885	686	433	0	29	438	745	0	0	0	331	243	0	0	0	1316	3500
2002	4323	0	0	423	1053	868	0	24	490	741	0	0	0	597	952	0	0	0	853	1973
2003	4040	0	0	676	1669	371	0	0	160	740	0	0	0	557	2645	0	0	0	1249	3653
2004	3930	92	0	299	1559	503	0	0	259	424	0	0	0	122	457	0	0	0	975	2532
2005	3258	156	0	290	1606	330	0	0	280	653	0	0	0	202	2349	0	0	0	797	3028
2006	2965	183	0	359	2035	320	12	0	103	825	0	0	0	59	1175	0	0	0	657	3468
2007	2676	383	0	301	1819	314	10	0	44	645	0	0	0	145	444	0	0	0	547	2161

Table 2
ANOVA table for the model selected by BIC under the assumption with 8 area strata by Nakano (1998).

(a) The model including the effect of G^*A (Model III-A)

	DF	Sum of square	Mean square	*F* value	Pr > *F*
Model	103	27578.77	267.755	487.52	<.0001
Error	29425	16160.76	0.54922		
Corrected Total	29528	43739.52			

R-Square = 0.631, BIC = 66,210

Effects	DF	Type III SS	Mean square	*F* value	Pr > *F*
Y (year)	32	347	11	19.8	<.0001
A (area)	7	1503	215	391.0	<.0001
Q (quarter)	3	17	6	10.5	<.0001
G (gear)	4	418	105	190.4	<.0001
A^*G	25	1879	75	136.9	<.0001
A^*Q	20	1515	76	138.0	<.0001
Q^*G	12	253	21	38.4	<.0001

(b) The model including the effect of Y^*A (Model III-B)

	DF	Sum of square	Mean square	*F* value	Pr > *F*
Model	296	27324.54	92.31263	164.39	<.0001
Error	29232	16414.98	0.56154		
Corrected Total	29528	43739.52			

R-Square = 0.625, BIC=67,050

Effects	DF	Type III SS	Mean square	*F* value	Pr
Y(year)	32	116	4	6.4	<.0001
A (area)	7	1390	199	353.5	<.0001
Q (quarter)	3	14	5	8.4	<.0001
G (gear)	4	5634	1409	2508.4	<.0001
Y*A	218	1625	7	13.3	<.0001
A*Q	20	1851	93	164.8	<.0001
Q*G	12	224	19	33.3	<.0001

excluding A^*G. Both Models III-A and III-B were subsequently analyzed using the GLM-tree algorithm. In addition, several simpler models including (a) no interaction terms (Model 0), (b) the single interaction term of A^*Q (Model I) and (c) the two interaction terms of A^*Q and Q^*G (Model II) were considered for comparison with the selected models.

Comparisons of the area stratification produced by the GLM-tree algorithm using the simplest model (Model 0) versus the most complex models (Models III-B and III-A) showed that making structural assumptions about the number of interaction terms included in the GLM can produce different area stratifications (Fig. 3). The GLM-tree algorithm using the three different GLM structures selected the same boundary at 25°N in the 1st iteration (Fig. 3a). However, the selected partition with Model 0 was different from those selected using Models III-A and III-B in the 2nd iteration (Fig. 3b). In particular, the second boundary was selected at 210°E from 25 to 40°N for Model 0, but at 30°N for Models III-A and III-B.

Differences in the area stratifications produced by the alternative models increased as the number of steps increased. By the 7th partitioning step, there were three distinct sets of 8 areas produced by the alternative model structures (Fig. 3c) that each differed considerably from the area stratification proposed by Nakano (1998). Further, the three area stratifications created using the GLM-tree algorithm produced substantially better fits to the CPUE data than the one used by Nakano (1998). In particular, BIC values for the area stratifications using the GLM-tree algorithm were 64,148 for Model III-A and 66,176 for Model III-B, while those derived from Nakano's stratification with the same data set were 66,210 for Model III-A and 67,050 for Model III-B. The BIC differences of over 1000 units indicated that the GLM-tree algorithm produced much more plausible model fits and provided very strong evidence that the area stratifications produced by the algorithm were significantly better (Kass and Raftery, 1995).

As expected, the optimal number of defined areas differed for the two goodness-of-fit criteria (Fig. 3d and e). Using the AIC criteria, 67 areas were selected using the simplest model with no interaction terms (Model 0), while 30 and 31 areas were selected for Models III-A and III-B, respectively. In comparison, the number of areas selected using the BIC criteria were much lower, being 42, 24 and 14 under Models 0, III-A and III-B, respectively.

The two goodness-of-fit criteria showed different patterns of model refinement as the number of strata increased (Fig. 4). For example, the relative trajectory of BIC values under Model III-B (Fig. 4b) differed from the trajectory of AIC values (Fig. 4a); this difference resulted from the parameter penalty and the large number of Y^*A parameters estimated for this model (Fig. 4e). Similarly, for both the AIC and BIC criteria, the resulting spatial structure for Model 0 with no interactions is seen to be more complex than those with interaction terms, indicating that the interactions account for some of the spatial heterogeneity in the CPUE data.

Our results indicate that alternative model structures can explain the CPUE data equally well (Fig. 4). This is because increasing the number of areas can compensate for the structural complexity of the GLMs. For example, AIC values of less 65,000 were achieved by each of the Models I, II, III-A and III-B when assuming more than 11, 10, 5, and 6 areas, respectively (Fig. 4a). However, as AIC or BIC, improved with increasing number of strata, the number of alternative models with similar goodness-of-fit values declined. In this case study, the minimum AIC of 58,140 and minimum BIC of 60,880 was achieved solely by Model III-B with 31 strata and by Model III-A with 24 strata, respectively, among all the alternative models and stratifications examined. The structural complexity of the best fitting model selected using AIC or BIC indicated the importance of including interaction terms when evaluating area stratifications, because structurally simpler models cannot explain the CPUE data as well as more complex ones as the number of area strata increase. In particular, the incorporation of the A^*Q, and Y^*A or Y^*G terms had a substantial impact on overall model fit in this case study.

It is also noteworthy that the model fits with a total of 8 areas produced by the GLM-tree algorithm were always better than that using the area stratification defined by Nakano (1998) for each model structure. For example, there was roughly a 5% difference in the amount of explained variance (Fig. 4d) between the models using the 8 spatial strata defined by Nakano (1998) and those using strata created with the GLM-tree algorithm. This result suggests that the algorithm is a robust and effective approach for creating adaptive area stratifications under alternative modeling assumptions.

The historical swordfish abundance indices estimated from the standardized CPUE exhibited roughly similar patterns across model structures and area stratifications, with peaks of abundances occurring in the mid-1970s, mid-1980s, and post-2000 (Fig. 5a), though each index was influenced by the structural model assumptions and the area stratification used (Fig. 5b). Indices for models with a small number of area strata, especially those with only 2–4 areas, tended to diverge from the indices based on models with a larger number of areas. On the other hand, indices for models with 5 or more areas were generally more similar, especially for Models 0, I and II. It is also noteworthy that indices based on the Nakano stratification (which incorporated 8 areas) were more similar to indices produced by the GLM-tree algorithm with only 2–4 areas. Consequently, the abundance indices based on Nakano's stratification were different from the corresponding GLM-tree based indices

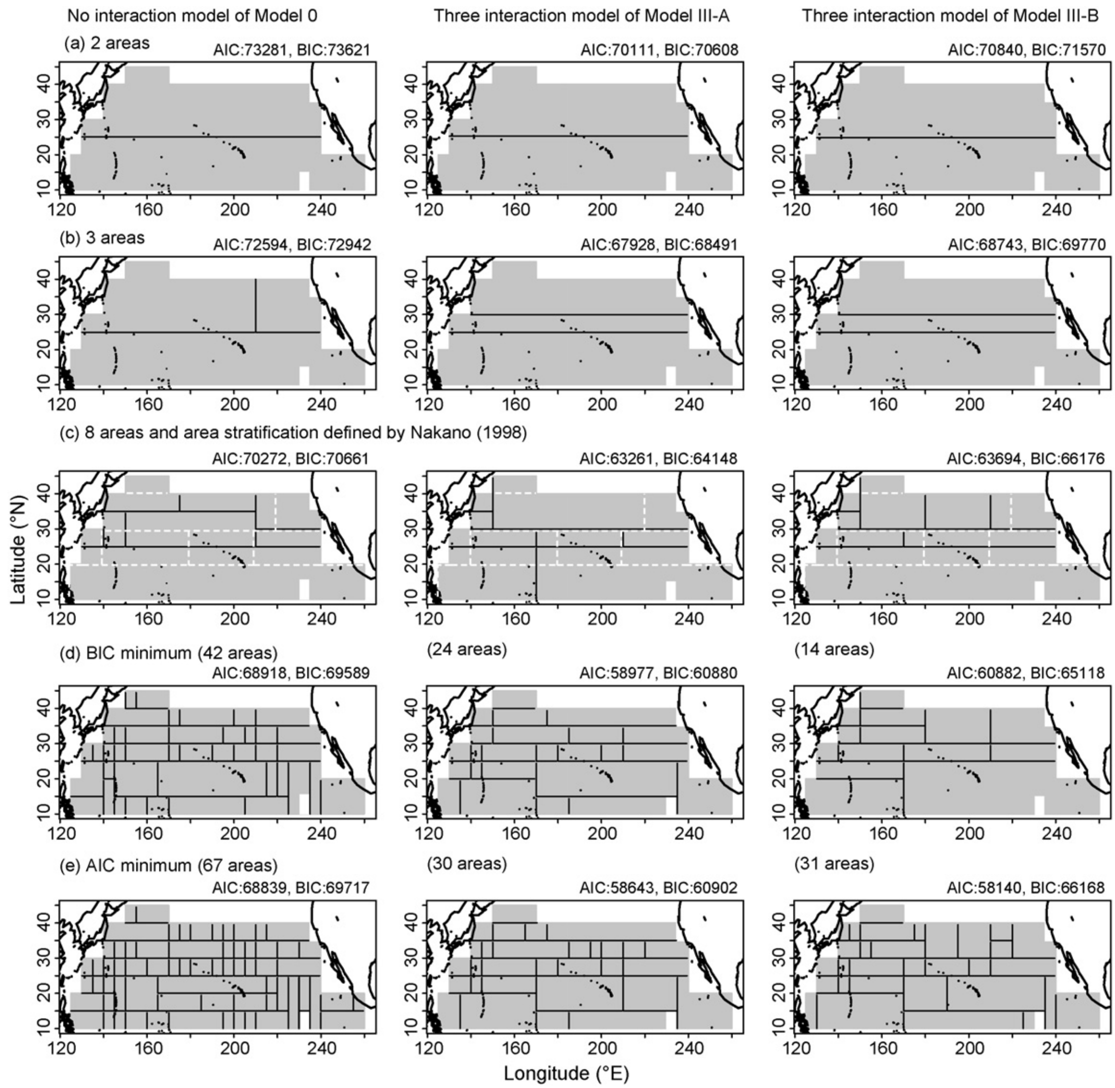

Fig. 3. Examples of the increasing number of areas created by the boundaries (solid black lines) selected by the GLM-tree algorithm with difference assumptions regarding model structures (Models II, III-A and III-B). White broken lines in (c) indicate the area stratification defined by Nakano (1998). The gray shaded region, same as the region surrounded by thick gray lines with fishing effort in Fig. 1, represents the actual spatial region to be stratified.

with the same number of areas. In addition, the GLM-tree indices with >5 areas displayed varying degrees of differences between models. For example, the estimated abundance indices using Model 0 were higher in the mid-1980s than those estimated using Models I and II, while the Model 0 estimates after 2000 tended to be smaller than those using Models I and II. In some cases, models with >5 areas were not estimable because of missing data. In particular, a convergent time trend was not observed for Model III-A because least squares means could not be calculated under this model structure with more than 6 areas. In addition, a convergent trend for Model III-B could not observed because the amount of CPUE data was insufficient to estimate the Y^*A interactions in some strata.

The divergence of the estimated abundance indices was found to be especially significant in the initial steps of the GLM-tree algorithm, particularly from the 1st to the 3rd iterations in Models 0, I and II (Fig. 6). The indices differed by around 2–4% on an annual basis during these initial steps. During subsequent iterations with approximately >5 areas, the estimated abundance indices converged to similar values, with the differences in the average abundance indices generally being less than 1%. This result suggests that the estimated indices tended to converge after sufficient number of areas have been created by the GLM-tree algorithm. On the other hand, observed differences between successive estimates of indices under Model III-B (which included the Y^*A term) were found to be larger than those for the other models, and suggests that the indices estimated for the model with the Y^*A term may take more iterations to converge.

The optimal number of areas determined by the GLM-tree algorithm was sensitive to the observed CPUE data used for model fitting, the chosen goodness-of-fit measures and the chosen model structure (Fig. 7). In particular, the optimal number of areas differed by each bootstrap data set even when the same model structure was assumed. While there was a general tendency for more complex models to have a lower optimal number of areas, the optimal number of areas also varied by goodness-of-fit criterion. For example, under Model III-B, the optimal number of strata in the bootstrap data sets ranged from 22 to 46 with an average of 35 using AIC, and ranged from 16 to 26 with an average of 20 using BIC. This suggested that the optimal number of areas produced using the BIC criterion will be lower than the number produced using AIC.

The bootstrap analysis also showed the potential variation of the area stratification boundaries produced by the GLM-tree algorithm (Fig. 8). A large variety of boundaries were produced from the 50

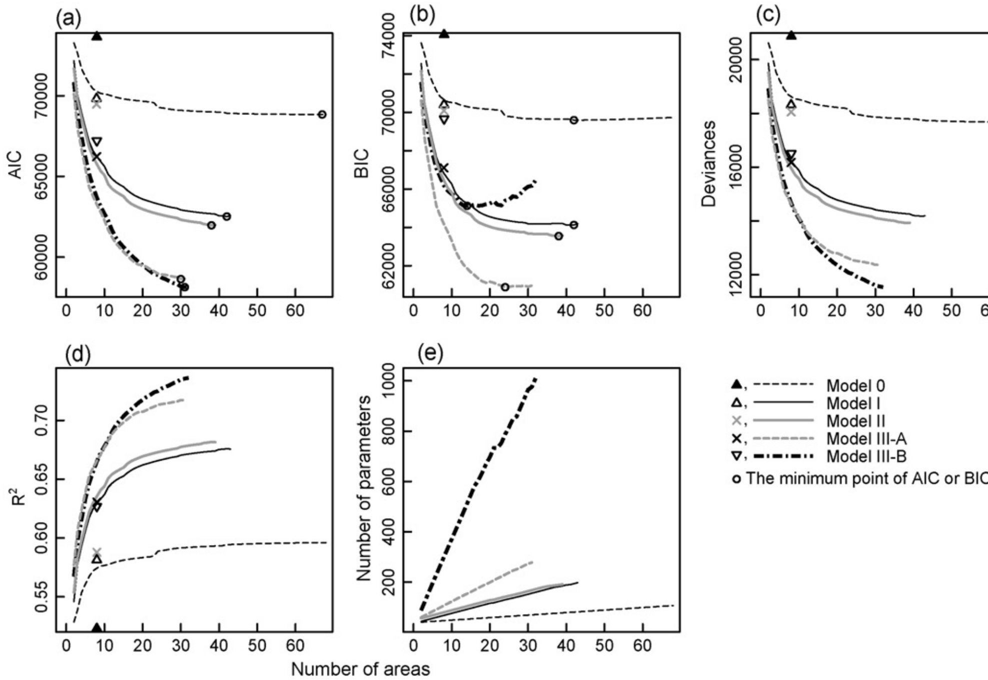

Fig. 4. Trajectories of (a) AIC, (b) BIC, (c) model deviances, (d) *R*-squares and (e) number of estimated parameters versus the number of areas. The trajectories were produced by the GLM-tree algorithm until AIC could not be improved by increasing the number of areas. The values of AIC, BIC, model deviances and *R*-squares estimated with 8 areas defined by Nakano (1998) are also shown by symbols of triangles or crosswise crosses. Circles on the trajectories show the minimum points of AIC and BIC.

bootstrap data sets fitted under Models III-A and III-B (Fig. 8, right panels), and the boundaries observed in the all of bootstrap replicates (black thick lines in Fig. 8) were a subset of the total number of boundaries that might have been created with the GLM-tree algorithm. In particular, the total length of the boundaries selected in all of the bootstrap replicates was approximately 1/3 of the total length of all of the fitted boundaries (Fig. 8, left panels). While the boundaries selected in all of the bootstrap replicates suggest the existence of a robust envelope of boundaries that was not sensitive to sampling variation, roughly half of the boundaries selected were highly dependent on the bootstrap realization. Interestingly, the latitudinal boundaries created with the GLM-tree algorithm seemed to be more robust than the longitudinal boundaries in this study. This pattern likely reflects persistent latitudinal variation in the spatial distribution of the swordfish population in the North Pacific.

4. Discussion

This study demonstrated the effectiveness of the GLM-tree algorithm to create area stratifications that produce better fits to CPUE data (Fig. 4) and also showed how different area stratifications affected estimates of abundance indices as a function of model complexity (Figs. 5 and 6). The trajectories of the goodness-of-fit measures as a function of the number of areas suggested that several alternative models can produce similar fits to the observed CPUE data by compensating model complexity with the number of areas assumed (Fig. 4). However, the goodness-of-fit values of the simpler models that lacked appropriate interaction terms could not be substantially improved by increasing the number of areas. Thus, the GLM-tree algorithm cannot compensate for an over-simplified model structure. This indicated that the selection of an appropriate model structure was very important for obtaining a good fit to the CPUE data. At the same time, comparing the estimates of abundance indices produced under various assumptions about area stratification with a sufficient number of strata (Fig. 6) suggested that strict optimization for adaptive area stratification until AIC or BIC minimum may not always be needed to derive robust estimates of abundance indices, from a practical point of view.

The new computer-intensive approach presented in this study exhibited good performance and was more effective than the previous ad hoc approach for choosing area strata for standardizing swordfish CPUE. This was evident in the smaller goodness-of-fit measures calculated from the model assuming 8 areas created by the GLM-tree algorithm compared to those assuming the ad hoc area stratification determined by Nakano (1998) (Fig. 4). In addition, the abundance indices estimated with the ad hoc area stratification diverged from the trends estimated by the GLM-tree algorithm with a sufficient number of strata, and were similar to those with a poor spatial structure based on only 2–4 areas (Fig. 5). This indicates that Nakano's area stratification was not adequate to explain the spatial structure in the data. In contrast, the GLM-tree algorithm was able to effectively identify area stratifications in order to estimate abundance indices that were relatively robust for further increases in the number of areas.

Regression trees have been previously applied to select area stratifications for analyzing longline catch data for bigeye tuna (Watters and Deriso, 2000) and blue shark (Walsh and Kleiber, 2001). Our method incorporated the advantages of the recursive-binary algorithm used in regression trees to automatically minimize a chosen objective function for stratifying spatial data. The method in this study also extended the concept of stepwise model selection to minimize an appropriate goodness-of-fit criterion for a chosen statistical model, such as GLM, to standardize CPUE. In general, the GLM-tree algorithm can be adapted to apply to any likelihood-based statistical model or goodness-of-fit criterion. Furthermore, although we do not develop the ideas in this study, it seems clear that other categorized factors used for standardizing CPUE, such as HPB and month which were categorized in an ad hoc manner in this study, could be adaptively stratified with a generalized version of this algorithm.

There are a few aspects of our method that might be improved in the future. The GLM-tree algorithm tends to subdivide areas

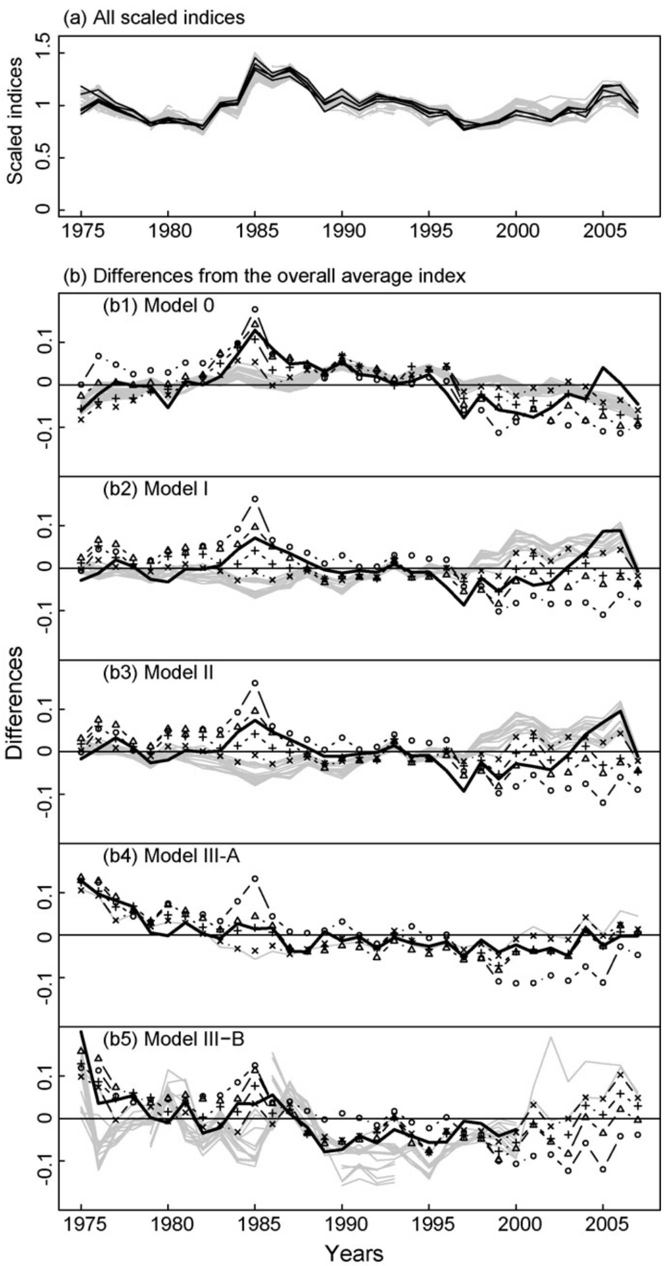

Fig. 5. (a) Estimated swordfish abundance index for each model scaled by the respective average for each index, where GLM-tree based indices are shown by gray lines, and those assuming the area stratification by Nakano (1998) are shown by black lines; (b) the difference between each index and the overall average index across all models, where gray lines indicate models with more than 5 areas, lines with symbols indicate models with 5 or fewer areas, and black lines indicate models based on the area stratification of Nakano (1998). The symbols of circles, triangles, crosses and crosswise crosses with lines show estimates from the models assuming 2, 3, 4 and 5 areas, respectively. (b1) Model 0, (b2) Model I, (b3) Model II, (b4) Model III-A and (b5) Model III-B. The number of series of the abundance indices shown in this figure are 66, 19, 19, 5 and 20 for Models 0, I, II, III-A and III-B, respectively, where least squares means could be calculated. Those indices are derived from area stratification produced by the GLM-tree algorithm with AIC criteria.

with a substantial amount of fishery data because the total model deviance is the sum of the deviance estimated for each data point. In order to overcome this potential problem, it may be possible to use a weighted deviance in the goodness-of-fit criterion where the weights correspond to the inverse of the number of data points in each cell. In this way, all year, quarter and area cells would have equal weightings in the calculation of the deviance function. In practice, creating areas with inadequate amounts of effort data to calculate least squares means should also be avoided, for example, by adding a penalty term to the objective function. The overall tendency of the GLM-tree algorithm to subdivide areas with a substantial amount of data may be advantageous to avoid creating strata with missing data. Nevertheless, in some cases, it will not be possible to calculate the least squares means as the number of areas increases due to a prevalence of strata with low sample sizes; this was especially true for Model III-A (Fig. 6), for example.

The implementation of the GLM-tree algorithm used in this study was programmed to create rectangular strata. However, the actual physical biogeographic provinces that affect the distribution and catchability of oceanic fishery resources are not always rectangular (cf. Longhurst, 1998). Thus, allowing more flexible shapes for constructing the area strata to represent relevant biogeographical and persistent oceanographic features would likely provide better fits to the swordfish longline fishery data. Given this, the 'optimal stratifications' estimated in this case study may not the best attainable because numerous alternative area stratifications constructed with more flexible shapes exist, which may be able to explain the CPUE data as well as or better than an optimal rectangular partitioning. In addition, the best model and selected boundaries were observed to be somewhat variable under the bootstrap resampling of the CPUE data (Figs. 7 and 8). This variability suggested that outliers or future updates of the fishery data would likely alter the best estimate of area stratification for swordfish in future GLM-tree analyses.

It was impractical to investigate alternative shapes in this case study due to the vast number of potential candidate shapes and the logistical constraints of limited time and computer resources. For example, in the case study with a sample size of $n = 29{,}527$ CPUE observations, the convergence of the GLM-tree algorithm required the evaluation of 5000 GLMs by 'biglm' packages in R to create approximately 60 areas under Model 0, which took about an hour of CPU time (Intel(R) Core(TM)2 CPU 6700 @ 2.66 GHz, 8GB DDR2-SDRAM). In comparison, it took about 15–23 h under Model III-B while creating roughly 20–30 areas with 2000–3000 evaluations of GLMs. Increasing the number of estimated model parameters and increasing the sample sizes can limit the tractability of the GLM-tree algorithm. In addition, the fact that some fishery data are specifically aggregated by rectangular grid would also make it difficult to use non-rectangular shapes for area stratification.

The concordance of estimates of swordfish abundance indices after some sufficient number of iterations of the GLM-tree algorithm (Figs. 5 and 6) suggested that the calculated abundance indices were robust and converged as the number of strata increased. In other words, even though abundance indices from some models and spatial stratifications achieving the minimum goodness-of-fit criteria could often not be calculated due to missing data, swordfish abundance indices could be reasonably well approximated by one or more models with some sufficient number of areas. However, the abundance indices estimated using Model III-B, which included the interaction between year and area effects, seemed to be more sensitive to the choice of area stratification than the other models even after a relatively large number of iterations. Including the Y^*A term in GLM was based on the assumption that variability of CPUE at the local spatial scale and during a specific year was representative of the abundance in that region. Therefore, fine-scale stratification of a large spatial region tended to match up with the local temporal changes in catch rates in the model that includes the Y^*A term while models without the Y^*A term tended to treat local spatial changes of CPUE as observation error. Determining how much area stratification will affect the estimated abundance indices will in general depend on characteristics of the fishery data used. However, localized variability of catch rates at small spatial scales have been observed not only in swordfish, but

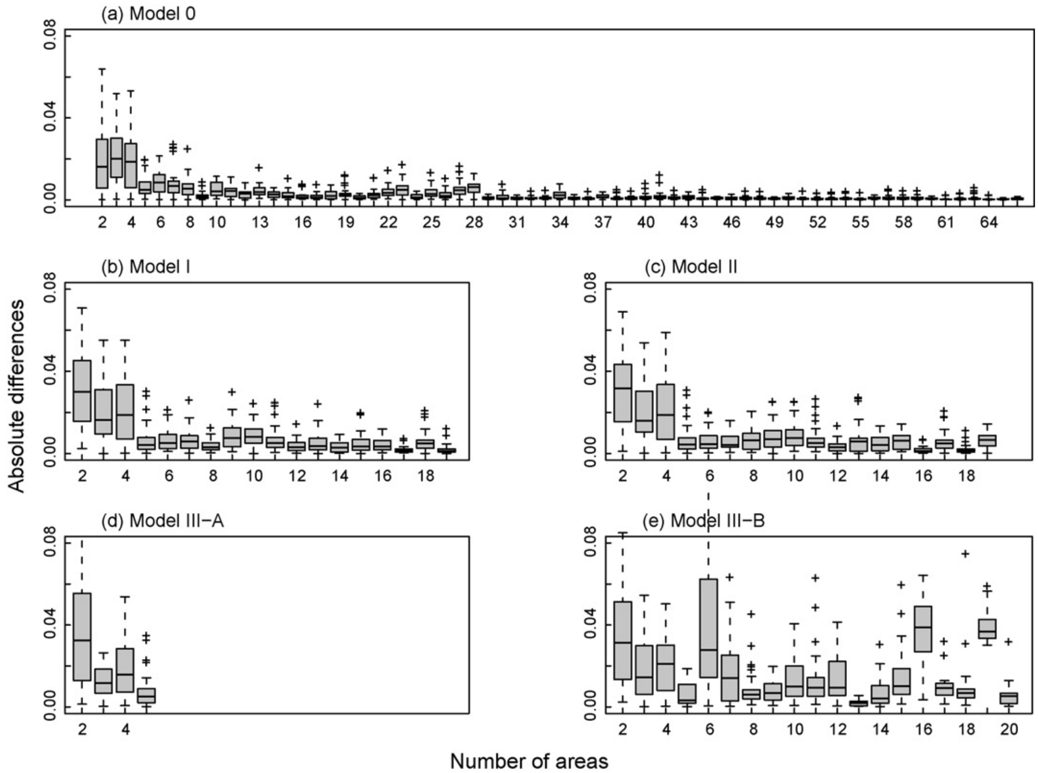

Fig. 6. Boxplot of the absolute difference of the scaled abundance index when successively adding one additional area stratum by the GLM-tree algorithm with AIC criteria. Each plot shows the distribution of the difference in each year. Outliers more than 0.08 are not shown.

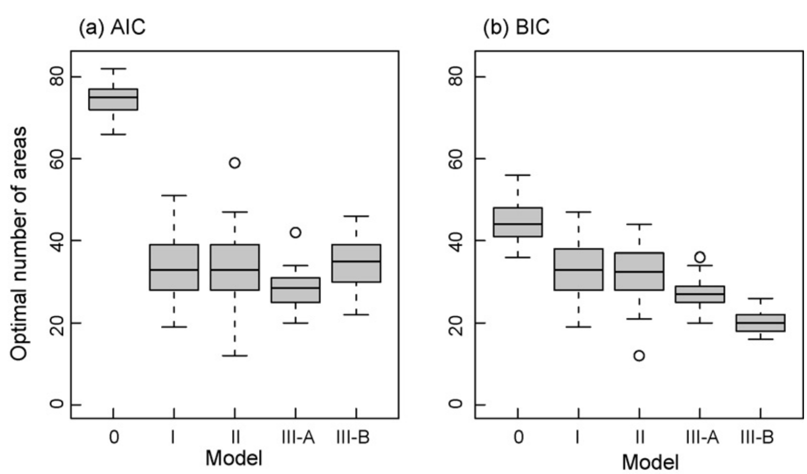

Fig. 7. Boxplot of the optimal number of area strata that archieved minimum (a) AIC and (b) BIC calculated with 50 bootstrap data sets.

also in many other fish populations (cf. Punt et al., 2000). Thus, the inclusion of the *Y**A* term seemed to have an important effect on the estimates of abundance indices.

Adaptive area stratification to standardize CPUE becomes more important when the goal is to estimate annual trends by region for a spatially structured stock assessment model in order to evaluate local depletion or other effects of spatial structure on relative abundance trends (Punt et al., 2000). If substantial differences in abundance trends exist among the defined area strata, which could be caused by temporal changes in migration or spatio-temporal differences in fishing effort, these patterns should be detected by selecting an appropriate area stratification. In general, analyses to standardize CPUE do not typically include a rigorous analysis of the appropriateness of an area stratification. Area stratifications determined in an ad hoc manner may cause misinterpretation of historical abundance trends and potential population structures because the assumed spatial structure is insufficient to standardize the CPUE data. In contrast, this study represents a first step to determining area stratification in an objective and systematic manner, while allowing for the flexibility to adapt to a selected objective function and model structure. In particular, the GLM-tree algorithm appears to be very useful for estimating abundance indices of species exhibiting heterogeneous spatio-temporal distributions. Future work will focus on investigating how to incorporate more detailed information on the biology of the target species and on the relevant oceanographic features of the fishing grounds.

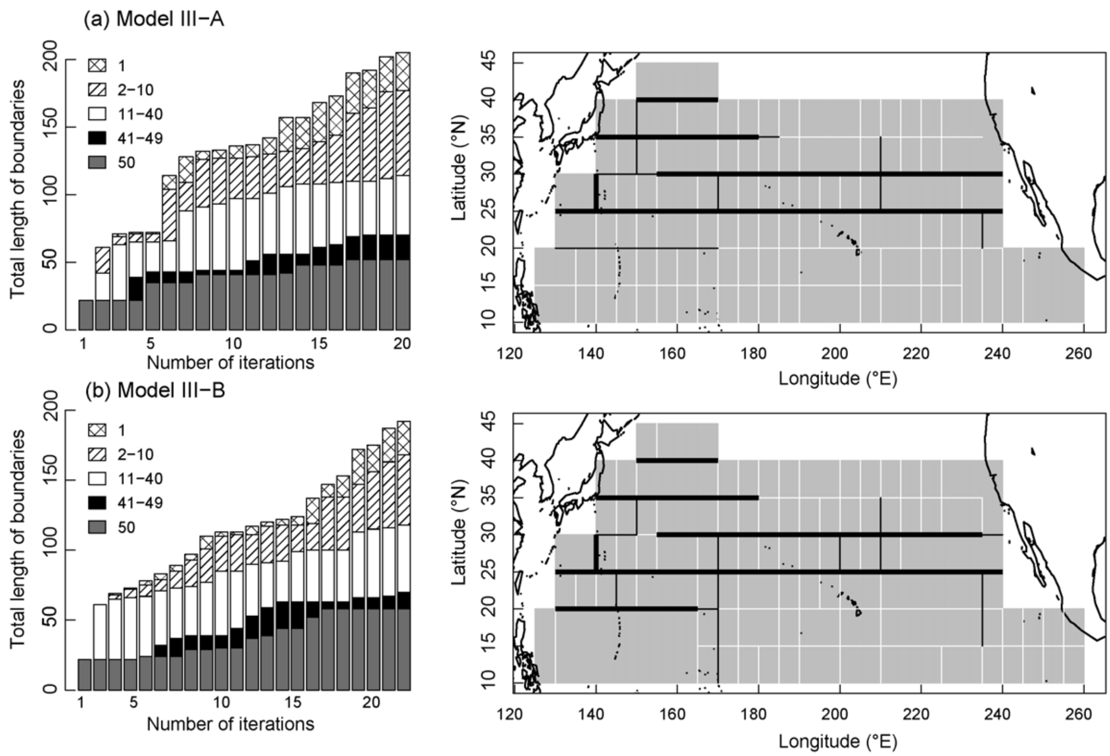

Fig. 8. Variability of the boundaries selected by the GLM-tree algorithm with 50 bootstrap data sets. Left panels show total length of boundaries selected in bootstrap replicates by the number of times selected as a boundaries. Unit of the length of boundaries is 1 side of $5^\circ \times 5^\circ$ square. Right panels show examples of the boundaries observed in every (black thick lines), 40–49 (black thin lines) and 0–39 (white lines) replicates after 20 iterations in Models III-A (a) and 22 iterations in Model III-B (b).

Acknowledgements

This study has been funded by the Fisheries Research Agency and National Marine Fisheries Service. Our colleagues in NRIFSF, Y. Takeuchi and H. Shono are thanked for their helpful discussion and statistical advices. K. Yokawa is thanked for encouraging this study and giving useful advices on fisheries targeting swordfish. Thanks are extended to the ISC Billfish Working Group members and chairman Dr. G. DiNardo, who encouraged this study and provided useful advices and discussion, and an anonymous reviewer whose helpful comments improved this manuscript.

References

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F.s. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest.

Bigelow, K.A., Boggs, C.H., He, X., 1999. Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. Fish Oceanogr. 8, 178–198.

Bishop, J., 2006. Standardizing fishery-dependent catch and effort data in complex fisheries with technology change. Rev. Fish Biol. Fish 16, 21–38.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Chapman & Hall/CRC.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-theoretic Approach. Springer, New York.

Campbell, R.A., Tuck, G., Tsuji, S., Nishida, T., 1996. Indices of abundance for southern bluefin tuna from analysis of fine-scale catch and effort data. In: Working Paper SBFWS/96/16 Presented at the Second CCSBT Scientific Meeting, Hobart, Australia.

Campbell, R.A., 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fish Res. 70, 209–227.

Eveson, P., Farley, J., Bravington, M., 1999. The Aerial Survey Index of Abundance: Updated Analysis Methods and Results. CCSBT-ESC/0909/12, Busan, Korea, pp. 34.

Gavaris, S., 1980. Use of a multiplicative model to estimate catch rate and effort from commercial data. Can. J. Fish Aquat. Sci. 37, 2272–2275.

Hastie, T., 2008. gam: Generalized Additive Models. R Package Version 1.0.

Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman and Hall, New York.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.

Kimura, D.K., 1981. Standardized measures of relative abundance based on modeling log(c.p.u.e.), and their application to Pacific ocean Perch (*Sebastes-Alutus*). J. Du Conseil 39, 211–218.

Kleiber, P., Bartoo, N., 1998. Standardizing swordfish, *Xiphias gladius*, longline catch per unit of effort using general additive models. NOAA Tech. Rep. NMFS 142, 181–193.

Longhurst, A.R., 1998. Ecological Geography of the Sea. Academic Press.

Lumley, T., 2005. biglm: Bounded Memory Linear and Generalized Linear Models. R package version 0.4.

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., Sibert, J.R., Fonteneau, A., Hampton, J., Kleiber, P., Harley, S.J., 2006. Interpreting catch per unit effort data to assess the status of individual stocks and communities. ICES J. Mar. Sci. 63, 1373–1385.

Nakamura, I., 1985. FAO species catalogue. vol. 15. Billfishes of the world: an annotated and illustrated catalogue of marlins, sailfishes, spearfishes and swordfishes known to date. FAO Fish Syno., 125.

Nakano, H., 1998. Stock status of Pacific swordfish, *Xiphias gladius*, inferred from CPUE of the Japanese longline fleet standardized using general linear model. NOAA Tech. Rep. NMFS 142, 195–209.

Punt, A.E., Walker, T.I., Taylor, B.L., Pribac, F., 2000. Standardization of catch and effort data in a spatially-structured shark fishery. Fish Res. 45, 129–145.

R-Development-Core-Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Australia.

Rodríuez-Marín, E., Arrizabalaga, H., Ortiz, M., Rdríuez-Cabello, C., Moreno, G., Kell, L.T., 2003. Standardization of bluefin tuna, *Thunnus thynuus*, catch per unit effort in the baitboat fishery of the bay of Biscay (Eastern Atlantic). ICES J. Mar. Sci. 60, 1216–1231.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Searle, S.R., Speed, F.M., Milliken, G.A., 1980. Population marginal means in the linear-model—an alternative to least-squares means. Am. Stat. 34, 216–221.

Shono, H., 2005. Is model selection using Akaike's information criterion appropriate for catch per unit effort standardization in large samples? Fish Sci. 71, 976–984.

Uozumi, Y., Uosaki, K., 1998. Review of the Japanese swordfish, *Xiphias gladius*, fisheries in the Pacific Ocean. NOAA Tech. Rep. NMFS 142, 133–146.

Venables, W.N., Dichmont, C.M., 2004. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. Fish Res. 70, 319–337.

Walsh, W.A., Kleiber, P., 2001. Generalized additive model and regression tree analyses of blue shark (*Prionace glauca*) catch rates by the Hawaii-based commercial longline fishery. Fish Res. 53, 115–131.

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, R.D., 2006. The importance of identifying spatial population structure in restocking and stock enhancement programmes. Fish Res. 80, 9–18.
Watters, G., Deriso, R., 2000. Catchs per unit of effort of bigeye tuna: a new analysis with regression trees and simulated annealing. Bull. Int. Am. Trop. Tuna Commun. 21, 531–552.
Yokawa, K., Uozumi, Y., 2001. Analysis of operation pattern of Japanese longliners in the tropical Atlantic and their blue marlin catch. In: Collective Volume of Scientific Papers, ICCAT, vol. 53, pp. 318–336.