Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery

Xi He [a,*], Keith A. Bigelow [a], Christofer H. Boggs [b]

[a] *Pelagic Fisheries Research Program, Joint Institute for Marine and Atmospheric Research, School of Ocean and Earth Science and Technology, University of Hawaii, 2570 Dole Street, Honolulu, HI 96822, USA*
[b] *Honolulu Laboratory, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2570 Dole Street, Honolulu, HI 96822, USA*

Abstract

In the Hawaii-based longline fishery, changes in fishing operations to target different species produce changes in the effectiveness of fishing effort units. Catch-per-unit-effort (CPUE) indices used in resource monitoring were improved by segregating dissimilar types of fishing effort. Cluster analysis was used to classify longline sets in relation to species composition of the catches. Based on proportions of eight species and three broader species groups in 46 961 longline sets from 4 years (1991–1994) of commercial fishery data, five effort clusters were identified. Spatial distribution of sets and differences in fishing operations among clusters were then compared to reveal apparent differences in fishing strategies. Three clusters comprised ~ 80% of the total sets, and the catch compositions suggested targeting for either broadbill swordfish (two clusters) or bigeye tuna. The other two clusters were most similar to the tuna cluster, but their catch compositions indicated a mixed-species fishing strategy. Fishing operations were most different between sets in the tuna and swordfish clusters. Swordfish sets were characterized by (1) the largest vessels, (2) the least number of hooks per set, (3) the greatest number of lightsticks, (4) the longest set duration, (5) the highest percentage of night sets, (6) a larger percentage of sets within the full moon phase, and (7) the lowest percentage of sets within the main Hawaiian Islands Exclusive Economic Zone. Time series of CPUE for three species (bigeye tuna, yellowfin tuna, and swordfish) based on different clusters were compared, and the most appropriate CPUE time series for resource monitoring are recommended. © 1997 Elsevier Science B.V.

Keywords: Cluster analysis; Bigeye tuna; Hawaii; Longline fishery; Spatial distribution of effort; Swordfish

1. Introduction

Fishery scientists rely heavily on catch-per-unit-effort (CPUE) data to analyze trends in population abundance (Hilborn and Walters, 1992) or availability of highly mobile populations to localized fisheries (He and Boggs, 1996). Fishery-independent estimates for tuna and billfish populations are virtually non-existent. In the Hawaii-based US longline fishery, catch and effort data from commercial fisheries provide the only available indicators of resource trends.

The usefulness of CPUE as a relative index depends on many assumptions including that the unit

* Corresponding author at: Cat Cove Marine Lab, 92 Fort Ave., Salem, MA 01970, USA.

measure of effort is uniformly effective for a given area, season, and set of environmental conditions. However, Hawaii's longline fishermen frequently switch target species, altering their operations and gear configuration to exploit the changing availability or marketability of different resources efficiently. The effectiveness with which the unit measure of effort (1000 hooks) catches different species is altered by the choice of fishing strategy.

Various models have been used to account for variation in effective effort by fishing strategies. For example, effort can be corrected for depth factors affecting longline gear capture efficiency (Suzuki, 1989), or gear depth can be included along with other factors in a generalized linear model (GLM) of longline CPUE (Punsly and Nakano, 1992). When critical aspects of fishing strategy are unknown, cluster analyses have been used to categorize subsets of fishing effort (e.g. trawl tows, trawling trips, fishing vessels) by their similarity in catch composition (Rogers and Pikitch, 1992; Lewy and Vinther, 1994; Ward et al., 1996). Clusters with low effective fishing effort for particular species can be segregated or eliminated when analyzing resource trends (Lewy and Vinther, 1994; Ward et al., 1996).

A domestic longline fishery has been operating in the Hawaii Exclusive Economic Zone (EEZ) since the late 1920s (Otsu, 1954; Shomura, 1959; Boggs and Ito, 1993). Traditionally, the fishery primarily targeted tuna species, including bigeye (*Thunnus obesus*), yellowfin (*Thunnus albacares*), and albacore tuna (*Thunnus alalunga*). Istiophorids (mostly marlins) also comprised a large fraction of the traditional catch. Fishing operations were conducted during daylight by setting hooks without lightsticks at depths ranging from the surface to 400 m. The fishery has undergone dramatic changes, expanding from 37 active vessels in 1987 to 140 by 1991. Swordfish (*Xiphias gladius*) became a major target species and vessels explored new fishing areas within and beyond the Hawaii EEZ (Pooley, 1993; Dollar, 1994; DiNardo and Kwok, 1996). Vessels targeting swordfish operated at night by setting hooks with lightsticks near the surface (0–100 m). Some vessels frequently switched target species and fishing techniques between fishing trips or between sets within a fishing trip.

The purposes of this study were to categorize individual fishing sets using cluster analysis based on similarity of catch compositions, examine the spatial distribution of effort and operational information to indicate differences in fishing strategies among clusters, and provide refined CPUE time series for three species (bigeye tuna, yellowfin tuna, and swordfish).

2. Methods

Detailed descriptions of the Hawaii-based longline fishery can be found in Shomura (1959), and Boggs and Ito (1993). Depending on fishing location, a typical longline trip including traveling and fishing time may last from 12 to 30 days. An average of 8.8 longline sets per trip (range 5–15 sets) were conducted from 1991 to 1994. Each longline operation can take as long as 24 h to set (deploy) and retrieve the main line, to which branch lines and hooks are attached. Often the end of the mainline deployed first is retrieved last so that individual hooks may fish anywhere from a few hours to 24 h.

Data from Hawaii-based longline fishermen were obtained from the mandatory logbook programme of the National Marine Fisheries Service (Dollar and Yoshimoto, 1991). Since October 1990, all longline vessels fishing in the Hawaii EEZ or landing fish in Hawaii are required to report their daily catches to the NMFS monitoring programme after completion of each fishing trip. Logbook data include vessel identification, fishing locations, starting times of setting and retrieval, number of hooks deployed, number of fish caught by species, and trip type categorized as tuna, swordfish, or mixed. The latter category is often missing. Upon receiving the logbook data, the NMFS monitoring programme staff checks and edits the reports, and fills in missing trip type information based on a variety of subjective criteria (Ito et al., 1996). Data on the species targeted with individual sets are not included and data on gear configuration needed to estimate set depths were not included until 1995.

A total of 47 344 longline sets were reported from 1991 to 1994. All clustering and comparisons of spatial distribution and operations were based on pooled (1991–1994) data. Some records (1.4%) in the logbook data, such as sets which caught no fish

and incomplete records of number of hooks, were deleted prior to the analysis. Cluster analyses were conducted on 46 961 sets. Twenty species or species groups were listed in the longline reports (Table 1). On the basis of preliminary analysis some species were infrequently caught or poorly identified. These species were combined as three species groups resulting in 11 types of catch (Table 1). Because each record consisted of catch data from each set, a matrix of 46 961 × 11 was used in the cluster analysis. Catch composition from each set was calculated as the proportion of each type of catch to the total catch of the set. Data were arcsin–square-root transformed before analysis to normalize their distribution (Snedecor and Cochran, 1980).

Clusters were developed using two steps because the large number of records precluded a direct hierarchical cluster analysis on the whole data set. First, a non-hierarchical cluster analysis (K-mean method) was used to group all records into 2500 clusters using the FASTCLUS procedure of the Statistical Analysis Systems (SAS) Institute Inc. program (Johnson and Wichern, 1988; Statistical Analysis

Systems Institute Inc., 1989). The FASTCLUS procedure was also used to identify and delete 203 outliers with very different catch compositions. These outliers usually consisted of sets with very low catches (one to three fish) of one or two species. Second, a hierarchical cluster analysis (Ward method, Ward, 1963; Statistical Analysis Systems Institute Inc., 1989) was applied to the 2500 non-hierarchical clusters. The choice for the number of clusters to produce was largely subjective. At least two clusters (tuna sets and swordfish sets) were expected. More than two clusters were produced to allow other possible categories to emerge. Additional clusters were considered until the smallest cluster contained less than 10% of the total number of sets.

After the cluster analysis, catch compositions (mean percentages of the eight species and three species groups) were calculated for each cluster and compared among clusters. Seasonal and spatial distribution of sets and fishing operation characteristics for each cluster were summarized. Spatial distribution of sets was represented as the number of sets per one square nautical degree and as the percentage of

Table 1
Species or species groups used in this study and corresponding species caught by the Hawaii-based longline fishery from 1991 to 1994

Species/group used	Species caught and scientific name
Albacore	*Thunnus alalunga*
Bigeye tuna	*Thunnus obesus*
Blue marlin	*Makaira mazara*
Blue shark	*Prionace glauca*
Mahimahi	*Coryphaena hippurus*
Other billfishes	Black marlin *Makaira indica*
	Sailfish *Istiophorus platypterus*
	Shortnose spearfish *Tetrapturus angustirostris*
Other fishes	Moonfish *Lampris guttatus*
	Northern bluefin tuna *Thunnus thynnus*
	Oilfish and Escolar [a] *Ruvettus pretiosus* and *Lepidocybium flavobrunneum*
	Skipjack tuna *Katsuwonus pelamis*
	Wahoo *Acanthocybium solandri*
	Unidentified fishes
Other sharks	Mako sharks *Isurus* spp.
	Thresher sharks *Alopias* spp.
	Unidentified sharks
Striped marlin	*Tetrapturus audax*
Swordfish	*Xiphias gladius*
Yellowfin tuna	*Thunnus albacares*

[a] These species are not distinguished by Hawaiian fishermen.

sets within the main Hawaiian Islands (MHI) EEZ. Spatial correspondence of sets between clusters was compared using the spatial correlation coefficient (Schneider, 1994). The r value, which ranges from −1 to 1, is calculated in an analogous manner as a correlation coefficient between any two variables. A higher r value indicates more correspondence in spatial distribution of sets among clusters. Maps showing the spatial distribution of effort excluded any one square nautical degree containing data from fewer than three individual vessels to preserve the confidentiality of fishing areas exploited by few individuals.

Characteristics of fishing operations included vessel length, number of hooks per set, number of lightsticks per set, duration of set, and the diurnal and lunar periodicity of fishing effort. Diurnal periodicity was characterized as day, day–night, or night based on the interval between the reported times when setting began and gear retrieval began. Day fishing occurred when this interval occurred mostly during the day, night fishing when it occurred mostly at night, and day–night when the interval between the initiation of setting and retrieval exceeded 17 h. These categories were not completely descriptive of the fishing period because gear retrieval may require many hours during which some of the gear continues to fish. Lunar periodicity was characterized in relation to the moon phase. Each set occurred within one of three equal time subsets within the lunar cycle (29.5 days): the 9.8 days centered around the new moon, the 9.8 days centered on the full moon, and the remaining two intervals centered around the first and third quarters.

The CPUE (number of fish caught per 1000 hooks) of bigeye tuna, yellowfin tuna, and swordfish at a monthly scale from 1991 to 1994 was calculated and compared among clusters, sets of clusters, and subsets of clusters. Time series of CPUE were compared in terms of relative temporal changes using Pearson product–moment correlation coefficients (Snedecor and Cochran, 1980). Because we were interested in temporal changes instead of absolute CPUE values, both x and y time series were transformed by differencing at a lag of 12 months. This also removed seasonal changes in the time series. A higher correlation coefficient (R) represents closer relationships between two time series.

3. Results

3.1. Cluster analysis and fishing strategies

Five clusters of longline sets were produced by the cluster analysis (Fig. 1). Based on the euclidean distance between cluster centroids, Clusters 1, 2, and 3 were closely related. These three clusters of sets caught high percentages of bigeye tuna (Table 2). Clusters 4 and 5 were closely related because these sets caught a large number of swordfish (Table 2). Cluster 1 had the highest percentages of bigeye and yellowfin tuna but also a high percentage of swordfish. Although trips targeting both tuna and swordfish (mixed trips) constitute a major sector of the local fishery, data from Cluster 1 indicate that both types of fish may be targeted by individual sets. Cluster 2 had high catches of bigeye tuna (24%), other fishes (17.5%), and striped marlin (12.6%), and these sets appeared to have specifically targeted tuna. Cluster 3, like Cluster 1, had an even mixture of tuna and swordfish but was dominated by catches of mahimahi (*Coryphaena hippurus*, 43.3%). A fishing strategy to target mahimahi seems unusual and has not previously been recognized in the local fishery. Clusters 4 and 5 had swordfish catches of 31.1% and 59.0%, respectively, and were clearly comprised of sets that primarily targeted swordfish. Composition of the blue shark catch was much larger in Cluster 4
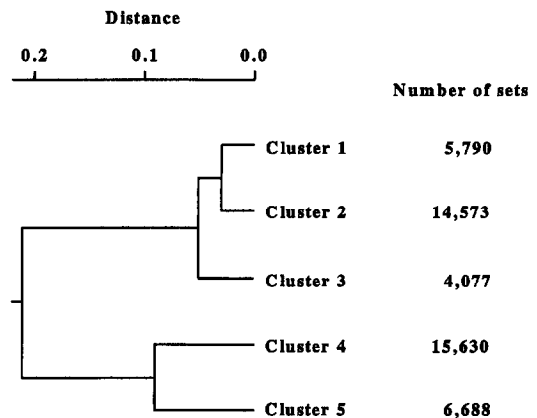

Fig. 1. Dendrogram of five clusters of longline sets from the Hawaii-based longline fishery showing Euclidian distance between clusters.

Table 2
Mean percentages of catches for eight species and three species groups within five clusters of sets from the Hawaii-based longline fishery (1991–1994)

Species group	Cluster				
	1	2	3	4	5
Albacore	2.6	9.4	1.4	6.0	4.5
Bigeye tuna	36.1	24.0	15.4	5.0	7.8
Blue marlin	2.5	3.8	2.2	1.0	2.7
Blue shark	7.4	9.3	11.1	45.1	0.3
Mahimahi	3.9	9.1	43.3	5.2	12.4
Other billfishes	2.9	4.9	3.2	0.6	2.0
Other fishes	5.3	17.9	4.0	1.2	2.2
Other sharks	3.3	2.8	0.9	1.0	0.8
Striped marlin	5.4	12.6	6.3	1.8	4.2
Swordfish	15.2	1.0	9.0	31.1	59.2
Yellowfin tuna	15.0	4.9	3.0	1.9	3.9

(45.1%) than in Cluster 5 (0.3%), but the shark catch was probably unintentional.

The seasonal distribution of sets from each of the clusters differed because of the seasonal availability of target species and the strategic responses of the fishermen (Fig. 2). Sets targeting swordfish (Clusters 4 and 5) usually peaked during March and April, while sets for bigeye tuna (Cluster 2) were most common during winter months. Clusters 1 and 3 had the fewest sets of the five clusters in almost every month. Cluster 1 sets had a semi-annual cycle, peaking in both winter and summer, while Cluster 3 sets showed no distinctive seasonal pattern.

Comparison of fishing strategies among clusters (Table 3) indicated that Cluster 2 (tuna) sets were the most different from Cluster 4 (shark–swordfish) sets. Average operational characteristics of Cluster 2 sets included (1) the smallest vessels, (2) the most hooks per set, (3) the fewest lightsticks per set, (4) the shortest set duration, (5) the highest percentage of day sets, (6) the most even distribution of sets among the lunar phases and (7) the largest percentage of sets within the MHI EEZ. In contrast, Cluster 4 sets were characterized by (1) the largest vessels, (2) the fewest hooks per set, (3) the most lightsticks per set, (4) the longest set duration (together with Cluster 5), (5) the largest percentage of night sets, (6) a larger percentage of sets within the full moon phase than Cluster 2, and (7) the lowest percentage of sets within the MHI EEZ. Cluster 5 (swordfish) sets had

similar characteristics to Cluster 4. Clusters 1 (tuna–swordfish sets) and 3 (mahimahi–tuna sets) appeared to represent mixed fishing strategies with intermediate characteristics between the tuna set and swordfish set clusters, except that the mixed fishing strategies had the highest percentage of sets during the full moon phase and the lowest percentage of sets during the new moon phase (Table 3).

3.2. Spatial distribution of effort

Significant correlation in the spatial distribution of sets between clusters was evidenced by highly significant ($P < 0.01$) r values (Table 4). Exceptions were between Clusters 2 (tuna) and 4 (shark–swordfish), which had an r value (0.052) not significantly different from zero, and Clusters 2 and 5 (swordfish), which had a low but significant ($P < 0.05$) r value (0.13, Table 4). Clusters 4 and 5 had a wider spatial distribution of sets than Clusters 1–3 (Fig. 3). Cluster 4 had effort distributed from the Hawaiian islands to the North Pacific subarctic frontal zone, with effort increasing zonally at ~ 30°N. Effort in Cluster 2 was largely concentrated around the
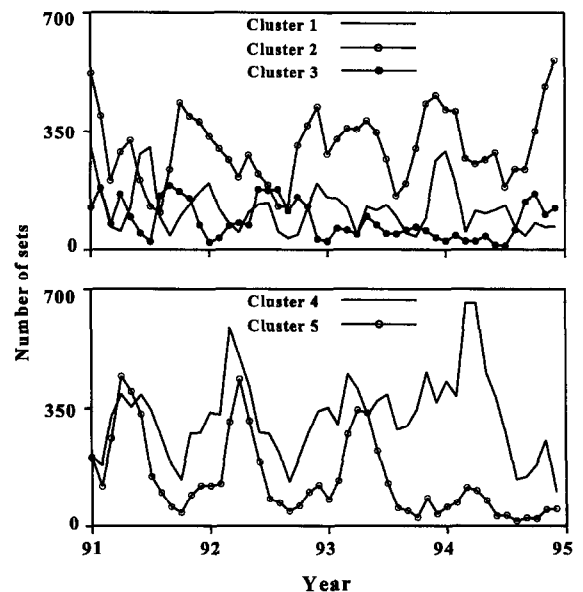

Fig. 2. Number of sets by month for the five clusters of sets in the Hawaii-based longline fishery.

Table 3
Characteristics of fishing operations for five clusters of sets from the Hawaii-based longline fishery (1991–1994). When two numbers are listed within a column they are the mean and standard deviation (in parentheses). Set duration is the interval between the start of deployment and the start of retrieval, not the entire duration of the fishing operation

	Cluster				
	1	2	3	4	5
Primary catch	Tuna–Swordfish	Tuna	Mahimahi–Tuna	Shark–Swordfish	Swordfish
No. of sets	5790	14 573	4077	15 630	6688
Vessel length (m)	65 (10.7)	60 (10.1)	69 (11.0)	75 (9.8)	71 (9.3)
Hooks per set	1051 (343)	1345 (331)	975 (279)	847 (283)	852 (176)
Lightsticks per set	174 (210)	21 (96)	221 (243)	477 (283)	359 (227)
Duration of set (h)	11.4 (2.9)	9.7 (2.9)	11.8 (2.4)	12.4 (1.6)	12.4 (2.2)
% time of fishing					
Day	37.0	84.6	24.5	3.9	6.5
Day–night	3.0	2.7	2.2	1.2	2.7
Night	60.0	12.7	73.3	94.9	90.8
% sets by lunar phase					
New	23.5	32.5	20.6	25.3	23.7
1st and 3rd quarter	33.6	34.1	32.0	35.7	35.7
Full	42.9	33.4	47.4	39.0	40.6
% sets within MHI EEZ	58.9	68.6	48.6	9.0	24.4

MHI. The closest correlation with Cluster 2 was shown by Clusters 1 (tuna–swordfish) and 3 (mahimahi–tuna).

3.3. CPUE time series

Time series of CPUE for three species (bigeye tuna, yellowfin tuna, and swordfish) were compared

at three levels. The lowest level compared subsets of the two largest clusters (2 and 4) to determine if CPUE calculated from longline sets within homogeneous trips (Subsets 2a and 4a) differed from longline sets within trips composed of longline sets from several clusters (Subsets 2b and 4b). At an intermediate level, tuna CPUE time series were compared among Clusters 1–3 and swordfish CPUE was com-

Table 4
Spatial correlation coefficient (r) between clusters of sets from the Hawaii-based longline fishery (1991–1994). Numbers in parentheses represent the number of one square nautical degrees that had at least one set from either cluster, including those not shown in Fig. 3 to preserve confidentiality. Asterisks indicate significant (* $P < 0.05$) and highly significant (** $P < 0.01$) values

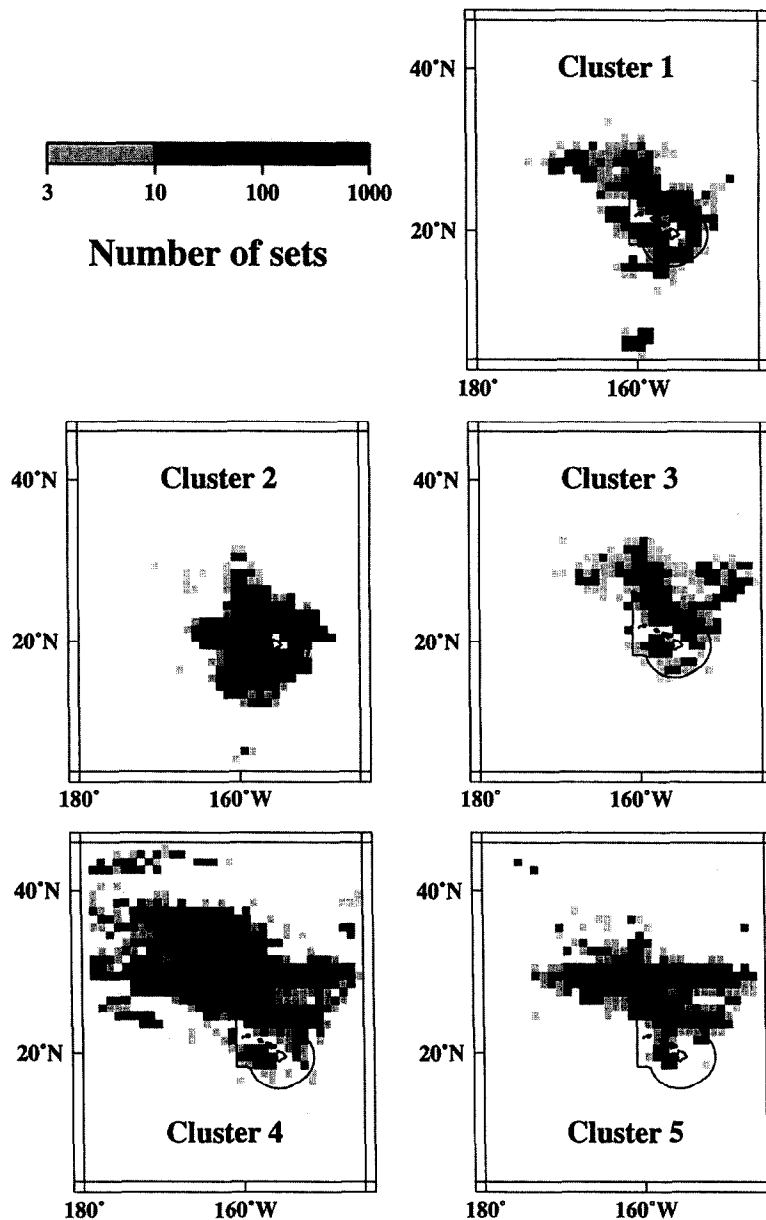| Cluster	Cluster			
	2 Tuna	3 Mahimahi–Tuna	4 Shark–Swordfish	5 Swordfish
1 Tuna–Swordfish	0.60 ** (539)	0.87 ** (536)	0.31 ** (853)	0.67 ** (636)
2 Tuna		0.55 ** (521)	0.052 (877)	0.13 * (669)
3 Mahimahi–Tuna			0.38 ** (821)	0.72 ** (587)
4 Shark–Swordfish				0.67 ** (816)

Fig. 3. Spatial distribution of fishing effort for five clusters of sets in the Hawaii-based longline fishery (1991–1994) showing the main Hawaiian Islands and the main Hawaiian Islands Exclusive Economic Zone. Spatial resolution of fishing effort is one nautical degree. Some fishing areas were deleted to maintain confidentiality.

pared between Clusters 4 and 5. At the highest level, CPUE time series were calculated by aggregating Clusters 1–3 into Cluster I and Clusters 4 and 5 into Cluster II, thus joining the most similar clusters.

Bigeye tuna CPUE time series were similar for Subsets 2a and 2b, (top panel, Fig. 4), both in terms of absolute values and relative changes ($R_{2a,2b} =$ 0.75, Table 5). Time series for Clusters 1–3 were
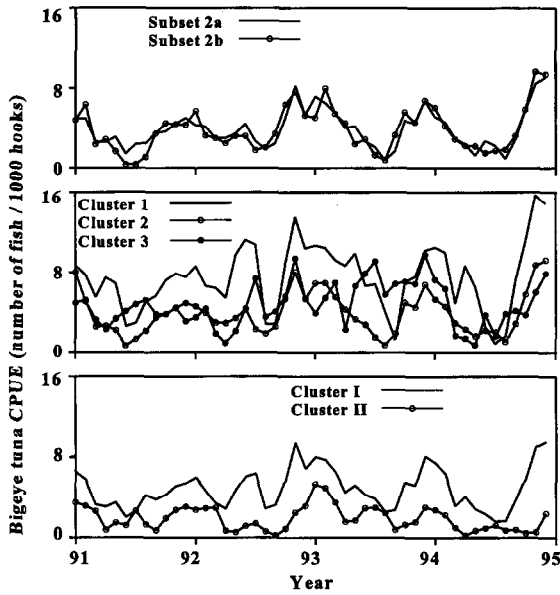
Fig. 4. Monthly mean CPUE of bigeye tuna for different clusters of sets in the Hawaii-based longline fishery. Subset 2a represents sets from homogeneous trips, and Subset 2b represents sets from trips with several set types. Cluster 1 represents tuna–swordfish sets, Cluster 2 tuna sets, Cluster 3 mahimahi–tuna sets, Cluster I represents Clusters 1–3, and Cluster II represents Clusters 4–5.
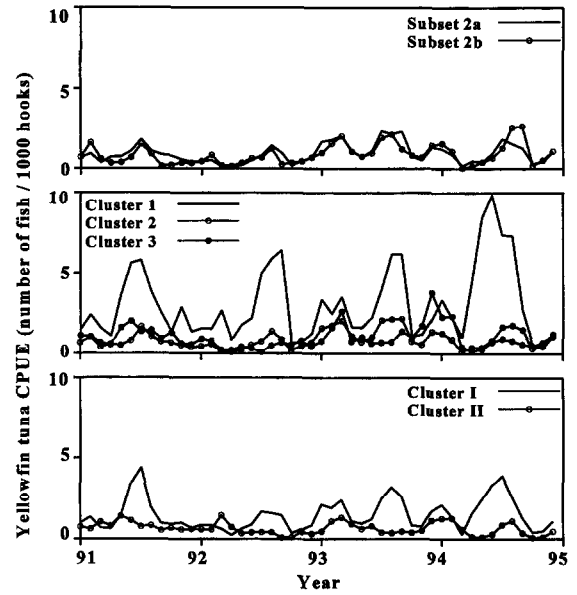

Fig. 5. Monthly mean CPUE of yellowfin tuna for different clusters of sets in the Hawaii-based longline fishery. Subset 2a represents sets from homogeneous trips, and Subset 2b represents sets from trips with several set types. Cluster 1 represents tuna–swordfish sets, Cluster 2 tuna sets, Cluster 3 mahimahi–tuna sets, Cluster I represents Clusters 1–3, and Cluster II represents Clusters 4–5.

different (middle panel, Fig. 4), although relative changes between Clusters 1 and 2 were closely correlated ($R_{1,2} = 0.77$, Table 5). Clusters I and II time series were also different in both absolute value and relative changes (bottom panel, Fig. 4).

Yellowfin tuna CPUE time series were similar for Subsets 2a and 2b (top panel, Fig. 5), both in terms of absolute values and relative changes ($R_{2a,2b} = 0.74$, Table 5). Time series for Clusters 1–3 were

different (middle panel, Fig. 5), although time series for Clusters 1 and 3 showed similar relative changes ($R_{1,3} = 0.61$). Cluster 1 time series indicated strong peaks during summer with absolute values higher than Clusters 2 or 3. Absolute values for Cluster I and II time series were different (bottom panel, Fig. 5) with weakly correlated relative changes ($R_{I,II} = 0.34$, $P = 0.041$, Table 5).

Table 5
Correlation coefficient (R) between time series of CPUE calculated at various levels for bigeye tuna, yellowfin tuna, and swordfish. Subscripts 2a and 4a indicate Subsets of Clusters 2 and 4 from homogeneous trip b only. Subscripts 2b and 4b indicate subsets of Clusters 2 and 4 from trips with several set types. Subscripts I and II indicate aggregations of Clusters 1–3 and 4–5, respectively

Species	Correlation coefficient						
	$R_{2a,2b}$	$R_{1,2}$	$R_{1,3}$	$R_{2,3}$	$R_{4a,4b}$	$R_{4,5}$	$R_{I,II}$
Bigeye tuna	0.75 **	0.77 **	0.31	0.25			0.29
Yellowfin tuna	0.74 **	0.14	0.21	0.61 **			0.34 *
Swordfish					0.78 **	0.86 **	0.33

Asterisks indicate significant (* $P < 0.05$) and highly significant (** $P < 0.01$) values.
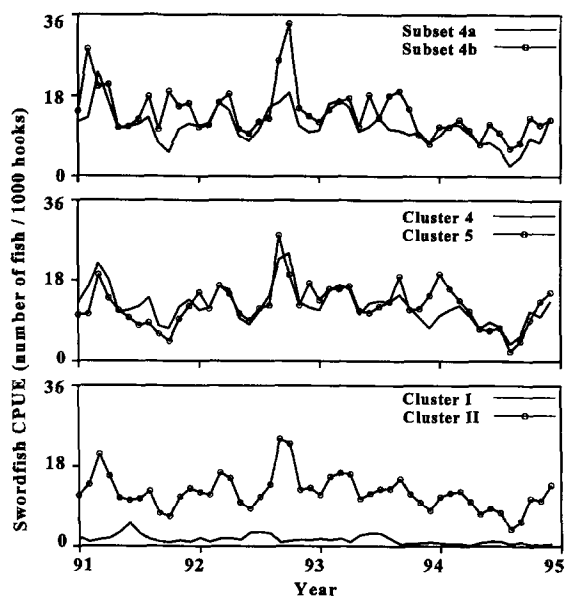
Fig. 6. Monthly mean CPUE of swordfish for different clusters of sets in the Hawaii-based longline fishery. Subset 4a represents sets from homogeneous trips, and Subset 4b represents sets from trips with several set types. Cluster 4 represents shark–swordfish sets, Cluster 5 swordfish sets, Cluster 1 represents Clusters 1–3, and Cluster II represents Clusters 4–5.

Swordfish CPUE time series were similar for Subsets 4a and 4b (top panel, Fig. 6). Although the absolute values of Cluster 4b were slightly higher than Cluster 4a, the relative changes were highly correlated ($R_{4a,4b}$ = 0.78, Table 5). CPUE time series for Clusters 4 and 5 were also similar (middle panel, Fig. 6), with highly correlated relative changes ($R_{4,5}$ = 0.86, Table 5). CPUE time series of Cluster I and II were different (bottom panel, Fig. 6).

3.4. Effort categorization by clustering and trip type

At the highest level of cluster analysis, the Hawaii longline fishery can be described as either a tuna and mixed (Cluster I) or swordfish (Cluster II) fishing strategy. Good correspondence was evident between these clusters of sets and the trip type given by the logbook data. The cluster analysis classified 94.8% of sets from logbook defined tuna trips as Cluster I (tuna and mixed) sets, and classified 90.7% of sets from logbook defined swordfish trips as Cluster II

(swordfish) sets. Sets from logbook defined mixed trips were split between Cluster I (44.8%) and Cluster II (55.2%).

3.5. Intra-trip switching of fishing strategies

Of the 5287 trips conducted from 1991 to 1994, 2054 trips (39%) were entirely composed of sets from Cluster I and 1108 trips (21%) were composed of sets from Cluster II. The remaining 2125 trips (40%) contained both Cluster I and II sets, and thus appeared to reflect switching fishing strategies within a trip.

4. Discussion

Cluster analysis has been used in other fisheries as an effective quantitative method to identify different fishing strategies (Rogers and Pikitch, 1992; Lewy and Vinther, 1994). This quantitative method is necessary, especially in multispecies fisheries because commercial fisheries data often do not provide enough information on fishing behavior and operations. Catch composition is an end-product of fishing that is usually reported by fishermen and contains valuable information on fishing operations for identification of fishing strategies.

Information on catch composition, spatial distribution of effort, and characteristics of fishing operations were examined to describe apparent fishing strategies within the Hawaii-based longline fishery. The actual intentions of the fishermen to target certain species on individual sets using specific methods are not known but may be revealed for a subsample of the fleet through analysis of data from an observer program initiated in 1994. The two dominant types of longline fishing identified by clustering sets according to catch composition were tuna (Cluster 2) and swordfish (Clusters 4–5). Although Clusters 1, 2, and 3 were all characterized by substantial catch proportions of tuna and two or more other types of catch, only Cluster 2 closely resembles traditional tuna longline fishing with a high proportion of tuna, istiophorid billfish, and few swordfish. Clusters 1 and 3 appeared to reflect 'mixed' strategies because they caught substantial proportions of both tuna and swordfish. Overall, clustering revealed categories

much like those used to describe whole trips in the logbook data base. However, a mixed trip may contain tuna sets and swordfish sets, mixed sets, or all three. The main benefit of the method of categorizing fishing strategies used in this study was that individual sets were objectively and quantitatively assigned to a particular category, allowing segregation of the data at a finer resolution.

Between the two most dissimilar fishing strategies, differences in catch compositions are closely related to spatial differences in effort distribution and fishing operations. Targeting tuna is generally conducted with small vessels confined to waters surrounding the MHI with no preference for a particular moon phase. Tuna fishing operations are characterized as daytime fishing with the use of many hooks per set but few lightsticks. Although not quantified in this study, tuna fishing gear in Hawaii is generally deployed in the upper 400 m of the water column. Thus, the fishing characteristics summarized for Cluster 2 should effectively target bigeye tuna based on diurnal movement patterns. Bigeye tuna occupy depths > 200 m during the day and move up to shallower waters at night (Grudinin, 1989; Holland et al., 1990). Hooks set in deeper water catch more bigeye tuna than those set shallow (Saito, 1975; Hanamoto, 1987; Boggs, 1992; Punsly and Nakano, 1992).

Contrary to tuna fishing effort, swordfish effort is distributed over a larger spatial area, from the Hawaiian Archipelago to 45°N. Fishing gear is usually deployed at night with many lightsticks and fewer hooks per set than tuna fishing. These operational characteristics also correspond well with the diurnal behavior of swordfish, as swordfish are typically distributed in deep (> 500 m) water during the day and near-surface waters at night (Carey and Robison, 1981; Carey, 1990).

Ward et al. (1996) clustered longline vessels by the similarity of time and area of operations and the similarity of catch composition. The two methods produced similar results except when several fishing strategies were applied in the same time and area. In the Hawaii-based fishery there appears to be much overlap in the spatial and temporal application of mixed fishing strategies and those for tuna and swordfish. Unlike the difference in spatial distribution between tuna and swordfish sets, spatial distri-

bution of the mixed sets (Clusters 1 and 3) was highly correlated with all the other types.

Although tuna catches exceeded swordfish catches in the mixed set categories, the operational characteristics most strongly resembled those for swordfish sets. The tuna catches of the mixed sets are again explained by diurnal behavior. Tunas' upward movement at night makes them somewhat vulnerable to shallow fishing methods at night. Operational differences between tuna sets and mixed sets are a good reason for not combining CPUE data from tuna and mixed sets despite the similarity in catch composition of these clusters.

Cluster analysis is useful in segregating dissimilar types of fishing effort. After cluster analysis, the CPUE index could be further improved using general linear modelling (GLM) techniques which incorporate additional sources of variability (e.g. gear configuration, area, time, environment). We recommend including data from several categories of set types if the CPUE trends for that species are similar when calculated for each type. For example, there were no appreciable differences in CPUE between sets from homogenous trips and from trips with several set types so it is not necessary to segregate these subsets. If the patterns are similar but differ in absolute value this can be accommodated in a further GLM analysis of the CPUE time series for the combined group using various independent variables. When the CPUE time series pattern is uncorrelated among set types, it is not recommended that a combined data set be used to describe the dynamics of that resource unless a GLM analysis accounts for those differences.

A primary goal of this study was to select the best representative CPUE time series for bigeye tuna, yellowfin tuna, and swordfish in the Hawaii-based longline fishery. For bigeye tuna, the Cluster 2 time series is the best choice because this cluster best represents traditional Hawaii tuna fishing strategies. This cluster can be used for comparison with pre-1990s data because it best reflects the traditional fishing method. Cluster 2 also has the advantage of containing a large number of sets. The next choice would be a combination of Clusters 1 and 2 because these series are highly correlated and the number of sets and fishing areas can be increased, which may prove valuable in studies considering the widest area

and finest spatial and temporal resolution that can be described by the data.

An appropriate CPUE time series for yellowfin tuna is not clearly evident. One choice is the time series calculated from Cluster 1 which has the highest CPUE but has a low number of sets. Cluster 1 does not represent a traditional strategy, which is more comparable to Cluster 2. Cluster 1 might best represent the dynamics of the yellowfin tuna resource from 1991 to 1994. Alternatively, a time series calculated from a combination of Clusters 1–3 (Cluster I) containing a larger number of sets will be useful in analyses considering other independent variables. Similar seasonal patterns exist in the time series from Cluster 1 and Cluster I, but absolute CPUE values are different.

The CPUE time series calculated from Cluster II (Clusters 4–5) is the best choice for swordfish. This time series not only corresponds well with both time series calculated from Clusters 4 and 5 but also represents a larger number of sets and fishing areas. Further analysis of differences in fishing strategies between Clusters 4 and 5 appears warranted because Cluster 5 contains a much smaller proportion of undesirable bycatch (e.g. sharks) than Cluster 4.

One could argue that clustering sets based on species proportions in the catch could cause biases in cluster-based CPUE indices. For example, if fishing intensified and the abundance of swordfish declined then an increased number of sets targeting swordfish but with low swordfish catch might not be included in the swordfish cluster. As a result a real decline in swordfish CPUE for sets targeting swordfish might be dampened or obscured. The severity of such bias was assessed by simulating a substantial decline (50% of observed) in the number of swordfish caught per set in 1993, and reclustering the entire data set. The resulting cluster structure was very similar to the original. The 'new' Cluster II correctly classified 93.2% of the sets defined by the original analysis as swordfish or swordfish–shark sets (Cluster II). Monthly mean swordfish CPUE based on the 'new' Cluster II was reduced 4.17% on average (range −0.02% to −14.11%) compared to monthly mean CPUE based on the original Cluster II (with the simulated 50% decline). However, the two CPUE time series showed the same trends and the same decline in 1993.

The clustering method of categorizing set types based on catch composition in the Hawaii-based longline fishery appears robust even when the abundance of an important species changes because clustering reflects differences in gear selectivity for many species. In a fishery where data on gear and operational characteristics are incomplete or unreliable, clustering based on catch composition provides a useful, objective, and quantitative alternative to traditional means of classification. Clustering methods can be further applied to analyze fishermen's fishing behavior by linking catch compositions and fishing operations to individual fishing vessels or captains. Other multivariate techniques, such as principle component analysis, may also be used to examine numerical relationships between catch compositions and fishing operations.

Acknowledgements

References

Boggs, C.H., 1992. Depth, capture time, and hooked longevity of longline-caught pelagic fish: timing bites of fish with chips. Fish. Bull., 90: 642–658.

Boggs, C.H. and Ito, R.Y., 1993. Hawaii's pelagic fisheries. Mar. Fish. Rev., 55: 69–82.

Carey, F.G. and Robison, B.H., 1981. Daily patterns in the activities of swordfish, Xiphias gladius, observed by acoustic telemetry. Fish. Bull., 79: 103–122.

Carey, F.G., 1990. Further acoustic telemetry observations of swordfish. In: R.H. Stroud (editor), Proceedings of Second International Billfish Symposium, Kailua-Kona, Hawaii, 1–5

August 1988. Part 2, Natl. Coalition for Mar. Conserv., Savannah, Georgia, USA, pp. 103–122.

DiNardo, G.T. and Kwok, W., 1996. Spatiotemporal dynamics of swordfish (*Xiphias gladius*) landings in the Hawaii-based North Pacific pelagic longline fisheries. Proceedings of the International symposium on Pacific swordfish. NOAA Tech. Rep. NMFS, in press.

Dollar, R.A., 1994. Annual report of the 1993 Western Pacific longline fishery. Honolulu Laboratory, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2570 Dole Street, Honolulu, HI, 96822, USA. Administrative Report, H-94-06, 38 pp.

Dollar, R.A. and Yoshimoto, S.S., 1991. The federally mandated longline fishing log collection system in the Western Pacific. Honolulu Laboratory, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2570 Dole Street, Honolulu, HI, 96822, USA. Administrative Report, H-91-12, 34 pp.

Grudinin, V.B., 1989. On the ecology of yellowfin tuna (*Thunnus albacares*) and bigeye tuna (*Thunnus obesus*). J. Ichthyol., 26: 22–29.

Hanamoto, E., 1987. Effect of oceanographic environment on bigeye tuna distribution. Bull. Jpn. Soc. Fish. Oceanogr., 51: 203–216.

He, X. and Boggs, C.H., 1996. Do local catches affect local abundance? Time series analysis of Hawaii's tuna fisheries. In: Shomura R.S., J. Majkowski and R.F. Harman (Editors), 1996. 1995 Status of interactions of Pacific tuna fisheries: Proceedings of the second FAO expert consultation on interactions of Pacific tuna fisheries, 23–31 January 1995, Shimizu, Japan. FAO Fisheries Technical Report 365. 612p. pp. 224-240.

Hilborn, R. and Walters, C.J., 1992. Quantitative Fisheries Stock Assessment. Chapman and Hall, New York, USA, 570 pp.

Holland, K.N., Brill, R.W. and Chang, R.K.C., 1990. Horizontal and vertical movements of yellowfin and bigeye tuna associated with fish aggregating devices. Fish. Bull., 88: 493–507.

Ito, R.Y, R.E. Dollar and Kawamoto, K., 1996. The Hawaii longline fishery for swordfish. Proceedings of the International symposium on Pacific swordfish. NOAA Tech. Rep. NMFS, in press.

Johnson, R.A. and Wichern, D.W., 1988. Applied Multivariate Statistical Analysis. Prentice Hall, NJ, USA, 607 pp.

Lewy, P. and Vinther, M., 1994. Identification of Danish North Sea trawl fisheries. ICES J. Mar. Sci., 51: 263–272.

Otsu, T., 1954. Analysis of the Hawaiian long-line fishery, 1948–52. Commer. Fish. Rev., 16: 1–17.

Pooley, S.G., 1993. Hawaii's marine fisheries: Some history, long-term trends, and recent developments. Mar. Fish. Rev., 55: 7–19.

Punsly, R. and Nakano, H., 1992. Analysis of variance and standardization of longline hook rates of bigeye (*Thunnus obesus*) and yellowfin (*Thunnus albacares*) tunas in the Eastern Pacific Ocean during 1975–1987. Bulletin of Inter-American Tropical Tuna Commission, La Jolla, California, USA, 20(4): 167–177.

Rogers, J.B. and Pikitch, E.K., 1992. Numerical definition of groundfish assemblages caught off the coast of Oregon and Washington using commercial fishing strategies. Can. J. Fish. Aquat. Sci., 49: 2648–2656.

Saito, S., 1975. On the depth of capture of bigeye tuna by further improved vertical longline in the tropical Pacific. Bull. Jpn. Soc. Sci. Fish., 41: 831–841.

Schneider, D.C., 1994. Quantitative Ecology: Spatial and Temporal Scaling. Academic Press, New York, USA, 395 pp.

Shomura, R.S., 1959. Changes in tuna landings of the Hawaiian longline fishery. Fish. Bull., 60: 87–106.

Snedecor, G.W. and Cochran, W.G., 1980. Statistical Methods, Seventh Edition. The Iowa State University Press, Ames, IA, USA, 507 pp.

Statistical Analysis Systems (SAS) Institute Inc., 1989. SAS/STAT User's Guide, Version 6, Fourth Edition. SAS Institute Inc., Cary, North Carolina, USA, 943 pp.

Suzuki, Z., 1989. Catch and fishing effort relationships for striped marlin, blue marlin, and black marlin in the Pacific Ocean, 1952 to 1985. In: Stroud, R.H. (Editor), Planning the Future of Billfishes, Research and Management in the 90's and Beyond. Part 1: Fishery and Stock Synopses, Data Needs and Management. Natl. Coalition Mar. Conserv., Savannah, GA. Mar. Rec. Fish., 13: 165–177.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc., 58: 236–244.

Ward, P.J., C.M. Ramirez and Caton, A.E., 1996. The types of longlining activities of Japanese vessels in the eastern Australian fishing zone during the 1980s. pp. 49–74, In: Ward, P.J. (Editor), Japanese Longlining in Eastern Australian Waters 1962–1990. Bureau of Resource Sciences, Canberra, 264 pp.