INTER-AMERICAN TROPICAL TUNA COMMISSION

1ST EXTERNAL REVIEW OF DATA USED OF STOCK ASSESSMENTS OF TROPICAL TUNA IN THE EASTERN PACIFIC OCEAN

La Jolla, California (USA)
02-06 October 2023
(by videoconference)

DOCUMENT RVDTT-01-01

FISHERY DEFINITIONS FOR BIGEYE TUNA IN THE EASTERN PACIFIC OCEAN

INTRODUCTION

The exploratory assessment models for bigeye tuna in the eastern Pacific Ocean (EPO) are not spatially structured and use the "areas-as-fleets" approach, which models geographic areas as separate fleets with different selectivity curves in a single-stock stock assessment model. Although this approach implicitly assumes that the stock is homogenously distributed throughout its range and any differences in composition data arise due to different contact selectivity (Hurtado-Ferro et al. 2014), it recognizes that fishing in different areas usually leads to different ages/sizes of fish being removed from the population due to spatial variation in size structure. As such, fisheries need to be defined spatially to achieve a homogeneous distribution of fish across each area. This approach ensures that the size composition of each fishery is not sensitive to the location of fishing activities (Punt 2019).

We use a regression tree approach for analyzing length frequency data to provide gear and set type-specific fishery definitions. The regression tree algorithm (Lennert-Cody *et al.* 2013, Lennert-Cody *et al.* 2010) uses recursive partitioning to search for hierarchical binary decision rules that divide the data into more homogeneous subgroups. The binary decision rules are selected to provide the greatest decrease in the heterogeneity of length composition data, which is measured based on the Kullback–Leibler divergence. The regression tree algorithm has been recently included in an R package *FishFreqTree*[1], where fisheries length-frequency data, separated by gear (longline/purse-seine) and purse-seine set type (floating object/unassociated/dolphin), are grouped by latitude, longitude, quarter, and cyclical-quarter.

Two main differences exist between the regression tree analysis conducted for the last benchmark assessment and this exploratory analysis. The previous analysis (BET-02-02) was initially conducted to define stock structures for a spatially explicit stock assessment model for bigeye tuna in the EPO. It is based on both catch-per-unit-effort (CPUE) and length frequency to find compromised spatial boundaries across gear and set type. In contrast, this analysis is based solely on length frequency and is conducted for each gear and set type to provide uncompromised gear and set type-specific fishery definitions. The habitat preference of bigeye tuna is size-specific, so fish caught by different gear types (longline catches predominately adult bigeye and purse-seine catches predominantly juvenile bigeye) are likely to have distinct spatial patterns of age/size composition. As such, independent fishery definitions are more appropriate for this assessment model that utilizes the "areas-as-fleets" approach.

The second difference is the source of longline composition data. The regression tree analysis conducted for the last benchmark assessment is based partially on the longline length composition data that Japan

[1] https://github.com/HaikunXu/FishFreqTree

submitted to the IATTC's public domain. This data is coarse and pre-aggregated by 5° latitude, 10° longitude, and 1 quarter (WSBET-02-02). This exploratory analysis uses the new longline length composition data submitted to the IATTC by Japan through a Memorandum of Understanding. This data has a much finer spatial and temporal resolution (1° latitude, 1° longitude, and month) and includes additional useful information, such as the bin size associated with each length measurement. In addition, we also consider the longline composition data submitted by Korea to evaluate the sensitivity of longline fishery definitions to the source of longline composition data.

LONGLINE FISHERIES

Longline fisheries are defined in this exploratory analysis using mainly the Japanese longline length composition data which covers the EPO between 1986-2020. The data were measured by fishers before 2011, by both fishers and observers between 2011-2014, and by observers after 2014. Before being analyzed by the regression tree algorithm, the data is filtered to include only commercial vessels' data collected at a spatial resolution of 1° x 1° and a bin size of 1, 2, or 5 cm, as training vessel data is deemed unrepresentative of the commercial catches (SAC-07-03d). The input length frequencies for the regression tree analysis are computed from this data for each 5° x 5° x year x quarter stratum with a bin size of 10 cm (60-70 cm, 70-80 cm, …, >190 cm). The main reason for aggregating the data at 5° x 5° is to match the spatial resolution of the longline catch data submitted to the IATTC. We remove the poorly sampled spatiotemporal strata with a sample size of less than 15 fish measured and poorly sampled spatial grids with less than 3 years of data between 1986 and 2020 (Figure 1).

We specify the regression tree algorithm to define four splits or five longline fisheries for the EPO excluding Hawaii, where a separate longline fishery is defined as in the previous benchmark assessment model (SAC-11-06). The regression tree is hierarchical and may exhibit a certain degree of instability. Instead of selecting only the best candidate for each split, we consider the top four and two competing candidates for the first and second splits, respectively, and rank the eight (4 x 2) 4-split combinations according to the proportion of variance in the length-frequency data explained.

Among the eight 4-split combinations, the best one selected for the longline fishery in the EPO explains 14.12% of the variance in the length-frequency data (Table 1). The first split (15°S) divides the EPO into tropical and temperate regions (Figure 2 and Table 1). The second split (105°W) divides the tropical EPO into the eastern and western portions (Figure 2 and Table 1). The third split (5°S) further separates the western tropical EPO into the northern and southern portions (Figure 2 and Table 1). The last split (90°W) divides the temperate EPO into the coastal and offshore portions (Figure 2 and Table 1). In general, the fish caught by longline are larger in tropical regions than in temperate regions and are larger in offshore regions than in inshore regions (Figure 2). The two offshore tropical regions have the largest mean size of bigeye for the longline, while the inshore region of Peru has the smallest mean size of bigeye for the longline (Figure 2).

We apply the same analysis to the Korean longline length composition data which covers only the offshore tropical EPO between 2011 and 2020 (Figure 3). Consistently, the first split selected by the regression tree algorithm for the longline fishery is 5°S, explaining 5.44% of the variance in the length-frequency data. However, the comparison of the longline length frequency data collected by Japan and Korea in the same area (offshore tropical EPO) and period (2011-2020) shows that Korean fishers catch a noticeably higher proportion of large bigeye than Japanese fishers (Figure 4). Therefore, the Japanese and Korean longline fleets should be treated as separate longline fisheries in the assessment models. There are eight longline fisheries defined for bigeye tuna in the EPO, of which six are Japanese longline fisheries and two are Korean longline fisheries (Korean longline vessels operate predominantly in the two offshore tropical fishing grounds). We assume that other CPCs (e.g., Chinese Taipei, China, United States), which catch a smaller amount of bigeye in the EPO than Japan and Korea, have the same selectivity as Japan and are grouped into the six Japanese longline fisheries.

Given that longline catches are reported to the IATTC in numbers by some fleets and in weight by others, two longline fleets, one in number and one in weight, are defined for each of the eight longline fisheries. This allows longline catches to be included in their original units and the stock assessment model to conduct the unit transition internally. In total, the exploratory assessment model includes sixteen longline fishery fleets (Table 4).

PURSE-SEINE FISHERIES ON FLOATING-OBJECTS

The definition of floating-object (OBJ) fisheries is based on length composition data collected by port samplers from the OBJ sets made by Class-6 vessels (Suter 2010). Port samplers collect data only from wells with catch from the same set type, sampling area, and year-month. We remove the data before 2000 because the sampling protocol used by the IATTC port-sampling program changed in that year and the OBJ fishery was not fully expanded across the EPO during the 1990s. The raw data has a 5° x 5° spatial resolution and a 1 cm bin size from 1 cm to 201 cm. Poorly sampled grids with less than 4 years of data available since 2000 are removed from the dataset. The remaining data is then aggregated into fifteen 10 cm length bins (<30 cm, 30-40 cm, 40-50 cm, …, >160 cm).

Same as in the last benchmark assessment model, the assessment model in this exploratory analysis includes five OBJ fishery fleets. We therefore specify the regression tree algorithm to find four splits. We also consider the top four and two competing candidates for the first and second splits, respectively, and rank the eight (4 x 2) split combinations according to the proportion of variance in the length-frequency data explained.

Among the eight 4-split combinations, the best one selected for the OBJ fishery in the EPO explains 13.93% of the variance in the length-frequency data (Table 2). The first three splits (110°W, 100°W, and 125°W) are all meridional (Table 2), which is consistent with the fact that the contrast in the mean size of bigeye for the OBJ is most obvious in the East-West direction (Figure 5). Notably, the first split (110°W), which is the most important split, is identical to that selected in the last benchmark assessment. The mean size of bigeye tuna for the OBJ increases continuously from the management boundary (150°W) to the coastline, which is almost the opposite of the spatial pattern of mean size for the longline. The fourth split (15°S) divides the inshore EPO by latitude into tropical and temperate inshore regions (Figure 5 and Table 2). The temperate inshore region off Peru has the largest mean size of bigeye for the OBJ, while the westmost region adjacent to the management boundary has the smallest mean size (Figure 5).

We also conduct a sensitivity analysis on the OBJ fishery definition by fitting the same regression tree model to standardized length frequency data. Standardized length frequency is defined as the raw length frequency divided by the length frequency averaged for the EPO in the same year. This standardization aims to remove the influence of factors other than selectivity and availability on length frequency distribution. For example, the highly variable recruitment of bigeye tuna in the EPO occasionally creates strong cohorts moving through the population and may distort the length frequencies observed in the OBJ fishery. The regression tree provides an identical fishery definition for the OBJ fishery based on the standardized length frequency data, further validating the credibility of the selected fishery definition for OBJ.

PURSE-SEINE FISHERIES ON FREE SCHOOLS

The definition of unassociated purse-seine (NOA) fisheries, as for the OBJ fisheries, is based on length composition data collected by port samplers, but from NOA sets, made by Class-6 vessels (Suter 2010). We remove the data before 2000 because the sampling protocol used by the IATTC port-sampling program changed in that year. The raw data is aggregated into fifteen 10 cm length bins (<30 cm, 30-40 cm, 40-50 cm, …, >160 cm).

The length frequency data for NOA sets are sparse both spatially and temporally (Figure 6), and NOA sets contribute to only a small percentage of bigeye catch in the EPO. Therefore, we include only two NOA fishery fleets in the exploratory assessment model. Due to the lack of length composition data and a negligible percentage contribution to total bigeye catch, we pool both pole-and-line and dolphin-associated purse-seine sets into the NOA sets in the exploratory assessment model for convenience (but not in the tree analysis). The same simplification was made in the last benchmark assessment.

The best split selected for the NOA fishery in the EPO (i.e., 130°W) explains 9.14% of the variance in the length-frequency data (Table 3). The mean size of bigeye tuna for the NOA is generally larger in the onshore region than in the offshore region, although the spatial pattern of mean size is very noisy (Figure 6).

SUMMARY

Twenty-four fisheries are defined for the exploratory assessment model for bigeye tuna in the EPO (Table 4 and Figure 7). These fisheries comprise sixteen longline fisheries, six OBJ fisheries (including one discard OBJ fishery), and two NOA+DEL fisheries. One unsolved issue identified in the last benchmark assessment is that some purse-seine fisheries have multiple modes in the aggregated length frequency. This is an indication that these fisheries are not well defined, which is not surprising given that the fishery definitions made in the last benchmark assessment are compromises among different gears and set types. In contrast, the fishery definitions made in this exploratory analysis are specific to gear and set type. As a result, no purse-seine fisheries in the exploratory assessment have a notable bimodal pattern in the aggregated length frequency (Figure 8), indicating that independent fishery definitions are more suitable for this assessment.

REFERENCES

Hurtado-Ferro, F., Punt, A.E., and Hill, K.T. 2014. Use of multiple selectivity patterns as a proxy for spatial structure. Fisheries Research **158**: 102-115.

Lennert-Cody, C.E., Maunder, M.N., Aires-da-Silva, A., and Minami, M. 2013. Defining population spatial units: Simultaneous analysis of frequency distributions and time series. Fisheries Research **139**: 85-92.

Lennert-Cody, C.E., Minami, M., Tomlinson, P.K., and Maunder, M.N. 2010. Exploratory analysis of spatial–temporal patterns in length–frequency data: An example of distributional regression trees. Fisheries Research **102**(3): 323-326.

Punt, A.E. 2019. Spatial stock assessment methods: a viewpoint on current issues and assumptions. Fisheries Research **213**: 132-143.

Suter, J.M. 2010. An evaluation of the area stratification used for sampling tunas in the eastern Pacific Ocean and implications for estimating total annual catches.

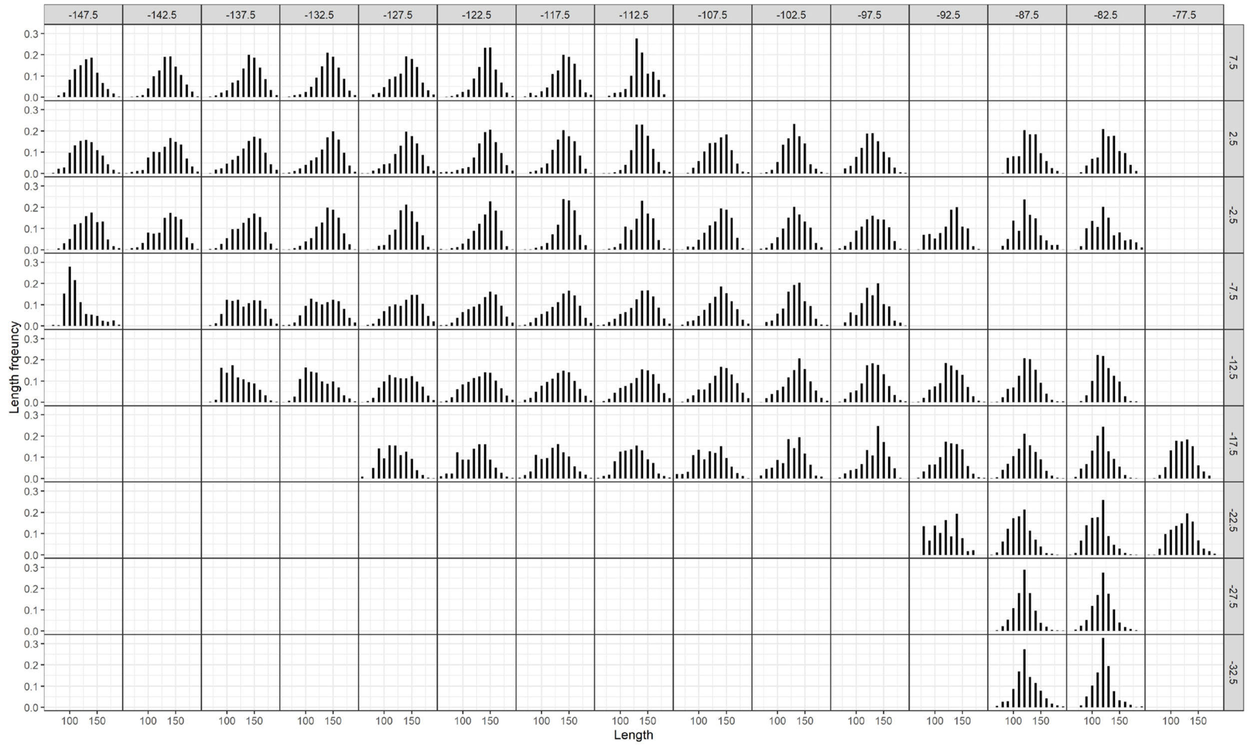

FIGURE 1. The spatial pattern of length frequency for bigeye tuna caught by the Japanese longline fishery.

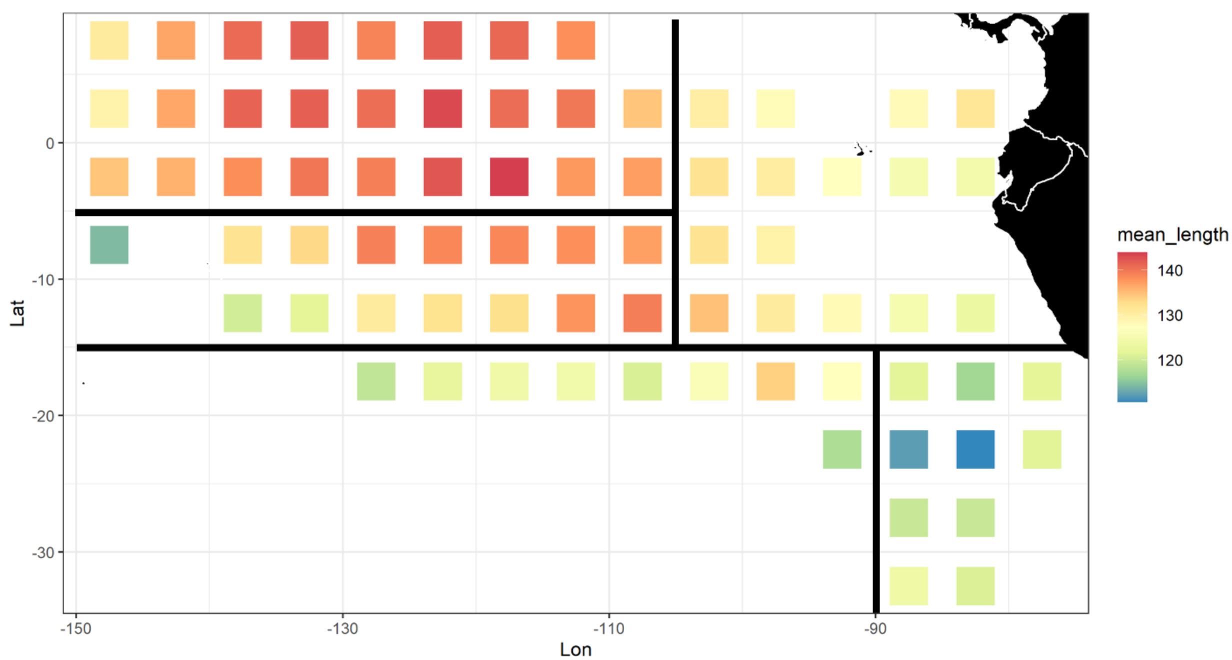

FIGURE 2. Map of average length (cm) of bigeye tuna caught by the Japanese longline fishery. The four solid black lines are the best four-split combination selected by the regression tree algorithm.

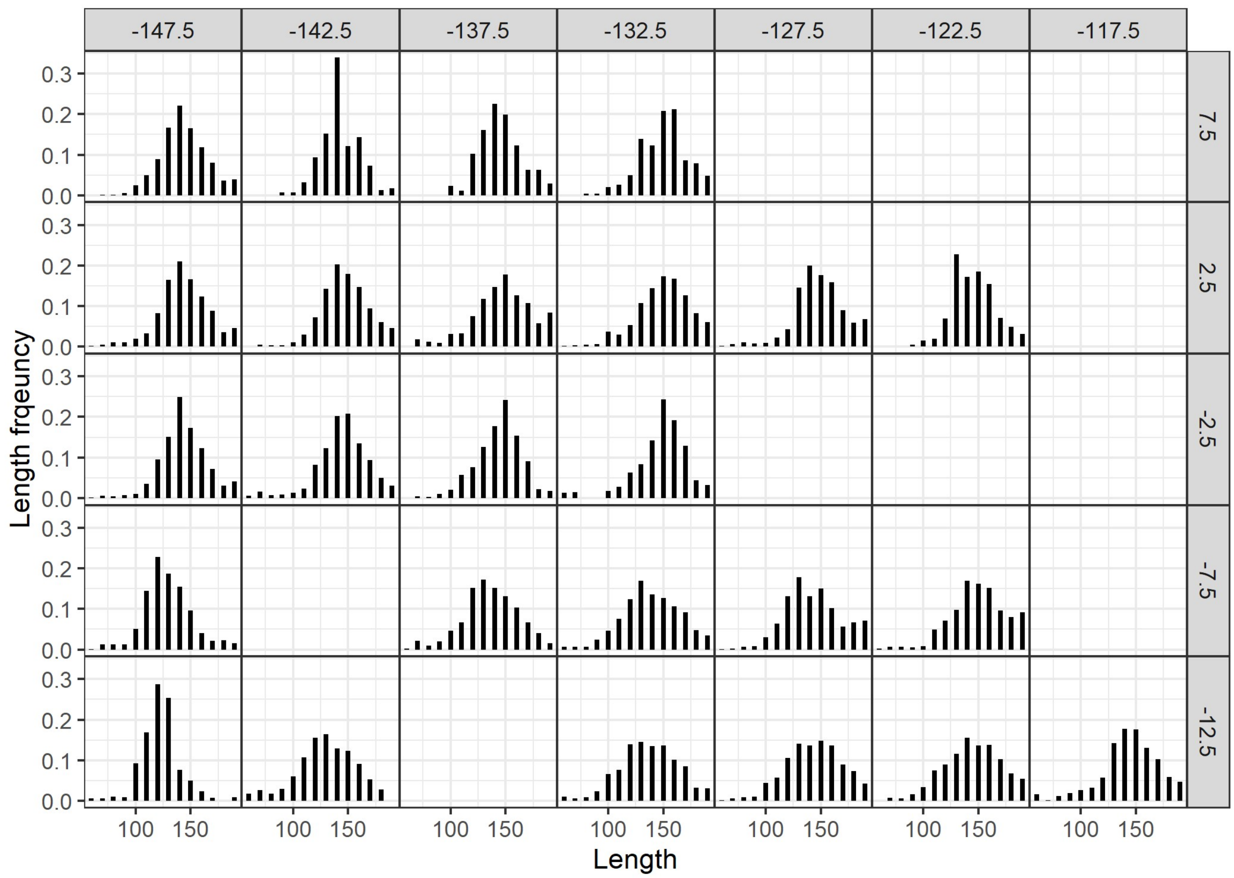

FIGURE 3. The spatial pattern of length frequency for bigeye tuna caught by the Korean longline fishery.

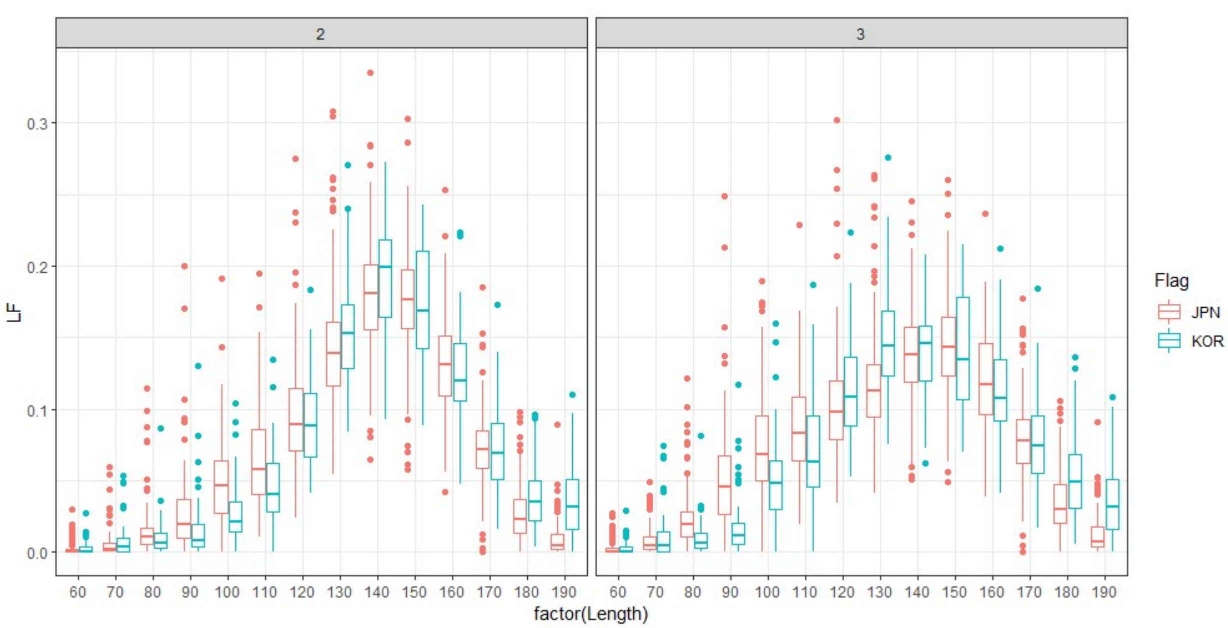

FIGURE 4. Comparison of Japanese and Korean longline length composition data in Areas 2 and 3 between 2011-2020.

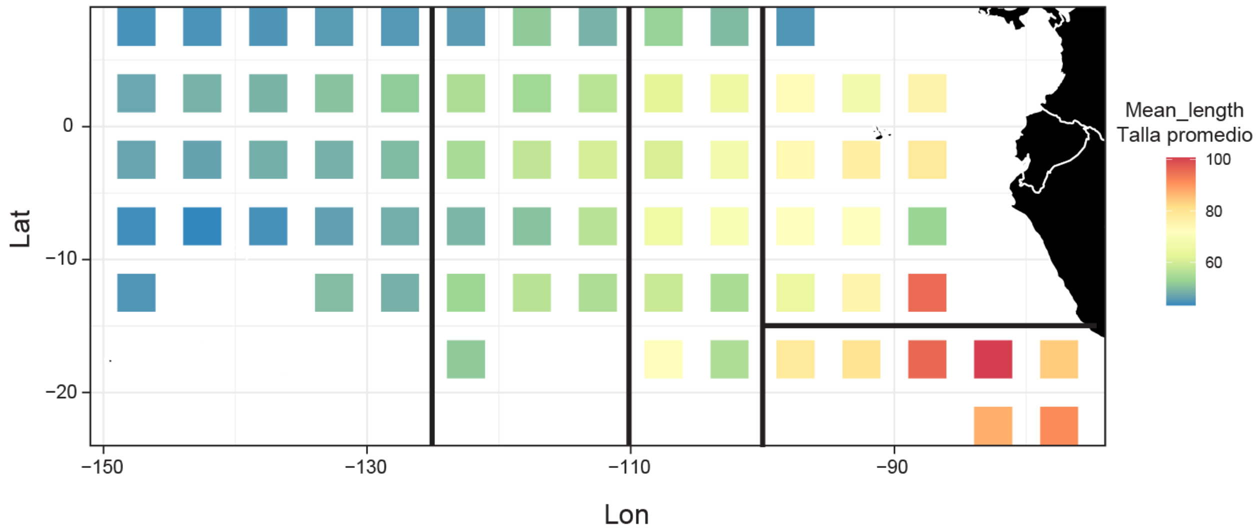

FIGURE 5. Map of average length (cm) of bigeye tuna caught by the floating-object fishery. The four solid black lines are the best four-split combination selected by the regression tree algorithm.

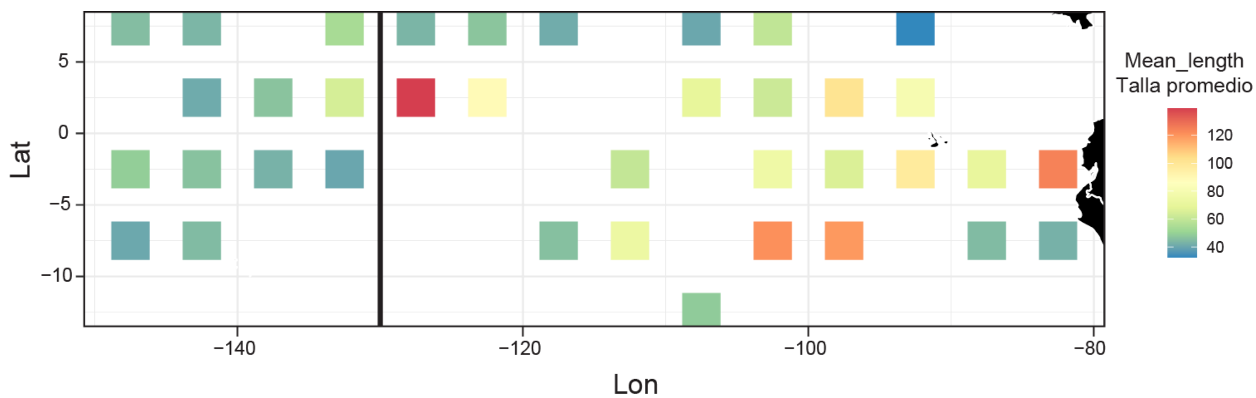

FIGURE 6. Map of average length (cm) of bigeye tuna caught by the unassociated fishery. The solid black line is the best split selected by the regression tree algorithm.

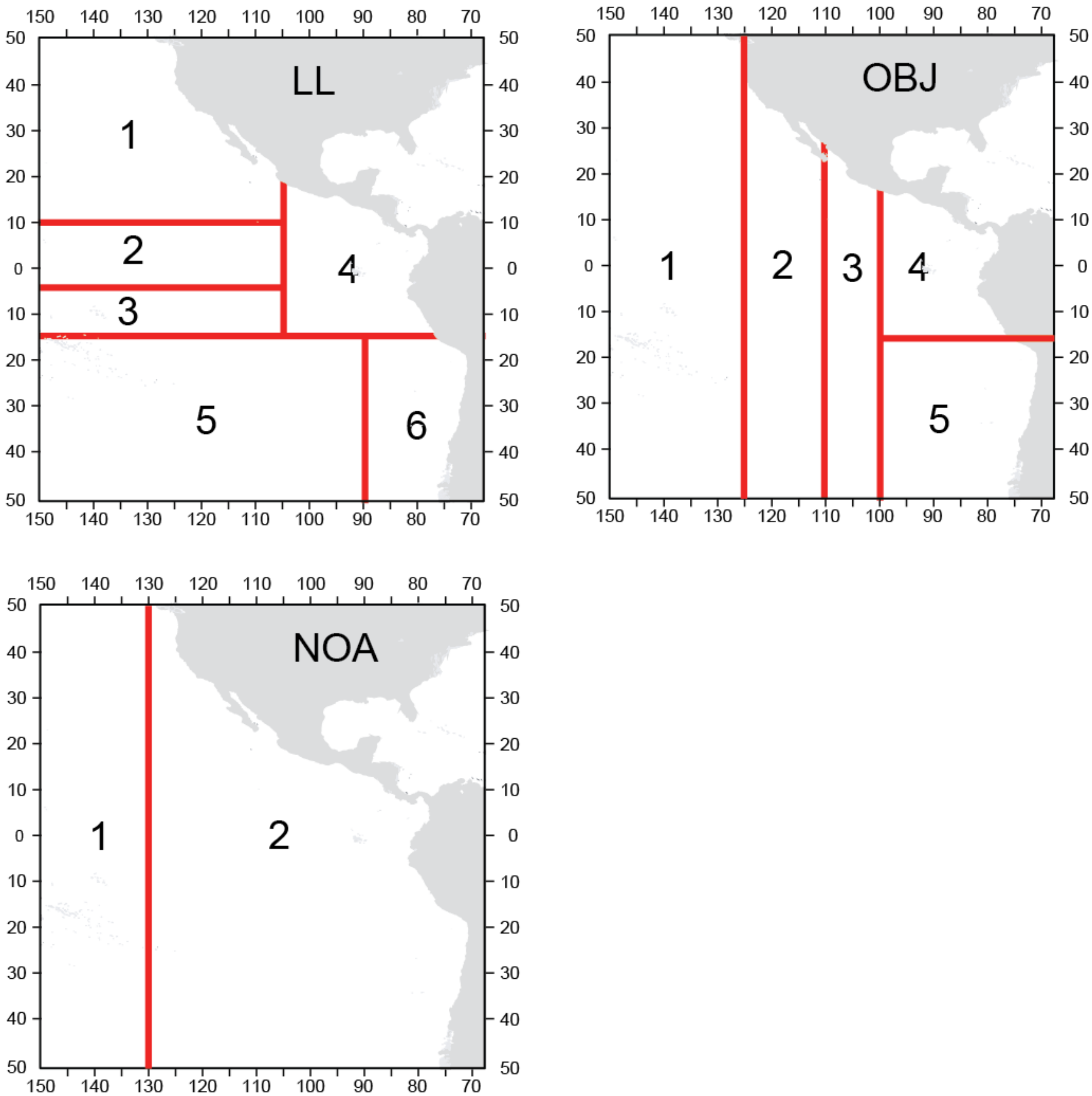

FIGURE 7. Summary of area definitions for the longline (LL), floating-object (OBJ), and unassociated (NOA) fishery fleets in the exploratory assessment models for bigeye tuna in the eastern Pacific Ocean.

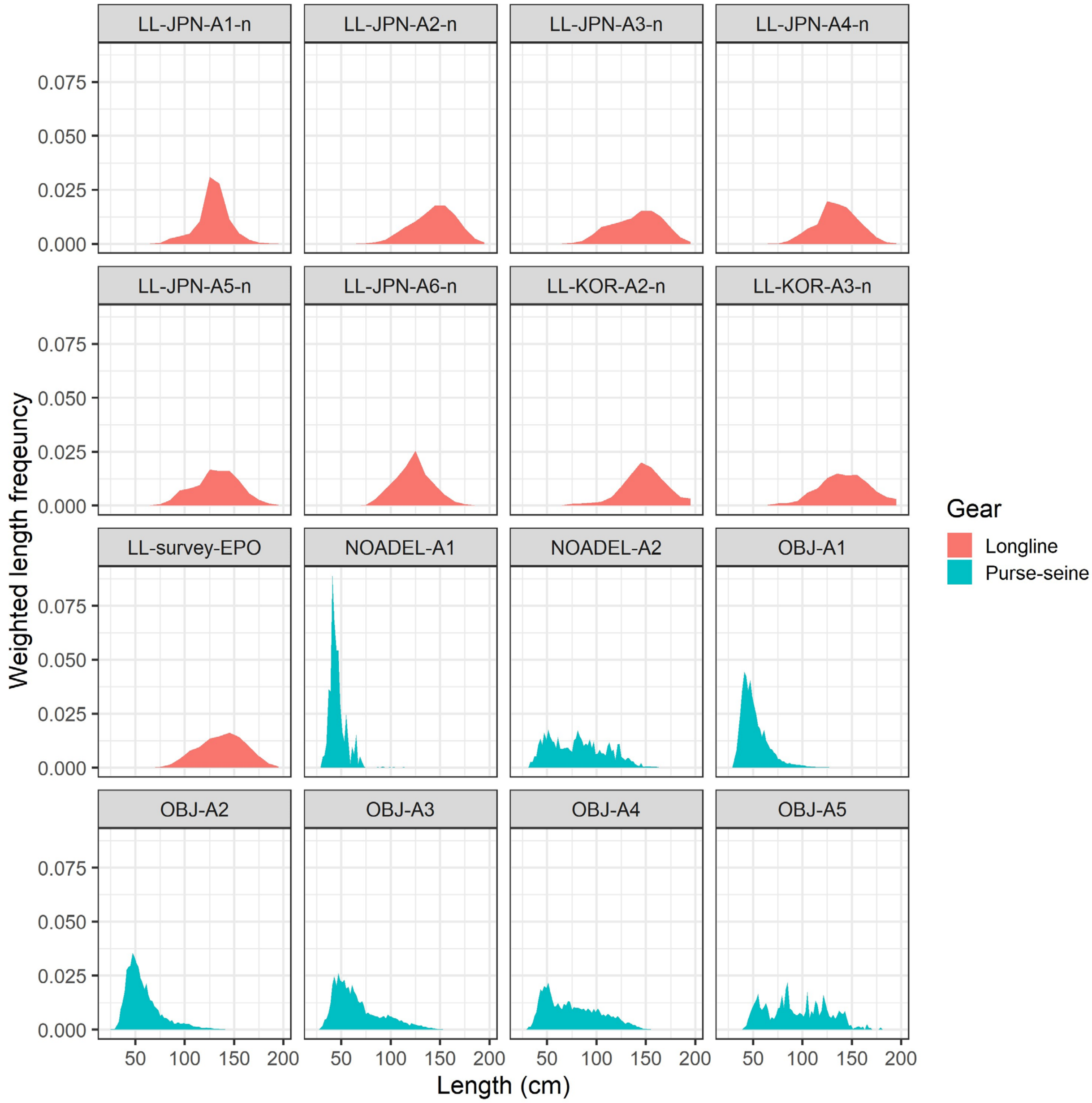

FIGURE 8. Sample-size weighted length frequency of bigeye tuna observed by each fishery and survey fleet in the exploratory assessment model.

TABLE 1. The best four-split combination selected by the regression tree algorithm for the longline fishery for bigeye tuna in the eastern Pacific Ocean. The last column shows the percentage of variance in the length-frequency data explained.

Split	Key	Value	Variance explained
Split1	Latitude	15°S	8.07%
Split2	Longitude	105°W	10.91%
Split3	Latitude	5°S	13.01%
Split4	Longitude	90°W	14.12%

TABLE 2. The best four-split combination selected by the regression tree algorithm for the OBJ fishery for bigeye tuna in the eastern Pacific Ocean. The last column shows the percentage of variance in the length-frequency data being explained.

Split	Key	Value	Variance explained
Split1	Longitude	110°W	10.53%
Split2	Longitude	100°W	11.92%
Split3	Longitude	125°W	13.21%
Split4	Latitude	15°S	13.93%

TABLE 3. The best four-split combination selected by the regression tree algorithm for the NOA fishery for bigeye tuna in the eastern Pacific Ocean. The last column shows the percentage of variance in the length-frequency data being explained.

Split	Key	Value	Variance explained
Split1	Longitude	130°W	9.14%

TABLE 4. A summary of the fishery fleets defined for the exploratory assessment of bigeye tuna in the EPO. PS = purse-seine; LL = longline; OBJ = sets on floating objects; NOA = sets on unassociated fish; DEL = sets on dolphins. Fleet-specific definition of column “Area” can be found in Figure 7.

<table>
<tr><th>Fleet Number</th><th>Gear</th><th>Flag/Set type</th><th>Area</th><th>Catch data</th><th>Unit</th></tr>
<tr><td>1</td><td rowspan="16">LL</td><td rowspan="6">Japan</td><td>1</td><td rowspan="16">Retained catch only</td><td rowspan="8">1,000s</td></tr>
<tr><td>2</td><td>2</td></tr>
<tr><td>3</td><td>3</td></tr>
<tr><td>4</td><td>4</td></tr>
<tr><td>5</td><td>5</td></tr>
<tr><td>6</td><td>6</td></tr>
<tr><td>7</td><td rowspan="2">Korea</td><td>2</td></tr>
<tr><td>8</td><td>3</td></tr>
<tr><td>9</td><td rowspan="6">Japan</td><td>1</td><td rowspan="8">tons</td></tr>
<tr><td>10</td><td>2</td></tr>
<tr><td>11</td><td>3</td></tr>
<tr><td>12</td><td>4</td></tr>
<tr><td>13</td><td>5</td></tr>
<tr><td>14</td><td>6</td></tr>
<tr><td>15</td><td rowspan="2">Korea</td><td>2</td></tr>
<tr><td>16</td><td>3</td></tr>
<tr><td>17</td><td rowspan="8">PS</td><td rowspan="6">OBJ</td><td>1</td><td rowspan="5">Retained catch + discards (inefficiency)</td><td rowspan="8">tons</td></tr>
<tr><td>18</td><td>2</td></tr>
<tr><td>19</td><td>3</td></tr>
<tr><td>20</td><td>4</td></tr>
<tr><td>21</td><td>5</td></tr>
<tr><td>22</td><td>1-5 (EPO)</td><td>Discards (size-sorting)</td></tr>
<tr><td>23</td><td rowspan="2">NOA+DEL</td><td>1</td><td rowspan="2">Retained catch + discards (all)</td></tr>
<tr><td>24</td><td>2</td></tr>
</table>