DRAFT

INTER-AMERICAN TROPICAL TUNA COMMISSION

EXTERNAL REVIEW OF IATTC BIGEYE TUNA ASSESSMENT

La Jolla, California (USA)
3-7 May 2010

DOCUMENT BET-01-02a (DRAFT)

PRELIMINARY ANALYSIS OF SPATIAL-TEMPORAL PATTERN IN BIGEYE TUNA LENGTH-FREQUENCY DISTRIBUTIONS AND CATCH-PER-UNIT-EFFORT TRENDS

Cleridy E. Lennert-Cody, Mark N. Maunder and Alexandre Aires-da-Silva

CONTENTS

1. SUMMARY

Multivariate regression tree methods were used to study spatial and seasonal pattern in bigeye tuna length-frequency distributions and catch per unit effort trends from the longline fishery and the purse-seine fishery on floating-objects of the eastern Pacific Ocean (EPO). Preliminary results suggest that spatial pattern often dominated over seasonal pattern. There was some similarity between the spatial structure identified from the longline length-frequency distributions and that from the catch per unit effort trends. Separate analyses of the two data types indicated a partition at ~150°W, consistent with the current western boundary definition of the EPO fishery area. In addition, latitudinal partitions of the EPO from analysis of one or both data types were indicated at ~10-15°N, at the equator, and at 15°S. The latitudinal partition ~10°-15°N is also consistent with the current stock assessment stratification for the longline fishery. A longitudinal partition at around ~95°-100°W was also suggested. Results of the analysis of purse-seine length-frequency data indicate an inshore/offshore partition of the EPO at 110°W. In addition, there was some evidence that the inshore region should be further divided into an equatorial area between 10°S-5°N, and an additional nearshore partition along 90°W. This stratification EPO shows some similarity to that currently used for the purse-seine stock assessment. The catch per unit effort data from the purse-seine fishery did not provide any information on spatial structure. An algorithmic procedure for simultaneous tree-based analysis of spatial pattern in the length-frequency distributions and catch per unit effort trend estimates is currently being developed.

2. BACKGROUND

Understanding the spatial-temporal distributions of fish populations is important for stock assessment and management. In particular, analysis of spatial pattern in fisheries catch data can be used to define sub-stocks, as well as sub-fisheries with different selectivities. The spatial sampling stratification from which current stock assessment areas for bigeye tuna in the EPO (Figure 1) are derived is based on consideration of the historical distribution of the longline and purse-seine fisheries and their catch characteristics (Tomlinson, 2004; Suter, 2008 and references therein; Aires-da-Silva and Maunder, 2009). The spatial stratification of the purse-seine fisheries were restricted to the 13 market sampling areas that were designed to optimize length-frequency sampling for yellowfin tuna and these may not be appropriate for bigeye tuna. Several studies have undertaken a re-evaluation of the existing spatial strata defined for the EPO. An annealing algorithm was used to group monthly 5°x10° estimates of bigeye CPUE from the longline fishery in 1963-1992 into larger areas (Watters and Deriso, 2000). Results suggested post-stratification of the data might be simplified to 9 areas from the 13 sampling areas, and indicated the need for additional stratification west of 120°W around the equator. Cluster analysis has been applied to purse-seine port-sampling and catch and effort data from 2000-2006 to derive revised boundaries of the 13 sampling areas currently in use (Suter, 2008). Two cluster analyses were performed, one on average weight and standard deviation of weight for the three main tuna species, and the other on species and set type composition of the catch. The revised 13-area stratification was similar to that currently in use; aggregating strata led to increased variance of the estimated total catch. Schaefer (2008) summarized support for spatial structure from catch data, biology, and genetics, and concluded that there are potentially spatially-segregated northern and southern sub-stocks, with little mixing between them. In this report, we present preliminary results of multivariate regression tree analyses of the spatial-temporal pattern associated with bigeye tuna length-frequency distributions and CPUE trends from the recent Japanese longline fishery and the purse-seine fishery of large vessels on floating-objects in the EPO.

3. DATA

Length-frequency and nominal CPUE data for bigeye tuna from the Japanese longline fishery and the international purse-seine fishery on floating-objects were used in this analysis. The different data sets are described in detail below.

3.1. Length-frequency

3.1.1. Longline

Length-frequency data from the Japanese longline fishery from 70°-160°W in 2002-2007 were used in this analysis. The data are counts of fish in 2-cm intervals (28-220cm), with a resolution of 5°-latitude by 10°-longitude by month. To be consistent with the stock assessment, fish were ‘grown’/‘shrunk’ to the middle month of each quarter by adding (subtracting) a monthly length increment, where appropriate. The monthly increments were derived from a Von Bertalanffy growth model. The 2-cm counts of each 5°-latitude by 10°-longitude by month were then grouped into larger length intervals and the proportion of fish per interval computed. Length intervals were chosen to be small enough to capture the structure in the data, but large enough to avoid a dominance of zero values in any given interval across samples, and to avoid selecting interval boundaries near length gaps created by growing/shrinking fish. Length intervals were: 0-86cm; 88-100cm; 102-114cm; 116-128cm; 130-142cm; 144-156cm; 158-170cm; 172-186cm; >186cm. Data prior to 2002 were not used because most of the available data were at a fairly coarse spatial resolution (10°-latitude by 20°-longitude).

3.1.2. Purse-seine

Purse-seine length-frequency samples from floating-object-set wells (size-class 6 vessels only) from the EPO in 1975-2008 were used in this analysis. Details of sample collection are described by Suter (2008) and references therein. The data set used was dominated by samples from 2000 (1,839 samples from 2000-2008 compared to 519 samples from 1975-1999). The analyses were conducted separately on data from 1975-1999 and 2000-2008 because of changes to the port-sampling protocol that were implemented beginning in 2000. The length-frequency data are counts of fish in 1-cm intervals, with a resolution of 5° latitude by 5° longitude by month. Length-frequency samples were first raised to the well catch in order to accommodate any samples collected from weight-sorted unloadings. As was done with the longline data, fish were then ‘grown’/‘shrunk’ to the middle month of each quarter. The 1-cm counts of each sample were then grouped into larger length intervals and the proportion of fish per interval computed. Length intervals were: 0-44cm, 45-54cm, 55-64cm, …, 125-134cm, and ≥ 135cm.

3.2. CPUE

3.2.1. Longline

Catch and effort (number of hooks) data from the Japanese longline fishery of the Pacific Ocean (105°E to 70°W) in 1975-2007 were used in the analysis. The data are raised monthly catches of bigeye tuna and numbers of hooks by 5° square area. For each quarter by 5° square area, this gives a times series of nominal CPUE (*i.e.*, unstandardized catch-per-hook), based on a maximum of 99 points (3 months x 33 years).

3.2.2. Purse-seine

Catch and effort (number of days fishing) for the purse-seine fishery on floating-objects of the EPO in 1993-2008 were used in the analysis. The data are nominal monthly catch-per-day fishing in each 5° square area for size-class 6 vessels. For each quarter by 5° square area, this gives a time series of up to 48 points (3 months x 16 years). Data prior to 1993 were not used because of the development of the fishery on fish-aggregating devices in the early 1990's.

4. METHODS

Multivariate regression trees (*e.g.*, De'ath, 2002), an extension of the classical regression tree approach (Breiman *et al*., 1984), were used to analyze the two types of data. The spatial and temporal partitions of the data ('splits') identified with this approach provide information about spatial and temporal structure. The methods are described in detail below.

4.1. Length-frequency

The multivariate regression tree approach for the length-frequency data used an impurity-based implementation of the Kullback-Leibler divergence (KLD; *e.g.*, Wang *et al*., 2005) to measure node heterogeneity ('impurity') instead of the classical squared error loss function. The KLD is a commonly used measure for quantifying the difference between distributions. Two different analyses were undertaken: a 'basic' analysis, and a 'within-year' analysis. The basic method follows that outlined by Lennert-Cody *et al*. (2010), with the KLD-based measure of impurity of a collection of samples (*e.g.*, on the node of a tree) given by:

$$I_{KLD} = \sum_{i=1}^{n} \sum_{j=1}^{m} p_i(j) \log\left(\frac{p_i(j)}{\overline{p}(j)} \right),$$

where $p_i(j)$ is the proportion of fish in the j^{th} length interval (j=1,..,m) of the i^{th} sample (i=1,…,n), and $\overline{p}(j)$ is the average proportion of fish in the j^{th} length interval (average computed over samples in the collection).

For the within-year analyses, the impurity measure above was modified to measure differences within each year:

$$I_{KLD\_within\_year} = \sum_{l=1}^{y} \sum_{k=1}^{n_y} \sum_{j=1}^{m} p_{lk}(j) \log\left(\frac{p_{lk}(j)}{\overline{p}_l(j)} \right),$$

where $p_{lk}(j)$ is the proportion of fish in the j^{th} length interval of the k^{th} sample of the l^{th} year, and $\overline{p}_l(j)$ is the average proportion of fish in the j^{th} length interval of the l^{th} year, for y years and n_y samples per year. This measure of impurity may be less susceptible to the influence of unusual events that occurred in only a few years (*e.g.* strong recruitment).

Three predictors were used in the two regression tree analyses: quarter, 5° latitude, and 5°- or 10°-longitude (depending on the data source). All predictors were treated as numeric. The multivariate response was the proportion of fish in the binned length intervals. The data unit of the analysis was a month-area (5° latitude by 5° longitude or 5° latitude by 10° longitude). Basic trees were trimmed using the '1-se' rule (Breiman *et al*., 1984), and within-year trees were built to the size of the corresponding 1-se basic tree.

The following sensitivity analyses were undertaken for the longline length-frequency data. Both basic and within-year trees were grown with different restrictions on the minimum number of samples required on terminal nodes (5, 10, and 20 samples). In addition, for basic trees, two hundred resampled data sets were created and trees were grown on these resampled data sets. Finally, basic trees were also built on data not grown/shrunk to the quarter, for all years combined and separately for individual years. For these trees, the response was the proportion of fish in each of 12 length intervals: 0-86 cm, 88-96 cm, 98-106 cm, …., 168-176 cm, 178-186 cm, and ≥ 188 cm. The predictors used were month, 5°-latitude, and 10°-longitude.

For purse-seine data, the following sensitivity analyses were undertaken. Basic and within-year trees were grown on years grouped into an El Niño category (years 1982-83, 1987, 1992-93, and 1997-98) and a non-ENSO category (years 1978-81, 1985, 1994-96, 2001, 2005, and 2007). Basic trees were also built on individual years for 2000-2008. Within-year trees for 2000-2008 were grown with different restrictions on the minimum number of samples required on terminal nodes.

4.2. CPUE

For each quarter-5° square area with sufficient data, smooth trends in nominal CPUE were estimated using penalized cubic regression splines (Wood, 2006). Quarter-5° square areas were considered to have sufficient data if they contained at least 50 data points over at least 25 years and if the total catch was at least 0.01% of that for the entire data set. The same basis definition, basis dimension (6 basis functions), knot locations and smoothing parameter were used for each quarter-5° square area. Because sample size and data spacing varied by quarter-5° square area, the common smoothing parameter was selected in the following manner. Preliminary smooth trends were first fitted to each quarter-5° square area with the smoothing parameter of each quarter-5° square area estimated by generalized cross-validation (GCV). The median of these individual smoothing parameter estimates (a value of 8.35) was then used as the common smoothing parameter to obtain final estimates of the smooth trends. The vector of estimated spline coefficients obtained for the final smooth trend of each quarter-5° square area, excluding that of the constant term, was used as the response variable to a multivariate regression tree (*c.f.* Yu and Lambert, 1999). Impurity was measured with the classical squared error loss function. Three predictors were used: quarter, 5° latitude, and 5° longitude. To assess the stability of the results, the analysis was repeated using several different values for the common smoothing parameter.

5. RESULTS AND DISCUSSION

Preliminary results suggest that spatial pattern often dominated over seasonal pattern. For this reason, the following presentation of results focuses on spatial pattern.

5.1. Length-frequency

5.1.1. Longline

Analysis of the length-frequency data indicate both latitudinal and longitudinal structure associated with the length distributions of the sampled catch. Data pooled over years and quarters show a predominance of small- to moderate-sized fish caught north of ~10°N, more medium- to large-sized fish caught to about 15°S, and mostly medium-sized fish in the catch further south (Figure 2). The basic 1-se tree and the comparably sized within-year tree are shown in Figures 3-4. Both analyses indicate spatial splits at 0° and 15°S, as well as an offshore partition at the western edge of the EPO (~140°-150°W; Figure 5). The split in the basic 1-se tree at 10°N was also found in the second most common tree structure of the resampled data (Figure 6). The reason this split was not found by the within-year analysis may be due to the fact that the split appears to reflect pattern present in some years, but not in others. Of the annual basic trees built on data not grown/shrunk to the quarter, only 2002 and 2003 had the first split at 10°N. The inshore-offshore longitudinal split in the within-year tree at 100°W (Figure 4) was supported to some extent by the second most common basic tree on the resampled data (Figure 6) which had a split at 90°W.

Basic 1-se trees built for all years on samples without growing/shrinking fish (predictors: month, latitude, longitude) were similar to those of built on samples grown/shrunk to the quarter, and are not shown. Increasing the minimum number of samples required on a terminal node had no affect on the dominate splits in either tree analysis, but in some instances did result in changes at the lower branches.

5.1.2. Purse-seine

In contrast to similarities between results of the two tree analyses for the longline length-frequency data, the two tree approaches yield generally different structure when applied to the purse-seine length-frequency data. Data pooled of years and quarters show an overall inshore gradient in the size of the sampled catch from small fish in the north and offshore to medium to large fish in the south and inshore (Figure 7). The basic 1-se tree and the comparably-sized within-year tree are shown in Figures 8-9. The first split in the trees built on data from 2000-2008 was at 110°W for both the basic and within-year trees, as well as for most of the annual trees (not shown). For 1975-1999, the 1-se basic tree had only one split at 95°W, while the main split of the within-year was at 100°W (not shown). The first splits of the El Niño and non-ENSO trees were at 105°W and 110°W, respectively (not shown). Thus, there seems to be reasonable support for an inshore-offshore partition around 100°-110°W.

Beyond this, results of the two types of tree analyses largely differ for 2000-2008 (Figures 8-10), possibly due to spatial-temporal confounding. (Results from analysis of data for 1975-1999 are not considered further because almost all 1-se basic trees involved only one split.) The within-year tree shows splitting west of 110°W (after first splitting on quarter), while more structure appears in the 1-se basic tree east of 110°W. Of the two main branches in the 1-se basic tree, the branch corresponding to inshore of 110°W would be selected by cross-validation over the one corresponding to west of 110°W. The within-year algorithm may be detecting more pattern to the west because data are present in most years for that area, increasingly so in later years. On the other hand, the basic algorithm detects more pattern east of 110°W because this algorithm is more likely to detect differences present in only a few years. For example, samples were only available east of ~90°W in some years; it appears that when samples existed from this nearshore area, the length-frequency structure was different from that of samples from sets made elsewhere.

Based on the 1-se rule for determining tree size, the spatial structure found for the purse-seine length-frequency data (Figure 10) is largely different from that found for the longline length-frequency data (Figure 5). This may be due in part to differences in the spatial resolution of the longline and purse-seine data and the fact that the latitudinal range of the purse-seine data was less than that of the longline data, affecting detection of spatial structure at lower and higher latitudes, particularly to the north.

5.2. CPUE

5.2.1. Longline

Large-scale patterns in CPUE trends are clearly evident both across the Pacific and within the EPO (Figure 11). West of ~165°W, trends were often increasing or flat, while east of ~150°W, trends were often decreasing or dome-shaped. In addition, variability in CPUE was not constant spatially. In particular, within the EPO, CPUE was quite variable in the inshore areas (*e.g.*, east of ~110°W; north of 15°N). The 1-se tree for the spline coefficient vectors is shown in Figure 12. The first split divides the Pacific at 150°W. A map of the spatial splits from this tree within the EPO is shown in Figure 13. Within the EPO, some of the spatial structure identified for CPUE trends is consistent with that found for the length-frequency data (Figure 14). For example, all three tree analyses (basic and within-year length-frequency, CPUE trends) indicate a longitudinal split at the equator, over part or all of the EPO. A latitudinal split that divides the EPO north of the equator at ~10-15°N is supported by the CPUE trends tree and the basic length-frequency tree. A longitudinal split of the EPO that creates an inshore region at ~95-100°W is supported by the with-in year length-frequency tree and, in part, by the CPUE trends tree. In addition, a longitudinal split in the northern part of the EPO at ~140°W is supported by the with-in year length frequency tree and the CPUE trends tree. On the other hand, there is less agreement among analyses on the location of a longitudinal split between ~100°-140°W.

Based on the spatial structure identified for the two data types, an example of a composite stratification was constructed (Figure 15). This subjective choice of partitions was obtained by giving more weight (consideration) to the same partitions that occurred in multiple analyses, spatially averaging the location of similar partitions across main analyses and sensitivity analyses (*e.g.*, the split at 95°W), and giving consideration to lower-level splits that might belong to only one analyses but were also present in sensitivity analyses (*e.g.*, the split at 120°W between the equator and 15°N). In this example stratification, the split at 95°W was not extended north of the equator because, having given priority to a partition at the equator, creating a small area north of the equator and east of 95°W could be problematic for the longline stock assessment; there is only limited fishing activity in this area. A more objective procedure for constructing a spatial stratification for stock assessment is being developed (see Section 6).

Decreasing the common smoothing parameter from 8.35 to 0.835 (~ the 0.27th quantile of the individual quarter-5° square area GCV-selected smoothing parameters) did not have much affect on the dominant splits. Fairly similar results for the dominant splits were also obtained when the smoothing parameter was increased to 35.1 (the 0.65th quantile of the individual quarter-5° square area GCV-selected smoothing parameters). Not surprisingly, as the smoothing parameter was further increased (forcing even more smoothing of the trends), splits below that at ~150°W were somewhat different. Further sensitivity analyses and procedures for smoothing parameter selection are being explored.

5.2.2. Purse-seine

The time series of nominal CPUE for the purse-seine floating-object fishery appear to be more variable than nominal CPUE for the longline fishery. There was a greater percentage of zero-valued observations for the purse-seine CPUE time series than for the longline CPUE time series, and the data did not seem to be well-summarized by the spline smooths (Figure 16). The tree algorithm did not find much meaningful pattern in the coefficient vectors. The 1-se tree had only one split, which was on quarter (quarters 1-3 *versus* quarter 4). These time series have not been analyzed further at this point.

6. IMPLICATIONS FOR THE STOCK ASSESSMENT

The goal of identifying spatial structure in the bigeye tuna fishery data is to improve the spatial stratification of the stock assessment. The stock assessment addresses spatial structure through either the stock dynamics (*e.g.* modeling sub-populations) or the fishery dynamics (*e.g.* defining fisheries with different selectivity and catchability by area). A general rule of thumb is that spatial differences in age or length structure requires modeling separate fisheries while spatial differences in CPUE requires modeling sub-populations. However, differences in CPUE could be due to differences in age/size structure. The size structure could differ among areas due to isolated populations with different biological characteristics or different localized depletion, size specific movement, or different fisher behavior (*e.g.* depth of longline gear).

The current stock assessment splits the longline fishery at 15°N (Figure 1) and this split (10°N or 15°N) is support by both the CPUE and length frequency data. In addition, there is support from both data sets for a split at the equator. The length frequency data also supports a split at 15°S and a split around 95°W. It is not clear if any of these splits support modeling sub-populations since the differences in CPUE may be due to differences in size/age. However, tagging data indicates low movement rates implying that differences in the length frequency and CPUE data is due to local processes and sub-stocks should be modeled.

The current assessment splits the purse seine floating object fisheries into four sub-fisheries (Figure 1). The length frequency data generally supports these four areas, particularly the splits surrounding the central area. However, due to the narrow latitudinal range of the data and the seasonal nature of structure west of 110°W it is difficult to evaluate the other area splits. A split at 125°W might be worth considering. The CPUE data does not provide any information on spatial structure. The current purse seine fishery definitions and those suggested by the data are generally different from the splits suggested for longline and it is therefore difficult to determine sub-populations that are consistent with the two datasets.

7. FUTURE WORK

An analytical procedure for performing a tree-based analysis simultaneously on length-frequency and CPUE data might lead to a means of obtaining a spatial stratification that was sensible for both data types. In as much as the types of tree algorithms use herein do not revisit partitions, such a procedure would require combining information from candidate splits for length-frequency distribution and CPUE trends into one impurity measure at each step in building the tree. For example, a combined impurity measure might take the following form:

$$I_{pooled}(s) = \gamma \frac{I_{L-F}(s)}{I_{L-F-m}} + (1-\gamma)\frac{I_{CPUE}(s)}{I_{CPUE-m}},$$

where $I_{L\text{-}F}(s)$ is the impurity from the length-frequency data at the current candidate split s (*e.g.*, based on the KLD), $I_{L\text{-}F\text{-}m}$ is the impurity for the best length-frequency tree at the current tree size of m when the length-frequency data were analyzed separately, $I_{CPUE}(s)$ is the impurity from the CPUE trends at the current split s (*e.g.*, based on the prediction error), $I_{CPUE\text{-}m}$ is the impurity for the best CPUE trends tree at the current tree size of m when the CPUE data were analyzed separately, and γ is a parameter between [0,1] which can be used to give added weight to one or the other of the two data types. γ might be defined based on data quality or based on a measure of stock assessment performance.

Additional improvements to the methods that are also underway include using prediction error as the loss function for the analysis of trends in CPUE (*i.e*., fully implementing the method of Yu and Lambert, 1999). Other methods for selecting a common smoothing parameter are also being considered. This is an important aspect of the analysis because there can be different numbers of data points by quarter-5° square area, and the spacing between data points can be irregular. Methods will also be adapted to categorical predictors. In addition, other methods for determining the optimal tree size (other than 1-se-type rules) will be considered. If estimates of error for the longline catch and effort data were made available, those could be incorporated into the spline algorithm went estimating trends in CPUE.

REFERENCES

Aires-da-Silva, A. and Maunder, M.N. 2009. Status of bigeye tuna in the eastern Pacific Ocean in 2007 and outlook for the future. Inter-American Tropical Tuna Commission Stock Assessment Report 9.

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

De'ath, G. 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. Ecology 83(4): 1105-1117.

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

Schaefer, K. 2009. Stock structure of bigeye, yellowfin and skipjack tunas in the eastern Pacific Ocean. Inter-American Tropical Tuna Commission Stock Assessment Report 9.

Suter, J.M. 2008. An evaluation of the area stratification used for sampling tunas in the eastern Pacific Ocean and implications for estimating total annual catches. Thesis for Master of Science in Statistics, San Diego State University, San Diego, California, U.S.A.

Tomlinson, P.K. 2004. Sampling the tuna catch of the eastern Pacific Ocean for species composition and length-frequency distributions. Inter-American Tropical Tuna Commission Stock Assessment Report 4.

Wang, Q., Kulkarnia, S.R. and Verdú, S. 2005 Divergence estimation of continuous distributions based on data-dependent partitions. IEEE Transactions on Information Theory 51, 3064-3074.

Watters, G. and Deriso, R. 2000. Catch per unit effort of bigeye tuna: A new analysis with regression trees and simulated annealing. Inter-American Tropical Tuna Commission Bulletin 21(8): 531-571.

Wood, S.N. 2006. *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC.

Yu, Y. and Lambert, D. 1999. Fitting trees to functional data, with an application to time-of-day patterns. Journal of Computational and Graphical Statistics 8 (4): 749-762.

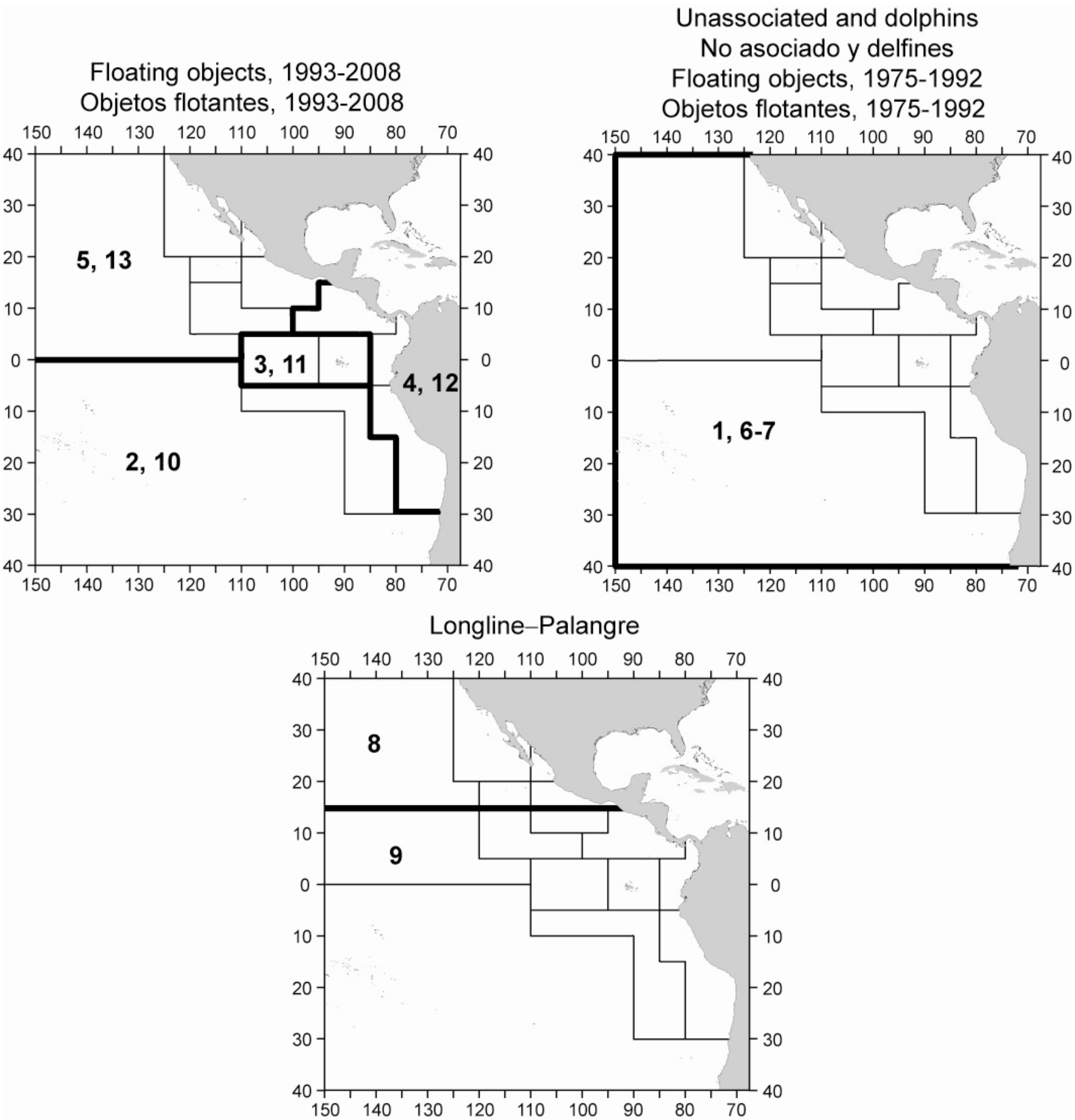


FIGURE 1. Areas used for stock assessment and sampling in the EPO (from Aires-da-Silva and Maunder, 2009). Thick black lines indicate the boundaries of the stock assessment areas; thin black lines indicate the boundaries of the areas used for collection of port-sampling data.

DRAFT

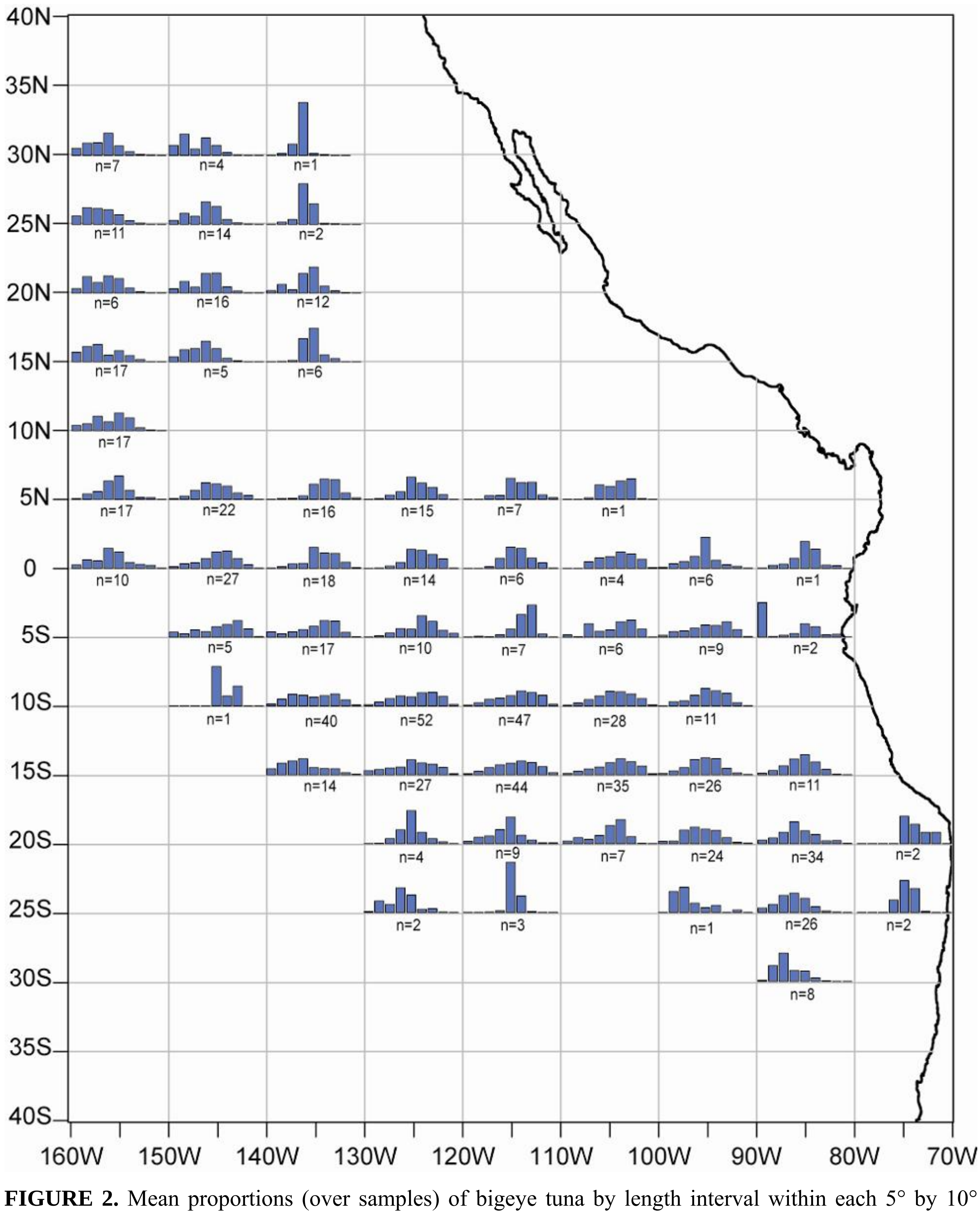


FIGURE 2. Mean proportions (over samples) of bigeye tuna by length interval within each 5° by 10° area (data pooled over quarters and years) for the longline fishery from 2002-2007. The number of samples (n) per 5° by 10° area is shown below each frequency distribution.

DRAFT

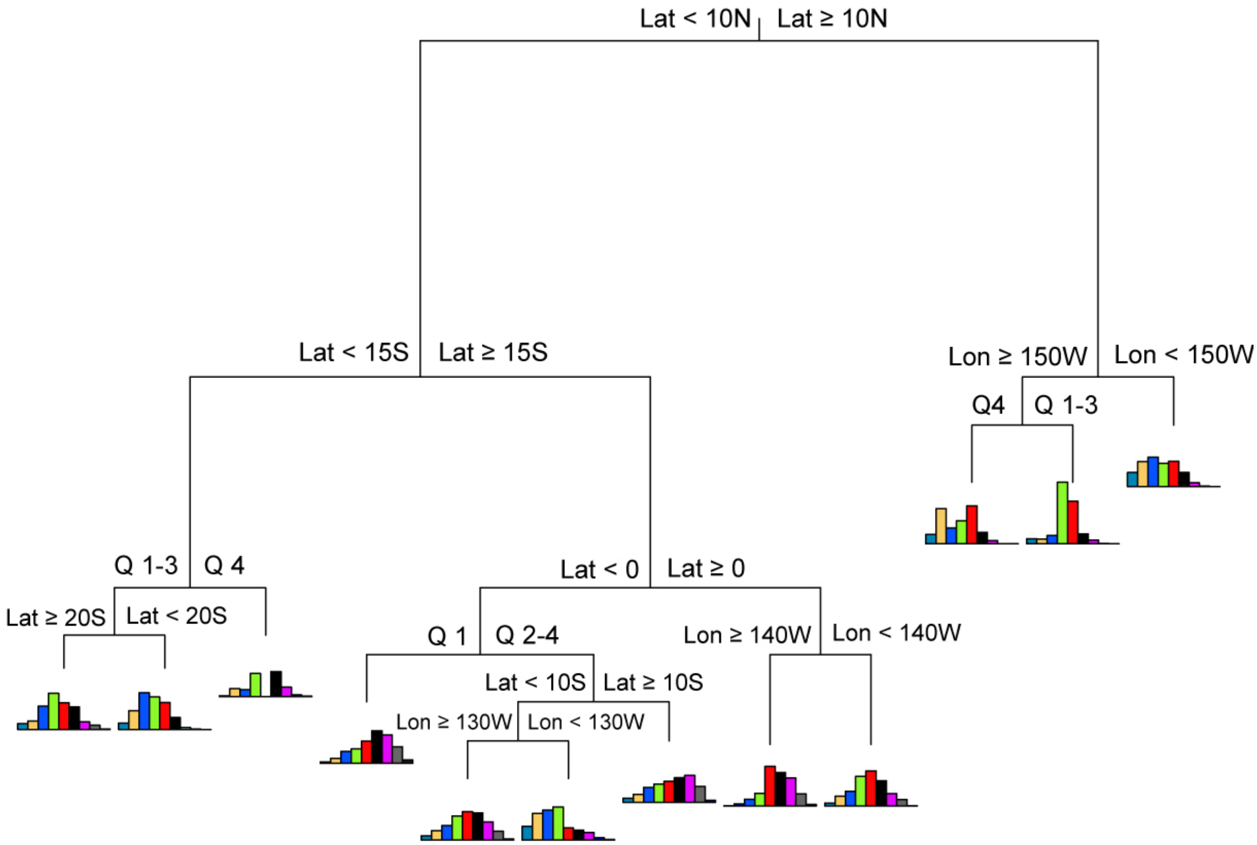


FIGURE 3. 1-se basic tree from longline length-frequency analysis. Predicted proportions in each length interval are shown at the terminal nodes. Length of branches in the tree is proportional to the error explained by the fit. ‘Lat’=latitude; ‘Lon’=longitude; ‘Q’=quarter. For longitudinal splits, ‘≥ ’ indicates longitudes eastward and ‘<’ indicates longitudes westward.

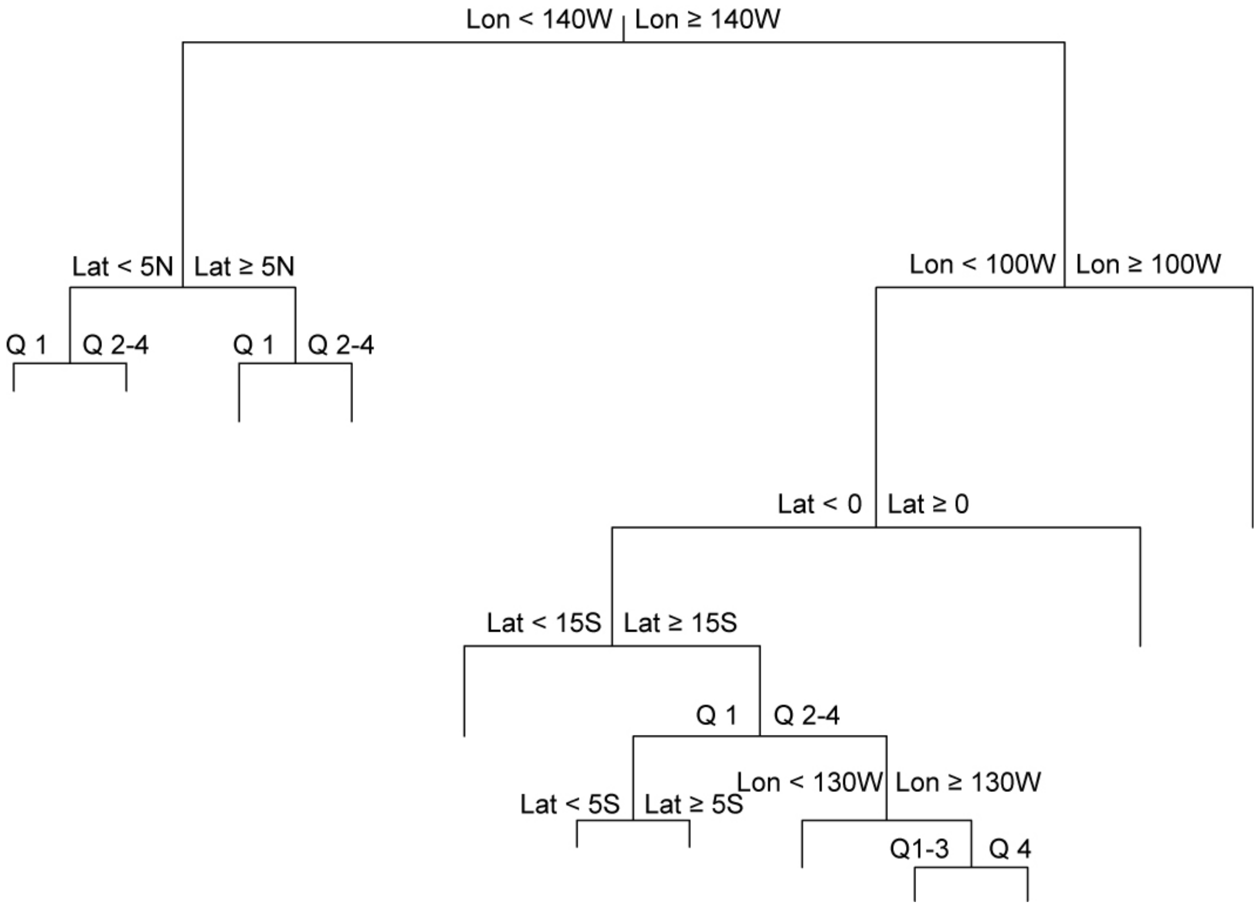


FIGURE 4. Within-year tree from the longline length-frequency analysis. Length of branches is proportional to the error explained by the fit. 'Lat'=latitude; 'Lon'=longitude; 'Q'=quarter. For longitudinal splits, '≥ ' indicates longitudes eastward and '<' indicates longitudes westward.

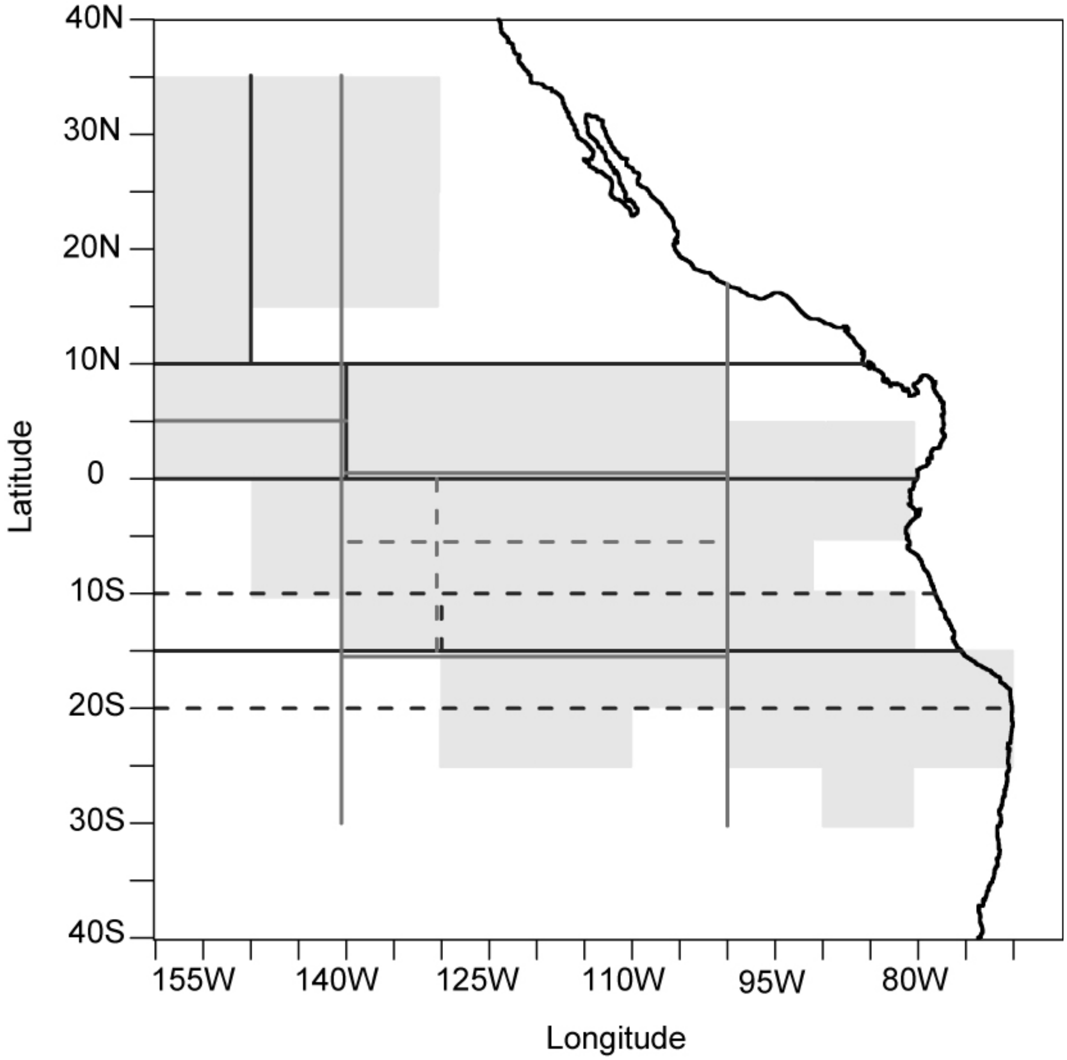


FIGURE 5. Map of spatial splits of trees shown in Figures 3-4. Dark gray = basic 1-se tree; light gray = within-year tree. Dashed lines indicate spatial splits that are seasonally-modified (see Figures 3-4). Light gray shading indicates the area of the EPO for which data were available (see also Figure 2). Spatial splits from Figures 3-4 were drawn to the maximum latitudinal and longitudinal extent of the data.

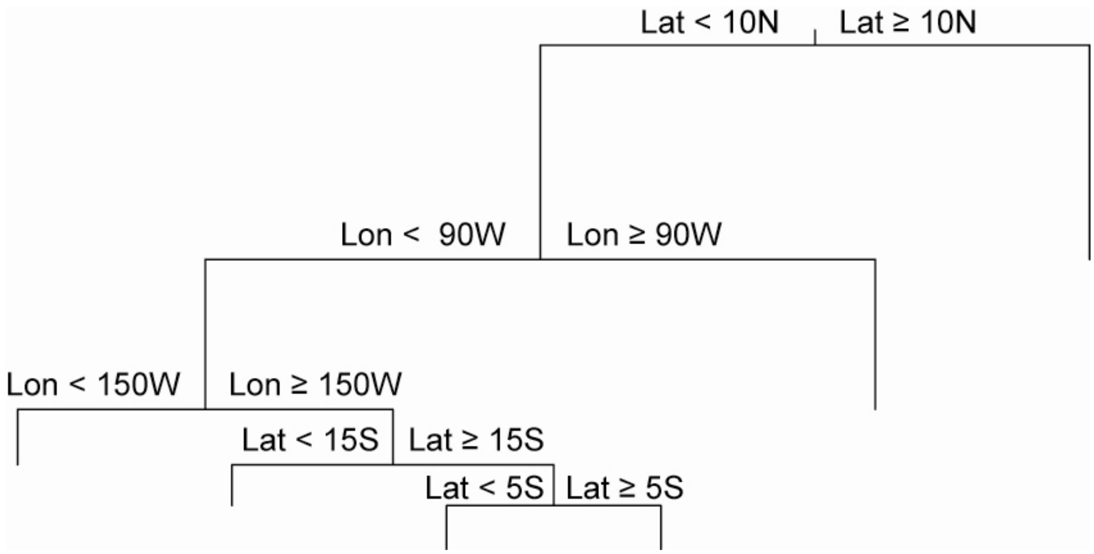


FIGURE 6. Example of second most common 6-split tree structure for the resampled longline length-frequency data. Length of branches in the tree is proportional to the error explained by the fit. 'Lat'=latitude; 'Lon'=longitude; 'Q'=quarter. For longitudinal splits, '≥ ' indicates longitudes eastward and '<' indicates longitudes westward.

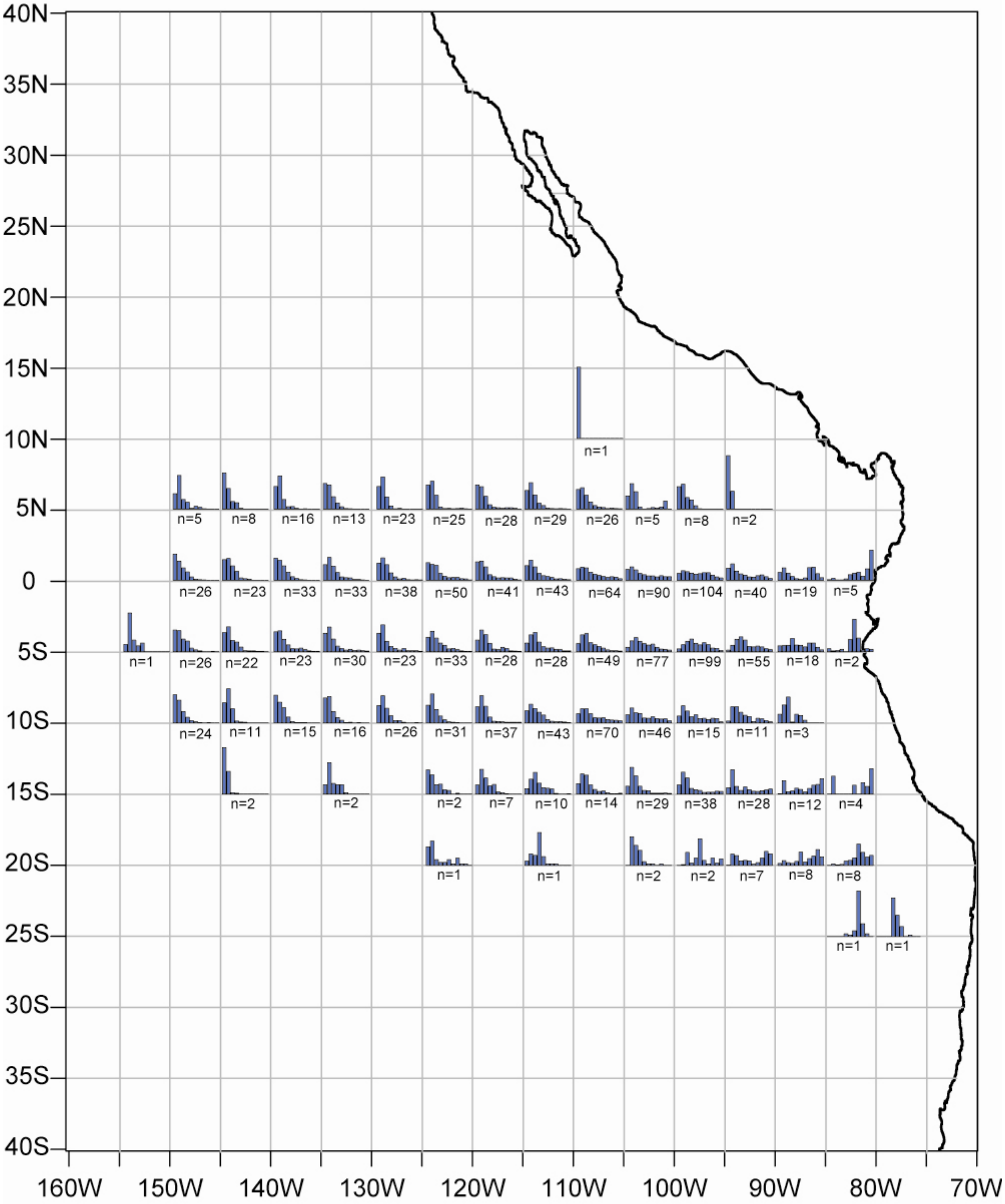


FIGURE 7. Mean proportions (over samples) of bigeye tuna by length interval within each 5° by 5° area (data pooled over quarters and years) for the purse-seine floating-object fishery for 2000-2008. The number of samples (n) per 5° by 5° area is shown below each frequency distribution.

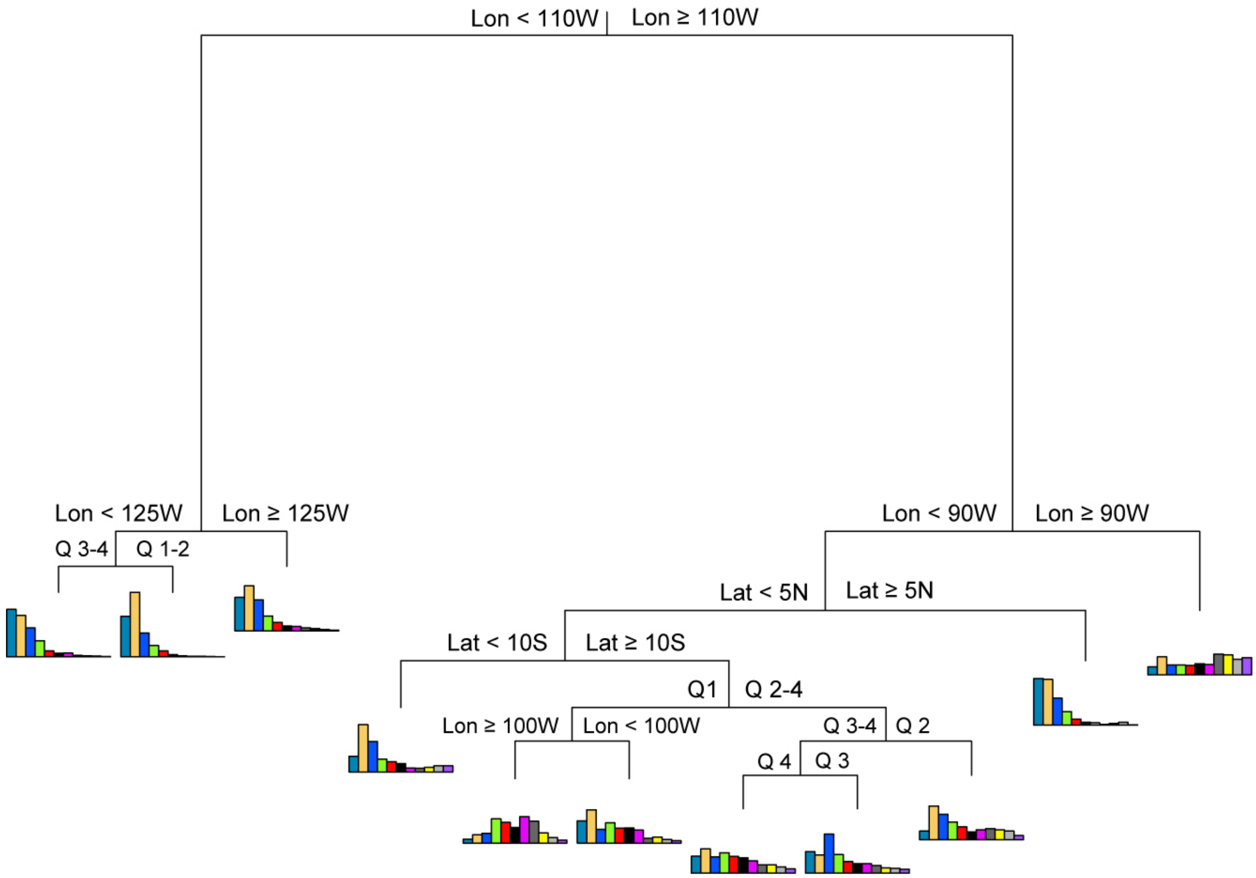


FIGURE 8. 1-se basic tree from purse-seine floating-object length-frequency analysis for 2000-2008. Predicted proportions in each length interval are shown at the terminal nodes. Length of branches in the tree is proportional to the error explained by the fit. 'Lat'=latitude; 'Lon'=longitude; 'Q'=quarter. For longitudinal splits, '≥ ' indicates longitudes eastward and '<' indicates longitudes westward.

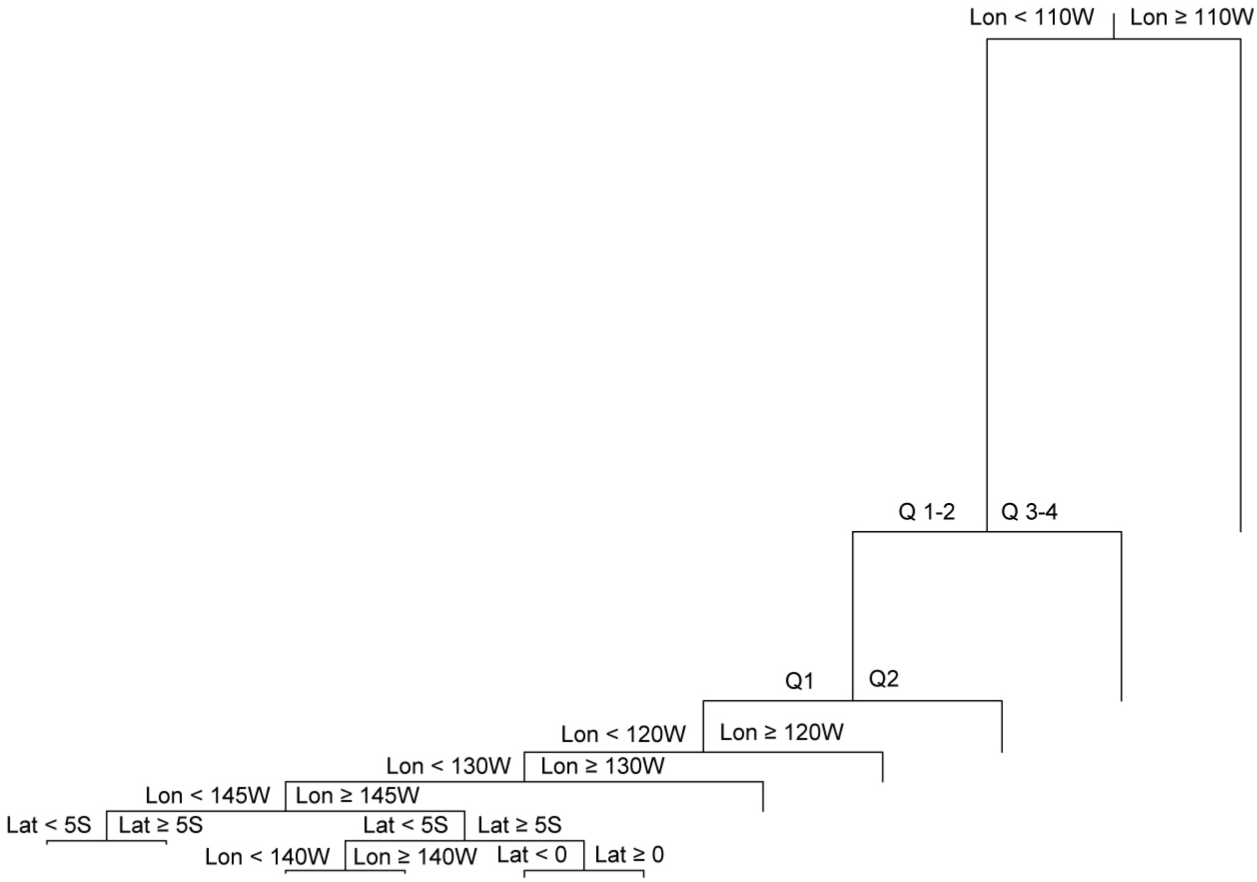


FIGURE 9. Within-year tree from the purse-seine floating-object length-frequency analysis for 2000-2008. Length of branches in the tree is proportional to the error explained by the fit. 'Lat'=latitude; 'Lon'=longitude; 'Q'=quarter. For longitudinal splits, '≥ ' indicates longitudes eastward and '<' indicates longitudes westward.

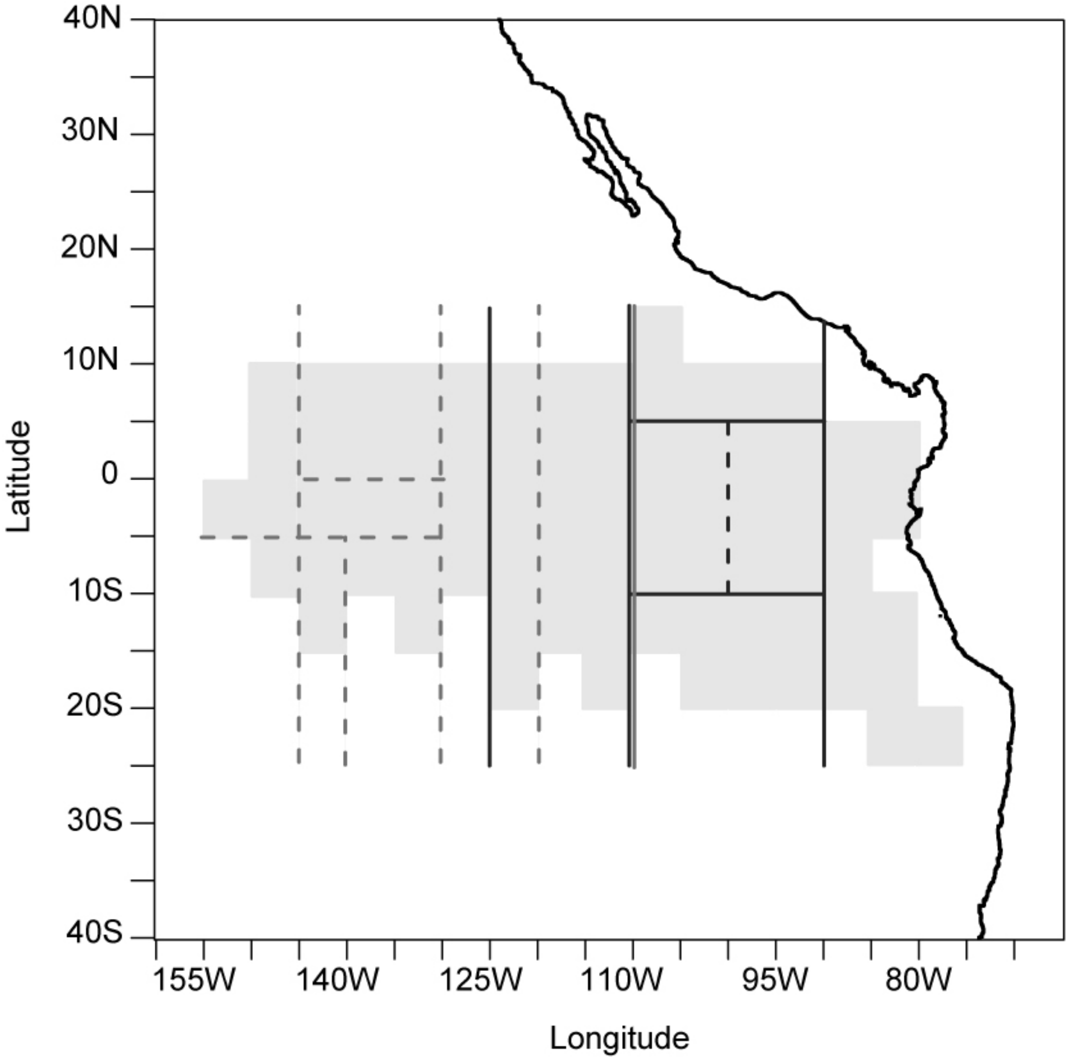


FIGURE 10. Map of spatial splits of trees shown in Figures 8-9. Dark gray = basic 1-se tree; light gray = within-year tree. Dashed lines indicate spatial splits that are seasonally-modified (see Figures 8-9). Light gray shading indicates the areas of the EPO for which data were available (see also Figure 7). Spatial splits from Figures 8-9 were drawn to the maximum latitudinal and longitudinal extent of the data.

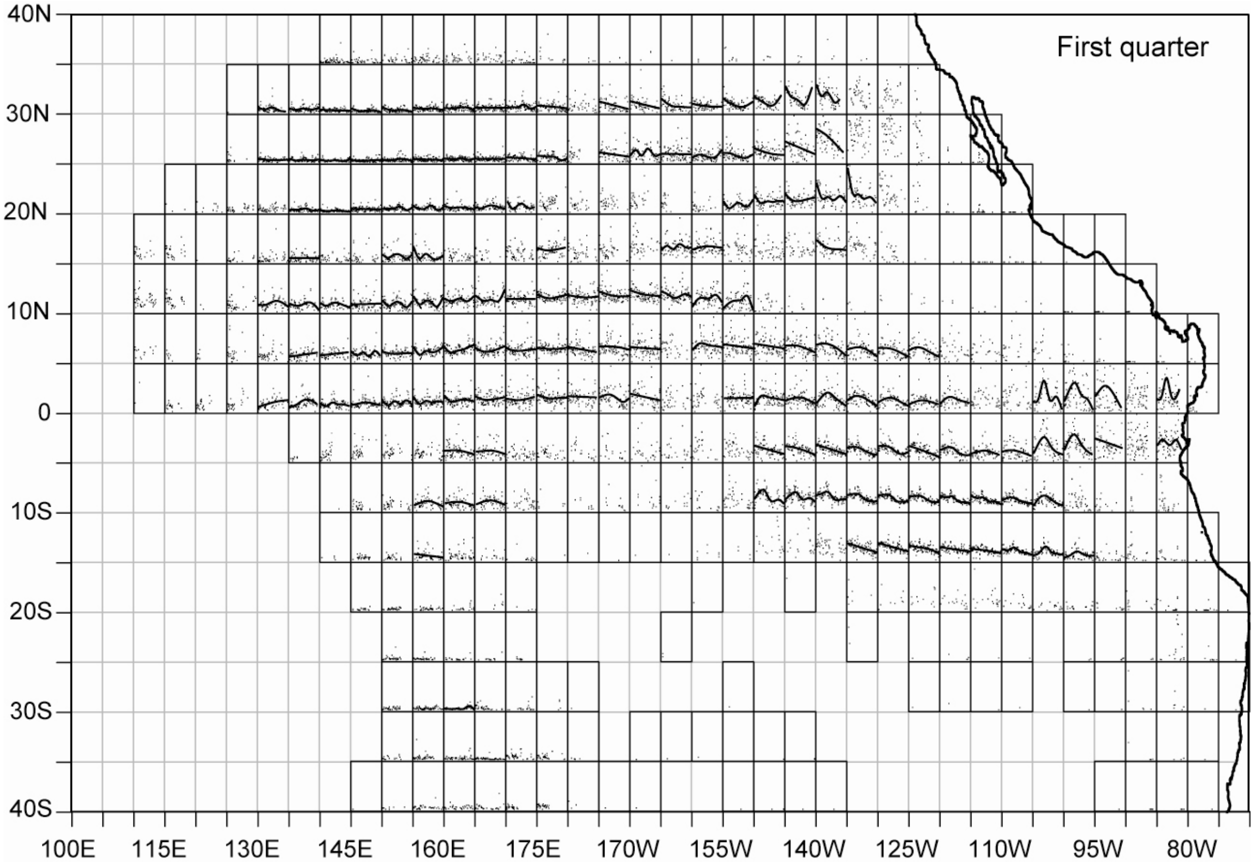


FIGURE 11. Maps of 5° square area smoothed trends of longline CPUE, by quarter. Each box in all maps has the same axes (x-axis = year; y-axis = CPUE) and ranges (year: 1975-2007; CPUE: 0-0.25*max(CPUE), maximum computed over all data). Points (black dots: monthly CPUE for a given year, within a 5° square area) were plotted for any 5° square area between 40°S and 40°N with data. If, in addition, the 5° square area had at least a minimum amount of data (≥ 50 data points, ≥ 25 years, ≥ 0.01% of the total catch), the data were also smoothed using penalized cubic regression splines (with smoothing parameter estimated by GCV within each quarter-5° square area) and the resulting smooth trends are indicated by black lines.

DRAFT

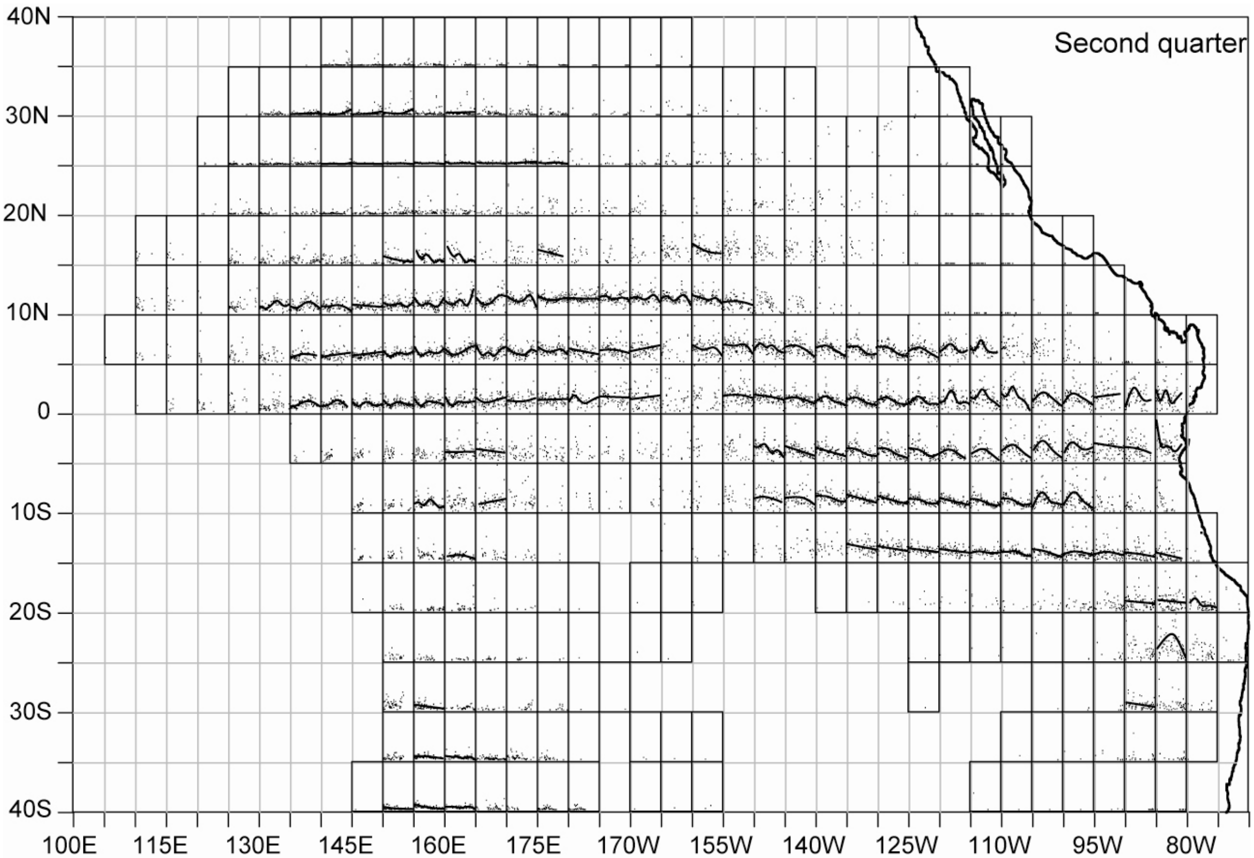


Figure 11. (continued)

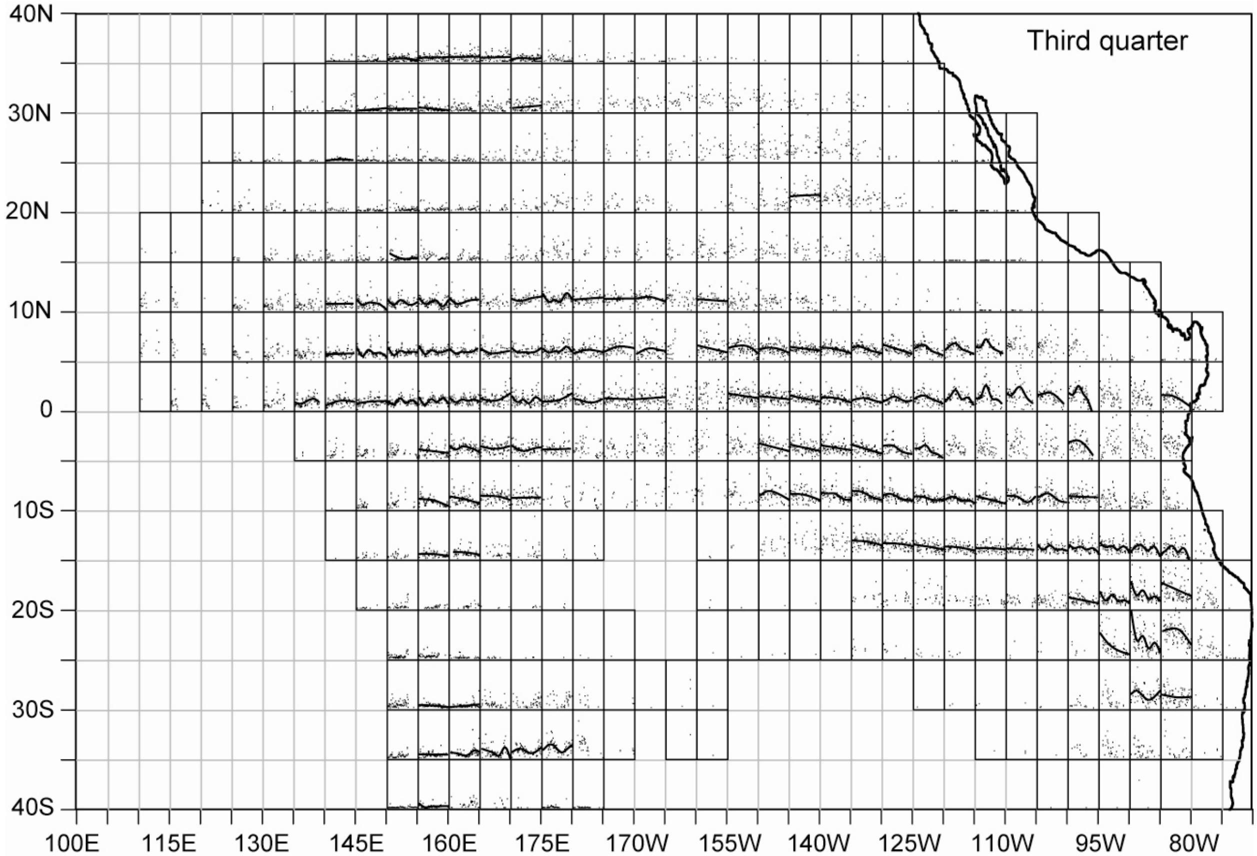


Figure 11. (continued)

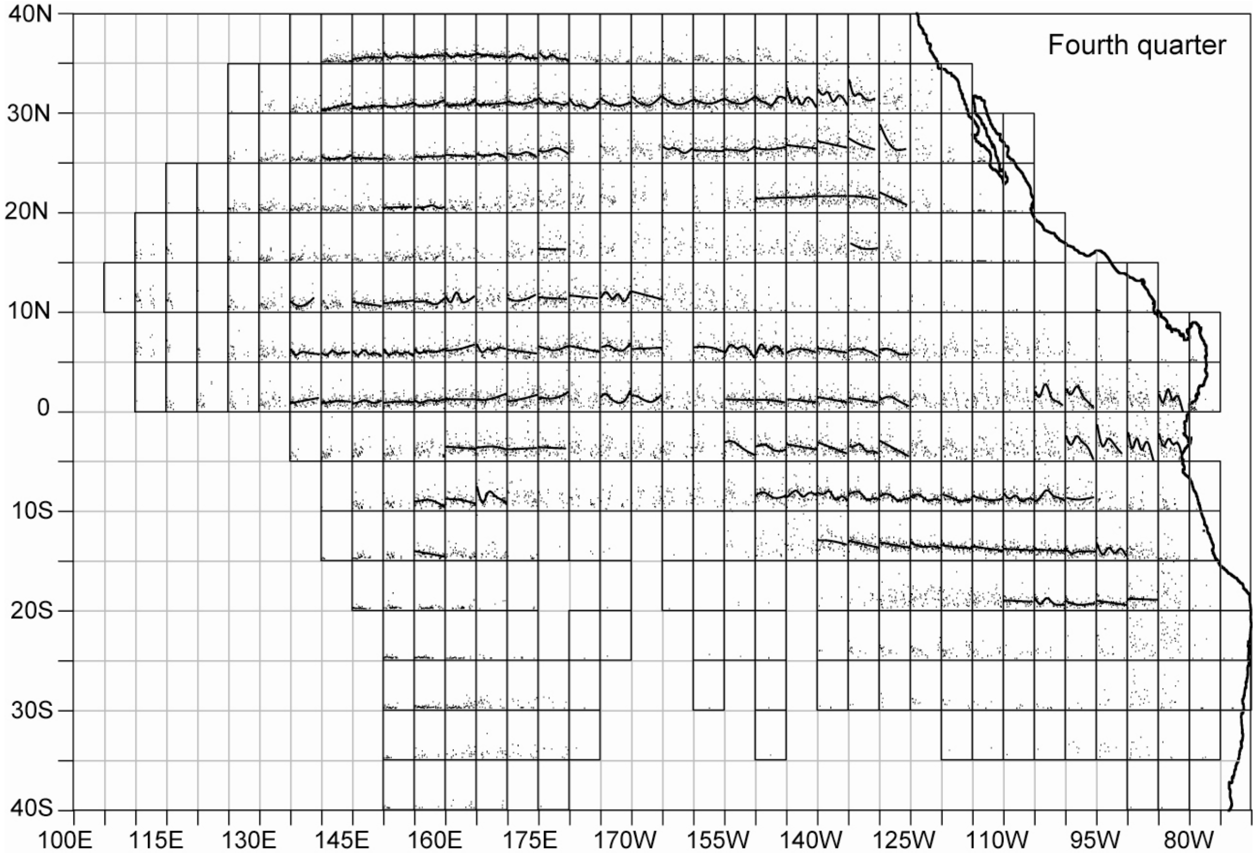


Figure 11. (continued)

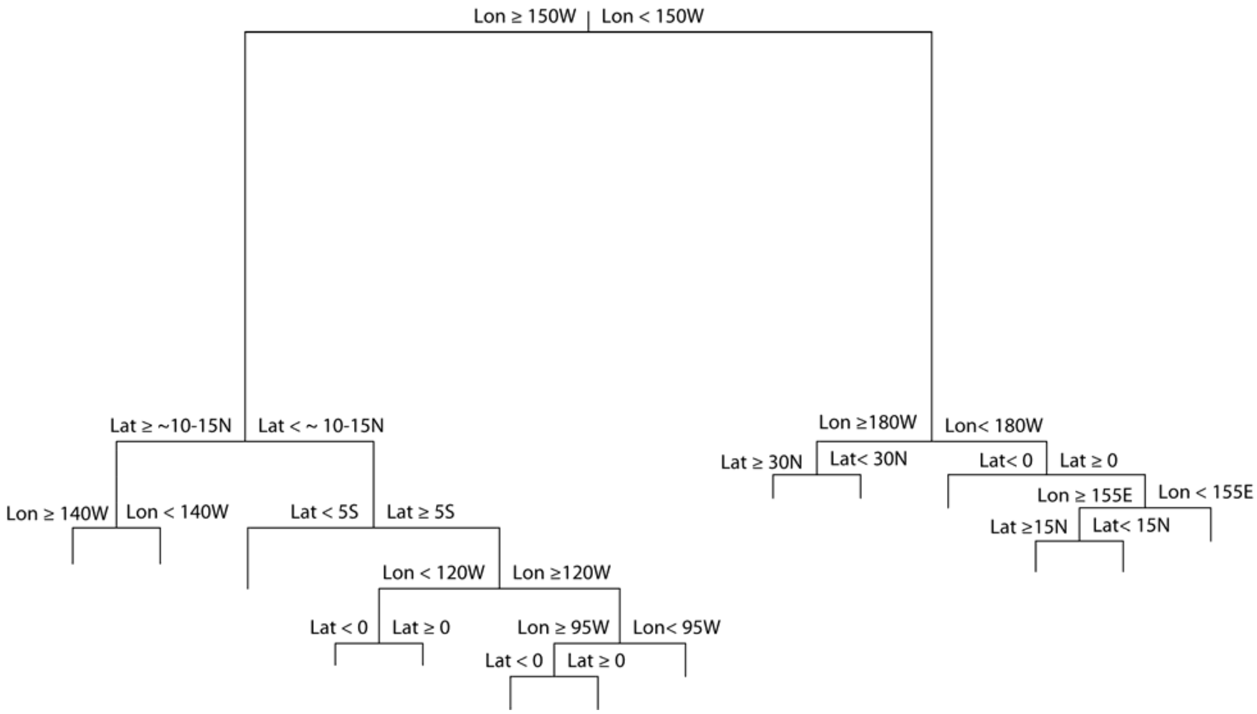
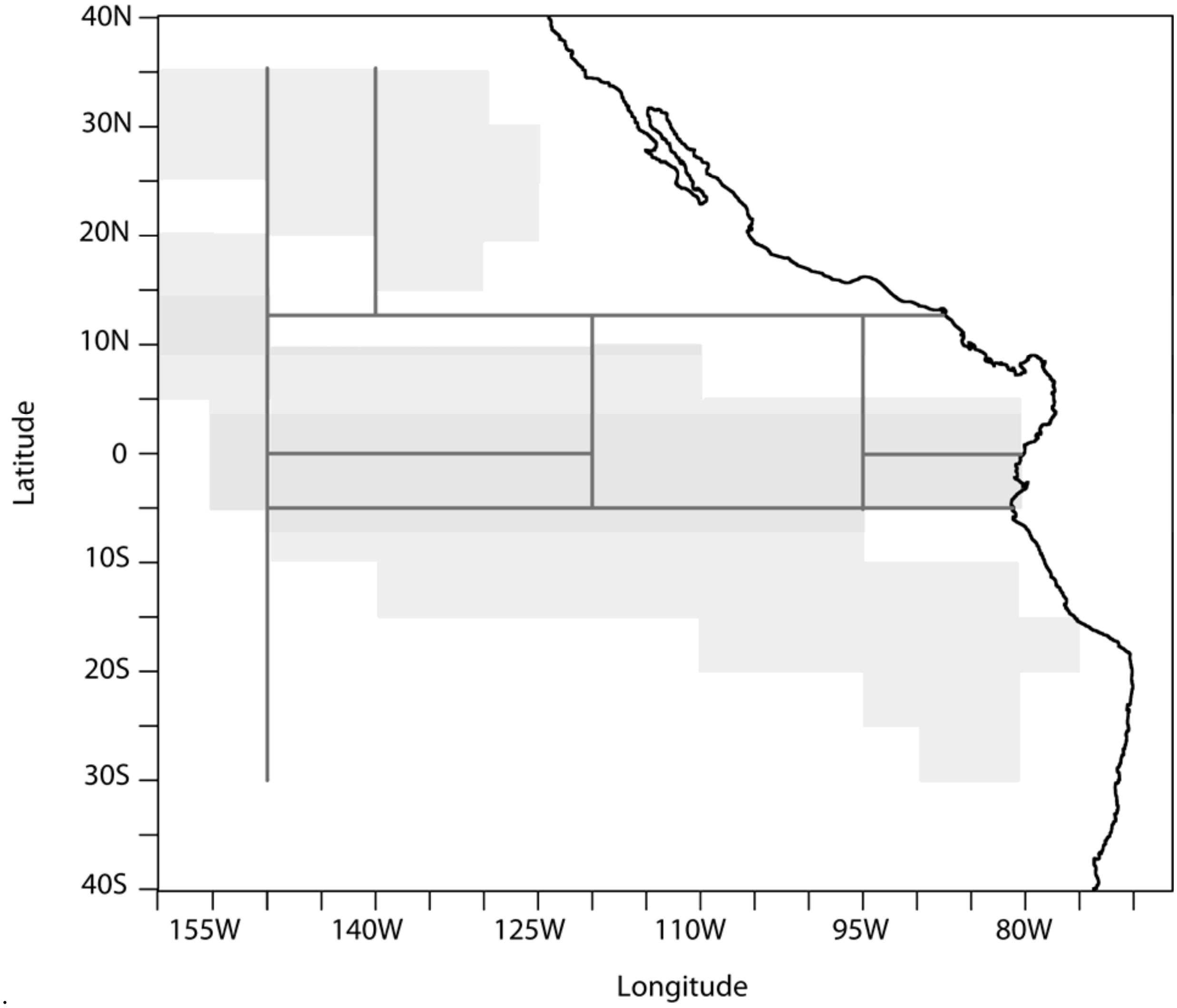


FIGURE 12. The 1-se tree for the longline CPUE trend coefficients (trends by quarter are shown in Figure 11). Length of branches in the tree is proportional to the error explained by the fit. 'Lat'=latitude; 'Lon'=longitude; 'Q'=quarter. For longitudinal splits, '≥ ' indicates longitudes eastward and '<' indicates longitudes westward. The first split east of 150°W is indicated as falling between 10°-15°N because of a data gap at those latitudes within the EPO (see Figure 13).



.

FIGURE 13. Map of spatial splits (EPO only) for 1-se CPUE trends tree shown in Figure 12. Light gray shading indicates the areas of the EPO for which data were available (see also Figure 11). Spatial splits from Figure 12 were drawn to the maximum latitudinal and longitudinal extent of the data. Note that the split along 12.5°N actually represents a split between 10°-15°N which falls within a data gap in the EPO at those latitudes.

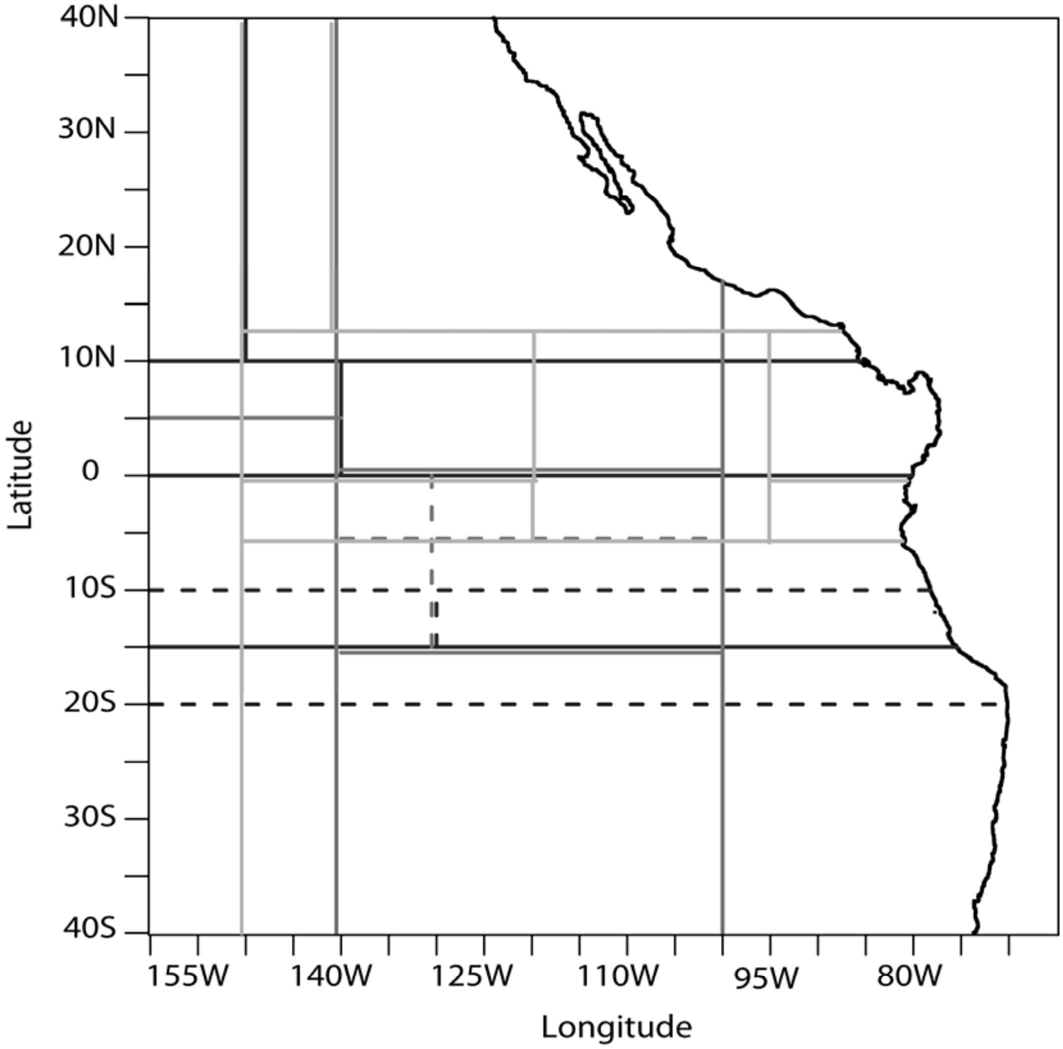


FIGURE 14. Composite of spatial splits shown in Figures 5 (longline length-frequency) and 13 (longline CPUE trends). Dark gray: 1-se basic length-frequency tree; medium gray: within-year length-frequency tree; light gray: 1-se longline CPUE trends tree.

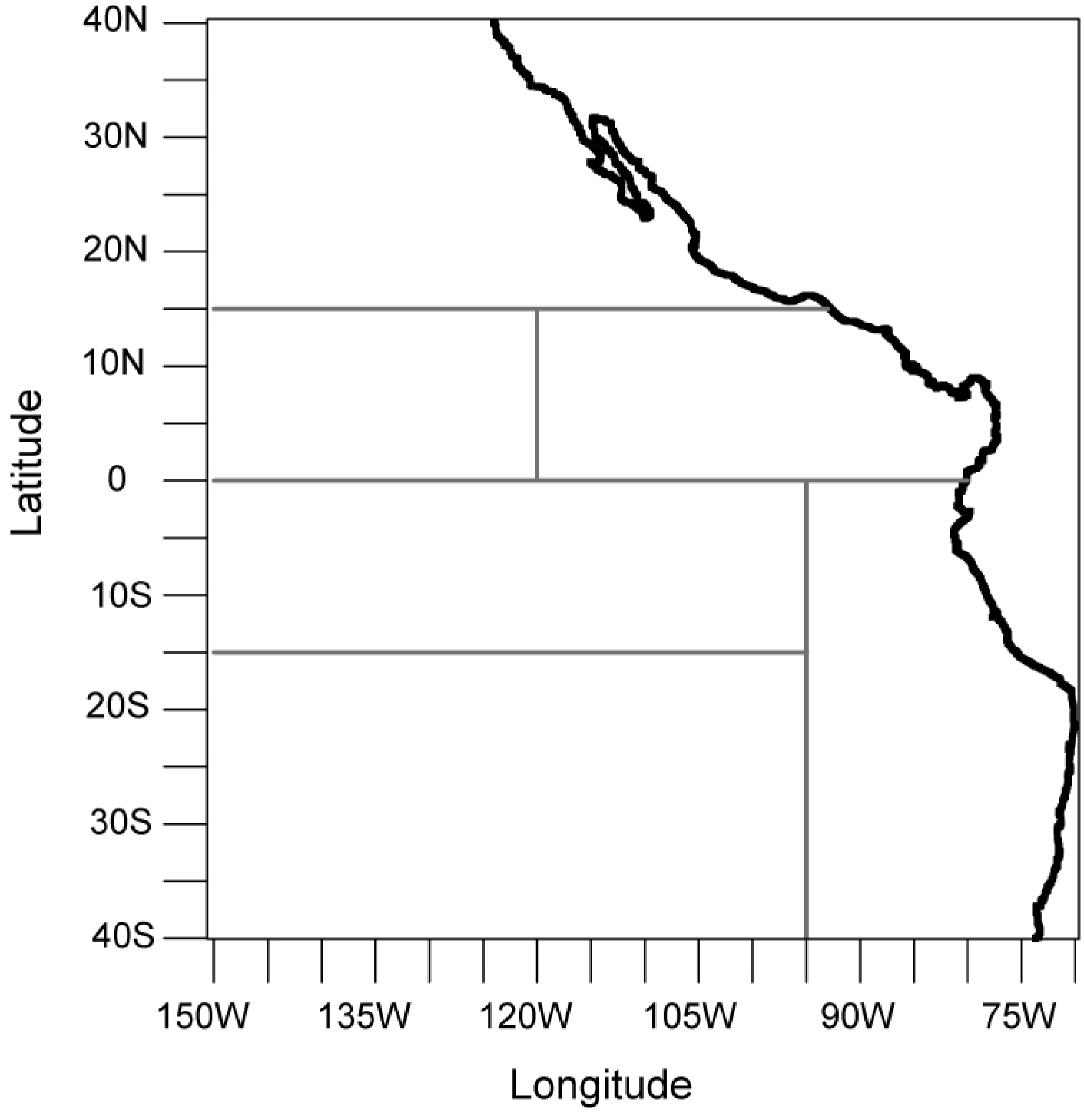


FIGURE 15. Subjective composite of spatial splits from tree analyses of longline length-frequency and nominal CPUE (Figure 14 and sensitivity analyses (not shown).

DRAFT

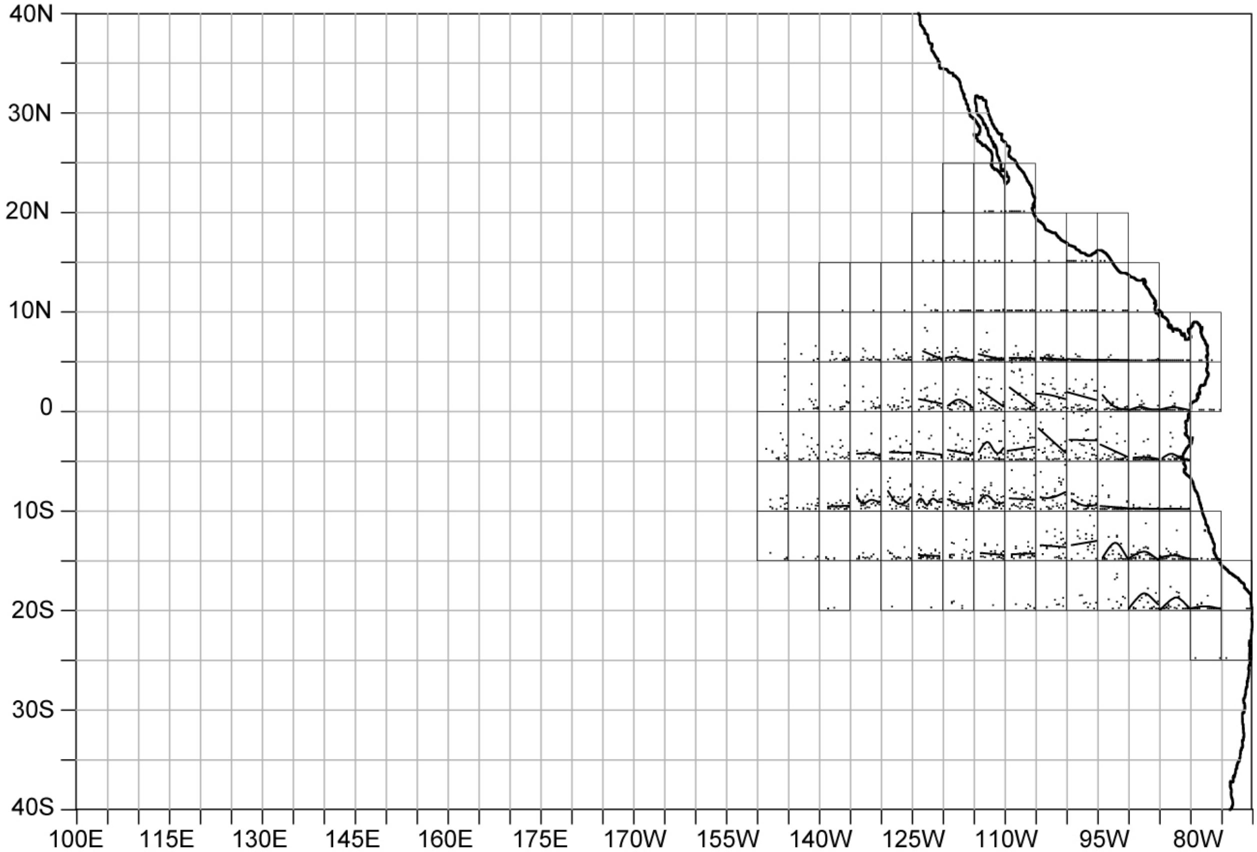


FIGURE 16. Map of 5° square area smoothed trends of purse-seine floating-object CPUE for the first quarter. Each box in the map has the same axes (x-axis = year; y-axis = CPUE) and ranges (year: 1993-2007; CPUE: 0-0.25*max(CPUE), maximum taken over all data). Points (black dots: monthly CPUE for a given year, within a 5° square area) were plotted for any 5° square area between 40°S and 40°N with data. If, in addition, the 5° square area had at least a minimum amount of data (≥ 20 data points, ≥ 10 years, ≥ 0.01% of the total catch), the data were also smoothed using penalized cubic regression splines (with smoothing parameter estimated by GCV within each quarter-5° square area) and the resulting smooth trends are indicated by black lines.