DRAFT

INTER-AMERICAN TROPICAL TUNA COMMISSION

EXTERNAL REVIEW OF IATTC YELLOWFIN TUNA ASSESSMENT

La Jolla, California (USA)
15-19 October 2012

DOCUMENT YFT-01-03

POSTSTRATIFICATION OF PURSE-SEINE PORT-SAMPLING DATA FROM DOLPHIN SETS

Cleridy E. Lennert-Cody, Mark N. Maunder, Alexandre Aires-da-Silva, Alejandro Pérez, JoyDeLee Marrow

CONTENTS

1. SUMMARY

This document summarizes further progress on development of a poststratification approach to the estimation of total tuna catch by species for the purse-seine fishery on tunas associated with dolphins. Previous analyses (Lennert-Cody *et al*. 2012) used linear and generalized linear models to study spatial-temporal variability in the average weight of yellowfin tuna and in the species fractions, which are components of the sample data used to estimate total catch. The work presented here includes an analysis of spatial-temporal variability in the average weight of skipjack tuna, and further evaluation of candidate poststrata using linear models. As done previously, all analyses were applied to purse-seine sets on tunas associated with dolphins. The results suggest that for both yellowfin and skipjack tuna, retaining the 13 sampling areas for estimation of total species catch in dolphin sets is likely not necessary. However, to make a final determination, a study of the variance and coefficient of variation of the total annual catch by species is needed.

Introduction

Stratification is used in stock assessment to address differences in stock and fishery dynamics. Stratification can also be used during data collection to guard against skewed sample allocations (which might lead to bias) and to minimize variance of the estimators of population totals (*e.g*., Thompson 1992). The goals of stratification for stock assessment and data collection are often in agreement. However, they may differ if the characteristics of the fisheries have changed over time. At present, the estimates of tuna catch by species for the purse-seine fishery in the eastern Pacific Ocean (EPO) (*e.g*., Aires-da-Silva and Maunder 2012) are only available for the same spatial-temporal stratification as that used to guide data collection for the port-sampling program (Suter 2010; Tomlinson 2004). This imposes the limitation that the areas used for stock assessment must be based on aggregation of the 13 sampling areas (Figure 1a).

A further complication associated with use of the present sampling strata for catch estimation in the purse-seine fishery is that every year there are a large number of strata with catch but limited or no sample data. The present stratification used for estimation of the species composition of the catch by set type is based on 13 areas, 12 months and two vessel size categories, or 312 potential strata. On average, annually from 2000-2010, approximately 20% of the total landed catch has been associated with sampling strata for which it was not possible to obtain port-sampling data (due to logistic constraints) or for which fewer than two samples were available per stratum. This "missing" or insufficient data problem corresponds to hundreds of strata with catch but insufficient sample data for species composition estimation. The current process of assigning sampling data from other strata to those strata with missing information amounts to a collapsing of the sampling strata. This process follows a complex set of rules for which simplification may be possible.

Poststratification (*e.g*., Holt and Smith 1979; Valliant 1993) is a technique used in data analysis to group samples, after the data have been collected, when estimates of population totals are desired for groups whose definitions were not expressly part of the data-collection protocol. Two poststratified estimators of species catch were proposed (Lennert-Cody *et al*. 2011) that would make relatively minor modifications to the form of the current estimator (Appendix), while allowing other stratifications to be used. The first estimator was based on the assumption that both the sampling strata and the poststrata contain important information that must be retained. The second estimator was based was on the assumption that only the poststrata matter. A preliminary evaluation (Lennert-Cody *et al*. 2012) found that the fine-scale stratification required by the first estimator may not be necessary, but that large-scale poststratification of the sample data (*e.g*., at the level of the present stock assessment areas, Figure 1b) may not be adequate. In addition, for the models considered, spatial structure in the data appeared to dominate over temporal (monthly, quarterly) structure. This document describes additional analyses done to evaluate the necessary level of spatial poststratification. As in the previous work, the analysis focuses on the port-sampling data collected during 2000-2011 for sets on dolphins by large purse-seine vessels (vessels with greater than 363 metric tons fish-carrying capacity). The two species considered in these analyses are yellowfin tuna and skipjack tuna because almost no bigeye tuna are caught in dolphin sets (IATTC 2012).

2. DATA

Data on the species and size composition of the catches of tuna by purse-seine vessels are collected when vessels arrive in port to unload (Tomlinson 2004; Suter 2010). To ensure that the samples collected are representative of the entire fishery, categories, or 'strata', have been established to guide sample collection. For a given type of purse-seine set, these sampling strata are defined by the location of fishing (13 areas, Figure 1a), the month of fishing and the size of vessel (small and large purse-seiners), for a total of 312 possible strata. Not all strata have fishing activity in any given year. Samples are collected by stratum according to a 'two-stage' approach, where the wells of a vessel are the first stage, and the fish within a well are the second stage. Because the number of wells in a stratum is not known in advance and because some vessels may unload in ports where logistics make sampling prohibitively difficult, wells to be sampled are selected opportunistically. However, a well is sampled only if all the catch it contains is from the same sampling stratum (*i.e*., same area, month, purse-seine set type and size of vessel). Over the course of a year, unequal numbers of wells will be sampled per stratum. A rough estimate of the sampling coverage for 2000-2010, based on amount of catch in sampled wells relative to the total catch, is 8%.

Once a well of a vessel has been selected to be sampled, individual fish are sampled from the well as the catch is unloaded. A number of fish of each species (typically 50) are measured for length. From the same well, and independently of the measured fish, several hundred fish are counted for species composition. The fish sampled from the well are selected one at a time, from an opportunistically established starting point, as circumstances permit. Depending on the port of unloading, some well catches may be sorted by species and weight category before the catches are accessible to IATTC staff for sampling. Catches from these types of unloadings are therefore sampled slightly differently; details of the port-sampling data collection procedures can be found in the appendix of Suter (2010).

A data feasibility analysis (Lennert-Cody *et al*. 2012) found that since 2000 both the sampling area and the 5° area were recorded for most samples. A comparison of the 5° areas of the port sampling data to the actual positions of the sets whose catches went into the sampled wells indicates that about 81% of all samples from 2000-2011 were in agreement with actual set positions at the 5°-area level, and about 97% of all samples were within one 5° area of the 5° area of the corresponding set. Thus, poststratification is feasible for the 2000-2011 data as long as poststrata are constructed from combinations of 5° areas.

The average weight of a species in each sample was computed from individual lengths, converted to weight (Tomlinson 2004). The individual lengths (in mm) are converted to weights (in kg) using the formulas: w = 1.77e-08$l^{3.02}$ for yellowfin tuna and w = 2.55e-09$l^{3.336}$ for skipjack tuna. So that data of sorted and non-sorted unloadings can both be used in the analyses, the average weight of a species in the sampled well is the estimate of the total weight of species in the well divided by the estimated total number of fish of that species in the well. For non-sorted samples, this reduces to the sample average weight (sum of weights of measured fish divided by the number of fish meausred). For samples from sorted unloadings, estimates of the total numbers and weights of each species in each sort were obtained from the method of Tomlinson (2004). The estimates were then summed across sorts to get the estimates of total number and weight of each species in the well.

Analysis

Both the current estimator of total species catch (Appendix) and alternatives (Lennert-Cody *et al*. 2012) are based on the average weights of fish in the sample and the proportion of each species in the sample counts. Thus, to further evaluate the benefit of various levels of spatial and temporal poststratification of the port-sampling data for estimation of total catch by species, analyses of spatial-temporal variability in both average weights and species proportions were conducted.

Several spatial stratifications were considered in these analyses. Three main spatial stratifications were used: the first spatial stratification was the 3-area stratification presently used for stock assessment (Aires-da-Silva and Maunder 2012; "stock assessment areas", Figure 1b); the second was the five-area tree stratification described in Document YFT-01-02 ("tree areas", Figure 1c); and the third was the 13-area stratification used for the port-sampling data collection ("sample areas", Figure 1a). Two of these three spatial stratifications were also modified (see below) to produce two alternative spatial stratifications. Two time periods were also considered: month and quarter. As an example of the spatial pattern in the average weight data, frequency distributions of the average weight of yellowfin tuna and skipjack tuna by tree area are shown in Figures 2-3.

Average weight

To explore spatial-temporal structure in average weight, two types of analysis were undertaken separately for each species (yellowfin, skipjack). In the first analysis, regression trees ('CART'; Breiman *et al*. 1984) were applied to the average weight data within each of the stock assessment areas, and separately, within each of the tree areas. These CART analyses were applied to the data of all years combined, with predictors 5º latitude, 5º longitude, quarter, and year. The spatial predictors were numeric and the temporal predictors categorical. Results of these CART analyses were used to sub-stratify the stock assessment areas and tree areas.
In the second analysis, multiple linear regression models ('LM') for average weight were fitted separately to the data of each year for the whole EPO, using the area (stock assessment, tree, CART-modified stock assessment or CART-modified tree), and quarter as independent variables. All independent variables were categorical. The LM analyses were used to compare the various levels of poststratification for the average weight data.

In both the CART and the LM analyses, the total catch in the well (all tuna species combined) was used as a weight, to be consistent with the estimator of total species catch (Appendix).

Before conducting the CART and LM analyses, the square root transformation was applied to the data of both yellowfin and skipjack to help the data conform to the assumption of equal variance. This transformation was selected based on inspection the residual diagnostic plots from fitting a linear model to average weight with predictors of sample area, month and year, and with total catch in the well as a weight. This transformation helped somewhat with the tail behavior of the data of both species, but later analyses suggested that in the future a Box-Cox transformation of the skipjack data would be more appropriate ($\lambda = 0.5$; Box-Cox(y) = $(y^{\lambda} -1)/\lambda$).

For the LM analysis, the following linear models were fitted annually to the data of each species (where sufficient data were available):

i) $sqrt(\overline{w}_j)$ = overall constant + stock assessment area effect + error
ii) $sqrt(\overline{w}_j)$ = overall constant + tree area effect + error
iii) $sqrt(\overline{w}_j)$ = overall constant + stock-CART area effect + error
iv) $sqrt(\overline{w}_j)$ = overall constant + tree-reducedCART area effect + error
v) $sqrt(\overline{w}_j)$ = overall constant + tree-CART area effect + error
vi) $sqrt(\overline{w}_j)$ = overall constant + sample area effect + error
vii) $sqrt(\overline{w}_j)$ = overall constant + stock assessment area*quarter + error
viii) $sqrt(\overline{w}_j)$ = overall constant + tree area*quarter + error

ix) $sqrt(\overline{w}_j)$ = overall constant + stock-CART area*quarter + error
x) $sqrt(\overline{w}_j)$ = overall constant + tree-reducedCART area*quarter + error
xi) $sqrt(\overline{w}_j)$ = overall constant + tree-CART area*quarter + error
xii) $sqrt(\overline{w}_j)$ = overall constant + sample area*quarter + error

where $\overline{w}_j$ is the average weight of a species in the j^{th} sampled well, 'sqrt' indicates square root, '*' denotes a model with main effect and first-order interactions, "stock assessment area" refers to the areas shown in Figure 1b, "tree area" refers to the areas shown in Figure 1c, "sample area" refers to the areas shown in Figure 1a, and the modified spatial stratifications "stock-CART", "tree-CART' and "tree-reducedCART" are the results of the CART analyses (see below) and are shown in Figures 1d-f, respectively. Models with area-quarter interactions were fitted for the purpose of illustration, even though not all area-quarter combinations were represented in the data. The Akaike Information Criterion (AIC; Burnham and Anderson 2002) was computed for each model, and the difference in AIC between each model and the model with the lowest AIC ($\Delta AIC = AIC - AIC_{min}$, Burnham and Anderson 2002) was used to compare models within each year and species. As a general rule, models that performed similarly to the model with the lowest AIC will have a low ΔAIC value (~2 or less), whereas those that performed poorly by comparison will have a high ΔAIC value (~ > 10) (but see Discussion Section).

Species counts

To look at the variability of species composition within the current stock assessment areas and the tree areas, classification trees were built for the presence/absence of skipjack in the samples. Less than 20% of samples had skipjack (Figure 4). Previous analyses (Lennert-Cody *et al*. 2012) of the proportion of skipjack in the samples with a logistic regression model (positives= counts of skipjack, negatives = counts of yellowfin) were problematic (models failed to convergence in some years), most likely because those samples that had skipjack had proportionally only a small amount (*i.e*., the probability of getting a skipjack in the sample count was typically 0 or small). For this reason, an analysis was done based on presence/absence of skipjack in the samples using a two-class classification tree. The predictors were the 5º latitude, 5º longitude, month (or quarter) and year. The spatial predictors were numeric and the temporal predictors categorical. The total catch in the well (all tuna species combined) was used as a weight. Only non-sorted samples were used in this analysis. The determination of presence/absence of skipjack for each sample was based on the sample counts (0: no skipjack in the sample count; 1: ≥ 1 skipjack in the sample count). The classification algorithm was run separately within each of the three current stock assessment areas and within each the five tree areas. Classification trees were built with default settings and with settings that doubled the cost of misclassifying samples with skipjack. However, this had almost no effect on the split variable-values selected, and so only the classification trees built with the default settings are presented below.

3. RESULTS

Average weight

CART analyses

In general, for yellowfin tuna the greatest decrease in the prediction error that could be achieved by the regression tree analysis occurred within the first few partitions of the data (Figures 5-6), both within each stock assessment area and within each tree area. This was the case whether month was included as the temporal predictor (not shown) or quarter was included as the temporal predictor. The first partition of the data within each area was either based on year or location (latitude, longitude), not on quarter (Figures 7-14). However, quarter was sometimes one of the first alternative split variables. This suggests that either spatial variability dominates over quarterly variability in this data set or that spatial variability was more easily detected with these analyses. In a few areas (*e.g.*, tree area 1, Figure 6) it is questionable whether the data should be partitioned at all.

Given that splits on year, when they occurred at the very top of the tree, were not consistent (within a level of spatial stratification; “stock assessment” or “tree”), and the low level of sampling coverage, in what follows it is assumed that partitions on year were most likely to reflect sampling variability (Table 1 of the Appendix shows the number of samples for average weight by 5º area for each year). Therefore, to further subdivide the stock assessment areas and tree areas (Figure 1b-c), the first partition of each tree was used if it was based on latitude or longitude, or the best competitor was used when year defined the first partition (compare first partitions of data in stock assessment areas and tree areas from Figures 7-14 to Figure 1d-f). An exception to this rule was made for stock assessment area 7 (Figure 7) because the partition on latitude would have resulted in a sub-area with relatively few samples, and so the best competitor split was used instead.

As with the analyses for yellowfin tuna, in general the first partition of the skipjack data set in each area achieved the greatest reduction in prediction error (Figures 15-16). However, there was more unexplained variability in the average weight data for skipjack and hence more areas where it was questionable whether the data should be partitioned at all (no analyses were done for tree area 4 due to the small number of samples). In addition, year defined the first partition of the data in every area (Figures 17-22). This may indicate that that sampling variability is more prevalent for the skipjack tuna data than for yellowfin tuna data, which would not be surprising given that there were fewer samples for skipjack tuna (Appendix Tables 1-2). Alternatively, this could be due to inter-annual variation in recruitment or spatial variation in recruitment combined with lack of movement of fish. Annual variability in recruitment will cause the mean weight to change from year to year. Since skipjack tuna is shorter-lived than yellowfin tuna, recruitment will compromise a large component of the total population and may cause mean weight to have larger temporal variability. Or, due to limited movement of these species, spatial variation in recruitment will cause spatial variability in mean weight. The spatial pattern of recruitment may also vary annually. Because of this, the modified stock assessment areas and modified tree areas from the yellowfin analysis (Figure 1 d-f) were also used for skipjack.

LM analyses

Stratifications based on fewer strata (*e.g*., the three stock assessment areas or the five tree areas) were not generally favored for either species (Table 1). On the other hand, the stratification based on the current sampling areas (13 areas) was only selected in three of 12 years for each species as the best model. There is also a lot of year to year variability in which type of stratification would be 'best' (Table 1), which is not surprising given the results of the tree analyses (Figures 7-14 and 17-22) and may be indicative of temporal variability in the spatial pattern of recruitment and/or year-to-year sampling variability. Overall, based on these analyses, for yellowfin tuna a poststratification based on sub-stratified tree areas or tree-like areas would seem preferable. Of the 12 models fitted for yellowfin tuna, eight favored some form of spatial stratification based on tree areas or sub-stratified tree areas (Table 1). This would be expected since the tree areas were based in part on analysis of large-scale pattern in the yellowfin tuna length-frequency data (Document YFT-01-02). For skipjack tuna there is less agreement across years: in four of the 12 years, a model based on sub-stratifying the stock assessment areas was favored, whereas in and five of the 12 years, a model based on sub-stratifying the tree areas was favored.

Proportion samples with skipjack

There was a lot of unexplained variability in the skipjack presence/absence data, to the point that the tree analyses are of questionable value (Appendix Figures 1-8). The best first partitions of the data in all areas were based on year, which is not suprising given previous analyses (Lennert-Cody *et al*. 2012). Thus, although there may be common spatial pattern, year to year variability is making detection of that spatial pattern problematic. The presence of skipjack in the samples was not considered further (but see Discussion section).

4. DISCUSSION

This document presented analyses of port-sampling data for yellowfin and skipjack tunas from dolphin sets. The purpose of the analyses was to determine an appropriate level of poststratification of the port-sampling data for estimation of total catch by species in dolphin sets. The results of the analyses of average weights of yellowfin tuna and of skipjack tuna suggest that retaining the 13 sampling areas for estimation of total species catch in dolphin sets may not be necessary. The results also suggest that of the two types of large-scale stratifications considered (current stock assessment areas and tree areas), a tree-type stratification would be favored for yellowfin, with less convincing results for skipjack. Given that the stock assessment models use a quarterly time step, these results suggest that a tree-like spatial stratification by quarter may be a practical compromise to the current stratification used for estimating total species catch in dolphin sets.

The challenges encountered with spatial-temporal analysis of these data, particularly the sample count data, suggest that a two-part procedure should be used in the final analyses for selecting the level of spatial-temporal poststratification. This is because selecting poststrata purely based on ΔAIC from the linear models for average weight of the dominant species in the catch may not yield satisfactory results for catch estimates of other species. Therefore, in addition to the linear model analyses of average weight, it would be useful to compare the performance of candidate poststrata based on annual estimates of the variance and coefficient of variation of the total catch by species. This would take into account variability in the species composition of the sample count data, as well as variability in the species weight data. This will be particularly important for the other types of purse-seine sets which catch all three species of tropical tunas (IATTC 2012).

Given that the least squares models for average weight were fitted with total catch as weights, the typical guidelines for meaningful changes in AIC (Burnham and Anderson, 2002) may not apply. Therefore, if changes in AIC are to be used as a criterion for selecting among candidate stratifications in the future, a simulation procedure should be used to establish the criteria for competitor and poor models.

Acknowledgements

Special thanks to Nickolas Vogel for data base assistance and to Patrick K. Tomlinson for helpful discussions.

References

Aires-da-Silva, A. and Maunder, M.N. 2012. Status of yellowfin tuna in the eastern Pacific Ocean in 2011 and outlook for the future. Inter-American Tropical Tuna Commission Stock Assessment Report 12, pages 3-110.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. 1984. *Classification and Regression Trees*. Chapman & Hall/CRC. 358 pp.

Burnham, K.P. and Anderson, D.R. 2002. Model Selection and Multimodel Inference. Springer.

Holt, D. and Smith, T.M.F. 1979. Post stratification. Journal of the Royal Statistical Society, Series A 142, Part 1: 33-46.

IATTC 2012. Fishery Status Report 10, Inter-American Tropical Tuna Commission. http://www.iattc.org/PDFFiles2/FisheryStatusReports/FisheryStatusReport10ENG.pdf

Lennert-Cody, C.E., Maunder, M.N., Tomlinson, P.K., Aires-da-Silva, A., Pérez, A., McCracken, M. 2011. Poststratified estimators of total catch for the purse-seine fishery port-sampling data. Document SAC-02-10 presented at the 2nd Scientific Advisory Meeting of the Inter-American Tropical Tuna Commission, 9-12 May, 2011, La Jolla, CA ,USA.

Lennert-Cody, C.E. Maunder, M.N., Tomlinson, P.K., Aires-da-Silva, A., Pérez, A. 2012. Progress report on the development of poststratified estimators of total catch for the purse-seine fishery port-sampling data. IATTC Document SAC-03-10, 3rd Meeting of the Scientific Advisory Committee, La Jolla, California, USA, 15-18 May, 2012.

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. Inter-American Tropical Tuna Commission Special Report 18.

Thompson, S.K. 1992. Sampling. Wiley.

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, pages 311-333.

Valliant, R. 1993. Poststratification and conditional variance estimation. Journal of the American Statistical Association 88: 89-96.

Table 1. ΔAIC from the linear model analyses (LM) of average weight of yellowfin tuna (YFT) and skipjack tuna (SKJ), by year and model. A value of 0 for ΔAIC indicates the 'best' model for that year and species. However, in some cases, because of data sparseness, several related models for the same year and species may all have a ΔAIC value of 0.

	YFT ΔAIC	SKJ ΔAIC
2000		
Stock (model (i))	56.9	25.2
Tree (model (ii))	28.0	29.8
StockCART (model (iii))	47.8	8.8
Tree-reducedCART (model (iv))	22.1	31.2
TreeCART (model (v))	12.0	31.2
Sample (model (vi))	16.2	31.1
Stock * quarter (model (vii))	31.8	34.0
Tree * quarter (model (viii))	13.1	34.9
StockCART * quarter (model (ix))	30.5	0.0
Tree-reducedCART * quarter (model (x))	16.7	36.6
TreeCART * quarter (model (xi))	6.0	36.6
Sample * quarter (model (xii))	0.0	36.6
2001		
Stock (model (i))	39.8	15.81
Tree (model (ii))	42.4	1.16
StockCART (model (iii))	30.3	16.06
Tree-reducedCART (model (iv))	35.6	1.55
TreeCART (model (v))	35.9	1.71
Sample (model (vi))	28.2	10.19
Stock * quarter (model (vii))	3.8	12.30
Tree * quarter (model (viii))	24.9	4.22
StockCART * quarter (model (ix))	9.6	12.44
Tree-reducedCART * quarter (model (x))	16.0	0.1
TreeCART * quarter (model (xi))	14.5	0.0
Sample * quarter (model (xii))	0.0	14.00
2002		
Stock (model (i))	57.6	6.6
Tree (model (ii))	32.8	9.1
StockCART (model (iii))	47.9	1.7
Tree-reducedCART (model (iv))	24.0	2.6
TreeCART (model (v))	26.3	0.0
Sample (model (vi))	35.7	8.2
Stock * quarter (model (vii))	33.7	6.3
Tree * quarter (model (viii))	12.9	9.1
StockCART * quarter (model (ix))	24.7	0.1
Tree-reducedCART * quarter (model (x))	0.0	5.0

TreeCART * quarter (model (xi))	5.3	4.4
Sample * quarter (model (xii))	3.6	10.2
2003		
Stock (model (i))	59.7	9.6
Tree (model (ii))	23.0	17.5
StockCART (model (iii))	43.1	11.5
Tree-reducedCART (model (iv))	2.8	5.7
TreeCART (model (v))	4.9	3.5
Sample (model (vi))	13.7	12.2
Stock * quarter (model (vii))	53.5	6.3
Tree * quarter (model (viii))	19.0	9.2
StockCART * quarter (model (ix))	42.1	0.0
Tree-reducedCART * quarter (model (x))	0.3	5.7
TreeCART * quarter (model (xi))	9.0	6.9
Sample * quarter (model (xii))	0.0	12.9
2004		
Stock (model (i))	25.7	32.7
Tree (model (ii))	17.1	39.0
StockCART (model (iii))	5.1	33.6
Tree-reducedCART (model (iv))	16.4	38.4
TreeCART (model (v))	0.0	40.1
Sample (model (vi))	1.1	38.1
Stock * quarter (model (vii))	32.0	17.1
Tree * quarter (model (viii))	25.6	7.1
StockCART * quarter (model (ix))	18.3	13.2
Tree-reducedCART * quarter (model (x))	31.2	0.0
TreeCART * quarter (model (xi))	19.3	0.0
Sample * quarter (model (xii))	14.3	21.1
2005		
Stock (model (i))	96.7	25.4
Tree (model (ii))	29.4	8.4
StockCART (model (iii))	48.8	14.2
Tree-reducedCART (model (iv))	0.3	0.0
TreeCART (model (v))	1.2	1.8
Sample (model (vi))	15.0	12.0
Stock * quarter (model (vii))	75.0	27.8
Tree * quarter (model (viii))	2.1	16.9
StockCART * quarter (model (ix))	44.9	10.2
Tree-reducedCART * quarter (model (x))	0.0	5.9
TreeCART * quarter (model (xi))	3.3	9.1
Sample * quarter (model (xii))	9.3	0.3

2006		
Stock (model (i))	29.7	12.4
Tree (model (ii))	4.6	14.2
StockCART (model (iii))	9.1	10.6
Tree-reducedCART (model (iv))	0.7	17.5
TreeCART (model (v))	0.6	13.6
Sample (model (vi))	19.0	17.6
Stock * quarter (model (vii))	24.9	1.7
Tree * quarter (model (viii))	0.6	6.4
StockCART * quarter (model (ix))	12.0	0.0
Tree-reducedCART * quarter (model (x))	0.0	7.6
TreeCART * quarter (model (xi))	1.7	9.4
Sample * quarter (model (xii))	17.2	7.0
2007		
Stock (model (i))	39.7	17.19
Tree (model (ii))	46.4	14.08
StockCART (model (iii))	38.6	15.19
Tree-reducedCART (model (iv))	36.6	16.93
TreeCART (model (v))	37.3	13.0
Sample (model (vi))	21.4	20.4
Stock * quarter (model (vii))	40.7	9.4
Tree * quarter (model (viii))	57.8	6.4
StockCART * quarter (model (ix))	24.2	6.0
Tree-reducedCART * quarter (model (x))	34.2	8.5
TreeCART * quarter (model (xi))	36.4	0.0
Sample * quarter (model (xii))	0.0	8.1
2008		
Stock (model (i))	62.0	42.92
Tree (model (ii))	52.5	25.35
StockCART (model (iii))	30.8	13.76
Tree-reducedCART (model (iv))	20.3	26.48
TreeCART (model (v))	0.0	28.46
Sample (model (vi))	36.8	0.73
Stock * quarter (model (vii))	53.9	18.92
Tree * quarter (model (viii))	45.2	6.77
StockCART * quarter (model (ix))	35.5	3.17
Tree-reducedCART * quarter (model (x))	20.4	8.86
TreeCART * quarter (model (xi))	3.0	10.93
Sample * quarter (model (xii))	36.5	0.0
2009		
Stock (model (i))	170.6	31.0
Tree (model (ii))	87.9	11.9

StockCART (model (iii))	54.8	31.0
Tree-reducedCART (model (iv))	52.2	11.9
TreeCART (model (v))	1.9	11.9
Sample (model (vi))	30.1	32.0
Stock * quarter (model (vii))	148.9	33.7
Tree * quarter (model (viii))	64.6	0.0
StockCART * quarter (model (ix))	40.4	33.7
Tree-reducedCART * quarter (model (x))	26.9	0.0
TreeCART * quarter (model (xi))	0.0	0.0
Sample * quarter (model (xii))	13.2	31.0
2010		
Stock (model (i))	76.0	3.3
Tree (model (ii))	57.5	8.9
StockCART (model (iii))	42.0	6.4
Tree-reducedCART (model (iv))	58.6	7.0
TreeCART (model (v))	49.8	8.8
Sample (model (vi))	33.9	0.0
Stock * quarter (model (vii))	69.7	8.7
Tree * quarter (model (viii))	64.0	11.9
StockCART * quarter (model (ix))	0.0	11.8
Tree-reducedCART * quarter (model (x))	44.7	11.8
TreeCART * quarter (model (xi))	27.8	13.7
Sample * quarter (model (xii))	10.9	8.1
2011		
Stock (model (i))	139.2	60.4
Tree (model (ii))	86.3	44.0
StockCART (model (iii))	56.2	63.8
Tree-reducedCART (model (iv))	11.7	43.1
TreeCART (model (v))	1.2	45.0
Sample (model (vi))	60.1	58.0
Stock * quarter (model (vii))	130.4	38.2
Tree * quarter (model (viii))	95.0	14.4
StockCART * quarter (model (ix))	35.8	42.2
Tree-reducedCART * quarter (model (x))	10.1	16.4
TreeCART * quarter (model (xi))	0.0	17.0
Sample * quarter (model (xii))	56.9	0.0

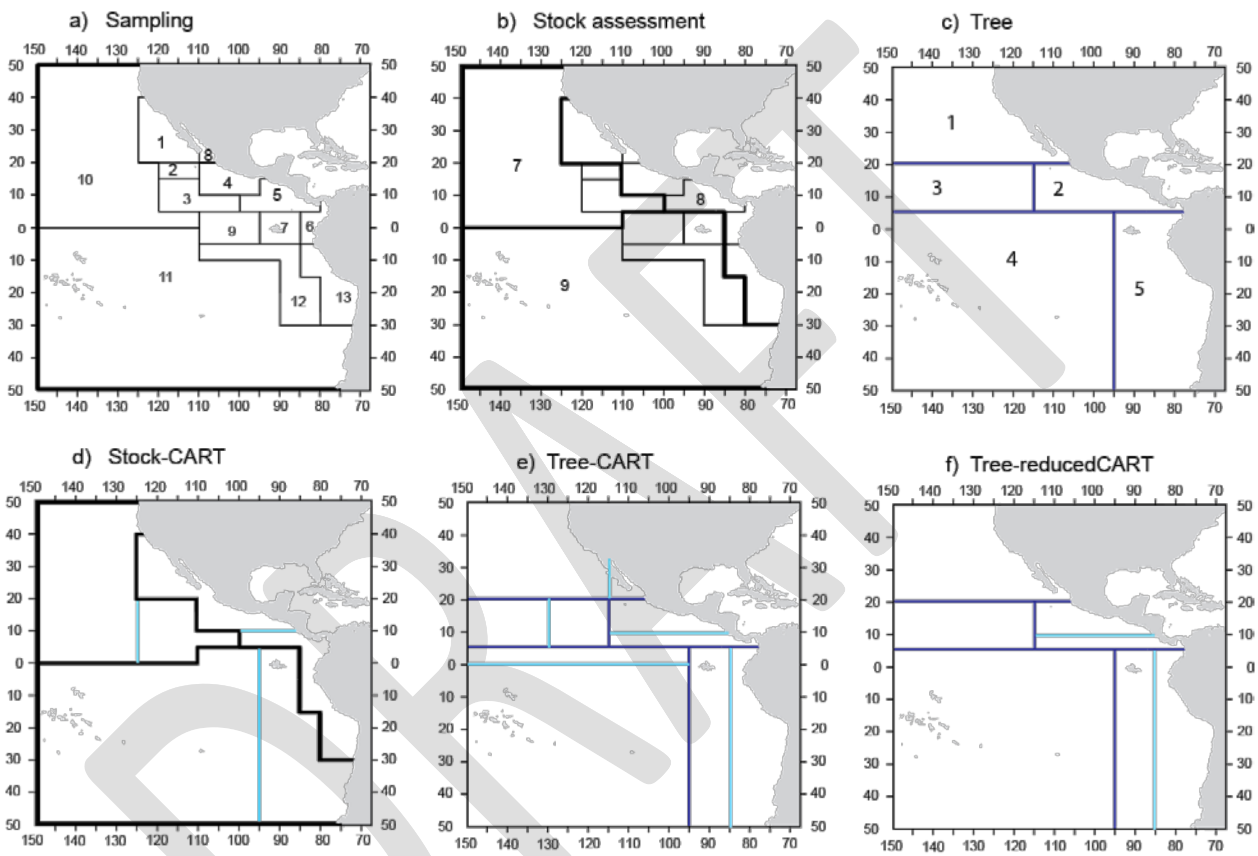

Figure 1. a) Sampling areas; b) stock assessment areas for yellowfin tuna in dolphin sets (thick black lines); c) stratification obtained from the tree analysis ("weighted simultaneous" analysis; Document YFT-01-02); d) results of CART analysis of average weight within stock assessment areas; e) results of CART analysis of average weight within tree areas; f) results of CART analysis of average weight within tree areas, excluding sub-stratification of areas 1, 3 and 4.

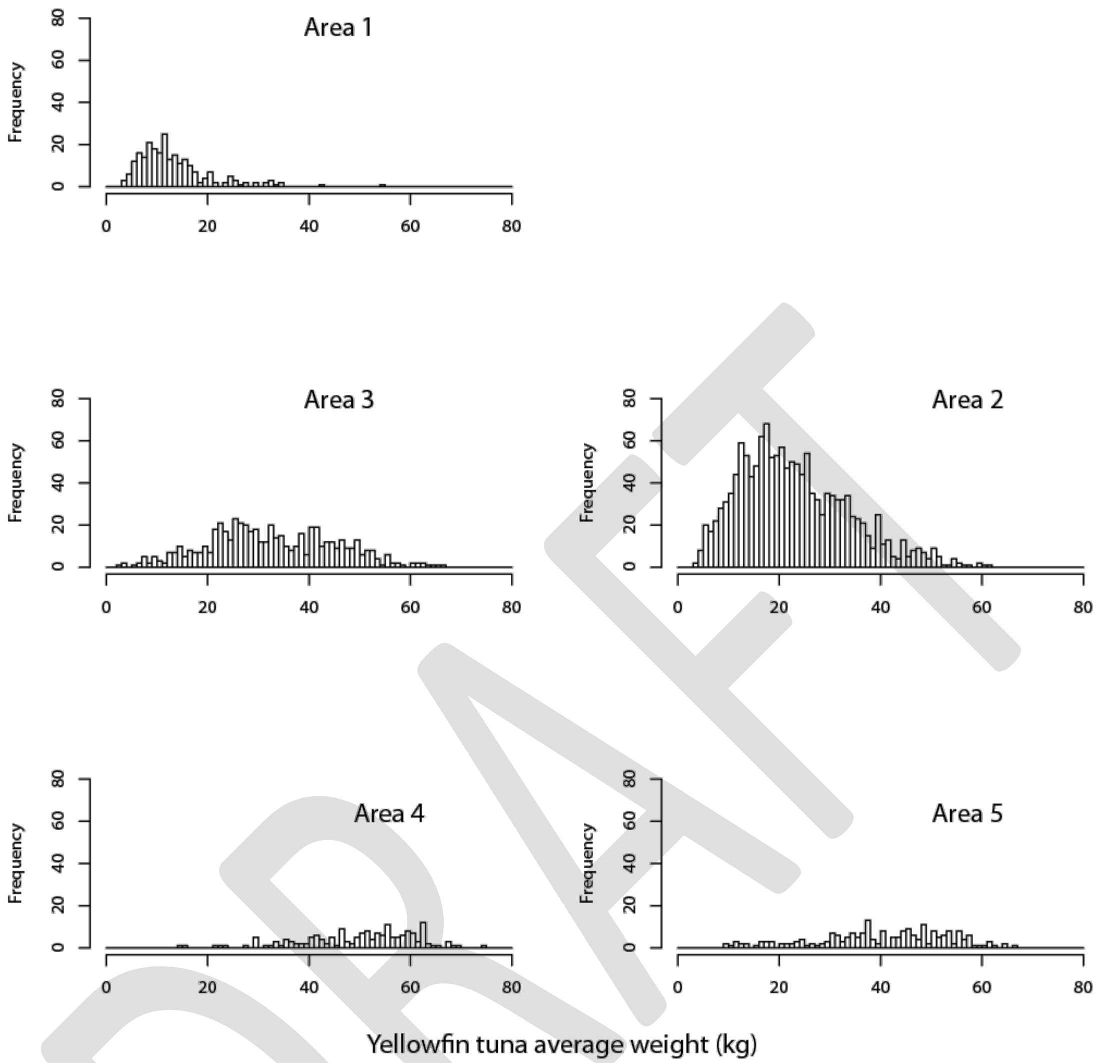

Figure 2. Frequency distributions of average weight per sample for yellowfin tuna ("Frequency" is in number of samples) by tree area (Figure 1c).

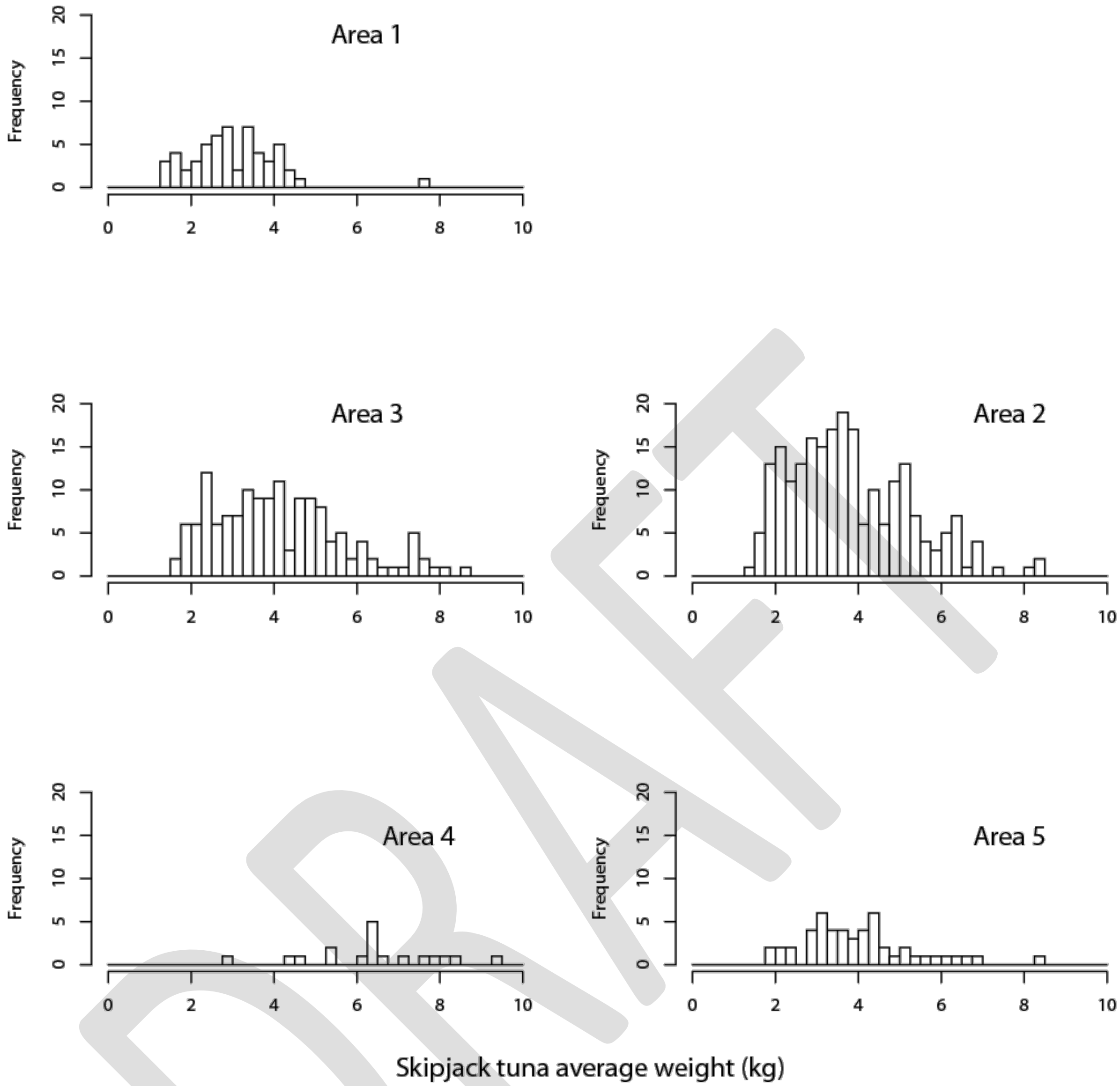

Figure 3. Frequency distributions of average weight per sample for yellowfin tuna (“Frequency” is in number of samples) by tree area (Figure 1c).

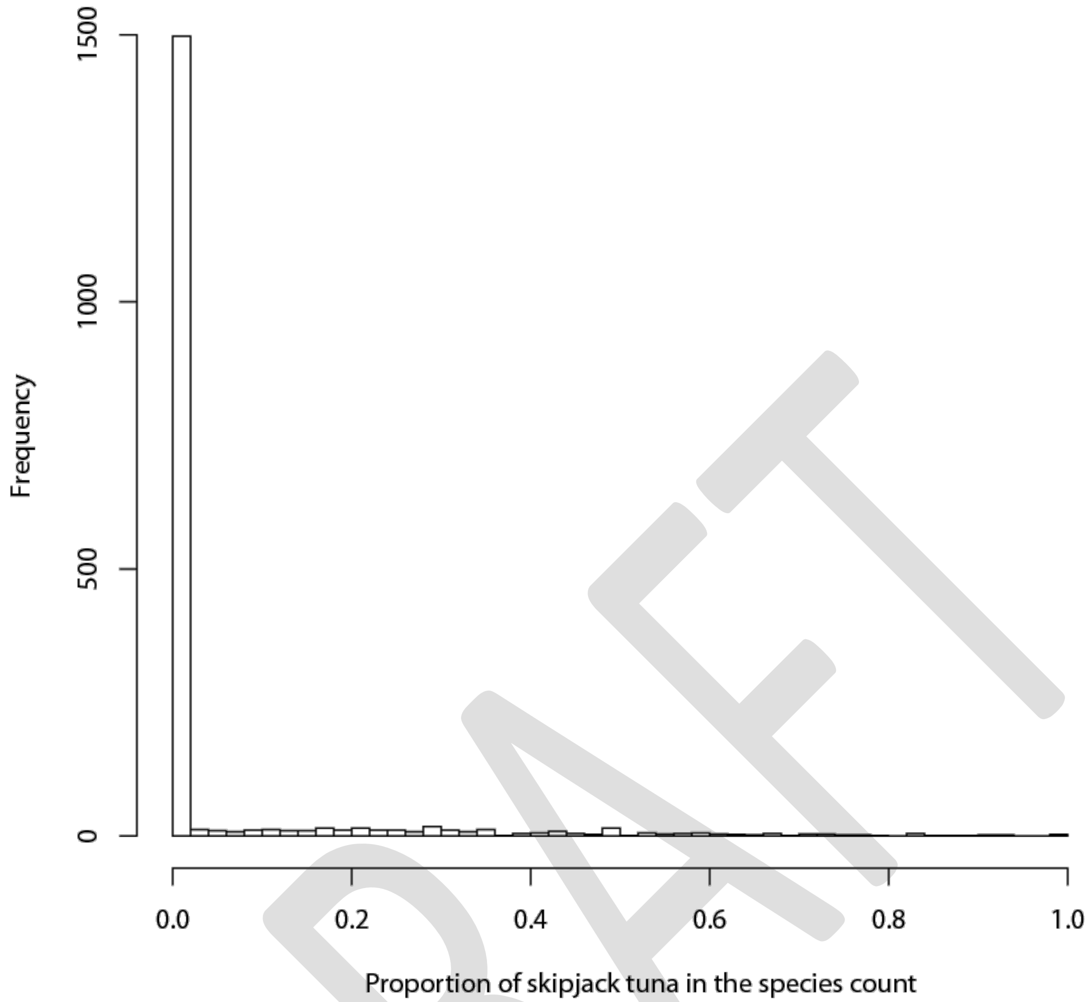

Figure 4. Proportion of samples (pooled across areas) with ≥ 1 skipjack tuna in the samle species count (“Frequency” is number of samples).

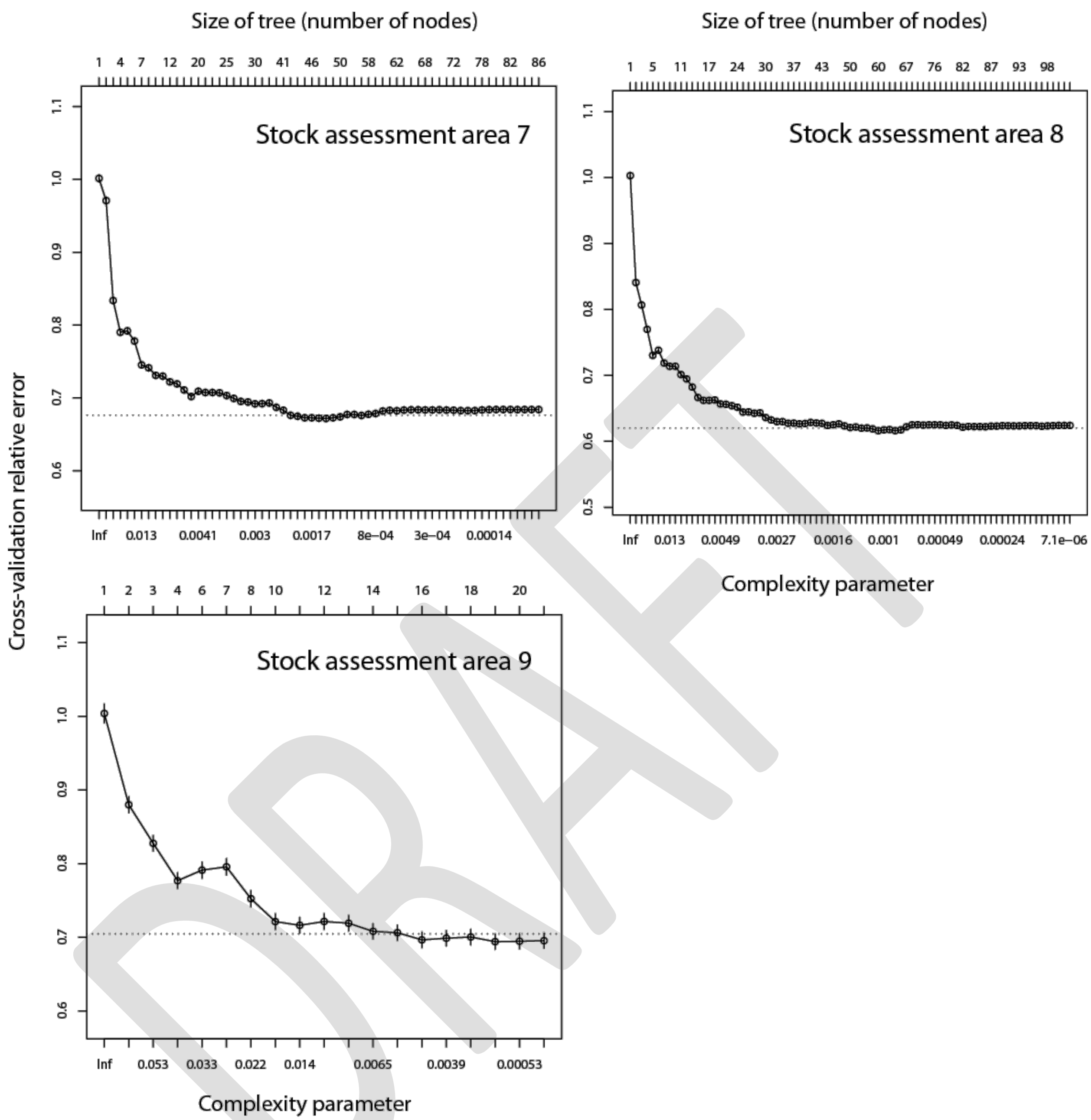

Figure 5. Plots of cross-validation relative error (cross-validation estimate of the tree error, scaled so that the first node has an error of 1.0) as a function of tree size for the regression tree analysis of average weight of yellowfin tuna, by stock assessment area (Figure 1b). The complexity parameter is the penalty on tree size. The dashed line indicates the minimum error plus one standard error.

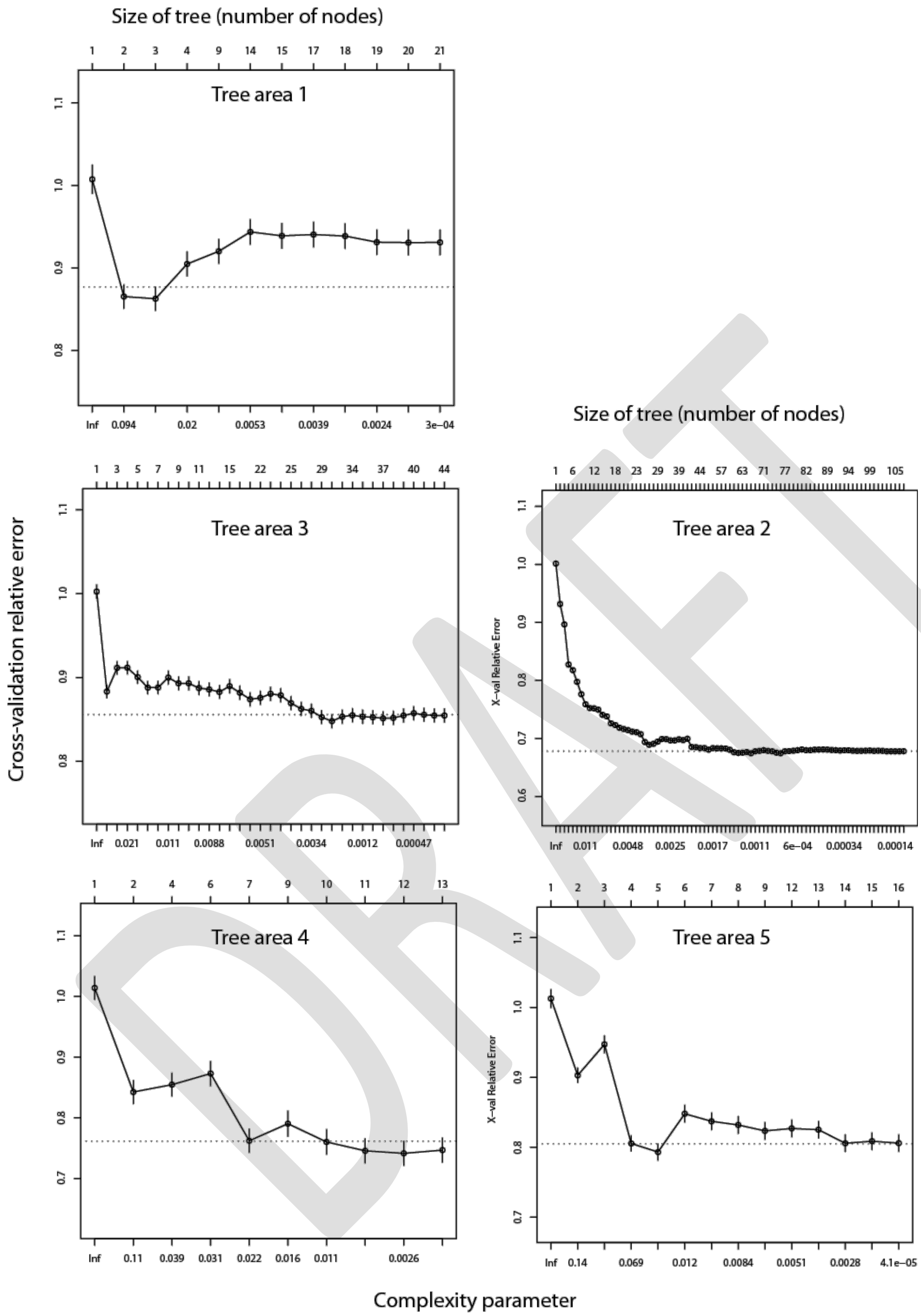

Figure 6. Plots of cross-validation relative error (cross-validation estimate of the tree error, scaled so that the first node has an error of 1.0) as a function of tree size for the regression tree analysis of average weight of yellowfin tuna, by tree area (Figure 1c). The complexity parameter is the penalty on tree size. The dashed line indicates the minimum error plus one standard error.

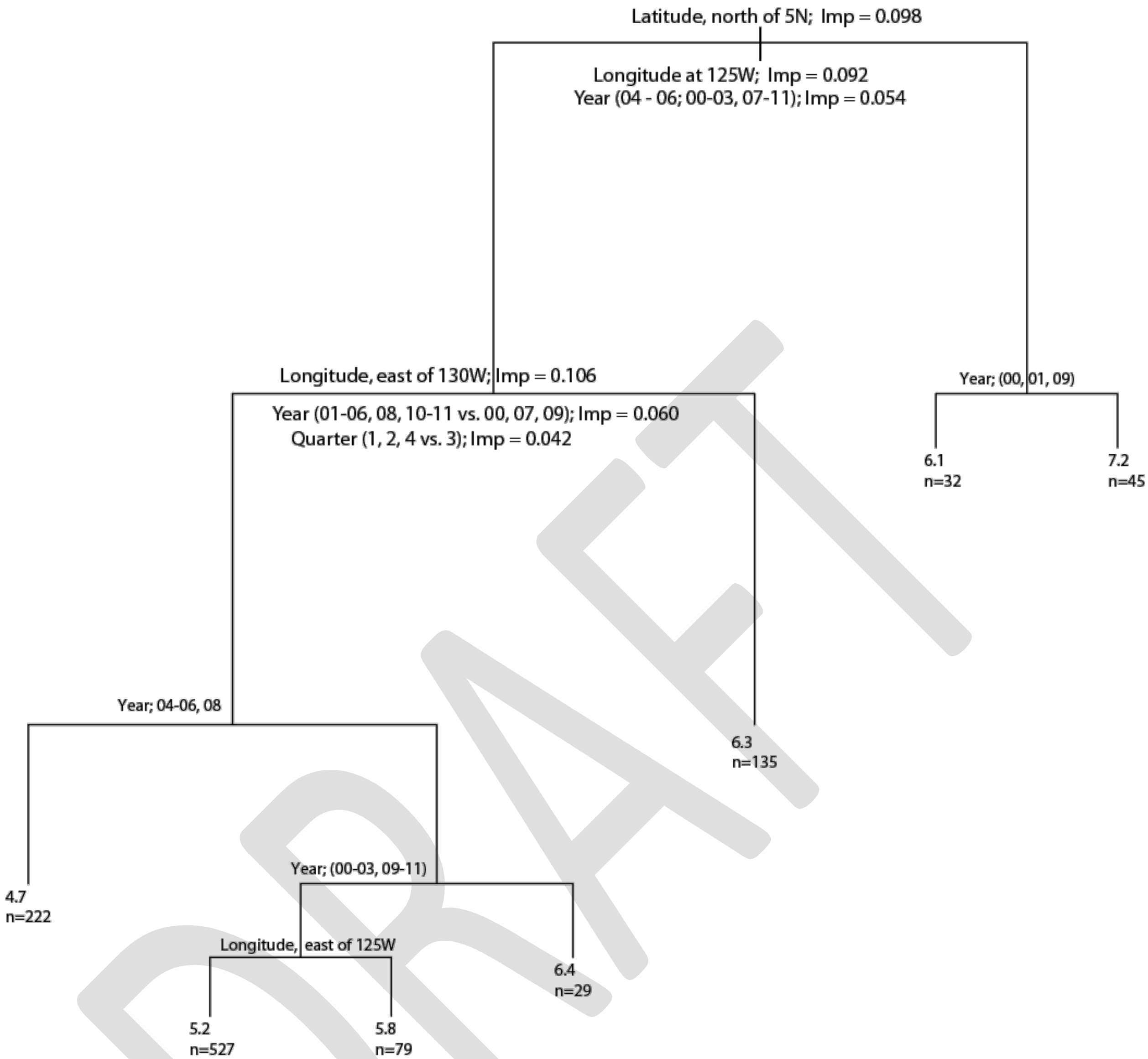

Figure 7. Regression tree obtained for yellowfin tuna average weight in stock assessment area 7 (Figure 1b). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

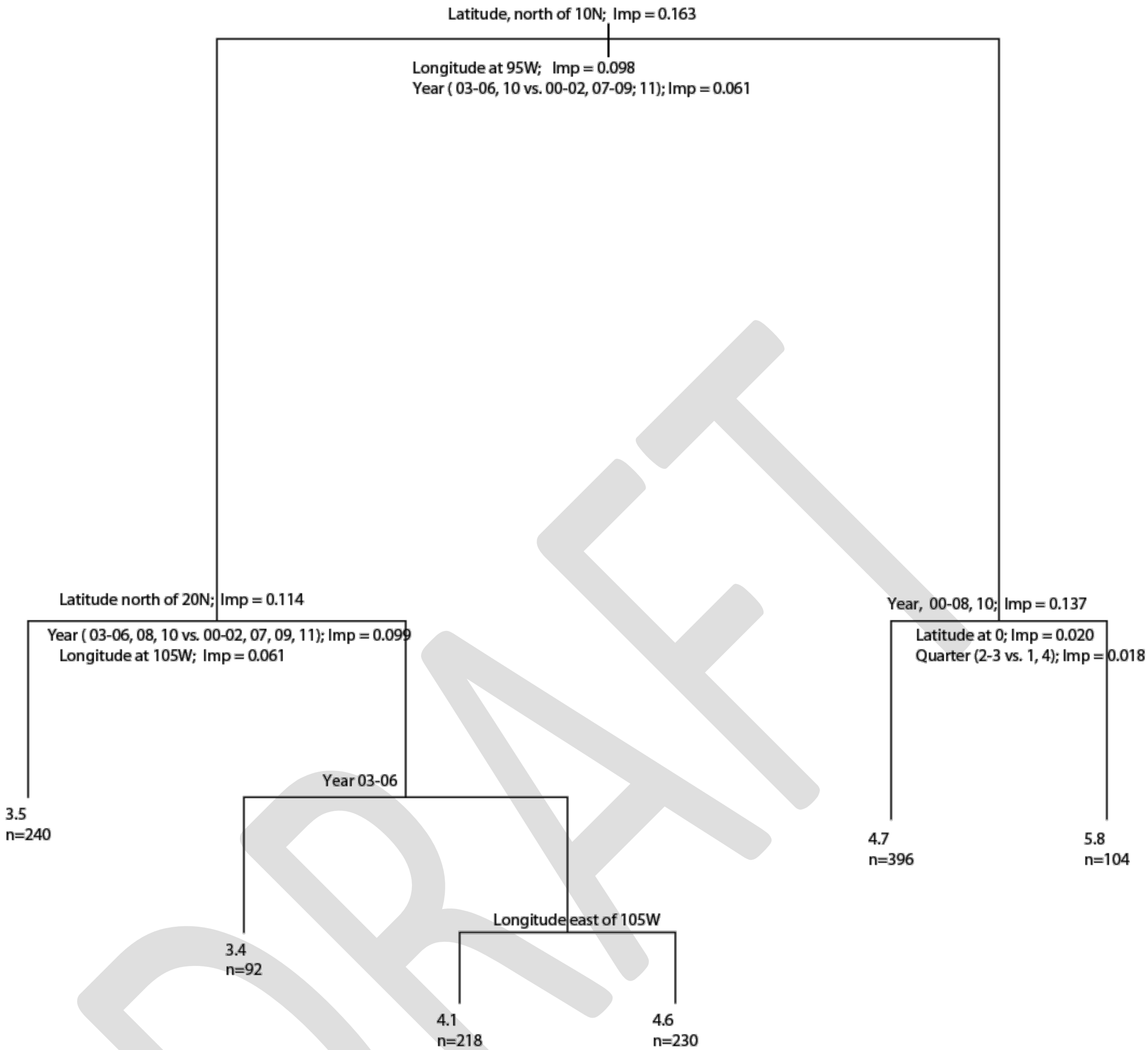

Figure 8. Regression tree obtained for yellowfin tuna average weight in stock assessment area 8 (Figure 1b). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

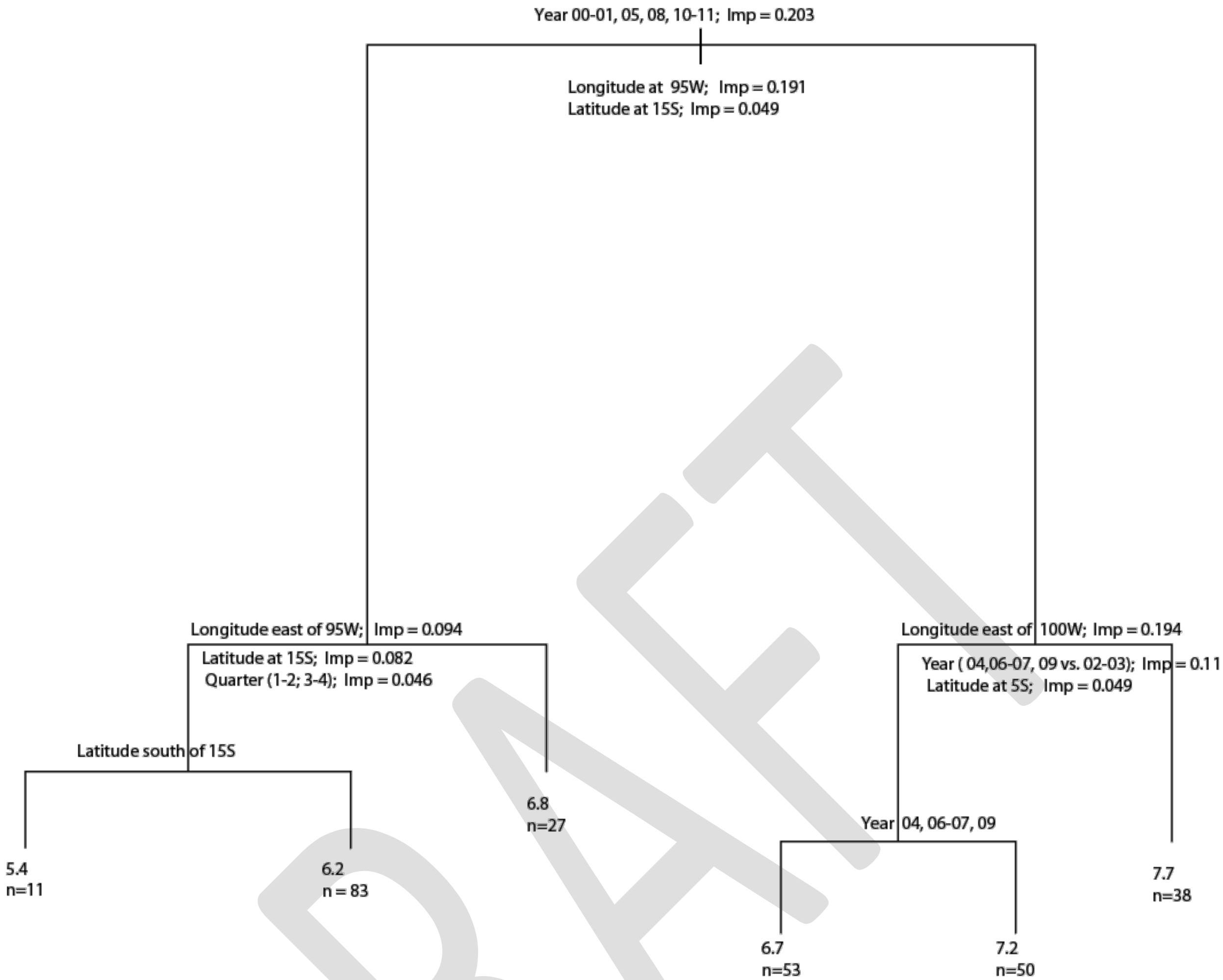

Figure 9. Regression tree obtained for yellowfin tuna average weight in stock assessment area 9 (Figure 1b). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

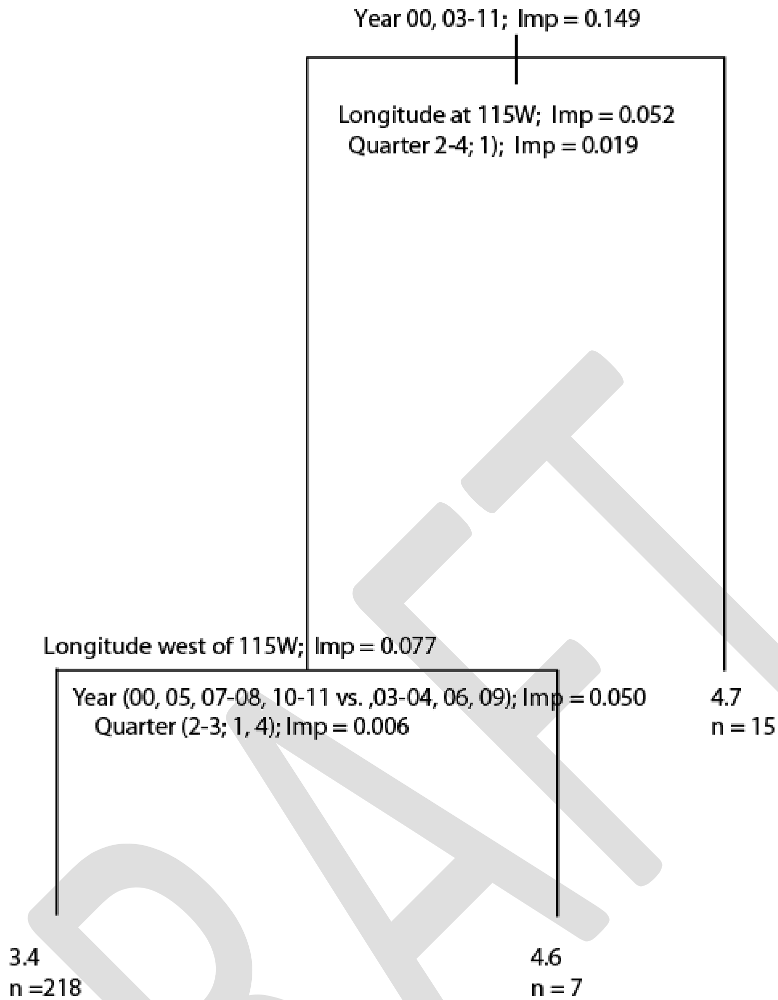

Figure 10. Regression tree obtained for yellowfin tuna average weight in tree area 1 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

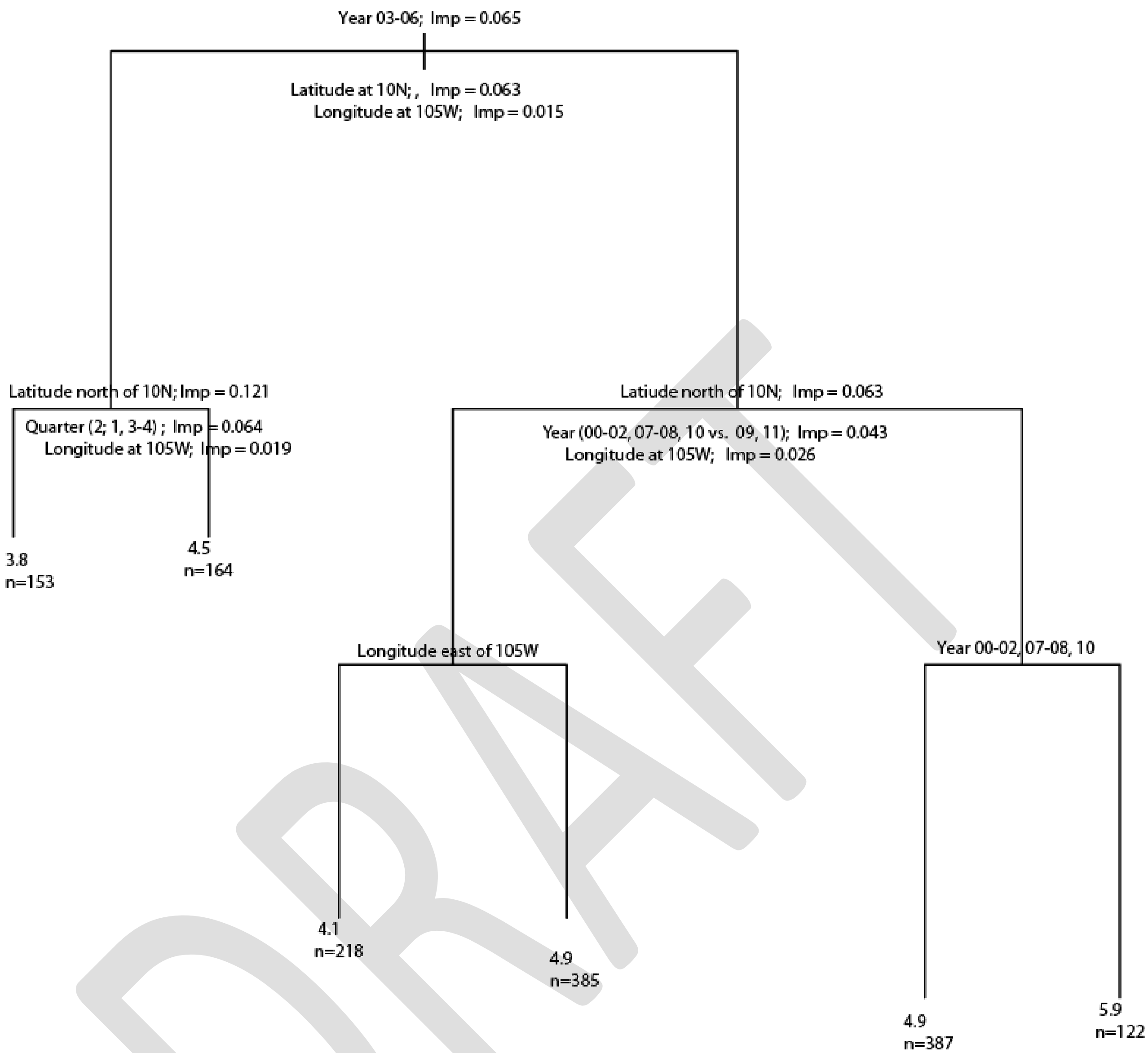

Figure 11. Regression tree obtained for yellowfin tuna average weight in tree area 2 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

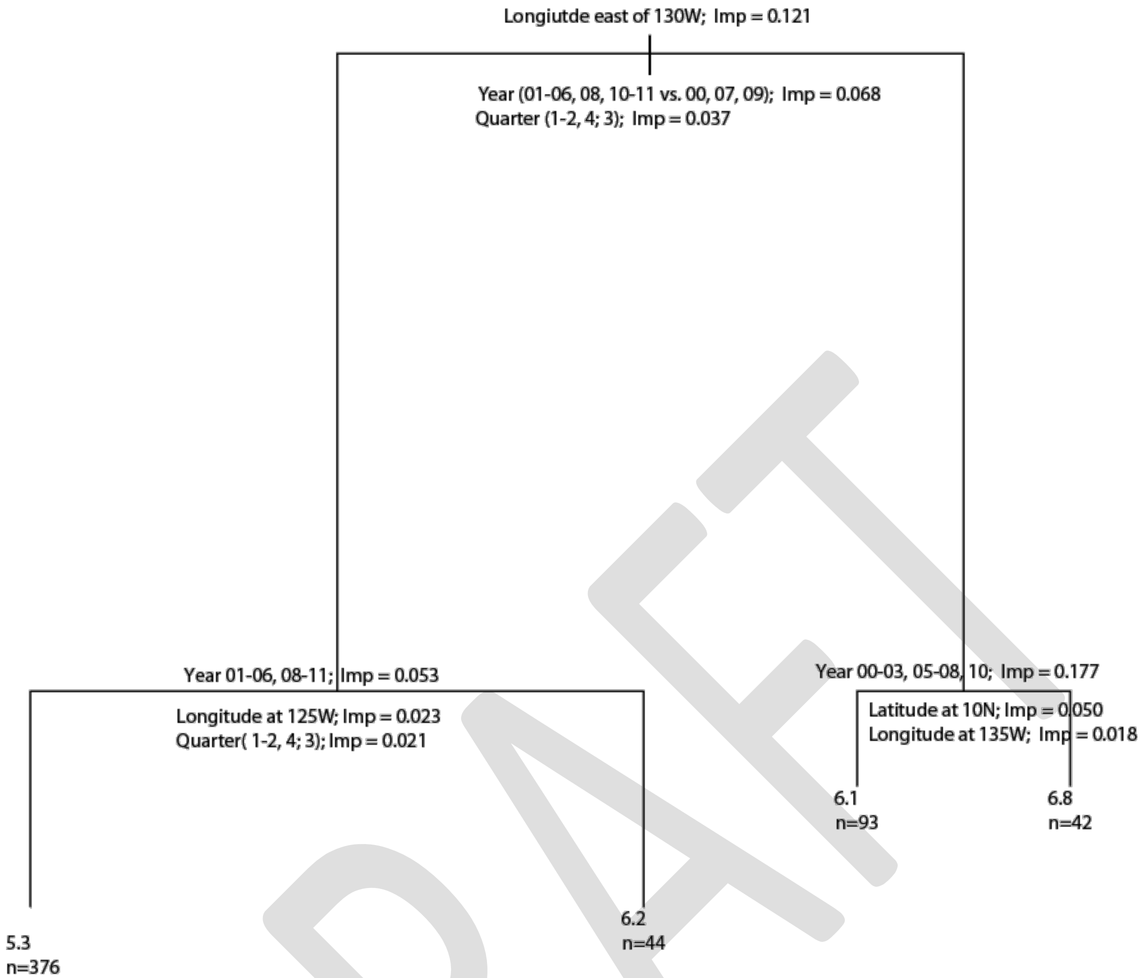

Figure 12. Regression tree obtained for yellowfin tuna average weight in tree area 3 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

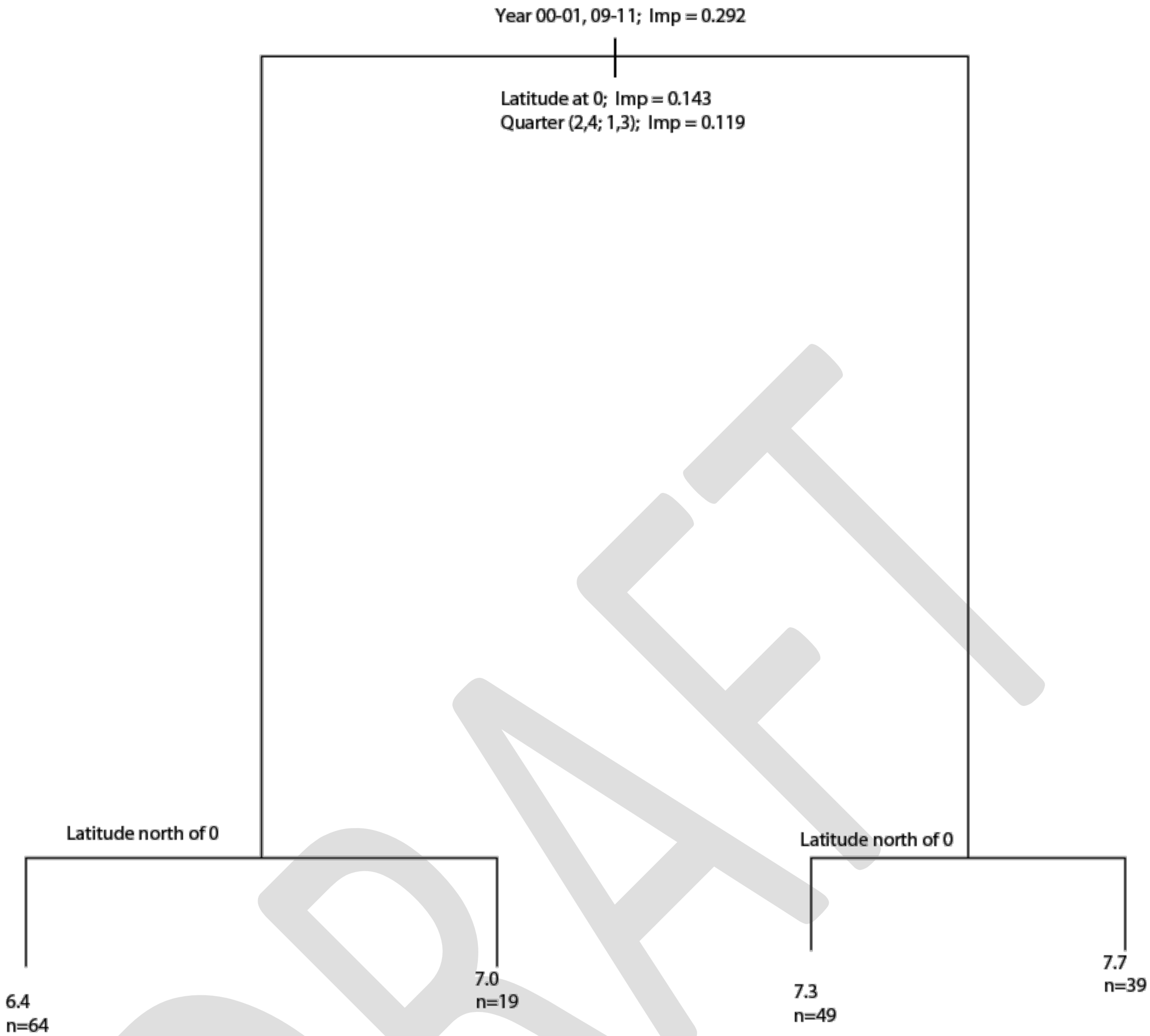

Figure 13. Regression tree obtained for yellowfin tuna average weight in tree area 4 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

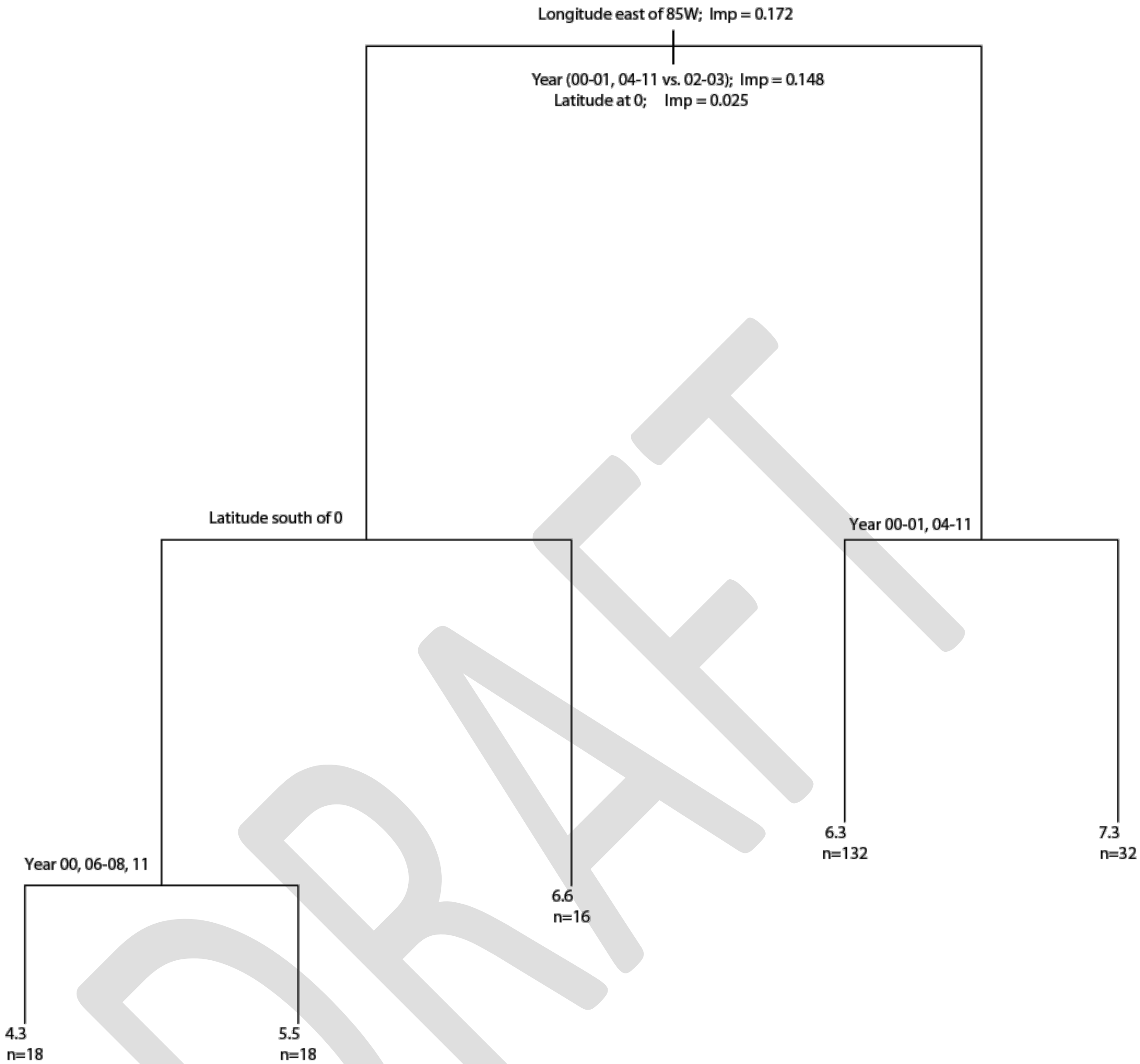

Figure 14. Regression tree obtained for yellowfin tuna average weight in tree area 5 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

DRAFT

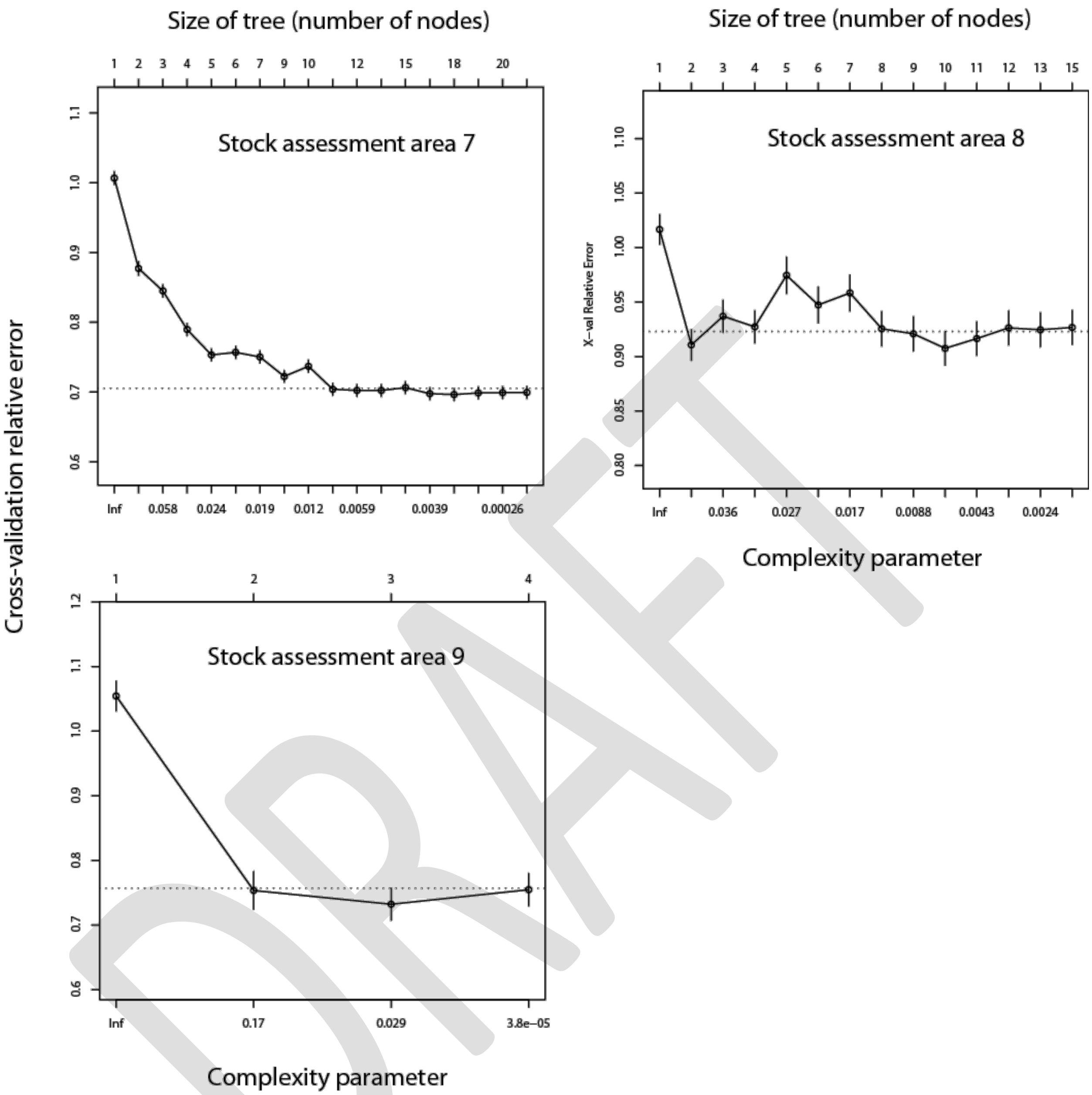

Figure 15. Plots of cross-validation relative error (cross-validation estimate of the tree error, scaled so that the first node has an error of 1.0) as a function of tree size for the regression tree analysis of average weight of skipjack tuna, by stock assessment area (Figure 1b). The complexity parameter is the penalty on tree size. The dashed line indicates the minimum error plus one standard error.

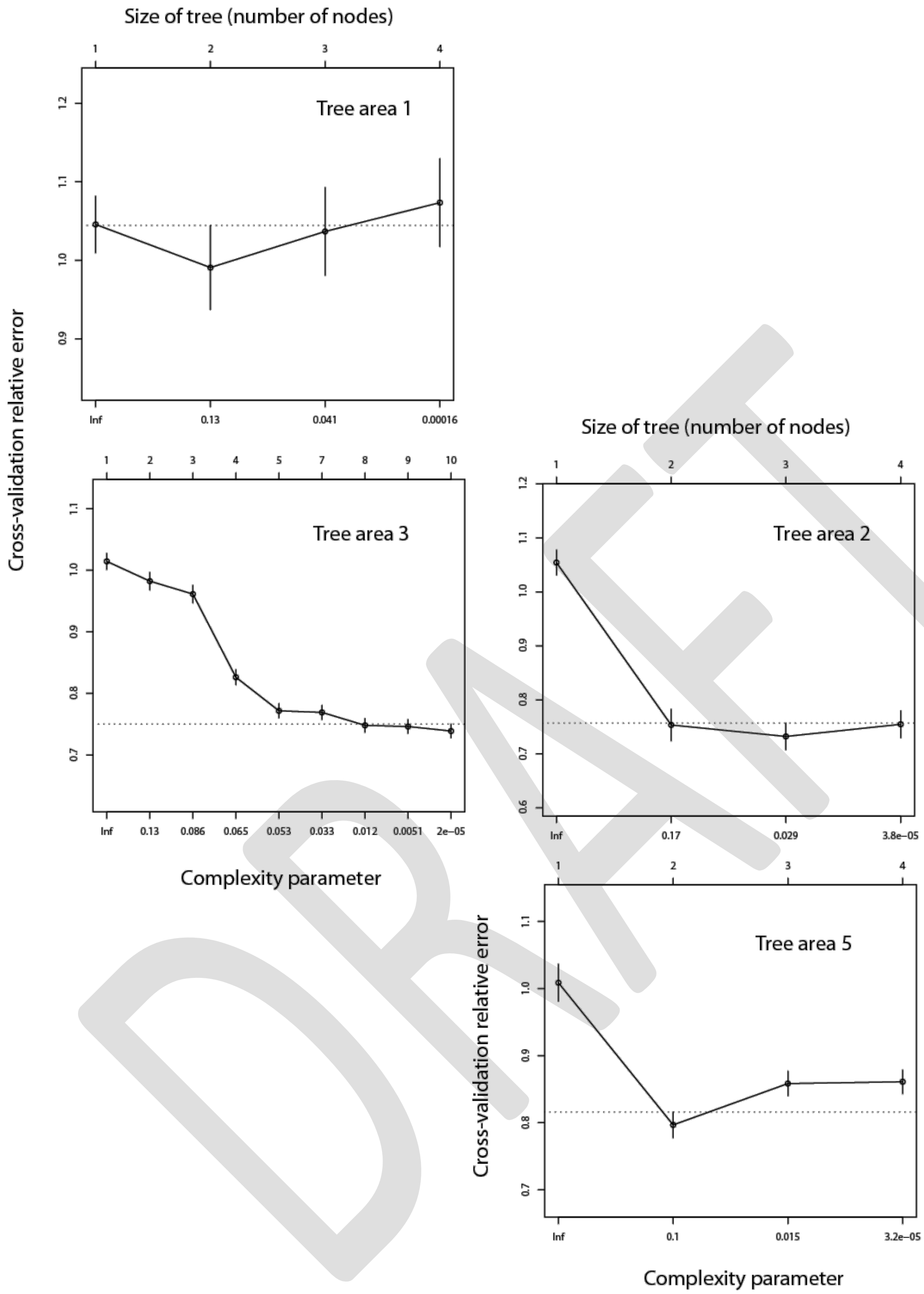

Figure 16. Plots of cross-validation relative error (cross-validation estimate of the tree error, scaled so that the first node has an error of 1.0) as a function of tree size for the regression tree analysis of average weight of skipjack tuna, by tree area (Figure 1c). The complexity parameter is the penalty on tree size. The dashed line indicates the minimum error plus one standard error.

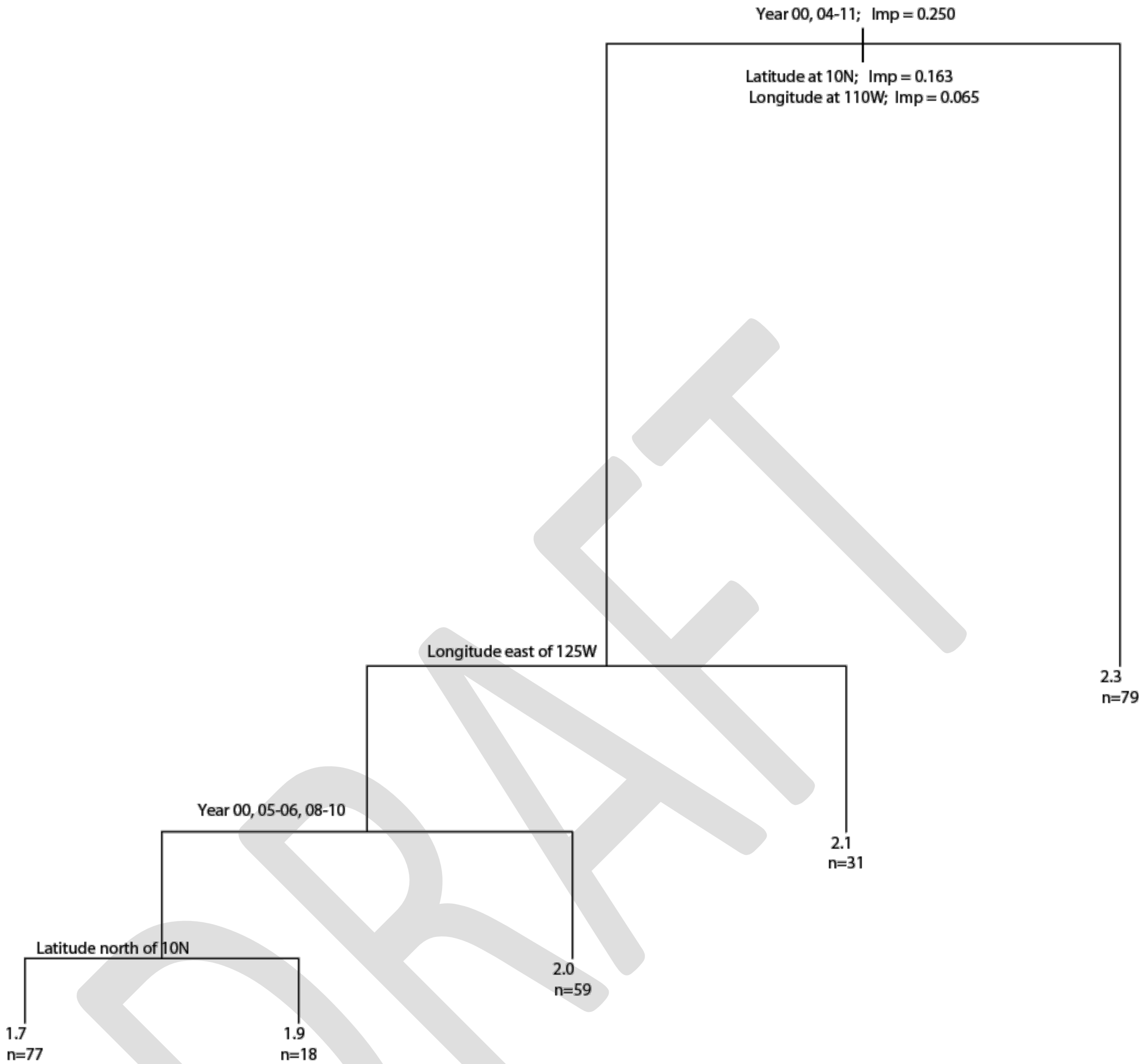

Figure 17. Regression tree obtained for skipjack tuna average weight in stock assessment area 7 (Figure 1b). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

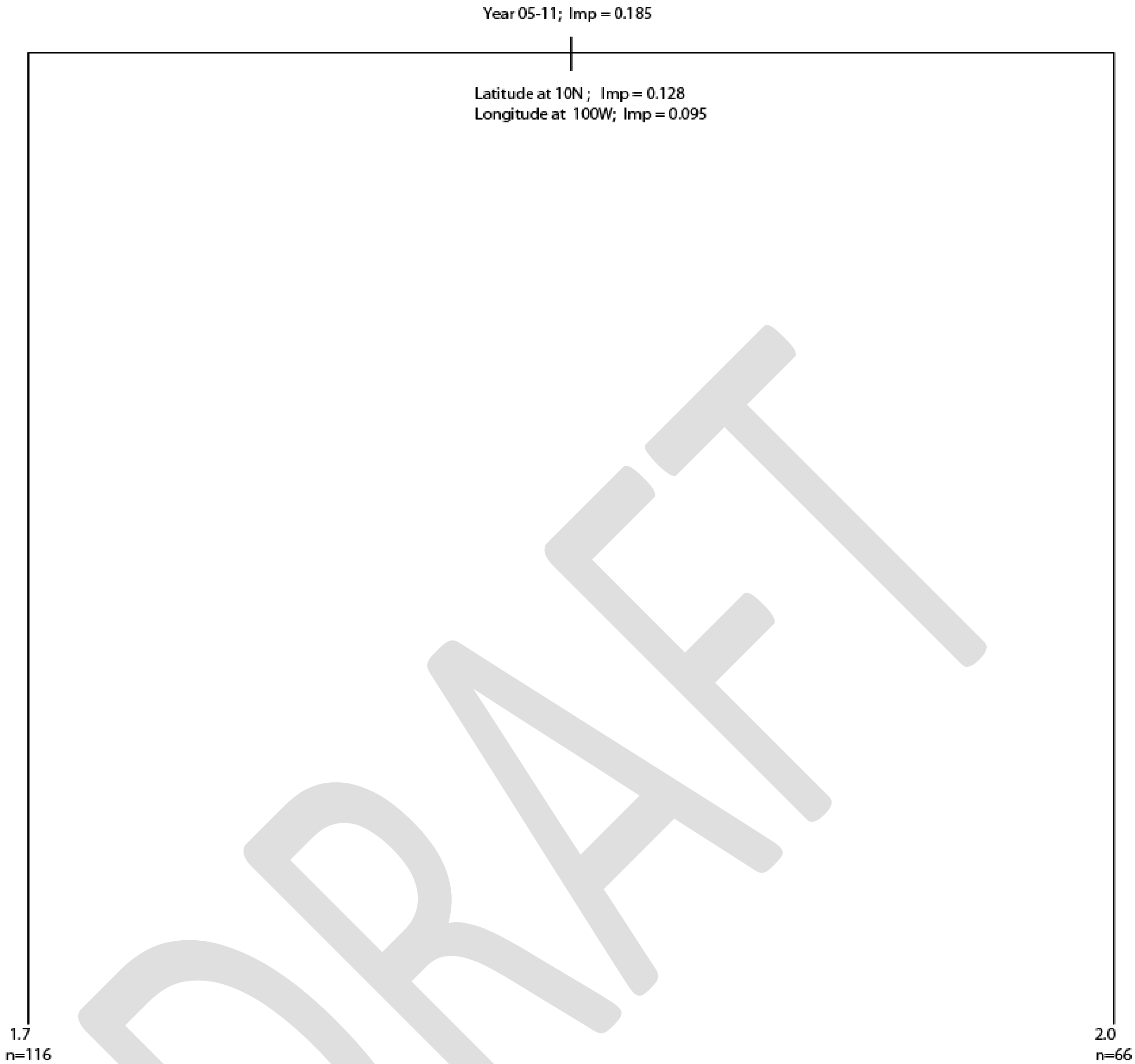

Figure 18. Regression tree obtained for skipjack tuna average weight in stock assessment area 8 (Figure 1b). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

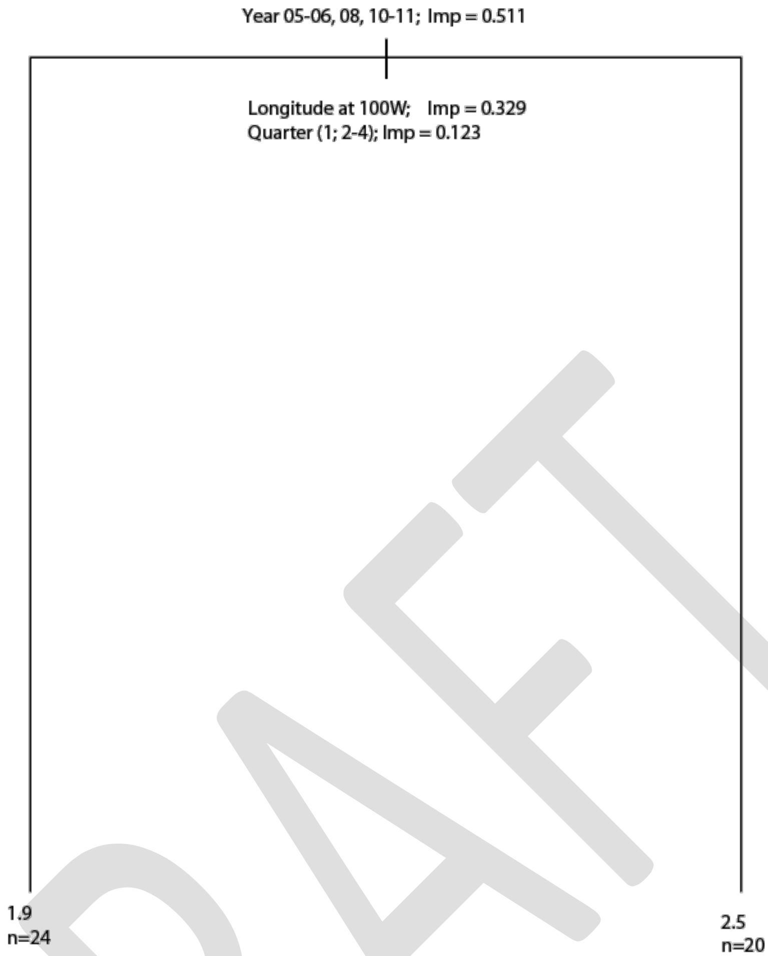

Figure 19. Regression tree obtained for skipjack tuna average weight in stock assessment area 9 (Figure 1b). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

Figure 20. Regression tree obtained for skipjack tuna average weight in tree area 2 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

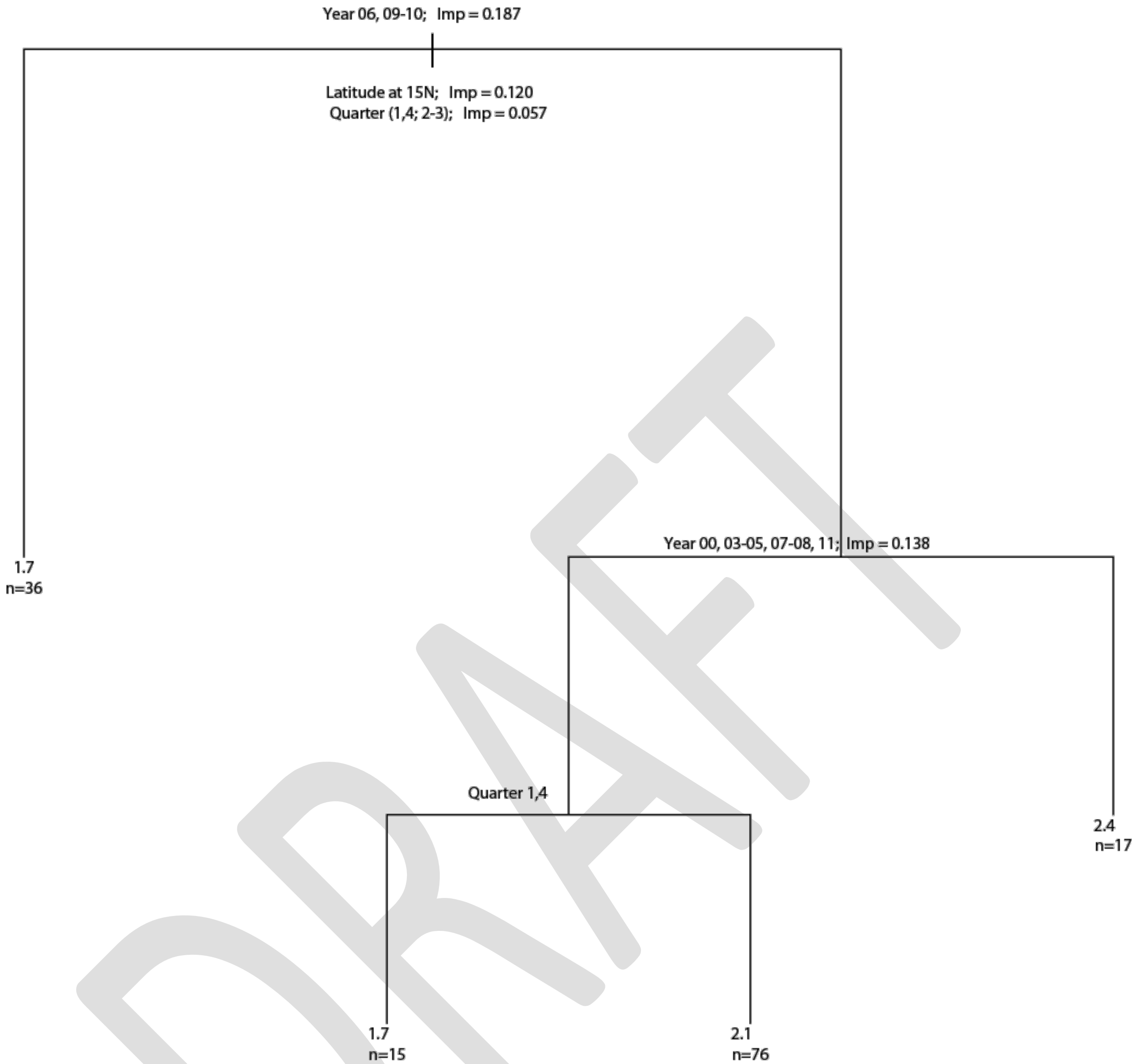

Figure 21. Regression tree obtained for skipjack tuna average weight in tree area 3 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

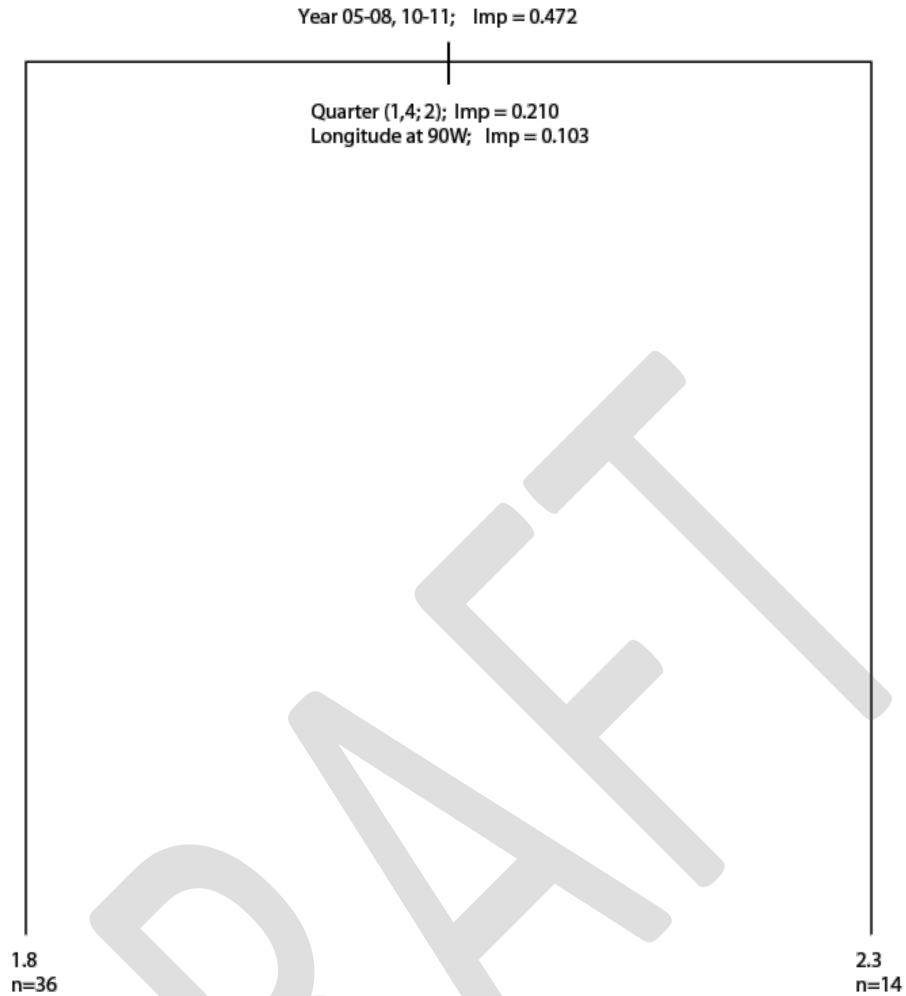

Figure 22. Regression tree obtained for skipjack tuna average weight in tree area 5 (Figure 1c). Text above each branch gives the variable chosen to define the split and the value of that variable that defines the left side of the branch. Text below the branch shows the first two competitor splits. Branch length is proportional to the error of the fit. “Imp” is the improvement achieved by that partition of the data. Shown below each terminal node are the mean square root average weight and the number of samples.

APPENDIX.

Current estimator of total catch by species

The current estimator of total catch by species (Tomlinson 2004) has the general form of a ratio-type estimator of the stratum total (*e.g.*, Thompson 1992) based on the amount of the catch in sampled wells. The total estimated annual catch (in weight) of species *i* (*i* = 1,…, 3 for yellowfin, skipjack and bigeye tunas) in sampling stratum h is given by:

$$\widehat{W}_{hi} = W_h \hat{p}_{hi}$$

$$= W_h \left[\frac{\sum_{j=1}^{q} W_{hj} \left(\frac{\frac{w_{hij}}{m_{hij}} \frac{n_{hij}}{n_{h.j}}}{\sum_{i=1}^{3} \frac{w_{hij}}{m_{hij}} \frac{n_{hij}}{n_{h.j}}} \right)}{\sum_{j=1}^{q} W_{hj}} \right]$$

$$= W_h \left[\frac{\sum_{j=1}^{q} W_{hj} \left(\frac{\overline{w}_{hij} f_{hij}}{\sum_{i=1}^{3} \overline{w}_{hij} f_{hij}} \right)}{\sum_{j=1}^{q} W_{hj}} \right]$$

$$= W_h \frac{\left[\sum_{j=1}^{q} W_{hj} \cdot g(\overline{w}_{hij}, f_{hij}, i = 1, \ldots, 3)\right]}{\left[\sum_{j=1}^{q} W_{hj}\right]} \qquad (1)$$

where W_h is the total weight of all species combined in sampling stratum *h* (assumed known), $\hat{p}$ is the estimate of the species fraction (derived from weight) in the stratum, W_{hj} is the total weight of all species combined in the j^{th} well sampled from sampling stratum *h* (also assumed known), *j*=1, …, *q* wells sampled, *w* is the sum of the weights of fish measured (converted from lengths), *m* is the number of fish measured, *n* is the number of fish counted for species composition, and *g* represents the function of the sample means ($\overline{w}$) and sample species fractions (f) shown in curved brackets (*i.e.*, a function of only the *w*'s, *m*'s and *n*'s).

Table 1. Number of samples available for yellowfin tuna average weight by 5º square area and year, 2000-2011.

2000														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	6	5	0	0	0	0	0	0
15ºN	0	0	0	0	0	2	1	3	5	0	0	0	0	0
10ºN	1	3	6	4	4	2	5	13	7	4	0	1	0	0
5ºN	3	8	7	4	8	0	0	7	15	17	12	19	4	1
0º	0	0	0	4	4	3	1	1	1	2	5	6	2	0
5ºS	0	0	0	0	0	0	1	1	2	0	0	0	1	0
10ºS	0	0	0	0	0	0	0	0	0	1	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	2	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2001														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	5	2	0	0	0	0	0	0
15ºN	0	0	0	0	0	4	13	11	5	3	0	0	0	0
10ºN	0	0	2	4	13	3	8	7	17	12	7	0	0	0
5ºN	0	1	3	9	2	2	9	8	14	19	25	19	3	0
0º	0	0	0	1	2	6	6	1	3	5	4	4	0	0
5ºS	0	0	0	0	0	0	1	5	0	0	0	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	7	12	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	1	4	3	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	4	0

2002														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	4	3	0	0	0	0	0	0
15ºN	0	0	0	0	1	3	14	26	0	3	0	0	0	0
10ºN	0	0	2	9	0	9	13	20	13	13	11	2	0	0
5ºN	0	1	0	1	4	1	11	10	31	36	12	8	0	0
0º	0	0	0	0	2	5	3	4	2	9	2	2	1	0

5ºS	0	1	0	0	0	0	1	1	3	3	2	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	2	2	11	0	0
15ºN	0	0	0	0	0	0	0	0	0	2	2	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2003														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	10	19	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	4	9	2	0	0	0	0	0
10ºN	0	1	0	5	0	9	11	19	7	0	0	0	0	0
5ºN	0	1	6	6	12	6	15	10	10	10	4	6	1	0
0º	0	0	0	1	6	4	1	1	2	2	1	0	0	0
5ºS	0	0	0	0	0	2	4	3	1	0	7	0	3	0
10ºS	0	0	0	0	0	0	0	0	0	0	2	1	1	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2004														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20ºN	0	0	0	0	0	1	13	7	0	0	0	0	0	0
15ºN	0	0	0	0	0	4	3	4	2	4	0	0	0	0
10ºN	0	0	5	11	1	8	9	9	2	0	0	0	0	0
5ºN	0	1	1	1	0	0	0	5	4	9	8	2	1	0
0º	0	0	0	0	1	0	0	0	0	0	8	0	0	0
5ºS	0	0	0	0	1	2	0	3	1	1	6	0	3	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	1	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	2	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2005														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	7	7	0	0	0	0	0	0
15ºN	0	0	0	0	3	7	16	2	3	0	0	0	0	0
10ºN	0	0	1	3	0	7	12	9	5	0	0	0	0	0
5ºN	0	0	4	5	1	4	5	7	10	8	5	7	4	0

0º	0	0	0	1	0	0	0	0	0	0	4	3	4	0
5ºS	0	0	0	0	0	0	0	0	0	0	14	1	2	0
10ºS	0	0	0	0	0	0	0	0	0	0	2	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2006														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	2	10	6	0	0	0	0	0	0
15ºN	0	0	0	1	6	2	5	3	3	0	0	0	0	0
10ºN	0	1	2	6	3	2	1	1	2	0	5	1	0	0
5ºN	0	0	5	5	0	1	1	4	2	3	9	10	3	1
0º	0	0	0	2	0	0	0	0	0	0	0	0	1	0
5ºS	0	0	0	0	0	0	0	0	0	1	5	0	1	0
10ºS	0	0	0	0	0	0	0	0	0	4	1	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2007														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	2	11	0	0	0	0	0	0
15ºN	0	0	0	0	1	2	4	8	1	0	0	0	0	0
10ºN	0	0	1	3	4	5	5	25	6	0	0	0	0	0
5ºN	0	4	3	4	0	1	0	0	0	13	13	2	1	1
0º	0	0	0	0	0	0	0	0	1	0	1	1	0	0
5ºS	0	0	0	0	0	0	0	1	1	0	0	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	3	2	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	1	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2008														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	10	5	0	0	0	0	0	0
15ºN	0	0	0	0	0	2	7	13	2	0	0	0	0	0
10ºN	2	0	6	6	1	13	6	7	4	0	8	2	0	0

5ºN	0	5	3	6	2	4	1	4	5	5	15	5	4	0
0º	0	0	1	1	0	0	0	0	0	0	0	3	0	0
5ºS	0	0	0	0	0	0	0	1	0	0	3	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	1	4	11	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2009														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	1	24	7	0	0	0	0	0	0
15ºN	0	0	0	0	3	11	7	3	5	0	0	0	0	0
10ºN	2	4	14	12	4	5	3	7	12	7	5	0	0	0
5ºN	1	2	3	6	3	5	1	0	1	22	21	15	3	0
0º	0	0	0	2	3	0	0	0	0	1	1	0	0	0
5ºS	0	0	0	0	0	0	0	2	1	4	7	0	1	0
10ºS	0	0	0	0	0	0	0	0	0	0	1	5	1	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2010														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	1	3	1	0	0	0	0	0	0
20ºN	0	0	0	0	0	2	22	10	0	0	0	0	0	0
15ºN	0	0	0	0	0	20	22	18	9	0	0	0	0	0
10ºN	0	0	3	2	6	9	14	14	15	21	2	0	0	0
5ºN	0	0	8	4	7	4	6	0	2	6	11	11	2	0
0º	0	0	0	0	1	0	0	0	0	1	0	0	4	1
5ºS	0	0	0	0	0	0	0	0	0	1	2	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	1	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2011														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	16	12	0	0	0	0	0	0
15ºN	0	0	0	0	1	13	21	36	8	0	0	0	0	0

10ºN	0	0	1	4	7	8	12	19	4	3	0	1	0	0
5ºN	0	2	6	9	13	2	4	8	11	15	3	8	8	2
0º	0	2	5	8	1	0	0	0	0	0	0	5	2	1
5ºS	0	0	0	0	0	0	0	0	0	0	2	0	1	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	1	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 2. Number of samples available for skipjack tuna average weight by 5º square area and year, 2000-2011.

2000														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	3	1	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10ºN	0	0	0	1	0	1	1	0	0	0	0	0	0	0
5ºN	0	0	0	0	1	0	0	0	0	1	1	1	0	0
0º	0	0	0	0	0	0	0	0	0	1	0	0	0	0
5ºS	0	0	0	0	0	0	0	0	0	0	0	0	1	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	1	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2001														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	3	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	1	1	0	0	0	0	0	0	0
10ºN	0	0	0	0	4	0	2	0	0	0	0	0	0	0
5ºN	0	0	0	1	1	1	1	1	2	2	5	3	0	0
0º	0	0	0	0	1	1	3	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	0	0	1	0	0	0	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	1	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	2	0

2002														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	1	2	1	0	0	0	0	0	0
10ºN	0	0	1	1	0	4	3	0	0	1	0	0	0	0
5ºN	0	0	0	0	1	1	3	6	5	2	0	0	0	0
0º	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0

15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2003														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	1	2	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	1	4	0	0	0	0	0	0
10ºN	0	0	0	1	0	0	2	5	4	0	0	0	0	0
5ºN	0	0	1	2	2	0	8	5	8	9	1	0	0	0
0º	0	0	0	0	0	0	1	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	1	0	0	1	0	0	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2004														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	1	5	1	0	0	0	0	0	0
15ºN	0	0	0	0	0	2	1	0	0	1	0	0	0	0
10ºN	0	0	0	7	1	4	1	0	0	0	0	0	0	0
5ºN	0	0	0	0	0	0	0	0	2	3	1	0	1	0
0º	0	0	0	0	0	0	0	0	0	0	2	0	0	0
5ºS	0	0	0	0	0	0	0	3	0	0	2	0	2	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2005														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	6	5	0	0	0	0	0	0
15ºN	0	0	0	0	3	7	9	1	3	0	0	0	0	0
10ºN	0	0	1	0	0	1	7	5	2	0	0	0	0	0
5ºN	0	0	2	1	1	2	2	0	3	2	1	0	0	0
0º	0	0	0	0	0	0	0	0	0	0	2	0	0	0
5ºS	0	0	0	0	0	0	0	0	0	0	10	0	0	0

10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2006														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	2	5	0	0	0	0	0	0	0
15ºN	0	0	0	1	5	2	4	0	0	0	0	0	0	0
10ºN	0	0	0	2	1	2	1	0	2	0	0	0	0	0
5ºN	0	0	1	1	0	0	0	1	1	1	1	0	0	1
0º	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	0	0	0	0	0	3	0	1	0
10ºS	0	0	0	0	0	0	0	0	0	1	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2007														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	2	3	0	0	0	0	0	0
15ºN	0	0	0	0	1	0	1	1	0	0	0	0	0	0
10ºN	0	0	0	0	1	1	0	5	1	0	0	0	0	0
5ºN	0	0	1	0	0	1	0	0	0	5	3	0	0	0
0º	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	0	0	1	0	0	0	1	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	2	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	1	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2008														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	6	1	0	0	0	0	0	0
15ºN	0	0	0	0	0	1	1	0	0	0	0	0	0	0
10ºN	0	0	1	3	1	6	5	0	0	0	7	1	0	0
5ºN	0	1	0	3	1	0	0	0	0	0	0	0	0	0
0º	0	0	0	0	0	0	0	0	0	0	0	0	0	0

5ºS	0	0	0	0	0	0	0	1	0	0	1	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	3	9	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2009														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	3	0	0	0	0	0	0	0
15ºN	0	0	0	0	1	4	1	0	0	0	0	0	0	0
10ºN	0	0	0	0	3	1	1	0	0	0	0	0	0	0
5ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0º	0	0	0	0	1	0	0	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	0	0	0	0	0	2	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	1	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2010														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	2	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	4	2	2	0	0	0	0	0	0
10ºN	0	0	0	1	1	1	0	0	5	2	0	0	0	0
5ºN	0	0	1	1	3	0	1	0	1	0	0	0	0	0
0º	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5ºS	0	0	0	0	0	0	0	0	0	1	1	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2011														
	140ºW	135ºW	130ºW	125ºW	120ºW	115ºW	110ºW	105ºW	100ºW	95ºW	90ºW	85ºW	80ºW	75ºW
25ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	2	0	0	0	0	0	0
15ºN	0	0	0	0	1	10	11	2	1	0	0	0	0	0
10ºN	0	0	0	0	5	8	6	4	0	3	0	0	0	0
5ºN	0	0	1	1	2	0	0	0	0	3	0	0	0	0

0º	0	0	0	0	0	0	0	0	0	0	0	0	0	1
5ºS	0	0	0	0	0	0	0	0	0	0	1	0	0	0
10ºS	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20ºN	0	0	0	0	0	0	0	0	0	0	0	0	0	0

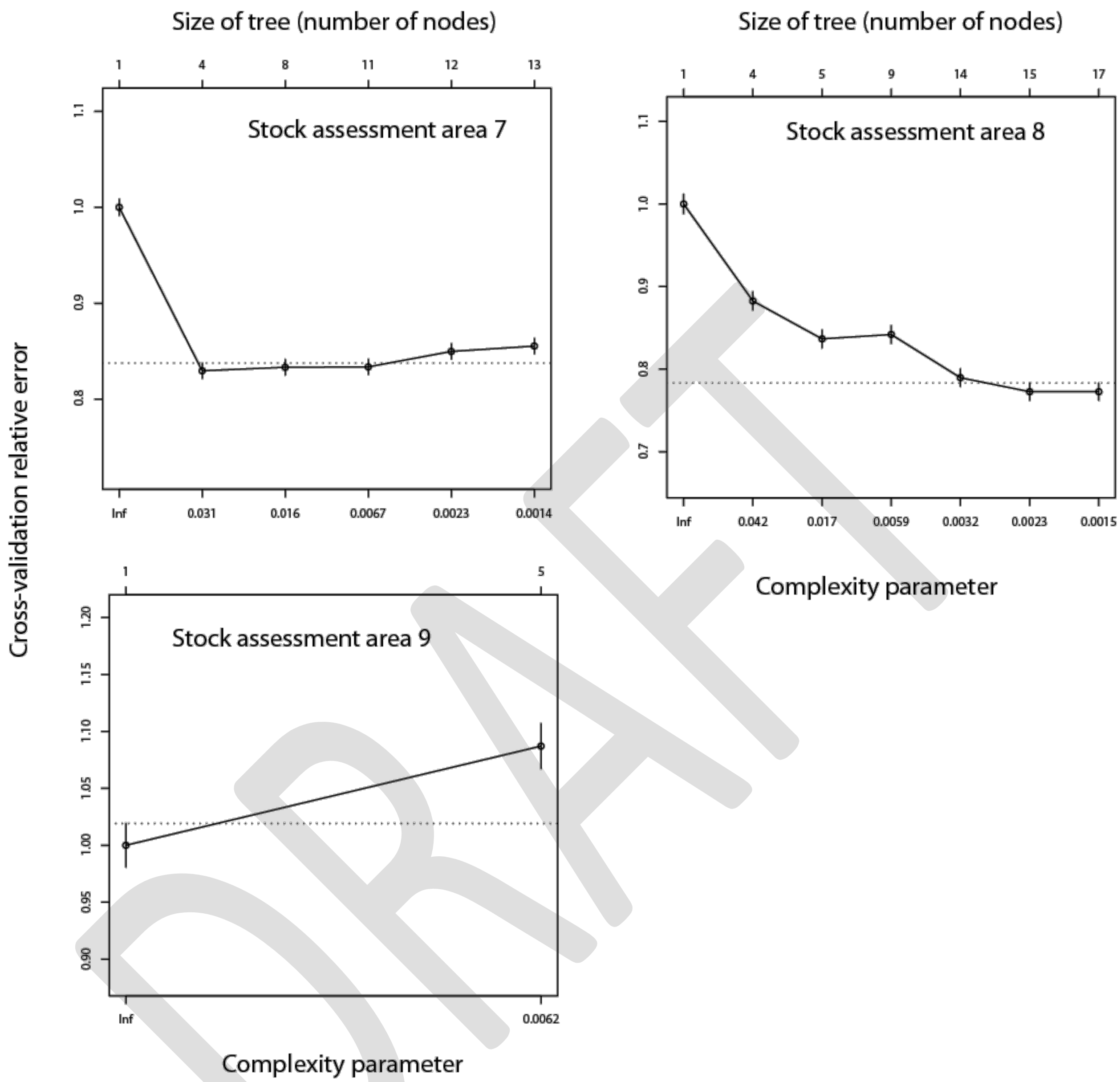

Figure 1. Plots of cross-validation relative error as a function of tree size for the classification tree analysis of the presence/absence of skipjack tuna, by stock assessment area (Figure 1b).

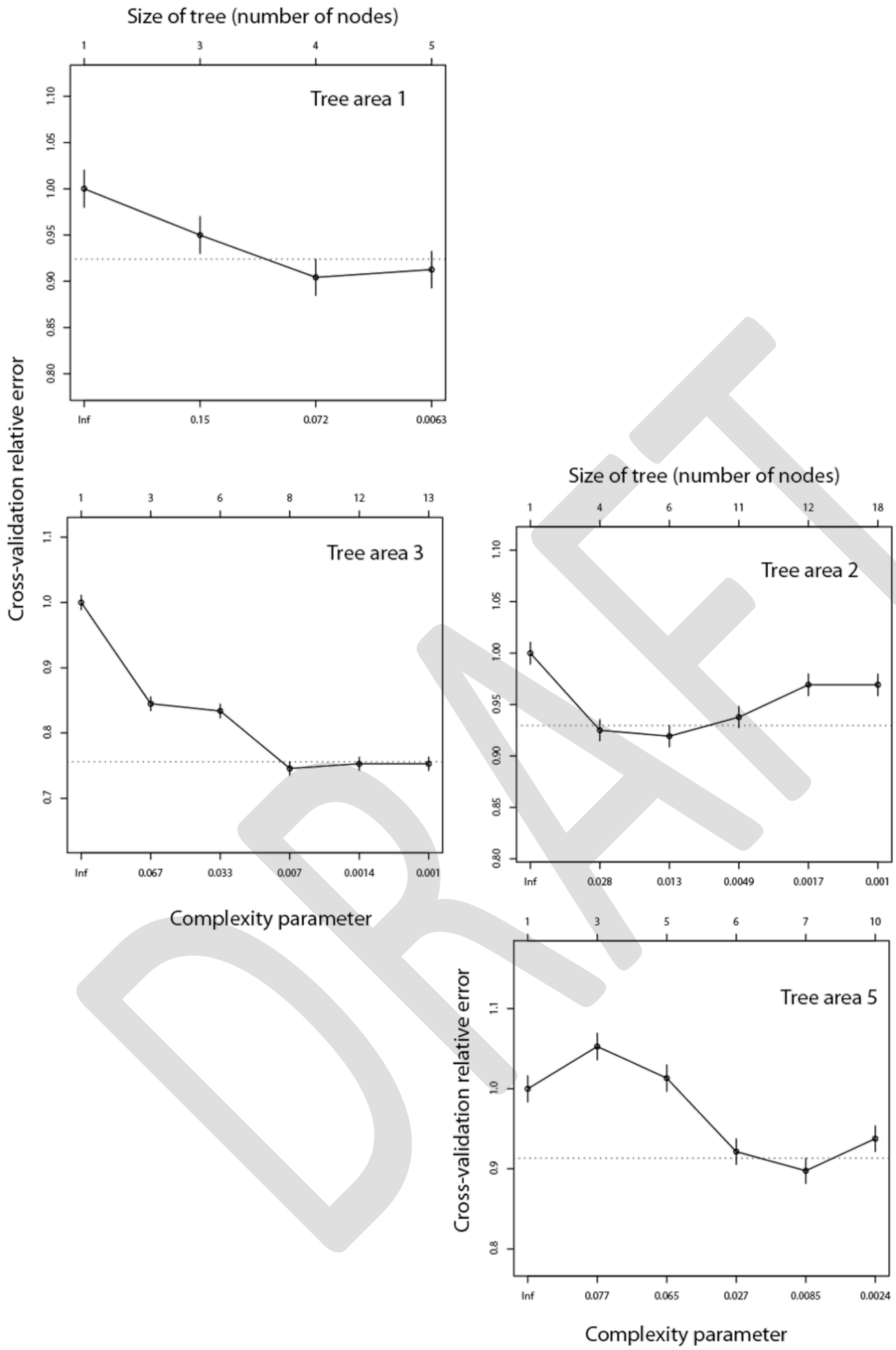

Figure 2. Plots of cross-validation relative error as a function of tree size for the classification tree analysis of the presence/absence of skipjack tuna, by tree area (Figure 1c).

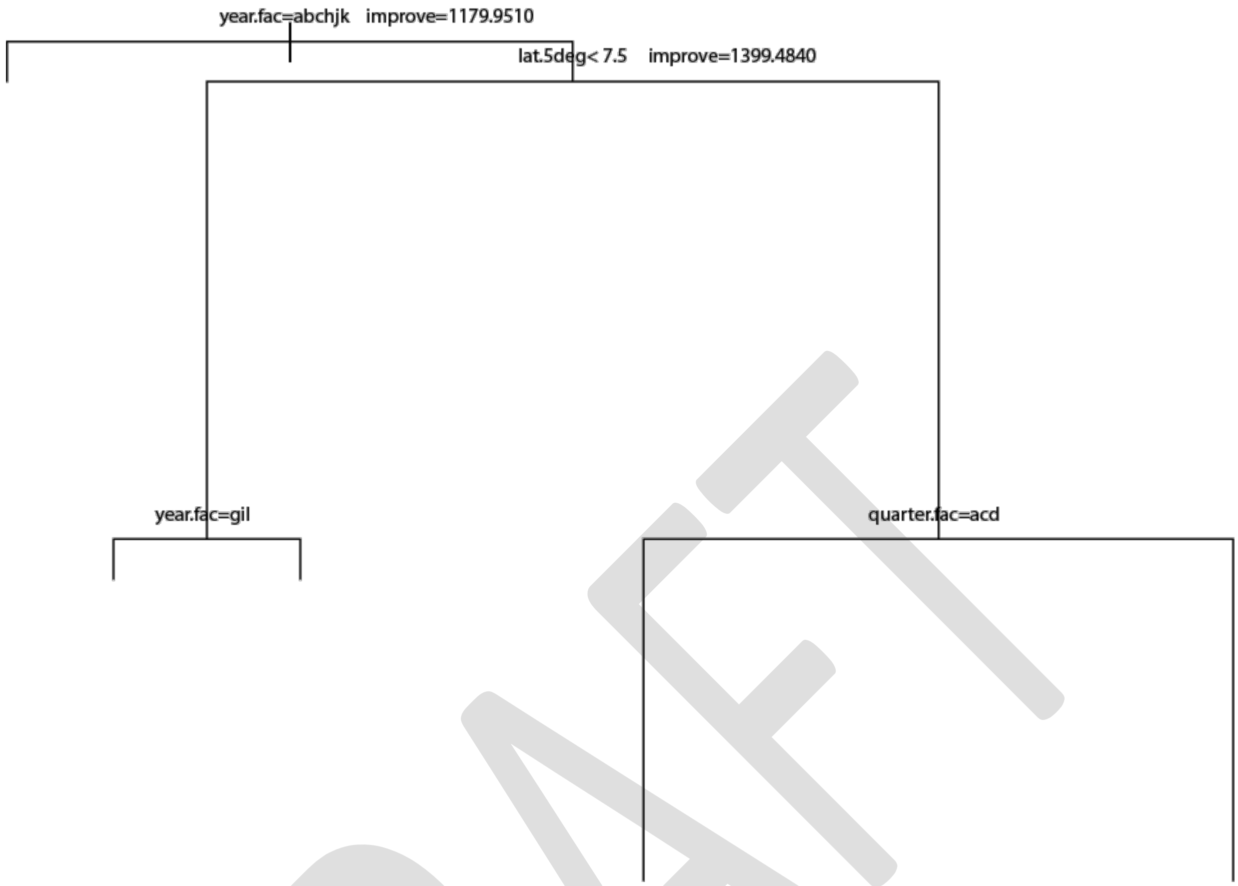

Figure 3. Example of the classification tree obtained for presence/absence of skipjack tuna in stock assessment area 7 (Figure 1b).

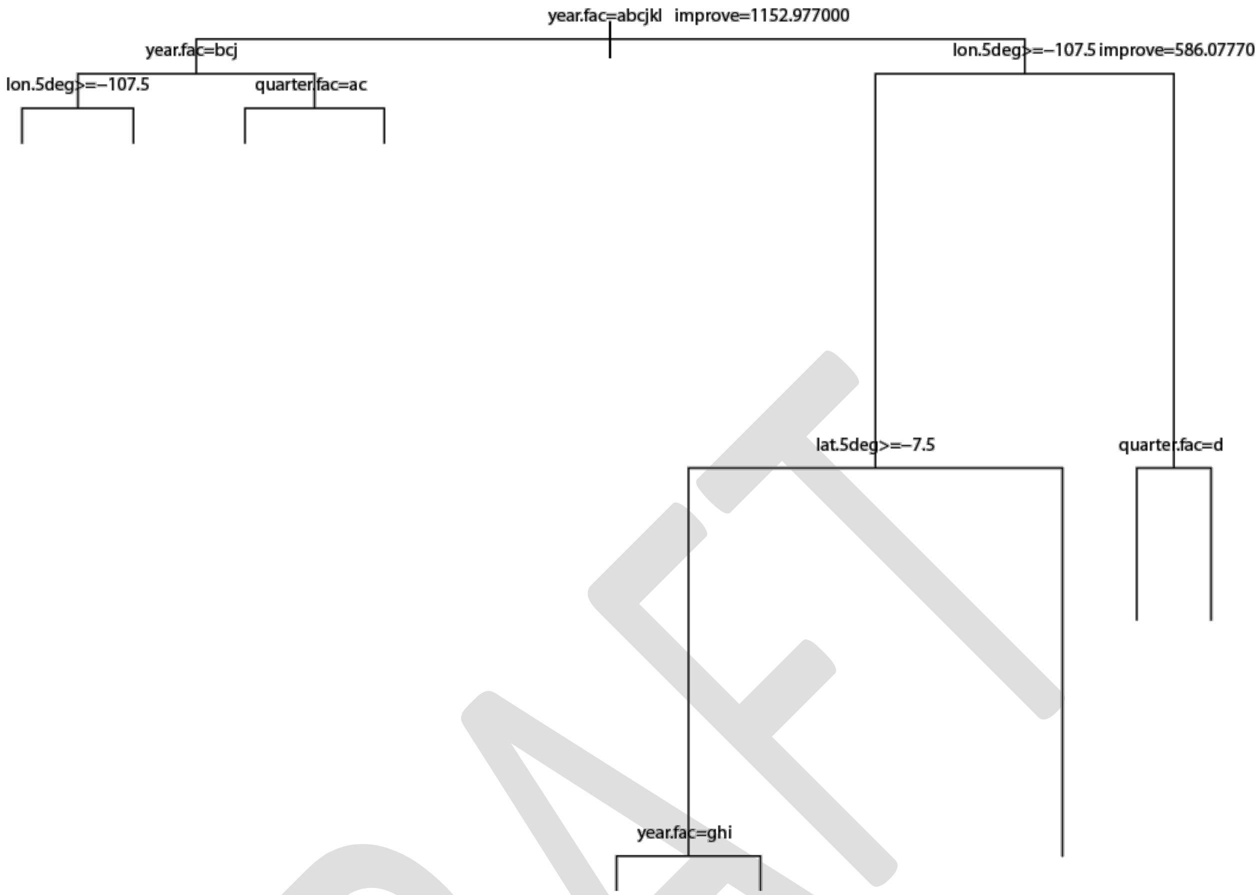

Figure 4. Example of the classification tree obtained for presence/absence of skipjack tuna in stock assessment area 8 (Figure 1b).

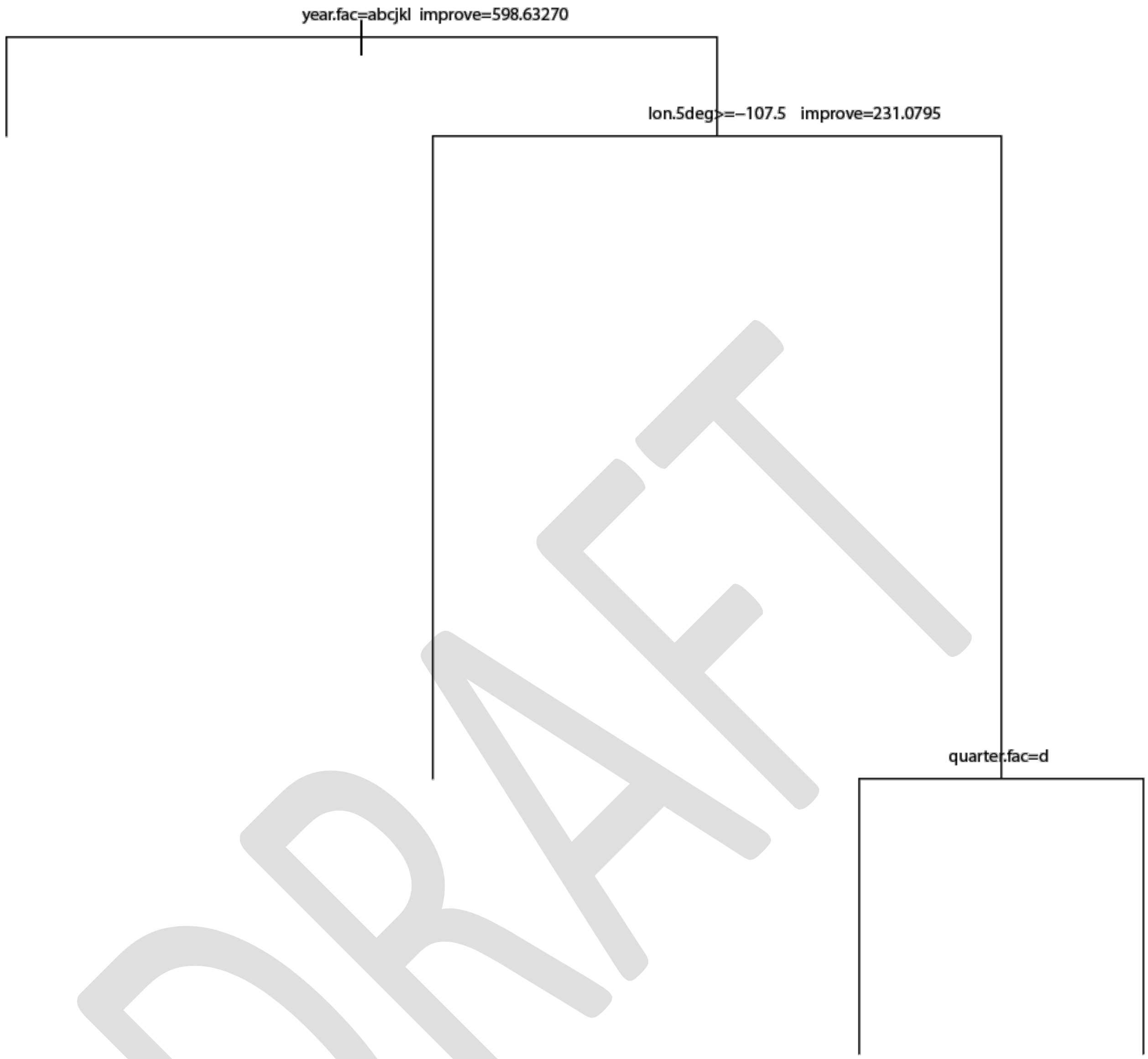

Figure 5. Example of the classification tree obtained for presence/absence of skipjack tuna in tree area 1 (Figure 1c).

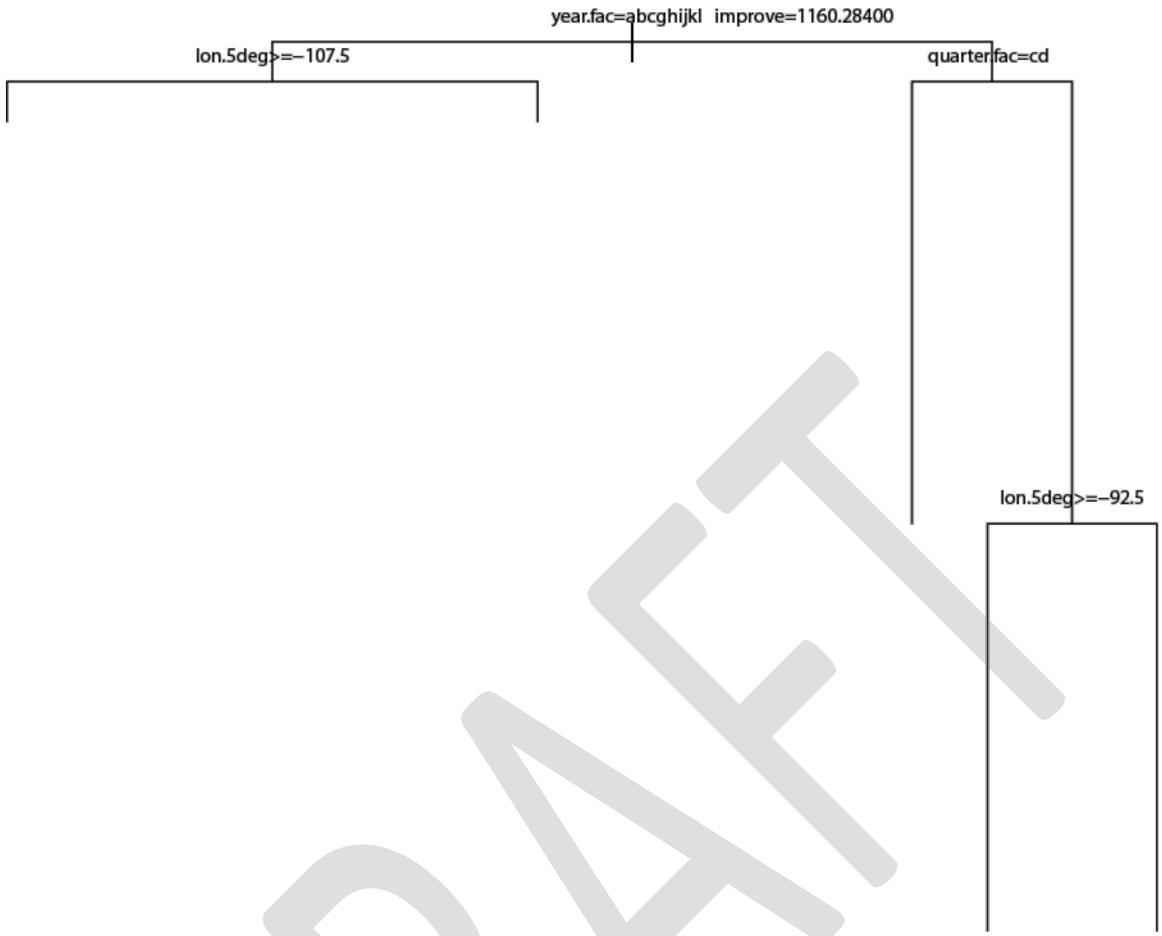

Figure 6. Example of the classification tree obtained for presence/absence of skipjack tuna in tree area 2 (Figure 1c).

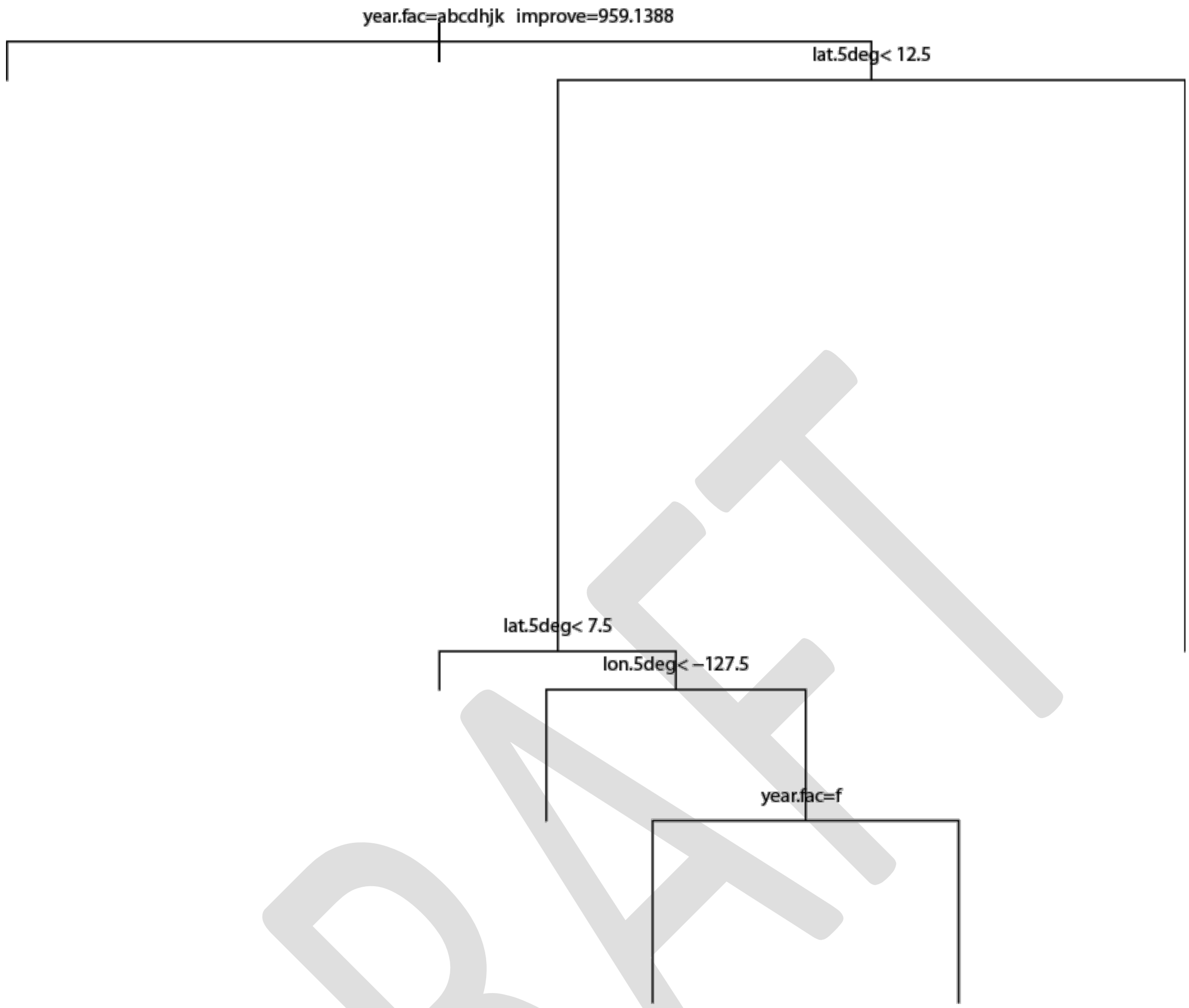

Figure 7. Example of the classification tree obtained for presence/absence of skipjack tuna in tree area 3 (Figure 1c).

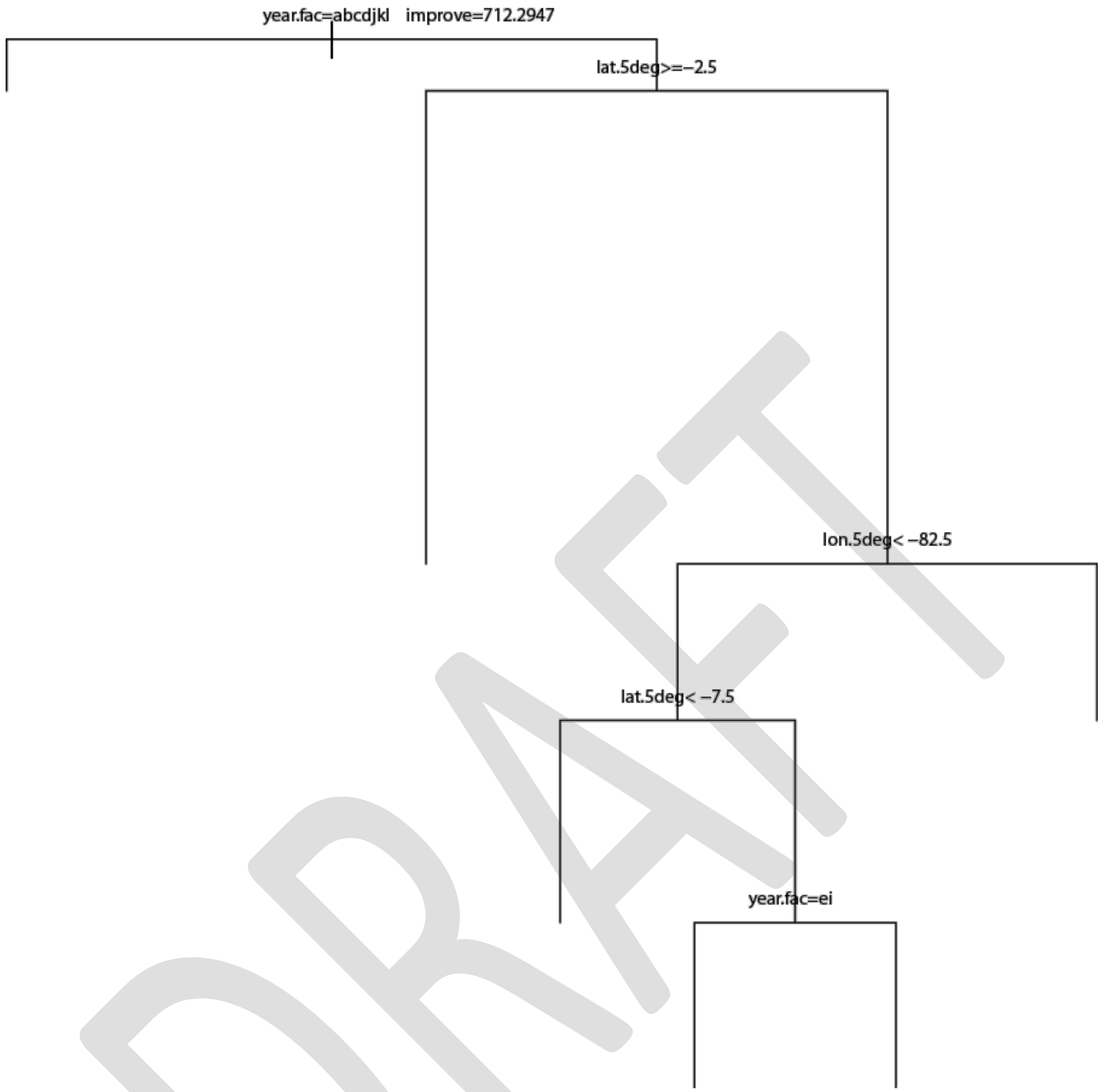

Figure 8. Example of the classification tree obtained for presence/absence of skipjack tuna in tree area 5 (Figure 1c).