Progress report on the development of poststratified estimators of total catch for the purse-seine fishery port-sampling data

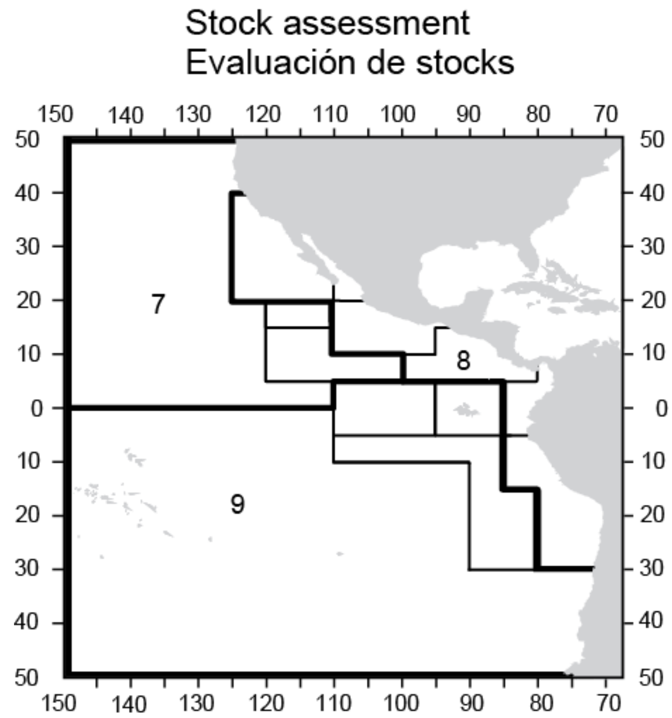
SAC-03-10

Outline of presentation

- Motivation
- Background
    - Port-sampling data
    - Estimation of total catch by species
- Preliminary analysis of stock assessment areas
    - Methods of analysis
    - Results
- Preliminary analysis of poststratification
    - Methods of analysis
    - Results
- Future work

Why consider poststratification?

- IATTC tuna stock assessments presently use large areas formed from sampling areas.
- As fisheries evolve over time, it is useful to consider alternative assessment areas.
- To do this, modifications to the methodology for estimating total catch by species are needed.
- Poststratification may simplify treatment of ‘missing data’ (strata with catch but no samples).
- The present work focuses on the purse-seine fishery for dolphin sets from 2000-2011, with emphasis on yellowfin tuna.



Background: port-sampling data

- A stratified two-stage sampling procedure is used to sample tuna catches in port:
    - 1st stage: vessel wells
        - sampled largely opportunistically;
        - only sampled if catch is from the same "stratum."
    - 2nd stage: fish within a well
        - individual fish selected from an opportunistically-established starting point during unloading;
        - approximately 50 fish of each species are measured;
        - independent of the measured fish, several hundred fish are counted for species composition;
        - sampling differs slightly for catches unloaded by species and size.

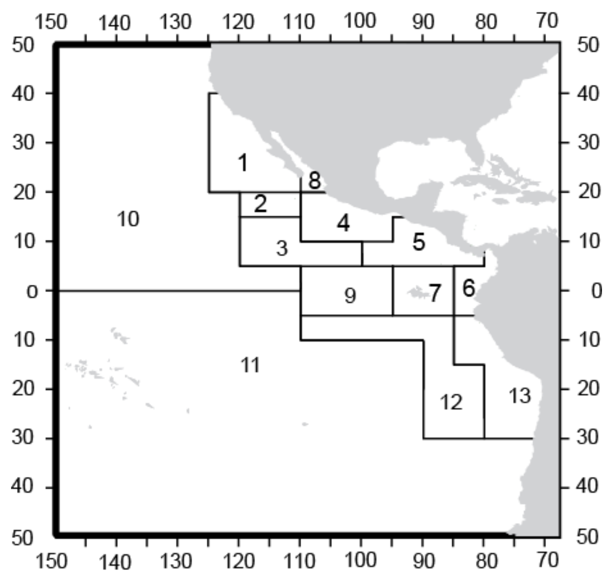
Background: port-sampling data

To obtain a representative data set, the fishery is divided into categories (“strata”) that are defined by:

Area

Month

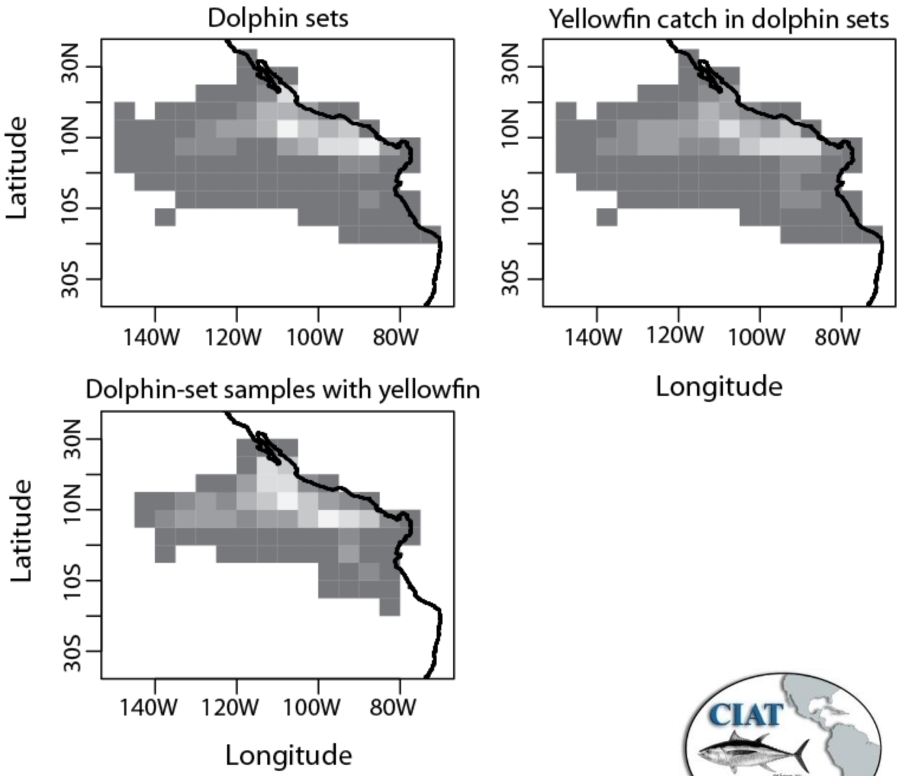
Mode of fishing



Type of vessel	Type of set
small purse-seiner	floating-object
"	unassociated
"	dolphin
large purse-seiner	floating-object
"	unassociated
"	dolphin

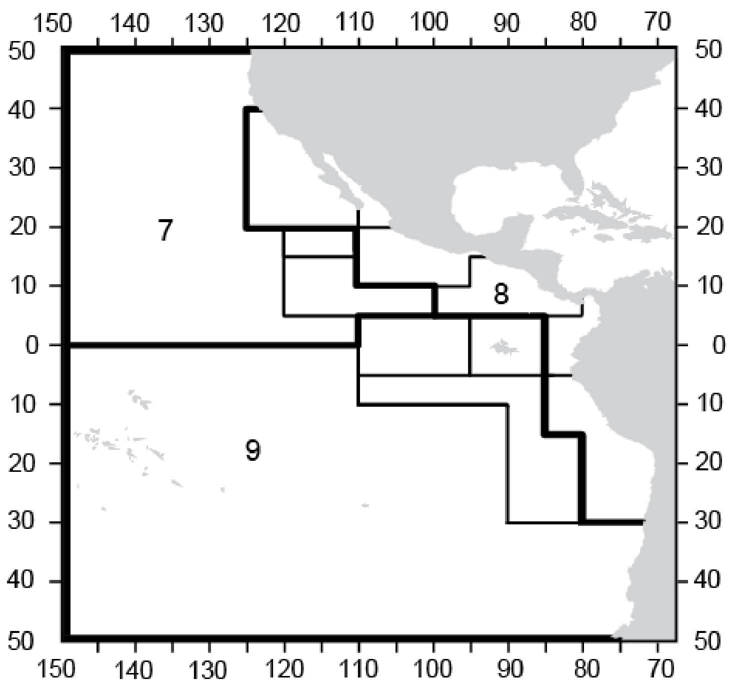
Background: port-sampling data

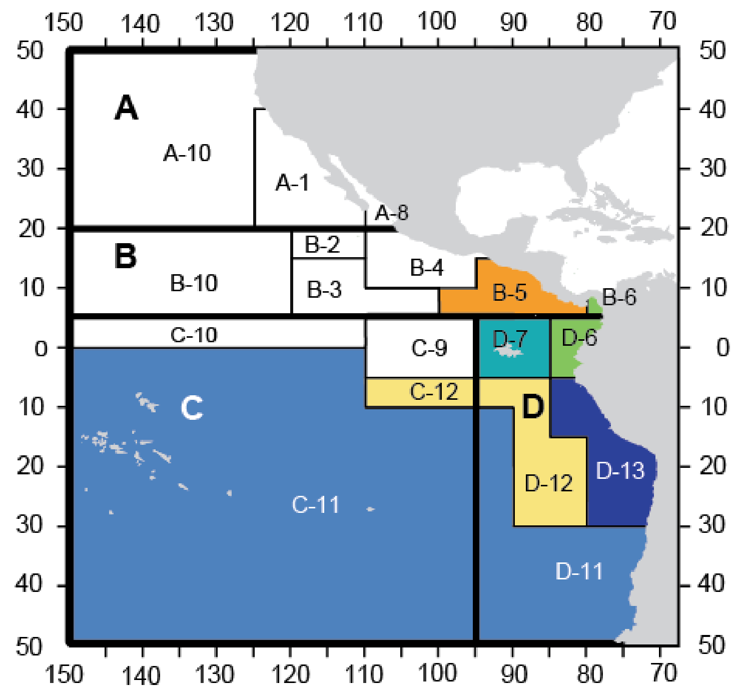
- Sample data contain information on both the sampling area and the 5º area.
- The detailed spatial data compare well with actual set positions:
  - 81% of samples agree at level of 5º area;
  - 97% are within one 5º area.
- Sampling is generally proportional to effort and catch.



Background: estimation of catch by species

- Sampling strata are defined by area, month and fishing mode.
- Alternative assessment areas may have boundaries that cross sampling areas.
- For this reason modification to the methodology for estimation of total catch by species may be necessary.

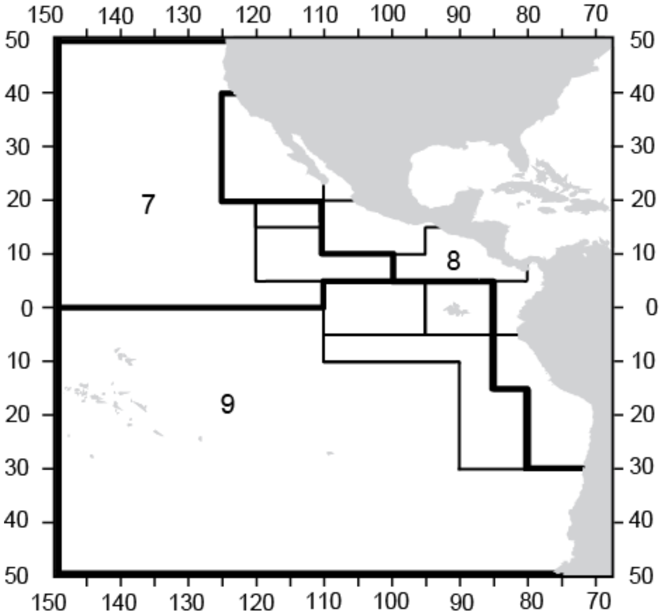




Background: estimation of catch by species

Current estimator of species catch (in weight):

$$\widehat{W}_{hi} = W_h \,\widehat{p}_{hi} \quad = W_h \left[\frac{\sum_{j=1}^{q} W_{hj} \left(\dfrac{\dfrac{w_{hij}}{m_{hij}} \dfrac{n_{hij}}{n_{h.j}}}{\sum_{i=1}^{3} \dfrac{w_{hij}}{m_{hij}} \dfrac{n_{hij}}{n_{h.j}}} \right)}{\sum_{j=1}^{q} W_{hj}} \right] \quad = W_h \frac{\left[\sum_{j=1}^{q} W_{hj} \cdot g(\ldots)\right]}{\left[\sum_{j=1}^{q} W_{hj}\right]}$$



W_h : total weight of all species combined in sampling stratum *h (h*: **area** x month x fishing mode);

$\hat{p}$: estimate of the fraction of species *i* (in weight);

W_{hj} : total weight of all species combined for the j^{th} well sample in *h*;

w : sum of the weights of fish measured (from lengths);

m : number of fish measured;

n : number of fish counted;

g : function of the sample means and species fractions.

Background: estimation of catch by species

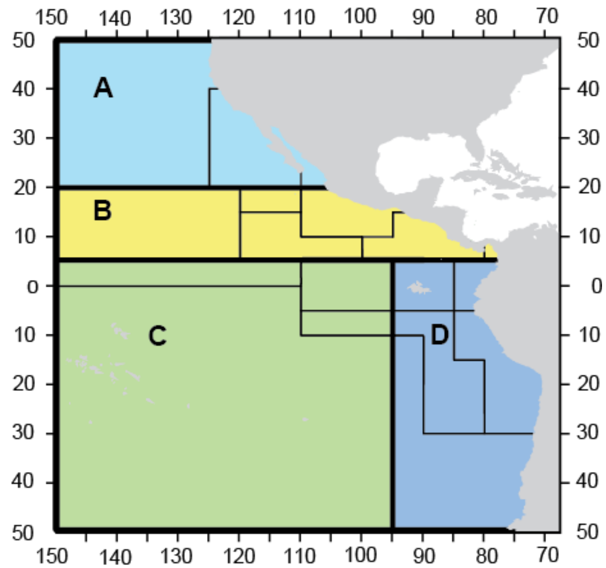
- Suppose new stock assessment areas have been defined.
- Two different ways to modify the current methods for estimating fishery totals:

1) Assume that only the stock assessment areas are important.

2) Assume both the sampling areas and the stock assessment areas need to be taken into consideration.

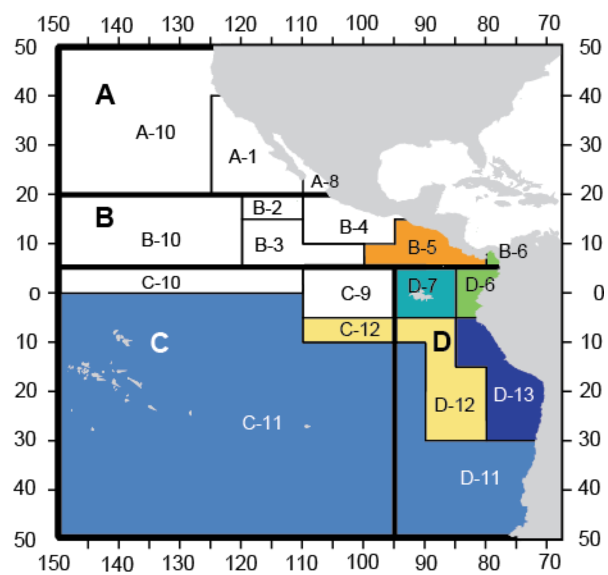
Background: estimation of catch by species

1) Stock assessment areas:



$$\widehat{W}_{ps-II;\,ci} = W_c \frac{\left[\sum_{j\in c} W_{cj}\cdot g(\ldots)\right]}{\left[\sum_{j\in c} W_{cj}\right]}$$

2) Sampling and stock assessment areas:



$$\widehat{W}_{ps-I;\,ci} = \sum\nolimits_{h:\,h\cap c} W_{h\cap c}\frac{\left[\sum_{j\in h\cap c} W_{hj}\cdot g(\ldots)\right]}{\left[\sum_{j\in h\cap c} W_{hj}\right]}$$

CIAT IATTC

Outline of presentation

- Motivation
- Background
  - Port-sampling data
  - Estimation of total catch by species
- Preliminary analysis of stock assessment areas
  - Methods of analysis
  - Results
- Preliminary analysis of poststratification
  - Methods of analysis
  - Results
- Future work

Preliminary analysis of stock assessment areas

Purpose: identify large areas with similar population characteristics.

Two types of data used:

Length-frequency distributions (port-sampling data), 2000-2011;

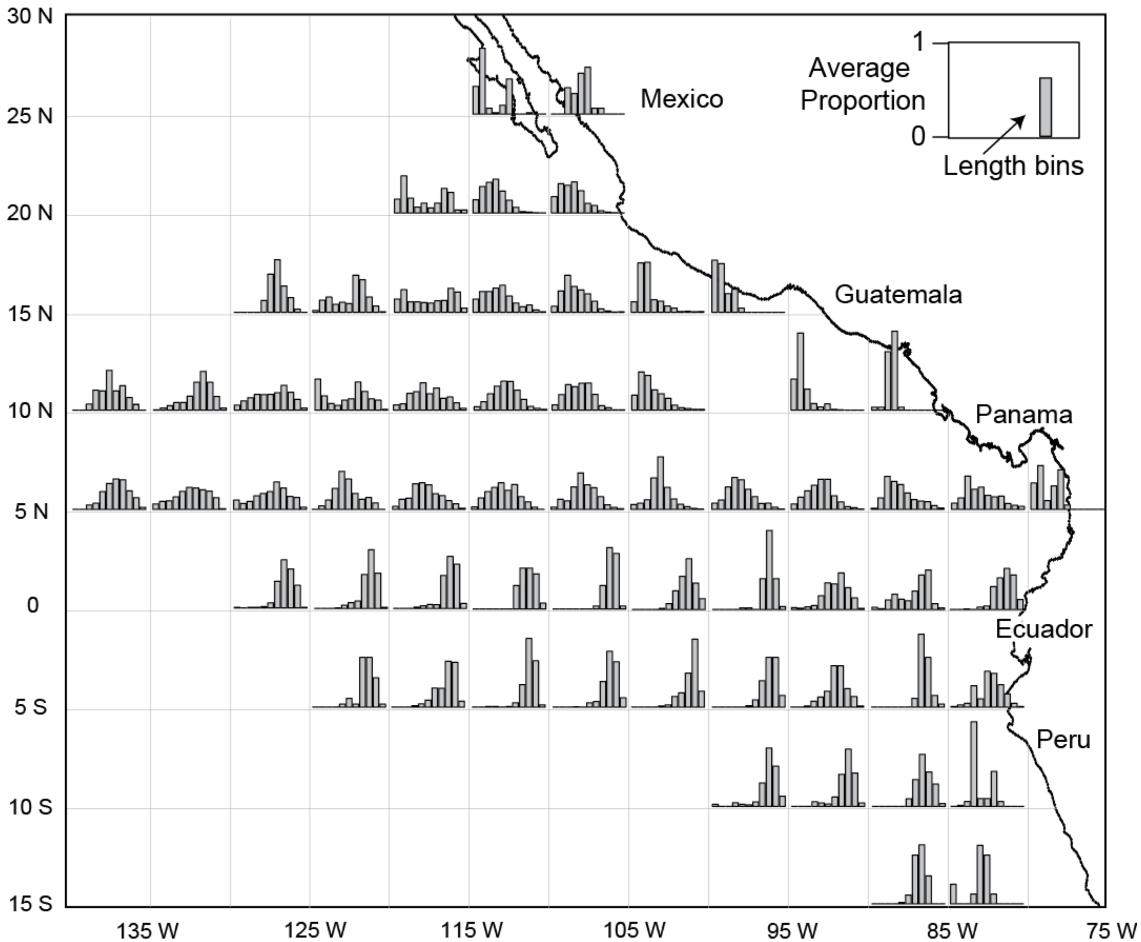
Nominal CPUE trends (observer and logbook data), 1975-2011.

Data processing for length-frequency distributions:

- Raise sample data to the well catch.
- Grown/shrink lengths to mid-month of each quarter (assessment has quarterly time step).
- Summarize each sample by proportion of fish, p_t, in 11 length intervals:
  ≤ 58cm, 59-69cm, …,136-146cm, 147-159cm, ≥ 160cm.

Preliminary analysis of stock assessment areas

Example of summarized length-frequency distributions, 2003-2007 (pooled over quarters and years):



Preliminary analysis of stock assessment areas

Data processing for nominal CPUE trends:

- Catch per day fishing for each month and 5º area was computed by the method used for the assessment.

- Within each 5º area, by quarter, estimate the temporal trend in nominal CPUE using regression splines:

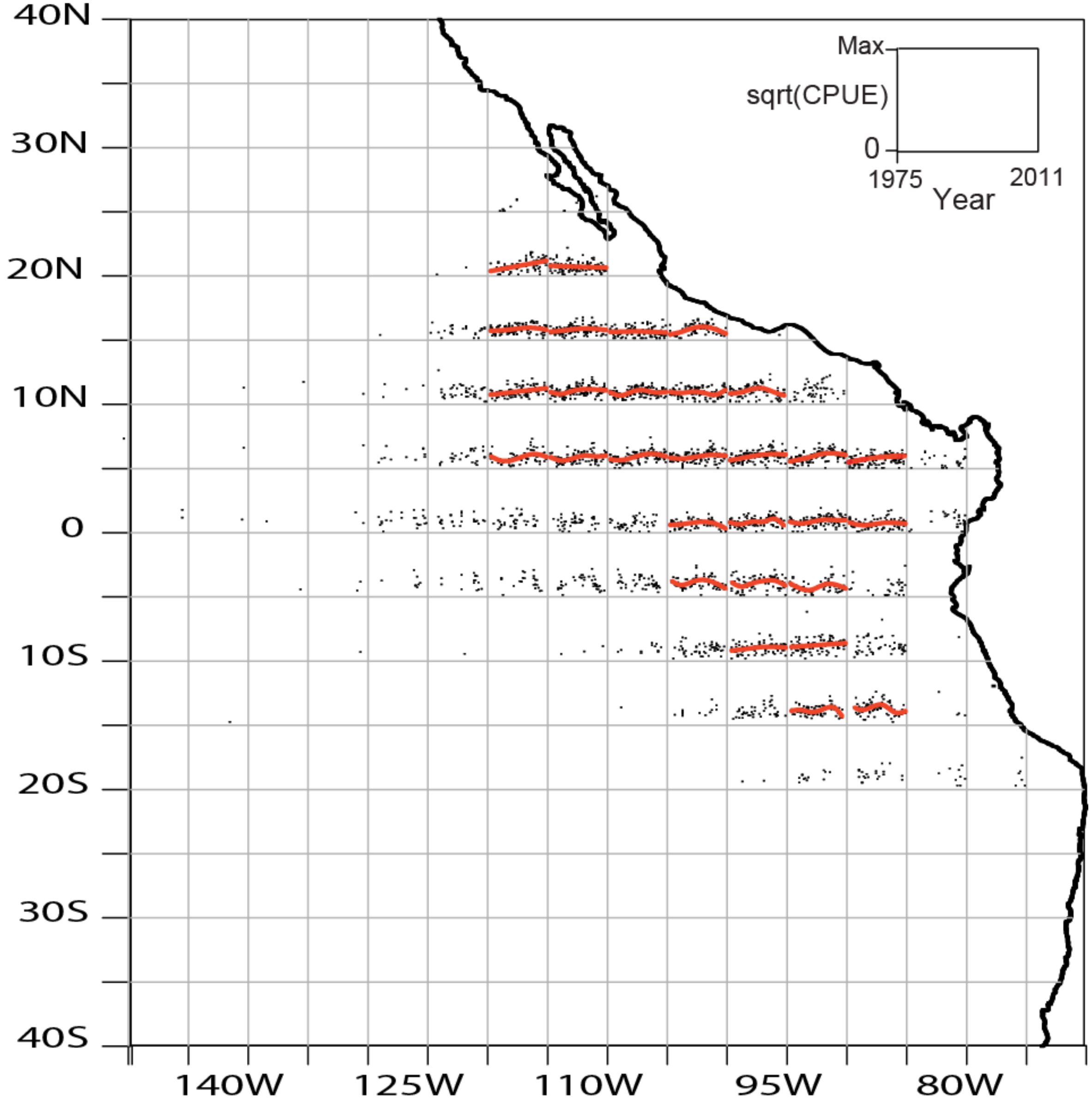
  sqrt(CPUE) = f(year) + error

  where f is a smooth function, “sqrt” = square root.

- The spline coefficients, β_k, (excluding that of the constant term) are used to summarize the trend in each 5º area-quarter.

Preliminary analysis of stock assessment areas

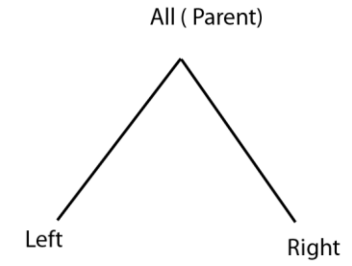
Example of CPUE trends, January-March:



Preliminary analysis of stock assessment areas

The method of analysis is based on multivariate regression trees.

This method allows us to explore pattern *simultaneously* in length-frequency distributions and CPUE trends.



Split criterion:

γ [scaled reduction in heterogeneity, length-frequency] +
(1- γ) [scaled reduction in heterogeneity, trend coefficients]

Measures of heterogeneity for a collection of *r* observations :

Length-frequency

Kullback-Leibler divergence: $\sum_r \sum_t p_r(t) log\left(\frac{p_r(t)}{\bar{p}_.(t)}\right)$

CPUE trend coefficients

Modified squared error: $\sum_r \sum_k \left(\beta_{rk} - \tilde{\beta}_k\right)^2$

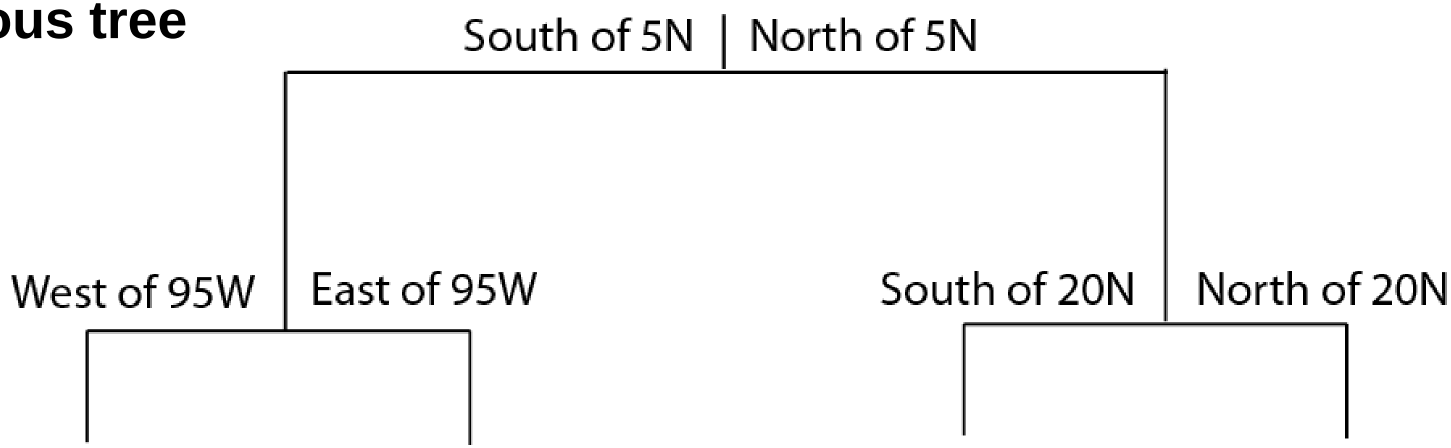
Preliminary analysis of stock assessment areas

Use the 5º latitude, 5º longitude and quarter of fishing as predictor variables.

Build a small tree.

All partitions of this small tree were spatial:

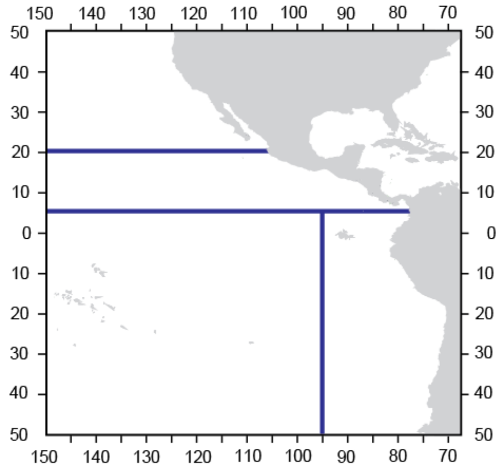
Simultaneous tree



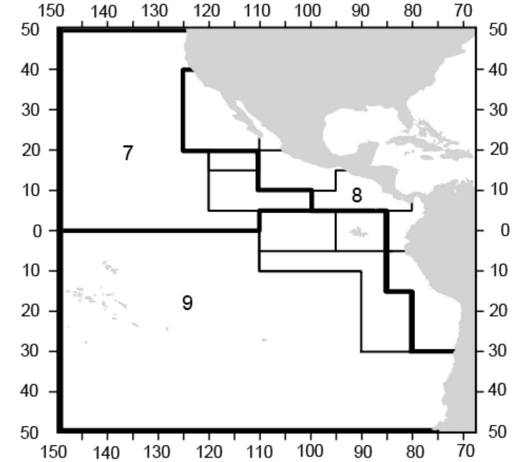
Preliminary analysis of stock assessment areas

Spatial strata produced by the simultaneous tree analysis:

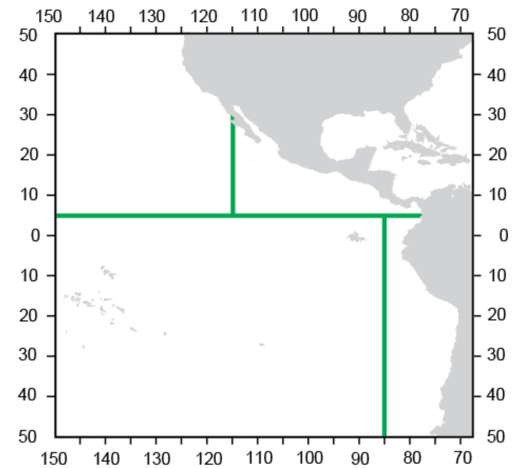
Simultaneous tree
Árbol simultáneo



Stock assessment
Evaluación de stocks



Length-frequency tree
Frecuencias de tallas



Outline of presentation

- Motivation
- Background
  - Port-sampling data
  - Estimation of total catch by species
- Preliminary analysis of stock assessment areas
  - Methods of analysis
  - Results
- Preliminary analysis of poststratification
  - Methods of analysis
  - Results
- Future work

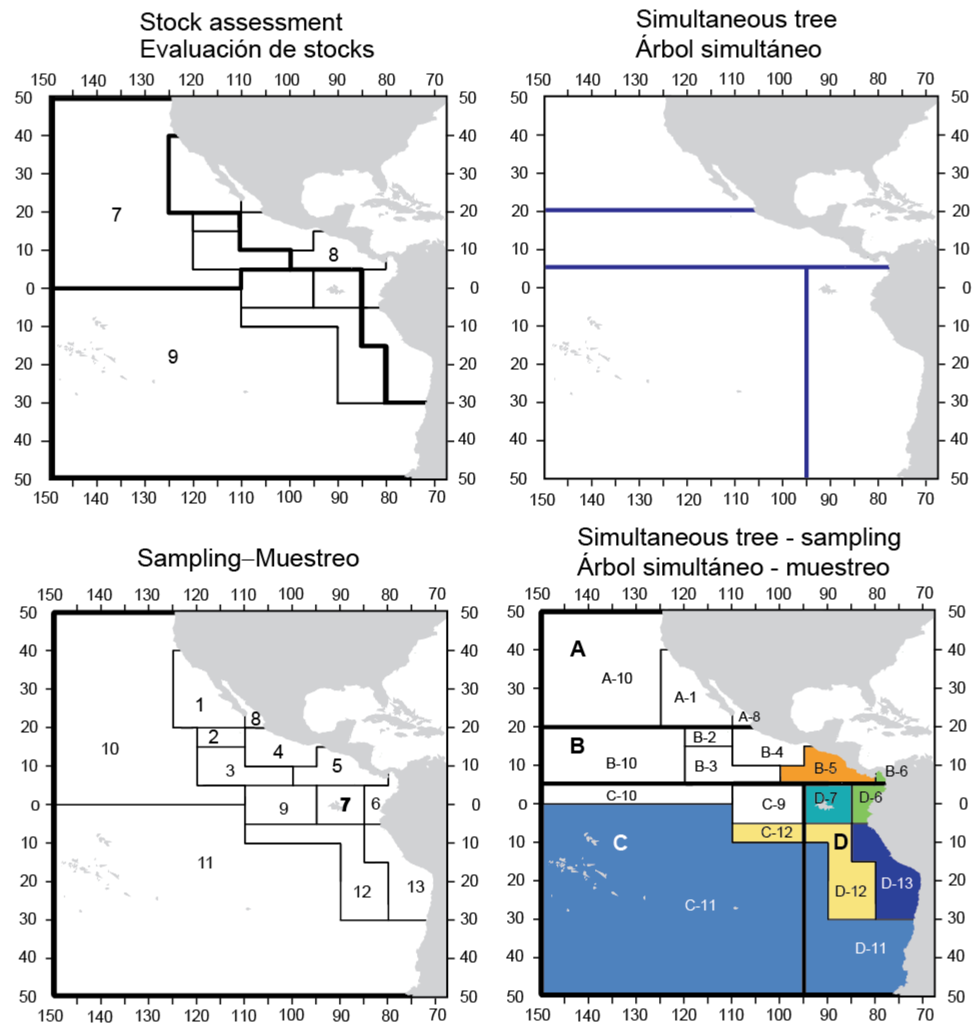
Preliminary analysis of poststratification

Purpose:
Determine level of poststratification of sample data for estimation of species catch.

Four spatial stratifications were considered.

Linear and generalized linear models were used to analyze spatial-temporal variability in:

- average weight;
- species fractions.



Preliminary analysis of poststratification

For average weight, linear models were fitted to the sample data for yellowfin tuna:

$sqrt(\bar{w}_j)$ = overall constant + stock assessment area effect + error
$sqrt(\bar{w}_j)$ = overall constant + tree area effect + error
$sqrt(\bar{w}_j)$ = overall constant + sample area effect + error
$sqrt(\bar{w}_j)$ = overall constant + sample-tree area effect + error
$sqrt(\bar{w}_j)$ = overall constant + month effect + error
$sqrt(\bar{w}_j)$ = overall constant + month effect + stock assessment area effect + error
$sqrt(\bar{w}_j)$ = overall constant + month effect * stock assessment area effect + error
$sqrt(\bar{w}_j)$ = overall constant + quarter effect * stock assessment area effect + error

Models were fitted separately by year (2000-2011), with weights equal to the individual-well total catch amounts.

Within each year, models were compared with: $\Delta AIC = AIC - AIC_{min}$, where AIC_{min} is the AIC of the best model in the collection.

Preliminary analysis of poststratification

For species counts, logistic regression models were fitted to the sample data:

$\log(r_j/[1-r_j]) =$ overall constant + stock assessment area effect
$\log(r_j/[1-r_j]) =$ overall constant + tree area effect
$\log(r_j/[1-r_j]) =$ overall constant + sample area effect
$\log(r_j/[1-r_j]) =$ overall constant + sample-tree area effect

where r_j is the probability that a fish drawn from the j^{th} well was a skipjack.

Models were fitted with weights equal to the individual-well total catch amounts.

Overall results were similar to those for average weight, but there was model instability caused by the predominance of yellowfin in the data.

Preliminary analysis of poststratification

Example of results of analysis of average weight:

	ΔAIC
2006	
Stock assessment (model (i))	12
Tree (model (ii))	1
Sample (model (iii))	1
Sample-tree (model (iv))	0
Month (model (v))	36
Stock assessment + month (model (vi))	7
Stock assessment * month (model (vii))	4
Stock assessment * quarter (model (viii))	7
2007	
Stock assessment (model (i))	18
Tree (model (ii))	75
Sample (model (iii))	0
Sample-tree (model (iv))	0
Month (model (v))	95
Stock assessment + month (model (vi))	22
Stock assessment * month (model (vii))	16
Stock assessment * quarter (model (viii))	19

Preliminary analysis of poststratification

Results of the analysis of yellowfin tuna average weight data can be summarized as follows:

- For most years, spatial structure appears to dominate over temporal structure.
- A spatial stratification somewhere between the 13 sampling areas and the 3 stock assessment areas may prove adequate for estimation of total species catch.

Future work

Preparation for the IATTC External Review in October will include:

1) A sensitivity analysis of the simultaneous regression tree results, including an evaluation of the effect of interannual variability in the length-frequency data.

2) Regression tree and generalized linear model analyses of the port-sampling data (by area, time period) *within* each candidate stock assessment area (from (1)), to provide poststrata for estimating species catch.

3) Comparison of estimates of catch by species, and variance of catch, for several stock assessment configurations and poststratifications (from (1)-(2), plus current).

References

Lennert-Cody, C.E., Maunder, M.N., Aires-da-Silva, A., Minami, M. Submitted. A simultaneous multivariate regression tree approach to defining population spatial units using frequencies of individual characteristics and times series of abundance. Fisheries Research, March 2012.

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, 2012, La Jolla, CA ,USA.

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.

Maunder, M.N., Lennert-Cody, C.E., Aires-da-Silva, A., Bayliff, W.H., Tomlinson, P.K., and Schaefer, K.M. 2010. Summary of data available for bigeye tuna in the eastern Pacific Ocean and its use in stock assessment. Document BET-01-07. External Review of IATTC Bigeye Tuna Assessment, 3-7 May, 2010, La Jolla, California, U.S.A.

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.

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.