IATTC’s approach

Background

The IATTC team builds a single-stock assessment model, using Stock Synthesis, to fit to the simulated dataset for yellowfin tuna in the Indian Ocean. The assessment model ignores all simulated tagging data and uses the “areas-as-fleets” approach to deal with spatially varying fishery selectivity. The assessment model includes sixteen fishery fleets and one survey fleet, which is a term in Stock Synthesis referring to the fleet that has data (e.g., abundance index and compositions) but takes no catch. The survey fleet is added to the model to reflect the abundance and associated length compositions at the population level. We use a spatiotemporal model (VAST) to fit separately to cell-level longline catch-per-unit-effort (CPUE) and length composition data simulated for the whole Indian Ocean, generating area-weighted standardized abundance index and associated length compositions for the survey fleet. In this way, the between-region movements of yellowfin within the Indian Ocean are implicitly accounted for, since the standardized abundance index and length compositions are computed for the whole Indian Ocean instead of a particular region.

Fishery fleets

The model has the same number of fishery fleets as the 4-area assessment model, in which fishery fleets are defined according to gear type (longline, troll, purse-seine), area of operation (Region 1-4), and purse-seine set type (associated and unassociated). However, we redefine three associated purse-seine fishery fleets by applying a regression tree algorithm to simulated cell-level composition data for the associated purse-seine fishery (Figure a). We also explore redefining longline fishery fleets using this approach but decide to reject the four longline fishery fleets redefined by the regression tree algorithm. The main reason is that the simulated composition data for the longline fishery is very sparse (10 lat \* 20 lon interval), leading to a lack of clear grouping of composition data in space.

The composition data for all associated purse-seine are recomputed by raising the observed length frequency by catch. In comparison to the associated purse-seine fisheries, other fisheries have much smaller catch amounts or/and much more limited number of composition observations, so they are not re-defined by the regression tree algorithm and their composition data are not raised by catch. The effective sample sizes of all fishery compositions are weighted by the Francis approach (citation).

To evaluate the impacts of applying the regression tree algorithm to redefine associated purse-seine fishery fleets and using catch-weighted length composition data as model inputs, we build three models in a stepwise manner:

Model 1: Default (region-specific) associated purse-seine fisheries fleets and default (sample-weighted) length compositions for all fishery fleets

Model 2: Regression tree-defined associated purse-seine fisheries fleets and default (sample-weighted) length compositions for all fishery fleets

Model 3: Regression tree-defined associated purse-seine fisheries fleets and catch-weighted length compositions for all associated purse-seine and longline fleets

Survey fleet

The survey fleet includes a longline abundance index and length compositions using the simulated catch rate and length composition data for longline. Both standardizations were conducted in VAST. To make the standardization for length compositions faster, the original 5-cm-resolution composition data are grouped to have a new bin size of 10 cm.

Abundance index

The time series of abundance index for the survey fleet is directly from VAST, which separately models encounter probability and positive catch rate to deal with zero-inflated catch rate observations. We use the *logit* and *log* link functions for the linear predictors of encounter probability and positive catch rate, respectively. Both the encounter probability and positive catch rate in the VAST for index standardization includes an intercept term (year-quarter effect), spatial term, and spatiotemporal term. We use 150 spatial knots to simulate the spatial and spatial random effects in VAST, both of which are assumed to be autocorrelated in space following the Matérn function. Geometric anisotropy in the spatial autocorrelation is turned off in VAST due to convergence issue. The VAST model estimates the abundance index by fitting to cell-level longline CPUE data simulated for yellowfin in the Indian Ocean.

The abundance index for the survey fleet is directly from VAST. The coefficient of variation (CV) of the abundance index is time-varying. It is the sum of the value estimated by VAST and a constant used to scale the mean CV to be 0.075:

Length compositions

Length compositions for the survey are estimated by using a length-specific spatiotemporal model built in VAST. This model is fit to length-specific CPUE data derived from matching two longline datasets: CPUE and length compositions. Specifically, the length-specific CPUE input to the spatiotemporal model is simply the product of CPUE and length frequency in the same stratum (year-quarter latitude longitude). Note that length compositions are sampled more sparsely than CPUE in space - only the strata which have both CPUE and length composition observations are retained as model input.

The VAST model for length compositions uses the same link functions for encounter probability and positive catch rate. An intercept term (year-quarter effect), spatial term, and spatiotemporal term is included for every length bin. All these terms are assumed to be independent and identically distributed among length bins Several simplifications are applied to make this VAST model computationally feasible, especially considering that it needs to be looped for 100 times in the future: 1) only 10 spatial knots are used to simulate the spatial and spatial random effects; 2) length bins are regrouped from the original 5 cm resolution to 10 cm resolution; 3) all hyperparameters are assumed to be shared among length bins. The length compositions for the survey are computed as , where is the standardized abundance index for bin and time from the length-specific VAST model.

The input sample size for the survey length compositions is the product of 5 (the effective sample size for one cell) and the number of cells sampled in the retained length-specific CPUE data. The effective sample sizes of survey compositions are weighted by the Francis approach (citation).

Building SS model

Building the data file for SS needs these three steps

1. Load the original SS data file from YFT\_SRD\_4A\_4.RData into the R environment
2. Modify the data file based on what is described in previous sections
3. Write the new SS data file using function *SS\_writedat* from the R package *r4ss*

Running SS model

Each model is run for three times to get final results

1. Run with LF weight = 1
2. Run with LF weight = Francis
3. Run with LF weight = Francis and suggested recruitment bias adjustment from *r4ss*

Model Results

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Chart

Description automatically generated

Chart, histogram

Description automatically generated

Model diagnostics

Jitter

10 jitter runs are conducted for Model 3 to make sure the total likelihood of the model reaches global maximum.

Runs test

Text

Description automatically generated with low confidence

R0 likelihood profile

Chart, line chart

Description automatically generated

ASPM

Chart

Description automatically generated with low confidence Chart

Description automatically generated

Chart

Description automatically generated

Retrospective pattern

Chart, scatter chart

Description automatically generated