The spatiotemporal dynamics of yellowfin and bigeye tuna in the Eastern Pacific Ocean

Haikun Xu and Peter Kuriyama
Inter-American Tropical Tuna Commission
Scripps Institution of Oceanography

CAPAM Spatiotemporal Workshop
Feb. 28, 2018
Acknowledgements

• Cleridy Lennert-Cody, Mark Maunder, and Carolina Minte-Vera
• Jim Thorson
• Jin Gao
Topics

• Apply VAST to the aggregate catch rate of yellowfin tuna in the EPO between 1975-2016

• Apply VAST to the length composition of yellowfin tuna in the EPO between 2000-2016
1. CPUE: Data

- $1^\circ \times 1^\circ$ catch (in biomass) and effort (in days) data from the vessels with >75% dolphin-associated sets
- Nominal CPUE = catch / effort
1. CPUE: Spatiotemporal Model

VAST separately models encounter probability ($p$) and positive catch rate ($\lambda$) for sample $i$:

1. $p_i = \log^{-1}(\beta_1(t_i) + \omega_1(s_i) + \epsilon_1(s_i, t_i))$
2. $\lambda_i = \exp(\beta_2(t_i) + \omega_2(s_i) + \epsilon_2(s_i, t_i))$

$\beta_1(t_i)$ and $\beta_2(t_i)$: intercept in year $t_i$
$\omega_1(s_i)$ and $\omega_2(s_i)$: spatial variation at location $s_i$
$\epsilon_1(s_i, t_i)$ and $\epsilon_2(s_i, t_i)$: spatiotemporal variation at location $s_i$ in year $t_i$
1. CPUE: Spatiotemporal Model

Autocorrelated spatial and spatiotemporal residuals:
\[ \omega_1 \sim \text{MVN}(0, \sigma_{\omega_1}^2 R_1) \]
\[ \omega_2 \sim \text{MVN}(0, \sigma_{\omega_2}^2 R_2) \]
\[ \varepsilon_1 (s, t) \sim \text{MVN}(0, \sigma_{\varepsilon_1}^2 R_1) \]
\[ \varepsilon_2 (s, t) \sim \text{MVN}(0, \sigma_{\varepsilon_2}^2 R_2) \]

where
\[ R_1(s, s') = \frac{1}{2^{\nu-1} \Gamma(n)} \times (\kappa_1 |H(s - s')|)^{\nu} \times K_\nu(\kappa_1 |H(s - s')|) \]
\[ R_2(s, s') = \frac{1}{2^{\nu-1} \Gamma(n)} \times (\kappa_2 |H(s - s')|)^{\nu} \times K_\nu(\kappa_2 |H(s - s')|) \]

- decorrelation distance
- geometric anisotropy
- Matern smoothness (=1)
1. CPUE: Spatiotemporal Model

The probability of catch data $c$ for sample $i$:

$$\Pr(c_i = C) = \begin{cases} 1 - p_i & \text{if } C = 0 \\ p_i \times \text{lognormal}(c_i | \lambda_i, \sigma_m^2) & \text{if } C > 0 \end{cases}$$

Joint Log-Likelihood = $\sum_i \log(\Pr(c_i = C))$

TMB + R
The spatiotemporal dynamics of yellowfin tuna in quarter 1

Large inter-annual variation in fishery locations

use knots to approximate spatial and spatiotemporal residuals
Predicted log catch rate for quarter 1

$$\hat{d}(k, t) = \hat{p}(k, t) \times \hat{\lambda}(k, t)$$

K: index of spatial knot
1. CPUE: Spatiotemporal Model

Total abundance for the entire domain

\[ I(t) = \sum_{k=1}^{n_k} \left( a(k) \times \hat{d}(k, t) \right) \quad k: \text{spatial knot} \]
Predicted log catch rate for quarter 2
Predicted log catch rate for quarter 3
Predicted log catch rate for quarter 4
Comparison the 4 quarterly indices of abundance

problematic for the quarter as year approach

change in catchability?
For fishery-dependent data:
How to quantify the change in catchability?
Long-term mean log catch rate

The locations of fisheries activities change from quarter to quarter, but the quarterly indices of abundance should be calculated based on the same domain in the EPO.

Discussion: How to predict the catch rate in the regions without data?
Long-term mean log catch rate

Discussion: How to predict the catch rate in the regions without data?
Solution 1: catch rate = 0
Problem: Predicted catch rate is high in some adjacent regions with data
Discussion: How to predict the catch rate in the regions without data?
Solution 1: catch rate = 0
Problem: Predicted catch rate is high in adjacent regions with data
Solution 2: interpolate catch rate based on spatial autocorrelation in residuals
Problem: spatial correlation in residuals may vary over space
The fishery follows the movement of warm waters.
Discussion: How to predict the catch rate in the regions without data?
Solution 1: catch rate = 0
Problem: Predicted catch rate is high in adjacent regions with data
Solution 2: interpolate catch rate based on spatial autocorrelation in residuals
Problem: spatial correlation in residuals may vary over space

Any suggestions?
2. Length Comps: Data

- $5^\circ \times 5^\circ$ number at length (1cm bin) and effort (in days) data from the vessels with >75% dolphin-associated sets
- Case study: Northern hemisphere; Quarter 2
2. Length Comp: Spatiotemporal Model

\[ p_i = \text{logit}^{-1}\left(\beta_1(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_1}(l_i, f) \omega_1(s_i, f) + \sum_{f=1}^{n_f} L_{\epsilon_1}(l_i, f) \epsilon_1(s_i, f, t_i)\right) \]

\[ \lambda_i = \exp\left(\beta_2(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_2}(l_i, f) \omega_2(s_i, f) + \sum_{f=1}^{n_f} L_{\epsilon_2}(l_i, f) \epsilon_2(s_i, f, t_i)\right) \]

\( l_i \): length bin (1-9);

\( L_{\omega_1}, L_{\epsilon_1}, L_{\omega_2}, L_{\epsilon_2} \): generate the loading matrixes \( (L^T L) \) for the spatial and spatiotemporal covariance among \( l_i \)

\( n_f \): 2
Spatial residuals

Mean predicted log catch rate between 2000-2016

Size segregation in spatial distribution:
- Small YTF > coastal region off Mexico
- Large YFT > pelagic and equatorial regions
Correlation matrix for spatiotemporal residuals

Size segregation in spatiotemporal distribution too:
- positive correlations between close length bins
- negative correlations between distant length bins

AR1 in length does not work for YFT

Hypothesis
Differing habitat preferences of the smallest and largest YFT are driven by env conditions
Problems in this length comp modelling

- Spatial resolution is low ($5^\circ \times 5^\circ$) -> 100% observed encounter rate occurs in some grid cells

My solution: $\text{RhoConfig} = \{\text{"Beta1"}=2, \text{"Beta2"}=0\}$ -> random walk
Increase the number of length bins from 9 to 20 leads to improved model fit.

Discussion:
• Bad fit or bad diagnose?
• How many number of bins?
• How many number of factors?
Bigeye tuna (*Thunnus obesus*)

- Assessed as one stock
- Decline affected by La Niña events
- Strong recruitment in 2012
Catches
Recruitment
Data (fishery-dependent)

• CPUE
  • #fish/#hooks
  • 5° latitude × 5° longitude

• Length compositions
  • Counts (100% encounters)
  • 5° latitude × 10° longitude

• Japanese catches only
  • Longline
  • Annual
CPUE

#fish/#hooks

![Map showing CPUE (catch per unit effort) with color scale indicating fish/hook density.](image)
VAST model - CPUE

Extrapolation (Lat-Lon)

Knots (North-East)
Index of abundance
Compared to spawning biomass ratio
Diagnostics
Anisotropy
VAST model – length comps

• Counts (100% encounters)
• 5° latitude × 10° longitude
• Length quartiles
Data-length compositions
30 knots
Problems
Open questions

• Handling 0% and 100% encounter rate

• Spatial modeling over ~500km by ~500km grids?
  • Bigger in some cases

• Discussion:
  • Bad fit or bad diagnostics?
  • How many number of bins?
  • How many number of factors?
• Backup slides
Problems in this length comp modelling

• How to eliminate the possibility that the size segregation is not caused by including only 2 loading factors for 9 length bins?

\[ p_i = \text{logit}^{-1}\left( \beta_1(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_1}(l_i, f) \omega_1(s_i, f) + \sum_{f=1}^{n_f} L_{\varepsilon_1}(l_i, f) \varepsilon_1(s_i, f, t_i) \right) \]

\[ \lambda_i = \exp\left( \beta_2(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_2}(l_i, f) \omega_2(s_i, f) + \sum_{f=1}^{n_f} L_{\varepsilon_2}(l_i, f) \varepsilon_2(s_i, f, t_i) \right) \]
The fishery follows the movement of warm waters.
Plan for the next step

1. Find the key env driver(s) that causes the size-segregation in spatial and spatiotemporal distribution
2. Calculate the mean predicted catch rate at length as a function of the driver
3. Find the SST range in which
   \[ \log(d_{large}) - \log(d_{small}) = \log\left(\frac{d_{large}}{d_{small}}\right) \]
   is largest
4. Improve spatial management: predicting preferred fishing grounds using real-time env observations