Modeling silky shark bycatch

Mihoko Minami (Institute of Statistical Mathematics, Tokyo)
Cleridy Lennert-Cody (IATTC)
Why model silky shark bycatch?

There is concern about possible negative effects of fisheries catch / bycatch on shark populations.

Silky shark bycatch per set may index abundance, all other things being equal....

Modeling shark bycatch is a first step towards IATTC’s mandate to provide preliminary advice on key shark species involved in the purse-seine fishery (IATTC Resolution C-05-03).

IATTC has more data on silky sharks than on other shark species.
Overall approach for modeling bycatch per set: use generalized linear and additive models to estimate trends.

This has proved somewhat challenging because of the characteristics of bycatch per set.

Proceeding in two steps:

1) Explore different probability functions for the “random component” of generalized linear/additive models.
   
   Started with floating object set data; will expand analysis to dolphin and unassociated sets.

2) Explore in detail spatial/environmental effects.
Average silky shark bycatch per set by size category (unstandardized)

with tons

without tons

Number of sharks per set

Year


Number of floating object sets
1994-2004
Silky shark bycatch per set
1994-2004
Bycatch per set – floating object sets

1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
Probability functions used for modeling bycatch per set:

- Poisson
- negative binomial
- zero-inflated Poisson
- zero-inflated negative binomial

The negative binomial is an extension of the Poisson distribution that can better model highly variable count data.

Similarly, the zero-inflated negative binomial can be considered an extension of the zero-inflated Poisson.

Zero-inflated probability functions are used to model data with both a large proportion of zero-valued observations and also large positive values.
Zero-inflated models

- **Perfect state (zero state):** bycatch never occurs
- **Imperfect state:** bycatch could occur

Frequency of bycatch per set:
- Frequency range: 0 to 30
- Frequency bins: 0, 1, 5, 10, 15, 20, 25, ≥30

Bycatch per set distribution:
Poisson and negative binomial regression models

Log-linear regression models were used to relate the mean bycatch per set \((\mu)\) to covariates:

\[
\log(\mu_i) = B_{i0} + B_{i1}\beta_1 + \cdots + B_{ik}\beta_k = \mathbf{B}_i \mathbf{\beta}
\]

where

- \(B_{i0}, B_{i1}, \ldots, B_{ik}\) values of covariates
- \(\beta_1, \ldots, \beta_k\) coefficients (parameters)
Zero-inflated regression models

Two stage regression model:

- Which state does a set take (\( p = \) probability of set being in “perfect” state)?

  **logistic regression model**

  \[
  \log \frac{p_i}{1 - p_i} = G_{i0} + G_{i1} \gamma_1 + \cdots + G_{ik_\gamma} \gamma_{k_\gamma} = G_i \gamma
  \]

- Amount of bycatch when set is in imperfect state (\( \mu = \) mean bycatch in imperfect state)?

  **negative binomial / Poisson regression model**

  \[
  \log(\mu_i) = B_{i0} + B_{i1} \beta_1 + \cdots + B_{ik_\beta} \beta_{k_\beta} = B_i \beta
  \]

where \( B_{i0}, B_{i1}, \cdots, B_{ik_\beta}, G_{i0}, G_{i1}, \cdots, G_{ik_\gamma} \) values of covariates
\( \beta_1, \cdots, \beta_{k_\beta}, \gamma_1, \cdots, \gamma_{k_\gamma} \) coefficients (parameters)
Data

- Floating object sets, 1994 – 2004 (32,148 sets)

- Data excluded:
  - sets with bycatch reported in tons
  - sets with no catch of target tunas
  - repeat sets on same floating object
  - sets missing data on predictors
  - data for 1993

- Predictor variables
  - location (latitude, longitude), year, calendar date, time
  - net depth, floating object depth
  - sea surface temperature
  - amount of tuna catch (log (tuna))
  - amount of non-silky shark bycatch (log(non-silky+1))
  - two proxies for floating object density
## Model comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood (training data)</th>
<th>Generalized Information Criterion (test data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson</td>
<td>-81849</td>
<td>&gt; 100000</td>
</tr>
<tr>
<td>Negative binomial</td>
<td>-32572</td>
<td>65280</td>
</tr>
<tr>
<td>Zero-inflated Poisson</td>
<td>-56389</td>
<td>&gt; 100000</td>
</tr>
<tr>
<td>Zero-inflated negative binomial</td>
<td>-32346</td>
<td>64827</td>
</tr>
<tr>
<td>(without smoothing)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero-inflated negative binomial</td>
<td>-31862</td>
<td>63921</td>
</tr>
</tbody>
</table>
Partial dependence plot

Number of animals per set

Year


Poisson with t.p.r.s.
NB with t.p.r.s. (theta=0.182)
NB with t.p.r.s. (theta=0.33)
ZIP with t.p.r.s.
ZINB
ZINB with t.p.r.s. (theta=0.555)
Interpretation of trends in bycatch per set is/will be complicated by...

Species identification concerns (pre-2005):
- Misidentification of silky sharks
- “Unknown” category: what proportion were silky sharks?

Floating object set data: true object density unknown.

Effect of 2000 IATTC bycatch resolution on live release unknown (pre-2005).

Existence of extremely large bycatches that are difficult to model.