Sloping Towards the Holy Grail

Combining Conventional Tagging and Tracking Data

John Sibert

sibert@hawaii.edu
Shhh!

- Conventional tagging data only rarely used directly in stock assessment.
  - MULTIFAN-CL and SS3 best examples.
  - May be used to estimate mortality for incorporation in assessment in non-integrated analyses.

- Tracking data are never directly used in stock assessment.
  - May be used indirectly to “standardize” longline fishing effort.
  - May be used qualitatively to identify habitats of particular interest.

- Need more data.

- Need appropriate models.
Spatial Models for Tagging and Tracking

- Advection-diffusion-reaction models for tuna populations since 1990 — “Eulerian”.
- State space Kalman filter models for individual fish since 2000 — “Lagrangian”.
- Really the same model with the same parameters.
- Maximum likelihood estimation of model parameters.

(Animation)

Is it feasible to integrate the Lagrangian models of individual fish movement into Eulerian model of fish populations?
SEAPODYM — General Overview

- 3-D Models
- Ocean Biogeochem.
- Primary Production from satellites
- Prey model
- Predator's population dynamics model
- Ocean Physics
- 3-D models

Blue = skipjack; Red = bigeye (kindly from A. Fonteneau)
SEAPODYM — Optimal Parameterization

Environmental forcing
- Epipelagic data: SST, PP, Surface forage
- 3-layer data: temperature, oxygen, currents forage biomass

Anthropogenic forcing
- Data by fishing fleets: Pole-and-line, Purse-seine, Long-line

ADR model
- Initial conditions
- 0-3 month juveniles model
  - $t = 0,\ldots,T$
- K-cohorts adults model

Output
- Computing predictions

Reseting model parameters
- No
- Quasi-Newton convergence?
  - Yes
- $t = T,\ldots,0$

Reverse model: Computing derivatives of cost function

Cost function

"Optimal" parameterization
State Space Model

Transition Equation

\[ \alpha_{ki} = \alpha_{ki-1} + c_i + \eta_i; \quad i = 1, \ldots, T_k \]

\[ c_i = \begin{pmatrix} u_{xy} \\ v_{xy} \end{pmatrix} \quad \eta_i = N(0, Q_i) \]

\[ Q_i = \begin{pmatrix} 2D_{xy} & 0 \\ 0 & 2D_{xy} \end{pmatrix} \]

Measurement Equation

\[ y_{ki} = z(\alpha_{ki}) + d_{ki} + \varepsilon_{ki}; \quad i = 1, \ldots, T_k \]

\[ d_{ki} = \begin{pmatrix} b_{xk} \\ b_{yk} \\ b_{\tau k} \end{pmatrix} \quad \varepsilon_{ki} = N(0, H_{ki}) \]

\[ H_i = \begin{pmatrix} \sigma_{xk}^2 & 0 & 0 \\ 0 & \sigma_{y_{ki}}^2 & 0 \\ 0 & 0 & \sigma_{\tau_{ki}}^2 \end{pmatrix} \]
Aggregate Model: SEAPODYM + State Space

Negative aggregate log likelihood

\[ \ell_A = -\left( w_\ell \cdot \log L_T + (1 - w_\ell) \cdot \frac{1}{n} \sum_k \log L_k \right) \]

Parameter Estimates

\( \hat{\Phi}_I, \hat{\Phi}_S = \arg\min_{\Phi_I, \Phi_S} \ell_A \)

SEAPODYM-TAGEST

\[ L_T(F, I, C_{obs}|\Phi_I, \Phi_F, \Phi_M) = \prod \left[ \frac{C_{\text{pred}}^{C_{\text{obs}}^e} e^{-C_{\text{obs}}}}{C_{\text{obs}}!} \right] \]

Unscented Kalman filter likelihood

\[ L_k(Y_k|\Phi_{Sk}) \quad k = 1, 2, \ldots, n; \quad \Phi_{Sk} = (b_{xk}, b_{yk}, b_{\tau k}, \sigma_{xk}, \sigma_{yk}, \sigma_{\tau k}, a_{0k}, b_{0k}) \]
First problem: scale incompatibilities

- SEAPODYM:
  - 0.5° grid
  - Monthly time steps
  - Grid-scale variability in $u, v, D$

- uKfsST:
  - Continuous in space
  - Daily time steps (sort of).
Second problem: “gappy” data.
Third problem: boundaries.
Fourth problem: automatic differentiation

• Need to connect state space model transition equation to the parameters of the habitat model through the chain rule.

• Matrix subscript operators are not differentiable with respect to the indices; i. e., $\frac{\partial z_{ij}}{\partial i}$ and $\frac{\partial z_{ij}}{\partial j}$ undefined.

• Bilinear interpolation might work;

\[ z_{xy} \approx f(z_{i,j}, z_{i+1,j}, z_{i,j+1}, z_{i+1,j+1}) \]
Solutions

1. Scale incompatibilities:
   bilinear interpolation between computational elements;
   daily time steps in state-space model between SEAPODYM iterations.

2. “Gappy” data:
   use reconstructed tracks from Royer’s ensemble Kalman filter.

3. Boundaries:
   reflect tags away from boundaries;
   penalize likelihood for violation of boundaries.

4. Differentiability:
   the interpolation and reflection operations may be inherently non-differentiable.
   But ...
Differentiability?
Preliminary results — 5 tracks
## Preliminary results — 5 tracks

### Estimated Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Value</th>
<th>Composite “Optimum”</th>
<th>Previous Optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma_a)</td>
<td>4.00</td>
<td>3.05</td>
<td>2.26</td>
</tr>
<tr>
<td>(T)</td>
<td>15.0</td>
<td>16.7</td>
<td>16.0</td>
</tr>
<tr>
<td>(\alpha_{oxy})</td>
<td>0.001</td>
<td>0.087</td>
<td>0.100</td>
</tr>
<tr>
<td>(\hat{\theta})</td>
<td>1.00</td>
<td>0.263</td>
<td>1.26</td>
</tr>
<tr>
<td>MSS</td>
<td>0.5</td>
<td>0.118</td>
<td>0.600</td>
</tr>
</tbody>
</table>
Preliminary Conclusions

Feasibility of combined model

- Model is sensitive to data and appears to shift parameter estimates in sensible directions.
- Model is computationally feasible. Minimization requires about 20 minutes 2.2GHz laptop.

Next Steps

- Improve reflective boundaries.
- Devise differentiable way of spatial interpolation.
- Relax constraint on identical errors for all tags.
Thanks

- Large Pelagics Research Center, University of New Hampshire — funding.
- Molly Lutcavage — data.
- CLS MEMMS Team — Inna Senina, Patrick Lehodey, François Royer — the difficult parts.
- Anders Nielsen, Chi Lam — for implementing and sharing uKfsST.