Status of State-Space Model for Coho Harvest and Movement

K. Newman, June 19, 1997

1 Data

- Three recovery years: 1985, 1986, 1987
- Stocks: Humptulips (Grays Harbor) coho, brood years 1982-84, CWT codes 632861, (632826-632827), (633138,633139,633163, 633201) with initial release numbers of 45,404, 50.876, and 63,360, respectively.
- Space: 11 recovery regions- southern Oregon (Brookings) to northern British Columbia (CDFO regions 9-12)

Region	South coordinate
Brookings	0.000000
Coos Bay	1.129760
Newport	2.478182
Tillamook	3.503198
Astoria	4.484178
Grays Harbor	5.268254
Quillayute	6.548769
C. Flattery	7.544993
SW Vanc Is	7.899463
NW Vanc Is	9.564474
N. BC (S)	11.607614
'top'	12.720993

Units are approximately 100 miles.

- Time: 16 weeks- 3rd week of June or 1st week of July to mid October
- Effort: commercial troll only (US- troll boat landings; Canada- troll boat days)
- CWT recoveries: all in the 9 regions and 16 weeks (with terminal area including those beyond 16 weeks)

2 Model Components and Parameters

1. **Initial**

   (a) The R fish released experience mortality of level γ_i upto beginning of catch period.

   (b) The $R\gamma_S$ surviving fish are distributed across the 11 recovery regions according to a Beta(γ_α, γ_β) distribution.

2. **Mortality**: within a given area fish are harvested at rate proportional to effort scaled by area; say for area a:

$$H_t[a,a](U.S.) = \frac{F_{a,t}(U.S.)}{M + F_{a,t}(U.S.)}\left(1 - \exp\left[-M - F_{a,t}(U.S.)\right]\right) \quad (1)$$

$$H_t[b,b](Canada) = \frac{F_{b,t}(Canada)}{M + F_{b,t}(Canada)}\left(1 - \exp\left[-M - F_{b,t}(Canada)\right]\right) \quad (2)$$

   where

$$M = 0.0001\gamma_n$$

$$F_{a,t}(U.S.) = \gamma_{f(U.S.)}\frac{\text{Effort}_{a,t}}{25,000(a_R - a_L)}$$

$$F_{b,t}(Canada) = \gamma_{f(Canada)}\frac{\text{Effort}_{b,t}}{50,000(b_R - b_L)}$$

3. **Movement**: a random walk with unequal direction and step size probabilities depending upon location on line at time t, p_t and time t itself.

   A fish within cell a cannot move beyond its natal area in its next step, nor can it move outside the 11 areas.

   The probability of moving to a cell b depends upon direction of and distance to cell b.

   (a) Direction: if cell b is in direction of natal stream, probability of moving toward stream is

$$\Pr(towards) = 1 - \exp[-0.1\gamma_d(|p_{t-1}| + t)],\ \gamma_d > 0 \quad (3)$$

   while if b is in the opposite direction use 1 minus the above probability.

   (b) Step size: follows a Beta(α, β) distribution with parameters depending upon direction of movement and distance from natal area. If at time $t - 1$ the fish is located at distance p_{t-1}, the Beta parameters for a step toward the natal area:

$$\alpha = 0.01\gamma_t(|p_{t-1}| + t) \quad (4)$$

$$\beta = 2.00, \quad (5)$$

   while for a step *away* from the natal area:

$$\alpha = \gamma_a/(1 + |p_{t-1}| + t) \quad (6)$$

$$\beta = 2.00. \quad (7)$$

3 State-Space Model and Estimation

With γ_n set $= 0$,

$$N_t = M_{t-1}(I - H_{t-1})N_{t-1} + w_t \quad (8)$$

$$C_t = H_t N_t + v_t \quad (9)$$

where

- N_t is the vector (over the 11 areas) of abundance at the beginning of interval $[t, t+1)$;
- C_t is the vector of CWT recoveries;
- H_t is a diagonal matrix of harvest rates during interval $[t, t+1)$ and is based on equations (1) and (2);
- M_{t-1} is a movement matrix giving probabilities of moving from cell i to cell j *after* harvest and based on equations (4-7);
- The error vectors w_t and v_t are multivariate Normal with mean zero and covariance matrix based on a convolution of multinomial rvs.

The Kalman algorithms (filtering and smoothing) do 2 things:

- provide a means of estimating the unobserved abundance N_t
- provide a means of calculating the likelihood equation, $\mathrm{L}(\gamma) \equiv \Pr(C_1, \ldots, C_T)$,

Maximum likelihood is used to estimate the parameters.

There are 9 parameters in the model for each of the 3 components:

- Initial: γ_i for initial survival and γ_α and γ_β for the Beta distribution at the beginning;
- Mortality: γ_n for natural mortality and $\gamma_f(U.S.)$ and $\gamma_f(Canada)$ for the fishing mortality
- Movement: γ_d for step direction and γ_t and γ_a for step sizes toward and away from natal area.

Based on simulation results it is difficult to estimate all 9 parameters precisely[1]. Here 3 of the parameters are fixed:

[1] And some are nearly non-identifiable- if the value of one parameter is increased, the value of another parameter can be decreased a certain amount with no loss of fit.

- $\gamma_b = 2.00$
- $\gamma_n = 0.00$
- $\gamma_a = 3.00$

4 Results

4.1 Parameter estimates

The estimates (with coefficient of variation in parentheses):

Module	Parameter	1985	1986	1987
Initial	γ_i	1.52 (9%)	4.11 (2%)	1.42 (5%)
	γ_α	5.38 (3%)	3.27 (1%)	3.70 (2%)
Mortality	$\gamma_f(U.S.)$	1.77 (9%)	4.94 (4%)	2.71 (6%)
	$\gamma_f(Canada)$	5.11 (6%)	2.93 (5%)	2.29 (10%)
Movement	γ_d	0.22 (5%)	0.24 (4%)	0.58 (16%)
	γ_t	56.58 (12%)	41.73 (6%)	20.38 (20%)

4.2 Interpretation

1. Initial

   - Initial survival estimates can be compared to crude 'cohort' analysis, namely take total recoveries and divide by release size. The model estimates of 1.52%, 4.11%, and 1.42% correspond to the simple estimates of 1.25%, 3.01%, and 1.32%, respectively.
   - Initial spatial distribution- 1985 further north than 1986 and 1987

2. Mortality: look at fitted harvest rates in U.S. and Canadian fisheries as a function of effort level (scaled into common units). Maximum rates about 45% for some Canadian fisheries (SW Vanc Island).

3. Movement

   - Expected step size towards natal area varies with distance from natal area and time in season and is defined by:

     $$E[\text{Step Size Homeward}] = \Pr(Towards)\frac{\gamma_t}{\gamma_t + 2.00} - \Pr(Away)\frac{3}{3+2}$$

     1985 expected step sizes tend to be smaller than those for 1986 and 1987.

   - Movement rates per week can be estimated by multiplying expected step size homeward by 100 $\Rightarrow$ rates of no more than 220 miles per week for 1985 up to 415 miles per week for 1987- maximums in the final time period in the regions furthest from the natal area.

4.3 Estimates of Abundance

The abundance by area-time cell can be estimated using the Kalman filter- the estimate is the expected number at time t based on all CWT recoveries upto time t:

$$\text{Kalman Filter } \hat{N}_t^t = \mathrm{E}[N_t | C_1, \ldots, C_t, \gamma]$$

where γ is the entire *vector* of parameter estimates.

An improved (post-season) estimate is the Kalman smooth, which uses all the CWT recoveries upto and beyond time t; say for T time periods (16 in this case):

$$\text{Kalman Smooth } \hat{N}_t^T = \mathrm{E}[N_t | C_1, \ldots, C_t, \ldots, C_T, \gamma]$$

4.4 Goodness of Fit

Can look at the observed (expanded) CWT recoveries and compare to model fits:

$$\hat{C}_t = \hat{H}_t \hat{N}_t^t,$$

and examine the residuals.

When the effort level is 0 the observed and expected catch should be 0- these cases are ignored in assessing goodness of fit.

5 Other Work

- Sensitivity/Robustness of Model Formulation to:
  - Initial seeds for parameter estimation (done w/ 6 params)
  - Estimating 9 instead of 6 parameters
  - Fixed value of the 3 parameters on estimates for remaining 6 (effect on parameter interpretations; e.g., movement rates)
  - Modeling of final time period.
- Simulation for management planning: the selection from the parameter space; stochastic versus deterministic components.
- Alternative spatial framework- inside fisheries
- Overlapping fisheries and mortality.