



Assessing different models for determining the effect of birth date on fish

Molly Dieterich, Dr. David Lonzarich, Dr. Jessica Kraker
University of Wisconsin – Eau Claire
Blugold Fellowship Research

Background/Motivation: Many models have been used to examine how natural selection acts on phenotypic traits. In this study we assessed different mathematical approaches for measuring selection on birth date in juvenile coho salmon (*Oncorhynchus kisutch*) from otolith data. Our goal was to find a model or group of models capable of modeling selective mortality in a mathematically tractable, robust, quantitative, and biologically meaningful way. A vigorous literature review yielded four distinctive modeling methods for estimating selective mortality; two models were compatible with our cross-sectional age-frequency data and both of them used maximum likelihood estimation to fit functions to the data. We used simulated date-of-birth frequency data to evaluate each model output. All of the models we examined discuss size-selective mortality, while we would like to evaluate birth-date-selective mortality using the same models.



Left[†]: Otoliths are small white structures found in the heads of fish that provide a sense of balance and aid in hearing. Growth rings on otoliths record growth beginning at hatching and can be used to determine age[†].
Right[†]: Coho salmon are also known as silver salmon. They live in marine environments but spawn in freshwater once before death around age three[†].

What are the data? The simulated data of birth date we used to assess the sensitivity and response of the models to different distributions were created in Microsoft Excel using a random number generator based on arbitrary theoretical distributions. Both before and after (the period of selection) data were created in this way. Frequency distribution graphs of example before and after data are shown below.



Situation 1: Strong selective mortality on fish born early in the season

Situation 2: Strong selective mortality on fish born in the middle of the season

Situation 3: No selective mortality on fish born early in the season

Models: After the examination of several modeling approaches, the two model groups we identified as fitting our criteria were the Schluter/Anderson^{1,2} fitness function and the Good, Dodson, Meehan, and Ryan³ fitness function.

Schluter/Anderson¹: In this mathematical approach for estimating size-selective mortality data, a spline regression model is used to fit $f(z)$, the relative fitness function. Relative fitness is a function of the sample sizes of fish caught before and after the period of selection (independent samples) and a measure of how much more likely a fish of class z will survive to the "after" population relative to other classes.

$f(z) = (S_0/S_1) * [h(z)/(1-h(z))]$ where:

- S_0 is the size of the sample from the "before" population and S_1 is the size of the sample from the "after" population and
- $h(z) = S_0(z)/[S_0(z)+S_1(z)]$ is the conditional probability that a fish of size z was caught alive in the "after" population
- $S_0(z)$ and $S_1(z)$ are raw frequency distributions for each size class z .

Although Anderson's article provided the math behind the fitness function, Schluter developed a model to use in actually fitting the fitness function.

The biological limitations of this model include:

- that it cannot be used if $h(z) = 0$, which occurs when no fish of a certain size class are present in $S_0(z)$ but fish of that size class are present in $S_1(z)$
- it assumes S_0 individuals all die (entire "before" sample)

This model is useful because:

- it uses spline regression
- for generation of curve and to obtain confidence intervals via bootstrapping; requires minimal assumptions
- sample size does not matter

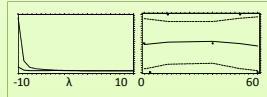
This modeling approach responds to the data as shown below, first in a simple example and followed by the outputs generated using the three situations of simulated data introduced above.

A simple example to follow Schluter/Anderson:

$z:$ 1 3 5 7
Time 0: 59 36 4 1 $S_1 = 100$
Time 1: 1 13 36 50 $S_2 = 100$

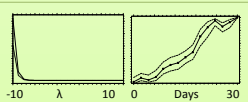
$$h(z) = \frac{S_0(z)}{S_0(z) + S_1(z)}$$

$$f(z) = \frac{S_0(z) * h(z)}{S_1 * (1 - h(z))}$$

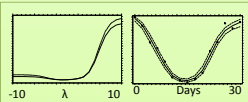


Schluter's fitness function[†] gives a range of possible lambda values (left) and a fitness function based on a specific lambda (right, $\lambda = 4$).

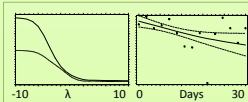
$h(1) = 1/(1+59) = 0.0167$ $f(1) = (100/100) * (0.0167/0.9833) = 0.0169$
 $h(3) = 13/(13+36) = 0.2653$ $f(3) = (100/100) * (0.2653/0.7347) = 0.3611$
 $h(5) = 36/(36+4) = 0.9$ $f(5) = (100/100) * (0.9/0.1) = 9$
 $h(7) = 50/(50+1) = 0.9804$ $f(7) = (100/100) * (0.9804/0.0196) = 50$



Schluter's fitness function[†]: Situation 1
Left: range of lambda values
Right: fitness function based on lambda = -6, confidence intervals based on 100 bootstraps



Schluter's fitness function[†]: Situation 2
Left: range of lambda values
Right: fitness function based on lambda = 0, confidence intervals based on 100 bootstraps



Schluter's fitness function[†]: Situation 3
Left: range of lambda values
Right: fitness function based on lambda = 6, confidence intervals based on 100 bootstraps

Good, et al.³: In this model for estimating size-selective mortality with regard to differing weather conditions the authors analyzed shifts in the distribution of hatching sizes. Frequency distributions of hatch sizes of fish in the "before" and "after" populations were modeled as the product of two multinomial distributions.

$$L(\pi_{01} \dots \pi_{0k} \pi_{11} \dots \pi_{1k}) = \binom{n_0}{y_{01} \dots y_{0k}} \pi_{01}^{y_{01}} \dots \pi_{0k}^{y_{0k}} \binom{n_1}{y_{11} \dots y_{1k}} \pi_{11}^{y_{11}} \dots \pi_{1k}^{y_{1k}}$$

- π_{ij} is the probability of a fish occurring in the j^{th} size-class at hatching ($i=0,1; j=1,\dots,k$),
- k is the number of size-classes,
- n_i is the total number of fish captured in the i^{th} time period, and
- y_{ij} is the number of fish captured in the i^{th} time period in size-class j at hatching.

The probability of a fish surviving to time period 1 if it was born in size-class j is $\pi_{1j} = (\pi_{0j} S_j R) / (\sum_{j=1}^k \pi_{0j} S_j R)$ where:

- S_j is the probability that a fish of hatch-size j survives all size-selective mortality
- $S_j = (e^{a \cdot \ln(j)}) / (1 + e^{a \cdot \ln(j)})$ is a monotonic, sigmoid curve; a and β are parameters that correspond to the level and rate of change of the curve, estimated using Maximum Likelihood Estimation; and x_j is standard length] and
- R is the probability that a fish survives all random, or non-size-selective, mortality during the period of selection [in this model, we assume $R=1$, or no random mortality, which allows us to estimate the maximum amount of selective mortality for a given size-class as S_j^3].

The biological limitations of this model include:

- that it assumes there is no random survival
- that it assumes monotonic selection [too limiting—provides no output if not monotonic (i.e. disruptive selection)]

This model is useful because it is:

- biologically intuitive and mathematically tractable
- works well with directional selection or no selection

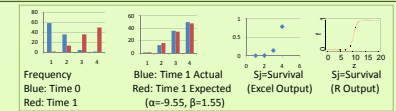
Good, et al.'s model responds to the data as is shown below, first with an example containing data simulated using four j categories and followed by the outputs generated using the three situations of simulated data presented above.

A simple example to follow Good, et al.:

$j:$ 1 3 5 7
Time 0: 59 36 4 1 $S_1 = 100$
Time 1: 1 13 36 50 $S_2 = 100$

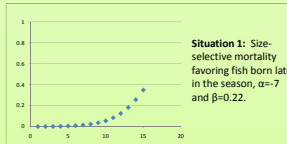
$$S_j = \frac{e^{a \cdot \ln(j)}}{1 + e^{a \cdot \ln(j)}}$$

($a = -9.55$ and $\beta = 1.55$ obtained based on trial and error)

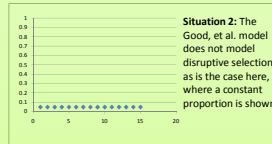


Time 1 Expected Frequency

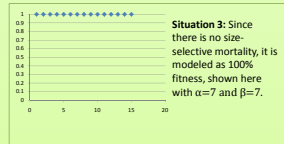
$S_1 = e^{(-9.55 \cdot \ln(1.55))} / (1 + e^{(-9.55 \cdot \ln(1.55))}) = 0.000335$ $\pi_{01} = 59/100 = 0.59$ $\pi_{11} = 1/100 = 0.01$ $\pi_{01} * S_1 = 0.000198$ $(0.000198/0.016391) * 100 = 1.2071$
 $S_3 = e^{(-9.55 \cdot \ln(1.55^3))} / (1 + e^{(-9.55 \cdot \ln(1.55^3))}) = 0.007392$ $\pi_{03} = 36/100 = 0.36$ $\pi_{13} = 13/100 = 0.13$ $\pi_{03} * S_3 = 0.002661$ $(0.002661/0.016391) * 100 = 16.234$
 $S_5 = e^{(-9.55 \cdot \ln(1.55^5))} / (1 + e^{(-9.55 \cdot \ln(1.55^5))}) = 0.141851$ $\pi_{05} = 4/100 = 0.04$ $\pi_{15} = 36/100 = 0.36$ $\pi_{05} * S_5 = 0.005674$ $(0.005674/0.016391) * 100 = 34.6164$
 $S_7 = e^{(-9.55 \cdot \ln(1.55^7))} / (1 + e^{(-9.55 \cdot \ln(1.55^7))}) = 0.785835$ $\pi_{07} = 1/100 = 0.01$ $\pi_{17} = 50/100 = 0.50$ $\pi_{07} * S_7 = 0.007858$ $(0.007858/0.016391) * 100 = 47.9425$
 $\sum \pi_{0j} * S_j = 0.016391$ $\Sigma = 100$



Situation 1: Size-selective mortality favoring fish born late in the season, $a = -7$ and $\beta = 0.22$.



Situation 2: The Good, et al. model does not model disruptive selection, as is the case here, where a constant proportion is shown.



Situation 3: Since there is no size-selective mortality, it is modeled as 100% fitness, shown here with $a = 7$ and $\beta = 7$.

Which model is better?

It is easier to produce a fitness function using Schluter's model (presented in Schluter 1988² and Anderson 1995¹ and modeled using Schluter's software[†]). The Good, et al.³ model is presented in such a way as to make it easier to understand from a biological viewpoint. A direct comparison of these models cannot be done because their outputs are not exactly the same.

Why this is important:

Modeling data allows us to develop an understanding of survival trends among a specific fish population. Examination of the function provided by the model shows how fitness is affected by birthdate (i.e. whether being born earlier or later in the season is more beneficial and/or promotes survival). Relative fitness can differ by year, location, and species; other factors such as weather can have an impact as well.

References:

- Anderson, C. S. 1995. Calculating Size-dependent Relative Survival from Samples Taken Before and After Selection. *The Belle W. Baruch Library in Marine Science* 19: 455-466.
- Schluter, D. 1988. Estimating the Form of Natural Selection on a Quantitative Trait. *Evolution* 42: 849-861.
- Good, S. P., Dodson, J. J., Meehan, M. G., and Ryan, D. A. J. 2001. Annual variation in size-selective mortality of Atlantic salmon (*Salmo salar*) fry. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1187-1195.
- <http://www.zoology.ubc.ca/~schluter/software.html>
- <http://www.famer.unsw.edu.au/research/otoliths/otoliths-main.jpg>
- <http://www.marinebiodiversity.ca/otolith/english/home.htm>
- http://globallyfisher.com/gallery/rod-sutterby/68_coho_salmon.jpg
- <http://www.nmfs.noaa.gov/pr/species/fish/cohosalmon.htm>

Acknowledgements:

This project was funded by the Blugold Fellowship program.