Some thoughts on model forms

Mark Scheuerell

14 March 2022

Background

The total return of adult salmon of species i in year t, $Y_{i,t}^A$, is the sum of natural- and hatchery-origin fish, $N_{i,t}^A$ and $H_{i,t}^A$, respectively, such that

$$Y_{i,t}^A = N_{i,t}^A + H_{i,t}^A. \tag{1}$$

Natural-origin fish

The number of natural-origin adults returning in some year t is the sum of age-specific returns of fish over previous years, the range of which depends on the species and location. In any given year t, adult spawners produce some number of offspring that survive to return as adult "recruits" over the following years ($R_{i,t}^A$). This process is generally assumed to be density dependent and affected by the environment (E_t), such that

$$R_{i,t}^A = f(N_{i,t}^A, E_t) \tag{2}$$

For example, one common choice for f is a Ricker model defined as

$$R_t = \alpha_{i,t} S_{i,t} \exp(-\beta S_{i,t} + w_{i,t}), \tag{3}$$

where α_t is the intrinsic rate of productivity (spawners per spawner) absent density-dependent effects, b is the per capita strength of density dependence, and w_t is some stochastic effect of the environment. In log-space the model becomes

$$\log(R_t) = \log(\alpha_{i,t}) + \log(S_{i,t}) - \beta S_{i,t} + w_{i,t}. \tag{4}$$

Furthermore, the intrinsic rate of productivity is often assumed to be a function of the environment, where

$$\log(\alpha_{i,t}) = \beta_0 + \sum_{i=1}^{M} \beta_{X,i} \; X_{i,t+h} \tag{5}$$

Here, β_0 is the log of the underlying mean productivity, and β_i is the effect of a covariate i at time t ($X_{i,t+h}$).

The estimated number of fish of age a returning in year t ($N_{a,t}$) is then product of the total number of brood-year recruits in year $t - a$ and the proportion of mature fish from that brood year that returned to spawn at age a ($p_{a,t-a}$), such that

$$N^A_{i,a,t} = R^A_{i,t-a} \; p_{i,a,t-a}. \tag{6}$$

The total number of natural-origin adults returning in some year t is then the sum of age-specific returns of fish over previous years, whereby

$$N^A_{i,t} = \sum_a N^A_{i,a,t}. \tag{7}$$

Hatchery-origin fish

Unlike natural-origin fish, the number of hatchery-origin adults in year t ($H^A_{i,t}$) is disconnected from the number of hatchery-origin adults in some years prior, and instead is merely the total number of hatchery-origin juveniles released some k years prior $H^J_{i,t-k}$ times their juvenile-to-adult survival

$$H^A_{i,t} = \sum_{k=1}^{K} s^H_{i,t-k} H^J_{i,t-k}. \tag{8}$$

Juvenile-to-adult survival

Juvenile-to-adult survival is typically thought to be a function of the river environment during the time the juveniles are migrating to sea ($E^R_{i,t-k}$), as well as conditions in the ocean during their time at sea ($\sum E^O_{i,t-k}$), whereby

$$\text{logit}(s^H_{i,t-k}) = \sum_{k=1}^{K} (\beta_{R,k} E^R_{i,t-k} + \beta_{O,k} E^O_{i,t-k}). \tag{9}$$

Simplifying assumptions

Several simplifying assumptions are necessary here given the lack of and species- and age-specific data. First, we can assume that there is a single (average) age of adults and therefore

k is fixed at a single value for each species. If so, then equation (4) can be rewritten in terms of total adults in return year t as a function of the total number of returning adults k years earlier:

$$\log(N^A_{i,t}) = \log(\alpha_{i,t}) + \log(N^A_{i,t-k}) - \beta_N N^A_{i,t-k} + w_{i,t}. \tag{10}$$

We can use the same assumption to simplify the model for hatchery-origin adults, such that equation (8) becomes

$$H^A_{i,t} = s^H_{i,t-k} H^J_{i,t-k}. \tag{11}$$

In log-space, the number of hatchery-origin adults is then

$$\log(H^A_{i,t}) = \log(s^H_{i,t-k}) + \log(H^J_{i,t-k}). \tag{12}$$

Combining equations (10) and (12) gives

$$\log(N^A_{i,t}) = \left[\log(\alpha_{i,t}) + \log(N^A_{i,t-k}) - \beta_N N^A_{i,t-k}\right] + \left[\log(s^H_{i,t-k}) + \log(H^J_{i,t-k})\right] + w_{i,t} \tag{13}$$

with

$$\log(\alpha_{i,t}) = \beta_0 + \sum_{i=1}^{M} \beta_{X,i}\ X_{i,t-kh}. \tag{14}$$

Here we can make the second assumption that the smolt-to-adult survival (SAR) of hatchery fish is known from independent studies and therefore fix the $\log(s^H_{i,t-k})$ term in equation (13) rather than estimate it.

Lastly, we can write out the effects of flow (F), ocean conditions (O) and spending (S) more explicitly, such that equation (14) becomes

$$\log(\alpha_{i,t}) = \beta_0 + \beta_{F,i} F_{i,t-l} + \beta_{O,i} O_{i,t-k} + \beta_{S,i} S_{i,t-m}. \tag{15}$$

Importantly, here the covariates are not log-transformed, as it's not clear to me why they should be given their presumed effect on productivity.