

Working document on reservoir survival theory for passage models

Date July 22, 2005 - JJA

Introduction

This is a working document describing the development of reservoir survival models for juvenile migrants.

Data

Pit tag data for calibrating and comparing reservoir survival model are available between 1995 and 2003. The data are stratified into groups representing three run types (yearling chinook, subyearling chinook, and steelhead), two rearing types (wild and hatchery), and two reach types: tributary reaches (releases to Lower Granite Dam) and hydrosystem reaches (Lower Granite to Little Goose, Little Goose to Lower Monumental, Lower Monumental to McNary, Lower Granite to McNary, McNary to John Day, John Day to Bonneville, and McNary to Bonneville).

Suggested stratifications are illustrated in Table 1 and Table 2. We may want to stratify hydrosystem data into wild and hatchery releases. The data could also be extended over the years 2004 and 2005.

Table 1. Daily groups of data for tributaries above hydrosystem for year 1995-2003

Species	Yearling Chinook				Steelhead				Subyearling Chinook	
Rear Type	Hatchery		Wild		Hatchery		Wild		Combined	
Years	1995 - 2003	Even years	1995 - 2003	Even years	1995 - 2003	Even years	1995 - 2003	Even years	1995 - 2003	Even years
Daily releases	469	213	203	61	581	257	92	42	220	89

Table 2. Daily groups of data within hydrosystem for year 1995-2003.

Reach	Yearling Chinook			Steelhead			Subyearling Chinook		
	LGR-BON	LGR-MCN	LGR-BON	LGR-BON	LGR-MCN	LGR-BON	LGR-BON	LGR-MCN	LGR-BON
Daily releases	3001	576	1222	1750	443	369	1463	218	662

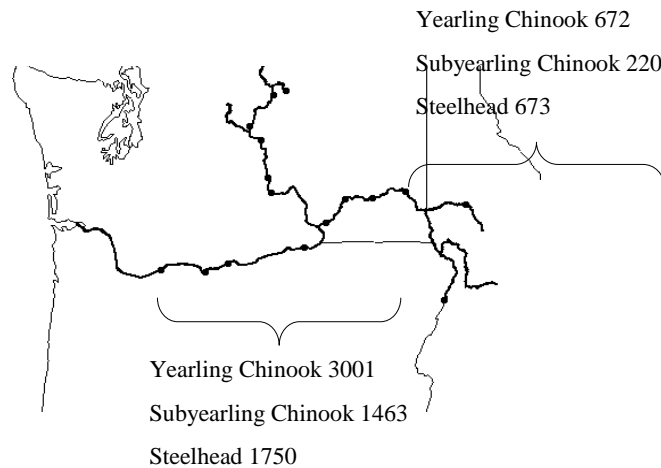


Figure 1. Number of survivals from daily releases of Pit tag data 1995-2003

Caveats on data

Numbers pre release

The number of daily release groups varied considerably. For example, for yearling Chinook through the hydrosystem, the release sizes of the 0, 25%, 50%, 75% and 100% quartiles for the 715 release groups were 3, 85, 234, 925, and 11,340 fish.

Tributaries

In the tributaries reaches studies on hatchery yearling chinook have been conducted by NOAA Fisheries (Muir et al. 2001). From these data, we obtained 469 survival estimates derived from eight hatchery stocks in the Snake River Basin released at 49 locations ranging from 52 to 630 km upstream of Lower Granite Dam. Wild spring chinook consisted of 203 groups released from 13 locations ranging from 52 to 621 km upstream of Lower Granite Dam. Ninety-three percent of the fish were captured with scoop or screw traps and tagged in-river. Eight-six percent originated from Imnaha River, Salmon River, and Johnson Creek traps. NOAA scientists (Muir et al. 2001) have measured hatchery steelhead survival since 1995; from this data, we obtained 581 survival estimates of fish released at 35 sites distributed over the Snake River Basin. Migration distances to Lower Granite Dam ranged from 49 to 696 km. Of the 92 survival estimates of wild steelhead, 79% are from the Imnaha trap, 142 km upstream of Lower Granite Dam. The remaining 19 survival estimates are distributed among 7 sites ranging between 52 and 436 km upstream of Lower Granite Dam. Subyearling chinook data originated from two studies: a NOAA study in which Lyons Ferry subyearling chinook were released predominantly in the Snake and Clearwater rivers (Smith et al. 2003), and a U.S. Fish and Wildlife Service study of wild subyearling fall chinook in the Snake River (Connor et al. 2003b). We combined the data from the two studies because the wild fish data exhibited a correlation between tagging date, tagging distance and fish length that biases the relationship of survival to model covariates. The NOAA study did not have this bias and the results from the combined dataset were very similar to the results of the hatchery fish alone. We updated these data through 2003 from the PITAGIS database, producing a group of 220 survival estimates released between 1995 and 2003 at multiple locations in the Snake and Clearwater rivers.

River covariates

The covariates associated with survival were extracted from DART (CRDART 1998). The indices, including temperature, transparency, total river flow, and fraction of water spill at dams, represent averages experienced by daily release groups between the day of release and the median day of detection downstream. Covariates for the hydrosystem include average of measures at dams between the release and recapture sites over the period of travel. Thus, for migration between Lower Granite and McNary dams between April 15 and May 1, index temperature is the average of temperatures at all dams over the reach between the two dates. For tributary releases, the indices were calculated as the average of the daily values at Lower Granite Dam and the release site or the nearest observation site.

Model forms

Our approach is to define one or more models describing the impacts of river covariates on smolt survival. We clearly do not have a complete understanding of the underlying processes and so models may be proposed with different parameters and different functional forms relating the covariates to survival. Here we discuss several of the considerations of model approaches. Ideally we would include several of these models in the overall structure of the Engine.

Models can be either empirical in which the relationship between covariates and survival are developed from empirical equations without a biological basis or they can be developed in terms of biological principles. We consider both.

Empirical models

An empirical model can be described as empirical regressions in which the coefficients have weak or no biological or mechanistic basis. Table 3 gives the empirical reservoir survival equations proposed for various reaches and species of the Columbia. Of this list models 3, 5, and 6 have some rudimentary biological processes. The other models are empirical fits to data. Models 3, 5, 10 and 11 confine S within the allowable range (0,1) others have no restrictions on the range of S .

Table 3. Columbia/Snake river smolt survival equations. Empirical parameters = a , b , c , d . Ecological parameters = m , n_1 , n_2 , ω , λ . S = migration survival. Covariates; S_{dam} = dam survival, B = cumulative spill fraction, D = migration day, F = flow, F^* = flow-survival threshold, N = number of dams in reach, T = fish travel time, T_w = water travel time, θ = temperature, V = migration velocity, X = migration distance, Z = transparency.

Model	Survival Equations	Reference
1	$S = (a + bF + cF^2)^N$	Sims and Ossiander (1981)
2	$S = \begin{cases} a + bF & F < F^* \\ bF^* & F \geq F^* \end{cases}$	Chapman et al. (1991), Williams et al. (2004)
3	$S = S_{\text{dam}} \exp(-ae^{b\theta}T)$	CBR (2003)
4	$S = S_{\text{dam}} (1 + a) \exp(-bT) / (1 + a \exp(-bT))$	Marmorek et al. (1996)
5	$S = S_{\text{dam}} \exp(-aX)$	NMFS (2000b)
6	$S = a + bD + cF + d\theta$	Smith et al. (2002)
7	$S = a + bT_w + cB + d\theta$	Petrosky et al. (2003)
8	$S = a + bT_w$	Petrosky et al. (2003)
9	$S = \exp(a - b/F)$	ISAB (2003)
10	$S = S_{\text{dam}} \exp(-aX\theta^m)$	Anderson (2003a)
11	$S = a / (1 + \exp(b - cF))$	Smith et al. (2003)
12	$S = a + bF + c\theta$	Connor et al. (2003b)
13	$S = a + bF + cF^2$	Smith et al. (2003)
14	$S = a + b\theta + c\theta^2$	Smith et al. (2003)
15	$S = a + bZ + cZ^2$	Smith et al. (2003)

Mechanistic models

In developing mechanistic models we need consider the system in terms of its first principles.

Interaction of River and Dam Passage mortality

Unlinked mortality

Because smolts in the Columbia River system pass through rivers (including reservoirs and free flowing reaches) and dams, we first consider how survival in the two environments is linked. Typically models assumed that dam passage mortality is wholly independent of river mortality. Mathematically this is expressed

$$(1) \quad S = \exp\left(-\left(r_{river} + r_{dam}\right)\right) = S_{river} S_{dam}$$

where S is total passage survival and the river and dam passage mortality rates, designated r_{river} and r_{dam} , are independent.

Dam passage mortality is typically characterized by the fraction of fish passing through a dam's spillways, turbines, and fish bypass system. Other bypass can be included in this structure also. In a three passage route example the fraction of fish passing each route depends on the fraction of flow spilled and the guidance efficiency of fish into the bypass system. Each passage route has an associated mortality and survival is lowest when the spill fraction is zero. The total dam passage survival is expressed in terms of the number of routes is

$$(2) \quad S_{dam,i} = \left(S_{bypass,i} * FGE_i + S_{turbine,i} (1 - FGE_i)\right) * (1 - B_i) + S_{spill,i} B_i$$

where the subscript i refers to a particular dam, FGE is the fish guidance efficiency, B is the fraction of fish passing with the spill, and S_{bypass} , $S_{turbine}$, and S_{spill} are smolt survivals through the routes. Passage route survivals, as well as FGE , can be estimated independently, such as defined in SIMPAS or given in Ferguson et al. (2004). Dam passage survival varies with the daily spill fraction, B , and the total dam passage survival is $S_{dam} = \prod_i S_{dam,i}$.

Linked mortality

Although eqs. (1) and (2) have been used in every passage model to date, biologists have hypothesized that the dam passage experience may stress fish, resulting in what has been termed delayed mortality that occurs downstream (Budy et al. 2002; Williams et al. 2004). If delayed mortality occurs from dam passage, then eq. (1) is incorrect. One alternative is to assume that the river and dam passage mortality rates are multiplicative, resulting in the equation relating reservoir and dam passage survival as

$$(3) \quad S = \exp(-r_{river}r_{dam})$$

where r_{river} represents a yet undefined function of river properties.

For simplicity and mathematical tractability, we might represent dam passage mortality rate as a function of the fraction of the river flow spilled, B , as

$$(4) \quad r_{dam} = r_0 \exp(bB)$$

The above formula has flexibility in that the mortality rate can either increase or decrease with spill and, in the range $B < 1$, the rate is essentially linear with spill. Additionally, the exponential form allows the link function in the regression equation for survival to be expressed as bB .

As a foot note if the dam passage and river mortalities are linked then estimation of the coefficients within the model cannot be estimated independently. This has both advantages and disadvantages when we consider calibration. Also further analysis of this proposed linkage is in order as in the consideration of the problems with unlinked mortalities. Another approach to linking stress in dam passage to mortality in the reservoir is through a vitality framework (Anderson 2000, Cobleigh and Anderson [to be submitted this summer]). In any case, these alternative approaches indicate more is to be learned in formulating mortality in hydrosystem passage.

Reservoir survival

Whether or not we consider dam and reservoir mortalities linked or unlinked we can develop a reservoir equation on first principles. Under the assumption that ultimately smolts mortality in the reservoirs occurs in the jaws of a predator, although the

probability of getting in the jaws may depend on the fish stress or vitality level then we may proceed. One framework is in terms of theory by Anderson, Gurarie and Zabel (2005) that characterizes smolt/predatory encounters with mean free-path lengths. To understand this theory, consider that smolts migrating directly through a field of stationary predators pass a predator gauntlet and survival depends on the distance traveled, i.e., the number of predators encountered, not the time it takes the smolts to travel the migration distance. However, if smolts and predators have random movement, then multiple encounters can occur and survival may also depend on the time it takes to complete the migration. The relative significance of distance and time on survival depends on the random predator/prey encounter speed, ω , and the smolt migration velocity, $V = X/T$, where T is migration time and X is migration distance. Reservoir survival equation is

$$(5) \quad S_{river} = \exp\left(-\frac{X}{\lambda} \sqrt{1 + \omega^2/V^2}\right)$$

where λ , the mean free-path length a prey travels before encountering a predator, depends on the density of actively feeding predators ρ , and the capture distance r , resulting in

$$(6) \quad \lambda = 1/\pi r^2 \rho$$

When $V > \omega$, smolts effectively pass a gauntlet of stationary predators and the mortality rate depends on distance so $S_{river} = \exp(-X/\lambda)$. Note this is the SIMPAS model equation (Equation 5 in Table 3). When $V < \omega$, the random velocity dominates and survival depends on the migration time so $S_{river} = \exp(-T\omega/\lambda)$. This is equivalent to Equation 3 in Table 3 if temperature is constant.

The parameters r , ρ , ω and V may vary with environmental conditions and the physiological condition of the fish. Therefore, we require biologically realistic equations linking conditions to the model parameters. Furthermore, the forms of equations must be amenable to statistical regression.

In mean free-path length based equations, we need to link the model parameters to environmental covariates. These in turn must enter through λ as defined by eq. (6)

Predator density

The predator density ρ is the density of actively foraging predators in the migratory path of the smolts. Whether a predator is active depends on its metabolic demand, which is depends on f temperature, and whether the fish is satiated. In turn, the evacuation rate of predators and the number of prey required to satiate a predator are functions of temperature (Cochran and Adelman 1982). Laboratory studies indicate that the number of smolts that northern pikeminnow, a major smolts predator in the Columbia River, consume to reach satiation, C_{max} , increases with temperature. Under the range of temperatures smolts encounter in migration, we can define the response function as $C_{max} = 0.05 \exp(0.2\theta)$ where θ is temperature (Vigg and Burley 1991). At 10°C, predators satiate with a third of a smolt; at 20°C, they require 2.7 smolts. Northern pikeminnows located 5 miles below John Day dam encountered 0.15 smolt/predator/day (Vigg et al. 1991), which is 40% of satiation at 10°C and 5% at 20 °C. Using a smallmouth bass evacuation rate (Essington et al. 2000), a predator evacuates half of a full stomach in 15 hrs at 20°C and in 50 hrs at 10°C. Thus, for the Columbia River, it is reasonable that a significant fraction of the predators would be partially satiated and therefore less active at lower temperatures than at higher temperatures. In contrast, Essington et al. (2000) concluded that satiation is unlikely in largemouth bass feeding in lakes because the stomach fullness was at a fraction of the satiation level (0 to 0.1). They suggested that the low feeding rates of the bass were due to behavioral responses between the predator and prey. However, if temperature, gut evacuation, and satiation work in concert to affect predator foraging behavior, then it is reasonable to expect that temperature, through several mechanisms, affects the density of active predators. For our purposes, we assume the fraction of the predator population actively foraging increases in proportion to the number of smolts required to satiate a predator. From Vigg and Burley's (1991) C_{max} equation, the effective predator density is then

$$(7) \quad \rho = \rho_x e^{m\theta}$$

where ρ_x is some active density reference, which may depend on the reach length or location, and m expresses how the active density increases with temperature θ . For

predation to increase with temperature we require $m > 0$, and from Vigg and Burley we hypothesize $m \sim 0.2$.

The predator density also depends on the reach. For smolts that migrate from different tributaries into Lower Granite Reservoir, ρ_X may reflect the distribution of predators between the reservoir and the upstream tributaries. If the density were greater in the reservoir than in the tributaries, then a migration in which the reservoir represents a small fraction of the total path would have a lower density per unit distance than a migration in which the reservoir represents a larger fraction of the path. We characterize possible distributions of predators in terms of reach length as

$$(8) \quad \rho_X = \rho_0 X^k$$

Capture distance

Laboratory predation studies indicate that a predator's reactive distance increases asymptotically with water clarity (Vogel and Beauchamp 1999; Sweka and Hartman 2001). However, water clarity also may work to the advantage of the prey, giving them more time to avoid a predator attack. Thus, we expect that predator/prey interactions with water transparency may be nonlinear. To allow for a complex response, we might represent the capture distance as an exponential function that is quadratic with water transparency distance Z as

$$(9) \quad r^2 = \alpha_0 \exp(n_1 Z + n_2 Z^2)$$

where α_0 is a scaling coefficient for the predator/prey capture area and n_1 and n_2 are coefficients that express how the capture area changes with transparency. We have had some success with this form but it would be more desirable to express reaction distance in fewer terms.

Combining terms

Combining these covariate terms, the river mortality rate is can be expressed

$$(10) \quad r_{river} = -aX^c e^{m\theta + n_1 Z + n_2 Z^2} \sqrt{1 + \omega^2/V^2}$$

where $a = \rho_0 \alpha_0$ is scaling coefficient and $c = 1 + k$.

Final Survival model

The final mechanistic survival equation depends on the assumption of how dam passage affects river mortality. If the dam passage mortality rate is additive, as expressed by eq. (1), the survival equation is

$$(11) \text{ Model 16} \quad S = S_{dam}(B) \exp\left(-aX^c e^{m\theta+n_1Z+n_2Z^2} \sqrt{1+\omega^2/V^2}\right)$$

Where the dam survival is expressed by eq. (2). If dam passage stress is multiplicative, as expressed by eq. (3), then survival is

$$(12) \text{ Model 17} \quad S = \exp\left(-aX^c Z^n e^{m\theta+n_1Z+n_2Z^2+b\sum B} \sqrt{1+\omega^2/V^2}\right)$$

Note that we express the cumulative spill $\sum B$ experienced by the fish in passing dams through the hydrosystem. Note also that spill at each dam occurs over the range $0 < B < 1$ corresponding to none of the water passing over the dam's spillway to all the water passing over the spillway.

Fitting models to data

Because the models in Table 3 plus model 16 and 17 have different forms, the regression equations require a variety of link functions (Table 4). Models 16 and 17 are particularly tricky because of the log square-root term as $k \log(1+\omega^2/V^2)$. We can use MLE techniques on these types of equations if we determined ω using a Newton-Raphson iteration of the regression equation to the condition $k = 1/2$. The approach is to use a trial value of ω in an MLE regression to derive the best estimates for the other coefficients. Each trial value of ω gives some value of k . We select the value of ω and the corresponding MLE estimated remaining coefficients which give $k = 1/2$. This algorithm converges nicely to a proper value in all but a few data sets. the datasets which do not

converges have structural problems that are likely the reason the model does not fit. In particular snake River wild fall Chinook survival.

Table 4. Transform functions for regression survival models. The regressions used the S-PLUS statistical software maximum likelihood regression routine (*lm*), and the nonlinear routine (*nls*) (S-PLUS 2002).

Model	Transform Function	Method	Comments
1	$S^{1/N}$	<i>lm</i>	
2	S	<i>lm</i>	linear regression on first part adjusting F^* to min SS of entire fit
3, 4 11	S	<i>nls</i>	
5	$\log(S/S_{dam})$	<i>lm</i>	If multiple reaches
		<i>mean</i>	If single reach
6-8, 12-15	S	<i>lm</i>	
9	$\log(S)$	<i>lm</i>	
10	$\log(-\log(S/S_{dam})/X)$	<i>lm</i>	
16	$\log(-\log S/S_{dam})$	<i>lm</i>	Ralphson-Newton iteration on ω in $k \log(1 + \omega^2/V^2)$ until $k = 1/2$
17	$\log(-\log S)$	<i>lm</i>	Ralphson-Newton iteration on ω in $k \log(1 + \omega^2/V^2)$ until $k = 1/2$

Comparing model fits to data

If we want to include several models in the Engine, we should select the ones that have the best fit to data. However, because models may have different link functions as is illustrated in Table 4 not all model can be compared with model inference methods such as AIC. However, we can compare models with the same link function. Across models with different link function, we need a different method of comparison. For simpler method, we can compare how close each model predicted survival to the data. For example we might use unweighted linear regression of model vs. observed survivals: $S_{\text{obs}} = a + bS_{\text{model}}$. Then a good fitting model should have the following characteristics

1. High r-squares
2. $a \rightarrow 0$ and $b \rightarrow 1$
3. residuals should be normally distributed across the range of S

We need consider if we want to penalize a fit because of its complexity. One approach is to us AIC to select the best model within a groups of common link forms and then compare the best of each group with the above scenario. There may also be some fancy way to compare across link forms in a AIC like manner.

We have some problem with counting the number of variables when comparing models. How do we handle models where we estimate dam parameters and dam survival outside the fit? Do we count all the model parameters as free variables? This gives model 16 about 50 to 100 more variable than model 17 that characterizes dam passage mortality in in terms of two coefficients!

We also need to consider data stratifications. Ideally, we would like the same model to fit all reaches and species and years so the only thing driving survival are the environmental variables.

It is also important that a model fit between year and within year data.