

Hard choices in assessing survival past dams — a comparison of single- and paired-release strategies

Joseph Zydlewski, Daniel Stich, and Douglas Sigourney

Abstract: Mark-recapture models are widely used to estimate survival of salmon smolts migrating past dams. Paired releases have been used to improve estimate accuracy by removing components of mortality not attributable to the dam. This method is accompanied by reduced precision because (i) sample size is reduced relative to a single, large release; and (ii) variance calculations inflate error. We modeled an idealized system with a single dam to assess trade-offs between accuracy and precision and compared methods using root mean squared error (RMSE). Simulations were run under predefined conditions (dam mortality, background mortality, detection probability, and sample size) to determine scenarios when the paired release was preferable to a single release. We demonstrate that a paired-release design provides a theoretical advantage over a single-release design only at large sample sizes and high probabilities of detection. At release numbers typical of many survival studies, paired release can result in overestimation of dam survival. Failures to meet model assumptions of a paired release may result in further overestimation of dam-related survival. Under most conditions, a single-release strategy was preferable.

Résumé : Les modèles de marquage-recapture sont largement utilisés pour estimer la survie de saumoneaux qui franchissent des barrages. Des lâchers jumelés ont été utilisés pour améliorer l'exactitude des estimations en excluant les composantes de mortalité non attribuables au barrage. Cette méthode s'accompagne d'une réduction de la précision parce que (i) la taille de l'échantillon est plus faible que pour un seul grand lâcher et (ii) les calculs de la variance accroissent l'erreur. Nous avons modélisé un système idéalisé comprenant un seul barrage afin d'évaluer les compromis entre l'exactitude et la précision, et comparé les méthodes en utilisant l'écart-type (RMSE). Des simulations ont été réalisées pour des conditions prédéfinies (mortalité due au barrage, mortalité de référence, probabilité de détection et taille de l'échantillon) afin de déterminer dans quels scénarios les lâchers jumelés sont préférables au lâcher unique. Nous démontrons que les lâchers jumelés n'offrent un avantage théorique par rapport à un schéma de lâcher unique que pour des échantillons de grande taille et des probabilités de détection élevées. Pour des nombres de lâchers caractéristiques de nombreuses études sur la survie, les lâchers jumelés peuvent se traduire par une surestimation de la survie après le passage dans un barrage. Le non-respect des hypothèses sur lesquelles repose le modèle pour les lâchers jumelés pourrait accroître encore davantage la surestimation de la survie après le passage dans un barrage. Dans la plupart des conditions, une stratégie de lâcher unique est préférable. [Traduit par la Rédaction]

Introduction

Mark-recapture methods are a mainstay for estimating demographic rates and trends in abundance of fish and wildlife populations. Advances in telemetry technology and in statistical-modeling techniques have led to widespread use of mark-recapture models to assess survival of downstream-migrating salmon smolts through the use of data collected from longitudinal telemetry arrays (Heupel et al. 2006). Examples of the application of these methods exist for both Pacific (*Oncorhynchus* spp.; e.g., Johnson et al. 2010; Clements et al. 2012) and Atlantic salmon (*Salmo salar*) species (e.g., Serrano et al. 2009; Stich et al. 2014). For modeling purposes, salmon are uniquely suited to these analyses as smolt migrations can be assumed to be "obligate". Fish that fail to leave fresh water of coastal systems during smolt migration likely make minimal contributions to populations (Horton et al. 2009). Software applications such as Program MARK (White and Burnham 1999), RMark (Laake and Rexstad 2009; R Core Team 2014), and USER (Lady and Skalski 2009) provide interfaces that readily allow analysis of spatially explicit

mark-recapture models for fish detected sequentially in rivers and coastal systems during out-migration.

Reaches of rivers through which smolt survival is estimated often include an anthropogenic structure of interest, such as a hydroelectric dam, a spillway, or a fishway. It is often desirable to understand the influence of these structures on fish survival. The criterion for most fish passage evaluations (e.g., to meet hydroelectric licensing requirements) is the proportion of smolts killed by a dam (dam mortality) or its converse, "dam survival" (Φ_{DAM}). Increasingly, hydroelectric operators have been required to assess direct effects of dams on smolt survival and to mitigate losses when dam-induced mortality exceeds a regulatory threshold.

While survival through a reach of river containing a dam (Φ_{REACH}) is estimated by releasing tagged fish upstream of a dam (a "single release"), this estimate includes natural attrition through the reach. This is the theoretical loss of individuals through a river reach if the dam were not present (i.e., "river" survival as described by Perry et al. 2012 and based on Burnham et al. 1987). We

Received 12 October 2015. Accepted 22 June 2016.

J. Zydlewski. US Geological Survey, Maine Cooperative Fish and Wildlife Research Unit, University of Maine, Orono, ME 04469, USA; Department of Wildlife, Fisheries and Conservation Biology, University of Maine Orono, ME 04469, USA.

D. Stich.* Department of Wildlife, Fisheries and Conservation Biology, University of Maine Orono, ME 04469, USA.

D. Sigourney. Integrated Statistics, 16 Summer Street, Woods Hole, MA 02543, USA.

Corresponding author: Joseph Zydlewski (email: josephz@maine.edu).

*Present address: Biology Department, SUNY College at Oneonta, Room 113A Science 1, Oneonta, NY 13820, USA.

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.nrcresearchpress.com/cjfas).

refer to this as “background mortality,” and for a single release we define this bias as “Bias_{BKGRND}”.

For many regulatory and conservation assessments, small differences in survival can be critical, and biased estimates could result in the loss of millions of dollars, undesired ecological consequences, or both. Therefore, resolving the actual impact of within-reach disturbances from the background signal has been the focus of many pioneering mathematical efforts in recent years. This effort has culminated in separate estimation of dam and background survival by releasing fish both upstream and downstream of dams (a “paired release”) and assuming that the difference (ratio) in survival estimates can be attributed to the structure (Skalski et al. 2001). This effort is akin to separating fishing mortality (F) from natural mortality (M) as components of total mortality (Van de Avyle et al. 1999).

The mathematics of the approach have been thoroughly vetted (Burnham et al. 1987; Skalski et al. 2001), and demonstration of the method even has been extended to include applications within multistate modeling frameworks in which there may be various routes of passage with varying survivals (Burnham et al. 1987; Skalski et al. 2009; Perry et al. 2012). Fundamental to the paired-release approach is the assumption that the survival of tagged fish released below the dam closely approximates the background survival of fish moving through the reach of interest. Survival through a dam (i.e., $\hat{\Phi}_{\text{DAM}}$) can then be derived from its relationship to the total survival through the reach of interest ($\hat{\Phi}_{\text{REACH}}$) when the survival associated with background river conditions ($\hat{\Phi}_{\text{BKGRND}}$) can be estimated:

$$(1) \quad \hat{\Phi}_{\text{REACH}} = \hat{\Phi}_{\text{DAM}} \cdot \hat{\Phi}_{\text{BKGRND}}$$

such that survival through the dam is calculated as

$$(2) \quad \hat{\Phi}_{\text{DAM}} = \frac{\hat{\Phi}_{\text{REACH}}}{\hat{\Phi}_{\text{BKGRND}}}$$

Exercising this approach can influence both accuracy and precision of survival estimates in undesirable ways. Because $\hat{\Phi}_{\text{DAM}}$ is derived as a ratio, there is reason to suspect that paired-release estimates of Φ_{DAM} will not be unbiased. This calculation has the potential to overestimate Φ_{DAM} particularly when estimates of the components are imprecise (e.g., at low sample sizes) or when $\hat{\Phi}_{\text{REACH}}$ and $\hat{\Phi}_{\text{BKGRND}}$ are close in value. When the denominator is less than the numerator, this calculation can even produce estimates of $\hat{\Phi}_{\text{DAM}}$ that are not bounded by the probability scale (i.e., survival estimates can be greater than 1.0) and thus are not representative of reality. This can occur in well-executed telemetry studies, and such survival estimates through the dam are set to 1.00 in practice (Skalski et al. 2009, 2014).

Despite this practical adjustment of obviously incorrect survival estimates, it has been noted that $\hat{\Phi}_{\text{DAM}}$ from paired-release studies can contain more subtle (i.e., uncorrectable) bias (Skalski et al. 2013). Commonly used maximum-likelihood estimators are only asymptotically unbiased. Because $\hat{\Phi}_{\text{DAM}}$ is a ratio of random variables, any directional departure from the expected value of $\hat{\Phi}_{\text{DAM}}$ must be non-negative (sensu Skalski et al. 2013), resulting in an overestimation of Φ_{DAM} . We refer to this positive, directional deviation from the expected value of Φ_{DAM} as *process bias* (Bias_{PRCS}). The impact of Bias_{PRCS} on accuracy of $\hat{\Phi}_{\text{DAM}}$ is also known to be greatest at small sample sizes ($n < 1000$) and when detection probabilities are low (Skalski et al. 2013). As accuracy and precision of Cormack–Jolly–Seber (CJS) models depend on “effective” sample sizes, survival and recapture probabilities have great influence. Systematic, positive error in survival estimates resulting from Bias_{PRCS} in paired-release study designs is no more desirable than systematic, negative error in survival estimates resulting from

Bias_{BKGRND} in single-release study designs. At large samples, the expected magnitude of Bias_{PRCS} is presumed to be less than that of Bias_{BKGRND}, but we expect that this may not be the case at smaller sample sizes that are commonly used for dam assessments (e.g., NMFS 2012). Thus, characterizing the trade-off between Bias_{BKGRND} and Bias_{PRCS} is important in understanding the relative expected accuracy of these two approaches at varying sample sizes.

Another computational trade-off in deriving $\hat{\Phi}_{\text{DAM}}$ occurs in the expected precision of $\hat{\Phi}_{\text{DAM}}$ from single- and paired-release methods. For the paired-release method, the standard error (SE) of $\hat{\Phi}_{\text{DAM}}$ is greater than either SE of $\hat{\Phi}_{\text{REACH}}$ or $\hat{\Phi}_{\text{BKGRND}}$. This is due to propagation of error in estimating the variance around the ratio of two random variables through the use of an appropriate Taylor series expansion (e.g., the Delta method; Seber 1982). One presentation of this calculation (among others) is as follows:

$$(3) \quad [\text{SE}(\hat{\Phi}_{\text{DAM}})]^2 \approx [\text{SE}(\hat{\Phi}_{\text{REACH}})]^2 \cdot (\hat{\Phi}_{\text{BKGRND}})^{-2} + [\text{SE}(\hat{\Phi}_{\text{BKGRND}})]^2 \cdot (\hat{\Phi}_{\text{REACH}})^2 \cdot (\hat{\Phi}_{\text{BKGRND}})^{-4}$$

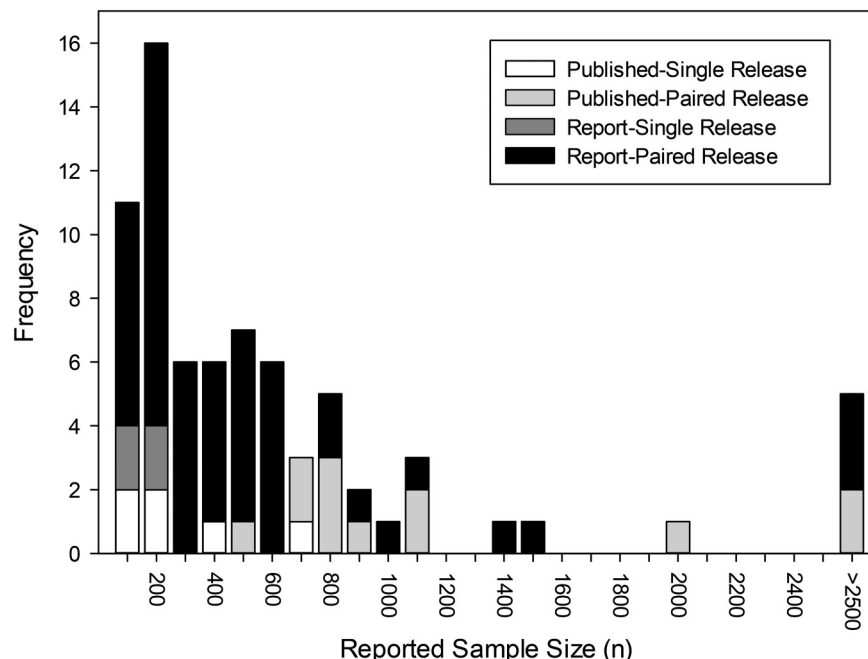
Specifically designed paired-release studies address this reduction in precision of $\hat{\Phi}_{\text{DAM}}$ through the use of larger sample sizes. Some studies have achieved highly precise estimates of Φ_{DAM} by releasing thousands of individuals for dam assessments (e.g., Skalski et al. 2014).

When using active telemetry, however, increasing tag numbers has a substantial cost (US\$200–\$350 for each tag). For large studies, such as those rigorously monitoring Pacific salmon, purchasing an additional 100 tags or more for an assessment may be a trivial expenditure. Indeed, some studies go so far as to release multiple groups downstream so as to reduce the bias associated with the handling effects on below-dam release groups (Skalski et al. 2009, 2013). In other systems, however, the purchase of tags may represent a large — or even the largest — part of a project budget.

We conducted an informal survey of both peer-reviewed literature and accessible reports from 2000 to 2015 that assessed juvenile salmonid survival past dams (Fig. 1). We noted the lowest sample size for which a survival estimate is reported and whether a single- or paired-release protocol was followed (for paired-release approaches, both control and test fish were added to calculate sample size). While this was not an exhaustive search, this survey demonstrated to us that the majority of these studies, regardless of protocol, used fewer than 500 tagged fish to estimate survival (e.g., Skalski et al. 2010; Holbrook et al. 2011; Norrgård et al. 2013; Karppinen et al. 2014). These more modest sample sizes represent a tremendous commitment of resources by researchers or hydropower managers and are representative of some federally prescribed passage assessments (e.g., a required paired release total sample size of 160; NMFS 2012).

Many agencies that regulate smaller operations understandably look to the design of some large-scale efforts to eliminate Bias_{BKGRND} through paired-release methods. This is because of the perceived validity of the technique and the desire for accurate assessment. However, are the release numbers common to studies of dam passage sufficient to warrant the use of a paired-release design in every situation? If we assume that there is a limit to funding for dam impact assessments, then the use of paired releases for estimating survival with a fixed number of tags means that sample size would be split between the two releases. In addition to the error inflation described above, the SEs of $\hat{\Phi}_{\text{REACH}}$ and $\hat{\Phi}_{\text{BKGRND}}$ would then be further increased through the reduction of sample size (compared with SE of $\hat{\Phi}_{\text{REACH}}$ obtained from a single release). This is because SE of a survival estimate is proportional to $1/\sqrt{N}$ (where N is the number of tags; Cormack 1964). Researchers, managers, and consultants are left with a hard choice: release all the tags upstream to minimize the SE of the reach-specific estimate

Fig. 1. Frequency histogram of reported sample sizes used to estimate survival of juvenile salmonids using active telemetry (acoustic or radio) past a dam. This is the result of a survey of both peer-reviewed literature and accessible reports from 2000 to 2015. Bar shading indicates sample sizes reported for each observation are the smallest size for which a survival estimate is reported from each publication. Where multiple releases are performed, the aggregate number is reported (e.g., five weekly releases of 100 fish is reported as 500). For paired-release approaches, both control and test fish (usually reported as “virtual releases”) are included. For study designs including multiple downstream releases, only the most immediate control group was included. This survey demonstrates that the majority of studies surveyed used fewer than 500 tagged fish to estimate survival.



(knowing there is an inherent $\text{Bias}_{\text{BKGRND}}$ resulting in an estimate of “minimum dam survival”) or split the tags into multiple releases that may reduce $\text{Bias}_{\text{BKGRND}}$, but inflate SE and risk $\text{Bias}_{\text{PRCS}}$ in producing an estimate of Φ_{DAM} .

This leads us to the crux of this paper. By viewing single-release Φ_{REACH} as an estimate of Φ_{DAM} , we assessed the theoretical advantage of paired-release strategy through a heuristic model, putting biological issues aside until the discussion. Such a process would not be possible to approach empirically. We stripped down a survival assessment to its bare components and constructed a simple hypothetical dam on a “Model River” and assigned known values for survivals and probabilities of detection. For varied release strategies (number of tags and paired or single releases), we generated capture histories for each permutation of a wide range of parameter values to provide estimates of reach survival with associated error. We drew from ongoing studies of smolt survival on the Penobscot River, Maine, as a rough baseline for the simulations to ensure we bracketed biologically meaningful values for model inputs (Holbrook et al. 2011; Stich et al. 2014). The estimates of reach survival from paired- and single-release designs were compared with the true value (those used to generate the original capture histories) to assess accuracy and precision.

In the end, these data were used to identify the conditions under which paired release is advantageous when minimization of root mean squared error (RMSE) of survival estimates was the criterion. These results, together with logistical considerations, are used to provide general and constructive guidance in study design.

Methods

Creation of the model river

A river section was conceptualized such that moving from upstream to downstream, three sections were delineated by release site R1 and three detection stations, D_1 , D_2 , and D_3 (Fig. 2). The “reach” is the interval of interest here and is bounded by D_1 and

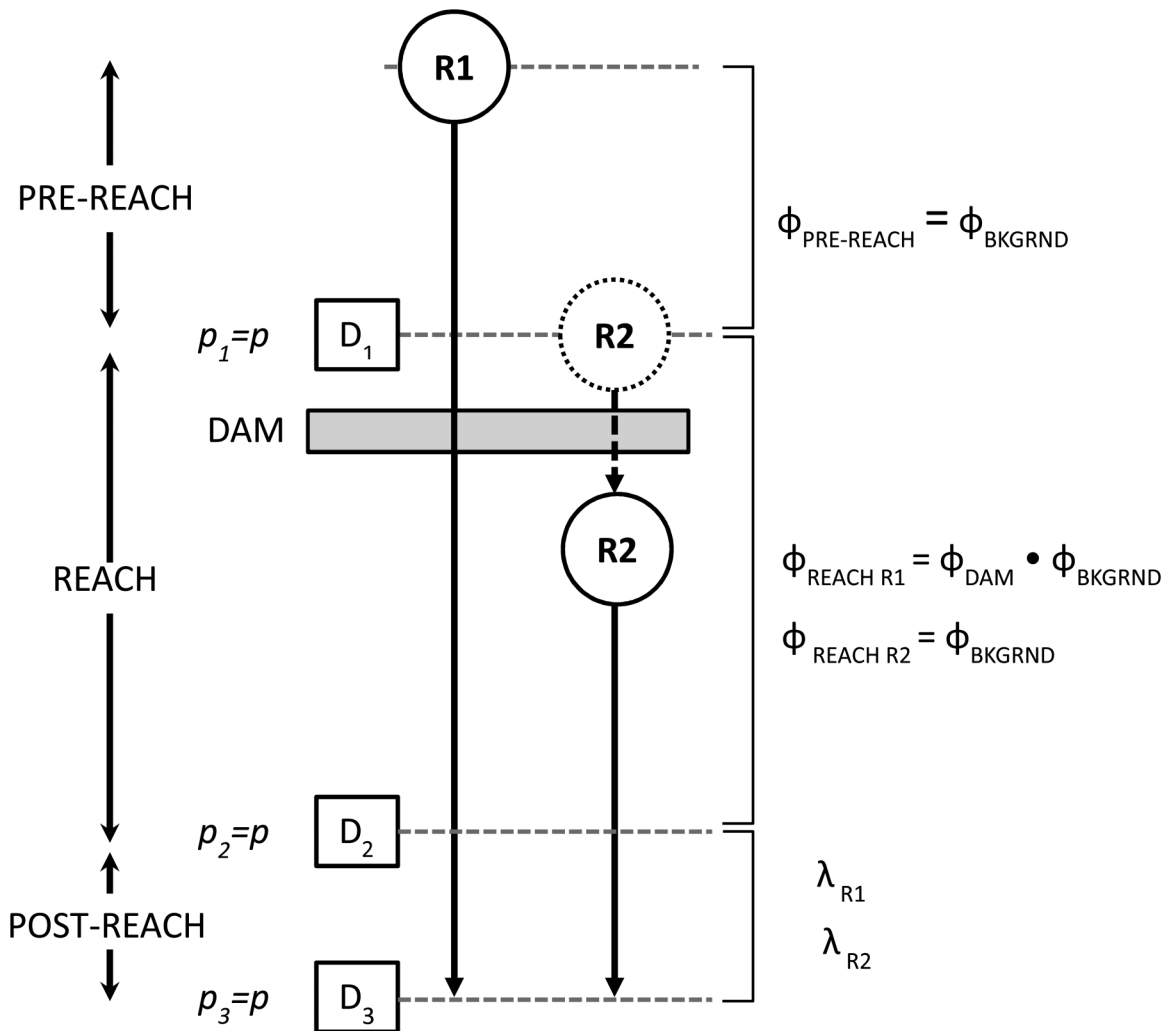
D_2 . The reach interval includes a dam and associated upstream and downstream areas of impact (i.e., head pond and tailrace). The length of this reach and the prereach and postreach intervals are undefined, but are assumed to be equal in length. For “fish” passing through the reach, only one path is allowed. Fish release sites are indicated in Fig. 2 as R1 and R2. We assume that there is no mortality associated with the release.

In a paired release, survival is estimated from the point of release (below the dam) to the next detection (further downstream, here D_2). This estimate is an approximation of the background survival in the river (Φ_{BKGRND}) from D_1 to D_2 . Thus, for modeling a paired release, fish released below the dam are treated as having been released above the dam in construction of capture histories (and the binary occurrence value is set as “1” at D_1 , reflecting a perfect probability of detection for this group). An implicit assumption is that the reach being assessed is sufficiently long so that the difference in length between the modeled release site (above the dam) and the true release site (below the dam) is negligible. Because our reach length is unspecified, this assumption was met by default in our simulations.

A priori definition of parameters to generate recapture histories

Although unknown in the real world, we defined a priori survival values and probabilities of detection to generate recapture histories. We simulated data using a wide range of values for each parameter and used nominal values of each parameter to construct a single-release and a paired-release “base model.” Importantly, all fish in our simulations have not been influenced by tagging or handling in any way. Though survival probabilities are more usefully presented as a rate per distance, our Model River has no length, so we applied interval values that correspond to published ranges of observed interval survival. Estimated survival probability of smolts travelling through free-flowing river sec-

Fig. 2. Idealized diagram of the “Model River” where the apparent survival through a reach with a dam is being estimated. R1 represents the release site for a single-release strategy where all fish are released upstream of the interval of interest. R2 represents the site where the second group of a paired-release strategy is located. For the paired-release strategy, the dotted location of R2 indicates the location assumed in models using a paired release. Locations of detection stations are indicated as “D”. Parameters estimated for detection probability and survival are indicated as p and Φ , respectively. λ is the unresolved estimates of probability of last detection (p_3) and of survival for the postreach (Φ).



tions during downriver migration is generally high and often at or near 0.99 km^{-1} but as low as 0.93 km^{-1} (Thorstad et al. 2012; Stich et al. 2015a). The contribution of survival through a free-flowing reach of river was set equal among intervals (prereach, reach, and postreach). This background survival (Φ_{BKGRND}) was allowed to range incrementally from 0.95 to 1.00, with the base model fixing Φ_{BKGRND} to 0.98 (Table 1).

Survival through the reach of interest (i.e., the reach containing the dam; $\Phi_{\text{REACH R1}}$) and Φ_{DAM} were allowed to vary proportionally to Φ_{BKGRND} . This was done to ensure that expected reach survivals (those values used to simulate the data) could not be greater than 1.0. For each permutation, Φ_{DAM} was calculated as

$$(4) \quad \Phi_{\text{DAM}} = F \cdot \Phi_{\text{BKGRND}}$$

where F (proportional influence of the dam on survival) was allowed to vary incrementally between 0.7 (low survival through the dam) and 1.0 (perfect survival), such that survival attributable through the reach was equal to background survival. For the base model, we assumed F was 0.90 as a moderately low survival such that Φ_{DAM} was 0.882 (0.90×0.98).

To define $\Phi_{\text{REACH R1}}$, we incorporated Φ_{BKGRND} (the defined bias due to background survival in single-release scenarios) and Φ_{DAM} (the survival attributed solely to the dam) as follows:

$$(5) \quad \Phi_{\text{REACH R1}} = \Phi_{\text{DAM}} \cdot \Phi_{\text{BKGRND}}$$

For a single-release scenario, $\Phi_{\text{REACH R1}}$ was used as an estimate of Φ_{DAM} . For a paired release, the estimate of Φ_{DAM} is resolved from $\Phi_{\text{REACH R1}}$ and Φ_{BKGRND} . While probability of detection can differ with water conditions and receiver characteristics, we set probability of detection (p) to be equal among the three detection stations in Model River and was held constant for each release scenario. We varied p incrementally among release scenarios from a low of 0.50 (poor) to 1.00 (perfect). For our base model, p was set to 0.90 (Table 1), representing a moderately high detection rate attained by many modern studies.

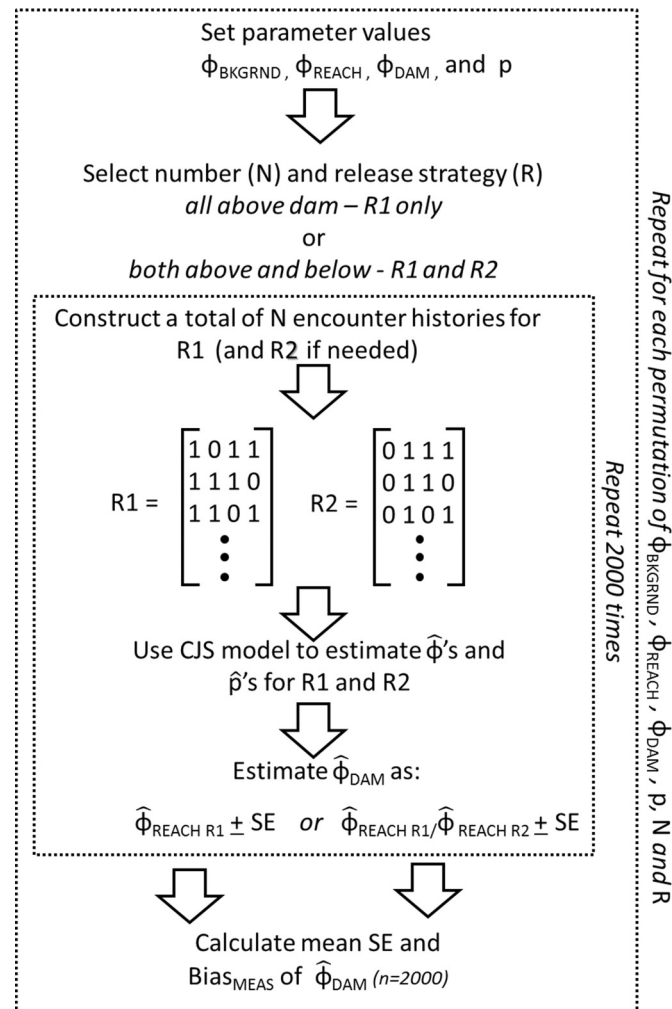
Generation of recapture histories from defined parameters

The a priori values for parameters used to simulate capture histories were combined in all possible ways, resulting in 144 possible scenarios for the parameters described above. Each of the 144 per-

Table 1. Parameters that were used to define simulations.

Parameter	Definition-influence	Value range	n	Interval	Base model
N	Total number of tags used for either release upstream (R1) or both upstream and downstream (R1 and R2)	35–5000	19	82 and 331	—
Φ_{RIVER}	Defined survival probability of fish moving through an interval of natural river habitat	0.95–1.00	6	0.01	0.98
Φ_{DAM}	Survival probability attributed to the dam, a component of $\Phi_{\text{REACH R1}} \cdot \Phi_{\text{DAM}}$ is calculated as a proportional difference from Φ_{RIVER} : $\Phi_{\text{DAM}} = F \cdot \Phi_{\text{RIVER}}$	—	—	—	0.882 ($F \cdot \Phi_{\text{RIVER}} = 0.90 \times 0.98$)
F	Proportional difference from Φ_{RIVER} (used to calculate both Φ_{DAM} and $\Phi_{\text{REACH R1}}$; this was done so that these Φ values could not exceed Φ_{RIVER})	0.70–1.00	4	0.1	0.9
$\Phi_{\text{REACH R1}}$	This is used as an estimate of Φ_{DAM} with a single release (R1), a biased estimator as $\Phi_{\text{REACH R1}} = \Phi_{\text{DAM}} \cdot \Phi_{\text{RIVER}}$	—	—	—	0.864 ($\Phi_{\text{DAM}} \cdot \Phi_{\text{RIVER}} = 0.882 \times 0.98$)
$\Phi_{\text{REACH R2}}$	Defined survival estimator of Φ_{RIVER} ; value set to Φ_{RIVER}	—	—	—	0.98
R	Proportion released above dam; 1.0 indicates only upstream release (R1), 0.5 indicates equality between upstream and downstream releases (R1 and R2)	0.5–1.0	6	0.1	1.0 (single release) or 0.6 (paired release)
p	Probability of encounter at detection stations	0.5–1.0	6	0.1	0.9

Fig. 3. Flow diagram of the process by which probability draws were used to produce simulated recapture histories for single- and paired-release scenarios. Simulated encounter histories were then used to produce estimates of apparent survival (with associated error and bias) through a dam. A total of 2000 replicates were performed for each set of defined parameters.



mutations of the a priori defined values (Φ_{REACH} , Φ_{BKGRND} , Φ_{DAM} , and p) was used to generate capture histories for experimental design conditions (Fig. 3). These conditions include the number of tagged fish released (N) and the proportion of tagged fish released above the dam (R). For N we used a low of 35 (an unacceptably small number for survival estimates; see Castro-Santos and Haro 2013) to a high of 5000 (representative of a logistical upper bound for assessment). For R, we used values from 0.5 to 0.9 (i.e., putting more than half of the tags downstream of the interval being studied) makes little intuitive sense and was not tested. For the paired-release version of the base model, R was set to 0.60 (60% of tagged fish were released above the dam and 40% were released below). For a single release, R was set to 1.0 to indicate that all tagged fish were released above the dam (Fig. 2).

All 144 combinations of survival and detection parameters (Φ_{BKGRND} , Φ_{DAM} , p, and R) were crossed with all release scenarios and sample sizes to result in a total of 16 416 predefined sets of parameter values used to simulate capture histories (Fig. 2). For each of these scenarios, 2000 sets of individual encounter histories were simulated. Encounter histories were simulated using methods from Kéry and Schaub (2012) as follows. All fish were assigned a “1” at release and “0” for occasions (detection stations) upstream of the release. For all possible encounters after release, true survival of each fish was determined by a random draw from a Bernoulli distribution (probability = Φ , $n = 1$). Values of Φ were specific to the interval and corresponded to Φ_{BKGRND} for pre-reach and post-reach intervals and Φ_{REACH} for the reach interval. If a fish survived to the next detection station, whether it was detected (or not) was determined by a second random draw from a Bernoulli distribution (probability = p, $n = 1$). A “1” was assigned to all encounter events where the fish both survived and was detected. If the fish survived but was not detected, a “0” was assigned. If the fish did not survive, a “0” was assigned for the encounter event and for each encounter event that followed.

This process was repeated for all N tagged fish for those released above the dam (R1) and, when applicable, below the dam as a paired release (R2). Each individual encounter history had four records: the first was the release point at R1, followed by the three detection sites (D_1 , D_2 , D_3) with eight possible outcomes:

- 1000
- 1001
- 1010
- 1100
- 1101
- 1110

1011
1111

As per the design of a paired release, though the fish are released below the dam, they are treated as having been released above the dam and detected at D_1 , and each individual is assigned a “1” for the release occasion. Thus, there were only four possible encounter histories. (We present these encounter histories as a string of four so that encounter instances between the releases are aligned.)

0100
0101
0110
0111

Once these encounter histories were generated, they could then be modeled to estimate interval survivals.

Modeling survival from simulated encounter histories

For each set of simulated capture histories, a CJS survival model was developed and analyzed to estimate survivals for each release group (Fig. 3). The R package RMark (Laake and Rexstad 2009; R Core Team 2014) was used for analysis of CJS models using maximum-likelihood estimation (MLE) in Program MARK (White and Burnham 1999). The logit-link function was used for all parameters in all models to restrict estimates to possible values on the probability scale (from zero to one). For releases above the dam (R1), the fitted model included independent estimates of survival (i.e., $\hat{\Phi}_{\text{PREREACH}}$ and $\hat{\Phi}_{\text{REACH R1}}$ were allowed to vary from one another). Encounter (i.e., detection) probability was assumed to be equal among all detection stations. The joint probability of survival and detection probability was estimated for the final interval (represented by λ), and model outputs for releases above the dam included only two survival estimates: $\hat{\Phi}_{\text{PREREACH}}$ and $\hat{\Phi}_{\text{REACH R1}}$. For releases below the dam (R2), models were similarly structured, though a single survival parameter ($\hat{\Phi}_{\text{REACH R2}}$) and a single detection probability (\hat{p}_2) were estimated. In this manner, an estimate of $\hat{\Phi}_{\text{REACH R1}}$ and, where a paired release was included (i.e., $R < 1.0$), $\hat{\Phi}_{\text{REACH R2}}$ was produced for each of the 2000 replicates of each of the 16 416 scenarios. Unless otherwise indicated, data from the MLE approach are those reported in the results.

When fish were only released above the dam, $\hat{\Phi}_{\text{REACH R1}}$ was used as the estimate of $\hat{\Phi}_{\text{DAM}}$, and SE of $\hat{\Phi}_{\text{DAM}}$ was derived using MLE. When paired releases were used, $\hat{\Phi}_{\text{DAM}}$ was calculated using

$$(6) \quad \hat{\Phi}_{\text{DAM}} = \frac{\hat{\Phi}_{\text{REACH R1}}}{\hat{\Phi}_{\text{REACH R2}}}$$

Note that both $\hat{\Phi}_{\text{REACH R1}}$ and $\hat{\Phi}_{\text{REACH R2}}$ were the result of independent draws from distributions (including the same background rate of mortality; Φ_{BKGRND}); therefore, our simulations explicitly meet the assumption that there is no statistical covariance between the two estimators. The SE of $\hat{\Phi}_{\text{DAM}}$ from paired releases was approximated using the Delta method (eq. 3). Because this is a simulation exercise and we defined Φ_{DAM} (an unknown in the real world), we could calculate measured bias ($\text{Bias}_{\text{MEAS}}$; more properly “deviance” for a single value) for each model:

$$(7) \quad \text{Bias}_{\text{MEAS}} = \hat{\Phi}_{\text{DAM}} - \Phi_{\text{DAM}}$$

For each individual CJS model, an SE and $\text{Bias}_{\text{MEAS}}$ were produced. (Recall that $\text{Bias}_{\text{BKGRND}}$ is the source of $\text{Bias}_{\text{MEAS}}$ for a single release, while $\text{Bias}_{\text{PRCS}}$ is the source of $\text{Bias}_{\text{MEAS}}$ for a paired release.) To assess the trade-off between bias and SE for each release strategy, we calculated RMSE as a measure of model fit. RMSE is a

well-established and widely used means of evaluating competing models that puts equal weight on bias and precision (e.g., Willmott et al. 1985; Ross 1996; Smith et al. 2007).

$$(8) \quad \text{RMSE} = \sqrt{\text{SE}^2 + \text{Bias}_{\text{MEAS}}^2}$$

For each set of conditions, results from the 2000 simulations were used to calculate the median, mean, and SE of all $\text{Bias}_{\text{MEAS}}$, SE, and RMSE for each $\hat{\Phi}_{\text{DAM}}$. Values of SE, $\text{Bias}_{\text{MEAS}}$, and RMSE were identified as “extreme” if mean values were greater than 1.0 (e.g., due to lack of model convergence). In these cases, a value of 1.0 was assigned for purposes of analysis and data visualization on a meaningful scale. Such cases only occurred in instances when $N = 35$ (we included these cases to demonstrated why the use of such small sample sizes was unreasonable in practice). The influence of each of the a priori defined parameters and release strategies on $\text{Bias}_{\text{MEAS}}$, SE, and RMSE was assessed.

To assess the conditions under which one strategy (paired or single release) might be advantageous in balancing the accuracy and precision of $\hat{\Phi}_{\text{DAM}}$, mean RMSE ($n = 2000$) of each strategy were directly compared at permutations of a priori and experimental conditions for the base models values ($R = 0.6$ or 1.0 , $\Phi_{\text{BKGRND}} = 0.98$). ΔRMSE was calculated as

$$(9) \quad \Delta\text{RMSE} = \text{mean}(\text{RMSE}_{\text{SINGLE}}) - \text{mean}(\text{RMSE}_{\text{PAIRED}})$$

such that a positive value indicated conditions under which a paired release was advantageous and a negative value indicated where a single release was advantageous.

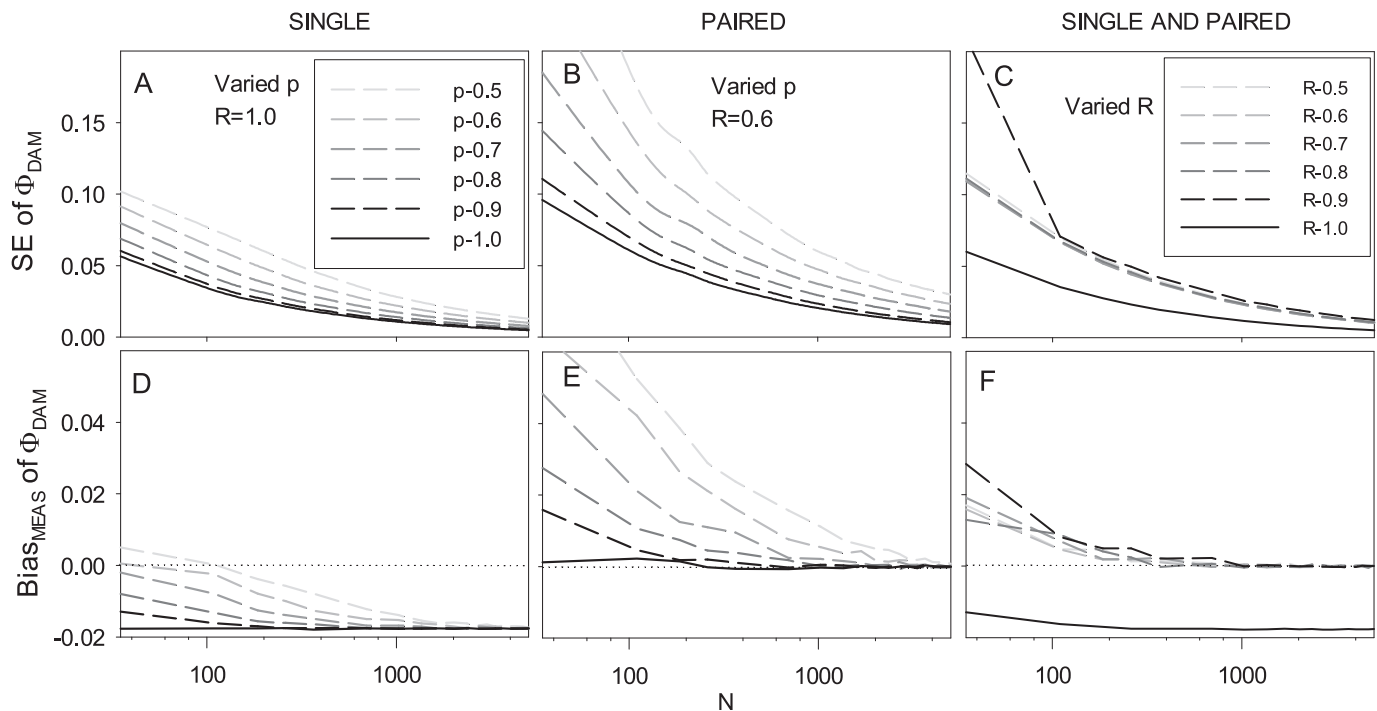
Comparison with Bayesian hierarchical modeling approach

The estimation of parameters for CJS models using the logit scale can result in abnormal behavior in MLE near the bounds of 0 and 1. Conversely, the use of the identity link function that avoids this approach can produce unrealistic survival estimates (outside of the probability scale; i.e., $\hat{\Phi}_{\text{DAM}} > 1.0$). We thought it reasonable to investigate whether these issues could potentially influence the results of this study, and we sought to verify that results were consistent across estimation methods. To do so, we applied a Bayesian hierarchical framework to estimate parameters and characterize measures of uncertainty to a subset of scenarios represented by the base model (Table 1). All structural components of the model (i.e., assumption of fixed detection probability and parameters estimated) were identical to the approach outlined for the CJS model using logit-link function. However, use of the Bayesian approach allowed us to estimate survival directly on a probability scale between 0 and 1 and avoid potential problems resulting from (1) failed model convergence in MLE at small sample sizes and (2) the use of link functions to estimate survival.

CJS models were fit using a Markov chain Monte Carlo (MCMC) approach in the program JAGS (using the R2jags package in R). Uniform, uninformed priors (from 0 to 1 on the probability scale) were used for all model coefficients. Each model was run using a burn-in of 600 samples followed by an adaptive phase of 3000 samples from the posterior distribution of each parameter. Markov chains were thinned by keeping every sixth sample to account for autocorrelation between successive samples and increase the number of effective samples. Convergence was assessed using the Gelman–Rubin statistic (\hat{r}) and by visual inspection of mixing (agreement) among chains. Effective size was sufficiently large to construct posterior estimations of all parameters. For each sample size used (from 35 to 5000) in the base model, 2000 replicates were run for single- and paired-release models.

To incorporate uncertainty of $\hat{\Phi}_{\text{DAM}}$ for paired-release models, we calculated $\hat{\Phi}_{\text{DAM}}$ for each MCMC iteration (as $\hat{\Phi}_{\text{REACH R1}}$ divided by $\hat{\Phi}_{\text{REACH R2}}$), allowing the estimation of the mean and SE to be

Fig. 4. Patterns of mean standard error (SE) and bias of $\hat{\Phi}_{\text{DAM}}$ (from 2000 simulation runs) in relation to probability of detection for a single-release strategy (A and D) or a paired-release strategy (B and E). Patterns of mean SE and bias of $\hat{\Phi}_{\text{DAM}}$ for different proportions of fish released as a second release are shown in panels C and F, respectively. When $R = 1.0$ in panels C and F, the approach by definition is a single-release strategy.



derived directly from the posterior distributions of parameters in each model. This direct approach is notable as it avoids the inflation of SE associated with the Delta method when estimating $\hat{\Phi}_{\text{DAM}}$ for paired-release estimates (eq. 3). Calculations of Bias_{MEAS} and RMSE were identical to the MLE approach described above.

Results

Patterns in SE

In all scenarios, SE of $\hat{\Phi}_{\text{DAM}}$ decreased as N increased (Figs. 4 and 5). The reduction in SE of $\hat{\Phi}_{\text{DAM}}$ was most notable at small to intermediate values of N , with diminishing returns at the largest N . For a single-release scenario (Fig. 4A), lower detection probabilities resulted in a pattern of higher SE at all values of N . However, it is notable that even when $N = 35$ and $p = 0.5$, SE remained less than 0.11 on average. While the same general trend was observed in a paired release where $R = 0.6$, estimated SE of $\hat{\Phi}_{\text{DAM}}$ was greatly inflated when compared with a single-release approach, particularly at low N values. At an N of 35, SE was 0.35 (Fig. 4B).

When p was held constant at 0.9 (base model), the proportion of fish released above the dam greatly influenced SE. Releasing all fish above the dam ($R = 1$) resulted in a notably lower SE than all other release strategies (Fig. 4C). This was apparent at all values of N , but the decrease in SE of $\hat{\Phi}_{\text{DAM}}$ was of greatest magnitude at small sample sizes. When few fish were released below the dam as a component of the paired release, the inflation of SE of $\hat{\Phi}_{\text{DAM}}$ was greatest (Fig. 4C). SE was always greater for paired releases compared with single releases using the same sample sizes. Among paired releases, SE decreased as R decreased from 0.9 to 0.6 (base model value), then increased slightly from 0.6 to 0.5 (Fig. 4C).

Patterns in Bias_{MEAS}

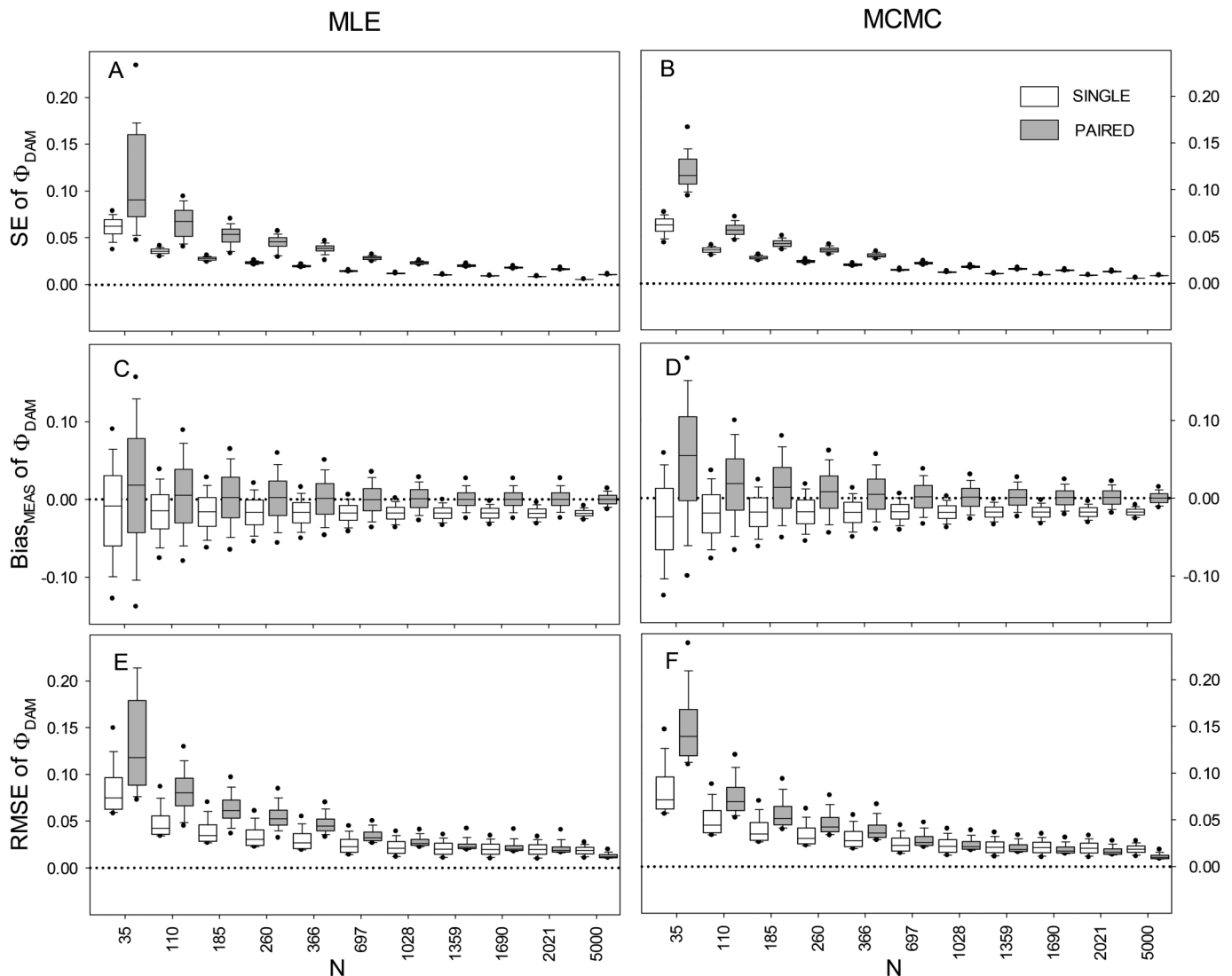
For single-release strategies, Bias_{MEAS} of $\hat{\Phi}_{\text{DAM}}$ was generally, but not always, negative (Figs. 4D and 5C). This difference (attributed to Bias_{BKGRND} for a single release) results in an underestimation of Φ_{DAM} under most conditions. At high N , mean Bias_{MEAS}

approached the defined influence of the a priori Φ_{BKGRND} component of survival (approaching $\Phi_{\text{DAM}} - \Phi_{\text{REACH}}$, or -0.01764 for the base model). At high probabilities of detection, survival is essentially defined by a binomial probability that does not change as a function of N (except for stochastic variation among the simulations). Thus, when $p = 1.0$, Bias_{MEAS} was negative and was not affected by increasing N (i.e., Bias_{MEAS} was equally low at $N = 35$ and 5000) in the single-release model. However, when detection probability decreased ($p < 1.0$), the magnitude of Bias_{MEAS} diminished (was less negative) as a result of increased uncertainty about the fate of fish. At very low detection probabilities (e.g., $p = 0.5$), a positive bias was observed when releases were fewer than 100 fish such that, on average, Φ_{DAM} was overestimated. At low release numbers, mean Bias_{MEAS} was reduced for the single-release models. For example, at an N of 110, Bias_{MEAS} was only 45% of theoretical Bias_{BKGRND} (0.008) at a p of 0.6. The mean Bias_{MEAS} was 90% of the theoretical Bias_{BKGRND} (0.013) at $p = 0.9$ (base model).

The pattern of Bias_{MEAS} for paired release differed from the single release in that any Bias_{MEAS} resulting from model estimates was generally positive (Figs. 4E and 5C). This difference (attributed to Bias_{PRCS} for a paired release) results in an underestimation of Φ_{DAM} under most conditions. The exception to this was at perfect detection probability ($p = 1.0$), for which there was negligible Bias_{MEAS} of $\hat{\Phi}_{\text{DAM}}$ at any release size (although there was a slight positive Bias_{MEAS} of ~ 0.002 at an N of less than 260 due to model stochasticity).

As probability of detection decreased, we observed trade-offs in the direction and magnitude of Bias_{MEAS}. In some scenarios ($p = 0.5$ and $N < 366$), the positive value of Bias_{MEAS} from paired-release models was equal to the negative value of Bias_{MEAS} that resulted from use of single-release models. At $N = 110$ and low probability of detection ($p = 0.5$), the positive value of Bias_{MEAS} resulting from paired releases was more than twofold greater in magnitude than the negative value of Bias_{MEAS} from a single release. At higher probabilities of detection, Bias_{MEAS} was reduced in the results of

Fig. 5. Box plots of standard error (SE), bias, and root mean squared error (RMSE) of $\hat{\Phi}_{\text{DAM}}$ for single- and paired-release strategies using release numbers of 35 to 5000 with parameters set to base model values. Trends resulting from maximum-likelihood estimation (MLE) are shown in the left panels; trends from Bayesian estimation (Markov chain Monte Carlo; MCMC) are shown in right panels. Box indicates 25th and 75th percentiles (central tendency is the median), whiskers are the 10th and 90th percentiles, and outlier symbols indicate the 5th and 95th percentiles for the 2000 simulations of each base model.



paired-release models, but even for $N = 110$ and $p = 0.9$ (base model value), the positive value of $\text{Bias}_{\text{MEAS}}$ from the paired release was more than 30% as great as the $\text{Bias}_{\text{MEAS}}$ from the single-release model (the error that is being corrected for through use of paired releases). It is notable that even at very large sample sizes ($N > 1000$), though the mean value of $\text{Bias}_{\text{MEAS}}$ approaches zero for the paired-release models, the distribution of $\text{Bias}_{\text{MEAS}}$ overlaps considerably with the distribution of $\text{Bias}_{\text{MEAS}}$ from the single-release models (Fig. 5B). Thus, while the paired-release estimate is unbiased at higher sample sizes (as demonstrated by thousands of simulations), individual estimates of Φ_{DAM} frequently deviate from the true value as much or more than estimates from a single release.

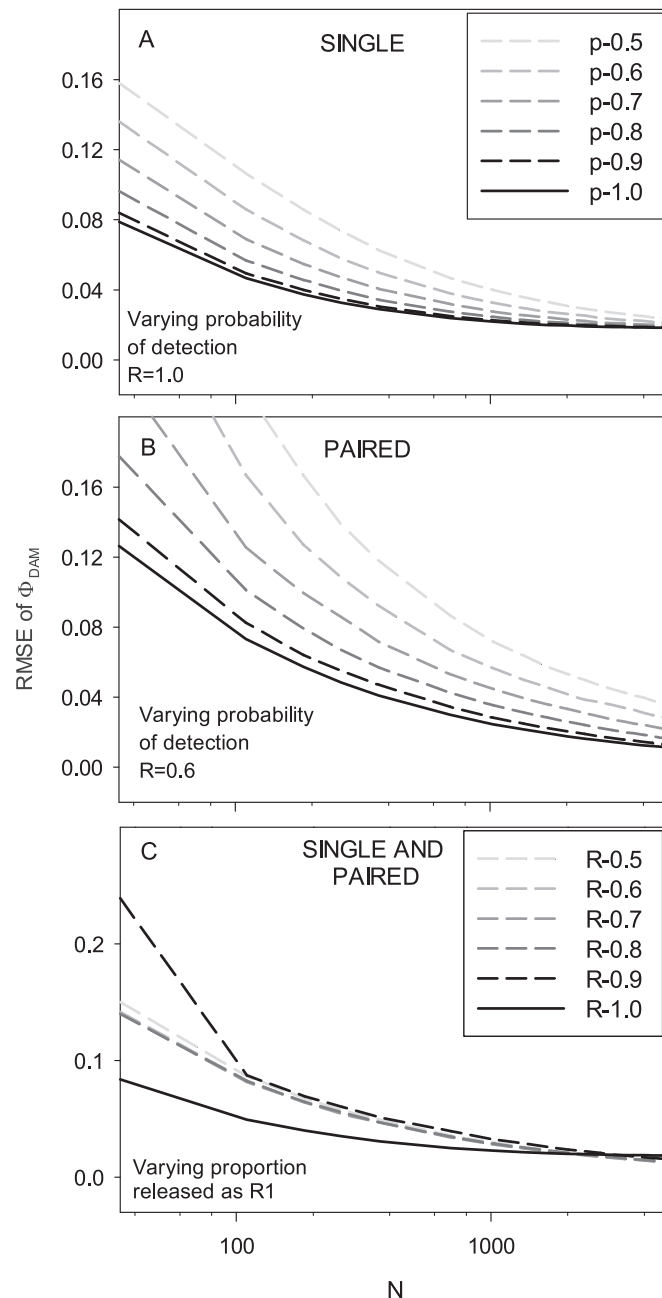
For a paired-release strategy, increasing the proportion of fish released below the dam reduced the mean value of $\text{Bias}_{\text{MEAS}}$. Regardless of the release proportion, release sizes of less than 110 fish resulted in $\text{Bias}_{\text{MEAS}}$ greater than 0.01. At the lowest release numbers, when a paired-release strategy produced the greatest positive values of $\text{Bias}_{\text{MEAS}}$, the negative values of $\text{Bias}_{\text{MEAS}}$ from single-release models were lowest (Fig. 4F), and Φ_{DAM} esti-

mated from the single release more accurately approximated the true value used in simulation. This situation was reversed as sample size increased.

Patterns in RMSE

For a single-release strategy, RMSE decreased with increasing sample size. This decrease was asymptotic to the influence of $\text{Bias}_{\text{MEAS}}$ at the largest values of N (Figs. 5E and 6A). Lower probabilities of detection were associated with higher RMSE in both single- and paired-release models. For single-release models, these curves converge from low N values (where SE dominates) to a minimum value that is dominated by the influence of the $\text{Bias}_{\text{MEAS}}$ as SE declines (due to the influence of $\text{Bias}_{\text{BKGRND}}$; Figs. 5E and 6A). For the paired-release scenario, there is likewise a general decline in RMSE as N increases at perfect probability of detection ($p = 1.0$; Fig. 6B). This results from converse patterns of SE and $\text{Bias}_{\text{MEAS}}$. At low sample size (but high detection, as in our base model), the magnitude of positive values of $\text{Bias}_{\text{MEAS}}$ from paired-release models is low. SE, however, is much higher than at the single-release strategy. As SE declines with increasing N , the RMSE curve asymp-

Fig. 6. Patterns of root mean squared error (RMSE) for $\hat{\Phi}_{\text{DAM}}$ in relation to probability of detection for a single-release strategy (A), a paired-release strategy (B), or for different proportions of fish released as the second release in a paired-release strategy (C). When $R = 1.0$ in panel C, the approach by definition is a single-release strategy.



totes towards zero in the paired-release models. These values are lower than RMSE of a comparable release size for the single-release strategy only at releases of more than 1000 fish. When the probability of detection falls, however, RMSE curves shift as increased SE inflates the RMSE (Figs. 5E and 6B). As decreased probability of detection disproportionately inflates SE in the paired release, RMSE in a paired release is similarly increased.

When holding all other parameters constant in the base model, RMSE was lower for single-release strategies when compared with all other possible paired-release strategies if N was less than 1359. At low N , the use of a small release below the dam resulted in

greatly inflated RMSE (Fig. 6C) resulting from the high SE and $\text{Bias}_{\text{MEAS}}$ generated from small sample size of the second release groups in these scenarios. RMSE was greatly reduced as R decreased from 0.9 to 0.8, with incremental gains being observed until R was reduced to 0.6. Owing to the slight increase in SE observed when R was reduced from 0.6 to 0.5, the RMSE was lowest when 60% of fish were released above the dam.

When using ΔRMSE as a comparison of model performance in estimating Φ_{DAM} , the data demonstrate that there was a limited set of conditions under which a paired release has a mathematical advantage over a single-release strategy (when survival is estimated using maximum likelihood methods and accuracy and precision are valued equally). Given a defined Φ_{BKGRND} of 0.98 as the base model value, these conditions indicate that a paired-release strategy outperforms a single-release strategy at an N greater than 1000 when natural mortality is low (Fig. 7). A single release was favored for all models with fewer fish released, regardless of pre-defined parameter values. It is notable that at $N = 1028$, an advantage to a paired release, was detected only if conditions were restricted to both probabilities of detection and Φ_{REACH} being near 1.0. Even with an N of 5000 fish, there was no mathematical advantage to a paired-release design when the probability of detection was less than 0.7. When the paired-release strategy was advantageous, the advantage (in ΔRMSE) was always less than 0.012.

Comparison of MLE and Bayesian hierarchical modeling of CJS parameters

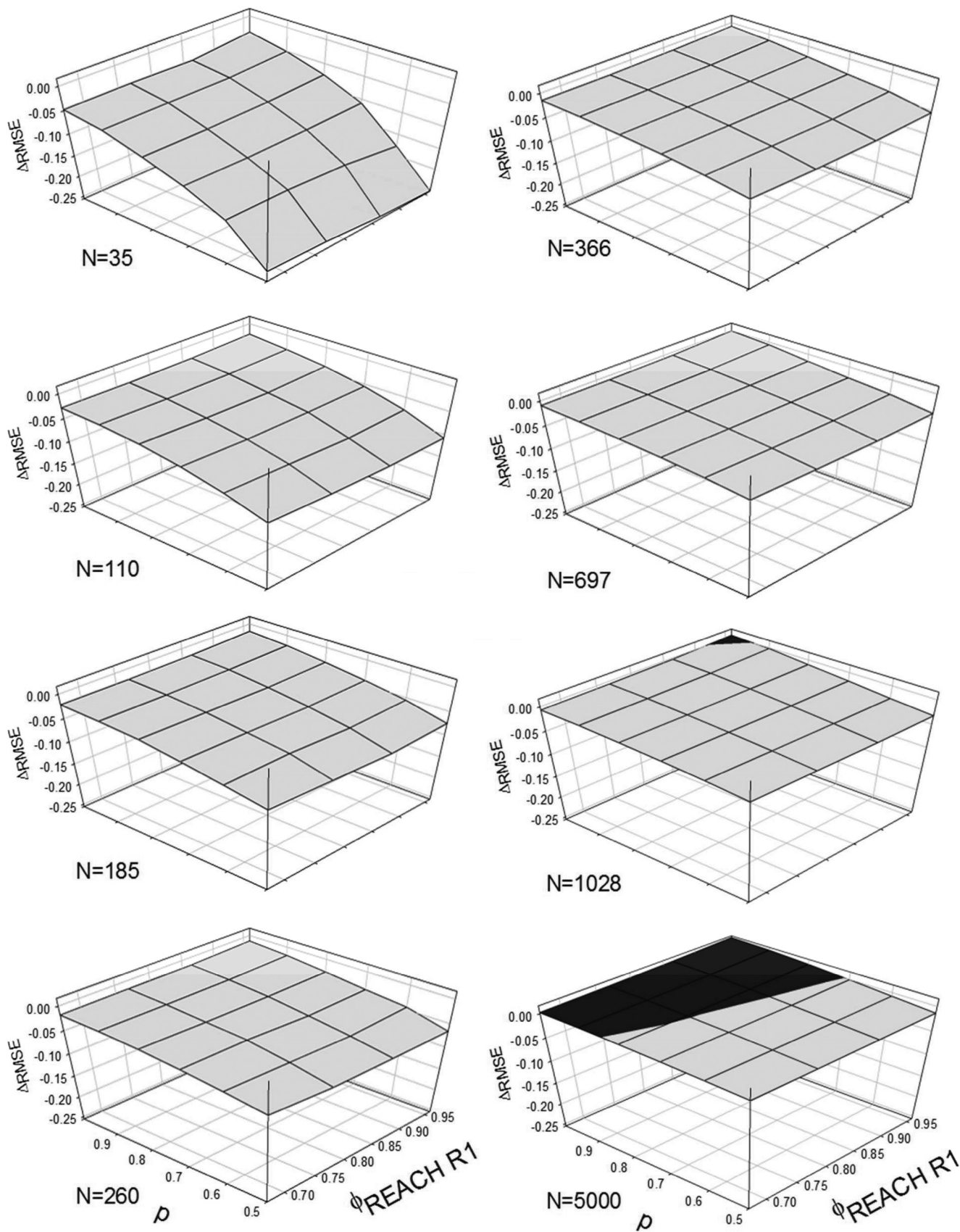
Use of the Bayesian hierarchical modeling resulted in qualitatively equivalent trends in SE, $\text{Bias}_{\text{MEAS}}$, and RMSE when compared with MLE results (Fig. 5). Both modeling approaches demonstrated consistently higher SE of Φ_{DAM} estimates for the paired-release versus the single-release approach at equal sample sizes. It is notable that while the mean of SE from paired-release models was similar between MLE and Bayesian approaches, the variability in estimated SE resulting from the use of Bayesian methods was substantially reduced compared with MLE (Fig. 5). This suggests that SE of parameters from the paired-release model were better estimated using the Bayesian approach at sample sizes of less than about 1000 fish.

This pattern in SE between modeling approaches was not evident in single-release models (Fig. 5). The median values of $\text{Bias}_{\text{MEAS}}$ were negative for the single release using both MLE and Bayesian estimation. We note that in contrast with MLE estimates, $\text{Bias}_{\text{MEAS}}$ for single-release models using a Bayesian approach was centered on the theoretical value of $\text{Bias}_{\text{BKGRND}}$ even at the smallest sample sizes. This indicates that the parameters of single-release models were estimated more accurately using Bayesian methods, likely due to issues with MLE convergence at small N . When compared with MLE, the Bayesian hierarchical modeling indicated a slightly larger positive median value of $\text{Bias}_{\text{MEAS}}$ for the paired-release models that approached zero at large N , but was still present even at the largest sample size used in this study (5000 fish). This is consistent with our expectations of $\text{Bias}_{\text{PRCS}}$ given that $\hat{\Phi}_{\text{DAM}}$ from paired-release models is a ratio of random variables and that ratio can never be less than zero (i.e., long-term, long-run mean values of $\text{Bias}_{\text{PRCS}}$ must be positive, no matter how small). When using RMSE to compare estimates from both paired- and single-release strategies with Bayesian hierarchical modeling, the paired release is advantageous only at very high sample sizes ($N > 1028$) in the base model.

Discussion

Bias in an estimate of survival is undesirable. In this exercise we purposefully imposed this undesirable component of error as river-associated mortality and confirmed that bias is present in estimates of survival through reaches containing dams in a single-release strategy. Lack of precision is likewise undesirable in estimates of

Fig. 7. Surface plots of the difference in root mean squared error (RMSE) between a paired-release strategy and a single-release strategy with variable probability of detection and $\Phi_{\text{REACH R1}}$. Each plot presents data for total number of fish released. Grey indicates conditions where a single release is advantageous over a paired release, while a black surface indicates where a paired release is advantageous.



survival. The inflation of SE due to both reduced numbers for the individual release groups above and below a dam and the inflation due to the approximation of SE using the Delta method were evident in our paired-release simulations. This simulation exercise is not needed to demonstrate the existence of this trade-off, but rather to quantify this balance and to understand where the trade-off might occur for a given species with specific study conditions.

By using RMSE as criteria for evaluation in our base model scenario, we demonstrated that a paired release is generally not advantageous at release sizes less than 1000. This result is consistent whether using MLE or Bayesian estimation. While it may seem intuitive to prioritize bias (accuracy) over error (precision), it is important to acknowledge the difference in fundamental objectives associated with running a simulation versus calculating a single estimate. In conducting this heuristic exercise, we sought to characterize the bias (central tendency of the distribution of deviances) and error (central tendency of the distribution of errors) by repeating a “study” thousands of times with known true values. Thus, the utility is in understanding the mechanics of the process and related implications. In the real world, where a single estimate is produced in a study, how far one has missed the mark is more important than “why”. RMSE is a standard tool that integrates bias and error to evaluate competing models. We have used a range of reasonable scenarios and demonstrated the drastic reduction in precision that can result from an effort to correct a small bias in estimates of true survival. At modest sample sizes, the paired-release strategy has the potential to be less accurate than a single-release strategy in estimating true dam survival.

This is a notable finding, as evaluations of survival for downstream-migrating salmon smolts generally use fewer than 500 fish in a release and most use fewer than 1000 (Fig. 1; e.g., Skalski et al. 2002; Holbrook et al. 2011; NMFS 2012). Investigators rarely report release numbers approaching 5000 (e.g., Skalski et al. 2010). Based on our simulations, only at the highest release sizes (1000–5000) and the highest probabilities of detection (more than 0.70) is a paired release advantageous if bias and precision are valued equally. In those cases, the relative advantage is quite limited in magnitude. These results demonstrate the data-hungry nature of mark-recapture models, even for a simple scenario. It is important to linger on the reality that researchers are infrequently addressing such a simple system as Model River, and path choice (multiple states) can be an integral part of models producing a de facto reduction in N where individual survival estimates are produced.

Maximizing detection efficiency is understood as a logistical goal of mark-recapture methods, but our simulations make clear the pivotal role of probability of detection on patterns of bias and of precision. We note that investigators typically strive for and often achieve high probabilities of detection (>0.90) as reflected in our base model. If, however, p is low, $\text{Bias}_{\text{MEAS}}$ of $\hat{\Phi}_{\text{DAM}}$ for a single-release strategy is reduced in MLE approaches. This is because of estimation error (as indicated by increased SE, particularly at low release numbers). Increasing p (or increasing N) reduces estimation error, as does use of Bayesian estimation. These solutions expose the underlying negative values of $\text{Bias}_{\text{MEAS}}$ of the single-release strategy (i.e., true $\text{Bias}_{\text{BKGRND}}$ becomes evident). This $\text{Bias}_{\text{BKGRND}}$ has a theoretical maximum, approaching the influence of a free-flowing reach at high N and high probabilities of detection (Fig. 4A). Using appropriate sample sizes and with reasonably high detection probabilities, the central tendency of $\text{Bias}_{\text{MEAS}}$ from single-release models will always approximate that maximum.

Though single-release methods have generally been considered inappropriate to estimate $\hat{\Phi}_{\text{DAM}}$, both single-release and paired-release methods produce unbiased estimators of $\hat{\Phi}_{\text{DAM}}$ when no background mortality is present (i.e., $\hat{\Phi}_{\text{DAM}} = \hat{\Phi}_{\text{REACH}}$ when

$\hat{\Phi}_{\text{BKGRND}} = 1.0$). Thus, when mortality is low, the single-release method is clearly advantageous (with respect to minimizing RMSE) because the SE of $\hat{\Phi}_{\text{REACH}}$ (from a single release) will always be smaller than the SE of $\hat{\Phi}_{\text{DAM}}$ (from a paired release) for a given sample size. Therefore, any advantage of paired-release methods is directly related to background mortality. $\text{Bias}_{\text{BKGRND}}$ is generally quite low in the “real world” where in-river mortality is only $0.003\text{--}0.07\text{-km}^{-1}$ as a global mean for Atlantic salmon (Thorstad et al. 2012).

For a paired release, the influence of detection probability and N on precision was as we expected. Precision was lower when using paired releases than in a single-release strategy because of lower N in each release group and because of inflation associated with the propagation of error through the Delta method. The observed trends are the same as for a single release with increasing sample size. A notable result from this work, however, was the high bias imposed by a paired release resulting in an overestimation of $\hat{\Phi}_{\text{DAM}}$. This is most evident at small sample sizes or when probability of detection is low. This potentially counterintuitive result is due to the asymmetrical distribution of survival estimates for the release below the dam and the derivation of $\hat{\Phi}_{\text{DAM}}$. To paraphrase the Mad Hatter from *Alice's Adventures in Wonderland* (Carroll 1865), with regard to the probability of mortality “one can have more than some, but one cannot have less than none.”

Because the background survival estimate is the denominator in the paired-release calculation of $\hat{\Phi}_{\text{DAM}}$ (eq. 2) at small sample sizes, stochastic events may result in a high (sometimes very high) positive error in $\hat{\Phi}_{\text{DAM}}$. Not only is this more clear at relatively low release numbers (hundreds of fish), but the magnitude of this $\text{Bias}_{\text{PRCS}}$ can be notably higher than the $\text{Bias}_{\text{BKGRND}}$ being corrected for in the single release. As such, the possibility exists for researchers to do more harm than good in terms of study design when sample sizes and detection probability are not adequate for the use of a paired-release approach. Similarly, at low N for a single release, the systematic $\text{Bias}_{\text{BKGRND}}$ due to river mortality is offset by the inherent variability in $\text{Bias}_{\text{MEAS}}$ associated with small samples sizes. Thus, for fewer than several hundred fish, if probability of detection is less than perfect ($p < 1.0$), there is less $\text{Bias}_{\text{BKGRND}}$ to correct for in a single release and more (sometimes much more) $\text{Bias}_{\text{PRCS}}$ imposed in a paired release if probability of detection is poor.

When using the logit scale and survival estimates are high, one might suggest that abnormal behavior in MLE near the boundaries of probability scale (e.g., 1.0) might be causal to this observation. Our use of a Bayesian hierarchical modeling approach confirms a positive $\text{Bias}_{\text{PRCS}}$ associated with the paired-release approach to estimate $\hat{\Phi}_{\text{DAM}}$ independent of MLE or link functions used for estimation. $\text{Bias}_{\text{PRCS}}$ is ameliorated by increasing sample sizes, but was never completely eliminated, even at the largest N used in our study.

Understanding the biases imposed by each of the release strategies examined here is important to both those who manage and those who evaluate survival through reaches of interest. For a single release, there is a theoretical negative bias ($\text{Bias}_{\text{BKGRND}}$) that has the potential to result in underestimation of $\hat{\Phi}_{\text{DAM}}$. This bias increases as both N and p increase, reaching a maximum equivalent to in-river mortality. In contrast, for a paired release, there is a systematic, positive bias ($\text{Bias}_{\text{PRCS}}$) that risks overestimation of $\hat{\Phi}_{\text{DAM}}$. For the paired release, this bias is lessened as N and p increase. Depending on one's perspective, there are different risks to each approach. Overestimating survival at a dam is just as distasteful for the manager trying to meet conservation objectives as is underestimating survival for the hydro-operator trying to meet regulatory criteria. Neither is desirable from an objective stance, and in either case, awareness is useful in guiding decisions. Simply put, knowing the direction of a bias can be important in effectively interpreting study results.

The Model River construct used for this study was exceptionally simple, by design, and presents the best conditions in assessing the theoretical trade-offs of precision and bias. For mark-recapture models, there are many assumptions that are adopted, and these assumptions have been effectively presented and discussed (e.g., Burnham et al. 1987). These general assumptions include (1) independence of individuals, (2) marked individuals are representative of population, (3) equal probability of detection, (4) absence of a tagging effect, and (5) correct interpretation of each encounter as a live fish. The paired-release strategy imposes additional assumptions, including that the distance between the R1 and the R2 release sites is negligible in comparison with the length of the interval of interest. Recall that while the second group (R2) is released below the dam, these fish are mathematically treated as if they were released above the dam (to estimate Φ_{BKGRND} for the appropriate length of river). Thus, failure to meet this assumption results in underestimation of $\Phi_{\text{REACH R2}}$ (used as the estimate Φ_{BKGRND}), resulting in overestimation of Φ_{DAM} . Meeting this assumption can be logistically difficult in places where dams are in close proximity to one another.

An added challenge is the assumption that paired-release groups experience the same survival probability in the interval of interest. For those released below the dam, variability in handling and tagging may affect survival through this interval of downstream migration. The smolt developmental stage has a heightened sensitivity to stress (Carey and McCormick 1998) and postrelease performance can vary. Often the first interval after release can have low apparent survival (Holbrook et al. 2011). If this is the case, Φ_{DAM} would also be underestimated based on the erroneous inclusion of tagging effects. Therefore, failure to meet assumptions can impose further positive bias, resulting in an overestimation of Φ_{DAM} . It is for this reason that a triple-release design has been implemented in some large studies (e.g., Skalski et al. 2009, 2013). We note, however, that for this simulation study we assumed our fish had not been influenced by tagging or handling, and therefore this study design could not have provided an advantage.

Managers and researchers face challenging decisions about how to best use limited resources to generate the most useful data. Radio, acoustic, and passive integrated transponder (PIT) telemetry are effective telemetry tools, but they are also expensive. Experimental designs that maximize mathematical efficacy can provide the optimal decision-making process for the resources available. The work here highlights several steps researchers and managers might take towards this goal.

- (1) Given limited resources, it is important to minimize model complexities and recognize that these models require sufficient numbers to reduce error to generate useful estimates of survival. In this study, models using 35 fish demonstrated poor model convergence and high RMSE, representing a clearly unacceptable sample size. Managers wishing to produce meaningful estimates of survival must match resources to increase sample size commensurate with a priori precision requirements and anticipated survival and detection probabilities.
- (2) At release sizes historically used to study smolts (hundreds or fewer fish per group), a single-release strategy is preferable to a paired-release strategy when precision and bias are penalized equally. This is because the risk of overestimating Φ_{DAM} with a paired-release design is greater than the risk of underestimating Φ_{DAM} using a single release (Figs. 5C and 5E).
- (3) The use of paired release is mathematically beneficial under limited and specific conditions (high detection) and sample sizes that often cannot be obtained even for studies of culturally important or critically endangered species owing to financial costs or scarcity of study specimens. Additionally, attention to systematic sources of positive bias in estimates of

Φ_{DAM} (both Bias_{PRCS} and failure to meet assumptions) is warranted to avoid underestimating the impact of dam-related mortality.

- (4) Regardless of release strategy used (single or paired), the tag type used, or the number of fish tagged, increasing probability of detection is exceptionally important in reducing the error associated with survival estimates. In some cases, adding additional receivers to telemetry networks may be far more cost-effective than increasing tag number to improve precision.
- (5) In a single-release strategy, identifying the reach of interest and minimizing the length being assessed is critical. This approach will minimize the inherent bias associated with this strategy (Bias_{BKGRND}). This must be accomplished while encompassing the entire reach of impact and being mindful of river conditions (e.g., minimizing the reach length too much could create positive bias if dead fish pass downstream receivers).
- (6) Managers wishing to describe survival thresholds at dams can identify an acceptable mortality rate through the reach of interest with an a priori acceptance that some component of mortality is attributable to natural attrition. Such an approach would obviate the need for a paired-release strategy that is far more expensive than a single release if producing equivalent estimates of equivalent precision.

Though these simulations provide guidance for mark-recapture modeling, decisions can be vastly different based upon a number of factors including available resources. Though this work is based on assessing survival for salmonid smolts, our findings and our cautions are equally important for the survival assessment of other species of interest. There are many confounding issues and objectives involved in assessment of dam passage, including the biological validity of some of the assumptions of these models. However, the expressed desire to estimate path-specific mortality rates remains a growing objective based on institutional mandates of legal compliance and conservation. These multistate models are even more data-hungry and also invite other methods such as MCMC approaches (as applied here for comparison). For conservation, it is also notable that biological impacts of dams are not limited to discrete, well-defined reaches. The influence of dams can certainly be acute (Holbrook et al. 2011; Stich et al. 2014), but dam passage can result in exposure to elevated predation (Blackwell et al. 1998) and result in nonlethal physical injuries (Stier and Kynard 1986; Mathur et al. 1996). Such experience can affect performance (Zydlewski et al. 2010) and survival in the estuary (Handeland et al. 1996; Halfyard et al. 2013; Stich et al. 2015b) and ocean (Budy et al. 2002; Schaller and Petrosky 2007; Haeseker et al. 2012). These factors contribute to delayed mortality, which remains a critical concern in conservation efforts and remains largely under-studied.

Acknowledgements

Financial support was provided by the University of Maine, and the US Geological Survey, Maine Cooperative Fisheries and Wildlife Research Unit, and the Department of Marine Resources. Mention of trade names or commercial products does not imply endorsement by the US Government. This manuscript benefitted from the reviews and quantitative expertise of Christopher Holbrook, Theodore Castro-Santos, Erik Blomberg, William Halteman, and several anonymous reviewers. The technical assistance provided by George Maynard is greatly appreciated.

References

- Blackwell, B.F., and Juanes, F. 1998. Predation on Atlantic salmon smolts by striped bass after dam passage. *N. Am. J. Fish. Manage.* 18(4): 936–939. doi: 10.1577/1548-8675(1998)018<0936:POASSB>2.0.CO;2.
- Budy, P., Thiede, G.P., Bouwes, N., Petrosky, C.E., and Schaller, H. 2002. Evidence Linking Delayed Mortality of Snake River Salmon to Their Earlier Hydrosystem

- Experience. N. Am. J. Fish. Manage. 22: 35–51. doi:10.1577/1548-8675(2002)022<0035:ELDMOS>2.0.CO;2.
- Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C., and Pollock, K.H. 1987. Design and analysis methods for fish survival experiments based on release–recapture. American Fisheries Society, Monograph 5, Bethesda, Maryland.
- Carey, J.B., and McCormick, S.D. 1998. Atlantic Salmon smolts are more responsive to an acute handling and confinement stress than parr. Aquaculture, 168: 237–253. doi:10.1016/S0044-8486(98)00352-4.
- Carroll, L. 1865. Alice's Adventures in Wonderland. MacMillan, New York.
- Castro-Santos, T., and Haro, A. 2013. Survival and behavioral effects of exposure to a hydrokinetic turbine on juvenile Atlantic salmon and adult American shad. Estuaries and Coasts, 38 (S1): S203–S214. doi:10.1007/s12237-013-9680-6.
- Clements, S., Stahl, T., and Schreck, C.B. 2012. A comparison of the behavior and survival of juvenile coho salmon (*Oncorhynchus kisutch*) and steelhead trout (*O. mykiss*) in a small estuary system. Aquacult. 362: 148–157. doi:10.1016/j.aquaculture.2011.11.029.
- Cormack, R.M. 1964. Estimates of survival from the sighting of marked animals. Biometrika, 51(3–4): 429–438. doi:10.1093/biomet/51.3-4.429.
- Haeseker, S.L., McCann, J.A., Tuomikoski, J., and Chockley, B. 2012. Assessing freshwater and marine environmental influences on life-stage-specific survival rates of Snake River spring–summer Chinook salmon and steelhead. Trans. Am. Fish. Soc. 141: 121–138. doi:10.1080/00028487.2011.652009.
- Halfyard, E.A., Gibson, A.J.F., Stokesbury, M.J.W., Ruzzante, D.E., and Whoriskey, F.G. 2013. Correlates of estuarine survival of Atlantic Salmon postsmolts from the Southern Upland, Nova Scotia, Canada. Can. J. Fish. Aquat. Sci. 70(3): 452–460. doi:10.1139/cjfas-2012-0287.
- Handeland, S.O., Järvi, T., Fernö, A., and Stefansson, S.O. 1996. Osmotic stress, antipredatory behaviour, and mortality of Atlantic Salmon (*Salmo salar*) smolts. Can. J. Fish. Aquat. Sci. 53(12): 2673–2680. doi:10.1139/f96-227.
- Heupel, M.R., Semmens, J.M., and Hobday, A.J. 2006. Automated acoustic tracking of aquatic animals: scales, design, and deployment of listening station arrays. Mar. Fresh. Res. 57: 1–13. doi:10.1071/MF05091.
- Holbrook, C.M., Kinnison, M.T., and Zydlowski, J. 2011. Survival of migrating Atlantic Salmon smolts through the Penobscot River, Maine: a pre-restoration assessment. Trans. Am. Fish. Soc. 140: 1255–1268. doi:10.1080/00028487.2011.618356.
- Horton, G.E., Letcher, B.H., Bailey, M.M., and Kinnison, M.T. 2009. Atlantic salmon (*Salmo salar*) smolt production: the relative importance of survival and body growth. Can. J. Fish. Aquat. Sci. 66(3): 471–483. doi:10.1139/F09-005.
- Johnson, S.L., Power, J.H., Wilson, D.R., and Ray, J. 2010. A comparison of the survival and migratory behavior of hatchery-reared and naturally reared steelhead smolts in the Alsea River and estuary, Oregon, using acoustic telemetry. N. Am. J. Fish. Manage. 30(1): 55–71. doi:10.1577/M08-224.1.
- Karppinen, P., Jounela, P., Huusko, R., and Erkinaro, J. 2014. Effects of release timing on migration behaviour and survival of hatchery-reared Atlantic salmon smolts in a regulated river. Ecol. Fresh. Fish, 23(3): 438–452. doi:10.1111/eff.12097.
- Kéry, M., and Schaub, M. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. Academic Press.
- Laake, J., and Rexstad, E. 2009. RMark: an alternative approach to building linear models in MARK. In Program MARK: a gentle introduction — Appendix C. Edited by E. Cooch and G. White.
- Lady, J.M., and Skalski, J.R. 2009. USER 4 — User specified estimation routine. University of Washington, Seattle.
- Mathur, D., Heisey, P.G., Euston, E.T., Skalski, J.R., and Hays, S. 1996. Turbine passage survival estimation for chinook salmon smolts (*Oncorhynchus tshawytscha*) at a large dam on the Columbia River. Can. J. Fish. Aquat. Sci. 53(3): 542–549. doi:10.1139/f95-206.
- NMFS. 2012. Endangered Species Act biological opinion. National Marine Fisheries Service. F/NER/2012/01568.
- Norrgård, J.R., Greenberg, L.A., Piccolo, J.J., Schmitz, M., and Bergman, E. 2013. Multiplicative loss of landlocked Atlantic salmon *Salmo salar* L. smolts during downstream migration through multiple dams. Riv. Res. Applic. 29: 1306–1317. doi:10.1002/rra.2616.
- Perry, R.W., Castro-Santos, T., Holbrook, C.M., and Sandford, B.P. 2012. Using Mark–Recapture Models to Estimate Survival from Telemetry Data. Telemetry techniques: a user guide for fisheries research. American Fisheries Society, Bethesda, Maryland, pp. 453–475.
- R Core Team. 2014. R: a language and environment for statistical computing, version 3.1.0. R Foundation for Statistical Computing, Vienna, Austria.
- Ross, T. 1996. Indices for performance evaluation of predictive models in food microbiology. J. App. Bact. 81: 501–508. doi:10.1111/j.1365-2672.1996.tb03539.x.
- Schaller, H.A., and Petrosky, C.E. 2007. Assessing hydrosystem influence on delayed mortality of Snake River stream-type Chinook salmon. N. Am. J. Fish. Manage. 27(3): 810–824. doi:10.1577/M06-083.1.
- Seber, G.A.F. 1982. The estimation of animal abundance. 2nd edition. Macmillan, New York.
- Serrano, I., Rivinoja, P., Karlsson, L., and Larsson, S. 2009. Riverine and early marine survival of stocked salmon smolts, *Salmo salar* L., descending the Testebo River, Sweden. Fish. Manage. Ecol. 16(5): 386–394. doi:10.1111/j.1365-2400.2009.00688.x.
- Skalski, J.R., Lady, J., Townsend, R., Giorgi, A.E., Stevenson, J.R., Peven, C.M., and McDonald, R.D. 2001. Estimating in-river survival of migrating salmonid smolts using radiotelemetry. Can. J. Fish. Aquat. Sci. 58(10): 1987–1997. doi:10.1139/f01-133.
- Skalski, J.R., Townsend, R., Lady, J., Giorgi, A.E., Stevenson, J.R., and McDonald, R.D. 2002. Estimating route-specific passage and survival probabilities at a hydroelectric project from smolt radiotelemetry studies. Can. J. Fish. Aquat. Sci. 59(8): 1385–1393. doi:10.1139/f02-094.
- Skalski, J.R., Buchanan, R.A., Townsend, R.L., Steig, T.W., and Hemstrom, M. 2009. A multiple-release model to estimate route-specific and dam passage survival at a hydroelectric project. N. Am. J. Fish. Manage. 29: 670–679. doi:10.1577/M07-227.1.
- Skalski, J.R., Townsend, R.L., Steig, T.W., and Hemstrom, S. 2010. Comparison of two alternative approaches for estimating dam passage survival of salmon smolts. N. Am. J. Fish. Manage. 30(3): 831–839. doi:10.1577/M09-103.1.
- Skalski, J.R., Seaburg, A.G., and Buchanan, R.A. 2013. The effects of high detection probabilities on model selection in paired release–recapture studies in the era of electronic tagging studies. Anim. Biotelemetry, 1: 12. doi:10.1186/2050-3385-1-12.
- Skalski, J.R., Eppard, M.B., Ploskey, G.R., Weiland, M.A., Carlson, T.J., and Townsend, R.L. 2014. Assessment of subyearling Chinook salmon survival through the federal hydropower projects in the main-stem Columbia River. N. Am. J. Fish. Manage. 34(4): 741–752. doi:10.1080/02755947.2014.910577.
- Smith, D.M., Cusack, S., Colman, A.W., Folland, C.K., Harris, G.R., and Murphy, J.M. 2007. Improved surface temperature prediction for the coming decade from a global climate model. Science, 317: 796–799. doi:10.1126/science.1139540. PMID:17690292.
- Stich, D.S., Bailey, M.M., and Zydlowski, J.D. 2014. Survival of Atlantic salmon *Salmo salar* smolts through a hydropower complex in the lower Penobscot River, Maine U.S.A. J. Fish Biol. 85(4): 1074–1096. doi:10.1111/jfb.12483. PMID:25130320.
- Stich, D.S., Bailey, M.M., Holbrook, C., Kinnison, and Zydlowski, J. 2015a. Catchment-wide survival of wild- and hatchery-reared Atlantic salmon smolts in a changing system. Can. J. Fish. Aquat. Sci. 72: 1352–1365. doi:10.1139/cjfas-2014-0573.
- Stich, D.S., Kocik, J.F., Zydlowski, G.B., and Zydlowski, J.D. 2015b. Linking behavior, physiology, and survival of Atlantic salmon smolts during estuary migration. Mar. Coast. Fish. doi:10.1080/19425120.2015.1007185.
- Stier, D.J., and Kynard, B. 1986. Use of radio telemetry to determine the mortality of Atlantic salmon smolts passed through a 17-MW Kaplan turbine at a low-head hydroelectric dam. Trans. Am. Fish. Soc. 115(5): 771–775.
- Thorstad, E.B., Whoriskey, F., Uglem, I., Moore, A., Rikardsen, A.H., and Finstad, B. 2012. A critical life stage of the Atlantic Salmon *Salmo salar*: behaviour and survival during the smolt and initial post-smolt migration. J. Fish Biol. 81: 500–542. doi:10.1111/j.1095-8649.2012.03370.x. PMID:22803722.
- Van den Avyle, M.J., and Hayward, R.S. 1999. Dynamics of exploited fish populations. In Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland, pp. 127–166.
- White, G.C., and Burnham, K.P. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study, 46(S1): S120–S139. doi:10.1080/00063659909477239.
- Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., and Rowe, C.M. 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res.: Oceans, 90: 8995–9005. doi:10.1029/JC090iC05p08995.
- Zydlowski, J., Zydlowski, G., and Danner, R. 2010. Descaling impairs the osmoregulatory ability of Atlantic salmon smolts. Trans. Am. Fish. Soc. 138: 129–136. doi:10.1577/T09-054.1.