Animal Biotelemetry

The effects of high detection probabilities on model selection in paired release-recapture studies in the era of electronic tagging studies

John R Skalski[*†], Adam G Seaburg[†] and Rebecca A Buchanan[†]

Abstract

Background: Acoustic-tag studies with their high to very high detection rates defy traditional statistical wisdom regarding analysis of tagging studies. Conventional wisdom has been to use a parsimonious model with the fewest parameters that adequately describes the data to estimate survival parameters in release-recapture studies in order to find a reasonable trade-off between precision and accuracy. This quest has generated considerable debate in the statistical community on how to best accomplish this task. Among the debated options are likelihood ratio tests, Bayesian information criterion, Akaike information criterion, and model averaging.

Results: Our Monte Carlo simulation studies of paired release-recapture, acoustic-tag investigations indicate precision is the same if a fully parameterized or a reduced parameter model is used for data analysis if detection probabilities are very high. In addition, the fully parameterized model is robust to heterogeneous survival and detection processes, while a reduced parameter model may be sensitive to misspecification.

Conclusions: Use fully parameterized, paired release-recapture models when detection probabilities are very high (≥0.90) to analyze acoustic-tagging data in order to retain both robustness and precision, and without the subjectivity and ambiguity introduced by the choice and application of model selection techniques.

Keywords: Tagging studies, Acoustic tags, Radio tags, Parsimony, Model selection

Background

In just the last two decades, fish tagging studies have seen major advances with the advent of satellite tags [1], biotelemetry tags [2], radio tags [3], and acoustic tags [4]. These active-tag technologies have improved the realism, accuracy, and precision of tagging studies and changed the ways tagging studies are analyzed [5,6]. In particular, the often high detection probabilities and detailed event data with active-tag technologies both enable and demand new data analysis methods, including such approaches as individual-based models [7], multistate models [8-10], and random-walk [11] models. These methods often use a large number of parameters to differentiate biological processes such as survival and movement from the detection process in order to provide absolute measures of biological parameters.

A general statistical principle is that there is a trade-off between the number of parameters estimated from a fixed set of data and the precision of the parameter estimates: the fewer parameters to be estimated, the more precise the estimates will be. This has been an important concept underlying analysis of fish tagging studies using traditional passive-tag technologies, as their typically low detection probabilities affect not only sampling precision but also estimability of parameters and model differentiation. In the landmark 1987 publication on the design and analysis of paired release-recapture studies to estimate fish survival, Burnham and colleagues [12], pp. 287–288 advocated parsimonious model selection as a means to improve sampling precision of fish tagging studies. Paired release-recapture studies are used to estimate fish survival between release locations, and have been used to estimate such measures as turbine passage mortality [13] and project passage survival at hydroprojects [14]. The principle of parsimony implies that, among models that satisfactorily fit the data, the model with the fewest parameters is

* Correspondence: skalski@uw.edu
[†]Equal contributors
School of Aquatic & Fishery Sciences, University of Washington, 1325 Fourth Avenue, Suite 1820, Seattle, WA, USA

preferred. Model parsimony has been seen as a means to counter low precision as a consequence of small sample sizes and low detection probabilities. Initially, likelihood ratio tests [15], then the Akaike information criterion [16,17], and, most recently, model averaging [17] have been recommended to find the best descriptions of tagging data.

The trade-off between the number of parameters and sampling precision translates into a trade-off between robustness of the estimation process and sampling precision. A wrongly identified and oversimplified model may introduce estimation bias in attempts to improve sampling precision, while an overparameterized model may have adequate accuracy but low sampling precision. Thus, the traditional search for parsimony raises the risk of model misspecification and bias. While these may be necessary risks in the case of low detection probabilities and sparse data, the case may be very different with the high detection probabilities available with new tag technologies.

In this paper, we review the trade-offs between precision and robustness of survival estimates in acoustic-tag investigations with typically high detection probabilities. We performed Monte Carlo simulations based on paired release-recapture investigations [12]. We examined survival estimates when downstream survival and detection probabilities are both correctly and incorrectly modeled.

Results
Simulations under H_o: homogeneous downstream parameters
Our simulation studies indicated that a paired-release estimate of survival S_0 was positively biased when the detection probabilities (P) were ≤0.30 and release sizes were as high as $R_1 = R_2 = 1,000$. This small sample bias results both from the fact that the maximum likelihood estimates of reach survival are only asymptotically unbiased [18], and from the fact that the estimate of S_0 is a ratio of random variables, which introduces an additional bias term of the form

$$E(\hat{S}_0) = E(\hat{S}_1) \cdot E\left(\frac{1}{\hat{S}_2}\right)$$
$$\doteq E(\hat{S}_1) \cdot \frac{1}{E(\hat{S}_2)}\left(1 + \frac{\mathrm{Var}(\hat{S}_2|S_2)}{E(\hat{S}_2)^2}\right),$$

based on a third-term Taylor series expansion [19]. Bias tended to increase as the release-recapture model became more parameterized at low detection probabilities (Table 1a). This bias coincided with the inability to find the maximum likelihood estimates for one or more of the eight alternative likelihood models under the null hypothesis (H_o) of downstream homogeneity of survival and detection processes (Table 1a). When model convergence was not an issue, neither was estimation bias. The

simulation studies also indicated the estimates for the standard errors of $\hat{S}_0$ $\left(\text{that is, } \widehat{\mathrm{Var}\,(\hat{S}_0|S_0)}\right)$ were also unbiased as long as model convergence was not a problem (that is, $P \geq 0.4$) (Table 1a versus Table 1c). When biased, the standard error from the inverse Hessian matrix tends to be overestimated, producing conservative (in other words, too wide) interval estimates.

At very high detection probabilities and release sizes $R_1 = R_2 = 1,000$, the standard errors had the same or nearly the same value across the eight alternative, paired release-recapture models under H_o (see Table 1b, $P \geq 0.90$). This suggested little or no precision advantages in using a parsimonious likelihood model under these circumstances. Precision gains from using valid models with reduced parameterization were the greatest when detection probabilities were the lowest ($0.1 \leq P \leq 0.3$) and where model convergence for some of the models became problematic (Table 1). For detection probabilities $0.4 \leq P \leq 0.7$, precision gains occurred only from the use of the simplest models (for example, M_6 and M_7).

Additional simulations were performed when release sizes were $R_1 = R_2 = 250$ or 500 for a $k = 2$- or 4-period study (Table 2). Results were similar, with little or no precision gain when detection probabilities were ≥0.90. At release sizes $R_1 = R_2 = 250$, precision gains became evident when detection probabilities were ≤0.80. For smaller, qualitative, paired release-recapture investigations, model parsimony may improve precision at the expense of robustness. Investigators should evaluate the properties of survival estimates when conducting studies outside of the range of simulations illustrated in this paper.

In general, our simulation results in the case of homogeneous downstream parameters indicate high detection probabilities (that is, $P \geq 0.90$) negated the potential precision advantages of finding a parsimonious paired release-recapture model. The fully parameterized model had the same precision as simplified models with fewer estimated parameters.

Simulations under H_a: heterogeneous downstream parameters
The previous section demonstrated that there were no precision gains from using a reduced paired release-recapture model when detection probabilities are very high (that is, ≥0.90), and downstream survival and detection processes are homogeneous. In this section, we examine the consequences on estimation bias of employing a reduced model when heterogeneity in survival or detection processes is misspecified.

$S_{12} \neq S_{22}$
In this set of simulations, the first shared reach below the paired releases had survival parameters that differed

Table 1 Simulation results under homogeneity

a. $\hat{\bar{S}}_0$

P	Failure rate	Model						
		M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.10	0.512	0.9629	0.9645	0.9637	0.9578	0.9554	0.9429	0.9414
0.20	0.279	0.9525	0.9523	0.9527	0.9513	0.9509	0.9376	0.9359
0.30	0.082	0.9385	0.9387	0.9389	0.9388	0.9386	0.9330	0.9327
0.40	0.010	0.9323	0.9324	0.9324	0.9324	0.9323	0.9310	0.9308
0.50	0.0	0.9313	0.9313	0.9313	0.9313	0.9313	0.9313	0.9313
0.60	0.0	0.9301	0.9301	0.9301	0.9301	0.9301	0.9302	0.9302
0.70	0.0	0.9305	0.9305	0.9305	0.9305	0.9305	0.9306	0.9305
0.80	0.0	0.9303	0.9303	0.9303	0.9303	0.9303	0.9304	0.9304
0.90	0.0	0.9302	0.9302	0.9302	0.9302	0.9302	0.9302	0.9302
0.95	0.0	0.9302	0.9302	0.9302	0.9302	0.9302	0.9302	0.9302
0.98	0.0	0.9301	0.9301	0.9301	0.9301	0.9301	0.9301	0.9301

b. $\widehat{SE}\left(\hat{S}_0 | S_0\right)$

P	Model						
	M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.10	0.2055	0.1971	0.1930	0.1654	0.1568	0.1013	0.0905
0.20	0.0970	0.0963	0.0958	0.0903	0.0885	0.0602	0.0553
0.30	0.0575	0.0576	0.0575	0.0562	0.0558	0.0415	0.0391
0.40	0.0383	0.0382	0.0382	0.0381	0.0380	0.0308	0.0296
0.50	0.0273	0.0273	0.0273	0.0273	0.0273	0.0240	0.0234
0.60	0.0210	0.0210	0.0210	0.0210	0.0210	0.0197	0.0194
0.70	0.0176	0.0176	0.0176	0.0176	0.0176	0.0171	0.0170
0.80	0.0158	0.0158	0.0158	0.0158	0.0158	0.0156	0.0156
0.90	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147
0.95	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.98	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0142

c. $s_{\hat{S}_0}$

P	Model						
	M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.10	0.1402	0.1338	0.1326	0.1175	0.1149	0.0750	0.0664
0.20	0.0750	0.0743	0.0741	0.0730	0.0730	0.0450	0.0414
0.30	0.0544	0.0545	0.0547	0.0548	0.0550	0.0344	0.0317
0.40	0.0379	0.0380	0.0380	0.0384	0.0385	0.0263	0.0250
0.50	0.0270	0.0270	0.0270	0.0271	0.0271	0.0219	0.0212
0.60	0.0210	0.0210	0.0210	0.0210	0.0210	0.0185	0.0181
0.70	0.0179	0.0179	0.0179	0.0179	0.0179	0.0168	0.0166
0.80	0.0156	0.0156	0.0156	0.0156	0.0156	0.0154	0.0153
0.90	0.0153	0.0153	0.0153	0.0153	0.0153	0.0152	0.0151
0.95	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.98	0.0139	0.0139	0.0139	0.0139	0.0139	0.0138	0.0139

Estimates of (a) paired survival ($\hat{S}_0 = \hat{S}_{11}/\hat{S}_{21}$), (b) average standard error from inverse Hessian (that is, $\widehat{SE}(\hat{S}_0|S_0)$), and (c) empirical standard errors (that is, $S_{\hat{S}_0}$) from Monte Carlo simulations for eight different likelihood models under H_o for varying levels of capture probability ($0.10 \leq P \leq 0.98$). Paired release-recapture survival was simulated at $S_0 = 0.93$. Failure rate for one or more models not converging reported in subtable (a).

Table 2 More simulation results under homogeneity

k (# of reaches)	P	$R_1 = R_2$	Failure rate		$\hat{S}_0$		$\overline{\widehat{SE}(\hat{S}_0|S_0)}$	
			M_{Full}	M_7	M_{Full}	M_7	M_{Full}	M_7
2	0.70	250	0.0180	0.0205	0.9311	0.9303	0.0513	0.0454
	0.80		0.0011	0.0022	0.9308	0.9304	0.0384	0.0365
	0.90		0.0000	0.0000	0.9303	0.9303	0.0311	0.0308
	0.95		0.0000	0.0000	0.9303	0.9303	0.0292	0.0291
	0.98		0.0000	0.0000	0.9302	0.9302	0.0285	0.0285
2	0.70	500	0.0020	0.0038	0.9306	0.9303	0.0360	0.0320
	0.80		0.0000	0.0002	0.9302	0.9300	0.0271	0.0258
	0.90		0.0000	0.0000	0.9301	0.9301	0.0220	0.0218
	0.95		0.0000	0.0000	0.9302	0.9302	0.0207	0.0206
	0.98		0.0000	0.0000	0.9302	0.9302	0.0202	0.0202
4	0.70	250	0.0003	0.0006	0.9303	0.9302	0.0353	0.0340
	0.80		0.0000	0.0002	0.9303	0.9302	0.0316	0.0311
	0.90		0.0000	0.0000	0.9307	0.9307	0.0295	0.0294
	0.95		0.0000	0.0000	0.9302	0.9302	0.0289	0.0288
	0.98		0.0000	0.0000	0.9305	0.9305	0.0285	0.0285
4	0.70	500	0.0000	0.0000	0.9303	0.9304	0.0249	0.0240
	0.80		0.0000	0.0000	0.9301	0.9300	0.0223	0.0220
	0.90		0.0000	0.0000	0.9300	0.9300	0.0209	0.0208
	0.95		0.0000	0.0000	0.9299	0.9299	0.0204	0.0204
	0.98		0.0000	0.0000	0.9304	0.9304	0.0202	0.0201

Failure rates, average value of $\hat{S}_0$, and average standard error from inverse Hessian (that is, $\overline{\widehat{SE}(\hat{S}_0|S_0)}$) for detection probabilities $P = 0.70, 0.80, 0.90, 0.95$ and 0.98, release sizes $R_1 = R_2 = 250$ or 500, and $k = 2$ or 4 periods in a paired release-recapture study.

between release groups (see Figure 1, $S_{12} \neq S_{22}$). All other model parameters were homogeneous across release groups. In this situation, the full model M_{Full} and models M_2 to M_5 were valid, but models M_6 and M_7 were invalid representations of the release-recapture data. When the upstream release R_1 had a higher survival value in S_{12} than the downstream release R_2 had in S_{22}, the estimate of survival from the paired release (equation (1)) was positively biased (Table 3) for the inappropriate models M_6 and M_7. The other models which did not assume homogeneity of this survival parameter remained unbiased. Levels of bias ranged from approximately 0.0009 to 0.0049 when survival differences $|S_{12} - S_{22}|$ ranged from 0.02 to 0.05. Alternatively, when S_{22} had a higher survival value than S_{12}, the estimate of survival from the paired release (equation (1)) was negatively biased for models M_6 and M_7; the other models which did not assume homogeneity of $S_{12} = S_{22}$ remained unbiased (Table 3).

Under both circumstances of $S_{12} > S_{22}$ and $S_{12} < S_{22}$, the precision of the models over the range of detection probabilities simulated (that is, $0.90 \leq P \leq 0.95$) was unaffected by the degree of parameter reduction of the release-recapture model (Table 3). The results of our

simulations suggest the paired release-recapture model was not robust to model violations concerning $S_{12} \neq S_{22}$, and there were no precision gains when a parsimonious model was used and detection probabilities were high. On the other hand, the fully parameterized model retained both the properties of robustness and precision, and should be used in the analyses of acoustic-tag survival studies in these circumstances.

$S_{13} \neq S_{23}$

When heterogeneity in the reach survival probabilities was simulated two reaches below the paired-release locations (that is, $S_{13} \neq S_{23}$), there was no effect on the estimation of S_0 (Table 4). Despite a difference of as much as $|S_{13} - S_{23}| = 0.05$, there was no effect on the unbiasedness of $\hat{S}_0$, even though models M_4 to M_7 had assumption violations (Table 4). Precision was also unaffected, with all eight models having essentially the same $\overline{\widehat{SE}(\hat{S}_0|S_0)}$ regardless of model violations or degree of parameter reduction. This result is a consequence of only consecutive reach survival estimates being (negatively) correlated in release-recapture investigations [18]. More than two reaches apart, survival estimates are
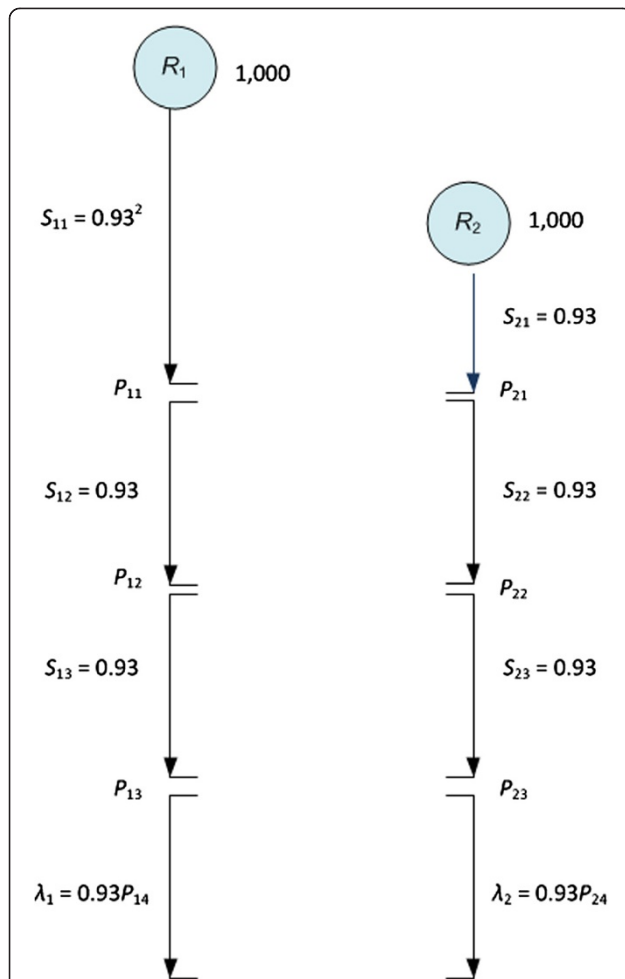
Figure 1 Schematic of paired release-recapture study. Schematic of the paired release-recapture design simulated under the null hypothesis of homogeneity with associated release sizes (R), survival (S), detection (P), and the joint probability of survival and detection in the last reach (λ). Detection probabilities $P_{11} = P_{12} = \ldots = P_{14} = P_{24}$ were varied at values of 0.10, 0.20, . . ., 0.90, 0.95, and 0.98.

uncorrelated. In addition, the correlation decreases as the capture probabilities increase [18].

$\lambda_1 \neq \lambda_2$

In the last period of a release-recapture study, the survival and detection probabilities are not separately estimable; only their product $Sp = \lambda$ is estimable. Simulations in the case of a $k = 4$-period, paired release-recapture study where $\lambda_1 \neq \lambda_2$ showed no bias in the estimation of S_0 despite models M_2 to M_7 having assumption violations. Again, precision was constant for $\hat{S}_0$ across the eight alternative models when detection probabilities were high (Table 5).

$P_{11} \neq P_{21}$

Setting detection probabilities across release groups different at the first downstream detection array (Figure 1,

$P_{11} \neq P_{21}$) resulted in the assumptions of only model M_7 being violated. A difference in detection probabilities of 0.05 (that is, $|P_{11} - P_{21}| = 0.05$) resulted in bias at the third decimal place (Table 6). Heterogeneity in detection probabilities further downriver that were improperly modeled had no effect on the survival estimate of S_0 from the paired release-recapture study.

Discussion

High detection probabilities are common in acoustic-tag studies as long as the ambient environment is relatively quiet. Exceptions occur, for example, in the tailrace environments of dams or riffle zones of streams or rivers where air bubbles can interfere with sound propagation. Other exceptions include noisy environments associated with construction, boat traffic, and mechanical operations (for example, turbines).

In a large industrial river reach such as the Columbia-Snake River, detection probabilities are typically very high. For example, during spring and summer acoustic-tag survival studies in 2012, 26 different cross-river arrays were used with 3 to 8 hydrophones per array [20,21]. River discharge ranged from 150 to 420 kcfs over river widths of 300 to 900 m. In one case, the estimated detection probability was as low as 0.90. In 94% of the cases, detection probabilities were ≥0.99, and in 59% of the cases, detection probabilities were estimated to be 1.0. At dense three-dimensional acoustic arrays at the faces of dams, $P = 1.0$.

In smaller rivers such as the San Joaquin River, Buchanan and colleagues [10] observed detection probabilities between 0.9 and 1.0 more than 90% of the time from detection arrays with as few as 1 to 4 hydrophones at river cross-sections less than 100 m. River discharges were in the range of 5 to 12 kcfs. Perry and colleagues [9] observed similar performance in the Sacramento River.

High detection probabilities not only simplify the statistical analysis of tagging studies but also permit smaller sample sizes for comparable precision. Skalski and colleagues [22] found sample sizes could be reduced by two orders of magnitude when performing release-recapture investigations using acoustic tags versus passive integrated transponder (PIT) tags in the Columbia River. Not only is study cost potentially reduced, but studies of threatened or endangered species become more feasible when the risk of take is substantially reduced.

Conclusions

Since the publication of Burnham and colleagues in 1987 [12], investigators have been encouraged to find a satisfactory model with the fewest number of parameters in paired-release investigations in order to improve sampling precision. Our Monte Carlo simulation studies using detection probabilities commonly encountered in

Table 3 Simulation results under heterogeneity, $S_{12} \neq S_{22}$

a. Bias $(\hat{S}_0 - S_0)$

P	S_{12}	S_{22}	Model						
			M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.90	0.95	0.93	−0.0002	−0.0002	−0.0002	−0.0002	−0.0002	0.0018	0.0017
0.95	0.95	0.93	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.0009
0.90	0.98	0.93	0.0000	0.0000	0.0000	0.0000	0.0000	0.0049	0.0047
0.95	0.98	0.93	0.0000	0.0000	0.0000	0.0000	0.0000	0.0024	0.0023
0.90	0.93	0.95	0.0003	0.0003	0.0003	0.0003	0.0003	−0.0017	−0.0016
0.95	0.93	0.95	0.0000	0.0000	0.0000	0.0000	0.0000	−0.0010	−0.0010
0.90	0.93	0.98	0.0002	0.0002	0.0002	0.0002	0.0002	−0.0047	−0.0045
0.95	0.93	0.98	0.0001	0.0001	0.0001	0.0001	0.0001	−0.0024	−0.0023

b. $\widehat{SE}(\hat{S}_0 | S_0)$

P	S_{12}	S_{22}	M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.90	0.95	0.93	0.0147	0.0147	0.0147	0.0147	0.0147	0.0147	0.0147
0.95	0.95	0.93	0.0144	0.0143	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.98	0.93	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146
0.95	0.98	0.93	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143
0.90	0.93	0.95	0.0147	0.0147	0.0147	0.0147	0.0147	0.0146	0.0146
0.95	0.93	0.95	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.93	0.98	0.0146	0.0146	0.0146	0.0146	0.0146	0.0145	0.0145
0.95	0.93	0.98	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143

Results of simulation studies of a paired release-recapture design (Figure 1) when all survival and detection parameters are homogeneous across release groups except $S_{12} \neq S_{22}$. In this case, models M_{Full}, and M_2 to M_5 are valid, but models M_6 and M_7 are invalid. Detection probabilities were simulated at either $P = 0.90$ or $P = 0.95$ at all locations.

acoustic-tag studies found little or no precision gains from selecting simpler yet valid models in describing paired-release investigations. Furthermore, reduced-parameter models for paired-release designs were found not to be robust to violations of the assumptions of homogeneous downstream survival or detection parameters in reaches immediately below the paired release. Heterogeneity in survival or detection probabilities in reaches two or more below the paired release had no effect on the unbiasedness of S_0 or its precision. Therefore, fully parameterized models are recommended in analyzing paired release-recapture, acoustic-tag studies with high detection probabilities (≥0.90) in order to retain both the properties of robustness and precision. Investigators should confirm these results when their studies are expected to be outside the range of simulations performed for this paper. Furthermore, it is generally good practice to perform preliminary ranging studies in the field to confirm anticipated detection rates.

Model selection should be reserved only for those instances where target precision was not achieved with the full model. The use of the fully parameterized model also eliminates the debate over the most appropriate model selection or model averaging techniques [17], thus removing a potential source of subjectivity and discrepancy among analysts, because there is only one fully parameterized model for everyone. Consequently, analyses are easily reproducible and consistent among analysts. The properties of consistency, objectivity, robustness, and precision are important when performing tagging studies that may influence resource management decisions.

Our findings pertain to paired release-recapture studies, which inherently involve model selection. Model selection is generally not an issue in standard single release-recapture studies [17]. Rarely are investigators interested in equating downstream reach survivals or detection probabilities across different locations or time periods. Nevertheless, high detection probabilities benefit these single release-recapture investigations in manners other than simply improving sampling precision. As mentioned previously, consecutive reach survivals from a release-recapture study are negatively correlated. Stochastic errors in the estimation of survival in one reach are compensated by estimation error in the opposite direction in the next reach. Patterns of overly high estimates preceding or following overly low estimates of survival can complicate biological interpretation of release-recapture results. To illustrate this point, we calculated the expected correlation between consecutive reach survival estimates in a single release-recapture investigation with two estimable reaches using the variance and covariance formula found in Cormack [18]. At

Table 4 Simulation results under heterogeneity, $S_{13} \neq S_{23}$

a. Bias $(\hat{\mathbf{S}}_0 - \mathbf{S}_0)$

P	S_{13}	S_{23}	Model						
			M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.90	0.95	0.93	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0002
0.95	0.95	0.93	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
0.90	0.98	0.93	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0004
0.95	0.98	0.93	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	0.0000	0.0000
0.90	0.93	0.95	0.0001	0.0001	0.0001	0.0001	0.0001	−0.0001	−0.0001
0.95	0.93	0.95	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
0.90	0.93	0.98	0.0001	0.0001	0.0001	0.0001	0.0001	−0.0004	−0.0003
0.95	0.93	0.98	0.0001	0.0001	0.0001	0.0001	0.0001	−0.0001	0.0000

b. $\overline{SE}(\hat{S}_0 | S_0)$

P	S_{12}	S_{22}	Model						
			M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.90	0.95	0.93	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148
0.95	0.95	0.93	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.98	0.93	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147
0.95	0.98	0.93	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.93	0.95	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147
0.95	0.93	0.95	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.93	0.98	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147
0.95	0.93	0.98	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144

Results of simulation studies of a paired release-recapture design (Figure 1) when all survival and detection parameters are homogeneous across release groups except $S_{13} \neq S_{23}$. In this case, models M_4 to M_7 are invalid. Detection probabilities were simulated at either $P = 0.90$ or 0.95 at all locations.

low-to-moderate detection probabilities (that is, $0.10 \leq P \leq 0.50$), there is a strong inverse correlation (that is, $-0.60 < Corr(\hat{S}_1, \hat{S}_2) < -0.50$) when reach survivals are as high as 0.93. However, at the high detection probabilities encountered with acoustic tags, the consecutive survival estimates from this single release-recapture investigation are nearly uncorrelated (Figure 2). The result is precise, uncorrelated survival estimates that are less prone to misinterpretation. Such properties of survival estimates are important if resource management decisions using such data are to be correctly informed.

Our simulation findings here relate to considerations of model selection and estimation for paired-release designs in a unidirectional linear migration scenario. For cases where fish have a choice of migration pathways or migration timing (for example, overwintering or delayed migration), the situation may grow more complicated as the full analysis model becomes more complex. In these cases, there may still be bias in parameter estimates even with high detection probabilities if the data are sparse relative to the complexity of the model. For example, the multi-state models of Buchanan and colleagues [10] and Perry and colleagues [9] incorporate multiple dynamically assigned states (that is, migration routes), any of

which may have few detections and thus may be subject to either inestimability or bias. The complexity of the analysis model in these cases may need to be reduced to better mirror the structure of the observed data. This type of model simplification is expected to both improve estimability and reduce bias caused by sparse data, and should not be confused with the issue of model selection from a paired-release design. The use of a paired-release design in these complex migration studies would face the same issue of model selection. Although such a scenario was not examined here, it is likely that similar conclusions apply, with an increased risk of bias under inappropriate homogeneity assumptions in return for little or no improvement in sampling precision. The interplay between the paired-release homogeneity assumptions and the inherent complexity and potential inestimability of a multistate model is not well understood, but can be assumed to be controlled partly by having an adequate sample size.

Methods

We used Monte Carlo simulation studies to simulate paired release-recapture investigations under a variety of conditions to assess robustness and precision of tagging

Table 5 Simulation results under heterogeneity, $\lambda_1 \neq \lambda_2$

a. Bias $(\hat{\bar{S}}_0 - S_0)$

P	λ_1	λ_2	Model						
			M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.90	0.95	0.93	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
0.95	0.95	0.93	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001
0.90	0.98	0.93	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
0.95	0.98	0.93	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001
0.90	0.93	0.95	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.95	0.93	0.95	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.90	0.93	0.98	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001
0.95	0.93	0.98	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003

b. $\widehat{SE}(\hat{S}_0 \mid S_0)$

P	λ_1	λ_2	M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
0.90	0.95	0.93	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147
0.95	0.95	0.93	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.98	0.93	0.0148	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147
0.95	0.98	0.93	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.93	0.95	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147	0.0147
0.95	0.93	0.95	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
0.90	0.93	0.98	0.0148	0.0148	0.0148	0.0148	0.0148	0.0147	0.0147
0.95	0.93	0.98	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144

Results of simulation studies of a paired release-recapture design (Figure 1) when all survival and detection parameters are homogeneous across all release groups except $\lambda_1 \neq \lambda_2$. In this case, models M_2 to M_7 are invalid. Detection probabilities were simulated at either $P = 0.90$ or 0.95 at all locations.

models. Initially, release sizes used in the simulations were set at $R_1 = R_2 = 1,000$ tags per release location. This sample size is reminiscent of release sizes of $1,000 \leq R_1 = R_2 \leq 3,000$ currently used to assess dam passage survival at federally operated dams in the Columbia/Snake River Basin [22]. A reduced set of simulations were also performed at $R_1 = R_2 = 250$ and 500, reminiscent of survival studies conducted at dams operated by public utilities in the Mid-Columbia River [23]. Detection probabilities (P) were initially examined over the range 0.10, 0.20, . . ., 0.80, 0.90, 0.95, and 0.98. This range included detection probabilities (for example, 0.10, . . ., 0.50) much lower than typically encountered in acoustic-tag studies but were used to assess extreme conditions. Subsequent simulations were performed with detection probabilities more commonly encountered in acoustic-tag studies (that is, >0.90).

Paired-release design

We simulated a paired-release design with four downstream reaches (Figure 1). The objective of the study was to estimate survival in the first reach bracketed by the two release groups:

$$S_0 = \frac{S_{11}}{S_{21}}. \tag{1}$$

The study design included four reaches, allowing estimation of six survival parameters (S_{ij}), six detection probabilities (P_{ij}), and two λ parameters in the full model (Figure 1). The focus of the simulations is on the precision and accuracy of the survival estimate $\hat{S}_0$.

We first simulated data under the null assumption of downstream homogeneity of survival and detection probabilities between the two release groups (H_0). Precision and unbiasedness of the S_0 estimate were examined for a set of eight alternative likelihood models under the null case (described below). Additional simulations were

Table 6 Simulation results under heterogeneity, $P_{11} \neq P_{21}$

P_{11}	P_{21}	Model						
		M_{Full}	M_2	M_3	M_4	M_5	M_6	M_7
a. Bias $(\hat{\bar{S}}_0 - S_0)$								
0.90	0.95	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	−0.0035
0.95	0.90	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0036
b. $\widehat{SE}(\hat{S}_0 \mid S_0)$								
0.90	0.95	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146	0.0145
0.95	0.90	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146

Results of simulation studies of a paired release-recapture design (Figure 1) when all survival and detection parameters are homogeneous across all release groups except $P_{11} \neq P_{21}$. In this case, only model M_7 is invalid. Detection probabilities were set at $P = 0.95$ at other locations.
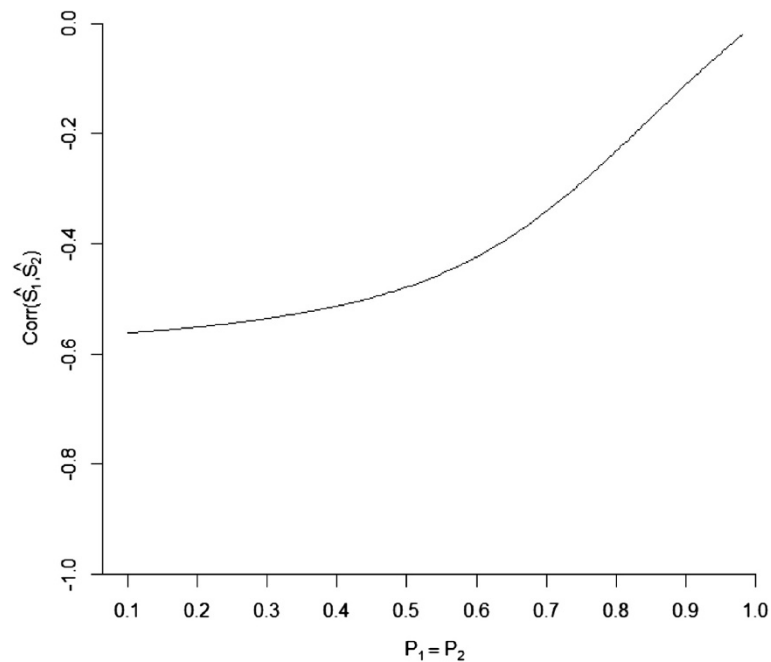
Figure 2 Negative correlation curve. Negative correlation in consecutive reach survival estimates (that is, $\hat{S}_1$ and $\hat{S}_2$) from a single release-recapture investigation with three reaches, $S_1 = S_2 = 0.93$, $\lambda = 0.93P$ over a range of detection probabilities ($0.10 \leq P \leq 0.98$). Results are independent of release sizes.

then performed to assess the sensitivity of the estimate of S_0 when the model assumptions were violated — in other words, to determine the presence and extent of estimation bias when the likelihood model incorrectly assumed homogeneity.

For each case, the survival probabilities in the first reaches after release were simulated at $S_{11} = 0.93^2$ and $S_{21} = 0.93$, thereby producing a survival value of $S_0 = 0.93$ in expectation (equation (1)). Unless otherwise specified (for example, in a nonhomogeneous case), the downstream reach survival probabilities (that is, $S_{12} = S_{13} = S_{22} = S_{23}$) were all simulated at 0.93, and the downstream detection probabilities (that is, $P_{11} = P_{12} = P_{21} = P_{22} = P_{13} = P_{23}$) were simulated over the range 0.10, 0.20, . . ., 0.95, and 0.98. In the last reach, the joint probability of survival and detection was modeled as $\lambda_i = S_{i4} \cdot P_{i4}$ with $S_{i4} = 0.93$ and $P_{i4} = P_{i3}$. Simulations for the nonhomogeneous cases were performed where (a) $S_{12} \neq S_{22}$, (b) $S_{13} \neq S_{23}$, (c) $\lambda_1 \neq \lambda_2$, and (d) $P_{11} \neq P_{21}$. In these simulations, only detection probabilities reminiscent of acoustic-tag studies (that is, >0.90) were evaluated. Differences in reach survivals were intentionally kept small (that is, ≤0.05) to assess the sensitivity of the paired release-recapture model to small model violations.

Ideally, an appropriate release-recapture model would produce unbiased point and variance estimates. Thus, for each scenario, 10,000 Monte Carlo simulations were performed, providing values for $\hat{\bar{S}}_0, s^2_{\hat{S}_0}$, and $\overline{\mathrm{Var}\,(\hat{S}_0|S_0)}$ for eight different likelihood models. The value $\hat{\bar{S}}_0$ is the average value of $\hat{S}_0$ derived from the 10,000 simulation runs and was compared to the actual value S_0 used to generate the release-recapture data (that is, $S_0 = 0.93$). Bias was then calculated as $\hat{\bar{S}}_0 - S_0$. The quantity $s^2_{\hat{S}_0}$ is the empirical variance among the 10,000 estimates of S_0. This quantity is a model-independent estimate of the sampling variance and was compared to the average variance estimate from the inverse Hessian for the likelihood model $\left(\text{that is,}\overline{\mathrm{Var}\,(\hat{S}_0|S_0)}\right)$. The difference between these two quantities is an estimate of the bias in the variance estimator.

The eight alternative likelihood models examined for each scenario were:

M_{Full}: All 14 parameters uniquely estimated (see Figure 1).

$M_2 = M_\lambda$: $\lambda_1 = \lambda_2$ parameters equated across releases; all other parameters unique.

$M_3 = M_\lambda, P_{\bullet 3}$: The $\lambda_1 = \lambda_2$ and $P_{13} = P_{23}$ parameters equated; all other parameters unique.

$M_4 = M_\lambda, P_{\bullet 3}, S_{\bullet 3}$: The $\lambda_1 = \lambda_2$, $P_{13} = P_{23}$, and $S_{13} = S_{23}$ parameters equated; all other parameters unique.

$M_5 = M_\lambda, P_{\bullet 3}, S_{\bullet 3}, P_{\bullet 2}$: The $\lambda_1 = \lambda_2$, $P_{13} = P_{23}$, $S_{13} = S_{23}$, and $P_{12} = P_{22}$ parameters equated; all other parameters unique.

$M_6 = M_\lambda, P_{\bullet 3}, S_{\bullet 3}, P_{\bullet 2}, S_{\bullet 2}$: The $\lambda_1 = \lambda_2$, $P_{13} = P_{23}$, $S_{13} = S_{23}$, $P_{12} = P_{22}$, and $S_{12} = S_{22}$ parameters equated; all other parameters unique.

$M_7 = M_\lambda, P_{\bullet 3}, S_{\bullet 3}, P_{\bullet 2}, S_{\bullet 2}, P_{\bullet 1}$: The simplest model; all parameters except $S_{11} \neq S_{21}$ equated across releases.

This set of models was examined under the presumption that the paired releases would eventually mix downriver producing homogeneous survival and detection processes among the two release groups. For each simulation, all eight of the above models were fit to the same dataset.

Authors' contributions

JRS directed the Monte Carlo simulation studies and wrote the initial drafts of the paper. AGS developed, tested, and performed the Monte Carlo simulation studies reported in this paper. RAB contributed to the writing and provided peer review of the final manuscript. All authors read and approved the final manuscript.

References

1. Costa DP, Robinson PW, Arnould JP, Harrison AL, Simmons SE, Hassrick JL, Hoskins AJ, Kirkman SP, Oosthuizen H, Villegas-Amtmann S, Crocker DE: **Accuracy of ARGOS locations of pinnipeds at-sea estimated using Fastloc GPS.** *PLoS One* 2010, **5**:1–9.
2. Donaldson MR, Arlinghaus R, Hanson KC, Cooke SJ: **Enhancing catch-and-release science with biotelemetry.** *Fish Fish* 2008, **9**:79–105.
3. Eiler JH: **Radio transmitters used to study salmon in glacial rivers.** *Am Fish Soc Symp* 1990, **7**:364–369.
4. McMichael GA, Eppard MB, Carlson TJ, Carter JA, Ebberts BD, Brown RS, Weiland M, Ploskey GR, Harnish RA, Deng ZD: **The juvenile salmon acoustic telemetry system: a new tool.** *Fisheries* 2010, **35**:9–22.
5. Spedicato MT, Lembo G, Marmulla G: **Aquatic telemetry: advances and applications.** In *Proceedings of the Fifth Conference on Fish Telemetry: 9–13 June 2003*. Ustica, Italy. Rome: COISPA tecnologia & ricerca, Food and Agriculture Organization of the United Nations; 2003.
6. O'Dor R, Boustany AM, Chittenden CM, Costello MJ, Moustahfid H, Payne J, Steinke D, Stokesbury MJW, Vanden Berghe E: **A census of fishes and everything they eat: how the census of marine life advanced fisheries science.** *Fisheries* 2012, **37**:398–409.
7. Hoffmann A, Skalski JR: **Inferential properties of an individual-based survival model using release-recapture data: sample size, validity, and power.** *J Appl Stat* 1995, **22**:579–595.
8. Buchanan RA, Skalski JR: **Using multistate mark-recapture methods to model adult salmonid migration in an industrialized river.** *Ecol Model* 2010, **221**:582–589.
9. Perry RW, Skalski JR, Brandes PL, Sandstrom PT, Klimley AP, Ammann A, MacFarlane B: **Estimating survival and migration route probabilities of juvenile Chinook salmon in the Sacramento-San Joaquin River Delta.** *North Am J Fish Manag* 2010, **30**:142–156.
10. Buchanan RA, Skalski JR, Brandes P, Fuller A: **Route use and survival of juvenile Chinook salmon through the San Joaquin Delta.** *North Am J Fish Manag* 2013, **33**:216–229.
11. Johnson GE, Hedgepeth JB, Skalski JR, Giorgi AE: **A Markov chain analysis of fish movements to determine entrainment zones.** *Fish Res* 2004, **69**:349–358.
12. Burnham KP, Anderson DR, White GC, Brownie C, Pollock KH: **Design and analysis methods for fish survival experiments based on release-recapture.** *Am Fish Soc Monogr* 1987, **5**.
13. Mathur D, Heisey PG, Euston ET, Skalski JR: **Turbine passage survival estimates for chinook salmon smolt (*Oncorhynchus tshawytscha*) at a large dam on the Columbia River.** *Can J Fish Aquat Sci* 1996, **53**:542–549.
14. Skalski JR, Townsend RL, Steig TW, Horchik JW, Tritt GW, McDonald RD: *Estimation of Rock Island Project Passage Survival of Yearling Chinook Salmon Smolts in 2003 Using Acoustic and PIT-tag Release-recapture Methods.* Wenatchee, WA: PUD No. 1 of Chelan County; 2003.
15. Hogg RV, Craig AT: *Introduction to Mathematical Statistics.* Englewood Cliffs, NJ: Prentice Hall; 1995.
16. Akaike H: **Information theory and an extension of the maximum likelihood principle.** In *Proceedings of the 2nd International Symposium on Information Theory, September 2–8, 1971, Tsahykadsor, Armenia.* Edited by Petrov BN, Csaki F. Budapest: Akadémiai Kiadó; 1973:267–281.
17. Burnham KP, Anderson DR: *Model Selection and Multimodel Inference.* New York: Springer-Verlag; 2002.
18. Cormack RM: **Estimates of survival from the sighting of marked animals.** *Biometrika* 1964, **51**:429–438.
19. Seber GAF: *The Estimation of Animal Abundance and Related Parameters.* New York: Macmillan; 1982.
20. Skalski JR, Townsend RL, Seaburg AG, McMichael GA, Oldenburg EW, Harnish RA, Ham KD, Colotelo AH, Deters KA, Deng ZD: *Compliance Monitoring of Yearling and Subyearling Chinook Salmon and Juvenile Steelhead Survival and Passage at Little Goose Dam, 2012.* Richland, WA: Pacific Northwest National Laboratory; 2013.
21. Skalski JR, Townsend RL, Seaburg AG, Ploskey GR, Weiland MA, Hughes JS, Woodley CM, Carlson TJ: *Compliance Monitoring of Subyearling Chinook Salmon Smolt Survival and Passage at Bonneville Dam, Summer 2012.* Richland, WA: Pacific Northwest National Laboratory; 2013.
22. Skalski JR, Townsend RL, Seaburg AG, Hughes JS, Weiland MA, Ploskey GR, Woodley CM, Carlson TJ: *Compliance Monitoring of Yearling and Subyearling Chinook Salmon and Juvenile Steelhead Survival and Passage at McNary Dam, 2012.* Richland, WA: Pacific Northwest National Laboratory; 2013.
23. Skalski JR, Steig TW, Hemstrom SL: **Assessing compliance with fish survival standards: a case study at Rock Island Dam, Washington.** *Environ Sci Pol* 2012, **18**:45–51.