FISH PASSAGE CENTER

847 NE 19th Avenue, #250, Portland, OR 97232
Phone: (503) 833-3900 Fax: (503) 232-1259
www.fpc.org/
e-mail us at fpcstaff@fpc.org

TO: Fish Passage Advisory Committee (FPAC)

FROM: Michele DeHart

DATE: February 16, 2016

RE: Review of *The Influence of Individual Fish Characteristics on Survival and Detection: Similarities across Two Salmonid Species*

The referenced article was circulated to FPAC by Paul Wagner. At my request, the Fish Passage Center (FPC) staff is providing this review of *The Influence of Individual Fish Characteristics on Survival and Detection: Similarities across Two Salmonid Species.* This article was published in the 35th edition of the North American Journal of Fisheries Management in 2015. The results of this study have implications for management recommendations (e.g., juvenile transportation effectiveness) and also for long-term monitoring studies relying on the detection of PIT tags within bypass systems at Snake River dams (e.g., the Comparative Survival Study (CSS)). For these reasons, we provide a close examination of the methodology and findings reported in this article.

In the discussion section of Hostetter et al. (2015), it states that "trait-selective (length and external condition characteristics) capture methods may lead to tagged populations that are not representative of the overall population of interest" and that the study "found consistent length selectivity and condition selectivity in both survival and detection probabilities."

An unrepresentative target population due to trait selectivity was first suggested in Zabel et al. (2005) with a specific focus on length selectivity. Hostetter et al. (2015) also examine length selectivity, and additionally examine selectivity due to degradative external condition characteristics (descaling, body injuries, and fin damage). In light of FPC's review of Hostetter et al. (2015), the FPC strongly cautions against any conclusion being made from this study regarding the effects of external condition characteristics on detection probability. Support for this cautionary statement is presented in the following. We also present other methodological concerns that if addressed, may lead the authors to different conclusions about trait selectivity.

In the review of Hostetter et al. (2015), the FPC found that this study:

- Overstated the importance and significance of external condition characteristics on detection probability at Little Goose and Lower Monumental dams. The best model in their candidate set included external condition characteristic variables, but none of these variables were statistically significant.
- Ignored results from other competing models in their candidate set. The second best fitting model for steelhead and yearling Chinook included only length effects on survival and detection probability, and did not include any external condition effects. Furthermore, this second best fitting model for steelhead differed from the best fitting model by only 1.2 AIC units which suggests that this model was equally supported by the data.
- Did not use multimodel inference techniques such as model averaging, or relative variable importance determination, which are widely recommended when there are other competing models in a candidate set of models, and when the primary objective of the modeling is to evaluate the importance of several predictor variables. Due to the lack of significance of many of the external characteristic variables, the conclusions of this study may have been different if multimodel inference techniques were used.
- Omitted important variables which are known to affect survival and detection probability at Snake River dams, namely flow (or water transit time), spill, and arrival day.

It is also important to remember that:

- The study and resulting conclusions are based on one year of collected data representing a single set of environmental conditions.
- Only hatchery fish were marked and evaluated for this study. No data are presented to demonstrate that daily mark groups were of consistent size over the time period. Thus, length may be confounded with other variables not included in this analysis that changed over the time period.
- Inferences about trait selectivity on detection probability at Lower Granite Dam cannot be made from this study since all the fish utilized in this study were tagged at this location. Results about detection probability from this study apply only to Little Goose and Lower Monumental dams. This is important to remember because the largest proportions of steelhead and yearling Chinook are transported from Lower Granite Dam. Thus any management implication about the effectiveness of transportation suggested by this study does not apply to Lower Granite Dam.
- Because the CSS tags fish above Lower Granite Dam, studies have been conducted to evaluate length selectivity on detection probability at Lower Granite Dam. These studies considered multiple years of data, provided only marginal support for any length-detection probability relationships, and found no clear differences in length distributions for detected and undetected fish at Lower Granite Dam.

Non-significant parameters and the best model

FPC's biggest concern with Hostetter et al. (2015) is that it overstates several conclusions that are not supported by their best fitting model. Specifically, in several sections of the paper, it concludes that the study found evidence of condition selectivity in both survival and detection probability (Abstract, Results 4th paragraph, and Discussion 1st paragraph). There appears to be some evidence of condition selectivity in survival, but very little evidence of condition selectivity in detection probability. Since Hostetter et al. (2015) does admit that the condition selectivity parameters for detection probability all contain zero (i.e., are not statistically significant), the false conclusion that the study makes is due to a misinterpretation of AIC. Model selection strategies, such as AIC, are parameter penalized goodness-of-fit measures, not variable significance measures.

An underappreciated facet of AIC-based model selection is that it has about a 1 in 6 chance of admitting a spurious variable based on lower AIC (Arnold 2010). Thus, it is entirely possible that the model with the lowest AIC in a candidate set contains variables that are not statistically significant (at the $\alpha = 0.05$ significance level). The statistical significance of the external condition variables (descaling, body injuries, and fin damage), and other variables included in the best fitting model, are indirectly presented with confidence intervals in Table 5 of Hostetter et al. (2015) and presented below in Table 1.

Table 1. Table 5 from Hostetter et al. (2015). Boxes are added to variables that are statistically significant ($\alpha = 0.05$ significance level) as indicated by 95% confidence that do not contain zero.

TABLE 5. Effects of individual fish characteristics on survival and detection probabilities for steelhead and yearling Chinook Salmon that were PIT-tagged and released at Lower Granite Dam in 2014. Values represent logit-scale parameter estimates (with 95% confidence intervals [CIs]) from the top model (see Table 3) for each species. Appendix Table A.1 provides the results for all parameters.

Probability or characteristic	Steelhead		Chinook Salmon	
	Mean	95% CI	Mean	95% CI
Survival				
Descaling	−0.41	(−0.72, −0.11)	−0.26	(−0.45, −0.06)
Body injuries	−0.51	(−0.93, −0.08)	−0.07	(−0.37, 0.23)
Fin damage	0.22	(−0.13, 0.56)	−0.12	(−0.29, 0.04)
FL	0.18	(0.02, 0.34)	0.16	(−0.03, 0.36)
Detection				
Descaling	0.08	(−0.02, 0.19)	−0.04	(−0.18, 0.10)
Body injuries	0.19	(0.00, 0.37)	0.10	(−0.12, 0.32)
Fin damage	−0.03	(−0.12, 0.06)	0.04	(−0.09, 0.16)
FL	−0.08	(−0.13, −0.04)	−0.20	(−0.25, −0.16)
Spill percentage	−0.36	(−0.39, −0.32)	−0.32	(−0.36, −0.28)

The parameter estimates for external condition in Table 5 of Hostetter et al. (2015) indicate the difference in mean detection probability (or survival) (on the logit scale) for an individual with a certain condition, compared to an individual without that condition. For these type of parameters, a 95% confidence interval that contains 0 indicates that the data cannot rule out that the difference in mean detection probability for a fish with a particular condition—compared to fish without that condition—is zero (at the $\alpha = 0.05$ significance level). The 95% confidence interval of every external condition-related detection probability parameter in Table 5 of Hostetter et al. (2015) contains 0.

It appears that Hostetter et al. (2015) has confused AIC model selection techniques with variable importance and significance techniques. A variable that is included in the best fitting model (i.e., the model with lowest AIC) does not necessarily suggest that this variable is statistically significant or important. The implications of this confusion are severe as it appears to have led the study to overstate conclusions and make broad statements about management recommendations and PIT-tag studies that may not be supported by the data.

Other competing models

Hostetter et al. (2015) chose to report the results only from the best fitting model (the model with the lowest AIC) for steelhead and Chinook. Reporting results from only the best model in the candidate set should be done only when no other models in the candidate set are supported as indicated by AIC weights equaling 0 for all other models (Anderson and Burnham 2002). This was not the case for the models reported in Hostetter et al. (2015) as shown below:

Table 2. Table 3 from Hostetter et al. (2015).

TABLE 3. Model selection results comparing seven a priori hypotheses used to investigate the influence of individual fish characteristics (FL and external condition characteristics) on survival and detection probabilities for juvenile steelhead and yearling Chinook Salmon. Models were compared based on the difference in Akaike's information criterion (ΔAIC_c) between the given model and the best-performing model (shown in bold italics) and based on model weight (w_i). The number of parameters (K) is also provided. Detailed models for each hypothesis are provided in Table 2.

		Steelhead[a]		Chinook Salmon[b]	
Hypothesis	K	ΔAIC_c	w_i	ΔAIC_c	w_i
Survival varies by reservoir; detection varies by dam and spill percentage	6	13.2	0.00	102.9	0.00
Length effect on survival	7	15.2	0.00	93.7	0.00
Length effect on detection	7	5.2	0.04	17.8	0.00
Length effect on survival and detection	8	1.2	0.34	3.1	0.18
Length and external condition effects on survival	10	12.0	0.00	86.6	0.00
Length and external condition effects on detection	10	8.9	0.01	22.1	0.00
Length and external condition effects on survival and detection	***14***	***0.0***	***0.61***	***0.0***	***0.82***

[a]The AIC_c of the top model for steelhead was 38,315.8.
[b]The AIC_c of the top model for Chinook Salmon was 26,845.5.

In fact, the model without any external condition effects on survival and detection probability (Hypothesis: Length effect on survival and detection) was the next best fitting for steelhead and Chinook. The differences in AIC from this model to the best fitting model (ΔAIC) was 1.2 and 3.1, resulting in AIC weights (w_i) of 0.34 and 0.18, for steelhead and Chinook respectively. The third best model in the candidate set for both steelhead and Chinook (Hypothesis: Length effect on detection) also did not include external condition effects.

According to Burnham and Anderson (2002, p. 70–72), when the difference in AIC between two models (ΔAIC) is < 2, then both models are equally well supported by the data. Thus, for steelhead, Hostetter et al. (2015) failed to recognize that the model without any external condition effects was also plausible and supported by the data. When there are other competing models in the candidate set, Burnham and Anderson (2002) recommends either (1) reporting the other competing models in the candidate set, or (2) using model averaging techniques, weighted by AIC weights. One of the benefits of model averaging is that this technique minimizes the effect of uninformative parameters. Hostetter et al. (2015) did not report other competing models in their candidate set and did not utilize model averaging techniques. If the study instead chose either of these approaches, it is possible that the interpretation of the results could have changed considerably. In the section below, we discuss how relative variable importance techniques that incorporate model weights should have been used.

Relative Variable Importance

The way in which Hostetter et al. (2015) decided to include models with external condition factors in their candidate set of models could be biasing their conclusions about the importance of these variables. When building models to express hypotheses about external condition factors, Hostetter et al. (2015) chose only to include all three external condition variables (descaling, body injuries and fin damage) or no external condition variables. The authors did not include any intermediate models with smaller subsets of the external condition variables (e.g., a model with only the descaling external condition variable, and not body injuries and fin damage). Intermediate models with fewer external condition variables should have been considered due to the lack of statistical significance of many of these variables (Table 1).

If Hostetter et al. (2015) wanted to assess the importance of the external condition variables, and other variables, they should have included more models that have subsets of the external condition variables in their candidate set and calculated the relative variable importance metric (Burnham and Anderson 2002: 167–169). Using this approach, the importance of a variable is made by making inferences from all the models in the candidate set by weighting the variables according to their AIC weights (w_i). This multimodel inference approach avoids only interpreting results from the best model and provides a better way to make statements about variable importance while accounting for uncertainty in the determination of model rankings.

Statistical and Practical Significance

Hostetter et al. (2015) does an adequate job of reporting results from their best fitting model. In this section, FPC focuses only on the results that were statistically significant (Table 1) and comment on the magnitude of the observed effect. FPC specifically focuses on the fork length variable, because it should be no surprise that descaled fish, and fish with body injuries or fin damage, would tend to be less likely to survive compared to fish without those conditions (Budy et al. 2002, Hostetter et al. 2011). Also, the relationship between spill (or power house passage) and detection probability is already well documented (Sandford and Smith 2002, McCann et al. 2015, Appendix J).

Zabel et al. (2005) was the first to discuss the potential of trait-selective capture methods evidenced by statistically significant relationships between (1) survival and fork length, and (2) detection probability and fork length, for PIT-tagged juvenile salmonids at Snake River dams. These results were also generally confirmed by Hostetter et al. (2015), except for the non-significant relationship between fork length and survival for Chinook. The significant fork length relationships that Hostetter et al. (2015) found indicated that:

- From Lower Granite to Little Goose Dam, steelhead survival increased from approximately 0.91 for a fish equaling the 2.5th length percentile, to approximately 0.95 for a fish equaling the 97.5th length percentile.
- At Little Goose Dam, steelhead detection probability decreased from approximately 0.36 for a fish equaling the 2.5th length percentile, to approximately 0.30 for a fish equaling to the 97.5th length percentile.
- At Little Goose Dam, yearling Chinook detection probability decreased from approximately 0.37 for a fish equaling the 2.5th percentile, to approximately 0.21 for a fish equaling the 97.5th percentile.

Without a statement of objectives or specific hypothesis to address, we cannot assert whether or not these relationships are practically significant. We would assert, however, that the overall effect size is not large except for the relationship between yearling Chinook detection probability and fork length. Because of the sample sizes utilized in this study—steelhead (n = 11,201) and yearling Chinook (n = 7,943)—it is not surprising that statistically significant results were found that may not be very practically significant.

Inference about Lower Granite Dam

It is worth noting that the largest proportion of steelhead and yearling Chinook that were collected for transportation in 2014 (and in most given years) occurs at Lower Granite Dam (www.fpc.org/smolt/transportation/). Hostetter et al. (2015) tagged their fish at Lower Granite Dam and thus were not able to make any inferences about detection probability at that site. The study is only able to make inferences about trait-selective detection probability at Little Goose and Lower Monumental dams, and only graphically present results for detection probability at Little Goose Dam. Thus, the management implications about transportation effectiveness that are conveyed from this study apply only to Little Goose and Lower Monumental Dam.

The CSS tags and releases fish above Lower Granite Dam and hence can estimate detection probability at this site. In response to Zabel et al. (2005)—a study that also did not estimate detection probability at Lower Granite Dam—the CSS examined whether the same length-detection probability relationship was supported at Lower Granite Dam for the same migration years in Zabel et al. (2005) (Berggren et al. 2006, Chapter 9). The AIC model-selection results provided only marginal support for any length-detection probability relationship across all sites and migration years. For Lower Granite Dam in particular, the estimated length-detection probability relationships were weak to nonexistent. Additionally, there were no clear differences in the length distributions for detected and undetected fish across all projects and migration years. For additional analyses regarding length selectivity and its implications for transportation effectiveness evaluation also see FPC (2008).

Important Variables not Included in this Analysis

Hostetter et al. (2015) chose to frame the focus of their study on individual trait (length and external condition) characteristics affecting survival and detection probability. This was reflected by their choice of models in their candidate set which included models with and without individual trait variables. However, the authors also decided to include spill (i.e., percentage of water spilled at a dam) as a covariate describing detection probability due to the known negative relationship between detection probability and spill (Sandford and Smith 2002, McCann et al. 2015, Appendix J). Thus, the Hostetter et al. (2015) extended the hypotheses about the factors affecting survival and detection probability to also include hypotheses about environmental variables. Since the authors chose to do so, it is unclear why the authors did not also include more variables that are also known to affect survival and detection probability. Specifically, the study should have also examined:

- The effects of flow (or water transit time) on detection probability (McCann et al. 2015, Appendix J, Plumb et al. 2012).
- An interaction between flow (or water transit time) and spill on detection probability (McCann et al. 2015, Appendix J, Plumb et al. 2012).
- The effects of arrival day, spill, and flow (or water transit time) on survival (Haeseker et al. 2012).

The choice not to include an arrival day variable is also particularly troublesome. This variable should have been included, along with interactions, to help account for biological, physiological and environmental changes that occur across the migration season. For instance, when Berggren et al. (2006, Chapter 9) examined the effect of length on detection probability at Lower Granite Dam, the study limited the data to a 30-day period extending from April 11 to May 10 because hydrological variables and fork length were relatively stable across this period during the years that they examined. This was necessary to do because an arrival day variable was not included in the model. Since there were trends in these variables over the entire migration season, it was important to limit the data to periods where these variables were stable; otherwise the effect of one variable (e.g., length) can be confounded with changes in the effect of another variable (e.g., flow).

If Hostetter et al. (2015) wanted only to examine models that address hypotheses about the effects of individual trait characteristics (length and external condition) on survival and detection probability, then the models considered in their candidate set should have focused only on these variables. Since the authors chose to extend their hypotheses to include environmental variables by modeling detection probability as a function of the percentage of flow spilled at a dam, then the study should have also considered other variables such as flow, spill, and arrival day when modeling survival and detection probability. A better set of candidate models would have also included models with flow, spill, and arrival day, and the full model in the candidate set should have included all these variables in addition to the variables already considered by the study.

References

Anderson, D.R. and K.P. Burnham. 2002. Avoiding Pitfalls When Using Information-Theoretic Methods. Journal of Wildlife Management (66):910–6.

Arnold, T.W. 2010. Uninformative parameters and model selection using Akaike's Information Criterion. Journal of Wildlife Management (74): 1175–1178.

Berggren T., P. McHugh, P. Wilson, H. Schaller, C. Petrosky, E. Weber, and R. Boyce. 2006. Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook and Summer Steelhead. BPA Contract # 1996-02-00.

Budy, P., G.P. Thiede, N. Bouwes, C.E. Petrosky, and H. Schaller. 2002. Evidence linking delayed mortality of Snake River salmon to their earlier hydrosystem experience. North American Journal of Fisheries Management (22): 35–51.

Burnham, K.P. and D.R. Anderson. 2002. Model selection and inference – a practical information-theoretic approach. Second edition. Springer-Verlag. New York.

FPC Memorandum. 2008. Size differences in transported fish due to size selectivity of smaller fish by juvenile bypass/collection system. www.fpc.org/documents/memos/52-08.pdf

Haeseker, S.L., J.M. McCann, J.E. Tuomikoski, and B. Chockley. 2012. Assessing freshwater and marine environmental influences on life-stage-specific survival rates of Snake River spring/summer Chinook salmon and steelhead. Transactions of the American Fisheries Society (141): 121–138.

Hostetter, N. J., A. F. Evans, D. D. Roby, K. Collis, M. Hawbecker, B. P. Sandford, D. E. Thompson, and F. J. Loge. 2011. Relationship of external fish condition to pathogen prevalence and out-migration survival in juvenile steelhead. Transactions of the American Fisheries Society (140):1158–1171.

Hostetter, N. J., A. F. Evans, R. R. O'Connor, B.M. Cramer, D. Fryer and K. Collis. 2015. The influence of individual fish characteristics on survival and detection: similarities across two salmonid species. North American Journal of Fisheries Management (35): 1034–1045.

McCann J., B. Chockley, E. Cooper, H. Schaller, S. Haeseker, R. Lessard, C. Petrosky, E. Tinus, E. VanDyke, R. Ehlke. 2015. Comparative Survival Study (CSS) of PIT-tagged Spring/Summer/Fall Chinook, Summer Steelhead, and Sockeye. BPA Contract # 1996-02-00.

Plumb, J.M., W.P. Connor, K.F. Tiffan, C.M. Moffitt, R.W. Perry, and N.S. Adams. 2012. Estimating and predicting collection probability of fish at dams using multistate modeling. Transactions of the American Fisheries Society (141):1364–1373.

Sandford, B. P. and S. G. Smith. 2002. Estimation of smolt-to-adult return percentages for Snake River basin anadromous salmonids, 1990–1997. Journal of Agricultural, Biological, and Environmental Statistics (7): 243–263.

Zabel, R. W., T. Wagner, J. L. Congleton, S. G. Smith, and J. G. Williams. 2005. Survival and selection of migrating salmon from capture–recapture models with individual traits. Ecological Applications (15): 1427–1439.