Comment: Testing the Independence of Microhabitat Preferences and Flow (Part 2)

Beecher et al. (1995) claimed to have validated an assumption of the instream flow incremental methodology (IFIM) and of its physical habitat simulation model (PHABSIM) that water depths and velocities preferred by fish are independent of streamflow. We disagree. First, the study compared habitat preference (of young steelhead *Oncorhynchus mykiss*) at low flow with habitat use (not preference) at high flow. Second, the statistical approach used can lead to the conclusion that depth and velocity preferences are the same at two flows even when most scientists would consider them different. Here, we address statistical issues that arise when comparisons of habitat distribution are based on fish preferences rather than on depth and velocity, and we recommend new techniques for comparing distributions.

Beecher et al. (1995) addressed the null hypothesis "fish distribution is independent of depth and velocity preferences determined at a different flow." They rejected the hypothesis on the basis of a goodness-of-fit test that compared the observed distribution of fish among preference quartiles at a high flow with the expected uniform distribution that would result if fish selected habitat without regard to preferences derived at a lower flow. (The test statistic was $P[dv]$, the product of preferred depth $P[d]$ and preferred velocity $P[v]$.) We do not believe that rejecting this hypothesis implies similarity in fish preferences over a range of flows.

The comparison was of low-flow preference with high-flow *use*; the question of whether preference changed with flow should have been addressed by calculating preference at both flows. It can be shown mathematically that both habitat use and habitat preference cannot be flow-invariant. Habitat use must change in response to flow (i.e., to habitat availability). The original question remains, because preference might be flow-invariant though habitat use may shift.

Unlike traditional comparisons of habitat preferences in which univariate depth and velocity preference distributions are compared between two fish populations, Beecher et al. used preference ranges instead of depth–velocity ranges. Their frequencies were numbers of fish, but expected fish numbers were calculated with the number of $P[dv]$ cells in each of four preference ranges: (expected number of fish) = (fraction of cells in preference range) × (total number of fish). This approach is unlikely to detect shifts in habitat preference because it compares the numbers of fish in each of four preference ranges (quartiles) between two flows. Collapsing data into ranges masks differences in depth and velocity preferences because many depth–velocity combinations can share the same $P[dv]$ value. Consider a unimodal depth preference curve for trout at low flow. Intermediate

depths are optimal, and both shallow and deep habitats are marginal. As flow increases, fish that had been forced to occupy shallow water may shift to deep water, as they did in Pert and Erman's (1994) study. If the fish distribution data are aggregated within preference ranges, shallow and deep fish would be lumped together in the low-preference category, and their joint proportion of the population might not change between low and high flows. In such a case, the test used by Beecher et al. would not detect the marked habitat shift that fish underwent. The same problem can arise with velocity, and the marginal preference range becomes an even greater catchall when both habitat factors are combined. Whether habitat is unsuitable because depth or another factor is too great or too slight is immaterial for IFIM calculations of weighted usable area (WUA) and instream flows. It does matter for testing and comparing preferences. Having the same proportions of fish in habitat deemed unsuitable, marginal, and optimal at different flows does not imply that the same depths and velocities were preferred.

The test used by Beecher et al. also had low power because their hypothesis was that high-flow use (assume it was preference) is independent of low-flow preference, rather than that preferences at the two flows are equal. The type I error rate is low for the hypothesis tested, and only large differences in preference would be diagnosed as real.

Thomas and Bovee (1993) used a test like that of Beecher et al. to evaluate transferability of IFIM habitat suitability curves. They quantified the relationship between type I and type II error rates and the number of occupied and unoccupied $P[dv]$ cells. This test is strongly influenced by habitat availability because it depends on cell frequencies instead of fish frequencies. Its dependence on the quantity and characteristics of unoccupied cells is undesirable, because it seems unreasonable that a difference in preference between two flows should depend on the index values assigned to empty cells. Fish density may influence the degree to which "suitable" cells are occupied.

We suggest alternative tests that can detect smaller differences between preferences than the one used by Beecher et al. (1995), and we propose a way to define ecologically significant differences. As an illustrative example, we use the frequency distributions of preference for depth and velocity shown by adult rainbow trout (nonanadromous *O. mykiss*) at low and high flows (Pert and Erman 1994). Preferences shifted to deeper and faster water when flow increased (Figure 3). We

chose this example because most people can agree without statistical confirmation that a clear shift in habitat preference occurred.

In two tests, we evaluated the habitat shift by resampling the joint depth–velocity preference distribution. Resampling provides confidence bounds of statistics with unknown distributional characteristics, such as the preference index. In our proposed tests and in an application of the Beecher et al. test, we used the true bivariate or joint preferences $(P[d,v])$ rather than the usual index $(P[dv] = P[d] \times P[v])$.

In our first test, we resampled the habitat use data for each flow, drawing fish observed in different depth–velocity combinations. For each of 50 replicate samples, we calculated the differences between preferences at high and low flows. The 1% and 99% confidence bounds for several depth and velocity classes did not include zero (zero implies no difference between flows; Table 1). According to this test, preferences were significantly different at the two flows, particularly in deeper habitat.

In the second test, we focused on defining an ecologically meaningful statistic to describe the preference distributions. The peak of the WUA curve would be a good ultimate endpoint, but we chose the peak of the joint preference distribution $P[d,v]$ as a simpler surrogate. We tested the hypothesis that the $P[d,v]$ peak did not shift in response to flow. For all the low-flow samples we drew, the peak occurred within the depth range of 96–120 cm and the velocity range of 15–30 cm/s. At the high flow, 36% of the samples peaked within these ranges, but 64% peaked in deeper (120–144 cm) and faster (30–60 cm/s) habitat. A binomial test rejected the hypothesis that the peaks were the same at both flows ($|z| = 3.75$; $P < 0.0001$).

Finally, we applied the goodness-of-fit test used by Beecher et al. to the joint preference data organized in the following form:

Low-flow preference range	High-flow percentages of:	
	Expected cells	Observed fish
0.0–0.1	54	40
0.1–0.3	22	26
0.3–1.0	24	34

The null hypothesis of independence from low-flow preferences was rejected ($\chi^2 = 8.46$; df = 2; $P = 0.014$).

These results appear to contradict one another. Although trout did not select habitat without re-
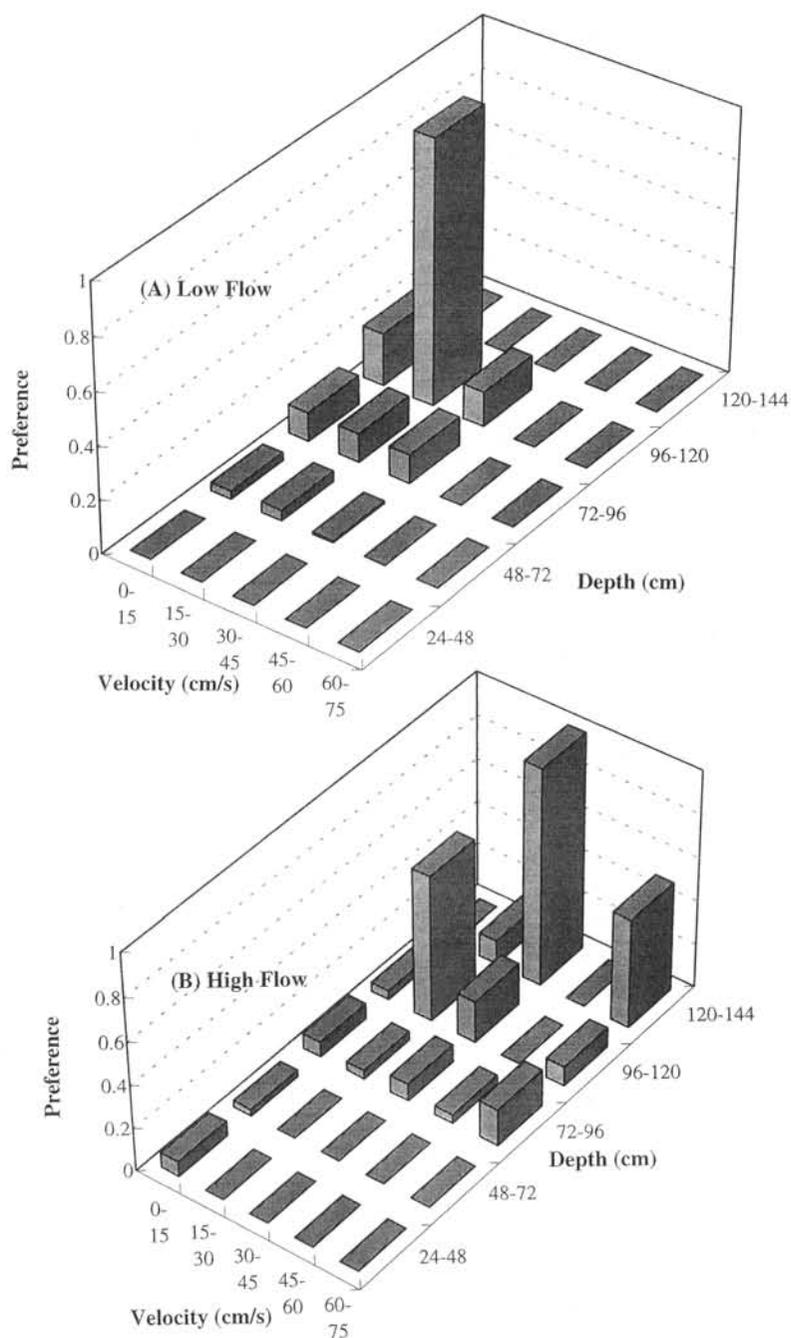
FIGURE 3.—Adult trout preferences calculated as a function of river depth and velocity for (A) a low flow and (B) a high flow.

gard to low-flow P[d,v], their habitat preferences shifted to greater depths and velocities with increased flow. This contradiction is possible because the tests are mirror images of one another and because the probability of a type I error (reject when true) is set to a low value ($\alpha \leq 0.10$) for each. Fish in these samples fell into the wide intermediate area between extremes of complete and no constancy in habitat preference with changes

in flow. Which test is better, and which level of type I error is acceptable? Ecologists are coming to realize that the balance between type I and type II errors should be reasonable in terms of ecological significance (Quinn and Dunham 1983; Roughgarden 1983; Toft and Shea 1983). In our case, it is misleading to use a test that rejects the null hypothesis at the slightest similarity and then claim that no shift in preference has occurred.

TABLE 1.—Resampling test of the hypothesis of zero difference between high-flow and low-flow joint preferences of rainbow trout for depth and velocity. Values are the sample differences in preference (high flow minus low flow) and (in parentheses) the nonparametric 1% and 99% confidence bounds determined by resampling. Asterisks indicate significant differences from zero ($P \leq 0.01$; i.e., 98% of the range of simulated differences failed to bracket zero). Parenthetic words in place of confidence bounds mean that a habitat combination was present only at low flow (low), at high flow (high), or at neither flow (neither).

Depth class (cm)	Velocity class (cm/s)				
	0–15	15–30	30–45	45–60	60–75
0–24	0.00 (0.00, 0.00)	(neither)	(neither)	0.00 (low)	0.00 (low)
24–48	0.07 (0.00, 0.18)	0.00 (high)	0.00 (0.00, 0.00)	0.00 (low)	(neither)
48–72	0.01 (−0.08, 0.08)	−0.04 (−0.14, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (high)
72–96	−0.04 (−0.27, 0.18)	−0.07 (−0.22, 0.01)	−0.03 (−0.28, 0.10)	0.09 (0.10, 0.54)*	0.04 (0.00, 0.19)
96–120	−0.16 (−0.56, −0.03)*	−0.34 (−0.77, 0.00)	0.09 (0.04, 0.80)*	0.08 (0.00, 0.26)	0.00 (0.00, 0.00)
120–144	(neither)	0.11 (high)	0.95 (0.55, 1.00)*	0.50 (high)	(neither)

However, small shifts in preference that do not influence the predicted relationship between WUA and streamflow may be tolerable.

How do we detect differences that are ecologically significant? One good way is to determine the magnitude of shift in depth or velocity preference that would significantly change peak WUA. Williams (1996) showed that variation in preference curves can cause large differences in peak WUA. Once the magnitude of a significant preference shift has been defined, one can design habitat studies with adequate power for detecting such a shift. If a compilation of IFIM studies allowed flow-related changes in habitat availability to be characterized, general guidelines might be developed that would circumvent the need for a new IFIM study on every regulated stream.

In summary, we recommend the following protocol for comparing habitat preferences. (1) Conduct comparisons with regard to bivariate depth and velocity distributions, not with regard to preferences. (2) Use resampling methods to obtain confidence bounds on indexes (such as preference) with unknown distributional properties. (3) Define a magnitude of preference change that is ecologically significant in terms of its effect on the predicted WUA–streamflow relationship.

HENRIETTE I. JAGER

Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831-6036, USA

EDMUND J. PERT

Department of Fisheries and Wildlife Sciences
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061, USA

References

Beecher, H. A., J. P. Carleton, and T. H. Johnson. 1995. Utility of depth and velocity preferences for predicting steelhead parr distribution at different flows. Transactions of the American Fisheries Society 124: 935–938.

Pert, E. J., and D. C. Erman. 1994. Habitat use by adult rainbow trout under moderate artificial fluctuations in flow. Transactions of the American Fisheries Society 123:913–923.

Quinn, J. F., and A. E. Dunham. 1983. On hypothesis testing in ecology and evolution. American Naturalist 122:602–617.

Roughgarden, J. 1983. Competition and theory in community ecology. American Naturalist 122:583–601.

Thomas, J. A., and K. D. Bovee. 1993. Application and testing of a procedure to evaluate transferability of habitat suitability criteria. Regulated Rivers: Research and Management 8:285–294.

Toft, C. A., and P. J. Shea. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. American Naturalist 122:618–625.

Williams, J. 1996. Lost in space: minimum confidence intervals for idealized PHABSIM studies. Transactions of the American Fisheries Society 125:458–465.