Restoration Ecology
THE JOURNAL OF THE SOCIETY FOR ECOLOGICAL RESTORATION

RESEARCH ARTICLE

A Mapping Technique to Evaluate Age-0 Salmon Habitat Response from Restoration

Damon H. Goodman,[1,2] Nicholas A. Som,[1] Justin Alvarez,[3] and Aaron Martin[4]

Abstract

To combat decades of anthropogenic degradation, restoration programs seek to improve ecological conditions through habitat enhancement. Rapid assessments of condition are needed to support adaptive management programs and improve the understanding of restoration effects at a range of spatial and temporal scales. Previous attempts to evaluate restoration practices on large river systems have been hampered by assessment tools that are irreproducible or metrics without clear connections to population responses. We modified a demonstration flow assessment approach to assess the realized changes in habitat quantity and quality attributable to restoration effects. We evaluated the technique's ability to predict anadromous salmonid habitat and survey reproducibility on the Trinity River in northern California. Fish preference clearly aligned with a priori designations of habitat quality: the odds of observing rearing Chinook or coho salmon within high-quality habitats ranged between 10 and 16 times greater than low qualities, and in all cases the highest counts were associated with highest quality habitat. In addition, the technique proved to be reproducible with "substantial" to "almost perfect" agreement of results from independent crews, a considerable improvement over a previous demonstration flow assessment. These results support the use of the technique for assessing changes in habitat from restoration efforts and for informing adaptive management decisions.

Key words: **demonstration flow assessment, habitat modeling, habitat validation, restoration effectiveness monitoring, salmon.**

Introduction

Global biodiversity is decreasing at alarming rates from anthropogenic degradation with few indications of improvement (Butchart et al. 2010). Freshwater ecosystems are among the most impacted creating an over-riding conservation priority (Malmqvist & Rundle 2002; Dudgeon et al. 2006). Recently, substantial levels of restoration have targeted freshwater environments (Osmerod 2003; Nakamura et al. 2006). Annual restoration expenditures exceed one billion dollars (Bernhardt et al. 2005; Brooks & Lake 2007), but surprisingly little effort has been apportioned to evaluate effects of restoration activities, particularly in large river systems (Palmer et al. 2007). Evaluating restoration outcomes is a key criterion for ecologically successful restoration and currently hampers progress of the field (Palmer et al. 2005).

Demonstration flow assessments (DFA) are a class of studies that evaluate habitats regarding a variety of management actions including establishment of environmental flow levels, stream channel manipulation, and other river restoration actions (Annear et al. 2004). These studies have been applied to evaluate habitats of a wide range of riverine species including frogs, toads, turtles, lampreys, and salmonids (Goodman et al. 2009). Initial applications of DFA compared conditions across streamflows via visual observations with little or no empirical measurements over large study areas and with low levels of effort. Panels of experts would apply professional judgment at study sites and develop a consensus-based numerical rating that was used as to inform streamflow management decisions.

Recently, DFA became critiqued for reliance on professional judgment and potential subjectivity that could vary across survey teams (Annear et al. 2004; Gard 2009). In an attempt to reduce the subjectivity of DFA, Railsback and Kadvany (2008) suggested altering DFA to include a judgment-based decision-making framework for classifying habitat during field surveys. Goodman et al. (2009) evaluated DFA as presented by Railsback and Kadvany (2008) and found the results to be irreproducible, despite the modifications. Furthermore, criticism extended to the metrics evaluated by DFA and questioned the ability of results to evaluate differences truly in habitat quality (Gard 2009). However, we believe that if DFA could be modified to measure metrics of interest accurately and be reproducible, it could be a useful and efficient tool for quantifying habitat response to a broad suite of restoration actions.

Herein, we describe a modified DFA addressing the primary shortcomings of previous applications, and evaluate its performance via an example from the Trinity River, a large regulated

[1]U.S. Fish and Wildlife Service, 1655 Heindon Road, Arcata, CA U.S.A.
[2]Address correspondence to D. H. Goodman, email: damon_goodman@fws.gov
[3]Hoopa Valley Tribe, Fisheries Department, Hoopa, CA, U.S.A.
[4]Yurok Tribe, Fisheries Department, Willow Creek, CA, U.S.A.

doi: 10.1111/rec.12148

river system in northwestern California, U.S.A. Modifications include (1) the development of clear and measureable habitat definitions derived from site-specific habitat use studies; (2) measuring habitat parameters during field surveys (Oswood & Barber 1982); and (3) implementing stringent global positioning system (GPS) survey techniques (Dauwalter et al. 2005; Radomski et al. 2011). We evaluated performance by (1) determining if the modified DFA was able to assess habitat quality via comparison with empirical fish-data and (2) evaluating if derived habitat maps were more reproducible than previous methods and sufficient to evaluate anticipated restoration effects.

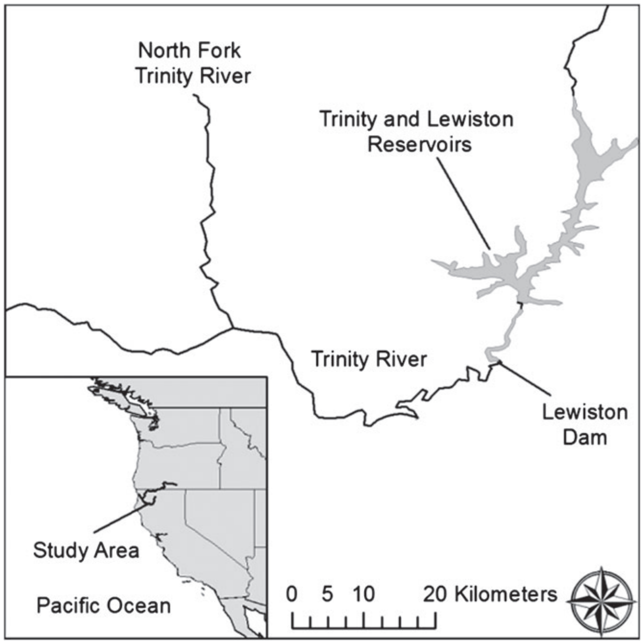

Figure 1. Map of the Trinity River restoration reach. The restoration reach spans from the base of Lewiston Dam to the confluence with the North Fork Trinity River.

Methods

Study Area

The Trinity River is the focus of a large-scale restoration program with management occurring on an annual basis (Fig. 1). The river was permanently altered in the 1960s with the construction of two dams that facilitate water export. The dams led to habitat degradation via removal of streamflow variability and magnitude, as well as interruption of coarse sediment and large wood supplies (U.S. Fish and Wildlife Service and Hoopa Valley Tribe 1999). In addition, the dams blocked anadromous fish access to approximately one-quarter of the 7,700 km^2 watershed. The loss and degradation of age-0 winter and early spring rearing habitat (hereafter denoted as age-0 habitat) prompted drastic declines in two salmon species present in the river, *Onchorynchus tshawysha* (Chinook salmon) and the federally endangered *O. kisutch* (coho salmon).

A large-scale effort was initiated in 2000 to improve conditions over a 64-km reach between Lewiston Dam and the North Fork Trinity River (hereafter referred to as the restoration reach; Locke et al. 2008). Stakeholders anticipate that restoration will benefit Chinook and coho salmon populations from substantial improvements in age-0 habitat area (approximately 400% increase as an interim goal). Restoration actions included increases in annual streamflow, seasonal streamflow variability, coarse sediment and large wood additions, and mechanical channel rehabilitation at 44 locations (Barinaga 1996). The first channel rehabilitation site was completed in 2005, and approximately half of the channel rehabilitation sites have been completed. Restoration actions are applied and modified on an annual basis through an adaptive management framework (Holling 1978) intended to improve future restoration based on performance of prior actions and improvements in restoration science. To facilitate this process, rapid feedback on age-0 habitat response is used to influence upcoming restoration planning and maximize benefits (Maddock 1999).

Modified DFA

Habitat definitions. We developed definitions of age-0 habitat through analysis of existing datasets that linked Chinook and coho salmon to microhabitat conditions. The analysis included observations of rearing Chinook and coho salmon at 2,124 locations within the restoration reach collected between 2003 and 2006 at streamflows ranging from 8.5 to 239.8 m^3/s. Each observation included physical parameters commonly used to describe age-0 habitat (Gard 2006; Hardy et al. 2006) including depth, mean column velocity (velocity), and distance to in-water escape cover (cover). Divers swam upstream while counting fish to minimize disturbance following Hillman et al. (1992). Fish counts were divided into size classes corresponding to two developmental phases to incorporate changes in behavior with growth: fry ($\leq$50 mm fork length) and presmolt ($>$50 mm fork length). Velocities were measured using hand-held flow meters on top-setting rods adapted from the study by Buchanan and Somers (1976). Physical parameters were summarized as habitat suitability indices to represent quality of a given parameter by species and size class (Bovee 1982).

We used habitat suitability indices and professional judgment to guide selection of parameter ranges and create habitat definitions (Table 1). Chinook salmon age-0 habitat included three qualities: (1) Optimal areas meeting depth, velocity, and cover criteria; (2) Suitable areas that meet either depth and velocity criteria or cover criteria; and (3) Low Quality area meeting none of the criteria. Observations of coho salmon were coupled to combined depth, velocity, and cover criteria, a behavior found in other river systems (McMahon & Hartman 1989). Therefore, we defined two qualities for coho salmon age-0 habitat, namely: (1) Optimal areas as defined above and (2) Low Quality for all areas not rated as Optimal.

Habitat Mapping Technique. To map habitat areas, physical parameters were located by measuring depth and velocity

Table 1. Criteria for Chinook and coho salmon age-0 habitat definitions and fish observed within each parameter range.

Fish stage	*Parameter*	*Range*	*Chinook*		*Coho*	
			n	*% obs.*	*n*	*% obs.*
Fry	Depth	>0–0.6 m	1034	79	321	87
	Velocity	0–0.15 m/s	971	69	299	80
	Cover	0–0.6 m	1037	84	321	92
Presmolt	Depth	>0–1 m	549	85	256	95
	Velocity	0–0.24 m/s	490	81	248	86
	Cover	0–0.6 m	550	79	257	96

Fry are fish ≤50 mm fork length and presmolt >50 mm fork length. Velocity was measured as mean column velocity and Cover as distance to in-water escape cover. Microhabitat use information was summarized as the number (n) and percentage (% obs.) of fish observations within each parameter range. Within each fish stage, Optimal age-0 habitat areas simultaneously met all parameter ranges whereas Sutiable met either Cover or Depth and Velocity ranges.

using hand-held flow meters on top-setting rods and outlining in-water escape cover. Geo-referenced locations were recorded to define the perimeter of each habitat area using a Pro XH GPS, Zephyr Geodetic™ antenna, and TerraSync™ survey software (Trimble Navigation Limited, Sunnyvale, CA, U.S.A.) on a tablet computer. A TruPulse™ 360 B laser rangefinder (Laser Technologies Inc., Centennial, CO, U.S.A.) was used to collect offsets between areas with good and poor GPS reception. Geospatial data were differentially corrected with H-Star carrier and code processing using proximal base file providers in Pathfinder Office software (Trimble Navigation Limited). Horizontal geo-positional error was assumed to be ±0.2 m from GPS (Wing et al. 2008) and ±0.3 m from Laser Rangefinder (Institute of Forest Ecosystem Research 2011). Small habitat areas (<2 m^2) were not mapped to reduce geo-spatial error following O'Connor and Rahel (2009). Exclusion of areas less than 2 m^2 may lead to a slight under estimate of age-0 habitat availability in study reaches.

Fish Use of Mapped Habitat Areas

We used relative abundance (fish counts) to assess if mapped age-0 habitat was associated with fish presence. This was tested by mapping age-0 habitat and then sampling mapped habitat areas with fish counts. In 2008, we mapped age-0 habitat in March, and counted fish later in March and April targeting specific developmental phases. Mapping and fish surveys were conducted at 8.4 m^3/s streamflow as measured by USGS gage 11525500. Age-0 habitat was mapped at a 0.5-km section in proximity to upstream spawning grounds. Next, maps were divided into fish sample units ranging from 12 to 31 m^2, a size range selected to reduce sample variance while facilitating efficient fish counts. Fish sample units were created from maps by: (1) preserving units within the desired size range; (2) dividing units greater than 31 m^2 into smaller units, and (3) removing all units smaller than 12 m^2. We applied a random sample stratified by age-0 habitat quality, targeting 20 units per quality. Fish sample unit perimeters were located using high-resolution aerial photography and GPS. In each unit, a diver would count fish by swimming upstream with a single-pass (Hillman et al. 1992).

We analyzed the presence and relative abundance of fish to assess their relationship with age-0 habitat quality with a suite of generalized linear models (GLM; Zuur et al. 2009). GLM models, and their corresponding distributions and link functions, were chosen based on the characteristics of the observed data and assumptions necessary to fit the respective models. When the average fish counts were small and the number of observed positive counts, relative to the sample size, was also small, analyses focused on the probability of fish observance, using a binomial distribution with a logit link function, where:

$$Y_i \sim \mathrm{Binomial}\,(1,\ \pi)$$

$$\log\left(\frac{\pi}{1-\pi}\right) = \ \mathrm{offset}\,(\log\,(\mathrm{area})) + \beta_0 + \beta_1 \times \mathrm{habitat}$$

where Y is 1 if fish were observed and 0 if not, π represents the probability of observing a positive count, area represents the area (m^2) of each age-0 habitat area, habitat is the categorical habitat qualities, β_0 and β_1 are regression coefficients, and offset indicates that the response is scaled by area, akin to density. In cases where samples contained a sufficient number of positive counts, hurdle models were used. Hurdle models are composite models that first assess the relationship between explanatory variables and the probability of positive counts, and then given positive counts (crossing the hurdle), assess the relationship between explanatory variables and the size of the counts (Zuur et al. 2009). All count distributions exhibited variance-to-mean relationships that were greater than expected (overdispersion) under a Poisson distribution. Accordingly, we opted to model the count component of the hurdle models as negative binomial (NB) random variables, with a log link, where:

$$Y_i \sim \mathrm{NB}\,(\mu, \theta)$$

$$\log\,(\mu) = \mathrm{offset}\,(\log\,(\mathrm{area})) + \beta_0 + \beta_1 \times \mathrm{habitat}$$

where $\mu = \mathrm{E}(Y)$ is the mean, Y is the fish count, θ scales the variance-to-mean relationship, and all other terms are as defined previously. We note that in the hurdle case, the NB distribution is truncated above zero. All models were fit using R statistical software (R Core Team 2011), with the hurdle models utilizing the Political Science Computational Laboratory (*pscl*) contributed package (Zeileis et al. 2008).

Reproducibility: Inter-rater Agreement

To assess the reproducibility of the modified DFA and to compare its performance relative to the previous DFA technique, we computed kappa (κ) statistics, which are used to quantify the similarity of categorical scores assigned via multiple raters (Gwet 2008). κ statistics were originally developed by Cohen (1960) to adjust mean agreement among raters by the expected agreement (i.e. agreement expected by random chance alone). In the context of categorical mapping, several issues require adjustments over the standard κ-type statistics. First, not all disagreements among raters are equal. Some age-0 habitat quality pairs can be more similar than others. For example, one crew assigning an age-0 habitat quality of Optimal is more similar to a second crew assigning Suitable than if the second crew had assigned Low Quality. Second, because of GPS error and independent GPS co-location, partial credit should be applied for near-cell agreement (Hagen-Zanker et al. 2005). Finally, raster-type categorical maps will most certainly exhibit spatial autocorrelation which can lead to biased κ values (Hagen-Zanker 2009). The improved fuzzy κ statistic (Hagen-Zanker 2009) addresses all the three issues described above and was used to quantify the strength of agreement among assessment crews (reproducibility). We assessed the quality of the fuzzy κ values according to Landis and Koch (1977).

For the modified DFA, improved fuzzy κ statistics were calculated at seven 400-m sample units within the restoration reach. Study units were selected from the restoration reach using generalized random tessellation stratified sampling (Stevens & Olsen 2004), which generally balances the units spatially within the restoration reach. Field surveys were conducted between July and October 2009. To reduce the influence of streamflow variation on comparisons, all samples were collected while the water release from Lewiston Dam was stable at 12.7 m^3/s with minimal tributary accretions. A single crew consisted of three people that measured depth, velocity, and to-cover distances within sampled units and rated polygons according to the age-0 habitat quality criteria described above. Two crews independently surveyed each sampled unit, and we assessed the similarity of each crew's ratings.

We calculated improved fuzzy κ statistics for the previous DFA using replicate survey data collected by Goodman et al. (2009) for comparison against the modified DFA. In the previous DFA, habitat had only two quality rankings: Suitable or Low Quality. Sample units were all within the restoration reach and data were collected in 2006. Water release from Lewiston Dam was stable at 12.7 m^3/s, the same streamflow as the modified DFA data collection. Sample unit lengths ranged from 1.16 to 1.63 km, but were divided into 400-m units to more align with this study. Several 400-m units did not contain age-0 habitat and were removed from analysis.

Crew-defined polygons will differ in size and orientation depending on how crews define the boundaries of regions that fall within the potential habitat classifications. To evaluate agreement between crews for both the modified and previous DFA, we first laid a 1-m^2 gridded mesh over each of the sites, and then computed agreement based on the centroids from each grid cell.

The improved fuzzy κ statistic requires two user-defined inputs: a similarity matrix among potential quality levels and a distance-decay function that describes the degree to which a cell belongs to the categories found near it. All values within the similarity matrix range between 0 and 1, where 1 indicates perfect agreement and 0 represents absolute disagreement. For the modified DFA age-0 habitat categories, the Optimal classification represents the case where both the depth and velocity criteria and cover criteria thresholds are met, so we opted for cell values of 0.5 when one crew rated a cell as Optimal and the other rated the same cell as either depth and velocity or cover. All other disagreements among crew ratings received the full disagreement value of 0. For the previous DFA technique, with only two levels, the similarity matrix consisted of only full agreement or full disagreement. Based on the discretized set of distances induced by our mesh, we opted to use an exponential decay function that resulted in high correlation within the first several meters and nonlinearly decayed to near zero correlation at distances approaching 25 m. More information on the similarity matrices, distance decay functions, and the improved fuzzy κ statistic can be found in the study by Hagen-Zanker (2009).

Results

For brevity, we present results below for the case of Chinook salmon fry, and note that results for Chinook salmon presmolts and both life stages of coho salmon are very similar. Interested readers can find modified DFA results in Appendices S1 and S2, Supporting Information.

Fish Use of Mapped Habitat Areas

There is very strong evidence that the probability of observing Chinook salmon fry and the mean number observed are related to age-0 habitat quality (Table 2). The odds of observing fish in habitats that are at least Suitable (i.e. Suitable or Optimal) is estimated to be almost 16 times higher (95% confidence interval [CI]: 3−85 times higher) than Low Quality habitats. Additionally, fish were observed in every Optimal habitat sample unit. Given that fish were observed, the mean number observed in Suitable habitats compared with Low Quality habitats is estimated to be 3.6 times greater (95% CI: 2−6.8 times higher), and the mean number observed in Optimal habitats is estimated to be 10.7 times greater (95% CI: 5.3−22.3 times higher), after accounting for the size of the habitat units.

Reproducibility: Inter-rater Agreement

All strength of agreement values for the modified DFA were higher than any from the previous DFA and demonstrate improved reproducibility. The improved fuzzy κ statistics for the seven modified DFA sites ranged from 0.624 to 0.828 with all values representing "substantial" to "almost perfect" strengths of agreement (Table 3). Figure 2 provides an example of spatial differences in modified DFA surveys from independent crews. For the seven previous DFA sites, improved fuzzy κ ranged from 0.149 to 0.344, with all values representing "Slight" or "Fair" agreement.

Table 2. Parameter estimates, standard errors, and tests of significance for the model evaluating the relationship between age-0 habitat categories and Chinook salmon fry use.

Parameter	*Estimate*	*SE*	*z value*	*p value*
Count model				
Low Quality	−0.32	0.278	−1.154	0.248
Suitable	1.302	0.317	4.103	<0.001
Optimal	2.367	0.352	6.721	<0.001
Log (θ)	0.175	0.193	0.905	0.365
Zero hurdle model				
Low Quality	0.539	0.476	1.133	0.257
Suitable or Optimal	2.757	0.863	3.195	0.001

The θ parameter accounts for overdispersion. For the count model, a truncated negative binomial with a log link function was used. For the zero hurdle model, a binomial distribution with a logit link function was used.

Table 3. Mean agreement and improved fuzzy κ statistics measuring reproducibility between crews.

Technique	*Site*	*Mean agreement*	*Improved fuzzy* κ
Modified	1	0.988	0.778
	2	0.987	0.742
	3	0.962	0.660
	4	0.971	0.624
	5	0.989	0.828
	6	0.976	0.657
	7	0.981	0.671
Previous	1	0.976	0.149
	2	0.981	0.285
	3	0.988	0.178
	4	0.981	0.179
	5	0.990	0.157
	6	0.996	0.237
	7	0.989	0.344

Results consist of age-0 habitat maps surveyed using the modified DFA technique (“modified”) and a previous method (“previous”). Crews assessed Chinook salmon fry habitat at seven sites under each technique, although the sites were not the same among technique applications or years. Under the modified technique, crews were to classify among four habitat categories, but under the previous technique only two categories were utilized. Landis and Koch (1977) associated κ values to the following levels of agreement: Almost perfect (0.81–1.00), Substantial (0.61–0.80), Moderate (0.41–0.60), Fair (0.21–0.40), Slight (0.00–0.20), and Poor (<0.00).

Discussion

We modified a stream habitat assessment technique based on shortcomings from previous applications and demonstrated improvements. The modifications of DFA included defining habitat quality through analysis of microhabitat use data, measuring physical parameters during surveys and geo-referencing measurements to develop maps. Improvements included alignment of target species with a priori designations of mapped habitat quality and superior reproducibility compared with a previous DFA.

Reproducibility is a central tenant of scientific study and essential for metrics evaluating change based on restoration actions. At the very least, to observe restoration-based changes in habitat quality, survey techniques must exhibit levels of among-crew error (in addition to other sources of variation) lower than the effect size of restoration actions. We made direct comparisons between a previous and modified DFA and found clear improvements in the reproducibility. Although reproducibility has been a problem for many commonly applied aquatic habitat assessment protocols (Whitacre et al. 2007; Roper et al. 2010), this work shows that techniques and protocols can be altered to achieve reproducible maps of habitat quality.

We evaluated reproducibility via the improved fuzzy κ statistic, and one issue that can plague κ type statistics is extremely rare or dominant levels of potential classification categories. For instance, the proportion of locations classified as Optimal

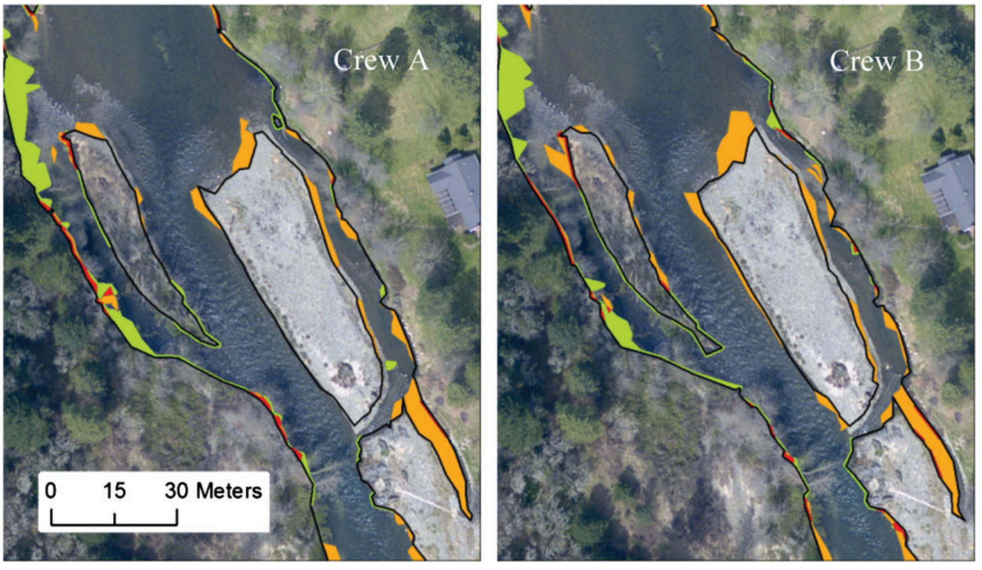

Figure 2. An example of Chinook salmon fry habitat map developed by independent crews using the modified DFA at a complex multi-channel sample unit. Green indicates areas that meet cover criteria only, orange areas meet depth and velocity criteria only, and red areas meet all habitat criteria. Black lines indicate the edge of the wetted channel.

fry habitat by the first crew at Site 1 was only 1.3%, whereas the same crew classified 87.9% of the locations as Low Quality for fry. If the true prevalence of some potential classification categories is either very high or low (or some of both), κ type statistics may not reflect the large extent of agreement among raters (Gwet 2008). As such, the improved fuzzy κ values that we report may be conservatively low, and hence bolster our claim that the habitat assessment methods are indeed likely reproducible. Future research should expand on this assessment through inclusion of more than two crews to further refine our understanding of the reproducibility of the modified DFA.

One of the primary advantages to DFA is its ability to assess habitat with a low level of effort. The modifications we developed require a higher level of effort than the previous DFA, particularly in the field effort, and require mapping grade GPS survey equipment. However, we found the modified DFA to require a much lower level of effort than other large river habitat assessment techniques (e.g. two-dimensional hydrodynamic habitat modeling). For example, the modified DFA mapping survey at a 400-m sample unit required 18–90 man-hours to complete field mapping, and 3–4 man-hours for post-processing. In comparison, two-dimensional hydrodynamic modeling of a 400-m sample unit with 0.5 m prediction resolution in the same study reach following methods as described by Wright et al. (2014) required 1,270 man-hours for field data collection and 960 hours for post-processing. The two-dimensional hydrodynamic models had the additional requirement of LiDAR base topography not included in the time estimates. However, the products from these techniques are drastically different and serve different purposes for restoration planning and assessment. The modified DFA provides a discrete assessment of habitat availability, whereas two-dimensional models additionally provide a framework to evaluate the consequences of various management actions (Hardy 1998).

Although we have addressed and improved several of the main critiques of the previous DFA methods, the methods have also been criticized for other issues related to the construction of streamflow-to-habitat relationships. Namely, Gard (2009) cited the need to assume linear relationships between observed streamflows, and the inability to extrapolate habitat amounts beyond those measured, as further weaknesses of the DFA technique. Although the details are not germane to the DFA improvements described herein, we do want to note that a suite of functions can be used to fit both linear and nonlinear flow-to-habitat curves, and that nonlinear curves do not impose a linear assumption between measured values. We further note that any method seeking to estimate relationships beyond the points of measurement should be applied cautiously, if at all, and substantial assumptions plague all streamflow-to-habitat construction methodologies.

Most restoration efforts lack formal evaluation of achievements (or failures), hampering progress in the developing field of restoration science (Brooks & Lake 2007). Whether efforts aim to improve the habitat of flora (Angelstam 1998) or fauna (Feunteun 2002), measuring response is a critical aspect of restoration monitoring. By identifying target species and implementing monitoring tools, such as the modified DFA, we can refine our understanding of how restoration affects habitat at a variety of spatial scales. Modified DFA can be applied using study designs such as before-after-control-impact to test design hypotheses of habitat response from restoration activities (Roni et al. 2005). Furthermore, the spatially explicit nature of the data can be used to evaluate the responses of specific features or track site evolution over time. This information can then be used by restoration planners to refine hypotheses for future management actions (Beechie et al. 2014). We developed and evaluated the modified DFA for two salmonid species in the Trinity River; however, with modifications to habitat definitions, we feel that the basic approach could be expanded to other fish species or river systems.

Implications for Practice

- Measuring changes in habitat can be used to evaluate restoration performance and inform future actions before population response occurs.
- Habitat metrics should have a demonstrated link to the specific physical requirements of species and life stages of interest to restoration programs.
- Habitat assessment protocols should utilize stringent survey techniques that prove reproducible.

Acknowledgments

Many individuals contributed to the success of this project. In particular, we would like to thank the field crew members for their support. Joe Polos and Nick J. Hetrick of U.S. Fish and Wildlife Service provided guidance for the study design and execution, and Jim Haas of the U.S. Fish and Wildlife Service National Training and Conservation Center for editorial support. Finally, we would like to thank Susan Yates, Robin DeBruyne, and the blind reviewers of *Restoration Ecology* for their thoughtful input. This project was funded by the Trinity River Restoration Program. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

LITERATURE CITED

Angelstam, P. K. 1998. Maintaining and restoring biodiversity in European boreal forests by developing natural disturbance regimes. Journal of Vegetation Science **9:**593–602.

Annear, T., I. Chisholm, H. Beecher, A. Locke, P. Aarrestad, N. Burkardt, et al. 2004. Instream flows for riverine resource stewardship, revised edition. Instream Flow Council, Cheyenne, Wyoming.

Barinaga, M. 1996. A recipe for river recovery? Science **273:**1648–1650.

Beechie, T., G. Pess, H. Imaki, A. Martin, J. Alvarez, and D. H. Goodman. 2014. Comparison of potential increases in juvenile salmonid rearing habitat capacity among alternative restoration scenarios, Trinity River, California. Restoration Ecology: DOI: 10.1111/rec.12131.

Bernhardt, E. S., M. A. Palmer, J. D. Allan, G. Alexander, K. Barnas, S. Brooks, et al. 2005. Synthesizing U.S. river restoration efforts. Science **308:**636–637.

Bovee, K. D. 1982. A guide to stream habitat analysis using the instream flow incremental methodology. Instream Flow Informational Paper No. 12, FWS-OBS 82/26, U.S. Fish and Wildlife Service, Washington, D.C.

Brooks, S. S., and P. S. Lake. 2007. River restoration in Victoria, Australia: change is in the wind, and none too soon. Restoration Ecology **15:**584–591.

Buchanan, T.J., and W.P. Somers. 1976. Discharge measurements at gaging stations. Techniques of Water-Resources Investigations of the USGS, U.S. Geological Survey, Washington, D.C.

Butchart, S. H. M., M. Walpole, B. Collen, A. vanStrien, J. P. W. Scharlemann, R. E. A. Almond, J. E. M. Baillie, B. Bomhard, C. Brown, and J. Bruno. 2010. Global biodiversity: indicators of recent declines. Science **328:**1164–1168.

Cohen, J. 1960. Coefficient of agreement for nominal scales. Educational and Psychological Measurement **20:**37–46.

Dauwalter, D. C., W. L. Fisher, and K. C. Belt. 2005. Mapping stream habitat with a global positioning system: accuracy, precision, and comparison with traditional methods. Environmental Management **37:**271–280.

Dudgeon, D., A. H. Arthington, M. O. Gessner, Z.-. I. Kawabata, D. J. Knowler, C. Lévêque, et al. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews **81:**163–182.

Feunteun, E. 2002. Management and restoration of European eel population (*Anguilla anguilla*): an impossible bargain. Ecological Engineering **18:**575–591.

Gard, M. 2006. Modeling changes in salmon spawning and rearing habitat associated with river channel restoration. International Journal of River Basin Management **4:**201–211.

Gard, M. 2009. Demonstration flow assessment and 2-D modeling: perspectives based on instream flow studies and evaluation of restoration projects. Fisheries **34:**320–329.

Goodman, D. H., A. C. Martin, P. P. Petros, and J. Klochak. 2009. Judgment based habitat mapping on the Trinity River, 2006. Arcata Fisheries Technical Report Number TR2009-12, U.S. Fish and Wildlife Service, Arcata, California.

Gwet, L. L. 2008. Computing inter-rater reliability and its variance in the presence of high agreement. British Journal of Mathematical and Statistical Psychology **61:**29–48.

Hagen-Zanker, A. 2009. An improved fuzzy kappa statistic that accounts for spatial autocorrelation. International Journal of Geographical Information Science **23:**61–73.

Hagen-Zanker, A., S. Geertman, and I. Uljee. 2005. Further developments of a fuzzy set map comparison approach. International Journal of Geographical Information Science **19:**769–785.

Hardy, T. B. 1998. The future of habitat modeling and instream flow assessment techniques. Regulated Rivers: Research and Management **14:**405–420.

Hardy, T. B., T. Shaw, R. C. Addley, and G. E. Smith. 2006. Validation of Chinook fry behavior-based escape cover modeling in the lower Klamath River. International Journal of River Basin Management **4:**169–178.

Hillman, T. W., J. W. Mullan, and J. S. Griffith. 1992. Accuracy of underwater counts of juvenile Chinook salmon, coho salmon and steelhead. North American Journal of Fisheries Management **12:**598–603.

Holling, C. S. 1978. Adaptive environmental assessment and management. John Wiley and Sons, Chichester, U.K.

Institute of Forest Ecosystem Research. 2011. Field-Map technology for area measurement. Technical report of Field-Map Stork set-up testing for area measurement, Strašice, Czech Republic.

Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics **33:**159–174.

Locke, A., C. Stalnaker, S. Zellmer, K. Williams, H. Beecher, T. Richards, C. Robertson, A. Wald, A. Paul, and T. Annear. 2008. Integrated approaches to riverine resource management: case studies, science, law, people, and policy. Instream Flow Council, Cheyenne, Wyoming.

Maddock, I. 1999. The importance of physical habitat assessment for evaluating river health. Freshwater Biology **41:**373–391.

Malmqvist, B., and S. Rundle. 2002. Threats to the running water ecosystems of the world. Environmental Conservation **2:**134–153, DOI: 10.1017/S0376892902000097.

McMahon, T. E., and G. F. Hartman. 1989. Influence of cover complexity and current velocity on winter habitat use by juvenile coho salmon (*Onchorhynchus kisutch*). Canadian Journal of Fisheries and Aquatic Sciences **46:**1551–1557.

Nakamura, K., K. Tockner, and K. Amano. 2006. River and wetland restoration: lessons from Japan. BioScience **56:**419–429.

O'Connor, R. R., and F. J. Rahel. 2009. A patch perspective on summer habitat use by brown trout *Salmo trutta* in a high plains stream in Wyoming, USA. Ecology of Freshwater Fish **18:**473–480.

Osmerod, S. J. 2003. Restoration in applied ecology: editor's introduction. Journal of Applied Ecology **40:**44–50.

Oswood, M. W., and W. E. Barber. 1982. Assessment of fish habitat in streams: goals, constraints, and a new technique. Fisheries **7:**8–11.

Palmer, M. A., E. S. Bernhardt, J. D. Allan, P. S. Lake, G. Alexander, S. Brooks, et al. 2005. Standards for ecologically successful river restoration. Journal of Applied Ecology **42:**208–217.

Palmer, M., J. D. Allan, J. Meyer, and E. S. Bernhardt. 2007. River restoration in the twenty-first century: data and experiential knowledge to inform future efforts. Restoration Ecology **15:**472–481.

R Core Team. 2011. R: A Language and Environment for Statistical Computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing, Vienna, Austria (available from http://www.R-project.org/).

Radomski, P., K. Wolzeschke, K. Carlson, and D. Perleberg. 2011. Reproducibility of emergent plant mapping on lakes. North American Journal of Fisheries Management **31:**144–150.

Railsback, S. F., and J. Kadvany. 2008. Demonstration flow assessment: judgment and visual observation in instream flow studies. Fisheries **33:**217–227.

Roni, P., M. C. Liermann, C. Jordan, and E. A. Steel. 2005. Steps for designing a monitoring and evaluation program for aquatic restoration. Pages 13–34 in P. Roni, editor. Monitoring stream and watershed restoration. American Fisheries Society, Bethesda, Maryland.

Roper, B. B., J. M. Buffington, S. Bennett, S. H. Lanigan, E. Archer, and S. T. Downie. 2010. A comparison of the performance and compatibility of protocols used by seven monitoring groups to measure stream habitat in the Pacific Northwest. North American Journal of Fisheries Management **30:**565–587.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association **99:**262–278.

U.S. Fish and Wildlife Service and Hoopa Valley Tribe. 1999. Trinity River flow evaluation, final report. U.S. Fish and Wildlife Service, Arcata, California.

Whitacre, H. W., B. B. Roper, and J. L. Kershner. 2007. A comparison of protocols and observer precision for measuring physical stream attributes. Journal of the American Water Resources Association **43:**923–937.

Wing, M. G., A. Eklund, J. Sessions, and R. Karsky. 2008. Horizontal measurement performance of five mapping-grade global positioning system receiver configurations in several forested settings. Western Journal of Applied Forestry **23:**166–171.

Wright, K. A., D. H. Goodman, N. A. Som, and T. B. Hardy. 2014. Development of two-dimensional hydraulic models to predict distribution of *Manayunkia speciosa* in the Klamath River. Arcata Fisheries Technical Report Number TR 2014-19, U.S. Fish and Wildlife Service, Arcata, California.

Zeileis, A., C. Kleiber, and S. Jackman. 2008. Regression models for count data in R. Journal of Statistical Software **27:**1–25.

Zuur, A. F., E. N. Leno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Fish use of mapped habitat area results for Chinook salmon presmolt, coho salmon fry, and coho salmon presmolt.

Appendix S2. Mean agreement and improved fuzzy κ statistics measuring reproducibility of modified DFA between crews for Chinook salmon presmolt habitat.