North American Journal of Fisheries Management 28:919–934, 2008

DOI: 10.1577/M06-171.1

[Article]

Detecting the Response of Fish Assemblages to Stream Restoration: Effects of Different Sampling Designs

B. P. BALDIGO*

U.S. Geological Survey, New York Water Science Center, 425 Jordan Road, Troy, New York 12180, USA

D. R. WARREN

Department of Natural Resources, Cornell University, Fernow Hall, Ithaca, New York 14853, USA

Abstract.—Increased trout production within limited stream reaches is a popular goal for restoration projects, yet investigators seldom monitor, assess, or publish the associated effects on fish assemblages. Fish community data from a total of 40 surveys at restored and reference reaches in three streams of the Catskill Mountains, New York, were analyzed a posteriori to determine how the ability to detect significant changes in biomass of brown trout *Salmo trutta*, all salmonids, or the entire fish community differs with effect size, number of streams assessed, process used to quantify the index response, and number of replicates collected before and after restoration. Analyses of statistical power (probability of detecting a meaningful difference or effect) and integrated power (average power over all possible α-values) were combined with before–after, control–impact analyses to assess the effectiveness of alternate sampling and analysis designs. In general, the more robust analyses indicated that biomass of brown trout and salmonid populations increased significantly in restored reaches but that the net increases (relative to the reference reach) were significant only at two of four restored reaches. Restoration alone could not account for the net increases in total biomass of fish communities. Power analyses generally showed that integrated power was greater than 0.95 when (1) biomass increases were larger than 5.0 g/m^2, (2) the total number of replicates ranged from 4 to 8, and (3) coefficients of variation (CVs) for responses were less than 40%. Integrated power was often greater than 0.95 for responses as low as 1.0 g/m^2 if the response CVs were less than 30%. Considering that brown trout, salmonid, and community biomass increased by 2.99 g/m^2 on average (SD = 1.17 g/m^2) in the four restored reaches, use of two to three replicates both before and after restoration would have an integrated power of about 0.95 and would help detect significant changes in fish biomass under similar situations.

Stream restoration has recently become a widespread, large-scale industry in the United States (Bernhardt et al. 2005). The reasons for restoring stream channels vary widely, but stated objectives commonly include enhancement of water quality, riparian vegetation, bank stability, channel stability, natural flows, instream habitat, fish populations, esthetics, and recreation (Bernhardt 2005; Palmer et al. 2005). Although these and closely related goals may be fixed or ephemeral, they typically coalesce stakeholder interests, establish funding sources and levels, and ultimately propel restoration projects to their completion. With a limited number of exceptions (Roni et al. 2002; Pretty et al. 2003; Baldigo et al. 2008a), few investigators have monitored, assessed, or documented the short- or long-term effectiveness of stream restoration projects in achieving declared goals (Palmer et al. 2005). It is unclear whether the scarcity of published information on responses results from a lack of interest, the extra costs for monitoring and analyses, or intricacies of designing and sampling targeted response variables; nevertheless, such scarcity prevents subsequent projects from learning from or improving upon the successes or failures of preceding, sometimes unique, restoration projects. For example, techniques applying natural channel design (NCD) restoration procedures have purportedly been used to enhance channel stability and the health of fish assemblages in streams and rivers across North America for almost 20 years (Rosgen 1994a, 1994b), yet the effects of NCD restoration on fish populations or communities have been reported only recently (Baldigo et al. 2008a, 2008b, this issue). Surprisingly, even when goals (e.g., decreases in bed scour, lateral channel migration, and bank erosion rates or increases in salmonid abundance and community biomass) are well defined, any monitoring of targeted parameters or indices typically is minimal, absent, or incomplete. These deficiencies may hinder the broader acceptance and application of successful restoration principles and techniques (Bernhardt 2005). In fact, monitoring and evaluation at any level have occurred only in roughly

* Corresponding author: bbaldigo@usgs.gov

Received June 21, 2006; accepted July 2, 2007
Published online June 16, 2008

10% of the 37,099 restoration projects conducted in the United States as of July 2004 (Bernhardt et al. 2005).

Several reasons may account for the lack of monitoring and analysis of effectiveness data, especially with respect to fish assemblages. The largest barrier to fish monitoring efforts may be the high interannual variability in fish populations and communities due to natural fluctuations in uncontrolled factors, such as precipitation, streamflow, air temperature, and water temperature, which could mask even large effects caused by major channel alterations. A study design that quantifies natural variations in fish populations or communities and the effects due only to restoration might be very intensive, complex, and expensive. The number and frequency of fish inventories and the sampling methods (study or sampling design), which determine costs and time commitments, may also be difficult to define and implement; therefore, such difficulties might deter efforts to document the effects of stream restoration on local fish assemblages. A lack of knowledge about local fish assemblages and the proper methods for surveying, quantifying, and assessing assemblage condition, and perhaps the lack of a good understanding of local fish populations and communities (given the high variability inherent in wild populations) may also provide a rationale for not attempting to document the potential effects associated with restoration. Costs of conducting fish studies are unlikely to be a major barrier considering that the mean cost of implementing 14,641 restoration projects reported to the National River Restoration Science Synthesis team before December 2005 was over US$800,000 (Bernhardt et al. 2005; Palmer et al. 2005) and the costs of quantitative fishery inventories in small- to intermediate-sized streams can range from $1,000 to $5,000 per year depending on the availability of sampling equipment, the experience of paid and volunteer personnel, the agency or organization that is conducting the work, and the size and complexity of the affected fish community.

Before–after, control–impact (BACI) sampling designs are particularly useful tools for evaluating fish assemblage responses to stream restoration because they address the problem of high natural variability and year-to-year changes in fish populations and community indices. The actual year-to-year differences in index values are termed "absolute" changes or responses so as to distinguish them from net (i.e., BACI) responses. The BACI analyses effectively separate the absolute year-to-year change in an index from the effects caused primarily by an impact or treatment, such as restoration (Stewart-Oaten et al. 1986; Underwood 1994). They quantify the net changes in population and community indices after treatment (restoration) of target reaches by adjusting or correcting direct responses to observed changes in corresponding indices at one or two types of control reaches (stable reference reaches and unstable control reaches). The net response is simply the increase or decrease in the difference between the index value at the treatment reach after restoration and the index value at a control reach. The reference and control reaches are not manually disturbed, but their fish communities may change naturally with interannual fluctuations in local climatic and environmental conditions. Fish communities at stable reference reaches simulate healthy target assemblages for restored treatment reaches, whereas communities at unstable control reaches are comparable to the degraded fish assemblages sometimes encountered at treatment reaches before restoration. In an evaluation of three restored stream reaches in the Catskill Mountains of southeastern New York (Figure 1), Baldigo et al. (2008a) applied a BACI design and found that fish community richness, biomass, and equitability generally increased when assemblages dominated either by sculpins (Cottidae) or by daces (Cyprinidae) and sculpins before restoration were replaced by assemblages dominated by one or more trout species after restoration. Large increases in the density and biomass of one or more trout species were the primary cause for shifts in the structure and function of fish communities at restored reaches (Baldigo et al. 2008b). Though the effects of restoration on resident fish populations and communities were not principal concerns of previous restoration efforts, BACI analyses showed that modified NCD restoration generally benefited fish communities and selected fish populations for as long as 4 years after treatment (Baldigo et al. 2008a, 2008b).

Various a posteriori power analyses that vary the number of streams and reaches, the number of years, the method of quantifying responses, and the statistical tests employed can be used to reanalyze results from Baldigo et al. (2008a; in combination with results from additional study reaches) and illustrate the effects that different sampling strategies and analyses have on the final interpretation of restoration-mediated changes in biomass of brown trout *Salmo trutta*, salmonids, and the entire fish community. Both statistical power and integrated power were estimated, and results of the latter calculation were used to judge the ability of various sampling or analysis strategies to detect significant changes in the three biomass indices. Statistical power is the ability or sensitivity of a statistical test to detect a meaningful difference between two groups if differences truly occur; generally, statistical power should be 0.80 (80%) or

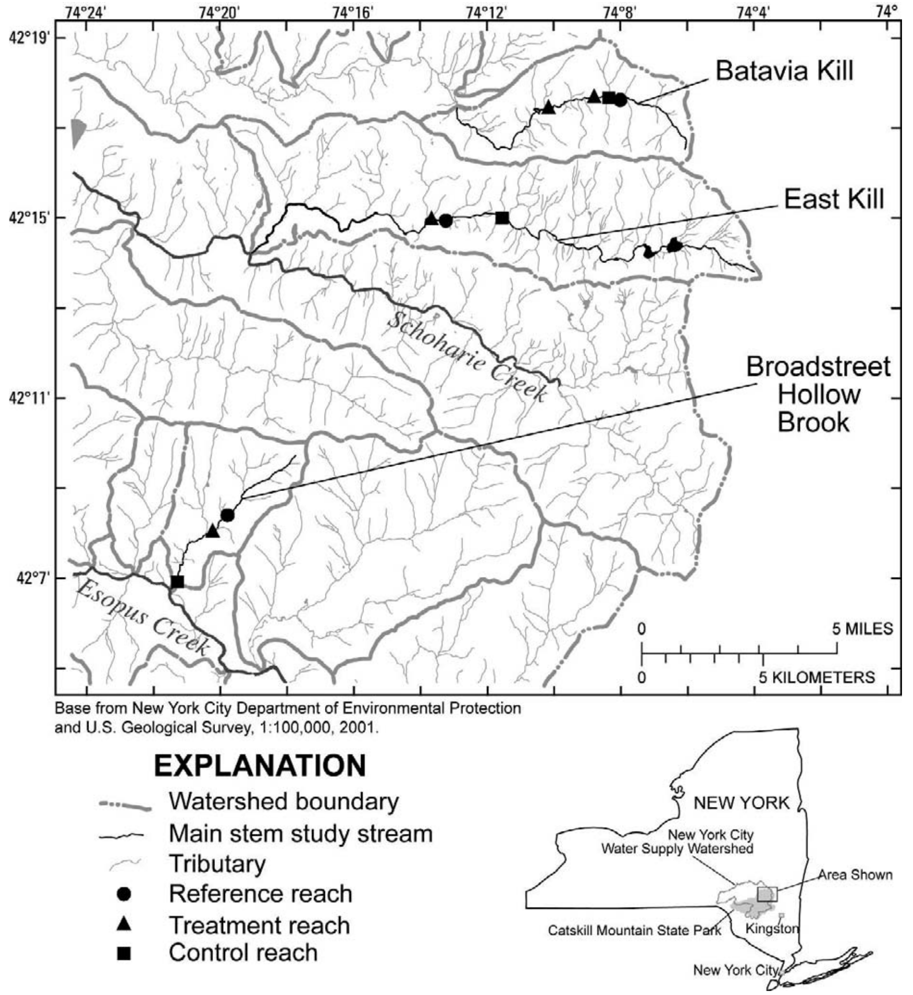

FIGURE 1.—Locations of stream restoration demonstration projects in three streams in eastern Catskill State Park, New York, and west of the Hudson River watershed (modified from Baldigo et al. 2008b).

higher. Specifically, it is the probability of correctly rejecting a false null hypothesis; therefore, it is defined as $1 - \beta$, where β is the probability of committing a type II error. Integrated power is a more-robust measure than statistical power because it quantifies the average power of a test design based on all possible values of α (Lenth 2005). Results of these analyses permit different subsets of real data (selected a posteriori) to be examined in a way that imitates common sampling designs and yields different results and interpretations concerning restoration effects on resident fish assemblages. Considering the small or moderate size of streams and the practical funding limits common to most restoration projects, information provided in this analysis may aid others in developing strategies to assess the effects of stream restoration, especially when the goals of such efforts are to improve the quality of local fish populations and related components of watersheds and stream ecosystems.

This paper summarizes the changes in brown trout, salmonid, and fish community biomass after restoration and the differences in statistical power and integrated power obtained from different combinations of surveys at four treatment reaches and their associated stable reference and unstable control reaches (i.e., one stable reach and one unstable reach per treatment reach). The ability of alternative sampling strategies to detect significant responses in the three indices is assessed. These efforts are only one component of a larger New York City Department of Environmental Protection (NYCDEP) program designed to determine whether

TABLE 1.—Survey numbers for fish community inventories performed at stable reference, unstable control, and restored treatment reaches in three streams of the Catskill Mountains, New York, during midsummer 1999–2004. Bold italic values indicate prerestoration surveys at the treatment reaches. General locations of streams and study reaches are shown in Figure 1.

		Year of survey					
Study stream	Study reach	1999	2000	2001	2002	2003	2004
Batavia Kill	Reference		1	2	3	4	5
	Control		6	7	8	9	10
	Lower treatment[a]		***11***	***12***	13	14	15
	Upper treatment[b]				***16***	17	18
East Kill	Reference		19		20	21	
	Control		22		23	24	
	Treatment[c]		***25***		26	27	
Broadstreet Hollow Brook	Reference	28	29		30	31	32
	Control		33		34	35	
	Treatment[c]	***36***	***37***		38	39	40

[a] Restored after the 2001 survey.
[b] Restored after the 2002 survey.
[c] Restored after the 2000 survey.

and how modified NCD techniques increase bed and bank stability, decrease sediment transport, and sustain or improve the quality of streams that supply drinking water to inhabitants of New York City (Baldigo et al. 2008b).

Methods

Study reaches.—All treatment (pre- or postrestoration), reference, and control reaches were on three streams located 35–42 km west to northwest of Kingston, New York (Figure 1). The drainage areas of treatment reaches ranged from 17 km^2 at the lower Batavia Kill to 45 km^2 at the East Kill. Primary reference reaches were 0.7, 1.9, and 0.1 km upstream from corresponding treatment reaches in Broadstreet Hollow Brook, lower Batavia Kill, and East Kill, respectively. Secondary reference reaches (control reaches) were also upstream from treatment reaches, except at Broadstreet Hollow Brook, where the control reach was located downstream from the treatment reach. Though channel stability, habitat quality, and resident fish communities at control and reference reaches were typically dissimilar, fishery data from both were used as controls to adjust changes in fish indices at restored treatment reaches for purposes of the BACI analyses, as described below. More detailed information on study streams and reaches is given by Baldigo et al. (2008b).

Fish surveys.—Fish inventories were generally done at paired (treatment and reference) reaches in July 1999, 2000, and 2002–2004 in Broadstreet Hollow Brook; July 2000–2004 in the Batavia Kill; and July or August 2000–2003 in the East Kill (Table 1). A second (upper) treatment reach was also surveyed in the Batavia Kill during July 2002–2004. In addition, fish assemblages were surveyed at control reaches on each stream during the first 3–4 years of the study. Restorations of large (300–1,600 m long) project reaches, which encompassed respective treatment reaches in each stream, were done in late summer or early fall of 2000 at Broadstreet Hollow Brook and East Kill, 2001 at lower Batavia Kill, and 2002 at upper Batavia Kill after fish surveys for each summer were completed. Fish survey methods are described in detail by Baldigo et al. (2008b).

Fish responses.—Several statistical tests were performed to assess the effects of restoration on each index under a variety of sampling design scenarios that ranged from simple to complex. The significance of responses was assessed using a range of replicates (total number of reaches, streams, and years), and inclusion of zero, one, or two reference reaches (to adjust data for BACI analyses). The most straightforward way to assess responses is to compare the overlap of 95% confidence intervals (CIs) and visually assess significant changes or differences ($P < 0.05$) in biomass estimates at the same treatment reach 1 year before and 1 year after restoration (Warren and Kraft 2003). Assessments of 95% CIs were analogous to evaluating absolute differences in population indices using one-tailed Student's *t*-tests; however, power analyses were not possible because sample sizes were fixed at one sample before and one sample after restoration at each reach. The Student's *t*-test was then used to assess absolute changes in each average index from individual treatment reaches or from pooled treatment reaches (as many as four) when two or more replicates were collected before and after restoration. The analysis of pooled data using Student's *t*-test was comparable to a one-factor analysis of variance

(ANOVA), but neither method accounts for unique responses within individual treatment reaches and streams or for natural changes in fish communities that may result from regional climatic and environmental fluctuations.

Analysis of data from a solitary treatment reach is typically inadequate to characterize real effects of restoration, because normal climatic variations can affect fish assemblages such that absolute changes in target indices after restoration can be either exaggerated or lessened. The BACI analyses adjust, correct, or standardize index measures at treatment reaches to those at untreated reference reaches (Stewart-Oaten et al. 1986; Underwood 1994) and thus have an advantage over other analyses that assess absolute responses to restoration. Estimates of biomass for brown trout, all salmonids, and the total fish community (indices) were adjusted or standardized to the same index measured at one or both control reaches to establish relative differences, or differentials, for each survey. A differential equals the index value at the treatment reach minus the index value at a control reach during the same survey period (year); it is defined as $(IT_y - IC_y)$, where IT is the index value at the treatment reach and IC is the index value at the corresponding control reach during year y. Any increase or decrease in a differential at the treatment reach after restoration is adjusted or corrected for changes in the same index at the untreated control reach and thus quantifies the net response of that index to restoration. The net response is defined as $(IT_a - IC_a) - (IT_b - IC_b)$, where the subscript a indicates values measured after restoration and the subscript b indicates values measured before restoration. When data for two or more treatment–control reach pairs or survey years in one, or more, streams are available, mean index differentials may be estimated and used to calculate the average net response and to evaluate whether the differentials from before and after restoration differ significantly. For example, for a restoration occurring in 2002, the following hypothesis can be tested: $\{[(IT_{2003} - IC_{2003}) + (IT_{2004} - IC_{2004})]/2\} = \{[(IT_{2000} - IC_{2000}) + (IT_{2001} - IC_{2001})]/2\}$. The average net response (change) in the index due to restoration is calculated as $\{[(IT_{2003} - IC_{2003}) + (IT_{2004} - IC_{2004})]/2\} - \{[(IT_{2000} - IC_{2000}) + (IT_{2001} - IC_{2001})]/2\}$.

Different statistical tools can be used to evaluate the specific effects of restoration within one treatment reach (or stream) or the general effects of a given restoration method on many streams within a particular region. The net responses for each of the three indices were compared by BACI analyses (1) using Student's t-tests to assess effects within each stream or among all streams (by assuming that streams are alike and ignoring temporal trends) and (2) using a mixed two-factor ANOVA across all four streams and choosing either a reference or control reach (balanced design) or both reference and control reaches (asymmetrical design) for each treatment reach. The example hypothesis above illustrates a balanced design, whereas an asymmetric analysis design would be described by the hypothesis $\{[(IT_{2003} - IC_{2003}) + (IT_{2003} - IR_{2003})]/2\} = \{[(IT_{2000} - IC_{2000}) + (IT_{2000} - IR_{2000})]/2\}$, where IR is the index value at a reference reach. The asymmetric design can double the number of replicates used to calculate mean differentials and net effects within a restored reach using data from only one prerestoration survey and one postrestoration survey. These results, however, can be biased if the increase in sample size relies solely on pseudoreplicates derived from an individual treatment reach and a small number of sampling years. Factors for the 2×2, 2×3, and 2×4 ANOVA designs were the fixed effect of period (before and after restoration) and the random effect of treatment reach (2–4 streams). Six prerestoration replicates (combining years and reaches) and 10 postrestoration replicates were used when only one control reach (i.e., either stable reference or unstable control) was related to each treatment reach. For asymmetric analyses, in which both reference and control reaches were related to each treatment reach, 11 prerestoration replicates and 17 postrestoration replicates were used. For each index, the analysis tested the (1) net differences before versus after restoration, (2) magnitude of the differentials among the four treatment reaches, and (3) factor interaction. Significant ($P < 0.05$) or marginally significant ($0.05 < P < 0.10$) interaction terms indicate that the analysis and the direction or significance of the restoration response could be confounded as a result of conflicting responses among streams. In general, the net responses used to test for effects of restoration were normally or nearly normally distributed (tested through a normal probability plot with 95% CIs); unless otherwise noted, differences were considered significant at P-values less than 0.05. The statistical tests and surveys (survey numbers from Table 1) used to test hypotheses that restoration produces an increase in the biomass of brown trout, salmonids, or the entire fish community and to estimate power for different study designs can be summarized as follows:

(1) 95% CIs for single-year evaluations of absolute changes in each index at each treatment reach (survey 12 versus 13, 16 versus 17, 25 versus 26, and 37 versus 38; Table 1);
(2) Student's t-test for multiple-year evaluations of

absolute changes at each treatment reach (surveys 11, 12 versus 13, 14 and surveys 36, 37 versus 38, 39; Table 1);

(3) Student's *t*-test for multiple-year evaluations of absolute changes in each index at four pooled treatment reaches (surveys 11, 12, 16, 25, 36, and 37 versus surveys 13, 14, 15, 17, 18, 26, 27, 38, 39, and 40; Table 1);

(4) Student's *t*-test of BACI (net) changes in each index at each treatment reach using one reference reach (e.g., surveys 11–1 and 12–2 versus surveys 13–3 and 14–4 at lower Batavia Kill; Table 1);

(5) Student's *t*-test of net changes in each index at each treatment reach using one reference reach and one control reach (e.g., surveys 11–1, 11–6, 12–2, and 12–7 versus surveys 13–3, 13–8, 14–4, and 14–9 at lower Batavia Kill; Table 1);

(6) Student's *t*-test of net changes in each index at pooled treatment reaches using one reference reach (surveys 11–1, 12–2, 16–3, 25–19, 36–28, and 37–29 versus surveys 13–3, 14–4, 15–5, 17–4, 18–5, 26–20, 27–21, 38–30, 39–31, and 40–32; Table 1);

(7) Student's *t*-test of net changes in each index at pooled treatment reaches using one reference reach and one control reach (surveys 11–1, 11–6, 12–2, 12–7, 16–3, 16–8, 25–19, 25–22, 36–28, 37–29, and 37–33 versus surveys 13–3, 13–8, 14–4, 14–9, 15–5, 15–10, 17–4, 17–9, 18–5, 18–10, 26–20, 26–23, 27–21, 27–24, 38–30, 38–34, 39–31, 39–35, and 40–32; Table 1);

(8) two-factor ANOVA of net changes in each index at two, three, and four pooled treatment reaches using one reference reach (surveys 11–1 and 12–2 versus 13–3, 14–4, and 15–5; surveys 16–3 versus 17–4 and 18–5; surveys 25–19 versus 26–20 and 27–21; and surveys 36–28 and 37–29 versus 38–30, 39–31, and 40–32; Table 1); and

(9) two-factor ANOVA of net changes in each index at two, three, and four pooled treatment reaches using two reference reaches (surveys 11–1, 11–6, 12–2, and 12–7 versus 13–3, 13–8, 14–4, 14–9, 15–5, 15–10; surveys 16–3 and 16–13 versus 17–4, 17–9, 18–5, and 18–10; surveys 25–19 and 25–22 versus 26–20, 26–23, 27–21, and 27–24; and surveys 36–28, 37–29, and 37–33 versus 38–30, 38–34, 39–31, 39–35, and 40–32; Table 1).

Power analyses.—Power analyses are typically used to establish the number of replicates needed to accurately assess statistical differences and to determine the likelihood that a sampling and analysis design will detect an effect under given or known variances and expected effect sizes. However, variability (before and after restoration) and mean changes in target indices (after restoration) are generally unknown before surveys are completed (as in our study); therefore, the a posteriori analyses of survey data presented herein are intended to demonstrate the effectiveness of alternative sampling designs and replicate numbers (different number of sample years or study streams) assessed to detect the effects of restoration. As such, these analyses are meant not to justify the statistical adequacy of our study design but rather to determine which practical sampling designs might be effective in characterizing the effects of stream restoration on biomass of fish populations and communities under circumstances similar to those encountered in our study streams. Statistical power at an α of 0.05 for two-sample *t*-tests (and two-way ANOVAs; see above) was calculated using observed sample variances, effect size (contrast), and number of replicates (Lenth 2005). Integrated power was also calculated for all *t*-tests and used as the primary tool to measure and interpret the observed responses or effects (Lenth 2005). Integrated power does not rely on a single α for hypothesis testing; therefore, it is a more-robust gauge (i.e., relative to statistical power) of the ability of a given sampling design or analysis to detect true differences. Lenth (2005) indicated that an integrated power target of 0.95 should be comparable to a statistical power target of 0.80 at an α-value of 0.05. The number of replicates (years of sampling at each reach); number of reaches inventoried before and after restoration within each stream; and use of no reference reach, one reference reach, or both a reference reach and a control reach (balanced and asymmetrical BACI designs) were varied for each biotic index to assess their effects on significance level (*P*), statistical power ($1 - \beta$) at an α of 0.05, and integrated power across all levels of α (Lenth 2005). Estimates of integrated power for all tests of brown trout, salmonid, and fish community biomass were summarized and used to (1) assess the effects of various study or analytical designs on the ability to accurately detect meaningful biomass responses to restoration in our streams and (2) evaluate various study designs that are typically employed to detect fishery responses under comparable situations.

Results

Fish Responses

The biomass of brown trout increased markedly at each of the four treatment reaches after restoration and although the significance level varied among analyses, no or very few brown trout were present at three of the four reaches before restoration (Figure 2). Absolute changes in brown trout biomass ranged from 0.88 to 5.33 g/m^2 at the four restored reaches, and increases were generally significant or the 95% CIs did not

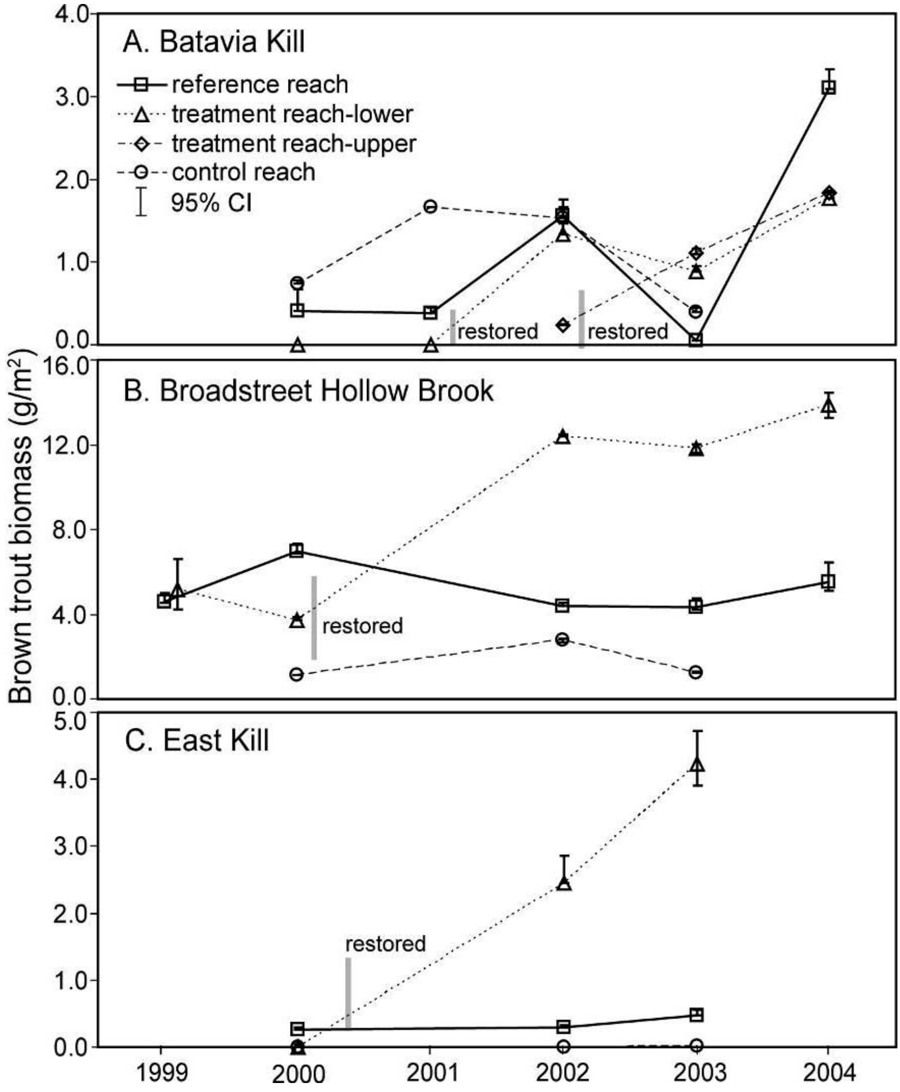

FIGURE 2.—Pre- and postrestoration estimates of total brown trout biomass (±95% confidence interval [CI]) at four restored (treatment) reaches, three stable reference reaches, and three unstable control reaches within three streams of the Catskill Mountains, New York, 1999–2004 (modified from Baldigo et al. 2008b; shaded vertical bars = approximate period in which restoration activities occurred). Two restored reaches (upper and lower) were present in the Batavia Kill.

overlap (Table 2, Figure 2). Net increases (relative to reference or control reaches) were typically less than 1.0 g/m^2 and were not significant at either restored reach on the Batavia Kill; net increase ranged from 3.26 to 6.17 g/m^2 and was generally significant for Broadstreet Hollow Brook, East Kill, and the four pooled treatment reaches (Table 2). The ANOVAs using data from two, three, or four streams and one reference reach or both reference and control reaches indicated that restoration in these streams generally accounted for a significant net increase of 1.81–2.69 g/m^2 in biomass of brown trout (Table 2). The increases in net biomass were only marginally significant when assessing responses at two or three treatment reaches with one reference reach (Table 2). Stream-to-stream differences in the size of differentials and factor interaction terms were significant only when both the reference and control reaches were used to calculate and assess net responses in brown trout population biomass at the four treatment reaches (Table 3).

The overall biomass of all salmonids increased at the four treatment reaches after restoration (Figure 3); the

TABLE 2.—Integrated power (IP), statistical power ($1 - \beta$), actual α-values (one-tailed), and the responses (effect sizes) of absolute biomass (actual year-to-year differences) and net biomass (differences in differential, calculated as biomass at the restored reach minus biomass at the control or reference reach during the same year) for brown trout, all salmonids, and the entire fish community at four treatment (restored) reaches in three streams of the Catskill Mountains, New York, based on alternative sampling designs (different numbers of years and treatment reaches) and analysis techniques (individual streams or pooled data; absolute or net biomass; 95% confidence interval [CI], *t*-test, or analysis of variance [ANOVA]). Values in bold italic indicate increases that were significant ($P < 0.05$).

Type of change assessed	Stream(s)[a] assessed	Reach(es) assessed[b]	Replicates before	Replicates after	Statistical test	Brown trout			
						Effect (g/m^2)	Actual α	$1 - \beta$ ($\alpha = 0.05$)	IP
Upper and lower Batavia Kill									
Absolute	BAT	T	1	1	95% CI	***1.34***	***<0.05***	na[c]	na
Absolute	UBAT	T	1	1	95% CI	***0.88***	***<0.05***	na	na
Absolute	BAT	T	2	2	*t*-test	1.11	0.065	0.55	0.79
Absolute	BAT	T	2	3	*t*-test	***1.33***	***0.018***	***0.93***	***1.00***
Absolute	BAT, UBAT	T	3	5	*t*-test	***1.31***	***0.001***	***1.00***	***1.00***
Net	BAT	T, R	2	2	*t*-test	0.70	0.205	0.17	0.70
Net	BAT	T, R	2	3	*t*-test	0.15	0.414	0.06	0.56
Net	BAT	T, R, C	3	5	*t*-test	0.52	0.207	0.19	0.73
Net	BAT, UBAT	T, R, C	6	8	*t*-test	***0.98***	***0.015***	***0.75***	***0.95***
Broadstreet Hollow Brook									
Absolute	BSH	T	1	1	95% CI	***3.62***	***<0.05***	na	na
Absolute	BSH	T	2	2	*t*-test	***3.69***	***0.024***	***0.86***	***0.97***
Absolute	BSH	T	2	3	*t*-test	***5.33***	***0.047***	***0.68***	***0.95***
Net	BSH	T, R	2	2	*t*-test	4.71	0.106	0.42	0.92
Net	BSH	T, R	2	3	*t*-test	***6.17***	***0.032***	***0.67***	***0.95***
Net	BSH	T, R, C	2	5	*t*-test	***5.19***	***0.009***	***0.90***	***0.98***
East Kill									
Absolute	EK	T	1	1	95% CI	***2.44***	***<0.05***	na	na
Net	EK	T, R, C	2	4	*t*-test	***3.26***	***0.004***	***1.00***	***1.00***
All streams (pooled)									
Net	ALL 4	T, R	6	10	*t*-test	***2.77***	***0.010***	***0.81***	***0.96***
Net	ALL 4	T, R, C	11	17	*t*-test	***4.11***	***0.001***	***0.94***	0.86
ANOVA (2 × 2, 2 × 3, and 2 × 4 designs)									
Net	BAT, EK	T, R	3	5	2 × 2 ANOVA	1.68	0.083	0.90	na
Net	BAT, UBAT, EK	T, R	4	7	2 × 3 ANOVA	1.53	0.093	0.94	na
Net	ALL 4	T, R	6	10	2 × 4 ANOVA	***2.69***	***0.022***	***1.00***	na
Net	BAT, EK	T, R, C	6	9	2 × 2 ANOVA	***1.98***	***0.001***	***1.00***	na
Net	BAT, UBAT, EK	T, R, C	8	12	2 × 3 ANOVA	***1.81***	***<0.001***	***1.00***	na
Net	ALL 4	T, R, C	11	17	2 × 4 ANOVA	***2.66***	***<0.001***	***1.00***	na

[a] BAT = lower Batavia Kill, UBAT = upper Batavia Kill, BSH = Broadstreet Hollow Brook, and EK = East Kill.
[b] Reach types were treatment (T), stable reference (R), and unstable control (C).
[c] Not applicable (na).

increases were closely linked with changes in brown trout biomass, as brown trout dominated the salmonid biomass at most study reaches (Table 2). Absolute increases in salmonid biomass ranged from 0.94 to 8.69 g/m^2 at the four restored reaches, and generally the changes were significant or the 95% CIs did not overlap (Table 2; Figure 3). The increases in net biomass of salmonids were typically less than 1.8 g/m^2 and were rarely significant at the two Batavia Kill restored reaches; however, the increases ranged from 3.26 to 9.30 g/m^2 and were generally significant for Broadstreet Hollow Brook, East Kill, and the four pooled treatment reaches (Table 2). The ANOVAs using data from two, three, or four streams and one reference reach or both reference and control reaches showed that restoration generally produced a significant net increase of 2.04–3.66 g/m^2 in salmonid biomass (Table 2). However, the magnitude of the differentials differed among streams ($P < 0.001$) and the interaction term was significant ($P = 0.001$–0.010; Table 3), indicating that salmonid biomass responses differed among streams and that the effects of restoration are best assessed within individual streams. Biomass increases were not significant when net changes were assessed using only two or three treatment reaches (Table 2).

Total community biomass decreased at both Batavia Kill treatment reaches and increased at the Broadstreet Hollow Brook and East Kill treatment reaches after restoration (Table 2; Figure 4). Absolute decreases in

TABLE 2.—Extended.

Type of change assessed	Salmonids				Fish community			
	Effect (g/m^2)	Actual α	$1 - \beta$ ($\alpha = 0.05$)	IP	Effect (g/m^2)	Actual α	$1 - \beta$ ($\alpha = 0.05$)	IP
Upper and lower Batavia Kill								
Absolute	***1.85***	***<0.05***	na	na	***−9.39***	***<0.05***	na	na
Absolute	***0.94***	***<0.05***	na	na	***−4.83***	***<0.05***	na	na
Absolute	1.51	0.071	0.51	***0.95***	−4.03	0.263	0.13	0.72
Absolute	***1.61***	***0.001***	***0.99***	***0.99***	−1.77	0.375	0.08	0.59
Absolute	***1.659***	***0.001***	***1.00***	***1.00***	−1.25	0.358	0.08	0.61
Net	1.75	0.216	0.16	0.94	−1.68	0.187	0.19	0.81
Net	1.03	0.289	0.11	0.66	−0.95	0.216	0.17	0.73
Net	1.06	0.162	0.23	0.77	−1.53	0.214	0.18	0.72
Net	***1.43***	***0.050***	***0.52***	***0.89***	−1.83	0.263	0.15	0.67
Broadstreet Hollow Brook								
Absolute	***8.69***	***<0.05***	na	na	***10.96***	***<0.05***	na	na
Absolute	***7.67***	***0.032***	***0.96***	***0.99***	***9.87***	0.060	0.74	***0.96***
Absolute	***8.26***	***0.006***	***1.00***	***1.00***	***10.91***	***0.018***	***0.90***	***0.98***
Net	9.09	0.066	1.00	***0.95***	***14.34***	***0.039***	***0.91***	***0.98***
Net	9.30	0.064	0.58	***0.95***	***14.98***	***0.039***	***0.96***	***0.98***
Net	***8.35***	***0.015***	***0.97***	***0.99***	***11.96***	***0.002***	***1.00***	***1.00***
East Kill								
Absolute	***2.44***	***<0.05***	na	na	***5.64***	***<0.05***	na	na
Net	***3.26***	***0.004***	***1.00***	***1.00***	2.41	0.141	0.28	0.80
All streams (pooled)								
Net	***4.00***	***0.019***	***0.70***	0.94	***4.23***	***0.045***	***0.54***	0.90
Net	***4.11***	***0.001***	***0.99***	***0.99***	3.45	0.067	0.44	0.86
ANOVA (2 × 2, 2 × 3, and 2 × 4 designs)								
Net	2.11	0.128	0.77	na	0.264	0.92	0.03	na
Net	1.78	0.113	0.91	na	0.86	0.687	0.11	na
Net	***3.66***	***0.002***	***1.00***	na	3.10	0.064	1.00	na
Net	***2.34***	***0.012***	***1.00***	na	0.17	0.949	0.03	na
Net	***2.04***	***0.009***	***1.00***	na	0.45	0.851	0.06	na
Net	***3.61***	***<0.001***	***1.00***	na	2.66	0.168	0.95	na

community biomass ranged from −1.25 to −9.39 g/m^2 at the two restored Batavia Kill reaches; however, most decreases were not significant (Table 2). Absolute increases in community biomass were significant and ranged from 5.64 to 10.96 g/m^2 (or the 95% CIs did not overlap) at the restored reaches of Broadstreet Hollow Brook (Table 2) and East Kill (Table 2; Figure 4). Net decreases in fish community biomass at the two Batavia Kill restored reaches were between −0.95 and −1.83 g/m^2 and were not significant (Table 2). In contrast, net increases at the restored reach of Broadstreet Hollow Brook ranged from 11.96 to 14.98 g/m^2 and were significant (Table 2). The net increase in community biomass was 2.41 g/m^2 at the East Kill (Table 2). The net increase in community biomass for the four pooled reaches ranged from 3.45 to 4.23 g/m^2 (Table 2) but was significant only when biomass data from the reference reaches were used to standardize responses in treatment reaches. The ANOVAs using data from two, three, or four streams and one reference reach or both reference and control reaches showed that restoration caused a net fish community biomass increase of 0.86–3.10 g/m^2, but the increases were not significant (Table 2). Significant factor interaction and differences in the size of differentials among streams (Table 3) suggest an analyses of total community biomass within individual streams would be most appropriate for evaluating responses to restoration.

Power Analyses

The results from both statistical and integrated power analyses were strongly correlated with each other ($r = 0.93$), and almost all values of statistical

TABLE 3.—Net biomass responses (mean change in biomass index differentials, calculated as biomass in the restored reach minus biomass in the control or reference reach during the same year) of brown trout, salmonids, and the entire fish community to restoration at four treatment reaches in three streams of the Catskill Mountains, New York, using only stable reference reaches or both reference and unstable control reaches. Also shown are the results (P-values) of two-factor analyses of variance assessing differences in pooled differentials (1) before versus after restoration, (2) among all four treatment reaches (regardless of restoration), and (3) for factor interaction (differing responses among the four reaches). Surveys were conducted between 1999 and 2004. Values in bold italic indicate significant differences ($P \leq 0.05$).

Index	Mean response (g/m^2)	P-value for differences in differentials		
		Before–after	Among streams	Interaction
Reference reaches only[a,b]				
Fish community	3.10	0.064	0.088	***0.005***
Salmonids	***3.66***	***0.002***	***0.001***	***0.010***
Brown trout	***2.69***	***0.022***	0.126	0.126
Reference and control reaches[a,c]				
Fish community	2.66	0.168	***0.012***	***0.030***
Salmonids	***3.61***	***<0.001***	***<0.001***	***0.001***
Brown trout	***2.66***	***<0.001***	***0.001***	***0.018***

[a] One stream had two restored sites that shared a single reference reach and a single control reach. Therefore, the total number of reaches was four treatment, three reference, and three control.
[b] $n = 16$; 6 replicates before and 10 replicates after (see analysis of variance [ANOVA], net change, all four streams, and treatment [T] and stable reference [R] reaches in Table 2).
[c] $n = 28$; 11 replicates before and 17 replicates after (see ANOVA, net change, all four streams, and treatment T, R, and unstable control [C] reaches in Table 2).

power greater than 0.80 corresponded to integrated power values greater than 0.95. Thus, the results of the integrated power analyses were used exclusively to define the relations between power and effect size, replicate number, and variability. Integrated power was strongly related to effect size and (with a few exceptions) was greater than 0.95 for t-tests when absolute and net biomass responses exceeded 5.0 g/m^2 for the fish community, 2.0 g/m^2 for all salmonids, and 1.0 g/m^2 for brown trout (Figure 5A). An effect size (biomass response) of 5.0 g/m^2 or larger always had an integrated power value greater than 0.95. The best-fit line indicated that biomass increases of 6.0 g/m^2 would always (in our tests) yield significant response determinations. Effect sizes of 2.0–5.0 g/m^2 would often indicate significant impacts if the variability in absolute or net responses were relatively low (coefficient of variation [CV = 100 × SD/mean] < 100%); effect sizes of 1.0–2.0 g/m^2 would occasionally denote significant impacts if the variability in responses was very low (CV < 30%).

The influence of variability on the ability of any particular sampling design to detect significant responses is illustrated by the strong relation between integrated power and CV (Figure 5B). Integrated power did not surpass 0.95 for any test design in which response CV was greater than about 90%. Integrated power was generally near or greater than 0.95 when the CV was less than 30% for total fish community biomass, less than 90% for salmonid biomass, or less than 80% for brown trout biomass (Figure 5B). Overall, the best-fit line for all absolute and net biomass responses showed that significant responses would be detected in (1) nearly all tests when the CV was less than 30% and (2) most (78%) of tests when the CV was less than 80%. The variability (i.e., CV) of net responses essentially characterizes the tendency of differences in biomass index values from paired (reference and treatment) reaches to vary due to the effects of restoration and the normal year-to-year changes in natural conditions. It is therefore important to consider that sampling precision within reference and treatment reaches and normal year-to-year changes in target indices both contribute to this variability; thus, for efficient detection of responses caused mainly by restoration, sampling precision should be maximized and interannual variation should be minimized.

The relations between the number of replicates and integrated (or statistical) power were either nonexistent or more variable than the relations between effect size or CV and integrated power (Figure 5C). The wedge- or wing-shaped distribution suggested a general increase in integrated power as the number of replicates increased; however, expected strong biomass associations with biomass indices did not occur. The absence of a relation appears to be related to the high variability in biomass responses and interannual variability in biomass within individual reaches. Regardless of the number of replicates, integrated power was generally near or greater than 0.95 when the response CV was less than 80%; integrated power was less than 0.95 when the response CV was greater than 80%.

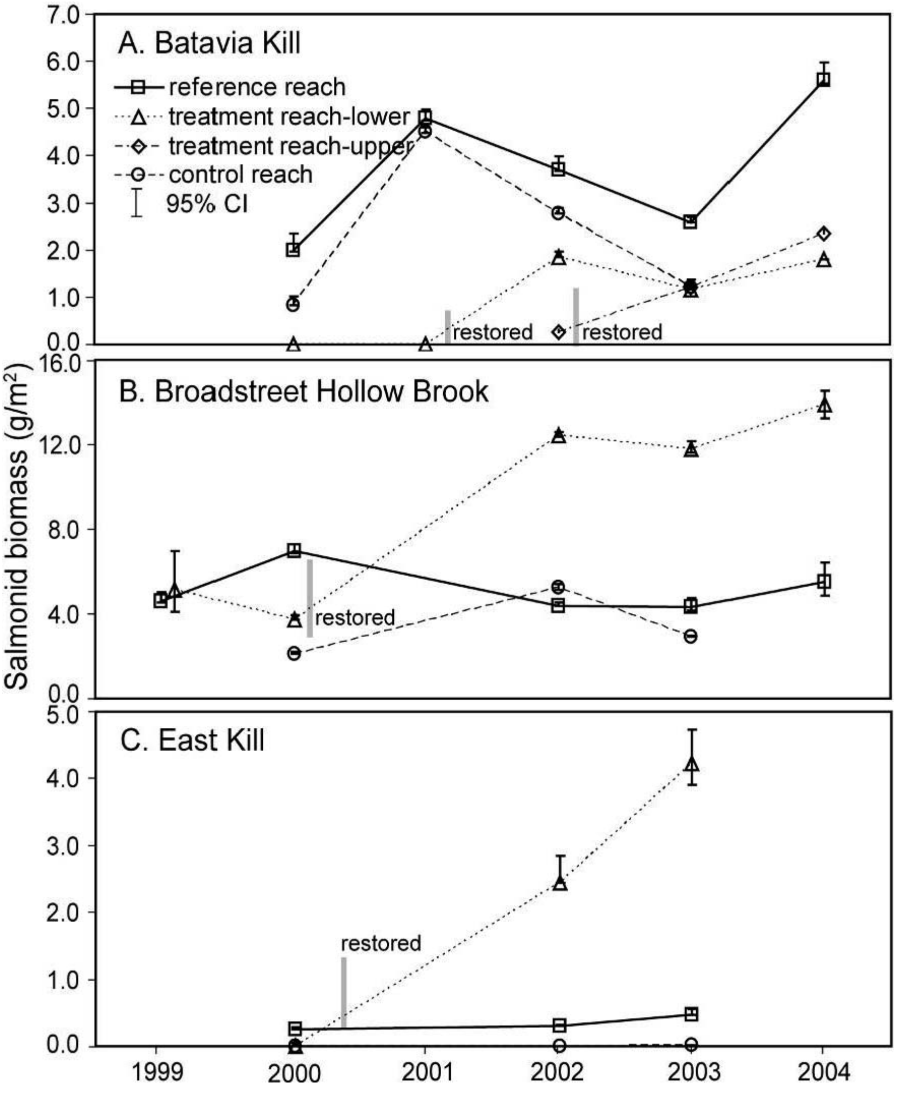

FIGURE 3.—Pre- and postrestoration estimates of total salmonid biomass (±95% confidence interval [CI]) at four restored (treatment) reaches, three stable reference reaches, and three unstable control reaches within three streams of the Catskill Mountains, New York, 1999–2004 (modified from Baldigo et al. 2008b; shaded vertical bars = approximate period in which restoration activities occurred). Two restored reaches (upper and lower) were present in the Batavia Kill.

Likewise, when four to eight replicates were assessed, integrated power was greater than 0.95 only if the CV was less than 40%. Thus, the variability in absolute and net responses is a critical design consideration. In fact, the CV and number of replicates can be used together to predict integrated power, as they accounted for 71% of integrated power variability in all test designs.

Discussion

Results from the present study suggest a number of ways in which the monitoring program design or data analysis can be tailored to allow efficient assessment and detection of the effects of stream restoration on targeted fish species or entire assemblages. The BACI sampling and analysis design was effective at isolating the true (net) effects of restoration from the normal year-to-year changes in the biomass of selected fish populations and communities. For example, absolute changes in brown trout population biomass at each of the two treatment reaches in the Batavia Kill illustrated large increases after restoration; however, net changes showed that the increases in brown trout biomass were

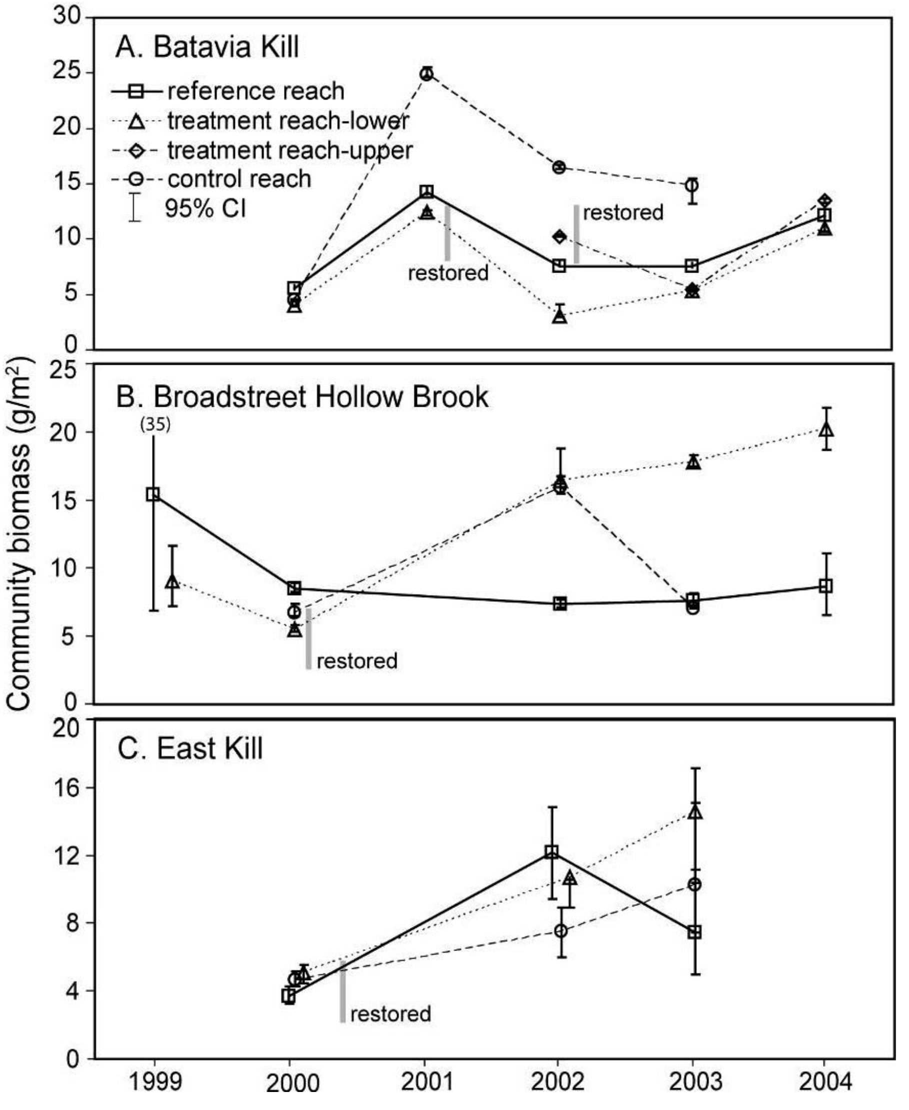

FIGURE 4.—Pre- and postrestoration estimates of total fish community biomass (±95% confidence interval [CI]) at four restored (treatment) reaches, three stable reference reaches, and three control reaches within three streams of the Catskill Mountains, New York, 1999–2004 (modified from Baldigo et al. 2008b; shaded vertical bars = approximate period in which restoration activities occurred). Two restored reaches (upper and lower) were present in the Batavia Kill.

smaller than absolute changes and generally were not significant (Table 2). Brown trout population biomass at the two treatment reaches did appear to be affected by the restoration, but either (1) the variability in data from treatment and references reaches was too great or (2) the basinwide (or regional) trends in other factors affected brown trout biomass similarly at both sets of paired study reaches. The first issue might alter the significance of statistical tests; the second could partially counteract absolute responses and reduce the net effect (or difference) attributed solely to restoration.

Regardless, the BACI analyses generally provided an unbiased measure of fish responses to restoration in this study, and the inclusion of such analyses in study designs for quantifying short- or long-term biological effects of stream restoration would help separate actual index changes caused by restoration from normal changes caused by natural trends in other environmental factors.

The utility of BACI sampling designs depends on the goals and scope of a restoration plan, and BACI designs might not be appropriate for all monitoring and evaluation programs. A number of studies have described the limitations or misuse of BACI study designs in various contexts (Hurlbert 1984; Smith et al. 1993; Underwood 1992, 1994; Murtaugh 2002). Hurlbert (1984) was concerned about pseudoreplication in larger ecological experiments. Underwood (1992, 1994) and Murtaugh (2002) also expressed concern about replication in BACI studies and recommended inclusion of additional control and treatment reaches when possible. Smith et al. (1993) and Murtaugh (2002) noted that the trends in biological communities at reference reaches in some cases might not parallel respective trends within treatment reaches, therefore leading to erroneous rejection of the null hypothesis. In the present study, we avoided common shortcomings by replicating treatments (restoration projects) across multiple systems and establishing multiple control reaches within the same streams in which restoration was conducted (typically within 1,000 m of our treatment sites). These strategies strongly reinforced the assertion that annual trends in fish assemblages (and indices) at control and reference reaches were comparable to those at the treatment reaches (movement of fish among study reaches within each stream was possible). In addition, all surveys were conducted within a few days of each other during the summer, thereby eliminating potential problems caused by seasonal variability, as noted by Smith et al. (1993) and Murtaugh (2002). These two papers further highlight the potential for a BACI study to separate effects that occur because of better conditions created by the treatment (impact) from those impacts possibly caused by the treatment. For example, in our study, one could argue that the act of weir placement, rather than the weirs themselves, induced the observed responses. We addressed this concern by collecting data for 3 years posttreatment. The direct impact of habitat rehabilitation was unlikely to persist for this length of time in these streams. We continue to monitor these sites periodically to address concerns about long-term responses. Finally, Murtaugh (2002) specifically recommended the use of tools in addition to statistical analyses (e.g., graphical displays, expert opinion, and common sense) to evaluate and interpret responses when using BACI analyses. In the present study, we used all three of these tools and additional statistical analyses to provide unbiased evidence for our conclusions. For example, treatment reaches in the Batavia and East kills showed an absence or near absence of salmonids before restoration and a measurable abundance of salmonids after treatment. Although the statistical analyses quantified a nonsignificant effect, graphical interpretation (Figure 2) and common sense also indicate that restoration influenced the presence of salmonids at both restored reaches relative to corresponding reference and control reaches.

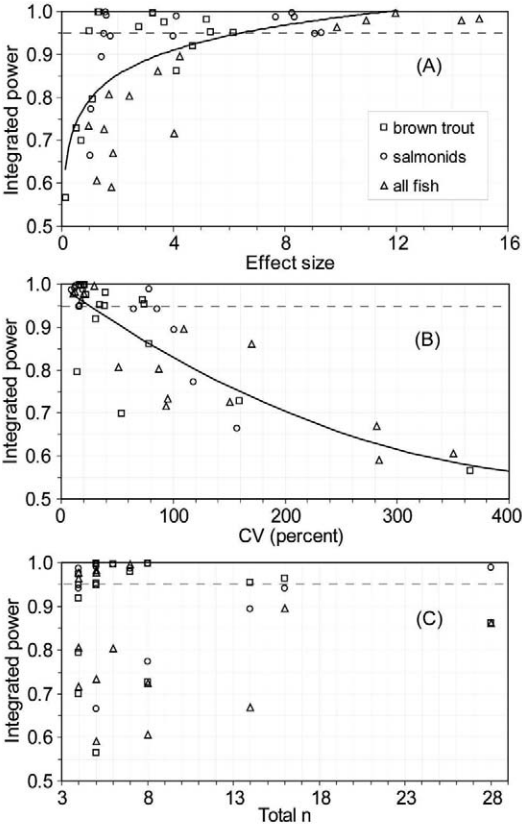

FIGURE 5.—Collective relations between integrated power (IP) and **(A)** effect size in terms of brown trout, salmonid, or fish community biomass response (g/m^2) to stream restoration, **(B)** the coefficient of variation (CV) of effect size, and **(C)** total number of replicates (*n*; years of sampling at each reach) for all *t*-tests evaluating absolute responses (actual year-to-year differences) and net changes (differences in the differential, calculated as biomass at the restored reach minus biomass at the control or reference reach during the same year) at two, three, or four restored reaches in streams of the Catskill Mountains, New York. The dashed horizontal line in each panel denotes the IP threshold of 0.95.

Integrated power for tests assessing the absolute and net biomass responses of brown trout, salmonid, and the fish community to restoration helped to identify important sampling criteria and to suggest study (sampling and analysis) designs for detecting significant fish biomass changes in small- to intermediate-sized streams. The various sampling designs determine the accuracy and precision with which net effects are determined, but the average net effect should be independent of the selected design. The analyses generally detected significant 5.0–6.0-g/m^2 biomass increases (integrated power > 0.95) due to restoration when the minimum number of replicates ranged from four to eight and when the CV for measured indices was less than 40%. In several cases where the effect size was as small as 1.0 g/m^2, integrated power exceeded 0.95 when the CV was less than 30%. Field surveys for the present study used salmonid counts to determine whether additional collection passes were needed to reduce 95% CIs to within 10% of the population estimate. Daces and sculpins are bottom-dwelling species and could not be sampled effectively; therefore, their presence increased the variability in estimates of fish community biomass and partly accounted for the lack of significant community responses at the Batavia Kill and the pooled treatment reaches. Findings for the pooled treatment reaches showed that restoration increased brown trout, salmonid, and fish community biomass (based on absolute and net changes) at the four restored reaches by an average of 2.99 g/m^2 (SD = 1.17 g/m^2). Five replicates (2–3 before treatment and 2–3 after treatment) would be required in this situation to generate an integrated power of 0.95. If one assumes that standard deviations for net responses (index differentials) vary widely, then sample sizes of 4, 6, and 14 (half before treatment and half after) will provide an integrated power greater than 0.95 if response CVs average 30, 50, and 80%, respectively. Sampling of either (1) two pairs of treatment and reference reaches (balanced or asymmetrical) for 1 year before restoration and 1 year after restoration or (2) one pair of treatment and reference reaches (symmetrical or asymmetrical) for 2 years before restoration and 2 years after restoration would be an appropriate minimal sampling design for fish populations that can be inventoried accurately and that change little from year to year. There are good arguments for both options. High variability in net response or small effect sizes would require an increased number of replicates. Because additional prerestoration data cannot be acquired after restoration is completed, it would be preferable to err on the side of too much data and collect at least 2 years of prerestoration data for each restoration project. Conversely, the first design may be the better "blind" option considering that year-to-year changes in differentials for any given index may be very large simply as a result of factors unrelated to the treatments. The power to detect significant effects due only to restoration could disappear if the variability for mean changes in a given index (i.e., the difference between treatment and control reach indices measured several years before and after restoration) becomes similar to or greater than the annual variability in the control reach index (measured over the same time period).

The objectives of any restoration project are of the utmost importance and must be carefully thought out and defined before a sampling and analysis plan that adequately characterizes and evaluates key responses can be designed. The actual size and location of sample reaches, number of reference or control reaches, survey periodicity and methods, total number of replicates before and after restoration, additional hydrologic and habitat variables, and data processing and analysis methods are only a few study design considerations. Selection of sampling design will depend on whether questions (and activities used to address such questions) focus on a broad or narrow issue across a wide region (many streams) or within a single stream. The results from such efforts could even be contradictory. For example, the ANOVAs indicated that restoration generally had no significant net effect on total fish community biomass at any of the four treatment reaches, yet total fish biomass increased significantly (by 12–15 g/m^2) at the Broadstreet Hollow Brook treatment reach after restoration. The ANOVAs are powerful tools for assessing the general effects of restoration techniques on species or species assemblages because they increase the number of samples and the power of respective analyses to detect significant responses. However, because data are pooled from several streams in which the initial habitat quality and fish community composition may vary widely, ANOVA results (whether significant or not) by definition disregard unique responses within individual streams.

In the present study, the prerestoration condition of stream habitat and fish populations appeared to affect the magnitude and direction (increase or decrease) of responses produced by restoration and thus should also be given some consideration. Streams with poor habitat quality, low fish species richness, low evenness, and low fish biomass before restoration were much more likely to exhibit significant (and larger) increases in each index after restoration than were stream reaches that had intermediate habitat quality and relatively healthy fish assemblages before restoration. In the three streams of the Catskill Mountains, restoration generally increased brown trout and salmonid biomass at most treatment reaches but did not strongly affect fish community biomass overall. Although the salmonid biomass responses were generally consistent, fish community biomass response was highly variable and was only significant within one stream. Information defining the health of preexisting fish assemblages is essential, not only for gauging any postrestoration effects but also for determining whether the stream ecosystem is degraded and whether there is any opportunity for improvement after restoration.

Increases in salmonid biomass constituted only one of the many responses indicating a general increase in stream ecosystem health due to restoration. For example, fish community biomass at the East and Batavia Kill treatment reaches before restoration consisted almost entirely of one or two small prey species (as high as 99% slimy sculpin *Cottus cognatus*, eastern blacknose dace *Rhinichthys atratulus*, and longnose dace *R. cataractae*) and few or no top-predator species; after restoration, biomass and relative proportions of brown trout, brook trout *Salvelinus fontinalis*, and rainbow trout *Oncorhynchus mykiss* increased and the relative proportion of minnow biomass decreased (Baldigo et al. 2008a). Fish communities at restored reaches often resembled the natural, more evenly balanced fish communities found in corresponding reference reaches of the region; therefore, restoration generally improved the overall health of local fish communities.

In summary, the ability to thoroughly assess the success or failure of a stream restoration project depends on several elements. First, there must be a sound foundation of relevant data (assessments) that illustrate or quantify substantial chemical, physical, or ecological perturbations and the potential for abatement via some form of remediation or restoration. Second, the primary objective (and associated goals) of any stream restoration effort should be clearly defined. Third, the restoration must be designed in a manner that addresses primary objectives and target goals. The selected design requires evaluation of results from comparable efforts to help predict responses within project areas; however, such information by and large is unreported in the scientific literature. Lacking this information, a fourth monitoring and analysis step will be necessary if a funding agency wants to determine whether the restoration project had the desired effect(s) on stated goals. Therefore, a sampling and analysis design would have to define the monitored factors, number and location of sample sites, duration of monitoring (both before and after restoration), sampling frequency, sampling methods, data analyses, and the degree or level of change that signifies success. Some of these strategies may be subjective; however, all require thoughtful consideration of their relation to project goals, monitoring costs, and available expertise. Though BACI methods can help separate restoration responses from normal year-to-year variability, the potential response size and index variability (and differentials) will also influence the choice of design. Our results suggest that an effective strategy to detect biological responses in low-order streams will include BACI analysis of data that are sampled from paired reference and treatment reaches at least twice before and twice after restoration. Lastly, the dissemination of results from such investigations can only help others to more clearly identify goals; effectively monitor, analyze, and detect significant biological responses; and judge whether stream and watershed restoration actually attains some desired level of success.

Acknowledgments

The authors extend their appreciation to Christiane Mulvihill, Anne Ernst, Britt Westergard, and Anne Finch of the U.S. Geological Survey (USGS); Christina Falk, Sarah Miller, Dan Davis, Phillip Eskeli, Mark Vian, and Elizabeth Reichheld of NYCDEP; Cliff Kraft, Milo Richmond, Walter Keller, Mathew Horn, Marshall Thomas, and Mike Compton of Cornell University; Douglas Dekoski, Joel DeBois, Jake Bucanan, and René VanShaack of the Green County Soil and Water Conservation District (GCSWCD); and many AmeriCorps and Ulster County Community College interns for technical support. This research was funded by the USGS, NYCDEP, and GCSWCD. Reference to trade names does not imply endorsement by the U.S. Government.

References

Baldigo, B. P., A. S. Gallagher-Ernst, W. Keller, D. R. Warren, S. J. Miller, D. Davis, T. P. Baudanza, D. DeKoskie, and J. R. Buchanan. 2008a. Restoring geomorphic stability and biodiversity in streams of the Catskill Mountains, New York, USA. Pages 1777–1790 *in* J. Nielsen, J. J. Dodson, K. Friedland, T. R. Hamon, J. Musick, and E. Verspoor, editors. Reconciling fisheries with conservation: proceedings of the Fourth World Fisheries Congress. American Fisheries Society, Symposium 49, Bethesda, Maryland.

Baldigo, B. P., D. R. Warren, A. G. Ernst, and C. I. Mulvihill. 2008b. Response of fish populations to natural channel design restoration in streams of the Catskill Mountains, New York. North American Journal of Fisheries Management 28:954–969.

Bernhardt, E. S. 2005. National River Restoration Science Synthesis (NRRSS) statistics page for all node areas. U.S. Geological Survey. Available: nrrss.nbii.gov. (December 2005).

Bernhardt, E. S., M. A. Palmer, J. D. Allan, G. Alexander, K. Barnas, S. Brooks, J. W. Carr, S. Clayton, C. N. Dahm, J. F. Follstad-Shah, D. L. Galat, S. G. Gloss, P. Goodwin, D. D. Hart, B. Hassett, R. Jenkinson, S. Katz, G. M. Kondolf, P. S. Lake, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano, B. Powell, and E. Sudduth. 2005. Synthesizing U.S. river restoration efforts. Science 308:636–637.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

Lenth, R. 2005. Java applets for power and sample size.

University of Iowa. Available: www.stat.uiowa.edu. (May 2005).

Murtaugh, P. A. 2002. On rejection rates of paired intervention analyses. Ecology 83:1752–1761.

Palmer, M. A., E. S. Bernhardt, J. D. Allan, P. S. Lake G. Alexander, S. Brooks, J. W. Carr, S. Clayton, C. N. Dahm, J. F. Follstad-Shah, D. L. Galat, S. Gloss, P. Goodwin, D. D. Hart, B. Hassett, R. Jenkinson, G. M. Kondolf, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano, and E. Sudduth. 2005. Standards for ecologically successful river restoration. Journal of Applied Ecology 42:208–217.

Pretty, J. L., S. S. C. Harrison, D. J. Shepherd, C. Smith, A. G. Hildrew, and R. D. Hey. 2003. River rehabilitation and fish populations: assessing the benefit of instream structures. Journal of Applied Ecology 40:251–265.

Roni, P., T. J. Beechie, R. E. Bilby, F. E. Leonetti, M. M. Pollock, and G. R. Pess. 2002. A review of stream restoration techniques and a hierarchical strategy for prioritizing restoration in Pacific Northwest watersheds. North American Journal of Fisheries Management 22:1–20.

Rosgen, D. L. 1994a. A classification of natural rivers. Catena 22:169–199.

Rosgen, D. L. 1994b. River restoration utilizing natural stability concepts. Land and Water 38(4):36–41.

Smith, E. P., D. R. Orvos, and J. Cairns. 1993. Impact assessment using the before-after-control-impact (BACI) model—concerns and comments. Canadian Journal of Fisheries and Aquatic Sciences 50:627–637.

Stewart-Oaten, A., W. W. Murdoch, and K. R. Parker. 1986. Environmental-impact assessment—pseudoreplication in time. Ecology 67:929–940.

Underwood, A. J. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable world. Journal of Experimental Marine Biology and Ecology 161:145–178.

Underwood, A. J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications 4:3–15.

Warren, D. R., and C. E. Kraft. 2003. Brook trout (*Salvelinus fontinalis*) response to wood removal from high-gradient streams of the Adirondack Mountains (New York, USA). Canadian Journal of Fisheries and Aquatic Sciences 60:379–389.