North American Journal of Fisheries Management 19:860–866, 1999

A Power Analysis on the Monitoring of Bull Trout Stocks Using Redd Counts

BRYCE A. MAXELL*

Wildlife Biology Program, University of Montana, Missoula, Montana 59812, USA

Abstract.—The bull trout *Salvelinus confluentus* is listed as a federally threatened species in the Columbia and Klamath river drainages. A priori establishment of levels of decline or increase in bull trout redd numbers that will be considered biologically significant and levels of statistical significance that will be used to identify changes in redd numbers is essential to the success of future recovery plans. A prospective statistical power analysis indicates that with standard significance levels and two-tailed testing procedures the yearly variation in redd numbers typifying many stocks of bull trout in Montana limits the power of detecting less than 50% of changes in population size per generation to less than 0.8 during the first 15 years of a monitoring program. The limitations of monitoring bull trout stocks with redd counts, coupled with the critical nature of identifying future population changes, justifies the need to (1) identify and reduce the level of measurement error involved in redd counts, (2) use levels of statistical significance that adequately balance the risks of committing type I and type II errors, (3) use one-tailed testing procedures for identifying population declines during the initial and other critical years of a monitoring program, and (4) explore the use of other methods of monitoring.

In response to declines in stocks of bull trout *Salvelinus confluentus* throughout their range the U.S. Fish and Wildlife Service (USFWS) has listed distinct population segments in both the Klamath and Columbia river drainages as threatened under the Endangered Species Act (ESA; USFWS 1998). As required under section 4(f) of the ESA, the USFWS must now develop a recovery plan that includes objective, measurable criteria that, when met, would allow the species to be delisted (USC 1973a). Similarly, habitat conservation plans that states, tribes, agencies, companies, or individuals may effect must, to receive an incidental take permit, include criteria to ensure the maintenance and recovery of the species (USC 1973b; USFWS 1998).

The temporal and site-specific nature of spawning, the homing of adults to natal streams, and the relative ease with which spawning redds can be counted make redd counts a valuable index for evaluating trends in the size of local stocks and regional populations of bull trout (Rieman and McIntyre 1993; Rieman and McIntyre 1996; Rieman and Myers 1997). In addition, because redds are a product of only the reproductive adults, they provide a useful index of the effective population size of a stock (Meffe 1986; Meffe and Carroll 1994). For these reasons the determination of the status of bull trout stocks (stable, decreasing, or increasing) and subsequent management policies regarding those stocks are likely to be primarily based on the number of spawning redds counted each fall.

Despite the apparent utility and widespread use of this method for evaluating population status, no evaluation has been made of its ability to detect changes in population size; that is, what are the probabilities of detecting different magnitudes of population declines or increases over different time periods? Similarly, no clear evaluation has been made of the degree of measurement error associated with the counting of bull trout redds or the effect this measurement error might have on our ability to detect changes in population size (Rieman and McIntyre 1996). Finally, no evaluation has been made of our ability to detect biologically significant changes in bull trout redds at various levels of statistical significance; that is, will we be able to detect increases or declines in redd numbers that are accepted as biologically significant at levels of statistical significance that are commonly accepted and may be used in recovery goals (e.g., $\alpha = 0.05$)?

I performed a prospective statistical power analysis on the monitoring of bull trout populations using redd counts over a simulated 3–30 year monitoring and recovery plan, such as that proposed by the state of Montana (K. McDonald, Montana Department of Fish, Wildlife and Parks, personal communication). My analysis was done to: (1) identify the time required to detect various changes in redd numbers at standard levels of statistical significance, based on yearly variations in redd numbers typical of many bull trout stocks in Montana; (2) identify the effect of different levels of

* E-mail: nathist@selway.umt.edu

Received October 1998; accepted April 1999

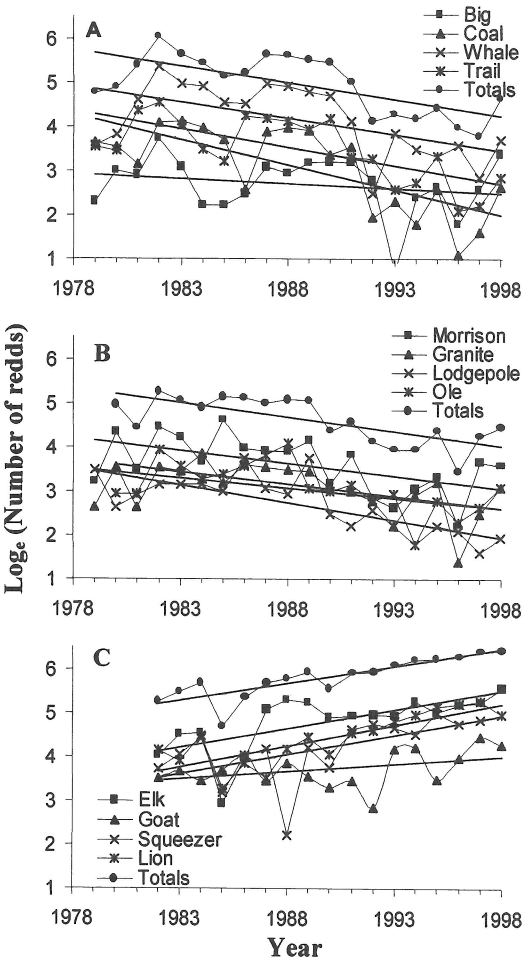

FIGURE 1.—Line graphs and linear regressions of the natural log of redd numbers over time for individual tributaries and drainage totals. Slopes of regression lines for all tributaries and drainage totals in the North (A) and Middle (B) forks of the Flathead River are significantly negative ($P \leq 0.05$), except for Big Creek ($P = 0.4$). Slopes of regression lines for tributaries and drainage totals in the Swan River (C) are significantly positive ($P \leq 0.05$), except for Goat Creek ($P = 0.09$).

measurement error on statistical power; and (3) identify the effects of using one-tailed testing procedures and lower levels of statistical significance ($\alpha = 0.2$) on statistical power.

Methods

Although several approaches are possible when attempting to detect population trends, I used linear regression with sample year as the independent variable and $\log_e$-transformations of redd number as the dependent variable. This approach assumes that the population increases or declines in an exponential manner and that the instantaneous rate of change (r) is constant over the monitoring period. The approach is convenient because the slope of a regression line fit to the $\log_e$-transformed data is equivalent to the intrinsic rate of change, r, for the population, which can then be compared against a null hypothesis that the slope is zero (Gerrodette 1987). In addition, the exponential model is appropriate because overall trends in redd numbers in the Flathead and Swan river drainages of Montana appear to have exponentially decreased and increased, respectively, over the last 18 years (Figure 1), and other bull trout stocks may be expected to decline or increase in a similar manner in the future.

TABLE 1.—Effect sizes used in the power analysis and the resulting change in redd numbers over 15 years (three generations of bull trout), based on an average minimum generation time of 5 years and a constant annual finite growth rate over 14 years of population change (15 years of monitoring).

Percent change in redd numbers per generation	Annual finite growth rate (λ)	Intrinsic rate of growth (r)	Percent change in redd numbers over 15 years
−50	0.871	−0.138	−95
−20	0.956	−0.045	−47
−10	0.979	−0.021	−26
0	1.000	0.000	0
+10	1.019	+0.019	+30
+20	1.037	+0.036	+66
+50	1.084	+0.081	+209

To model different rates of change, I used three increasing and three declining effect sizes: 50, 20 and 10% change in population size per generation (Table 1). These effect sizes were modified from the International Union for Conservation of Nature (IUCN) criteria for assessing a species' threat of extinction (Mace and Lande 1991). Using these effect sizes, I calculated the finite annual growth rate (λ), the associated intrinsic capacity for growth (r) and total percent change in population size that would result over a 15-year monitoring program (Table 1). In these calculations I assumed that the rate of change in population size was constant (Gerrodette 1987, 1993; Hayes and Steidl 1997) and that the average generation time for bull trout is 5 years (Rieman and McIntyre 1993).

To model the effect of measurement error on the power of detecting population change over 3–30 years, I established bounds of 30, 40, 50, 60, 70, and 80% variation around a model population of 100 redds. Although these levels of variation encompass variation in redd numbers resulting from a combination of measurement error and environmental and demographic stochasticity, the effects of reducing measurement error on statistical power can be evaluated by comparisons between the different levels of variation considered. A random-number generator was then used to select a number of redds from a uniform distribution between the upper and lower bounds of each level of variation considered. This was repeated 1,000 times for each level of variation, and the means and standard deviations were then used to calculate the coefficient of variation (CV = SD/mean) associated with a given level of variation in redd numbers:

Percent variation	CV
30%	0.178
40%	0.236
50%	0.291
60%	0.351
70%	0.414
80%	0.472

The six effect sizes (Table 1) and six CV values are largely representative of the average rates of change per generation (−43% to +61%) and CV values (0.415–0.672) associated with stocks of bull trout in the Flathead and Swan river systems of Montana over the last 16–19 years (T. Weaver, Montana Department of Fish, Wildlife and Parks, unpublished data). The effect sizes and CV values were then used to evaluate the statistical power associated with 3–30 years of monitoring data. I used the software package TRENDS (Gerrodette 1993) to perform the power analysis. This software allows the user to input different rates of change or effect sizes, different levels of statistical significance (levels of α and β), initial coefficients of variation (CV), and years of sample data (N). The user can input any four of these variables, and TRENDS will calculate the value for the fifth. In addition, the user can specify whether to model population growth as linear or exponential, how the CV varies with population abundance, whether a z- or t-distribution should be used for the calculations, and whether a one- or two-tailed test is to be used (Gerrodette 1987, 1993). For this analysis, population growth was modeled as exponential (Figure 1) and CV was assumed not to vary with different population sizes because there was no significant relationship between CV and mean number of redds for streams in the Flathead and Swan river systems from 1979 or from 1982 to 1998 respectively (Weaver, unpublished data). A t-distribution was specified because redd counts currently do not provide an estimate of the variance associated with any given annual estimate of population size (Gerrodette 1991, 1993). Finally, a two-tailed test with α = 0.05 was used for all considerations of the power of detecting population increases, whereas both a two-tailed test with α = 0.05 and a one-tailed test with α = 0.2 were used for the detection of population declines to reflect the asymmetry of the importance of detecting population declines over population increases (Rice and Gaines 1994).

Results

For each of the three rates of population change, the power of detecting a trend was greater for declines because declining and increasing effect size criteria do not result in the same absolute values of r for each absolute percentage decline and increase (Figures 2, 3; Table 1). Also, power increases with increasing numbers of sampling periods, increasing rates of population change and decreasing CVs (Figures 2, 3). For a 50% change per generation, all simulations except those using the largest CV value resulted in detecting a statistically significant change at or above a power of 0.8 with 10 or fewer years of redd counts for declines and 15 or fewer years of redd counts for increases (Figures 2C, F). Power associated with detecting 20% increases or declines in population size per generation using a two-tailed test with α = 0.05 remained below 0.8 throughout the first 15 years of monitoring for all but the lowest CV levels (Figures 2B, E). Power for detecting a 10% change in population size per generation, using a two-tailed test with α = 0.05, did not rise above 0.45 within the first 15 years of monitoring for any of the CV levels examined, and 19 and 35 years of redd counts were required before the power of detecting a statistically significant decline would rise to 0.8 at the lowest (0.178) and highest (0.472) CV values, respectively (Figures 2A, D; power for 35-year period not shown graphically).

Increasing α to 0.2 and using a one-tailed test raised the power of detecting 50% and 20% declines in population size per generation associated with all CV levels to 0.8 or higher with 7 (Figure 2C) and 15 (Figure 2B) years of redd counts, respectively. Using a one-tailed test with α = 0.2 also increased the power to detect a 10% decline per generation (Figure 2A), but power still remained below 0.8 during the first 15 years of monitoring for all but the smallest CV level. At a CV of 0.472 (characteristic of many bull trout popu-

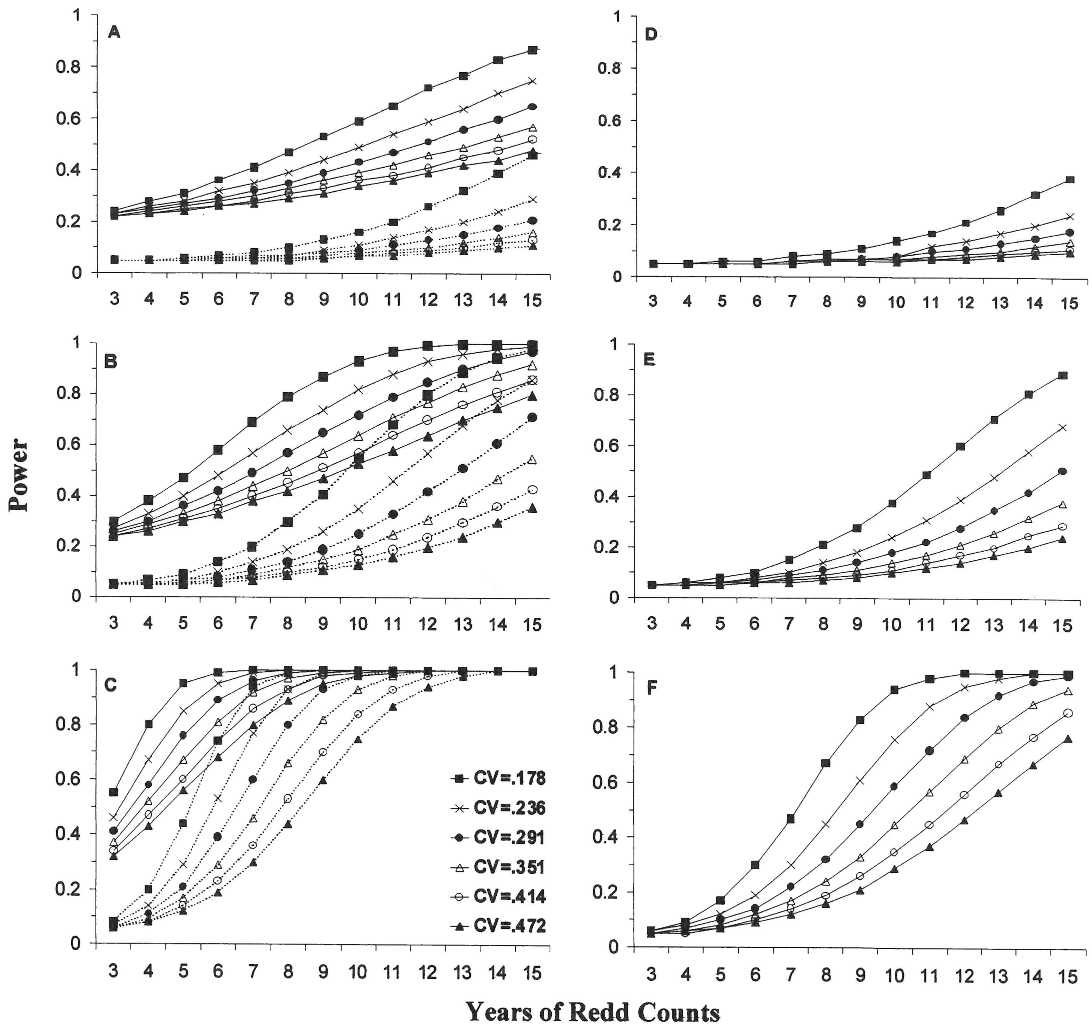

FIGURE 2.—Power of detection associated with six coefficient of variation (CV) values (0.178–0.472) for **(A)** 10%, **(B)** 20%, and **(C)** 50% declines in population size per generation, calculated with a two-tailed test and α = 0.05 (dashed lines) and with a one-tailed test and α = 0.2 (solid lines); also, power of detection for the same CV values with a **(D)** 10%, **(E)** 20%, and **(F)** 50% increase in population size per generation, calculated with a two-tailed test and α = 0.05.

lations in Montana), at least 25 years of monitoring were needed before power equaled 0.8 (Figure 3).

If one-tailed testing with $\alpha = 0.2$ was used and reduction of measurement error reduced CV levels from 0.472 to 0.291, time for detecting 10, 20 and 50% declines per generation could be reduced by 6, 4, and 2 years, respectively (Figure 3).

Discussion

Given the current status of bull trout populations throughout their range, the need to ensure rapid detection of further declines, and the need to evaluate the effects of recovery programs, it is critical for managers to establish or continue monitoring programs that will allow detection of trends in local stocks and regional populations. Redd counts are, and will probably continue to be, the best available means of detecting population trends (Rieman and McIntyre 1993; Rieman and McIntyre 1996; Rieman and Myers 1997). However, this prospective power analysis shows the limits of redd count data to detect changes in population size. With CV values that currently characterize many bull trout stocks in Montana and with traditional significance levels (α = 0.05) and two-tailed testing procedures, the power of detecting

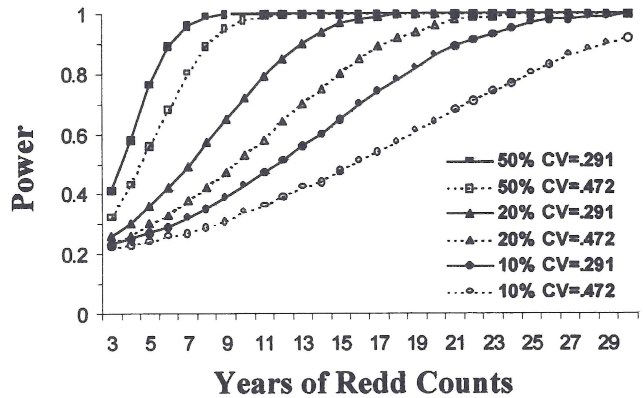

FIGURE 3.—Power of detecting three different declines in population size per generation (10, 20, and 50%) over 30 years of sampling as calculated with a one-tailed test, $\alpha = 0.05$, and coefficient of variation (CV) = 0.291 (solid lines) and CV = 0.472 (dashed lines); simulates the increases in power that might result from the identification and subsequent elimination of measurement error.

changes in population size remains low throughout the first 15 years of monitoring, unless the decline or increase is as high as 50% per generation (Figures 2, 3). Unfortunately, if declines are small and steady, populations could decline by more than 47% before the decline is detected (Figures 2, 3; Table 1).

These limitations indicate that, when preparing recovery plans, managers need to carefully consider the level of change that will be considered biologically significant, as well as the levels of statistical significance and the statistical tests that will be used to identify population trends. Lowering significance levels to $\alpha = 0.2$ and foregoing the ability to detect population increases by using a one-tailed test for only detecting population declines causes a significant increase in the power of detecting all levels of population decline associated with a variety of CV levels (Figures 2A–C) (Rice and Gaines 1994). More importantly, this allows population declines associated with a given CV value to have an 80% probability of being detected 2–10 years earlier. Increasing the power of detection by raising the value of alpha will require that managers weigh the relative risks of committing type I and type II errors (e.g., Mapstone 1995). However, the current status of bull trout stocks would seem to warrant lowering the traditional significance levels ($\alpha = 0.05$) and accepting the greater risk of type I errors because of the need to lower the risk of committing type II errors (Peterman 1990). Similarly, increasing statistical power by foregoing the ability to detect increases in population size when using one-tailed tests that only detect population declines will have to be carefully considered. For instance, without the ability to detect population increases, managers will only be able to evaluate the effectiveness of habitat restoration measures by testing to see if those measures did not result in population declines; managers would not be able to test for population increases. How will managers know whether bull trout have met the criteria for delisting or have recovered to fishable levels if they have given up the ability to detect population increases? One solution would be to use one-tailed testing procedures for a certain number of years (possibly 15 years or 3 bull trout generations) during the beginning of a monitoring program and then subsequently switch to using two-tailed testing procedures as long as no significant declines have been detected and redd numbers actually indicate an upward trend over time.

The only way to simultaneously reduce the risk of type I and type II errors is to reduce CV levels. Unfortunately, no extensive evaluation of the degree of measurement error involved in redd counts has been made to date (Rieman and McIntyre 1996). Thus, annual redd counts consist of only a single point estimate without any associated confidence intervals (Figure 1). This is of some concern because a variety of factors could potentially lead to measurement error. Redd numbers may be overestimated if (1) enumerated redds were made by other species, such as brown trout *Salmo trutta*; (2) enumerated redds involved brook trout *Salvelinus fontinalis* × bull trout hybridization or introgression; (3) spawning effort shifts from unmonitored to monitored reaches; (4) normal stream hydraulics create redd-like structures that are enumerated as redds; (5) bull trout test digs, in which eggs are not deposited, are enumerated as redds; or (6) enumerated redds are made by resident life history forms. On the other hand, redd numbers may be underestimated if (1) spawning activity occurs after an area has been surveyed; (2) redds are obscured by vegetation, periphyton, or high discharge; (3) spawning effort shifts from monitored to unmonitored reaches; or (4) redds are superimposed on one another. Evaluation of the extent of these measurement errors would identify the degree of precision associated with redd counts and enhance the ability of managers to make decisions that may ultimately determine whether or not the species will persist (Peterman and Bradford 1987; Peterman 1990; Reed and Blaustein 1997).

Finally, this power analysis shows that even if lower levels of statistical significance and one-tailed tests are used and the level of measurement error is identified and subsequently reduced, the time required for detecting declines in population size is still considerable. For example, if all of the above precautions are taken and measurement error was somehow reduced enough to lower the CV value to the lowest value considered (CV = 0.178), it would still require 14 years of monitoring before the power of detecting a steady 10% decline in population size per generation would rise to 0.8 (Figure 2A). Three generations of bull trout would have been produced during this period and the population could have declined to 73% of its original size before a statistically significant decline was detected. This lag suggests that alternative monitoring criteria should be explored, including: (1) monitoring trends at multiple spatial scales in order to identify underlying causes at the reach, tributary, or system scale; (2) monitoring adult abundance directly to increase the precision of population estimates; (3) using habitat-based criteria to ensure that habitat is being protected during the years in which declines have a low power of being detected; and (4) monitoring less-used and possibly lower-quality spawning reaches where declines may be likely to first occur as fish switch, because of less competition, to primary spawning areas. Unfortunately, each of these methods may be associated with a variety of assumptions, measurement errors, and logistical constraints, which may make their individual application no more beneficial than counting redds (e.g., Poole et al. 1997).

Another approach, suggested by Steidl et al. (1997), may also prove useful. They suggest that hypothesis tests only assess statistical significance, whereas biological significance may be better evaluated with confidence intervals. This approach identifies whether the effect is statistically significant (i.e., when confidence intervals do not overlap with zero) versus biologically significant (i.e., when the entire confidence interval is greater than the a priori established minimum biologically significant effect). This approach would still require that managers establish a priori levels of biological significance and evaluate the degree of measurement error in order to establish confidence intervals, but it would properly focus attention on the magnitude of the effect size and the precision of population estimates.

Acknowledgments

Mike Bader at the Alliance for the Wild Rockies pointed out the need for this and other projects involving bull trout in the state of Montana. Tom Weaver of the Montana Department of Fish, Wildlife and Parks, Kalispell, Montana, provided data on redd counts. Len Broberg, Chris Frissell, Scott Mills, Karen Pratt, Bruce Rieman, Andy Sheldon, Brad Shepard, Mark Taper, Mike Young, and an anonymous reviewer all made valuable suggestions that improved the manuscript.

References

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68:1364–1372.

Gerrodette, T. 1991. Models for power of detecting trends—a reply to Link and Hatfield. Ecology 72: 1889–1892.

Gerrodette, T. 1993. Program TRENDS: user's guide. Southwest Fisheries Science Center, La Jolla, California.

Hayes, J. P., and R. J. Steidl. 1997. Statistical power analysis and amphibian population trends. Conservation Biology 11:273–275.

Mace, G. M., and R. Lande. 1991. Assessing extinction threats: toward a reevaluation of IUCN threatened species categories. Conservation Biology 5:148–157.

Mapstone, B. D. 1995. Scalable decision rules for environmental impact studies: effect size, type I and type II errors. Ecological Applications 5:401–410.

Meffe, G. K. 1986. Conservation genetics and the management of endangered fishes. Fisheries 11(1):14–23.

Meffe, G. K., and C. R. Carroll. 1994. Principles of conservation biology. Sinauer, Sunderland, Massachusetts.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2–15.

Peterman, R. M., and M. J. Bradford. 1987. Statistical power analysis of trends in fish abundance. Canadian Journal of Fisheries and Aquatic Sciences 44: 1879–1889.

Poole, G. C., C. A. Frissell, and S. C. Ralph. 1997. Instream habitat unit classification: inadequacies for monitoring and some consequences for management. Journal of the American Water Resources Association 33:879–896.

Reed, J. M., and A. Blaustein. 1997. Biologically significant population declines and statistical power. Conservation Biology 11:281–282.

Rice, W. R., and S. D. Gaines. 1994. 'Heads I win, tails you lose': testing directional alternative hypotheses in ecological and evolutionary research. Trends in Ecology and Evolution 9:235–237.

Rieman, B. E., and J. D. McIntyre. 1993. Demographic and habitat requirements for conservation of bull trout. U.S. Forest Service General Technical Report INT-302.

Rieman, B. E., and J. D. McIntyre. 1996. Spatial and temporal variability in bull trout redd counts. North American Journal of Fisheries Management 16: 132–141.

Rieman, B. E., and D. L. Myers. 1997. Use of redd counts to detect trends in bull trout (*Salvelinus confluentus*) populations. Conservation Biology 11: 1015–1018.

Steidl, R. J., J. P. Hayes, and E. Schauber. 1997. Statistical power analysis in wildlife research. Journal of Wildlife Management 61:270–279.

USC (U. S. Code). 1973a. Endangered Species Act of 1973. U. S. Code, Section, Title 16, Chapter 35, Part 1533 (d).

USC (U. S. Code). 1973b. Endangered Species Act of 1973. U. S. Code, Section, Title 16, Chapter 35, Part 1533 (f).

USFWS (U.S. Fish and Wildlife Service). 1998. Endangered and threatened wildlife and plants; determination of threatened status for the Klamath River and Columbia River distinct population segments of bull trout. Federal Register 63:111(10 June 1998):31647–31674.