PAIRED– A tool to calculate the power of a paired design for monitoring fish response to habitat actions

Richard A. Hinrichsen

and

Rishi Sharma

August 29, 2011

Introduction

A simplified power analysis is presented to determine the sample sizes (numbers of sites and years) necessary to deliver a statistically reliable estimate of the effect of habitat actions on smolt abundance. We use the statistical technique of Liermann and Roni (2008) for a paired study used to explore designs that are least expensive while delivering the desired precision of the treatment effect estimate. In the paired design, it is assumed that each of *N* sites contains a control watershed, where no habitat action is taken, and a treatment watershed, where habitat is deliberately altered. The response variable is assumed in Liermann and Roni (2008) to be log(smolt abundance) and the nine assumptions are given in Table 1. The designs considered are before-after-control-impact (BACI) designs that have multiple pairs of control and treatment watersheds and (2) Before and After treatment collection of data (see for example, Keeley and Walters 1994; Solazzi et al. 2000). One important assumption is that the true treatment effects from adjacent sites are independent, and the treatment effect estimators from adjacent sites are also independent. Another important assumption is that the treatment effect is of a "press" type (Bender et al. 1984). That is, before treatment, the treatment-control difference at a given site has a fixed mean, then after treatment, the mean immediately shifts and persists at this new level. Though the analysis was developed with smolt abundance in mind, other response variables are possible and the model we consider fits into the framework of an odds ratio design (Skalski and Robson 1992).

The website www.onefishtwofish.net contains a web-based tool that implements this *a priori* power analysis. The assumptions of the analysis are given in Table 1. The code for implementing this power analysis (Appendix A) was implemented in R, a system for statistical computation and graphics (Venerables et al. 2010).

Table 1.—Assumptions used in the power analysis.[1]

A1	*Normally distributed errors and common variance*. At each site, the difference in log(smolts) between the treatment and control watersheds is normally distributed. The error variance is common to each site, but the mean differences are allowed to differ between sites.
A2	*"Press" type effects*. At each site *s*, during the Before period, the mean difference in log(smolts) between the treatment and control watersheds is constant, but then immediately shifts by θ_s with treatment.
A3	*Independent errors*. The random errors in the site-specific models describing the treatment and control differences are independent between sites.
A4	*Serially independent errors*. The random errors in the site-specific models describing the treatment and control differences are not serially correlated.
A5	*Normally distributed true effects*. The true site-specific treatment effects are allowed to differ between sites, but follow a normal distribution with fixed mean θ.
A6	*Independent error terms*. The random errors ($v_{t,s}$) in the site-specific models describing the treatment and control differences are independent of the errors (ε_s) describing the variability in the true treatment effect among sites, i.e., $\text{cov}(v_{t,s}, \varepsilon_s) = 0$.
A7	*Maximum likelihood estimator.* The estimator of each site-specific treatment effect is a maximum likelihood estimator (MLE).
A8	*Estimate of mean treatment effect*. The estimator of the mean treatment effect over all sites is the average of the site-specific treatment effect estimates.
A9	*Measurement error*. Measurement error in smolt abundance is ignored.

[1] These are assumptions for an idealized power analysis. If any of these assumptions change, the model equations will also need to be changed.

Methods

Mean treatment effect estimator.—The methods follow those of Liermann and Roni (2008) for a paired design, generalized to handle the case of different numbers of Before and After years. Let $smolts_{t,s,treatment}$ and $smolts_{t,s,control}$ be the observed number of smolts in year *t* at site *s* at the treatment and control sites, respectively. The observations are the site-specific differences in the log(smolts):

$$y_{t,s} = \log(smolts_{t,s,treatment}) - \log(smolts_{t,s,control}) . \quad (1)$$

The model equations used to describe the variability are

$$y_{t,s} = \mu_s + k_t \theta_s + v_{t,s}\,, \tag{2}$$

where μ_s is the mean difference between the treatment and control at site *s* before treatment, k_t is an indicator that is 0 before the restoration action and 1 after the restoration action, θ_s represents the true change in the mean change in log(smolts) at site *s* after the restoration action, $v_{t,s}$ is a normal random error with mean 0 and variance σ^2_{time}. The maximum likelihood estimator (MLE) of θ_s is the difference between the mean After observations and mean Before observations:

$$\hat{\theta}_s = \frac{\sum_{t=n1+1}^{n} y_{t,s}}{n_1} - \frac{\sum_{t=1}^{n_1} y_{t,s}}{n_2}\,, \tag{3}$$

where n_1 is the number of Before years, and n_2 is the number of After years. The conditional variance of the MLE is derived by the usual technique of inverting the Fisher information (Mood et al. 1974). This yields a variance of

$$\mathrm{var}\left(\hat{\theta}_s \mid \theta_s\right) = \sigma^2_{time}\left(\frac{1}{n_1} + \frac{1}{n_2}\right). \tag{4}$$

Notice that the conditional variance of the treatment effect is identical at each site because the time variance and numbers of Before years and After years are the same for each site.

The unknown parameter θ_s represents the true treatment effect size at site *s*. Theoretically, a true treatment effect at site *s* would be obtained if an infinite number of samples were taken from a site. It is assumed that true treatment effects follow a normal distribution with mean θ and fixed variance σ^2_{site}. The model for the true treatment effect may be written as

$$\theta_s = \theta + \varepsilon_s\,, \tag{5}$$

where ε_s is a normal random variable with mean zero and variance σ^2_{site}, and it is assumed that the site error is independent of the year error, i.e., $\mathrm{cov}(v_{t,s}, \varepsilon_s) = 0$.

The estimator of the true mean treatment effect is given by the sample mean of the estimated site-specific treatment effects:

$$\hat{\theta} = \sum_{s=1}^{N} \hat{\theta}_s / N \,. \tag{6}$$

This is the average log odds ratio statistic proposed by Skalski and Robson (1992) in an impact study design.

The variance of $\hat{\theta}$ is derived by applying the law of total expectation, which decomposes the variance into a weighted sum of the conditional variance of $\hat{\theta}_s$ and the variance of θ_s as follows:

$$\mathrm{var}\left(\hat{\theta}\right) = \mathrm{var}\left(\sum_{s=1}^{N} \hat{\theta}_s / N\right) \tag{7}$$

$$= E\left(E\left(\left(\sum_{s=1}^{N} \frac{\hat{\theta}_s}{N} - \theta\right)^2 \mid \theta_1, \theta_2, \ldots, \theta_N\right)\right)$$

$$= E\left(E\left(\left(\sum_{s=1}^{N} \frac{\hat{\theta}_s - \theta_s}{N} + \sum_{s=1}^{N} \frac{\theta_s}{N} - \theta\right)^2 \mid \theta_1, \theta_2, \ldots, \theta_N\right)\right)$$

$$= E\left(E\left(\sum_{s=1}^{N} \left(\frac{\hat{\theta}_s - \theta_s}{N} \right)^2 + 2\sum_{s=1}^{N} \left(\frac{\hat{\theta}_s - \theta_s}{N} \right) \left(\sum_{s=1}^{N} \frac{\theta_s}{N} - \theta \right) + \left(\sum_{s=1}^{N} \frac{\theta_s}{N} - \theta \right)^2 \mid \theta_1, \theta_2, \ldots, \theta_N \right) \right)$$

$$= \frac{\sum_{s=1}^{N} \mathrm{var}\left(\hat{\theta}_s \mid \theta_s\right)}{N^2} + \frac{\sum_{s=1}^{N} \mathrm{var}(\theta_s)}{N^2}$$

$$= \frac{\sigma_{time}^2}{N} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) + \frac{\sigma_{site}^2}{N}$$

This variance formula differs from that of Liermann and Roni (2008) because we allow the number of Before years (n_1) and After years (n_2) to differ. When $n_1 = n_2 = n/2$, then equation (7) is equal to the variance formula of Liermann and Roni (2008). Notice that in this formula, the variance of the true mean treatment effect may be made arbitrarily close to zero by increasing the number of sites (N), but not by increasing the number of years ($n = n_1+n_2$).

Statistical power. — The statistical power of a design was defined as the probability of rejecting the null hypothesis of no treatment effect when the true mean treatment effect differs from zero. We use the approach of fixing the type I error probability at α and using a two-sided hypothesis. Liermann and Roni (2008) use a one-sided alternative hypothesis, but we opt for a two-sided alternative because treatments may result in negative effects. The other necessary inputs for the power analysis are: the number of sites, the number of Before years, the number of After years, the conditional within-site variance (σ_{time}^2), and the variance in the true treatment effects over all sites (σ_{site}^2). Following the method of Liermann and Roni (2008), we assume that under the null hypothesis, $\hat{\theta} / SE(\hat{\theta})$ follows a central t-distribution with N-1 degrees of freedom. The critical t-value is then equal to the value of q such that $\Pr\{T_{N-1} >= q\} = \alpha / 2$ where T_{N-1} is a random variable that follows a central t-distribution with $N-1$ degrees of freedom. Power is then equal to

$$\Pi = F_{N-1,\mu}(-q) + 1 - F_{N-1,\mu}(q) \qquad (8)$$

where $F_{N-1,\mu}(x)$ is the cumulative distribution function of a non-central t distribution with *N*-1 degrees of freedom and noncentrality parameter $\mu = \theta / SE(\hat{\theta})$ (Appendix A). Experimenters often chose designs that will deliver a power of 0.80 or greater.

Example. — Take for example a paired design with N sites, where N is varied from 1 to 20, n1=2 before years, and n2=10 after years, the site-specific variance is equal to $\sigma^2_{time} = 0.47$, between-site variance equal to $\sigma^2_{site} = 1.19$, $\alpha = 0.05$ and true effect sizes range from 1.5 to 5.0. These variances were taken from Liermann and Roni (2008). Statistical power for this example is depicted in Figure 1. In each case a power of 0.80 may be obtained by increasing the number of sites to 8 or greater.

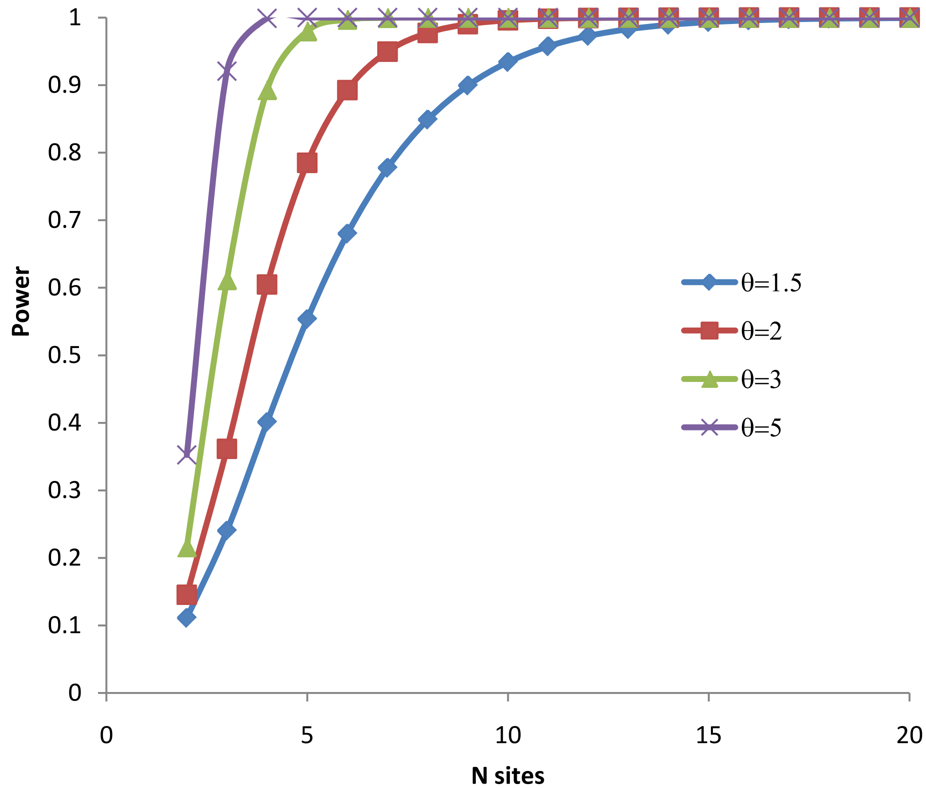

Figure 1. — Statistical power for the example varying the number of sites and the true effect size θ. In this example, type I error is $\alpha = 0.05$, n1=2, n2=10, $\sigma^2_{time} = 0.47$, $\sigma^2_{site} = 1.19$. These variances were taken from Liermann and Roni (2008). A power of 0.80 is achieved for all true effect sizes examined by using at least 8 sites.

Acknowledgements

This work was supported by the Bonneville Power Administration. We thank Martin Liermann for supplying documentation for calculating statistical power of the paired watershed design in Liermann and Roni (2008). Amber Parsons reviewed this documentation and tested the software at www.onefishtwofish.net.

References

Bender, E.A., T.J. Case, and M.E. Gilpin. 1984. Perturbation experiments in community ecology: theory and practice. Ecology 65:1-13.

Keeley, E. R., and C. J. Walters. 1994. The British Columbia watershed restoration program: summary of the experimental design, monitoring, and restoration techniques workshop. British Columbia Ministry of Environment, Land, and Parks and Ministry of Forests, Watershed Restoration Management Report 1, Victoria.

Liermann, M., and P. Roni. 2008. More sites or more years? Optimal study design for monitoring fish response to watershed restoration. North American Journal of Fisheries Management. 28:935-943.

Mood, A.M, Graybill, F.A., and D.C. Boes. 1974. Introduction to the theory of statistics, Third Edition. McGraw-Hill, New York, New York.

Skalski, J. R., and D. S. Robson. 1992. Techniques for wildlife investigations: Design and analysis of capture data. Academic Press. 237 pp.

Solazzi, M. F., T. E. Nickelson, S. L. Johnson, and J. D. Rodgers. 2000. Effects of increasing winter rearing habitat on abundance of salmonids in two coastal Oregon streams. Canadian Journal of Fisheries and Aquatic Sciences 57:906–914.

Venerables, W.N., Smith, D.M., and R Development Core Team. 2010. An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics Version 2.11.1 (2010-05-31). http://www.r-project.org/

Appendix A. R code used to calculate power to detect a significant shift in log(smolts) between treatment and control watersheds

```
# Program to implement a power analysis of a paired experiment
# as in Liermann and Roni (2008)
#N is the number of sites
#n1 is the number of Before years
#n2 is the number of After years
#theta is the true effect size which is the true shift in the difference
# in treatment and control log(smolts) between the Before and After periods.
#
lr<-function(vartime=1,varsite=1,N=20,n1=2,n2=8,theta=.02,alpha=0.05){

sd<-vartime*(1/n1+1/n2)+varsite
sd<-sqrt(sd)
se<-sd/sqrt(N)
cv<-se/theta
delta=theta
q<-qt(p=1-alpha/2,df=N-1)
power2<-1-pt(q,ncp=delta/se,df=N-1)+pt(-q,ncp=delta/se,df=N-1)

return(list(vartime=vartime,varsite=varsite,N=N,n1=n1,n2=n2,theta=theta,alpha=alpha,
se=se,cv=cv,power=power))
}
#outputs
#se -- standard error
#cv -- coefficient of variation
#power -- probability of rejecting the null hypothesis of no treatment effect
```