

**Pacific Lumber Company Habitat Conservation Plan:
A power analysis of the marbled murrelet inland monitoring
methods.**



**SUSTAINABLE
ECOSYSTEMS
INSTITUTE**

*0605 SW Taylor's Ferry Rd.
Portland, Oregon 97219*

Tel: 503-246-5008

Fax: 503-246-6905

e-mail: sei@sei.org

Website: <http://www.sei.org>

Prepared By

David Bigger
Sustainable Ecosystems Institute

and

Steven P. Courtney
Sustainable Ecosystems Institute

May, 2000

Introduction

Monitoring programs are set up to give an early warning, or to assess the effectiveness of a management plan. Unfortunately, many monitoring programs are unsatisfactory because they are not set up so as to detect changes efficiently and early. Our intention is to avert such problems by taking a proactive approach that builds on the current murrelet effectiveness monitoring program protocols. The first step is to establish whether or not the current design can detect any change in murrelet numbers. This step involves conducting a statistical power analysis that uses information collected in 1999 (see Courtney 1999 for details). It is understood that the 1999 data represent the baseline condition against which change will be assessed.

First, the effectiveness monitoring program is designed to address a single question:

Since the implementation of the PalCo HCP, is there a decline in the number of marbled murrelets in the Marbled Murrelet Conservation Areas (MMCAs) compared to murrelets in Headwaters and Humbolt State Park (reserves)?

In this report, we asked:

How does increasing the number of visits/stations in the MMCAs improve the probability of detecting a series of predefined decreases in the number of detections over time?

Background

The most effective way to test the above question is to design a well replicated experiment that has an experimental and a control group. In this experiment, murrelets in the reserve areas are in the control group, and the murrelets in the MMCAs are in the experimental group. Within each group, there are stands containing multiple stations (Table 1). Each station is visited several times during the field season. The assumption is that detections are relative to murrelet numbers.

During the 1999 field season, there were more murrelets detections in the control stands than experimental stands. In particular, murrelet detections were over six times greater in the control stands than experimental (Table 2a). In addition, the number occupied detections were 10 times greater in the control stands than in the experimental stands (Table 2b). In addition, the proportion of occupied detections was twice as high in the control stands than in the experimental stands (Table 2c).

To add to these results, there was a weak but significant positive correlation between the Julian day of the visit and detection (Figs. 1a & 1b). In other words, as one visits the stations later and later into the season, one is more likely to detect murrelets. This appears to be a common pattern observed in other studies (Miller and Ralph 1995, O'Donnell *et al.* 1995, White 1998).

An example of what might happen

Although there are several possible outcomes, we are primarily interested in cases where there is a decline in the number of detections in the MMCAs, particularly when there is little or no change in the number of detections in the reserves (see Figure 2). For now, let's assume that when we graph our results we discover a similar pattern as in Figure 1 -- a decline in the number of murrelet detections in the MMCAs with little or no change observed in the reserves. We then perform a conventional statistical test (in this case, a two-factor Analysis of Variance (ANOVA)) and find that there is a significant interaction between treatment and time. This result suggests that the murrelet detections are indeed declining in the MMCAs. However, what if there was not a statistically significant interaction? There are two possible explanations:

- 1) There is no decline in the number of detections in the MMCAs compared to the controls.
- 2) Alternatively, the power of the statistical test was too low to reject the null hypotheses that there is no significant decline over time. This is also known as a Type II error (Zar 1999).

The second explanation can be easily be tested by performing a power analysis. However, rather than waiting a few years to find out whether or not there was enough power to reject the null hypotheses, we decided to take a proactive approach. Below, we report the results of series of power analyses that we performed using the 1999 monitoring data to specifically explore how increasing sample size may reduce the chances of making a Type II error.

What is power analysis?

For starters, one needs an objective criterion for rejecting or not rejecting the null hypothesis. The power of a statistical test is defined as the probability of rejecting the null hypothesis when it is false and should be rejected (Zar 1999). In this case, the null hypothesis is that there is no change in murrelet detections or observed site occupancy behaviors in the HCP area. For the murrelet monitoring program, we need to know whether there are sufficient samples to detect a decline in murrelet numbers at the end of a monitoring period. In other words, if we do not find a statistically significant change, how confident are we that murrelet numbers are really not decreasing? In particular, are there enough visits (samples) to detect a 15%, 25%, 40% decline? If not, how many more samples do we need?

To estimate sample size, we have to know two things:

- 1) The minimum detectable difference between means. If we wish to detect a very small difference, we need a larger sample than if we wanted to detect a small difference.
- 2) An estimate of variation. If the variability between populations is great, then we will need a larger sample to detect differences between the means.

ANOVA models and Power analyses

We conducted two sets of power analyses that used the numbers of murrelet detections and of occupied detections from stands in the monitoring study (Table 1). The first set of analyses used information from every stand in the monitoring study. The second set of analyses only used information from stands where murrelets were observed with even medium frequency in 1999 (thus excluding the Shaw, Cooper Mill and Grizzly Creek stands).

The murrelet monitoring program is setup as a simple 2 x 2 ANOVA factorial design where the dependent variable is the number of detections per visit. The first factor consists of stands in control treatment where the control stands are outside of the HCP area (the reserves), and the experimental stands are inside the HCP area (the MMCAs). The second factor is time (before & after), where *before* represents the first survey conducted after the signing of the PalCo HCP and *after* is sometime in the future (*e.g.* one year, three years, etc...). The relationship between detection and treatment and time as a two-factor ANOVA while the relationship between detections and Julian days (Fig. 1) may be thought of as a regression analysis. These concepts are combined into a single analysis called an Analysis of Covariance (ANCOVA) where Julian days are a covariate in the model (Zar 1999). To insure that the assumption of normality and homoscedasticity of the variances for ANOVA were satisfied, a square root transformation $((x + 0.5)^{0.5})$ was used on the detection data (Fig. 3).

Note that this design is simply a before and after (BACI) experiment where the percent decline in detections represents the change from initial conditions. For instance, a 50% decline represents a 50% change from year 1 (1999). Clearly, this approach will not detect small changes over a short time-period. However, the cumulative small annual declines would be detectable over a longer period. For instance a 6% annual rate of decline would result in an overall decline of 27% after 5 years, or 46% decline after 10 years. The proposed monitoring program is expected to extend for 50 years.

It is also important to note that there is a family of ANOVA models to choose from, and it is our intention to explore a variety of alternative ANOVA models including nested models. Hence, the final model used to test the primary question in the monitoring program may differ from the one used in this report. In addition, other analyses may be appropriate after sufficient data have been collected. For instance, a trend analysis (or other regression analyses or using minimum likelihood approaches) may be possible after several years have elapsed. It is probable that such methods would yield additional information, and might have more explanatory power.

For this power analysis, we explored how increasing sample size (the number of stations and number of visits) in the experimental treatment improves the probability of a detecting a decrease in the proportion of occupied detections in the experimental plots. In this case, the number of control samples remains constant ($n = 60$). Thus, there will be unequal replication between treatments when sample size within the experimental treatment is increased. To account for unequal sample sizes, we calculated the harmonic mean of the two sample sizes:

$$n = 2n_c n_e / (n_c + n_e)$$

where n is the harmonic mean of the two sample sizes, n_c is the sample size for the control treatment, and n_e is the sample size experimental treatment.

To estimate the power of an ANOVA when a minimum detectable difference is specified, we computed

$$phi = (nd^2/2ks^2)^{1/2}$$

where n is the number samples (visits), d is the difference between treatment means, k is the number of treatments, and s^2 is the error MS. Note that this is a decline in the population means (see plus signs in Figure 1). Next, we compute the degrees of freedom

$$v_1 = k - 1$$

$$v_2 = k(n - 1).$$

After obtaining all of the necessary information, we consulted Appendix Fig. B1a (Zar 1999) to determine the power of the test for $\alpha = 0.05$. Although there are no universally accepted levels for statistical power, the standard is 0.80 (Cohen 1988).

Results

How does increasing the number of visits/stations in the MMCAs improve the probability of detecting a series of predefined decreases in the number of detections over time?

Based on the current design of 60 samples in both controls and experimental areas, the power of detecting at minimum a 24% decline in all detections is very high, 0.93, and well above the “standard” of 0.80 (Table 3a). Further, the addition of 110 visits (by adding more stations) in the MMCAs sharply improves the chances of detecting a 13% decline (Table 3a). Although the same trend holds for detecting declines in the number of occupied detections (Table 3b), the current design may not have enough power to detect a 20% in the number occupied detections. However, the addition of more stations in the MMCAs clearly increases the probability of detecting a 20% decline (Table 3b).

Analytical approaches that reduce variation while increasing sample size made an improvement towards increasing power. In particular, the exclusion of the Shaw, Cooper Mill and Grizzly Creek stands (stands with few or no detections) improved the probability of detecting a 12% decrease in all detections and occupancy detections (Tables 4a & 4b). Again, adding more stations in the MMCAs increased power.

Recommendations

Overall, the results are very encouraging. These series of power analyses suggest that the *current* murrelet monitoring study is capable of detecting a moderate decrease in the number of detections. However, we stress that increasing the effort to include 110 additional visits to experimental plots would result in increased ability to detect change (for instance a 50% increase in such ability at the 17% level in Table 4b). However, these benefits do not increase steadily: adding a further 80 visits increases our ability to detect changes by only another 1%.

We recommend increasing the number of stands used in the experimental treatment while maintaining at least five visits per station per year. Currently, we are exploring various methods to increase power by reducing variation, including the use of additional covariates (repeated measure designs), and of non-parametrics.

Finally, the power analysis suggests that one may need 60-170 visits between treatments to detect a decline. Due to the amount of effort and limited size of individual stands, it is not feasible to address similar types of questions within a stand.

References

- Cohen, J. 1988. Statistical power analysis for the behavioral sciences. 2nd. Edition. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey. Pp. 567.
- Courtney, S.P. 1999. Pacific Lumber Co. Habitat Conservation Plan: Marbled murrelet inland monitoring report.
- Miller, S. L. and C. J. Ralph. 1995. Relationship of marbled murrelets with habitat characteristics at inland sites in California. *In*: Ralph C. John; Hunt, George L., Jr.; Raphael, Martin G.; Paitt, John F., Technical Editors. 1995. Ecology and conservation of the Marbled Murrelet. Gen. Tech. Rep. PSW-GTR-152. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 420 p.
- O'Donnell, B. P., N. L. Naslund, and C. J. Ralph. 1995. Patterns of seasonal variation of activity of marbled murrelets in forested stands. *In*: Ralph C. John; Hunt, George L., Jr.; Raphael, Martin G.; Paitt, John F., Technical Editors. 1995. Ecology and conservation of the Marbled Murrelet. Gen. Tech. Rep. PSW-GTR-152. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 420 p.
- White, G. C. 1998. Logistic regression model of detection probability for marbled murrelets. *In*: Volume IV: Habitat conservation plans: Sustained yield plan/habitat conservation plan, Public Review Draft, July 1998.
- Zar, J. H. 1999. Biostatistical Analysis. 4th. Edition. Prentice-Hall, Inc., Upper Saddle River, New Jersey. Pp. 663.

Table 1. Treatments, stands, stations and number of visits during the 1999 field season.

CONTROL			EXPERIMENT		
Stand	Station	Visits	Stand	Station	Visits
Headwaters	C-M2A-05	8	Allen Creek	H-M01-07	8
	C-M2A-07	7		H-M01-24	6
	D-M01-03	10	Shaw	H-M04-05	6
	D-M01-04	8		H-M04-07	6
	D-M01-05	6	Cooper Mill	H-M08	6
State	Z-M01-01	7	Bell	H-M12-04	7
	Z-M01-02	8		H-M12-06	7
	Z-M01-03	8	Grizzly creek	I-M02-13	7
	Z-M01-07	8		I-M02-05	7
	Z-M01-08	9		I-M02-16	6

Table 2a. Total murrelet detections in 1999. In this and all following tables, *n* is the number of visits.

Treatment	Mean (SD)	<i>n</i>	Median	Range
Control	18.6 (20.0)	79	11	0-73
Experiment	3.3 (7.1)	66	0	0-44
Experiment^a	7.7 (9.3)	28	4	0-44

^a Shaw, Cooper Mill, and Grizzly stands excluded.

Table 2b. Occupied murrelet detections in 1999.

Treatment	Mean (SD)	<i>n</i>	Median	Range
Control	5.5 (8.5)	79	3	0-45
Experiment	0.3 (0.8)	66	0	0-4
Experiment^a	0.6 (1.1)	28	0	0-4

^a Shaw, Cooper Mill, and Grizzly stands excluded.

Table 2c. Proportion of murrelet detections that were classified as occupied in 1999.

Treatment	Mean (SD)	<i>n</i>	Median	Range
Control	0.31 (0.24)	67	0.3	0-1.00
Experiment	0.17 (0.17)	27	0.0	0-0.50

Table 3a. Probabilities of detecting a difference for all detections as a function of the total number of visits (n) in experimental stands. Julian days were a covariate. Detections were $(x + 0.5)^{0.5}$ transformed (population mean = 2.72, error MS = 2.76). The difference is a percentage of the back transformed population mean (10.3). The number of visits in the control stands in this and all following tables is 60, and $\alpha = 0.05$.

Difference	Exp.	
	$n = 60$	$n = 170$
13%	0.55	0.78
24%	0.93	0.99
33%	0.99	0.99

Table 3b. Probabilities of detecting a difference in the number of occupied detections as a function of the total number of visits (n) in experimental stands. Julian days were a covariate. Detections were $(x + 0.5)^{0.5}$ transformed (population mean = 1.49, error MS = 0.92). The difference is a percentage of the back transformed population mean (4.0).

Difference	Exp.	
	$n = 60$	$n = 170$
19%	0.50	0.72
30%	0.91	0.98
39%	0.99	0.99

Table 4a. Probabilities of detecting a difference for all detections (except for Shaw, Cooper Mill, and Grizzly stands) as a function of sample size in the experimental treatment. Julian days are a covariate. Detections were $(x + 0.5)^{0.5}$ transformed (population mean = 3.41, error MS = 3.08). The difference is a percentage of the back transformed population mean (15.3).

Difference	Exp.	
	<i>n</i> = 60	<i>n</i> = 170
12%	0.74	0.89
23%	0.99	0.99
32%	0.99	0.99

Table 4b. Probabilities of detecting a difference in the number of occupied detections as a function of the total number of visits (*n*) in experimental stands. Shaw, Cooper Mill, and Grizzly stands were not included. Julian days are a covariate. Detections were $(x + 0.5)^{0.5}$ transformed (population mean = 1.77, error MS = 1.19). The difference is a percentage of the back transformed population mean (5.2).

Difference	Exp.	
	<i>n</i> = 60	<i>n</i> = 170
17%	0.58	0.76
28%	0.94	0.99
37%	0.99	0.99

Figure 1a. Relationship between all detections and Julian date (Pearson $r = 0.382$, $n = 145$, $p < 0.001$).

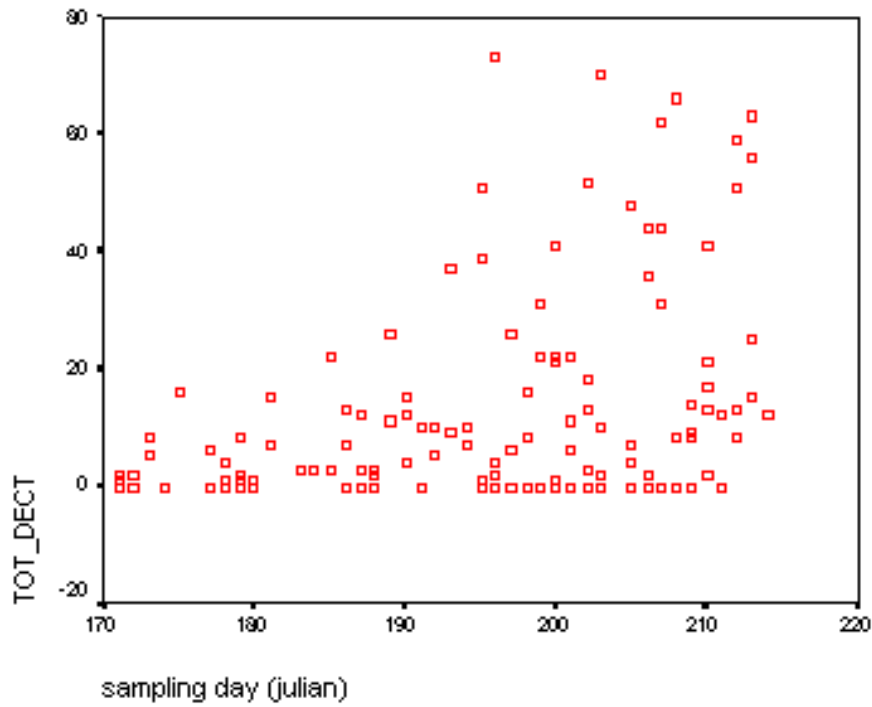


Figure 1b. Relationship between all detections and Julian date (Pearson $r = 0.263$, $n = 145$, $p < 0.001$).

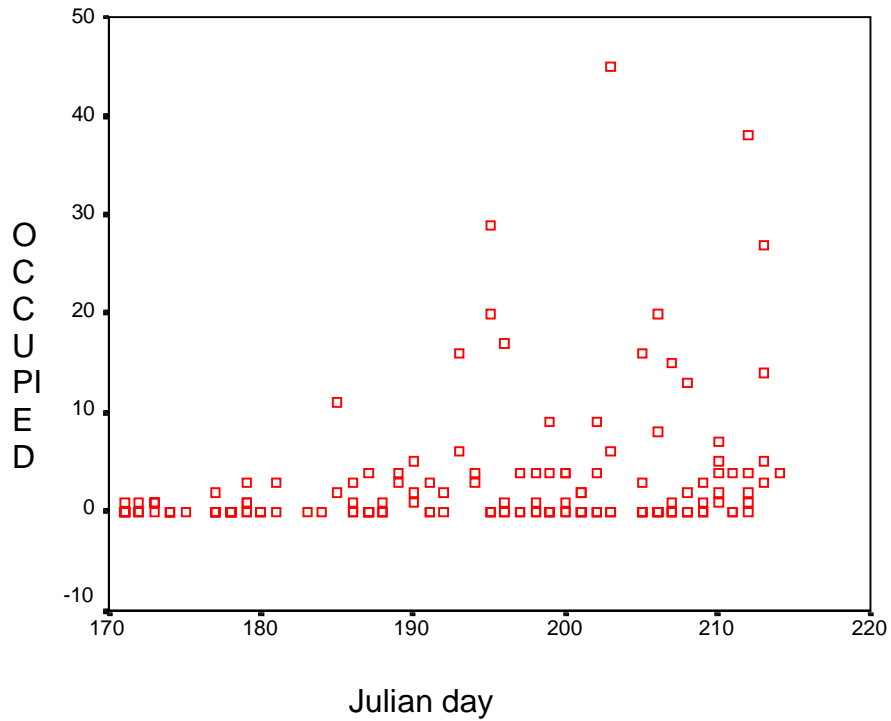


Figure 2. Means in a two-factor ANOVA, showing the effects of the two factors and their interaction. In this case, there is a large effect of Treatment (note the difference between Control and Experiment), a large effect of Time (note the difference between plus signs between Before and After time intervals), with a large interaction (Treatment \times Time). This interaction represents a decrease in the number detections in Experiment group and no change in the Control group over Time. The plus sign represents the population mean.

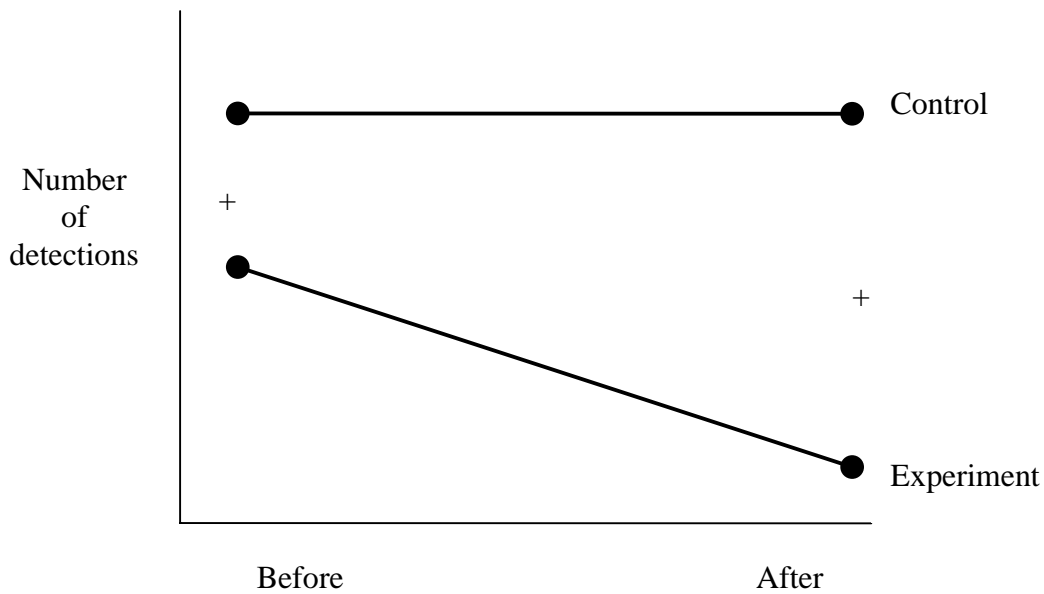


Figure 3a. Histogram showing the frequency of all detections from the 1999 study.

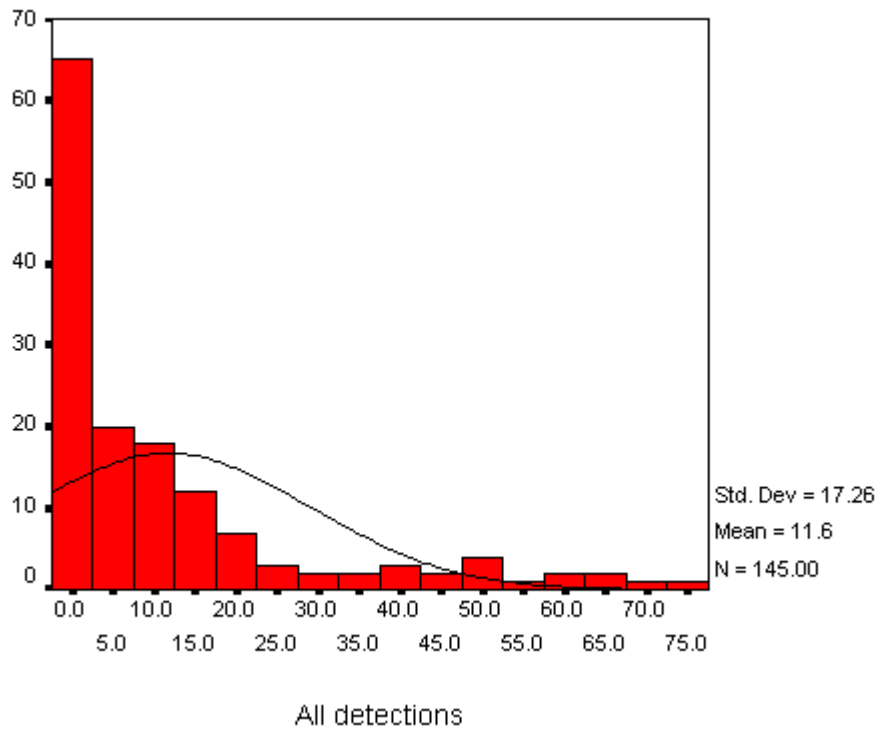


Figure 3b. Histogram showing the frequency of the square root transformed detections from the 1999 study and a normal curve.

