THE ANALYSIS OF THE TWO-PERIOD REPEATED MEASUREMENTS CROSSOVER DESIGN WITH APPLICATION TO A FORESTRY PROBLEM

Armando Garsd

Maria C. Fabrizio

Maria V. López

Recommended Citation

THE ANALYSIS OF THE TWO-PERIOD REPEATED MEASUREMENTS CROSSOVER DESIGN WITH APPLICATION TO A FORESTRY PROBLEM

by Armando Garsd, María C. Fabrizio and María V. López

Department of Statistics, School of Agricultural Engineering, University of Buenos Aires

Abstract

The two-period repeated measurements crossover design is not often used in agricultural studies. It is, however, an attractive model, involving the confluence of two powerful statistical ideas, treatment crossover and repeated measurements on the same experimental unit. This paper presents one approach for the statistical analysis of such design based on the work of Wallenstein and Fisher (1977). It is shown how the data may be transformed so that it can be analyzed under the framework of a completely randomized repeated measurements design. We formalize the analysis in the context of a forestry experiment conducted on poplar trees (*Populus sp.*), to compare the efficacy of two treatments to prevent damage by the coleopteran insect *Platypus sulcatus* (ambrosia beetle). Two insecticides were applied in a crossover fashion to two groups of 8 poplar trees each. Each tree was treated with one insecticide and evaluated on three occasions during the first year, received no treatment during the following one-year washout phase, and then (in the third year) received the other treatment and was evaluated on three occasions. One of the parameters analyzed to test for treatment differences was the number of tree lesions attributed to the insect. We present the results of our work and discuss the potential usefulness as well as the limitations of this interesting design.

1. Introduction

We present in this paper a real example of a forestry experiment involving a two-period repeated measures crossover (RMC) design. Since this type of design is not very common in agricultural studies, we thought our approach to analyze it might be of some interest to an audience of agricultural statisticians.

Some sixty five thousand hectares (about 161,000 acres) of land in the Parana river delta in Argentina are currently forested with poplar trees (*Populus sp.*) and related species. Despite the commercial promise of these forests their productivity is seriously threatened by various pests, including a coleopteran insect of the ambrosia beetle type (*Platypus sulcatus*). The intertwined dynamics of insect and forest populations pose permanent challenges to the

forest scientist, demanding periodical tests for alternative, more efficacious insecticides, as well as other forms of environmentally sound pest control.

The present study was designed to evaluate the comparative efficacy of two treatments to prevent damage to poplar trees caused by these beetles. Two insecticides were applied in a crossover fashion to two groups of 8 poplar trees each. Each tree was treated with one insecticide and evaluated on three occasions during the first year, received no treatment during the following one-year washout phase, and then (in the third year) received the other treatment and was evaluated on three occasions. One of the parameters analyzed to test for treatment differences was the number of active tree sites (lesions) attributed to the insect.

The RMC model comprises elements from two powerful experimental designs, treatment crossover and repeated measures on the same experimental unit. The statistical treatment of the resulting RMC model dates back at least to Lucas (1950), although we follow the unifying approach developed by Wallenstein and Fisher (1977).

In section 2 we introduce the basic RMC model. In section 3 we provide more details on the experimental methods. In section 4 we present the experimental results and apply the RMC model to the poplar data. Finally, in section 5, we discuss the usefulness as well as the limitations of the RMC design.

2. The RMC model

Data from repeated measures on the same experimental unit induce a multidimensional space. Therefore, general statistical models for this type of data correspond to multivariate analysis, possibly under homogeneous covariance structure. For small sample sizes, this analysis is not very powerful. However, under certain simplifying assumptions the generalized repeated measures model can be reduced to a powerful univariate model. When each experimental unit has been exposed to both treatments, statistical power can be further enhanced by comparing treatment results within each experimental unit. This is accomplished by taking the appropriate difference between period 1 and period 2.

Following Wallenstein and Fisher, 1977, let Y_{ijkm} be the observation for the jth experimental unit in the ith sequence of administration group at the kth period and mth time point within the period ($j=1,...,8$; $i=1,2$; $k=1,2$; $m=1,2,3$). Experimental units in sequence 1 receive treatment A in period 1 and treatment B in period 2; experimental units in sequence 2 receive the treatments in reverse order.

A repeated measurements model with carryover effects can be written

as

$$Y_{ijkm} = \mu + \delta_{2k}\lambda_i + \epsilon_{ij} + \phi_l + e_{ijk} + \tau_m + \delta_{2k}(\lambda\tau)_{im} + \omega_{ijm} + (\phi\tau)_{lm} + f_{ijkm} \quad (1),$$

and if carryover effects are not present, the model can be written as

$$Y_{ijkm} = \mu + \xi_i + \epsilon_{ij} + \pi_k + \phi_l + e_{ijk} + \tau_m + (\xi\tau)_{im} + \omega_{ijm} + (\pi\tau)_{km} + (\phi\tau)_{lm} + f_{ijkm} \quad (2),$$

where, in this example, $i = 1, 2$; $j = 1, \ldots, 8$; $k = 1, 2$; $l = i - (-1)^i\delta_{2k}$; $m = 1, 2, 3$; $\delta_{2k} = 1$ if $k = 2$, $\delta_{2k} = 0$ if $k = 1$.

The fixed sources of variation are:
μ: the overall mean,
ξ_i: the sequence effects,
λ_i: the carryover treatment effects,
ϕ_l: the direct treatment effects,
τ_m: the time effects,
π_k: the period effects,
and the interaction of time with sequence, period, and treatment (direct and carryover) effects.

The random effects in the repeated measurements crossover design are:
ϵ_{ij}: the effect of the jth experimental unit in the ith sequence,
e_{ijk}: the effect of the jth experimental unit in the ith sequence at the kth period,
ω_{ijm}: the effect of the jth experimental unit in the ith sequence at the mth time,
f_{ijkm}: the random fluctuation of the jth experimental unit in the ith sequence at the kth period and mth time.

All terms of fixed sources of variation except carryover effects follow the standard constraints, that is, main effects sum to zero, and the interaction effects sum to zero over both the t-levels (t=3) of time and the two levels of the main treatment effect. For carryover effects, the only constraint is that the interaction effects sum to zero over the three time points.

The error terms are distributed independently and identically between experimental units, and the set of four error terms are independently distributed.

By taking differences of the observations at each time point,

$$d_{ijm} = Y_{ij1m} - Y_{ij2m} \quad (3),$$

the repeated measurements crossover design is thus transformed to a completely randomized repeated measurements design, and in this way, it allows for a clear understanding of the methods of analysis and their assumptions.

Substitute expression (3) into equation (2) to obtain

$$d_{ijm} = \pi + \Psi_i + a_{ij} + (\pi\tau)_m + (\Psi\tau)_{im} + b_{ijm} \qquad (4),$$

where
$\pi = \pi_1 - \pi_2$,
$\Psi_i = (-1)^i(\Phi_2 - \Phi_1)$,
$a_{ij} = e_{ij1} - e_{ij2}$,
$(\pi\tau)_m = (\pi\tau)_{1m} - (\pi\tau)_{2m}$,
$(\Psi\tau)_{im} = (-1)^i((\Phi\tau)_{2m} - (\Phi\tau)_{1m})$,
$b_{ijm} = f_{ij1m} - f_{ij2m}$.

Equation (4) is the model for the t-time completely randomized repeated measurements design.

The carryover effects and the sequence and time effects are evaluated in terms of s_{ijm}, sums of the observations at each time point,

$$s_{ijm} = Y_{ij1m} + Y_{ij2m} \qquad (5),$$

substitute the expression (5) into equation (1),

$$s_{ijm} = 2\mu + \lambda_i + \alpha_{ij} + 2\tau_m + (\lambda\tau)_{im} + \beta_{ijm} \qquad (6),$$

where
$\alpha_{ij} = 2\epsilon_{ij} + e_{ij1} + e_{ij2}$
$\beta_{ijm} = 2\omega_{ijm} + f_{ij1m} + f_{ij2m}$

Equation (6) is not the model for the completely randomized two repeated measurements design since it does not assume $\lambda_1 + \lambda_2 = 0$, but it can be transformed to such setting

$\eta = 2\mu + (\lambda_1 + \lambda_2)/2$
$\upsilon_i = (-1)^i(\lambda_2 - \lambda_1)/2$
$\theta_m = 2\tau_m + ((\lambda\tau)_{1m} + (\lambda\tau)_{2m})/2$
$(\upsilon\theta)_{im} = (-1)^i((\lambda\tau)_{2m} - (\lambda\tau)_{1m})/2$

to obtain
$$s_{ijm} = \eta + \upsilon_i + \alpha_{ij} + \theta_m + (\upsilon\theta)_{im} + \beta_{ijm} \qquad (7)$$

Equation (7) is used to test for differences in carryover effects and for a carryover effect by time interaction.

For the complete univariate analysis for carryover effects, it is necessary to assume that each set of error terms $\{\epsilon_{ij}\}$, $\{e_{ijk}\}$, $\{\omega_{ijm}\}$, $\{f_{ijkm}\}$ be independent and identically normally distributed with variances σ_ϵ^2, σ_e^2, σ_ω^2, σ_f^2, respectively. This assumption is satisfied if the variance-covariance matrix of $(Y_{ijk1}, Y_{ijk2}, Y_{ijk3})$ is uniform, wich implies that the time interval between measurements does not affect the correlations. The tests described by Box (1950)

can be used to test for the validity of assumptions in the tests for carryover effects (based on sums of observations) and the tests for period and direct treatment effects (based on differences of observations).

For the analysis of treatment effects to be performed by univariate analysis of variance, it is only required that the variance-covariance matrix of the differences over time be uniform (spherical), that is,

$$\text{Var}(a_{ij}) = \sigma_a^2 \qquad\qquad \text{Var}(b_{ijm}) = \sigma_b^2$$
$$\text{Cov}(a_{ij}, a_{ij'}) = \text{Cov}(b_{ijm}, b_{ij'm}) = 0$$

where $i = 1, 2$; $j = 1,\ldots,8$; $m = 1, 2, 3$ and $j \neq j'$.

3. The Experiment

Poplar trees from a homogeneous cohort were randomly assigned to one of two groups, or sequences. The insecticides tested were a piretroid derivative (A) and a carbamic acid compound (B). Both insecticides target adult populations. Trees were sprayed in Spring (southern hemisphere), covering timber and branches up to 7 m high from the ground. To assess insect damage, the experimental unit was restricted to the tree trunk, from the base to a height of 2.5 m. The efficacy of each treatment was evaluated by counting (and marking) the number of active insect sites, or lesions. Previous baseline counting was performed on the same day the corresponding treatment was applied. Table 1 summarizes the schedule of experimental activities. Evaluations were not equally spaced.

4. Experimental results and statistical analysis

Results from the poplar experiment are presented in Table 2. Randomization resulted in balanced groups with respect to baseline number of active sites. Figure 1 shows the average number of active sites as a function of time. The parallel nature of the lines within each group and the similarity of the levels between the groups point to the absence of both carryover effect and periodicity. Figure 2 displays the average difference in the number of active sites between treatments A and B. The Figure shows no convincing evidence of time-related interactions. These observations go well with the expectations for a one year insecticide-free washout period.

We found the square-root transformation more suitable for the analysis of this data set, as it tended to make the variances more homogeneous.

A cursory repeated measures analysis of variance on the baselines (one per period) showed no significant difference in baseline

levels, between sequences or between periods. Likewise, a preliminary test (not shown) for the significance of differential carryover effects based on a derivation from equation 2 was not significant (p= 0.80).

The analysis based on equation 1 is presented in Table 3. As expected, the analysis shows a significant overall advantage for treatment A. No other effect is statistically significant. Inclusion of difference in baselines did not alter this conclusion.

5. Discussion

Studies involving repeated measures design presumably imply that the investigator has an interest not only in overall differential treatment effects but also in the evolution through time of any relative treatment advantage. In our data we found no evidence of overall time effects or treatment by time interactions that might account, for example, for progressive loss of insecticide residual effects. We found evidence that the advantage established in favor of one treatment (A) within a week of application remains tangible throughout each period. This effect may just express the separate, direct and persistent action of two compounds of different performance or, alternately, an early but durable differential impact of compound A on the population dynamics of the targeted insects. Finally, it should also be kept in mind, perhaps as a technicality, that due to the small sample sizes involved in this particular experiment, the lack of power in the statistical tests applied may have impaired the detection of any time effects.

The data met the sphericity and homogeneity assumptions required by the repeated measures model in general. But the time scale was not equally spaced. Notice, however, that the test for overall treatment differences is always valid, regardless of these assumptions (Arnold, 1981).

The baselines were not included as responses in the main repeated measures analysis, although this is in principle possible. However, overall treatment differences in such model correspond to interactions, which may be cumbersome to interpret in the context of equation 1.

Our approach to hypothesis testing is the standard one to show relative efficacy. It is also possible to analyze this experiment in terms of testing for the statistical equivalence of efficacy between the two formulations. There has been recent useful work on this topic for crossover designs (Chow and Liu, 1992).

Crossover designs afford distinct advantages over parallel designs. In particular, they require less experimental units and produce estimates with smaller residual variance. This advantage stems from

the availability under crossover of within-unit comparisons.

However, the crossover design has drawbacks (Senn, 1995). One aspect that deserves careful consideration in the poplar experiment is whether the experimental units at the end of the washout period are at a level comparable to their status at the initial baseline. While we found no evidence of a (virtual) carryover effect, in general this possibility must be explicitly accounted for (Jones and Kenward, 1989). For example, we have assumed that damage to the trees during the second period does not depend on existing damage from the first period or on the density of insects at the end of the first period. If, however, due to disparate treatment effects during the first period, the design becomes severely unbalanced in the number of active sites per group at the second baseline, then the interpretation of the overall treatment differences from the whole experiment, as in equation 1, would be more problematic one could use data from the first period only, but defeating the intended crossover design may impair the power of the ensuing tests.

Acknowledgements

References

Arnold, S. F. 1981. *The theory of linear model and multivariate analysis*. John Wiley, pp. 432.

Box, G. E.P. 1950. Problems in the analysis of growth and wear curves . *Biometrics* 6: 362-389.

Chow, S. C., and Liu, J. P. 1992. *Design and analysis of biovailability and bioequivalence studies*. M. Dekker, pp. 88-125.

Jones, B., and Kenward, M. G. 1989. *Design and analysis of cross-over trials*. Chapman and Hall, pp. 16-88, 242-266.

Lucas, H. L. 1950. Design and analysis of feeding experiments with milking dairy cattle. Mimeo Series #18, Institute of Statistics, North Carolina State University.

Senn, S. J. 1995. Crossover at the crossroads. *Applied Clinical Trials*, 4: 24-31.

Wallenstein, S., and Fisher, A. C. 1977. The analysis of the two-period repeated measurements crossover design with application to clinical trials. *Biometrics*, 33: 261-269.

Table 1
Schedule of experimental activities

	Period 1	Period 2
Treatment group 1	A	B
Treatment group 2	B	A
Date of treatment	9-10-92	9-10-94
Evaluation dates	9-15-92, 5-15-93, 9-15-93	9-15-94, 5-15-95, 9-15-95

*Baseline counting took place on this date, just before treatment proper.

Table 2
Results from poplar experiment: number of active insect sites

Untransformed data

Group 1, $n_1 = 8$

Treatment A					Treatment B			
EU	D0	D1	D2	D3	EU	D0	D1	D2	D3
1	7	4	4	4	1	8	2	2	2
2	7	6	5	5	2	8	4	4	3
3	8	7	7	6	3	8	4	5	6
4	7	4	4	4	4	8	5	4	3
5	6	5	5	4	5	7	4	4	3
6	6	3	2	3	6	6	4	3	3
7	7	4	4	4	7	5	3	4	4
8	7	5	5	4	8	6	4	3	3
Mean	6.88	4.75	4.50	4.25		7.00	3.75	3.63	3.38
SD	0.64	1.28	1.41	0.89		1.20	0.99	0.92	1.19

EU: Experimental unit. SD: Standard deviation.
D0, D1, D2, D3: Baseline, first, second and third
evaluation dates, respectively;
see Table 1.

Transformed data[1]

$$d_{1jm} = \sqrt{x_{ij2m}} - \sqrt{x_{ij}}$$

d_{1j0}	d_{1j1}	d_{1j2}	d_{1j3}	$d_{1j.}$
-0.18	0.59	0.59	0.59	1.77
-0.18	0.45	0.24	0.50	1.19
0.00	0.65	0.41	0.00	1.06
-0.18	-0.24	0.00	0.27	0.03
-0.20	0.24	0.24	0.27	0.75
0.00	-0.27	-0.32	0.00	-0.59
0.41	0.27	0.00	0.00	0.27
0.20	0.24	0.50	0.27	1.01
$d_{1.m:}$ -0.13	1.93	1.66	1.90	

$d_{1..}$: 5.49

(1) m= 0 denotes baseline.

(continues on next page)

Table 2 (continued)

Untransformed data

Group 2, $n_2 = 8$

	Treatment B				Treatment A			
EU	D0	D1	D2	D3	EU	D0	D1	D2	D3
1	6	4	3	2	1	7	4	4	3
2	7	4	3	3	2	7	4	4	3
3	8	5	5	3	3	7	5	5	4
4	8	3	3	3	4	7	5	4	4
5	7	4	4	3	5	8	5	5	5
6	7	4	4	3	6	6	6	5	4
7	8	3	4	3	7	7	2	3	3
8	8	5	4	5	8	6	4	5	5
Mean	7.38	4.00	3.75	3.13		6.88	4.38	4.38	3.88
SD	0.74	0.76	0.71	0.84		0.64	1.19	0.74	0.84

EU: Experimental unit. SD: Standard deviation.
D0, D1, D2, D3: Baseline, first, second and third
evaluation dates, respectively;
see Table 1.

Transformed data [1]

$$d_{2jm} = \sqrt{x_{2j2m}} - \sqrt{x_{2j1}}$$

d_{2jo}	d_{2j1}	d_{2j2}	d_{2j3}	$d_{2j.}$
-0.20	0.00	-0.27	-0.32	-0.59
0.00	0.00	-0.27	0.00	-0.27
0.18	0.00	0.00	-0.27	-0.27
0.18	-0.50	-0.27	-0.27	-1.04
-0.18	-0.24	-0.24	-0.50	-0.98
0.20	-0.45	-0.24	-0.27	-0.96
0.18	0.32	0.27	0.00	0.59
0.38	0.24	-0.24	0.00	0.00
$d_{2.m}$ 0.74	-0.63	-1.26	-1.63	

$d_{2..}$: -3.52

(1) m= 0 denotes baseline.

Table 3
Analysis of Variance Table

Source	df	MS	F	p
Treat.	1	0.85	11.52	0.004
Periods	1	0.04	0.55	
Error(a)	14	0.07		
Time*Treat.	2	0.01	0.47	0.63
Time*Per.	2	0.01	0.61	
Error(b)	28	0.02		

Treat.: treatments. Per.: period. df: degrees of freedom. MS: Mean square. F: Snedecor's F. p: p-value. Letters a and b denote errors, as described in equation 1.
__Note__: Mauchly's sphericity test and Box variance-covariance homogeneity test were both nonsignificant (p= 0.31 and 0.78, respectively). The Shapiro-Wilk test for nonnormality of residuals was nonsignificant (p=0.87).

Figure 1
Average number of active sites as a function of time (in months)
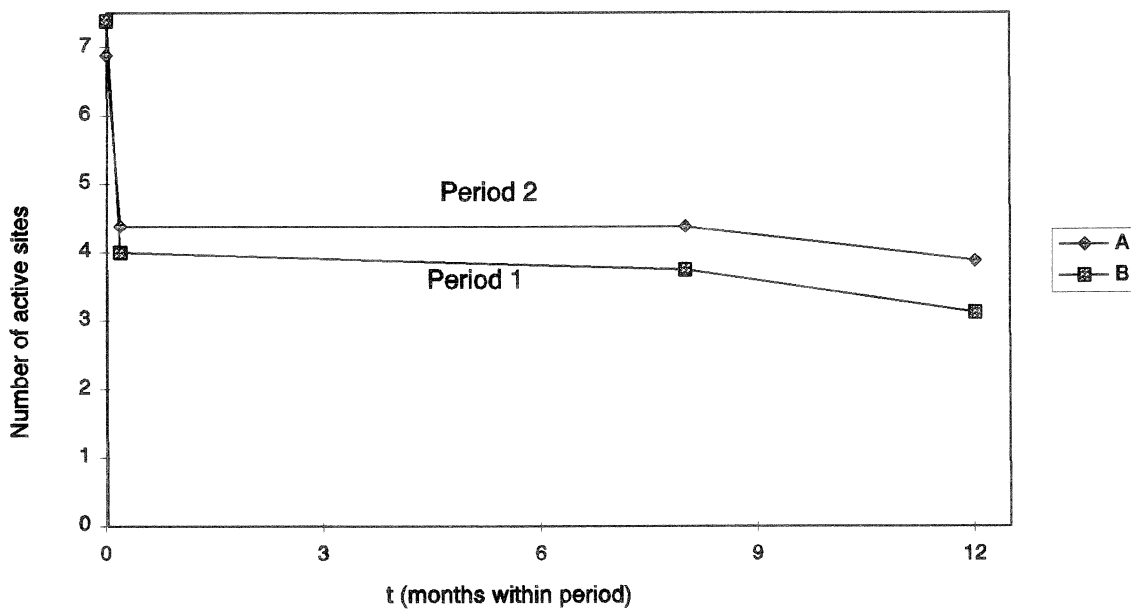## Untransformed data

Group 1

Group 2

Figure 2
Average difference (d) in number of active sites
Treatment A minus Treatment B (within sequence)
Transformed data

t (months after respective spraying)