Analysis and Interpretation of Ecological Field Data Using BACI Designs: Discussion

Loveday L. CONQUEST

McDonald, Erickson, and McDonald (2000) and Murtaugh (2000) have presented two interesting views of analysis and interpretation of data from before-after control-impact (BACI) studies. In my opinion, the general message they convey to the reader is that analysis of data from a BACI design, in combination with professional expertise, can add useful evidence as long as one is prudent about the resulting interpretations. Whether scientists who analyze BACI data will indeed use caution in interpretation, particularly in making statements regarding cause and effect, remains to be seen.

McDonald et al.'s (2000) study is application specific, taking the reader through an analysis of bird count data from the *Exxon Valdez* oil spill (EVOS). The authors prudently phrase their interpretation in terms of a weight of evidence argument rather than the much more difficult one of causation. I see the primary usefulness of this paper as that of encouraging scientists to use more complex models to realistically describe count data through time. (Using these models in data analysis has been considerably enhanced by the increasing availability of various types of computer software.) Particularly with biological data, a conventional first step is to try an additive model on untransformed data or an additive model on log-transformed data with the hope that the log transform restores the assumptions of variance homogeneity and normally distributed error terms. The presence of data recorded on the same units through time necessitates that the Huynh-Feldt condition (also known as sphericity) holds in order to use a conventional repeated measures approach. In case this sphericity condition does not hold, there are references to make conservative adjustments to the usual F-tests. (I do wonder just how many people would take the trouble to test these assumptions and do the proper corrections if such corrections were not already built into the

Loveday L. Conquest, Fisheries/Quantitative Science, University of Washington, Campus Box 355020, Seattle, WA 98195-5020 (E-mail: conquest@cqs.washington.edu).

Journal of Agricultural, Biological, and Environmental Statistics, Volume 5, Number 3, Pages 293–296

software program.) The authors also point out the effect that the constant c in the $\log(x + c)$ transformation has on the results—sometimes very little, sometimes a lot. (This leads me to believe that, if one is going to use the $\log(x + c)$ transformation to test ANOVA-type null hypotheses, it should be done for several values of c to assess the effects on the final conclusions.) The use of the generalized mixed model approach gives much cleaner results in that it meets certain assumptions and provides for a natural interpretation of parameter estimates. The authors end with the appropriate caution that statistical inference for this particular BACI design must be tempered with professional judgment to decide whether the disturbance caused the observed differences. Nonetheless, I still wonder how many scientists would be so cautious in their interpretations.

Murtagh's (2000) paper considers general issues with the BACI design. One of the issues is that of the presence of serial correlation, which tends to depress p-values and may therefore lead to Type I errors. Again, with the availability of software programs to handle increasingly complex models, if serial correlation is suspected, then it should be included in the model, even if some terms are difficult to estimate. Even an overly simple correlation model may be better than the conventional assumption of complete independence: This is in keeping with a primary rule of modeling, "all models are wrong, but some models are useful" (Anderson and Woessner 1992). Murtagh's time series analysis of the $(C - I)$ differences allows for both serial correlation and for a changing mean response through time.

Murtagh (2000) raises a further question: Even if there is a difference in two trajectories, how do we know that this has any relationship with the intervention? The only way to tell is to look at more trajectories. Without true replication, there really is no statistical inference. The use of multiple control units is suggested or the monitoring of multiple sites both near and far from the location of the intervention. Both of these are strategies to obtain some information on the natural variability of similar trajectories.

The lack of replicate trajectories can be very problematic; this is illustrated in Figures 1 and 2. Figure 1 shows a response through time of the single control unit and the single impact unit under a BACI design. The two curves are reasonably parallel up until the time of the disturbance. It appears that, following the disturbance, the impact curve decreases through time while the control curve increases. Figure 2 also displays the time trajectory of the impact and control units, but now the time trajectories of other, replicate units have been added. Here it appears that both the control and impact units are simply displaying the natural random variation that occurs in the time trajectories of these replicate units and that there is no longer anything special about the control or impact label. This is one of Murtagh's (2000) points: Even if serial correlation is properly modeled, comparing a single impact trajectory to a single control trajectory (and discovering that they are statistically different) may simply uncover the fact that replicate trajectories can display different behavior through time that has nothing whatsoever to do with the disturbance. The only proper way to deal with this is to have replicate control and impact trajectories. If this is impossible, then final interpretation of statistical results must be tempered with a great deal of caution and

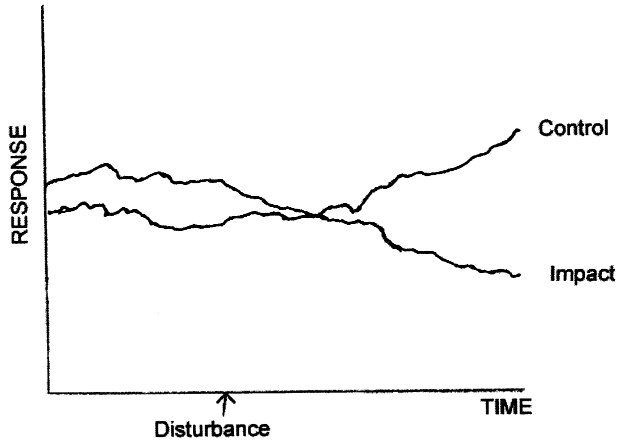

Figure 1. Response Time Traces for a Control Unit and an Impact Unit in a BACI Design.

professional expertise. Underwood (1994) discusses consequences of using unreplicated versus replicated BACI designs in the context of spatial and temporal monitoring.

What about case studies—are they simply an n of one and therefore of little or no use? On the contrary, case studies can reveal information about the physical, biological, and ecological processes in a system. For example, many useful results have come out of the Hubbard Brook Experimental Forest in New Hampshire (Likens and Bormann 1995). Another example is a before and after case study to assess the effects of forestry practices in a coastal stream ecosystem in British Columbia over 17 years (Hartmann and Scrivener 1990). A case study like this added to the general understanding of effects of silvicultural activities on features of stream condition and how variability in physical conditions and fish populations is induced by logging. In these two examples, the goal was to add to the base of process knowledge rather than to engage in specific hypothesis testing (Conquest 1998).

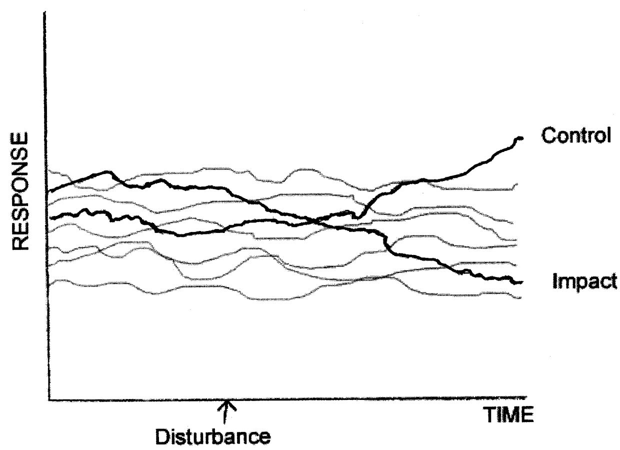

Figure 2. Response Time Traces for Several Replicate Units in Addition to the Ones Labeled Control and Impact.

Despite their associated problems regarding proper data analysis and interpretation, BACI-based designs are here to stay. In the absence of replicate trajectories, one needs to (1) stare hard at good graphs of the data, (2) use process knowledge in one's argument, and (3) avoid having to use a p-value. The phenomenon where data submitted for journal publication must always be accompanied by a p-value from a statistical test (one for which statisticians are partly responsible by the way statistical methods courses are taught, particularly to nonmajors) has been referred to by Stephen Hurlbert as "manic quantitation" (Hurlbert 1984); this is something to be avoided. In making sense of data recorded through time, we need to ensure that replication occurs whenever possible, that proper statistical models are used in the analysis, that hypothesis tests are done only when they are truly warranted, and that results are appropriately interpreted.

[Received March 2000.]

REFERENCES

Anderson, M. P., and Woessner, W. W. (1992), *Applied Groundwater Modeling: Simulation of Flow and Advective Transport,* San Diego, CA: Academic Press.

Conquest, L. L. (1998), "Statistical Design and Analysis: Considerations for Monitoring and Assessment," in *River Ecology and Management: Lessons From the Pacific Coastal Ecoregion,* eds. R. J. Naiman and R. E. Bilby, New York: Springer-Verlag, 455–475.

Hartman, G. G., and Scrivener, J. C. (1990), "Impacts of Forestry Practices on a Coastal Stream Ecosystem, Carnation Creek, British Columbia," *Canadian Bulletin of Fisheries and Aquatic Sciences,* 223, 148.

Hurlbert, S. H. (1984), "Pseudoreplication and the Design of Ecological Field Experiments," *Ecological Monographs,* 54, 187–211.

Likens, G. E., and Bormann, F. H. (1995), *Biogeochemistry of a Forested Ecosystem* (2nd ed.), New York: Springer-Verlag.

McDonald, T. L., Erickson, W. P., and McDonald, L. L. (2000), "Analysis of Count Data From Before-After Control-Impact Studies," *Journal of Agricultural, Biological, and Environmental Statistics,* 5, 262–279.

Murtagh, P. (2000), "Paired Interventaion Analysis in Ecology," *Journal of Agricultural, Biological, and Environmental Statistics,* 5, 280–292.

Underwood, A. J. (1994), "Spatial and Temporal Problems With Monitoring," in *The Rivers Handbook, Hydrological and Ecological Principles* (Vol. 2), eds. P. Calow and G. E. Pitts, London, Oxford: Blackwell Scientific, 101–123.