*Ecological Monographs*, 61(1), 1991, pp. 53–73


# DESIGNING ENVIRONMENTAL FIELD STUDIES[1]

L. L. EBERHARDT AND J. M. THOMAS
*Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352 USA*

*Abstract.* Field experiments in ecological and environmental research usually do not meet the criteria for modern experimental design. Subsampling is often mistakenly substituted for true replication, and sample sizes are too small for adequate power in tests of significance. In many cases, field-study objectives may be better served by various kinds of sampling procedures, even though the resulting inferences will be weaker than those obtainable through controlled experimentation.

The present paper provides a classification and description of methods for designing environmental studies, with emphasis on techniques as yet little used in ecology. Eight categories of techniques for field studies are defined in terms of the nature of control exerted by the observer, by the presence or absence of a perturbation, and by the domain of study. The first two categories include classical experimental approaches, replicated and unreplicated. Sampling for modelling provides efficient designs for estimating parameters in a specified model. Intervention analysis measures the effect of a known perturbation in a time series. Observational studies contrast selected groups from a population, while analytical sampling provides comparisons over the entire population. Descriptive survey sampling estimates means or totals over an entire population, while sampling for pattern deals with spatial patterns over a selected region.

We propose that the statistical concept of a "superpopulation" may be useful in ecology, and that it may be desirable to approach community and ecosystem studies in a sampling framework, with experimentation used for a fairly narrow range of subsidiary investigations. Much more attention to processes for drawing inferences about cause and effect is needed, in any case.

*Key words: analytical sampling; experimental design; field experiments; inferences; observational studies; pattern; populations; pseudoreplication; sampling design; sampling for modelling; Type I and II errors.*

## INTRODUCTION

Much of what we know as the scientific method is based on the idea of experimental investigation of an hypothesis. Beginning in the 1920s, statistical methodology was developed for the design and analysis of such experiments, so that the bulk of current research reports in many fields involve statistical analyses of the data. However, a strictly experimental approach to field studies is both difficult and expensive to achieve in the environmental sciences, particularly if worthwhile statistical analyses are to be incorporated. In an especially compelling demonstration, various court decisions forced detailed field studies under the National Environmental Policy Act. Many "impact studies" were subsequently conducted. Most of these confused subsampling and treatment replication (Eberhardt 1976).

The basic problem in impact studies is that evaluation of the environmental impact of a single installation of, say, a nuclear power plant on a river, cannot very well be formulated in the context of the classical agricultural experimental design, since there is only one "treatment"—the particular power-generating station. In a wider context, good experimental control may make it impossible to study the essential phenomena in an ecological system. If the system is more than the sum of its parts, ultimate understanding requires observation of the intact, functioning whole. Many fields of scientific endeavor have the same sorts of difficulties, including astronomy, economics, medicine, and sociology.

Recently Hurlbert (1984) reviewed a sizable number of ecological studies, and again pointed out that subsampling is often mistakenly assumed to constitute replication. He also provided valuable details of the frequency of experiments without replication in ecological field research. Hurlbert classed all such studies as being either "manipulative experiments" or "mensurative experiments." The manipulative class included situations in which the investigator controls circumstances of the study. Most scientists regard these as "experiments." In such investigations, the scientist assigns different treatments to experimental units in a specified manner, usually including untreated "controls."

The second class involved only passive observation of some process not under the investigator's control. As Hurlbert noted, this category is essentially concerned with sampling. His term "mensurative exper-

[1] Manuscript received 6 March 1989; revised 5 December 1989; accepted 14 May 1990.

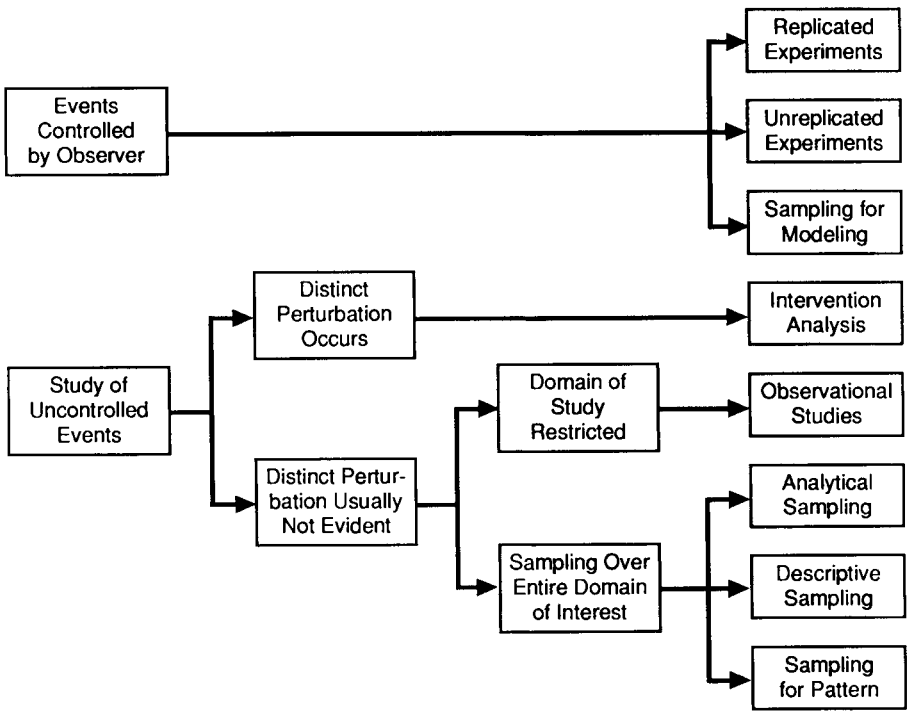


FIG. 1. A classification of the methods considered in this study.

(descriptive) sampling, and is mainly concerned with estimation of means or totals. A dozen or more texts are available, with the most widely used reference in environmental work being that of W. G. Cochran (1977: 4). He distinguished two methods:

> *Sample surveys can be classified broadly into two types—descriptive and analytical. In a descriptive survey the objective is simply to obtain certain information about large groups: for example, the numbers of men, women, and children who view a television program. In an analytical survey, comparisons are made among different subgroups of the population, in order to discover whether differences exist among them and to form or verify hypotheses about the reasons for these differences.*

Observational studies may be distinguished from analytical sampling mainly by the deliberate selection of contrasting portions of populations for study. The name comes from another book by Cochran (1983), who there suggested that analytical surveys have broader and more exploratory objectives. "Sampling for pattern" (Eberhardt 1978*a*) has emerged from the practical need to assess the extent and value of a body of ore, or the volume and extent of an oil field. Research on methods and results tends to be published in the geological literature, so that "geostatistics" may be used in titles of books and papers. In many applications, sampling locations may be uncontrolled by the observer (e.g., drilling of oil wells), and the methodology may be most useful as a way to reduce biases resulting from haphazard sampling.

The balance of this paper will be concerned with further details of the eight methods of Fig. 1 as they relate to existing and potential applications in environmental studies. A synopsis of the methods is:

1) Experiments with replication—strong inferences; preferred approach when feasible;

2) Experiments without replication—cost or circumstances prohibit replication;

3) Sampling for modelling—efficient experimentation for parameter estimation in specified non-linear models;

4) Intervention analysis—retrospective assessments of time-series data;

5) Observational studies—deliberate selection of contrasting groups in lieu of experimentation;

6) Analytical sampling—inferences from sampling over entire population of interest;

7) Descriptive sampling—efficient estimation of means and totals;

8) Sampling for pattern—description of spatial pattern; interpolation to reduce bias from haphazard sampling.

## REPLICATED EXPERIMENTS

Strong inferences depend on controlled experiments. Confirmation that two experimental outcomes are indeed different depends on randomization and replication to provide a measure of variability in units treated alike. Since these topics and many related aspects

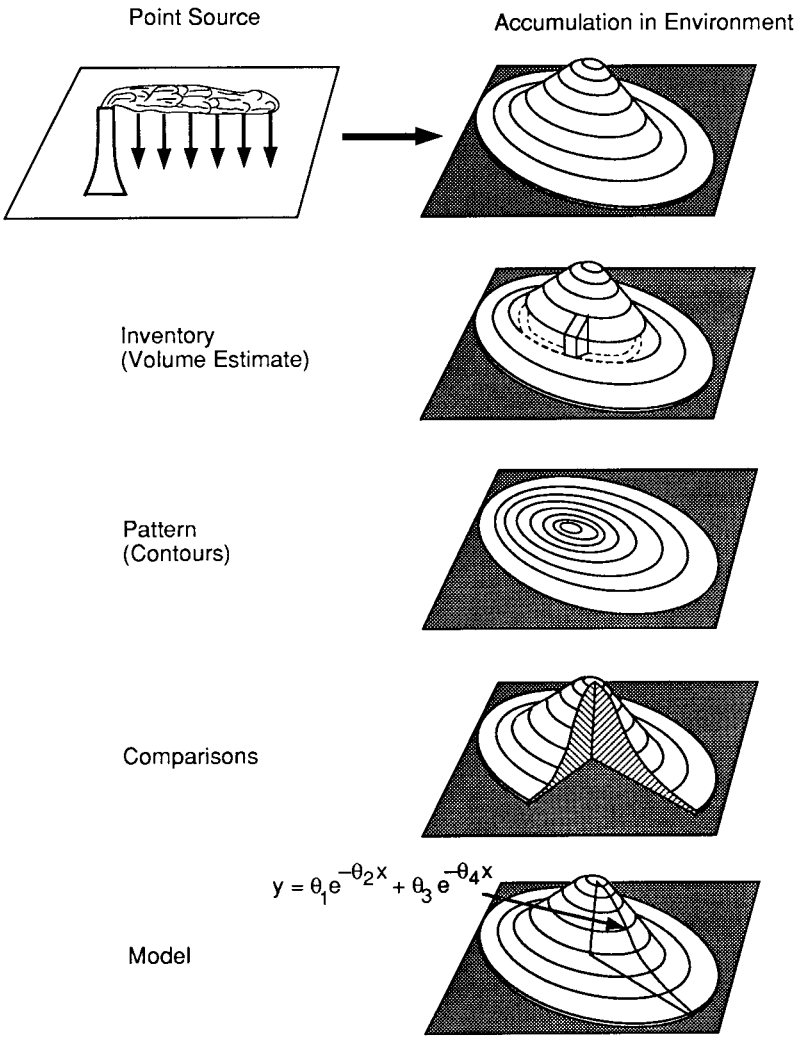


FIG. 2. An illustration of the four categories of sampling. The diagrams at the right are hypothetical concentrations in soil around a point source, here suggested as a smokestack. In practice, it is likely that the elliptical concentration pattern might be more strongly skewed in the direction of the prevailing winds.

pling can be described as an "inventory" method, sampling for pattern produces a contour map, while analytical sampling provides a basis for comparisons. Sampling for modelling stands somewhat apart due to the focus on parameter estimation for a specific model.

ACKNOWLEDGMENTS

This research was supported by the Environmental Research Division of the Office of Health and Environmental Research of the United States Department of Energy under contract DE-AC06-76-RLO1830. Helpful comments from S. H. Hurlbert, D. Meeter, and three anonymous referees are gratefully acknowledged.

LITERATURE CITED

Aitken, M. 1981. Regression models for repeated measurements. Biometrics **37**:831–832.

Atkinson, A. C. 1969. A test for discriminating between models. Biometrika **56**:337–347.

———. 1972. Planning experiments to detect inadequate regression models. Biometrika **59**:275–293.

Atkinson, A. C., and W. G. Hunter. 1968. The design of experiments for parameter estimation, Technometrics **10**: 271–289.

Bard, Y. 1974. Nonlinear parameter estimation. Academic Press, New York, New York, USA.

Bartlett, M. S. 1975. The statistical analysis of spatial pattern. Chapman and Hall, London, England.

Beck, J. V., and K. J. Arnold. 1977. Parameter estimation. John Wiley & Sons, New York, New York, USA.

Blau, G. E. 1975. Mathematical model building with an application to determine the distribution of Dursban insecticide added to a simulated ecosystem. Advances in Ecological Research **9**:133–163.

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. Journal of the Royal Statistical Society **B26**: 211–243.

Box, G. E. P., and N. R. Draper. 1987. Empirical model-building and response surfaces. John Wiley & Sons, New York, New York, USA.