PNL-4138
1-

NUREG/CR-2803
PNL-4138

Improved Field Experimental Designs and Quantitative Evaluation of Aquatic Ecosystems

Prepared by D. H. McKenzie, J. M. Thomas

Pacific Northwest Laboratory
Operated by
Battelle Memorial Institute

Prepared for
U.S. Nuclear Regulatory
Commission

REFERENCE COPY

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability of responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 1717 H Street, N.W.
Washington, DC 20555
2. The NRC/GPO Sales Program, U.S. Nuclear Regulatory Commission,
Washington, DC 20555
3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the NRC/GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations*, and *Nuclear Regulatory Commission Issuances*.

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Division of Technical Information and Document Control, U.S. Nuclear Regulatory Commission, Washington, DC 20555.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

Improved Field Experimental Designs and Quantitative Evaluation of Aquatic Ecosystems

Manuscript Completed: April 1984
Date Published: May 1984

Prepared by
D. H. McKenzie, J. M. Thomas

Pacific Northwest Laboratory
Richland, WA 99352

Prepared for
Division of Health, Siting and Waste Management
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN B2072

ABSTRACT

We used the paired-station concept and a log transformed analysis of variance as methods to evaluate zooplankton density data collected during five years at an electrical generation station on Lake Michigan. To discuss the example and the field design necessary for a valid statistical analysis, we provide considerable background on the questions of selecting 1) sampling station pairs, 2) experimentwise error rates for multi-species analyses, 3) levels of Type I and II error rates, 4) procedures for conducting the field monitoring program, and 5) a discussion of the consequences of violating statistical assumptions. We include details for estimating sample sizes necessary to detect changes of a specified magnitude.

Both statistical and biological problems with monitoring programs (as now conducted) are addressed; serial correlation of successive observations in the time series obtained was identified as one principal statistical difficulty. Our procedure reduces this problem to a level where statistical methods can be used confidently.

CONTENTS

ABSTRACT	v
INTRODUCTION	1
MONITORING PROGRAM CONSIDERATIONS	4
GENERAL DESIGN AND MONITORING CONSIDERATIONS	4
STATISTICAL CONSIDERATIONS	5
Error Rates and Sample Sizes	9
APPLICATION: ZION NUCLEAR POWER PLANT ZOOPLANKTON MONITORING . .	13
SUMMARY	20
REFERENCES	21

FIGURES

1	Relationship Between Sample Size and Detectable Difference in Preoperational vs Operational Means Using a Multiplicative Log Transformed Model Based on Paired Stations	12
2	Field Sampling Locations for Plankton Near Zion Nuclear Power Station	14
3	Total Zooplankton Comparison for Control Station 5 and Treatment Station 28, Zion Nuclear Power Plant	15
4	Station Mean for Total Zooplankton Densities at the North Zion Stations 5 and 28	16

TABLES

1	Analysis of Variance for the Log Difference Between Total Zooplankton Densities at the Zion Control and Treatment Stations	17
2	Summary of Estimated Variances, paired and Independent Station Densities for the Components of the Zooplankton Communities Sampled at Zion	19

INTRODUCTION

Nearly all ecological studies conducted today include some consideration of statistical design principles or analytical procedures, although many statisticians would probably place the majority of studies in the category of "too little, too late." However, a meaningful application of statistical principles and techniques to the investigation of ecological phenomenon is not simple, even though applied statistical textbooks may seem to portray usual experimental designs as adequate.

The objective of this paper is to discuss some statistical approaches suggested by earlier research (Eberhardt 1976, 1978, Gore et al. 1978, Thomas et al. 1978) and to propose a design for ecological studies which we suggest can detect changes in some aquatic population densities. We also stress an approach that emphasizes initial planning that will result in improved impact monitoring programs with a good chance of detecting environmental changes, propose reasonable values for statistical errors, and an affordable level of effort. The proposed design is illustrated using zooplankton data from an aquatic impact monitoring program conducted at the Zion Nuclear Power Station on Lake Michigan.

Despite an abundance of good statistical texts that provide the statistical machinery for testing hypotheses about data from designed experiments, applications in field aquatic biology are generally a reflection of the statistician's and field biologist's expertise and experience. This situation arises from two contributing factors. First, each ecological field survey is sufficiently unique that the well developed methods and examples from the fields of psychometrics, agriculture, or medical statistics are never quite appropriate. Thus, individual applications involve considerable interpretation and value judgment regarding the selection and application of appropriate statistical design and analysis machinery. The second contributing factor is the biologist's ability to translate his knowledge of the ecosystem into testable hypotheses. Quantitative ecology has not advanced to a state where statistical plans are available. Some notable exceptions include the books of Seber (1973) and Ricker (1975) dealing mostly with the estimation of population size. Most field data sets are collected under constraints and circumstances that either violate or, at least, stretch requirements and assumptions underlying most classical statistical procedures. Usually it is difficult to evaluate the underlying statistical assumptions and they even may go unquestioned (Fickeisen et al. 1975 and Pirie and Hubert 1977). Thus, the success of current aquatic research efforts in collecting data to evaluate a hypothesis appears to be highly dependent upon the resourcefulness, experience and subjective decisions of the individual researcher. However, good scientific procedures dictate that a more objective decision-making process is usually required.

Quantitative problems are frequently encountered when a researcher attempts to quantify a population or ecosystem response to a perceived or actual change in the environment. Examples include the quantitative evaluation of population changes due to pollution levels, harvesting policies, habitat

manipulations and water diversions. Field surveys designed to investigate such situations are generally characterized by lack of suitable controls and utilize a scheme whereby before-after treatment comparisons must be made using time series data. Generally no replication or random assignment of experimental units is used. Even though the remainder of this paper focuses on the field sampling and statistical problems of quantifying the impact of thermal electric power plants on zooplankton densities, the ideas are also relevant to the preceding situations.

Implementation of the National Environmental Policy Act (NEPA) in the United States has resulted in large expenditures to collect huge amounts of environmental data. The purpose of such collections is to evaluate or predict impact for new facilities, but instead has generated large volumes of difficult to locate and, we believe, unread literature i.e., Environmental Impact Statements (EIS). Our perusal of this literature shows that the level of applied ecological knowledge and field sampling technology and/or applied statistical expertise has been minimal. Most cases are characterized by a failure to apply the available methods and state-of-the-art methodology has been used only at a few sites. The major problem in these studies has been to separate changes induced by the construction and operation of the facility from natural and other man-induced variation and changes in population densities. Since the quantitative evaluation of the impacts of man's activities is a laudable and essential goal needed for rational management of our natural resources, the best available techniques should be used.

The primary objective of an environmental monitoring program is to measure the environmental changes induced by the operation of a power plant or other facility. In the case of power plants, potential impacts could result from thermal and chemical effluents as well as mechanical stresses, impingement and entrainment, all associated with the passage of water through the facility. Whether a monitoring program can achieve such a goal is dependent on many things, the most important of which is the initial design or planning stage. Usually, an effort to characterize and describe the aquatic ecosystem according to a traditional experimental design and scientific research framework is attempted. Another approach is to base the monitoring design on a previously conducted program, resulting in a repeat of unacceptable studies. Even though ecological change has been correlated with power plant operation (not necessarily resulting in a conclusion of cause) using either of these approaches, we believe circumstances have been largely fortuitous and the change probably could have been observed with considerably less expenditure of effort. We acknowledge that there is no single, best or even superior sampling design or statistical analysis technique for power plant impact assessment. Numerous design and analysis schemes probably could be invoked which would represent current state-of-the-art. We hope our approach will help aquatic ecologists avoid the more serious statistical faults of the past studies and also will aid them in detecting aquatic community changes resulting from the operation of power plants.

Many of our recommendations would seem intuitively obvious and simplistic both to statisticians and biologists. However, the results of recent studies indicate that many nuclear power plant monitoring programs have failed to consider one or more basic principles (Adams et al. 1977, Gore et al. 1977, Murarka et al. 1976b).

MONITORING PROGRAM CONSIDERATIONS

GENERAL DESIGN AND MONITORING CONSIDERATIONS

A nearly universal characteristic of monitoring programs for impact assessment is the designation of both impacted/treatment and nonimpacted/control areas or stations. However, establishment of such stations often does not appear to be based on experimental design criteria. In general, these stations are not selected in a random manner. In contrast, classical experimental designs used extensively for industrial and scientific applications require random assignment of experimental materials to treatment and control areas, plot and animals. Treatment and control groups are fundamental to the classic approach, but cannot be attained in the impact setting since the power plant affects only one area, and one or more suitable control areas may not be available. At best, randomization can only be applied to the selection of subsamples (stations) from within an area, so classical statistical methodology is not usually appropriate. However, we believe that a preferred approach can be developed which considers the ecological realities when sampling at treated and nontreated stations is necessary (i.e. control stations can be used to estimate density changes that result from normal environmental changes that are not associated with the power plant operation). Our recommended design and analysis scheme includes the notion of paired control-treatment stations (Eberhardt 1976).

The proper coupling or pairing of potentially impacted/treatment stations with nonimpacted/control stations should be a first consideration. We assume population densities at both locations will respond in a similar way to changes in environmental variables (i.e. amount of insolation, precipitation, winds, etc.). If this assumption is invalid, it cannot be concluded that changes at the treatment station result from stresses introduced by the operation of the power plant. Note that identical station pairs are not required. Identical stations are usually nonexistent due to the multitude of environmental variables and their differential effects on habitats of varying characteristics. In fact, environmental heterogeneity makes each station ecologically unique. Therefore a qualified biologist must use survey data, local knowledge, or personal observations to select station pairs likely to respond in similar ways to most expected environmental changes.

In addition, a biologist must balance two factors when selecting control sites: 1) they must be far enough away from the power plant so that there is only a small probability that populations will be impacted by the plant operations and 2) they must be close enough so distance between control and treatment sites will not effect their similarity. Environmental gradients (e.g., salinity gradients often found in estuaries) accentuate the problem. The final selection of sites and conclusions about "far enough away" must be qualitative.

After establishing sampling stations, some simple rules must be followed (Gore et al. 1978): when practical, sampling at control and treatment stations should occur simultaneously, and relocation or moving stations and changes in

collection techniques should be avoided since this greatly reduces the possibility that a resulting data base can be interpreted. It is usually impossible to disentangle changes caused by a power plant from those resulting from the changes in the sampling program. In general, consistent data bases, even though not based on an optimal number of stations or proper gear choices, are superior to inconsistent data bases for analytical evaluation.

Where possible, balanced data sets should be obtained. This rule includes the admonition to avoid missing data. In addition, equal numbers of samples should be collected within each level of the experimental factors (time) or dimensions (habitat). Habitat dimensions are commonly used to characterize or describe strata for the station pairs, i.e. depth of water, substrate type, current velocities or patterns, and are commonly incorporated into analysis of variance models. Thus, when designing the field sampling program, equal numbers of samples should be allocated to each level of such factors. For example, equal numbers of samples should be collected from each time period (monthly, bimonthly, quarterly). Information gained in the preoperational sampling may provide a logical basis for reducing the number of stations, samples, time periods, in the operational phase; however, this should not be allowed to result in unbalanced data sets. We suggest that preoperational data resulting from the deleted efforts should not be included in the analysis of the impact hypothesis. While the site specific characteristics may prevent complete balancing, the greater its degree, the easier the task of data analysis and interpretation.

Another factor to consider during the design of the monitoring program is the expected frequency of zero organism counts in the samples. These can be avoided by increasing the volume sampled and/or decreasing the number of taxonomic categories. The rationale for this recommendation is primarily statistical, although the biological interpretation can also be complicated by large numbers of samples with zero counts. One statistical assumption generally made when analyzing data is the assumption of normality or that the data can be transformed into the normal distribution. Substantial numbers of zero counts invalidate the normality assumption and alternative techniques, often more complicated and/or less powerful (in a statistical sense), must be utilized. If the true density of organisms is low and zero counts occur frequently, a low probability of detecting a plant impact that reduced organismal density exists.

STATISTICAL CONSIDERATIONS

When constructing testable hypotheses about impact, several statistical factors should be considered. These include the underlying probability model, acceptable error rates, definition of an experimental unit and the time series aspects of the data. Recent reviews of several case histories on assessment of power plant impact indicate that little, if any, attention is given to these factors (Gore et al. 1977, Murarka et al. 1976b, Adams et al. 1977). Our proposed analysis method and other state-of-the-art procedures rely heavily on analysis of variance (ANOVA) methodology (McCaughran 1977). However, to properly use ANOVA methods for hypothesis testing, the above factors must be considered.

In the following paragraphs we address four issues concerned with analysis of variance methodology: 1) the probability density function thought to describe the distribution of the organisms; 2) whether fixed, random or mixed variance components should be included in the analysis of variance model; 3) the sampling level that can be considered to provide replicate observations; and 4) the time series characteristic of the data, i.e., lack of statistical independence or the serial correlation of organism counts in successive sampling periods. Unfortunately, resolution of these issues is rendered impossible because no consensus on proper procedures exists (McKenzie et al. 1977). Thus, the following discussion is based upon a combination of qualitative information and the current literature.

Three probability distributions, normal, Poisson and lognormal, are commonly used to describe the occurrence of organisms; numerous authors have discussed their properties and applicability (Pielou 1969, Seal 1964, Eberhardt 1976). The common observation that the sample mean and variance are not independent seems to indicate that either the Poisson or lognormal model is appropriate for densities of aquatic organisms. However, McCaughan (1977) invokes the central limit theorem and the robustness of ANOVA and proposes the normal distribution whenever the sampling effort can be arranged to obtain "large" data values. The Poisson distribution is generally appropriate when the data are in the form of counts of rare events (Snedecor and Cochran 1967). This is not a very realistic assumption for aquatic organisms since they usually exhibit clumped or patchy distributions. For samples of aquatic organisms, very often the variance increases faster than the mean, exceeding the constant proportional relationship exhibited by data which are described by the Poisson. The nearly constant coefficient of variation often observed in ecological data implies that the lognormal distribution may be more appropriate (Eberhardt 1978) and that the treatment effects are proportional rather than additive (Snedecor and Cochran 1967). Power plant impacts are potentially of this form, i.e., a percentage reduction in the annual and seasonal densities of aquatic organisms. Therefore, we have assumed the lognormal model is appropriate and a logarithmic transformation is used prior to the application of ANOVA methodology in this paper.

Our use of a logarithmic transformation implies that the underlying analysis of variance model includes the factors and the error term in a multiplicative form. Thus, the model is additive after transformation but the change of scale may create or destroy significant interaction effects (Tukey 1949). Since it is nearly impossible to ascertain whether a significant interaction is real or an artifact of the choice of scale (Winer 1971), tests of hypotheses about such interactions should be avoided. Use of a lognormal model results in additional difficulties because the usually understood statistics (i.e. mean, standard deviation) used to present the results are in logarithmic units. However, similar problems result with other common transformations: square root and arcsine square root. The temptation to "untransform" the statistical results should be resisted. The antilog of the mean represents the geometric mean, but the antilog of confidence limits and variance do not have the familiar properties and interpretation associated with these estimates when they are based on the normal distribution (Steele and Torrie 1960, Autchison and Brown 1969).

A multiplicative model, typically applicable to power plant impact analysis, can be written as:

$$Y_{ijk} = \mu A_i B_j AB_{ij} \epsilon_{ijk}$$

where:

Y_{ijk} = observed organism density

μ = mean effect

A_i = effect of factor A level i

B_j = effect of factor B level j

AB_{ij} = interaction effect of factors A and B

ϵ_{ijk} = multiplicative error (lognormally distributed)

k = 1 to n replicates or repeated measurements

Typically the main factors (A,B) represent the effects of station location, sampling months, depth of sampling, etc. One decision is whether to include these effects in the analysis as fixed or random factors, in order to develop valid tests of hypotheses. Random effects are levels of factors selected randomly from some larger population of possible levels of factors. When a random factor hypothesis is tested, the conclusion is applicable to the entire population of factors. For example, in a field experiment to evaluate the effect of various pollutants on fish growth, the factor may be the ponds or lakes in which the fish are located. The number of different ponds in an area may be very large. If the ponds are randomly selected for use in the experiment, then the pond factor is a random factor. The results of testing the significance of the pond effect would be applicable to all ponds in the area. Alternatively, for a fixed factor it is assumed that there is no large population of levels for a factor. Therefore, levels are not randomly chosen and any hypothesis test relates only to the experimental conditions, i.e., all potential factors are included in the experiment. When selection of experimental factors from a larger population of potential factors proceeds in a systematic, nonrandom manner, then the factor should also be considered fixed (Winer 1971). The latter situation is more applicable to impact assessment, so we generally regard all factors in the above model as fixed when constructing hypothesis tests or interpreting results.

An important consideration when applying ANOVA methods is the definition of what constitutes a replicate or repeated observation (Y_{ijk}) on the experimental material. The statistical requirement that must be met is independence of the error term (ϵ_{ijk}), i.e., ϵ_{ijk+1} is not influenced by the value of ϵ_{ijk} . Translating this into ecological terms requires careful delineation of the experimental material or units. Within the context of impact assessment, it appears that experimental units are populations of aquatic organisms at a sampling station. This implies that each sampling occasion at a station produces one replicate and that repeated collections on one occasion, multiple deployments of the sampling gear, are subsamples and not replicates. The error term associated with such subsamples represents the measurement error and not the error associated with applying treatments to experimental units. This line of reasoning leads to an ANOVA with one observation per cell ($k = 1$) and an inability to obtain an estimate of the residual error variance ($\hat{\sigma}^2$). Consequently, hypothesis testing must then be based on the interaction error

variance ($\sigma^2 + n\sigma_{AB}$), which reduces the sensitivity of the tests. The lack of replicates also prevents application of further associated statistical analysis techniques, i.e., orthogonal contrasts.

We are proposing the use of observations collected on successive years to gain replication. Observations separated by one year logically appear to represent independent responses by experimental units to applied treatments (impact). However, aquatic populations (experimental units) are likely to have different densities between years in response to changes in environmental variables. Hence the need for and utility of the proposed paired station approach to remove this effect.

Finally, the serial correlation or lack of independence between the observations collected over time must be addressed. Intuitively, the density of organisms present at a subsequent sampling time depends somewhat on the density present during previous sampling times. This suggests that intervention analysis of time series models (Box and Tiao 1975) is an appropriate approach. However, data describing many annual cycles of the population are necessary to use these techniques. Data set length, as well as frequency of observation, severely limit the applicability of time series analysis for monitoring data. Box and Jenkins (1970) suggest that sampling intervals should be fairly short compared with the time constants expected for the system. The time constants are a measure of the speed with which the system approaches a new state following a shift in the input conditions. Since many aquatic organisms can respond fairly rapidly to certain changes, (i.e. zooplankton, phytoplankton), a sampling interval of daily or every few days appears needed. With the possible exception of impingement and physical-chemical monitoring, such sampling frequencies are not economically or biologically feasible.

To apply ANOVA techniques to monitoring data, a method of circumventing the problem of serial correlations must be found. Scheffe (1959) states that departures from the assumption of independence can have a serious effect on inferences. He examined a simple case where the adjacent observations had a correlation coefficient of 0.50, and concluded that the actual significance level was 0.17, instead of the nominal 0.05 level. Murarka et al. (1976a) report serial correlation coefficients for zooplankton sampling as large as 0.60 for observations taken on successive months, 0.40 for observations taken two months apart, and -0.6 for observations taken six months apart. We have observed serial correlation coefficients as high as 0.9 for benthic samples collected at two week intervals. This evidence indicates that the significance level of inferences based on analysis of variance techniques applied to serially correlated environmental data is often incorrect.

Changing sampling frequency can affect serial correlation. Clearly, as the frequency of sampling increases, serial correlations increase. Conversely, as the time interval between samplings decreases, organism densities should be less dependent on the previous values. The rate of increase or decrease would appear to be a function of organism behavior, life span, turnover rates, seasonality and rates of change for driving variables. Total phytoplankton and zooplankton densities observed near Zion Nuclear Power plant showed that

samples collected three months apart yielded densities with low serial correlation coefficients, while samples collected six months apart resulted in the large negative coefficients mentioned above (Murarka et al. 1976a). We expect that samples of aquatic insects with one reproductive cycle per year and a uniform mortality rate would have a high serial correlation coefficient between all cohort samples. In general, when analysis of variance techniques are used, data analysis based on less than monthly sampling intervals seems risky.

The use of differences between organism densities at control and treatment stations appears to be another method of reducing serial correlation. This procedure is based on the assumption that the relationship between control and treatment station densities is relatively constant and that the difference represents the random fluctuations about that relationship. Empirical evidence from the difference in log transformed values for several data sets collected at monthly intervals fails to reject the hypothesis that serial correlation coefficients are zero. Thus, we are led to conclude that the serial correlation coefficients are sufficiently small and the ANOVA procedures can be applied. However, additional research should be conducted to establish unequivocally the effectiveness of using the paired differences to reduce serial correlation(s).

Another possible benefit which may occur when the paired station approach is used is a reduction in the error variance. The variance of the difference of two random variables can be written as:

$$V(X_1 - X_2) = V(X_1) + V(X_2) - 2 \text{ COV } (X_1 X_2).$$

Thus, whenever densities from a control-treatment pair occur in a similar pattern, the covariance of X_1 and X_2 will be positive. This, in turn, will reduce the error variance in the ANOVA and provide more powerful statistical tests. If random pairing of stations is used, and the densities have zero correlation, then the paired approach is only slightly less efficient than a random station selection. However, poor pairing (i.e., a negative correlation between the pairs) will result in substantial loss in ability to detect significant differences when they occur (Snedecor and Cochran 1967).

Error Rates and Sample Sizes

To estimate the number of samples needed for a monitoring program three important factors must be considered. First, the consequences of falsely concluding that an impact exists must be examined. Statisticians call such an event a Type I error and assign the Greek letter α to its probability. A Type II error results when the conclusion is made that an impact does not exist, when one actually is present. The Greek letter β is assigned to Type II error probability. Finally, a decision must be made about the magnitude of an impact or change that the monitoring program should detect. After deciding on values for α , β , and the magnitude of change, and given a preliminary estimate of the variance or coefficient of variation and a proposed experimental design, a statistician can answer the often asked question: How many samples should be collected?

Perhaps the most universal (and unsatisfactory) way sample size has been established in the past has been to start with a level of effort or budget range and calculate how many samples can be collected within those constraints. This approach relies heavily on approximating the expenditures of previous monitoring programs or "what the market (utility) will bear." If previous studies were well designed and conducted with consideration of the statistical principles above, and had similar objectives and ecosystems, then a good estimate of an adequate sample size might be attained. Such an occurrence is unlikely.

An α error rate of 0.05 for testing hypotheses is commonly used in research studies by scientists and engineers. Their experiments have been characterized by: 1) controlled experimental conditions, 2) easy measurement of consequences of a wrong decision, 3) uncomplicated significance and/or interpretation of results, 4) some good evidence for the form of the underlying distribution, 5) small coefficients of variation, and 6) easy and inexpensive sampling. Monitoring programs designed to test impact hypotheses usually have few of these characteristics. As a consequence, we have used a reduced error rate ($\alpha = 0.10$) for testing impact hypotheses. This decision reduces the β error rate when sample size is constant, and reduces overall sample size for a fixed β .

It is important to define whether a taxon or a trophic level assemblage (phytoplankton, zooplankton, benthos) is the focus of the study because this choice influences the error rates in the statistical analysis. One possible criterion to use in choosing the experimental unit is to select the minimum number of categories for which environmental action will be initiated when a significant change is detected in any one category. Regardless of the choice, the experimentwise error rate must be considered. If taxonomic groups are the experimental unit, then the error rates (α and β above) are appropriate for each hypothesis tested. However, if the experimental units are trophic level assemblages, then the error rates implicit in the analysis for each of the individual taxonomic groups within trophic level must be adjusted to control the overall experiment error rates. Thus, if impact hypotheses about more than two taxonomic groups are to be tested, the investigator must be concerned about the experimentwise error rates.

An example will illustrate the problem. If an α level of 0.10 is used (i.e. a 10% chance of making a Type I error and a 90% chance of not making a Type I error on each analysis of a single taxon), then the formula: $p = 1 - (1 - \alpha)^n$ where n = number of analyses, can be used to calculate the probability of making at least one Type I error for the n taxa groups. When $\alpha = 0.1$ and $n = 6$, then $p = 0.49$. Simply stated, there is nearly a 50% chance of concluding there is a significant impact on at least one taxa group when, in fact, no impact exists. A similar calculation can be made for the probability of making at least one Type II error. Clearly, when a trophic level is the experimental unit, error rates must be adjusted on individual analyses to control the experimentwise error rate. One possible way to control the experimentwise error rate in these circumstances is to use a significance level of α/m , where m is the number of taxa examined (Dunn 1961).

Another way to approach the Type I/Type II error problems involves the use of multivariate techniques whereby individual taxa groups are simultaneously treated as dependent variables in one analysis. Such techniques require the additional assumption that taxa groups are correlated in an invariate manner throughout the sampling period. However, the robustness (i.e., sensitivity to serial correlations) and power of these methods are not well documented. We recommend that a thorough investigation of this and all other assumptions be made prior to using these techniques for impact analysis.

A calculation of the β error rate has not been made in any power plant impact study we have reviewed (McKenzie et al. 1977). In fact, both α and β error rates have generally been ignored in the evaluation of impact hypotheses. McCaughan (1977) infers by example that a rate of 0.10 may be acceptable. Thomas (1977) presents sample sizes based on β values of 0.10, 0.20 and 0.30. It appears to us that the consequences of making either a Type I or II error are similar. A Type I error results in power customers paying for unneeded environmental controls at power plants. On the other hand, when a Type II error is made, the public incurs an added loss due to the undetected environmental impact. One way to protect against serious Type II errors is to require continuing monitoring programs to be evaluated periodically, so that a larger Type II error rate can be used. Even then, it is not advisable to make decisions based on tests that have a β error rate greater than 0.20. We have used α and β levels of 0.10 and 0.20 respectively, and believe that these levels will usually achieve a realistic balance among sample sizes and possible consequences. Even though our choice of error rates is partly qualitative, we believe that if regulatory bodies were to require that these values be addressed, fewer equivocal studies would result.

A discussion of what constitutes a significant biological impact is clearly beyond the scope of this paper even though a judgment about the population size difference a study should detect is inextricably linked to this question. What constitutes a significant change probably is dependent upon many factors, site specifics, species, economic value, compensation mechanisms, etc., and no scientifically acceptable procedure for its quantification is available. However, we have prepared Figure 1 to illustrate the relationship between sample size, error rates and statistically detectable differences. By adopting the error rates above and the paired station approach, monitoring programs can detect smaller changes with increasing sample size. Figure 1 shows the sample size which results for four hypothetical standard deviations (which include our observed range). The abscissal values (Δ log units) can be converted to percent change ($\Delta\%$) (i.e., preoperational control/treatment vs operational control/treatment) by the relationship, $\Delta\% = (e^{\Delta \log \text{ units}} - 1) \times 100$. For example, a difference of 0.5 log units is comparable to a 65% change in original arithmetic ratio.

FIGURE 1. Relationship Between Sample Size and Detectable Difference in Preoperational vs Operational Means Using a Multiplicative Log Transformed Model Based on Paired Stations.

APPLICATION: ZION NUCLEAR POWER PLANT ZOOPLANKTON MONITORING

We have used data obtained at Zion Nuclear Power Plant (Murarka et al. 1976a) to demonstrate the paired station ANOVA procedure discussed above. The plant (two units of 1100 MW) is located on the western shore of Lake Michigan, and sampling stations were established both north and south of the site and at several depths, Figure 2. Our a posteriori pairing of stations, for example Station 28 with 5, is of unknown ecological validity. However, the similarity of the depth and distance transects leads us to believe our pairings were reasonable. Monthly samples were collected from January 1972 through December 1976 and the data presented as estimated density per m^3 for total zooplankton.

Zero zooplankton densities were found on three sampling occasions. These counts were not consistent with either the counts at the nearby stations or with the general pattern of zooplankton densities in the other years of collection. In addition, our statistical analysis of a similar set of phytoplankton data showed that ANOVA cells containing zero values had cell means that were significantly different from all other cell means (McKenzie et al. 1977). Even though the three zero counts are probably statistically significant, they were deleted from the subsequent analysis. Moreover, we did not have access to field crew log books in order to examine the possibility of sampling gear malfunction or other abnormal circumstances.

In order to demonstrate the effectiveness of using paired station differences as a method for removing seasonal cycles and thereby reducing the serial correlations in density, we have reproduced Figure 3 from Murarka et al. (1976a, appendix Z-2). Seasonal cycles in zooplankton density are illustrated with Station 5 and 28 representing a treatment-control station pair in the north zone on the 3 m depth contour (Figure 2). The log transformed differences for five years between the total densities show these differences fluctuate randomly about the mean, approximately zero, with no obvious cycles (Figure 4). A test of randomness (Ryan et al. 1976) for the series of log differences (by examining the number of runs) supports this visual appraisal ($P = 0.89$). Tests for each of the other five station pairs indicated that all were apparently random ($P \geq 0.59$). Murarka et al. (1976a) also reached a similar conclusion based on autocorrelation analysis. Finally, a comparison of Figures 3 and 4 shows that the paired difference procedure caused the considerable oscillations to be either damped and/or removed or displaced.

The log transformation of the multiplicative analysis of variance model used to analyze the zooplankton data can be written:

$$\log y_{ijklm} = \log \mu + \log S_i + \log M_j + \log Z_k + \log D_l + \log SM_{ij} + \log SD_{il} + \log MZ_{jk} + \log MD_{jl} + \log ZD_{kl} + \log \epsilon_{ijklm}$$

FIGURE 2. Field Sampling Locations for Plankton Near Zion Nuclear Power Station (From Murarka et al. 1976a).

FIGURE 3. Total Zooplankton Comparison for Control Station 5 and Treatment Station 28, Zion Nuclear Power Plant (From Murarka et al. 1976a).

FIGURE 4. Station Mean (Log Difference) for Total Zooplankton Densities at the North Zion Stations 5 and 28 (Figure 2).

Where:

- y_{ijklm} = observed ratio C/T (C = density at the control station,
T = density at the treatment station)
 μ = mean effect
 S_i = status effect (1 - preoperational, 2 - operational)
 M_j = month effect (j = 1 to 12)
 Z_k = zone effect (1 - north, 2 - south)
 D_l = effect of the depth of water at station pairs
(3, 9, 18 meters)
 SM_{ij} = interaction between status and month
 ϵ_{ijklm} = multiplicative error (lognormally distributed)

Our results (Table 1) show that the status main effect was not statistically significant. However, the status by month interaction was found to be significant ($P = 0.06$). Further examination of the interaction showed that data obtained in January was the major reason for this judgment. Preoperational data obtained during January 1972 and 1973, especially for the south zone station pairs for 1972, were much higher at the treatment stations, i.e., near the plant site. An opposite relationship, (i.e. where control station densities exceeded those at treatment stations) typifies the operational January samples. Assuming the interaction term is not a function of the scale of measurement, the log difference approach was useful in detecting statistically significant changes during the preoperational and operational periods, but interpretation should be the province of a qualified ecologist.

TABLE 1. Analysis of Variance for the Log Difference Between Total Zooplankton Densities at the Zion Control and Treatment Stations.

Source	SS	df	MS	F	Probability
Main Effects					
Status	0.04	1	0.04	0.16	0.69
Month	3.80	11	0.35	1.36	0.19
Zone	0.70	1	0.70	2.76	0.10
Depth	0.42	2	0.21	0.82	0.44
Interactions					
S X M	5.00	11	0.46	1.79	0.06
S X Z	0.04	1	0.04	0.14	0.71
S X D	0.22	2	0.11	0.42	0.65
M X Z	1.13	11	0.10	0.40	0.95
M X D	7.44	22	0.34	1.33	0.15
Z X D	0.64	2	0.32	1.26	0.28
Residual	71.76	283	0.25		
Total	91.56	348			

The density changes which could be detected when results are similar to those in Table 1 were estimated using the variance (0.254) and degrees of freedom (283) from that table. Our results indicate that differences of approximately 0.16 log units (about a 20% ratio change) in status means could be detected when $\alpha = 0.10$, and $\beta = 0.20$. This indicates that fairly powerful monitoring programs can be conducted with a reasonable expenditure of funds (sampling effort).

The Cladoceran component of the zooplankton community was characterized by the largest seasonal cycles. Densities were usually zero or near zero during February and March and as large as 1.6×10^6 organisms per cubic meter during August. Densities at control and treatment station pairs also exhibited the largest differences. The range of values for $\log(y)$ obtained were about ± 4 log units compared to ± 2 for total zooplankton. However, the ANOVA indicated that the error variance was not affected by the increased variability and that there were no significant changes in Cladocera numbers related to the operational status of Zion. Thus, we were able to deal with these larger season range of densities and differences between station pairs using the methods above.

The effectiveness of the station pairing in reducing error variance can be evaluated by comparing the error variance terms from both paired and unpaired analyses of variance. In general, with paired samples of size n , the standard error of the mean difference $\bar{D} = \bar{X}_1 - \bar{X}_2$ is $\sigma_D/\sqrt{n}$ where σ_D is the standard deviation of the population of paired differences. When the control and treatment stations are treated as independent groups, the standard error of the difference in means, $\bar{X}_1 - \bar{X}_2$, is calculated by; $\sqrt{2} \sigma/\sqrt{n}$, where σ is the pooled standard deviation of the original population from which the sample of size $2n$ was obtained. Thus, an approximate comparison of the two approaches can be made in terms of the variances i.e., σ_D^2 and $2\sigma^2$ (Snedecor and Cochran 1967). In the present case this comparison may underestimate the advantages of pairing because of the effects of serial correlation on the unpaired analysis discussed earlier. Estimates of σ_D^2 were available from our previous analyses and we obtained estimates of $2\sigma^2$ from an analysis of variance of the data without station pairing. These estimates of σ_D^2 and $2\sigma^2$ for the zooplankton data are in Table 2.

Pairing resulted in a much smaller variance of mean differences, $0.254/n$, versus $1.390/n$. This implies that when treated as independent samples, the sample size would have to be increased from n pairs to $(1.390)/(0.254) n$ pairs, or about 5.5 n pairs, in order to give the same variance of the mean difference as the paired analysis. The magnitude of the variance determines the size of density differences that can be detected at fixed α and β error rates for samples of size n . Thus, 5.5 times more samples are needed when using the non-paired station approach to assess impact as compared to the paired station approach. For the Zion monitoring program, this implies that instead of the nearly 700 observations obtained, approximately 3800 observations are needed to detect differences at the same levels of statistical significance.

TABLE 2. Summary of Estimated Variances, Paired and Independent Station Densities (Log Transformed), for the Components of the Zooplankton Communities Sampled at Zion

Zooplankton Component	Variance (Mean Difference)	Variance (Independent)	Ratio
	σ_D^2	$2\sigma^2$	$2\sigma^2/\sigma_D^2$
Total	0.25	1.39	5.5
Crustacea	0.31	1.41	4.5
Copepoda	0.29	1.32	4.5
Cladocera	0.77	3.06	4.0
Rotifera	0.62	4.67	7.5

SUMMARY

The concept of control and treatment station pairs is the fundamental basis for the proposed experimental method. This concept is based on the hypothesis that the relationship between biota at two stations can be estimated from the preoperational period and that the relationship continues during the operational period. Any changes observed in this relationship during the operational period are assumed to be the result of power plant impacts. Therefore, it is important that selected station pairs respond to natural environmental changes in a consistent manner.

A major problem identified for establishing station pairs is the location of the control stations. The universal heterogeneity in the environment prevents the establishment of identical station pairs. The requirement that the control station remains unaffected by the operation of the power plant dictates a spacial separation with associated differences in habitat. Thus, selection of the control station must be based on a balance between two criteria: 1) it must be far enough away from the plant site to be beyond the plant influence, and 2) it must be close enough to the treatment station so that biological communities respond similarly and consistently to natural environmental changes.

It appears that the paired station design can be utilized for some components of the aquatic ecosystem, i.e., zooplankton. Results from the Zion monitoring program allowed us to evaluate our procedures with data from a large lake environment. Its application to riverine, estuarine and marine environments is complicated by varying hydrodynamic factors and environmental gradients, so additional research will be required to evaluate the procedures for use with such data.

Guidelines were proposed for problems encountered during the statistical analysis of the data. An experimentwise error rate was discussed and a criterion for selecting the experimental unit given. Levels of the statistical error rates for testing hypotheses were proposed. An analysis of data showed that plant-induced changes of about 50% or smaller can be detected at an α level of 0.10 and β level of 0.2. Additional study is needed in order to determine the magnitude of density changes that can be considered significant ecological impacts.

The proposed methodology appears to be useful in a variety of field experimental situations. Although developed and a posteriorily applied to assessing power plant impact, the problems of quantifying a response in natural populations are similar. The treatment control station or area approach should be beneficial in separating the treatment effect from the natural variability and normal changes in an ecological system.

SUMMARY

The concept of control and treatment station pairs is the fundamental basis for the proposed experimental method. This concept is based on the hypothesis that the relationship between two stations can be estimated from the pre-treatment period and that the relationship can be estimated from the post-treatment period. Any changes observed in this relationship during the operational period are assumed to be the result of power plant impacts. Therefore, it is important that selected station pairs respond to natural environmental changes in a consistent manner.

A major problem identified for establishing station pairs is the location of the control stations. The universal heterogeneity in the environment prevents the establishment of identical station pairs. The requirement that the control station remain unaffected by the operation of the power plant dictates a station separation with associated differences in habitat, inlet, selection of the control station must be based on a balance between the criteria: 1) it must be far enough away from the plant site to be beyond the plant influence, and 2) it must be close enough to the treatment station to that biological communities respond similarly and consistently to natural environmental changes.

It appears that the paired station design can be utilized for some components of the aquatic ecosystem, i.e., zooplankton. Results from the monitoring program allowed us to evaluate our procedures with data from a large lake environment. Its application to riverine, estuarine and marine environments is complex and requires hydrodynamic factors and environmental gradients. Additional research will be required to evaluate the procedures for use with such water.

Guidelines were proposed for problems encountered during the statistical analysis of the data. An experimental design was discussed and a criterion for selecting the experimental unit given. Levels of the statistical error rates for testing hypotheses were presented. An analysis of data showed that plant-induced changes of about 20% or greater can be detected at a level of 0.10 and a level of 0.05. Additional study is needed in order to determine the magnitude of density changes that can be considered significant ecological impacts.

The proposed methodology appears to be useful in a variety of field experimental situations. Although developed and a posteriori applied to assess power plant impact, the problem of quantifying a response to natural populations are similar. The treatment control station or area approach should be beneficial in separating the treatment effect from the natural variability and natural changes in an ecological system.

REFERENCES

- Adams, S. M., P. A. Cunningham, D. D. Gray, K. D. Kumar and A. J. Witten. 1977. A Critical Evaluation of the Nonradiological Environmental Technical Specifications, Vol. 1. Program Description, Summary and Recommendations. ORNL/NUREG/TM-69, Oak Ridge National Laboratory, Oak Ridge, TN.
- Autchison, J. and J. A. Brown. 1969. The Lognormal Distribution. Cambridge University Press, New York.
- Box, G. E. P., and G. M. Jenkins. 1970. Time Series Analysis Forecasting and Control. Holden Day, San Francisco, CA., 553 pp.
- Box, G. E. P., and G. C. Tiao. 1975. Intervention analysis with applications to economic and environmental programs. J. Amer. Stat. Asso. 70:70-73.
- Dunn, O. J. 1961. Multiple comparison among means. J. Amer. Stat. Asso. 56:52-64.
- Eberhardt, L. L. 1976. Quantitative ecology and impact assessment. J. Environ. Manage. 4:27-70.
- Eberhardt, L. L. 1978. Appraising variability in population studies. J. Wildl. Manage. 42(2):207-238.
- Fickeisen, D. H., M. J. Schneider and J. C. Montgomery. 1975. A comparative evaluation of the Weiss saturometer. Trans. Amer. Fish. Soc. 104(4):816-820.
- Gore, K. L., J. M. Thomas, L. D. Kannberg and D. G. Watson. 1977. Evaluation of Nuclear Power Plant Environmental Impact Prediction, Based on Monitoring Programs, Summary and Recommendations. BNWL-2153/NRC-1, Pacific Northwest Laboratory, Richland, WA. 33 pp.
- Gore, K. L., J. M. Thomas and D. G. Watson. 1978. Quantitative evaluation of environmental impact assessment, based on aquatic monitoring programs at three nuclear power plants. J. Environ. Manage. (6).
- McCaughran, D. A. 1977. The quality of inferences concerning the effects of nuclear power plants on the environment. Pp. 229-242. In: W. Van Winkle (ed.), Proceedings of the Conference on Assessing the Effects of Power Plant Induced Mortality on Fish Populations. Pergamon Press, NY.
- McKenzie, D. H., L. D. Kannberg, K. L. Gore, E. M. Arnold and D. G. Watson. 1977. Design and Analysis of Aquatic Monitoring Programs at Nuclear Power Plants. PNL-2423/NRC-10, Pacific Northwest Laboratory, Richland, WA. 126 pp.

- Murarka, I. P., A. Policastro, E. Daniels, J. Ferrante and F. Vaslow. 1976a. An Evaluation of Environmental Data Relating to Selected Nuclear Power Plant Sites: The Zion Nuclear Power Station Site. ANL/EIS-5, Argonne National Laboratory, Argonne, IL.
- Murarka, I. P., A. J. Policastro, J. G. Ferrante, E. W. Daniels and G. J. Marmer. 1976b. An Evaluation of Environmental Data Relating to Selected Nuclear Power Plant Sites: A Synthesis and Summary With Recommendations. ANL/EIS-8, Argonne National Laboratory, Argonne, IL.
- Pielou, E. C. 1969. An Introduction to Mathematical Ecology. Wiley-Interscience, NY. 286 pp.
- Pirie, W. R., and W. A. Hubert. 1977. Comments and assumptions in statistical analyses. Trans. Amer. Fish. Soc. 106(6):646-648.
- Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin 191, Fisheries Research Board of Canada. Ottawa, Can. 382 pp.
- Ryan, T. A., B. L. Joiner and B. F. Ryan. 1976. MINITAB. Student Handbook. Duxbury Press, North Scituate, MA. 341 pp.
- Scheffe, H. 1959. The Analysis of Variance. John Wiley and Sons, Inc., NY. 477 pp.
- Seal, H. 1964. Multivariate Statistical Analysis for Biologists. Mathuen and Co., LTD, London. 209 pp.
- Seber, G. A. F. 1973. The Estimation of Animal Abundance. Griffin and Company LTD, London. 506 pp.
- Snedecor, G. W., and W. G. Cochran. 1967. Statistical Methods, Sixth Edition. Iowa State University Press, IA. 593 pp.
- Steele, R. G. D., and J. H. Torrie. 1960. Principles and Procedures of Statistics. McGraw-Hill Book Company, Inc., NY. 481 pp.
- Thomas, J. M. 1977. Factors to consider in monitoring programs suggested by statistical analysis of availability data, pp. 243-255. In W. Van Winkle (ed.), Proceedings of the Conference on Assessing the Effects of Power Plant Induced Mortality on Fish Populations. Pergamon Press, NY.
- Thomas, J. M., J. A. Mahaffey, K. L. Gore and D. G. Watson. 1978. Statistical methods used to assess biological impact at nuclear power plants. J. Environ. Manage. (6).

Tukey, J. W. 1949. One degree of freedom for nonadditivity.
Biometrics 5:232-242.

Winer, B. J. 1971. Statistical Principles In Experimental Design. Second
Edition. McGraw-Hill, NY. 907 pp.

Jurk, G. W. 1949. One degree of freedom for nonadditivity.
Biometrics 5: 233-245.

Winer, B. J. 1971. Statistical Principles in Experimental Design. Second
Edition. McGraw-Hill, NY. 507 pp.

DISTRIBUTION

No. of
Copies

No. of
Copies

OFFSITE

OFFSITE

10 U. S. Nuclear Regulatory
Commission
Environmental Specialists Branch
Bethesda, MD 20014

Frank Swanberg
Division of Safeguards, Fuel
Cycle and Environmental
Research
U.S. Nuclear Regulatory
Commission
Washington, DC 20555

Jack Mattice
Electric Power Research Institute
P. O. Box 10412
Palo Alto, CA 94303

ONSITE

30 Pacific Northwest Laboratory

Helen McCanmon
EV-34, GTN
U.S. Department of Energy
Washington, DC 20545

K. E. Byers (4)
N. R. Hinds
D. H. McKenzie (10)
J. M. Thomas (6)
W. L. Templeton
B. E. Vaughan
Publishing Coordination (2)
Technical Information (5)

Paul Rago
U. S. Fish and Wildlife Service
1451 Green Road Room 105
Ann Arbor, MI 48105

U.S. Nuclear Regulatory
Commission
Division of Technical
Information and Document
Control
7920 Norfolk Avenue
Bethesda, MD 20014

NRC FORM 335 (11-81)		U.S. NUCLEAR REGULATORY COMMISSION BIBLIOGRAPHIC DATA SHEET		1. REPORT NUMBER (Assigned by DDC) NUREG/CR-2803 PNL-4138	
4. TITLE AND SUBTITLE (Add Volume No., if appropriate) Improved Field Experimental Designs and Quantitative Evaluation of Aquatic Ecosystems				2. (Leave blank)	
7. AUTHOR(S) D. H. McKenzie and J. M. Thomas				3. RECIPIENT'S ACCESSION NO.	
9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Pacific Northwest Laboratory P. O. Box 999 Richland, Washington 99352				5. DATE REPORT COMPLETED MONTH YEAR April 1984	
12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Division of Health, Siting and Waste Management Office of Nuclear Regulatory Research U.S. Nuclear Regulatory Commission Washington, DC 20555				DATE REPORT ISSUED MONTH YEAR May 1984	
13. TYPE OF REPORT				6. (Leave blank)	
15. SUPPLEMENTARY NOTES				8. (Leave blank)	
16. ABSTRACT (200 words or less) <p>We used the paired-station concept and a log transformed analysis of variance as methods to evaluate zooplankton density data collected during five years at an electrical generation station on Lake Michigan. To discuss the example and the field design necessary for a valid statistical analysis, we provide considerable background on the questions of selecting 1) sampling station pairs, 2) experimentwise error rates for multi-species analyses, 3) levels of Type I and II error rates, 4) procedures for conducting the field monitoring program, and 5) a discussion of the consequences of violating statistical assumptions. We include details for estimating sample sizes necessary to detect changes of a specified magnitude.</p> <p>Both statistical and biological problems with monitoring programs (as now conducted) are addressed; serial correlation of successive observations in the time series obtained was identified as one principal statistical difficulty. Our procedure reduces this problem to a level where statistical methods can be used confidently.</p>				10. PROJECT/TASK/WORK UNIT NO.	
17. KEY WORDS AND DOCUMENT ANALYSIS				11. FIN NO. FIN B2072	
17a. DESCRIPTORS				14. (Leave blank)	
17b. IDENTIFIERS, OPEN-ENDED TERMS				19. SECURITY CLASS (This report) Unclassified	
18. AVAILABILITY STATEMENT Unlimited				21. NO. OF PAGES	
20. SECURITY CLASS (This page)				22. PRICE \$	

1. REPORT NUMBER NRC-4138 NRC-4138-2003		2. U.S. NUCLEAR REGULATORY COMMISSION BIBLIOGRAPHIC DATA SHEET	
3. TITLE AND SUBTITLE Improved Field Experimental Designs and Quantitative Evaluation of Aquatic Ecosystems		4. AUTHOR(S) D. W. McKenzie and J. M. Thomas	
5. REPORTING ORGANIZATION NAME AND ADDRESS Pacific Northwest Laboratory P. O. Box 999 Richland, Washington 99352		6. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research Division of Health, Safety and Waste Management 1155 Rockledge Drive Arlington, Virginia 22204	
7. DATE REPORT COMPLETED May 1984		8. DATE REPORT ISSUED May 1984	
9. FUNDING NUMBERS NRC-4138-2003		10. PROJECT/WORK UNIT NO. NRC-4138-2003	
11. TYPE OF REPORT Technical Report		12. TYPE OF REPORT Technical Report	
13. ABSTRACT We used the paired-station concept and a log transformed analysis of variance as methods to evaluate zooplankton density data collected during five years at an electrical generation station on Lake Michigan. To discuss the example and the field design necessary for a valid statistical analysis, we provide considerable background on the question of selecting 1) sampling station pairs, 2) experimental error rates for multi-series analysis, 3) levels of Type I and II error rates, 4) procedures for conducting the field monitoring program, and 5) a discussion of the consequences of violating statistical assumptions. We include details for estimating sample sizes necessary to detect changes of a specified magnitude.		14. ABSTRACT Both statistical and biological problems with monitoring programs (as now conducted) are addressed; serial correlation of successive observations in the time series obtained was identified as one principal statistical difficulty. The procedure reduces this problem to a level where statistical methods can be used confidently.	
15. KEY WORDS AND SUBJECT ANALYSIS Aquatic ecosystems; Field experiments; Statistical analysis; Zooplankton density		16. DISTRIBUTION STATEMENT Unlimited	
17. DISTRIBUTION STATEMENT Unlimited		18. DISTRIBUTION STATEMENT Unlimited	