University of Washington
Department of Civil and Environmental Engineering

SAMPLING DESIGN FOR AQUATIC ECOLOGICAL MONITORING PHASE II REPORT VOLUME I

Brian W. Mar
Dennis P. Lettenmaier
Richard R. Horner
Joanna S. Richey
Richard N. Palmer
Steven P. Millard
Gary L. Thomas

Water Resources Series
Technical Report No. 86
September 1983

Seattle, Washington
98195

Department of Civil Engineering
University of Washington
Seattle, Washington 98195

SAMPLING DESIGN FOR AQUATIC ECOLOGICAL MONITORING PHASE II REPORT VOLUME I

Brian W. Mar
Dennis P. Lettenmaier
Richard R. Horner
Joanna S. Richey
Richard N. Palmer
Steven P. Millard
Gary L. Thomas

Water Resources Series
Technical Report No. 86

Setember 1983

SAMPLING DESIGN FOR AQUATIC ECOLOGICAL MONITORING

PHASE II REPORT - METHODS DEVELOPMENT

EPRI Project RP 1729-1

VOLUME I

Prepared by
Environmental Engineering and Science Program
Department of Civil Engineering
University of Washington
Seattle, Washington 98195

Brian W. Mar
Dennis P. Lettenmaier
Richard R. Horner
Joanna S. Richey
Richard N. Palmer
Steven P. Millard
Gary L. Thomas

Prepared for the
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94303

EPRI Project Manager
J. S. Mattice

Ecological Studies Program

FEBRUARY 1984

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

TABLE OF CONTENTS - Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

These appendices are located in Volume II.

ACKNOWLEDGMENTS

This research was conducted by the members of the University of Washington Electric Power Aquatic Monitoring research team. The major contributors to the basic components of this project are:

Conceptual framework for ecological assessment: Richard R. Horner, Joanna S. Richey, Gary L. Thomas

Delphi process: Joanna S. Richey, Richard R. Horner

Statistical methods and modeling: Steven P. Millard, John R. Yearsley

Optimization: Richard N. Palmer, Gary J. Weins, Silvia C. Vega-Gonzales

Cost Data: Gail G. Mattson

Richard R. Horner served in the dual role of scientific contributor as well as project coordinator and budget manager.

We wish to acknowledge gratefully the guidance and assistance of Dr. Jack S. Mattice in establishing and maintaining our contacts with members of the electric power utility community, and in coordinating the activities of the project advisors and the review panel. Dr. Robert G. Otto of Robert G. Otto Associates (Baltimore, MD) and Dr. Daniel H. McKenzie of Battelle Pacific Northwest Laboratories (Richland, WA) made numerous valuable contributions in their role as advisors. Furthermore, members of the Delphi Panel, made a unique contribution to the research effort. Without their commitment, Phase II of this project could not have been successful.

Douglas M. Johnston provided skillful computer programming. Jan H. Spyridakis served very capably as technical editor for this report. Jane A. Lybecker performed the word processing with her usual excellence. These individuals deserve a large measure of credit and our sincere gratitude for their efforts.

This project is truly a team effort, and we wish to recognize the cooperation and dedication of all the participants.

Dennis P. Lettenmaier
Brian W. Mar
Principal Investigators

EXECUTIVE SUMMARY

The research results described in this report are the product of Phase II of a four phase effort to develop and test methods to improve the design of cost-effective aquatic monitoring programs. The PHASE II work statement (Appendix A) defined the following tasks: (1) to develop a handbook and software containing a framework and methods to improve the design of aquatic monitoring programs, (2) to evaluate and select methods to include in the handbook and software, (3) to develop and parameterize a method to perform cost-effective optimization on monitoring program elements, and (4) to identify potential case studies that can be used to evaluate the handbook and software. The first phase of this research project, completed in April 1982, reported the existing state-of-the-art and the outstanding issues in aquatic monitoring of thermal electric power plants. The third phase of this effort will evaluate the methods presented in this report by using selected case studies from utilities active in aquatic monitoring. The final phase of this research will be the improvement of the handbook methods based on the evaluation results, the final documentation of the project results, and the technology transfer of these results.

The framework for the handbook and software has been designated the Electric Power Aquatic Monitoring (EPAM) model. Chapter 1 of this report describes the basic concepts inherent in EPAM and covers prior efforts which form a portion of its foundation. Chapter 2 presents a user's guide to EPAM, and describes the hypothesis generating procedures, the sampling design alternatives, and the optimization methods. At each step in the EPAM process, the user is asked to consider the value of additional information versus the cost of obtaining such data. Chapter 3 describes the approach used to develop and refine the major components of EPAM. A nationwide group of 65 experts operating as a Delphi panel was employed to evaluate and resolve alternative frameworks to be incorporated in EPAM and to reduce uncertainty in the specific issues addressed by EPAM. The Delphi process is a method of structuring group communication to deal with a complex problem, often one with no generally accepted decision criteria, in an attempt to develop a consensus. The Delphi exercise was implemented with a series of questionnaires developed by the research team and mailed to panel members. Use of the Delphi panel permitted expert opinion to be incorporated into EPAM in an effective manner.

Chapter 4 presents the theory and application of several computational algorithms that perform specific functions in EPAM. One algorithm, termed CHOICE, provides a means of ranking options according to designated criteria. CHOICE may be activated in EPAM to select various elements of a monitoring program design on a priority basis. The remainder of Chapter 4 concerns statistical aspects of the EPAM framework.

A substantial portion of this research investigated certain statistical problems that have limited the success of aquatic monitoring programs. This inquiry focused on the linking of optimization methods with factorial treatment and time series analysis, plus the explicit recognition of the spatial and temporal correlation of errors. A multivariate framework for sampling program optimization was developed for those situations in which spatially-correlated errors must be considered. A pairwise comparison concept was selected as the general prioritization and optimization procedure to be used in EPAM. Parametric cost data were collected for use in cost optimization in EPAM.

A prototype handbook, a demonstration of the EPAM model, details of the analysis and theoretical derivations supporting this research, and work statements for the various research phases are found in the appendices of this report. This report is the source book for the Phase III evaluation of the EPAM handbook, software, and methods.

CHAPTER 1

OVERVIEW: BASIC CONCEPTS OF EPAM DEVELOPMENT

INTRODUCTION

This project is a nontraditional attempt by the research team to bring new perspectives to the improvement of aquatic monitoring program design. The research team was assembled to provide skills and insights developed from the statistical analysis and modeling of hydrologic and water quality data, the optimization of multiobjective natural resource allocation problems using subjective inputs, and general system analyses of complex ecosystems.

The overall goal of this project is to incorporate the most accepted aquatic monitoring design practices into a well-documented and easily used framework. Existing practice is being extended and improved for several purposes: (1) to formalize the recognition and use of subjective information, especially to define objectives and criteria to measure these objectives; (2) to provide practical methods to prioritize and increase the cost-effectiveness of monitoring activities; and (3) to focus on the tradeoff between the costs and precision of monitoring short- and long-term effects.

PHASE I SUMMARY

The Phase I report "Sampling Design for Aquatic Ecological Monitoring-Phase I Report," prepared in April 1982, summarized a review of approximately 500 publications and 50 interviews with experts in various aquatic monitoring activities. A workshop was held to review this summary and to permit some of these individuals to discuss and assess Phase I activities. While the details of this workshop are not repeated in this document, the conclusions are summarized to provide background and a point of departure for the Phase II efforts to improve and integrate the existing knowledge.

1. The methodology used in aquatic monitoring design varies in quality and tends to be fragmented by discipline. At one end of the spectrum are outstanding pioneering efforts that provide leadership and innovation to improve the process of monitoring design, while at the other extreme are myopic, biased, and obsolete efforts.

2. In most cases the design of a monitoring program is based on the experience and subjective decisions of a few individuals. Usually these individuals emphasize their areas of expertise and interests and select measurement techniques that they have used in the past. This practice creates a difficulty in the sharing of this expertise that stems from the inability of monitoring program designers to capture subjective insights and then to transfer this knowledge.

3. Potential aquatic effects of thermal power generation are well-recognized, but the magnitude and scope of these effects are site-specific and difficult to estimate.

4. Specific hypotheses of impact and expected levels of change caused by plant operation are rarely defined as part of the design of a monitoring program. Moreover, formalized regulatory specifications for aquatic monitoring programs are very general, and regulatory agencies tend to operate on a case-by-case basis.

5. The development of statistically valid monitoring programs requires not only the identification of hypotheses, but also the definition of treatments, controls, and variables that can be measured. If observations are correlated in space and/or time, the analysis of the monitoring data must be adjusted.

6. While water and pollutant movement can be modeled with some precision, most scientists lack confidence in using ecosystem models.

7. The costs of conducting monitoring programs are seldom reported, and the few data available cannot be standardized because of inconsistent assumptions and formats used in presenting the data. Cost-effectiveness has not been an important criterion in most designs.

The consensus, developed at the workshop, concluded that Phase I provided several new directions for Phase II. The workshop panel particularly redirected the Phase II work toward the development of a conceptual framework to generate hypotheses to be tested in monitoring designs. On the other hand, the proposed investigation of models to design monitoring programs and the risk/benefit analysis of monitoring program design were deemphasized by the panel. The panel believed that efforts to compile and define the effectiveness of techniques to measure a given variable would require much greater resources than those available in this project and should be the subject of another investigation. These recommendations altered the Phase II goals; rather than being concerned equally with what to measure in a cost-effective aquatic monitoring program, as well as where, when, how, and how often, Phase II goals now center most fundamentally on what to measure. Setting aside the question of how to measure, the research proceeded to devise means of determining where, when, and how often. Also as a result of the workshop deliberations, the emphasis on an effort to assess expert opinion was increased substantially beyond the initially anticipated level for Phase II. The review panel concurred with the recommendation that this assessment be conducted by use of the Delphi process (Linstone and Turoff, 1975).

PHASE II WORK STATEMENT

The contractual statement of work for Phase II, developed from the Phase I surveys and the workshop panel recommendations, is presented in Appendix A. Following is a brief summary of the tasks and references to pages on which discussions of the respective issues begin:

Task 1A Develop a methods handbook and interactive computer software containing:

a. A framework to evaluate and prioritize elements of a monitoring program (p. 4)
b. Alternative approaches to monitoring programs design (p. 7)
c. A method to generate hypotheses to be tested (p. 8)
d. Methods to test the hypotheses (p. 8)
e. Guidelines for selection of statistical methods for application to the data (p. 58)
f. Methods to select the optimized set of hypotheses and tests and an application of the methods to be demonstrated in an example (p. 49, Appendix C.3)

Task 1B Develop a prototype interactive software system that can be used in the application of the Task 1A products (Appendix C.2)

Task 2 Compare the effectiveness of the following methods to detect both long- and short-term impacts

a. Use of expert judgment (p. 50)
b. Use of statistical/optimization techniques (p. 58)
c. Use of simulation models (p. 114)

Task 3 Develop optimization techniques and cost data for use in designing aquatic monitoring programs (p. 86)

Task 4 Identify data sets generated by actual monitoring programs that can be used in evaluating the results of the Phase II effort (Appendix E)

OVERVIEW OF EPAM

Aquatic monitoring design must integrate concepts from many disciplines. The framework developed in Phase II, EPAM (a model for Electric Power Aquatic Monitoring), attempts to integrate: (1) the concepts that aquatic scientists use to generate hypotheses concerning the impact of thermal electric generating facilities on the aquatic environment with the methods statisticians use to design experiments to test hypotheses and demonstrate significance, (2) the methods experimentalists use to obtain and assure the quality of sampling data, and (3) the methods systems analysts use to optimize and allocate resources in a cost-effective manner. Since much of the information employed in the design of an aquatic monitoring program may be subjective or require value judgments, the EPAM framework includes methods to accommodate qualitative as well as quantitative information.

EPAM was designed to follow the classical pattern of allocating resources to a set of scientific inquiries. The basic steps in this pattern are: (1) identifying the potential causes and effects of interest, (2) formulating hypotheses that can be tested to determine if cause and effect are related, (3) designing experiments that can observe the proposed changes and have statistical rigor, and (4) balancing the cost of each experiment against the value of the information gained from that experiment. Since aquatic monitoring design may be an iterative process where increasing precision and detail must be provided as the process continues, EPAM is designed to process a spectrum of input formats ranging from general subjective information to precise and replicated observations.

NOMENCLATURE

Several terms have been used in EPAM and in this report to describe specific components in the EPAM framework. MONITORING PROGRAM is the term used to describe a set of SAMPLING PROGRAMS that collect data to test HYPOTHESES concerning potential responses of the aquatic environment to the activities related to a thermal electric generating facility. Each SAMPLING PROGRAM is designed to test a specific HYPOTHESIS. A HYPOTHESIS proposes that an IMPACT AGENT (a factor causing change) is affecting a TARGET (a factor proposed to be changed).

With this nomenclature the EPAM framework can be described as a sequence of procedures and methods to perform the following functions: (1) to identify IMPACT AGENTS associated with thermal electric generating facilities, and to identify TARGETS present that have a potential to be changed by these IMPACT AGENTS; (2) to combine and rank IMPACT AGENT and TARGET pairs in order to generate HYPOTHESES that may need to be tested in a MONITORING PROGRAM; (3) to determine the cost and effectiveness of alternative SAMPLING PROGRAMS that will test a given HYPOTHESIS, or to design the most effective SAMPLING PROGRAM given a cost constraint; and (4) to optimize or find the most cost-effective set of SAMPLING PROGRAMS to include in a MONITORING PROGRAM for site-specific goals.

BASIC FRAMEWORK

EPAM is a hierarchical framework containing four major components or levels. Figure 1.1 presents a flow chart representing the EPAM structure and components. LEVEL 0 contains only the optimization process to identify the most cost-effective set of SAMPLING PROGRAMS needed to meet the goals of the user. If the alternative SAMPLING PROGRAMS are not well known or need better definition, LEVEL 1 is summoned. LEVEL 1 contains information and methods to design SAMPLING PROGRAMS (the locations and timing of data collection) to

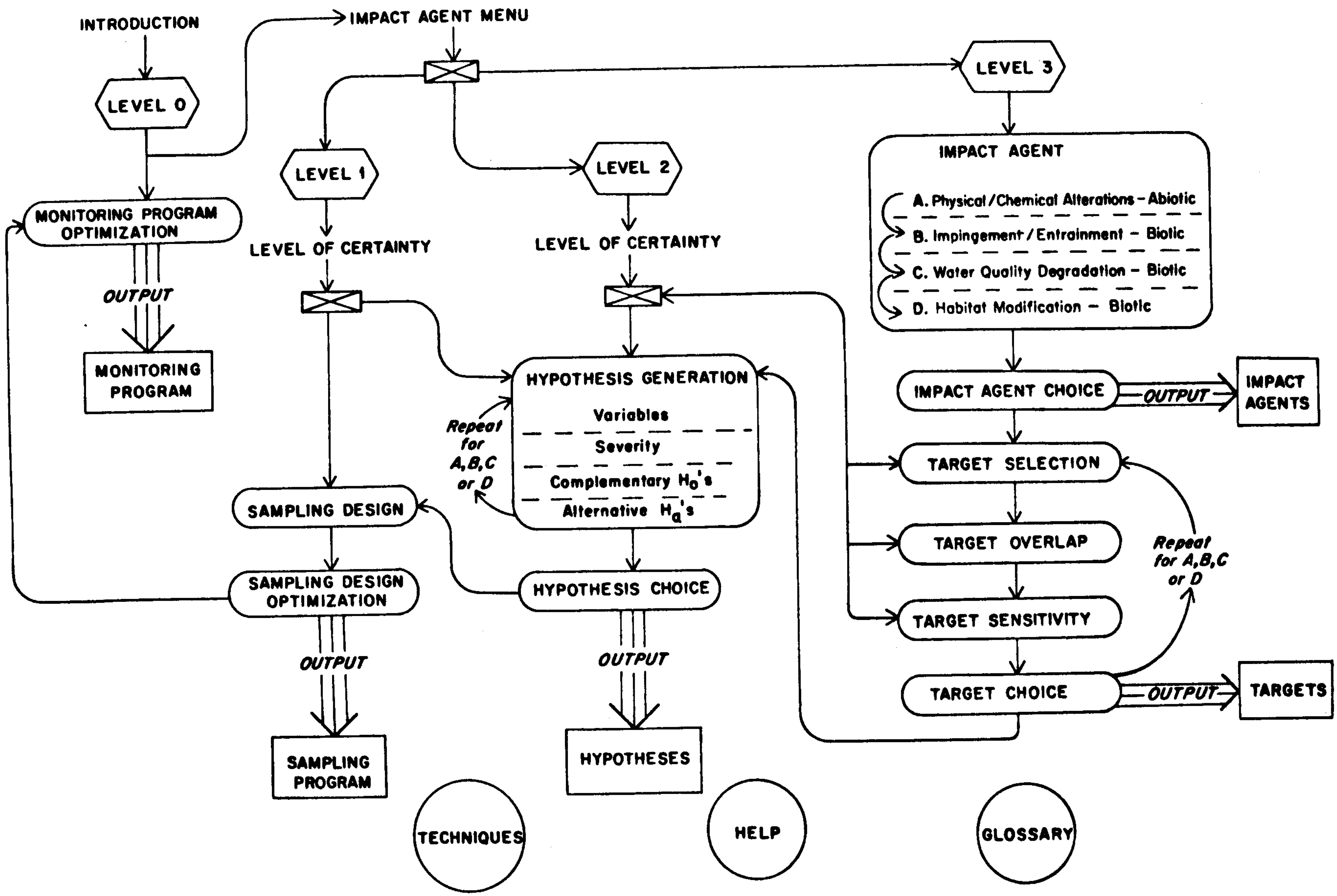

Figure 1.1. EPAM Flow Chart

test specific HYPOTHESES. If HYPOTHESES are not well defined, then methods in LEVEL 2 are available to assist in the generation of HYPOTHESES, as well as the ranking of a list of candidate HYPOTHESES in order to select those to be tested. If the user needs assistance to establish IMPACT AGENTS or TARGETS that will form the basis for the generation of HYPOTHESIS, LEVEL 3 contains methods to identify and rank these factors. The research and synthesis conducted to develop the EPAM model are described in the subsequent chapters and appendices of this report. A brief overview of the concepts incorporated in EPAM is presented in this chapter for readers seeking only a user's perspective of the model.

LEVEL 0

The ranking process for identifying the most cost-effective set of sampling programs is based on a pairwise comparison method suggested by Saaty (1977). Other multiobjective optimization schemes were much more complex and did not offer a reference point for subjective comparisons of qualitative information, which may enter into aquatic monitoring program decision-making. In LEVEL 0, the user characterizes each sampling program by its cost, the quality of the hypothesis test produced, the importance of the information to the user, and any other attributes of concern. The user then compares each of these CRITERIA pairwise using a scale of 1 to 9, where 1 indicates that the CRITERIA are equally important and 9 indicates that one CRITERION has been demonstrated to be much more important than the other. A 5 is used to indicate an apparent but not demonstrated difference in importance, and intermediate values may be assigned as appropriate.

The user then compares each SAMPLING PROGRAM with all others using each CRITERION as the sole basis for the comparison. These comparisons are tabulated in a matrix format; eigenvectors and eigenvalues are computed and used to establish the tradeoffs and ranking of each SAMPLING PROGRAM in terms of its contributions to the goals of the overall MONITORING PROGRAM as specified by the CRITERIA. This same tradeoff process is modified for use at each level of EPAM whenever optimization or ranking is needed. The theoretical basis for this method is described in Chapter 4.

LEVEL 1

LEVEL 1 contains the optimization and statistical algorithms for SAMPLING PROGRAM DESIGN. The user enters LEVEL 1 after defining the hypotheses to be tested, the variables to be measured, and the sampling costs and the quality of data that can be obtained from each observation. LEVEL 1 guides the user in defining an initial set of sampling locations, frequencies and replicates for testing each hypothesis, and in identifying appropriate statistical tests. Methods are included to examine existing sampling data to determine if the residuals in the data are correlated in either space or time. If the residuals in the data are independent, then methods to optimize a classical analysis of variance (ANOVA) are provided. If the residuals are correlated, a modified ANOVA optimization is available to compensate for such dependence. When ANOVA is inappropriate, alternative statistical designs are examined. Chapter 4 describes the theoretical basis for these methods and supporting information is provided in Appendix D.

LEVEL 2

LEVEL 2 contains methods to guide the user in generating and ranking hypotheses. LEVELS 2 and 3 are conceptually different from LEVELS 0 and 1, where computational algorithms developed by this research are used to optimize MONITORING PROGRAM components. In LEVELS 2 and 3 the user is guided to define HYPOTHESES (LEVEL 2) and IMPACT AGENTS and TARGETS (LEVEL 3), and then prepare tables of data or information that characterize these parameters using existing knowledge. At LEVELS 2 and 3, the user must decide whether to conduct specific studies to improve the quality of the information in the tables, or to use the existing data for the LEVEL 0 and 1 considerations. The user can examine the possible improvement in the MONITORING PROGRAM design if more information or better information is used, versus the cost of obtaining the additional data.

Since the generation of hypotheses requires judgment and experience, as well as knowledge of aquatic monitoring practices, the development of the LEVEL 2 and 3 framework and methods employed a Delphi panel of experts. The process used to obtain the panel's expert judgment and the results obtained are summarized in Chapter 3 of this report. The questionnaires are presented in Appendices B.2 through B.5.

The results of this research, as reviewed and augmented by the Delphi panel, created the LEVEL 2 process where the user generates and ranks hypotheses. The pattern of generation is to examine an IMPACT AGENT/TARGET pair and to develop a PRIMARY HYPOTHESIS of impact. The analysis also attempts to identify any ALTERNATIVE HYPOTHESES that may cause the same TARGET response and COMPLEMENTARY HYPOTHESES (indirect responses) that may result if the PRIMARY HYPOTHESIS occurred. Next the user defines the VARIABLES that could be measured to test the various types of HYPOTHESES. Each HYPOTHESIS is then characterized by the cost and precision associated with the VARIABLES selected to measure the TARGET response, the number of ALTERNATIVE HYPOTHESES that could cause the same outcome, and the number of COMPLEMENTARY HYPOTHESES that may be triggered by the PRIMARY HYPOTHESIS. HYPOTHESES are ranked using the pairwise comparison process to define those that are most cost-effective as defined by the CRITERIA selected.

LEVEL 3

LEVEL 3 guides the user in identifying and selecting IMPACT AGENTS and TARGETS that can potentially serve as the basis for the HYPOTHESES of impact to be defined for a particulate site. Four categories of IMPACT AGENTS associated with thermal electric generating facilities have been defined: (1) physical or chemical alterations, (2) impingement or entrainment, (3) water quality degradation, and (4) habitat modification. Similarly, categories of potential TARGETS (e.g., decomposers, producers, consumers, chemical indices, etc.) have been identified for various receiving water types. This information is available to guide the user in initially identifying IMPACT AGENTS and TARGETS for his or her particular site. The user evaluates each TARGET in terms of ecological, socioeconomic, and regulatory importance. The concentrations of each TARGET in time and space are overlayed with the temporal and spatial distributions of IMPACT AGENTS. Potential severity of the TARGET response to the IMPACT AGENT effect is also evaluated in LEVEL 3. The general philosophy is that the larger the overlay of TARGET domain and IMPACT AGENT area, and the higher the potential severity of the effect, the greater the probability is that any change observed would be caused by the plant.

For application at other EPAM levels, the user is asked to estimate the quality of the data underlying the analysis at LEVEL 3 and the cost to improve the quality of these data. Information on data quality and cost is used in

LEVEL 2 to consider the strength of the foundation for high ranking hypotheses. If the foundation is weak, the user can elect to collect more data for the estimated cost and return to LEVEL 3 to reconsider the IMPACT AGENT and TARGET definitions.

PRIOR WORK

The EPAM approach has its roots in the classical scientific method of inductive inference. As set forth theoretically by Platt (1964), this method attempts to strengthen the inferences drawn by applying the inductive procedure systematically to devise hypotheses and experiments to test them.

Several previous investigators have endeavored to simplify the process of determining what to measure in a monitoring program by providing a set of guidelines. These guidelines use data to identify components of an ecosystem likely to be impacted by a proposed development. Other investigators have specified that particular organisms or biotic components be included in the study regardless of the site-specific conditions. Such mandates have usually been based upon ecological perceptions or the socioeconomic value of the identified components. In general, the criteria for determining what to measure in a monitoring program have been as follows:

1. All biotic components of the ecosystem because of their high degree of interdependence and interaction (Meyers and Bremer, 1975)
2. Species determined to have critical importance based on one or more of the following:
   a. commercial or recreational value
   b. rare or endangered classification
   c. relevance to the well-being of a species in a or b above
   d. importance to the structure or function of the ecosystem (Munn, 1975; U.S. Nuclear Regulatory Commission, 1976; U.S. Environmental Protection Agency, 1977a,b)
3. Species designated by legislative mandate (Tetra Tech, 1981)
4. Species likely to be impacted, based on the interaction in time and space with the characteristics of the plant (Munn, 1975; Fritz et al., 1980; Rago et al., 1983)
5. A combination of the above with consideration given to the existence and feasibility of reliable measurement techniques (States et al., 1978)

6. A combination of the above plus some consideraton for the expected magnitude of the effect and the time scale of recovery. For example, a long-term effect on a commercial fishery would cause much greater concern (and hence, should be quantified more exactly) than a short-term effect, even if the same fishery were affected (States et al., 1978).

This last criterion, particularly, formed the basis for a portion of the EPAM framework. That is, the selection of the subjects of a monitoring scheme (outside of legal mandate) would be based primarily upon three issues: (1) the expected IMPACTS of plant operation, (2) the extent to which the temporal and spatial distribution of these IMPACTS overlap with the measured distribution of the susceptible organisms, and (3) the sensitivity of the TARGET to the IMPACT. The necessity to measure a given component will vary with the importance of the role that organism or biotic component (i.e., TARGET) plays in the ecosystem and in the socioeconomic sphere of the area. Extensive overlap of the distribution of an economically unimportant and highly redundant organism (in terms of function) with the field of influence of the intake structure of a large power plant might not justify the expense and complications of a major monitoring program. However, the partial overlap of a commercial species of critical importance with the same intake structure might justify an extensive program (e.g., an overlap extending only for two months a year during the period of migration of young fry). Thus, the nature of a given program becomes a function of the role that the interacting components play both in the ecosystem and in the political, socioeconomic, and legal spheres.

Other researchers also have considered the definition of hypotheses of impact on monitoring program subjects, as well as the design of experiments to test hypotheses. States et al. (1978) noted that the responses of ecosystem components to impact agents may or may not be linked in the context of formal hypotheses of impact. Holling (1978) and Fritz et al., (1980) generally adhered to the concept of hypothesis-based design, and advocated developing such a design through the assembly of a multidisciplinary panel of specialists familiar with the site to be monitored. Proctor et al. (1980) also favored basing monitoring efforts on hypothesis testing. Rago et al. (1983) believed that the logical process of monitoring program design could be standardized, although specific elements of a given program should be selected by investigators at the site. They recommended proceeding by developing a conceptual model of the aquatic system in question, leading to identification of hypotheses for testing.

Jeffers (1978) and Green (1979) treated experimental design and data analysis and management in detail in response to the difficulties of testing formal hypotheses with field data, and of providing statistical inference to assess impact. Green (1979), particularly, emphasized the selection of response variables that can be measured to evaluate the potential impacts and provided extensive guidance in this area. Comiskey and Brandt (1982) also were concerned with the need to identify appropriate data analysis techniques, including, in some cases, data management systems.

A major shortcoming of the existing guides to monitoring design is their dependence on in-house expertise. The resulting assumption is that these experts would be available and would be able to make a wide variety of value judgments about diverse subjects, such as the nature of functional linkages between two different trophic groups or the appropriateness of a particular variable or statistical or optimization technique. In practice, in-house expertise rarely covers the complete range of these needed skills. Even when extensive catalogs of elements to consider and open-ended sample impact matrices are presented (States et al., 1978), development of a viable set of hypotheses and appropriate test variables becomes an artful process at best. When little guidance information is presented, as, for example, in the selection of representative important species (RIS) advocated by the U.S. Environmental Protection Agency (1977 a,b), the process can be very cumbersome.

A further shortcoming exists in the complexity of the presentation format of many of the existing approaches. With the exception of the lists prepared by Jeffers (1978) and the efforts of Comiskey and Brandt (1982), the guides cited have used textual and diagrammatic formats that are cumbersome or confusing to use. There is frequently no obvious point of initiation, nor any means of transition between parts of the framework. Other shortcomings include: a lack of criteria to make decisions, a focus on short-term impacts, a lack of concern for cost-effectiveness, and a lack of consideration of alternative but attractive methods. An additional major drawback is that the utility of previous guides has not been demonstrated through application in actual case studies. As a result, most of these guides have not been implemented by aquatic monitoring program designers.

EPAM has been formulated to minimize these shortcomings encountered in the existing frameworks for monitoring program design. In the following chapter the basic components of EPAM are described in detail, and the following features of EPAM will be made apparent:

* Methods to present pertinent information and solicit opinions from experts by means of the Delphi technique in order to develop a consensus on pertinent aquatic monitoring issues where no objectively based or generally accepted opinion existed previously.

* Methods that can use both subjective and objective information to rank alternatives, given a multiobjective set of criteria.

* Improved statistical methods to design experiments that can examine both long- and short-term impacts, even when the measurements may be correlated in space or time.

* Data and methods to optimize monitoring programs for cost-effectiveness.

* A tiered approach that can be used by individuals with a broad range of knowledge and skill. The user can enter the framework at any point and elect either to conduct a tentative assessment, basing program design decisions on available information, or to perform more detailed studies prior to making decisions.

* Software that can be used interactively to record and manage all the information (whether subjective or objective) available for the design of a monitoring program.

CHAPTER 2

EPAM USER'S GUIDE

GENERAL EPAM FRAMEWORK

This chapter emphasizes the content of the interactive EPAM hierarchy (LEVELS 0 through 3), and presents other information pertinent to the EPAM structure and format. Appendix C.3 illustrates the interactive use of EPAM with actual data from the San Onofre Nuclear Generating Station.

As described in Chapter 1, EPAM contains four levels corresponding to the optimization of a MONITORING PROGRAM (LEVEL 0), design of a SAMPLING PROGRAM (LEVEL 1), generation of HYPOTHESES (LEVEL 2), and identification of IMPACT AGENTS and TARGETS (LEVEL 3). If the user has insufficient information at any given level, he or she is directed to the next level to develop the required information. In each level the user is guided with questions and prompts to define and select a set of CRITERIA that will be used to rank the subjects addressed by each level. For example, in LEVEL 0 the subjects are the SAMPLING PROGRAMS that are candidates for the MONITORING PROGRAM, and in LEVEL 3 the subjects are TARGETS and IMPACT AGENTS. The general procedure at all levels is to construct a table of subject characteristics, where the column headings of each table are the CRITERIA, and the row headings are the subjects. Each table entry should include indices of the quality or precision of the estimate of the characteristics, as well as the cost to increase the quality of the estimate.

Once the user completes a table, the data in each table become the input and basis for generation of pairwise comparisons of each subject with all other subjects in that table. The ranking procedure (termed CHOICE) requests pairwise weightings based on the data in each column (i.e., for each CRITERION). Whether the data are quantitative or qualitative, the following scale, as well as interim values, is used for ranking characteristics at all levels of EPAM:

1 = Subjects are demonstrated to be equal.

3 = One subject is thought to rank slightly higher.

5 = One subject is judged to rank higher.

7 = One subject is demonstrated to rank higher.

9 = One subject demonstrably dominates the other.

Values 2-5 represent subjective judgments in the pairwise comparisons, while 1 and 6-9 represent objectively demonstrated ratings. Appendix C.3 provides an example of the use of this ranking procedure.

Only three computational routines are used in the four levels of EPAM. At LEVEL 0, CHOICE is used to prioritize the SAMPLING PROGRAMS developed in LEVEL 1. In LEVEL 1 there are two other major computational routines, one that optimizes the frequencies and numbers of sampling locations and replicates, and another that analyzes sampling data for correlation of residuals. Classical sampling data analysis routines that compute means, variances, etc. have not been incorporated into EPAM since the literature review and surveys of utility practices indicated that extensive investment has been made in Statistical Analysis System (SAS) software with this capability. LEVELS 2 and 3 also use the CHOICE ranking scheme, but no other computational routines. Instead, LEVELS 2 and 3 involve an interactive narrative that prompts and queries the user in table building protocols.

LEVEL 0

LEVELS 0 and 1 are designed for users who have identified the major element of a MONITORING PROGRAM design. These levels permit such users to design and optimize individual SAMPLING efforts and to optimize an overall MONITORING PROGRAM incorporating a number of SAMPLING efforts. An algorithm labeled CHOICE is the basic component of LEVEL 0; it is also applied in several places at higher EPAM levels to rank sets of alternatives given multiobjective criteria. The output of LEVEL 0 is a monitoring program optimized according to the established CRITERIA.

A user entering LEVEL 0 of EPAM must have defined all the SAMPLING PROGRAMS required for the MONITORING PROGRAM. The table that will be constructed in LEVEL 0 will have as CRITERIA: (1) the cost of each SAMPLING PROGRAM, (2) the anticipated statistical power or other measure of effectiveness of the SAMPLING PROGRAM, and (3) any measure of the importance or need for the particular results it will yield. Representative sampling cost data have been compiled to assist the user in estimating such costs, if local data are lacking. The narrative in EPAM presents expert judgment of useful CRITERIA, as well as suggestions for characterizing and quantifying these characteristics for each SAMPLING PROGRAM. The user must provide a pairwise comparison of the relative importance of these CRITERIA to the MONITORING PROGRAM, and then rate each SAMPLING PROGRAM with respect to all others using each CRITERION independently. The CHOICE algorithm then will compute the priority of each SAMPLING PROGRAM and inform the user on the consistency of the pairwise comparisons. Chapter 4 develops the theoretical basis for this analysis.

LEVEL 1

LEVEL 1 is designed to assist individuals attempting to establish the components of a MONITORING PROGRAM (defining the individual SAMPLING PROGRAMS to be integrated into an overall monitoring program). Users of LEVEL 1 are assumed to have defined the HYPOTHESES to be evaluated and the VARIABLES to be measured. The user can employ the algorithms included in LEVEL 1 to produce a sampling design which minimizes cost for a given level of precision or, alternatively, maximizes precision for a given cost.

The user must have a set of HYPOTHESES to be tested and must have defined the VARIABLES that will be measured to test each hypothesis before LEVEL 1 can be exercised. Furthermore, the user must have an estimate of the cost to conduct each experiment and the variability of the data that will be obtained. The user can either conduct exploratory studies to improve these estimates or use existing information to determine whether the SAMPLING PROGRAMS are robust and insensitive to such estimates, or whether further refinement of the input data would be effective. If any of these data are inadequate, the user is directed to LEVEL 2 to develop the underlying HYPOTHESES more fully.

Data representing the variable to be measured are analyzed using the RESIDUALS algorithm to determine if the residuals are correlated in space or time. If no prior information exists, the value of an exploratory program will be weighted against the use of the first set of monitoring data for the residuals analysis. If the residuals are not independent, the design may have to be altered to improve the power of the test.

Two OPTIMIZATION algorithms are available and each has two options. The algorithms differ in that one assumes the samples are independent, while the other accepts the interdependence and corrects for it. The options are: (1) to maximize statistical effectiveness given a monetary budget, or (2) to minimize cost in order to obtain a given level of statistical power. After the user is prompted by EPAM to specify the algorithm and the option, the interactive prompts identify all necessary input data for the optimization. The theoretical basis for these methods is discussed in Chapter 4.

LEVEL 2

LEVELS 2 and 3 are designed for users who lack experience in comprehensive monitoring and sampling program design or the necessary knowledge of the circumstances at a given site. The various components of these levels consist of questions and prompts drawn from the collective knowledge of the project team and refined by the Delphi process. The data obtained are presented in a structured sequence designed to identify important IMPACTS, TARGETS, or HYPOTHESES for a specific site. An important function of LEVELS 2 and 3 is to provide a framework for recording data gathered and decisions reached during the analysis. New knowledge can be inserted as it develops, and its effect on the overall monitoring design can be evaluated at a small cost.

LEVEL 2 is designed to generate HYPOTHESES OF IMPACT. The user is requested to complete a LEVEL OF CERTAINTY MENU to evaluate the adequacy of information on IMPACT AGENTS and TARGETS of interest in order to use LEVEL 2. Furthermore, items on this list should be characterized by their distribution in time and space, and each TARGET should be further characterized by its sensitivity to each IMPACT AGENT. The user is referred to LEVEL 3 to develop missing data.

The next task consists of an eight-step definition of each HYPOTHESIS:

1. Identify the potential IMPACT AGENT.
2. Identify the TARGETS most likely to be affected by the IMPACT AGENT.
3. Define the spatial domain of the response.
4. Define the temporal domain of the response.
5. Define the severity of the response.
6. Select a VARIABLE to measure the response and define the cost, precision, and characteristics of such measurements.
7. Determine if other hypothesized changes could result as an indirect outcome of the subject hypothesis (COMPLEMENTARY HYPOTHESES).

8. Define ALTERNATIVE HYPOTHESES that could falsify the proposed hypothesis test.

The LEVEL 2 narrative leads the user through these steps and helps the user create a table of these data for each potential HYPOTHESIS. Information on the domains of IMPACTS and TARGETS forms the basis for defining the sampling pattern, while the characteristics of the variable to be measured provide the information to select the numbers of stations, frequencies, and numbers of replicates for the SAMPLING effort represented in each HYPOTHESIS test. These data are supplied by the user or developed with the assistance of the LEVEL 3 narrative.

Appendix C.3 contains an example which applies LEVEL 2 to develop HYPOTHESES for an aquatic monitoring program at San Onofre Nuclear Generating Station. Chapter 3 discusses the use of the Delphi panel and the results of the literature research used to create the LEVEL 2 framework.

LEVEL 3

LEVEL 3 is designed to assist the user when IMPACT AGENTS or TARGETS must be defined. In most cases, the informed user will not need to use this section of EPAM. The narrative in LEVEL 3 presents categories of IMPACT AGENTS and TARGETS commonly associated with different environments and different types of power generating facilities. The user is requested to examine his or her own situation and complete tables of IMPACT AGENTS and TARGETS with information on their characteristics. The CHOICE algorithm is used to rank the entries in these tables and provide input for the LEVEL 2 HYPOTHESIS GENERATION process. The user must estimate spatial and temporal domains for each IMPACT AGENT and each TARGET, as well as characterize the sensitivity of each TARGET to the associated IMPACT AGENT. If these data are incomplete or not available, the user must develop them. The user always will be guided to compare the improvement of an input value to the resulting improvement in the design of the MONITORING PROGRAM.

Appendix C.1 reprints a paper entitled, "A Conceptual Framework to Guide Aquatic Monitoring Program Design for Thermal Electric Power Plants," which was presented to the American Society for Testing and Materials Symposium on Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems. This paper discusses the development and content of LEVELS 2 and 3 in further detail.

Support Modules

The interactive version of EPAM contains support modules, such as HELP, GLOSSARY and TECHNIQUES to provide assistance in responding to queries at any level of EPAM. These texts are also published in the handbook format (see Appendix C.2). A user can either use the interactive computer format to complete the tables used by EPAM or a manual procedure, using the handbook as a reference guide; however, the computational algorithms such as CHOICE, RESIDUALS, and OPTIMIZATION, must be executed on a computer.

CHAPTER 3

THE DELPHI PROCESS IN EPAM DEVELOPMENT

This chapter reviews the results of the Delphi process as exercised in Phase II.

The Delphi technique is one of several communication methods developed during the last twenty-five years in the general field of technological forecasting. Initially developed by the Rand Corporation, its usage has expanded to include many contexts in which judgmental information is needed. The Delphi technique has been used successfully in the evaluation of educational issues, socio-cultural questions, environmental management options, and others (Bakus et al., 1982). Linstone and Turoff (1975) have characterized the Delphi technique as:

> ... a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem.

The Delphi technique was incorporated into this project to fulfill two objectives:

1. To test the use of an expert panel as a means of reaching a consensus about issues in the assessment of environmental impact, caused by electrical power generation, for which no objective criteria exist

2. To allow an evaluation and validation by an interdisciplinary group of experts of the proposed method for the design of aquatic monitoring programs

The conventional Delphi process is essentially a combination and extension of a polling and conference procedure. In this technique a small monitor team designs a questionnaire to address the range of issues to be considered by the Delphi process. The questionnaire is then sent to a larger group of expert respondents (the Delphi panel), each of whom submits answers to the monitor team. The monitor team reviews, and statistically summarizes the responses to this initial questionnaire. The monitor team then develops a second questionnaire, which includes both a reiteration of questions from the first questionnaire for which no clear consensus was evident, and new questions, which present new issues and options that were raised by the respondents during the first round. The expert respondents are given a chance to change their initial responses based upon the knowledge gained from the group responses and the new options presented.

This process of response and reiteration is usually repeated three or four times. The intermediate rounds in the Delphi process allow a formal exploration of areas of disagreement and an evaluation of the underlying reasons for these disagreements. This chance for exploration of dissenting or minority views represents one of the strengths of the Delphi process. In a more traditional conference or small group meeting, minority views frequently receive less than adequate consideration because of the overriding influence of dominant personalities, lack of time, or other logistical shortcomings.

The greatest potential weakness in the Delphi method lies in the ability of the monitor team to present correctly the developing consensus and dissenting views to the respondent group. Assuming that these views are represented properly, the Delphi method represents a unique technique which was felt to have potential utility for the field of environmental assessment. In consequence, the Delphi method was incorporated as a central feature in this project.

SELECTION OF THE DELPHI PANEL

In the absence of well established criteria for the selection of expert panels (Delbecq et al., 1975; Linstone and Turoff, 1975), the University of Washington researchers chose a straightforward selection procedure. In the first step, an initial list of 105 candidates was established from recommendations based on personal experience or from a review of the literature.

The second step of the selection process involved two levels of organization. In the first level, a set of six objective criteria was established as a basis for candidate selection. These six criteria were:

1. Past experience in at least two or more of ten specialty areas judged by the researchers to represent different subjects that would be considered during the development of the proposed research

2. Past experience in at least one of six water body types

3. Representation of each of the major geographic areas of North America

4. A reasonable balance of at least four professional environments

5. At least seven years of combined professional experience in the fields of ecology and environmental assessment of power generation effects or ten years in either field.

6. A high level of professional productivity

Appendix B.1 discusses the specific considerations exercised in applying these criteria.

The second level of selection evaluated each candidate so as to obtain a reasonable balance among panel members of the six criteria developed in level one. Each candidate received a Pre-Delphi Survey (PDS) that questioned the candidate about the six criteria and other issues. A copy of the PDS can be found in Appendix B.2.

The final panel composition required: (1) that all speciality areas be represented by at least two persons; (2) that all waterbody types and geographic areas be represented by at least three persons; (3) that panelists be approximately equally distributed among utilities, consulting firms, and academic institutions, with a smaller component from regulatory agencies; and (4) that all members meet criteria 5 and 6. It was decided that if this composition could not be met with the initial set of 105 candidates, more candidates would be identified and contacted. It was also decided that the minimum acceptable number of members on the Delphi panel was 20, although a membership of 50 or more was considered optimal.

IMPLEMENTATION OF THE DELPHI TECHNIQUE

Initial implementation of the Delphi technique required a three-stage effort:

1. All candidates (105 persons) received an introductory letter from the EPRI project manager (Dr. J.S. Mattice) explaining the

objectives of the research project, the rationale for invoking the Delphi technique, and the potential benefits to EPRI of their participation.

2. One week later, all candidates received an invitation letter from the University of Washington. This letter outlined the objectives of the project in greater detail, gave background information about the Delphi technique, explained the time schedule for invoking the technique, and requested that the candidate accept or reject the invitation by returning an enclosed, stamped postcard.

3. Following receipt of the consent postcard, candidates were sent the PDS and were asked to return completed PDS's (in enclosed, stamped envelopes) within a month of initial receipt.

Following receipt of the completed PDS's, the monitor team reviewed candidate credentials and finalized membership for the Delphi panel. The composition and formal credentials of the panel are presented in the Results section of this chapter.

The second step in implementing the Delphi technique involved the development of a set of questions by the monitor team and presentation of these questions to the Delphi panelists in questionnaire form. The response process was simplified by dividing the basic issues to be addressed into two questionnaires; a third questionnaire reiterated selected questions from the initial questionnaires. Questionnaire 1 (Q1) concentrated primarily on the conceptual basis of the proposed methodology, while Questionaire 2 (Q2) emphasized specific scientific issues raised in the methodology and an initial reiteration of questions from Q1. Questionnaire 3 (Q3), the final questionnaire, reiterated questions from both Q1 and Q2 and presented new options raised by the expert panel. All questionnaires were sent with return stamped envelopes. Copies of the three questionnaires are presented in Appendices B.3, B.4 and B.5. Figure 3.1 indicates the relationship between the question sequence and the development of EPAM.

Seventy-seven persons agreed to join the Delphi panel. All of these persons met the established criteria and received Q1 during the first week of April 1983. Returns were requested by April 30, 1983. A reminder letter was

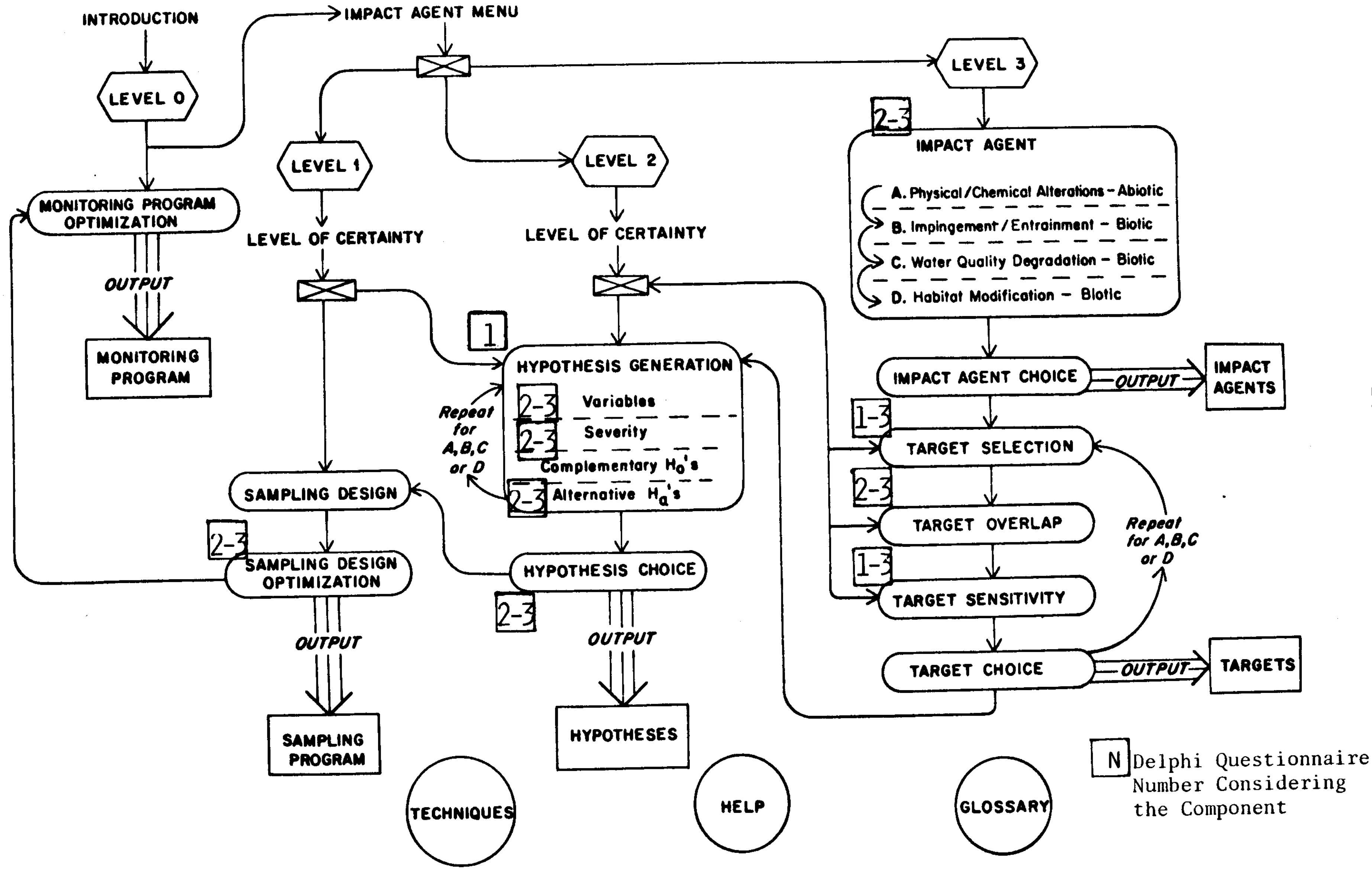

Figure 3.1. Coverage of EPAM Components in Delphi Questionnaires

mailed on May 2, 1983, to those who had not returned their questionnaire. A total of 65 questionnaires were received by mid-May, and further mailings were limited to this group. Q2 was mailed to the 65 panelists during the second week of May 1983. Panelists were asked to return it by May 31, 1983. A reminder letter was sent to those who had not returned the questionaire by June 8, 1983. Fifty-six panelists returned Q2 by mid-June. These results were used to develop the questions for Q3. Q3 was mailed to 62 panelists during the last week of June 1983. Panelists were to return the questionnaire by July 31, 1983. They were informed that they would receive a copy of the results of the full Delphi review later in the year, along with a survey in which they would evaluate their experience in participating on the panel. A reminder letter was sent to those who had not responded by August 7, 1983. Sixty-two persons returned the questionnaire.

RESULTS OF THE DELPHI REVIEW

Pre-Delphi Survey

The Pre-Delphi survey was concerned primarily with the biographical data of the invited panel members, their opinions of the Delphi process, and some general issues in monitoring program design. This section summarizes the responses. The data represent the responses of the 62 persons who continued as Delphi panel members through Questionnaire 3.

Appendix B.6 presents tabular data about the characteristics of the Delphi panel membership. In terms of areas of professional expertise, the majority of panelists reported experience and expertise with the ecological effects of electrical generation (82 percent), evaluation of monitoring programs (75 percent), and monitoring program design (73 percent). The majority of panelists reported experience with finfish, and nearly half had worked with benthic macroinvertebrates.

The panelists' experience in various water body types was fairly evenly distributed among estuaries, small lacustrine systems, and coastal environments, with a somewhat larger distribution of experience on large rivers and a smaller distibution on small riverine systems and the Great Lakes.

The geographic distribution of experience largely followed the concentrations of population and thermal electric power plants, with more than half of the members reporting experience in the Northeastern, Northcentral, and Southeastern United States or Eastern and Central Canada. The Southcentral United States was represented rather lightly on the panel (8 percent with experience in that area).

The panel membership was divided primarily among university, consulting and electrical utility affiliations, with a sizable component from government agencies. As stated earlier, six criteria were used to establish a balanced panel membership.

The PDS included a number of questions designed to measure the panelists' familiarity with and opinion of the Delphi process. They were generally unfamiliar with the process but most believed, based on their knowledge at the time, that it would be worthwhile to try it in the aquatic monitoring design area. A large majority (82 percent) believed the product of the proposed effort would be somewhat useful, useful or very useful. Thought has been given to implementing a small, specialized Delphi panel review via interactive computer mail during Phase III of the research. A majority of the respondents were familiar with interactive computers and were willing to participate in a review which used electronic mail.

The final section of the PDS questioned the panel about several general issues in aquatic monitoring. One issue concerned criteria for designation of abiotic or biotic characteristics (TARGETS) to be monitored. A consensus developed on the advisability of considering ecological (weighted mean score 1.51) and economic (2.02) importance, while tradition scored rather low (3.60).[1] Panelists suggested twelve additional criteria; thus, the question was repeated in Q1. A second issue explored in the PDS was the quality of certain aspects of past monitoring programs. Table 3.1 summarizes responses. None of the seven aspects listed rated higher than 2.77. There was almost complete dissatisfaction with the ability of monitoring results to establish significance and hence infer the existence of power plant effects. Because panelists brought up ten additional aspects of monitoring programs, this question also was repeated in Q1.

---

[1] A 1 to 5 scale was provided for Delphi responses, with 1 signifying a high rating and 5 a low rating.

Table 3.1. Delphi Panelists' Opinions of Monitoring Program Design

Aspect	Weighted Mean Score[a]
Design of Statistical Aspects	3.53
Establishment of Goals and Objectives	3.56
Implementation of Goals and Objectives	3.47
Execution of Program Design	2.77
Availability of Technology	2.85
Establishment of Significance of Results	4.37
Ability to Distinguish Change	4.47

[a] 1 = Satisfactory, 5 = Unsatisfactory.

Questionnaire 1

The initial questions in Q1 were reiterations of the two points for which PDS responses produced additional considerations. Table 3.2 presents Q1 responses to the question concerning criteria for designating biotic characteristics to be monitored (TARGETS), and lists those criteria for which a consensus developed (weighted mean score $\leq$ approximately 2.5). Table 3.2 also indicates the EPAM routine that addresses each point. All of the consensus issues are represented in the EPAM structure. The TARGET SELECTION subroutine, particularly, is strongly based in the Delphi panel review of monitoring TARGETS. These criteria were re-evaluated in Q2 and the final ranking is included in Q2 results, below.

Table 3.3 provides the responses to the Q1 question reiterating the degree of satisfaction with various aspects of monitoring program design, along with the consideration of these points in EPAM. The major areas of dissatisfaction, the ability to detect change and establish the significance of results, are addressed primarily in the HYPOTHESIS GENERATION subroutine. This subroutine guides the development of primary HYPOTHESES OF IMPACT, which incorporate the IMPACT AGENT, TARGET, and approach to testing. Expected severity of the effect is a major criterion for the adoption of a HYPOTHESIS.

HYPOTHESIS GENERATION also assists the user in developing ALTERNATIVE HYPOTHESES (H_a's) associated with conditions extraneous to the power plant, which may falsify the PRIMARY HYPOTHESIS, i.e., lead to a decision to accept or reject the PRIMARY HYPOTHESIS based on an observed response created by a factor or factors external to the power plant. Strategies to assess long-term change are considered at various points in EPAM, including TARGET SENSITIVITY, HYPOTHESIS GENERATION, SAMPLING DESIGN, and OPTIMIZATION.

The remainder of Q1 consisted of questions concerning the structure of monitoring programs generally and the overall structure of EPAM. Figure 3.1, page 24, illustrates the correspondence between this questionnaire and the EPAM layout. Discussion of Q1 results will center on four topics: (1) general monitoring program structure, (2) hypotheses of impact, (3) criteria for monitoring program optimization, and (4) potential for long-term effects.

Table 3.2. Delphi Ranking of Criteria for Target Selection

Criterion	Weighted Mean Score[a]	Consideration in EPAM
Ecological Importance	1.44	TARGET SELECTION
Economic Importance	1.95	TARGET SELECTION
Probability of Exposure*	2.11	TARGET OVERLAP
Duration of Exposure*	2.11	TARGET SENSITIVITY
Dominant Role in Community Metabolism*[b]	2.41	TARGET SELECTION
Ease and Accuracy of Quantification*	2.55	VARIABLES (HYPOTHESIS GENERATION)
Indicator Species*	2.58	VARIABLES (HYPOTHESIS GENERATION)
Designation By Law	2.58	TARGET SELECTION
Susceptibility to Change	2.58	TARGET SENSITIVITY

*Suggested by Delphi Panel

[a] 1 = Most Important, 5 = Least Important.

[b] Incorporated in definition of ecological importance in TARGET SELECTION.

Table 3.3. Delphi Panelists' Satisfaction with Aquatic Monitoring Programs

Aspect	Weighted Mean Score[a]	Consideration in EPAM
Detect change relative to other agents and natural variation	4.58	Falsification routine (Holling, 1978) (ALTERNATIVE HYPOTHESES) Criteria for PRIMARY HYPOTHESES
Establish significance of results (especially long-term)	4.35	Explicit consideration of long-term stategies and severity of potential effect
Establish goals and objectives	3.58	Identification of HYPOTHESES
Apply statistical design	3.51	Identification of statistical design (including assumptions, data needs, etc.)
Implement goals and objectives	3.42	Responsibility of monitoring team
Access technology	2.8	TECHNIQUES subroutine
Execute program design	2.75	Responsibility of monitoring team

[a] 1 = Satisfactory, 5 = Unsatisfactory.

Regarding general monitoring program structure, the panel reached a consensus on the following points:

1. Monitoring programs should be founded on quantifiable and scientifically valid hypotheses.

2. Methods for design of monitoring programs should be standardized.

3. Methods for collection of field samples should be standardized (or intercalibrated).

4. Alternative hypotheses should be evaluated.

5. Cost and error criteria should be used for choosing which hypotheses to evaluate.

6. Data not related to a formal hypothesis should be collected in some cases.

The final point had not been presented in the questionnaire but was raised by the panel. Examples of such cases given in the responses were: (1) the occurrence of unexpected conditions that may contribute to understanding a non-plant condition, and (2) the need to collect information for public relations purposes.

Questions relative to hypotheses of impact considered criteria for deciding to test primary hypotheses (H_0's) and attributes of a well-formulated hypothesis. In order of preference, the following criteria for selecting H_0's for testing were recommended by the panel.

1. Availability of good measuring techniques
2. Suitability of controls
3. Low variance in the measured element
4. Total cost of the test

Concerning the attributes of a hypothesis, a consensus was reached that it should incorporate the following elements:

1. Impact agent (effect mechanism)
2. Target (component to be observed for change)
3. Physical overlap
4. Sensitivity to impact agent
5. Quantifiable variable
6. Spatial and temporal dimensions of hypothesized change

The final element received the least support. Overall, the recommendations constitute strong support for the approach taken in EPAM.

With reference to optimization of monitoring program design, the consensus criteria in order of preference were as follows:

1. Efficiency in the sampling network
2. Priority of the primary hypothesis (based on established criteria)
3. Reduction of error (sampling variance)
4. Cost

EPAM incorporates these criteria in various ways. Efficiency is of central concern in the OPTIMIZATION subroutine, while CHOICE provides a mechanism for ranking HYPOTHESES according to an established priority. The structure provides an accounting system for the variance and cost associated with various options so that these can be considered in the final monitoring design.

Gauging the potential for long-term change was an area of considerable controversy in the panel's deliberations. A consensus developed on three situations likely to produce long-term effects: (1) presence of extremely vulnerable target(s), (2) extensive impact agent/target overlap, and (3) operation of an extremely intense impact agent. The panel failed to reach agreement on one issue of great interest, however: whether multiple power plants and other human activity, which introduce similar impact agents in the vicinity, create a high potential for long-term change. A number of panelists commented that site-specific considerations are too important to permit arriving at a general conclusion in these cases. These points were raised again in the subsequent questionnaires and will be discussed further in connection with their results.

Questionnaire 2

Questionnaire 2 consisted of two parts: (1) a reiteration of certain questions from Q1, and (2) questions about specific subroutines in EPAM. Repeated questions dealt primarily with the general monitoring program structure and the potential for long-term effects, with several other miscellaneous issues being represented.

There was a fairly strong consensus (weighted mean score 2.23) that collection of data other than that defined by a formal hypothesis should be considered. Relative to the purposes of data collection, the panel took a rather pragmatic view that documenting and explaining conditions for a specific purpose should take precedence over enhancing regional data bases. One question concerned means of making results of hypothesis testing believable, in addition to evaluating alternative hypotheses. Panelists ranked the various alternative approaches presented in the following order:

1. Correlation of observed biological change with physical/chemical changes and/or known tolerance limits
2. Laboratory testing
3. Replication at other sites
4. Demonstration that observed changes at one trophic level have indirect effects at another trophic level
5. Simulation model development

The majority of the panelists responding to Q2 (63 percent) believed that the issue of long-term change could only be addressed at the site-specific level. About 10 percent of the survey group was persuaded that a concentration of power plants or other sources is not likely to increase the potential for long-term effects. The remainder thought separation distance may be of interest, with the majority believing multiple sources must be within 10 miles (in a range of choices 10-200 miles) to increase the potential of long-term change.

In the portion of Q2 dealing with development of EPAM subroutines, panelists suggested options in addition to those presented in the questionnaire in a number of instances. In most of these cases, the questions were repeated in Q2. In other cases a consensus was not apparent after one round of consideration, and the questions involved also were repeated. This discussion will be confined to those areas which were not reiterated in the subsequent questionnaire.

Points of agreement arising in the respective sections included:

Section A (General Development Questions)

1. If the spatial extent of a hypothesized effect is equal to the local range of a target, it may be critical to the monitoring program. If it is greater than the target's local range, it should definitely be monitored; if it is less than the target's local range, the hypothesized effect should be considered if other criteria (e.g., sensitivity) indicate a high probability of impact.
2. Eliminating a primary hypothesis from a monitoring program should be considered if the degree of confidence in the underlying data base is low, at least one H_a with a high probability of occurring exists, or a total of more than three H_a's exists.

Section B (Target Choice)

1. The public interest criteria for target selection, in order of importance are those (1) designated by law (weighted mean score 1.54), (2) designated as representative important species, (3) designated as rare or endangered species, (4) considered to have other unique societal value, (5) noted to have a history of monitoring (weighted mean score 2.65).
2. The economic value of targets, in order of importance, is: (1) primary commercial value (weighted mean score 1.44), (2) primary sport value, (3) major food source to commercial or sport fishery, (4) secondary commercial value, (5) secondary sport value (weighted mean score 2.72).
3. Data bases likely to be most useful in judging relative vulnerability to an impact agent are onsite studies and historical records pertinent to a given target. Published tolerance data were judged to be rather marginal in value. Unfortunately, the best data bases are least available.
4. Definitions of overlap and relative vulnerability will help to prioritize monitoring program elements to a greater degree than otherwise possible.

Section C (Assessing Stock Distribution)

1. Physical overlap in time and space between an organism and an impact agent should and can be measured cost-effectively.
2. When overall geographic overlap between an impact agent and a target is low, but the local population spends considerable time in the affected zone, the target should be monitored. Support existed for expressing geographic overlap in a probabilistic form.

Sections D and E (Impingement/Entrainment and Physical/Chemical Problems)

1. Very few panelists knew of accepted models capable of deriving a probability density function of the intake field, whereas 38 percent were quite familiar with such models for discharge plumes.

Section F (Statistical Characteristics)

1. The panel rated the following statistical tests as being useful in past data analysis efforts: analysis of variance (weighted mean score 2.23), time series analysis (2.52), non-parametric tests (2.55), cluster analysis (3.05), and factor analysis (3.13). Respondents listed nine additional techniques which they had used successfully.
2. Most respondents have considered spatial and temporal correlation in statistical analysis.
3. The most successful technique reported for handling zeros in data sets is aggregation by space (weighted mean score 2.25). Other means of aggregation were considered to be less useful (scoring 2.47 - 3.00).
4. Space-time tradeoff for station allocation in assessing long-term change is determined most often by personal judgment and less often by an optimization model, although several respondents listed other approaches.
5. A number of panelists was aware of adequate data sets to assess the space-time tradeoff. These panelists were contacted as part of the Phase III effort.

Section G (Sampling Costs)

1. Fixed costs represent 25-50 percent of total sampling costs in the case of fish and 10-25 percent for benthic invertebrates, water quality and plankton.
2. The incremental cost of a sample at a different station, compared to the same station, is 50-300 percent greater.

   Note: Only 11 of the 56 respondents rated their experience in this area at a high level (1 or 2 on a 1 to 5 scale).

Questionnaire 3

Part I. Reiteration of questions from Questionnaire 1

The questions in Part 1 of Q3 re-evaluated four areas of concern that had been identified in Q1. These were: (1) alternative methods for evaluating hypotheses, (2) an operational definition of ecological value for EPAM, (3) the characteristics of sites and types of ecological changes associated with long-term effects and (4) the characteristics of sites demonstrating resiliency to adversity. The results of these evaluations are presented below.

(1) The panel suggested five possible alternatives for hypothesis evaluation in Q1. These alternatives were rated in Q2 and again in Q3. The re-iteration did not change the panel consensus. The panel felt that correlation of observed biological change with physical/chemical changes and/or known tolerance limits was the best approach among the various alternatives. Laboratory testing and replication of observed results at other sites ranked equally, although slightly lower than the use of correlation analysis. Panel members were neutral about the utility of demonstrating that indirect effects exist (at a different trophic level) and the utility of simulation models as a means of making primary hypotheses believable.

(2) A strong consensus supported an operational definition of the ecological value of an ecosystem component. In Q2, 72 percent of the panel members felt that this value was equally dependent upon the general public's opinion (including social, political or economic definitions of value or function) and the scientific

opinion (including the more traditional ecological definition of ecosystem structure and function). In Q3, 96 percent of all respondents agreed that the inclusion of both sets of values was appropriate. The operational definition of ecological value used in EPAM will include consideration of both social and scientific definitions of value.

(3) The questions of how to characterize sites that have an increased potential for long-term ecological change and how to predict which types of ecological change have the potential to continue for long time periods (defined herein as being equal to or greater than plant life) were considered in all three questionnaires. In Q1 and Q2 a great deal of confusion arose as a result of poorly defined nomenclature, particularly the failure of the University of Washington team to distinquish among long-term, cumulative, and severe ecological change, and as a result of the relative paucity of data that describe the characteristics of long-term ecological effects. Some of the confusion was alleviated in Q3 by the use of a clearcut definition, and the Delphi panel reached a consensus on several issues (see below). Nevertheless, because precise characterization of long-term effects is necessary before appropriate monitoring programs can be designed, further questions on this subject will be considered in the Delphi round to be conducted during Phase III.

The Delphi panel reached a strong consensus that it is difficult to observe long-term effects in the absence of preoperational data regardless of other characteristics of the effects. This result suggests that baseline data must be collected if the potential for long-term effects exists at a given site. The panel also felt that long-term effects were more likely to appear some time after the plant starts operating than immediately after operation begins. The panel was evenly divided about the site-specificity of long-term effects. About half of the panel felt a priori predictions could not be made because of this problem, while the other members felt some predictions were possible.

Specific conditions likely to coincide with an increased potential for long-term effects were agreed upon. Eighty-three percent of the panel believed that the presence of multiple power plants or other industrial water users on the same water body increased the potential for long-term effects. The panel felt that if conditions at a new site were similar to conditions at an existing site that had experienced long-term effects, then the new site had a high potential for long-term change. Changes in the basic habitat structure of the ecosystem and in the reproductive strategy or growth habits of resident organisms were also judged as being important diagnostic indicators of an increased potential for long-term change.

The panel was divided on the issue of defining an allowable level of change (i.e., that which could be tolerated without resulting in long-term change). Fifty-three percent felt that less than a 10 percent change in any chemical or physical attribute could be tolerated, while 47 percent felt that consideration of this type of criterion was not useful.

(4) The panel agreed that certain target or site characteristics would be likely to decrease the potential for long-term effects or, in other words, created conditions in which either a given target or the site itself was resilient to adverse effects. These were, in order of importance:

a. The affected population is ubiquitous in distribution.
b. The affected population has broadly based environmental requirements for reproduction.
c. The affected population has exhibited resiliency in historical studies.
d. Refuge areas are available near the affected area.

The panel did not believe that the option of artificial replacement (e.g., hatchery, planting, etc.) was a viable means of decreasing the potential for long-term effects.

Part II. Development of Subroutines for EPAM

The remaining sections of Q3 were a re-iteration of questions about the development of EPAM and its various subroutines. Since this section of the questionnaire considered so many different issues, the majority of the results are presented in list form. Choices that received high consensus scores (usually greater than 2.5) were considered to represent viable alternatives for inclusion in EPAM. Any choice that received more than 10 percent of its total score at the opposite extreme (either positive or negative) from the mean score are noted as representing an issue for which dissenting opinions still exist. All such choices will be evaluated in detail in Phase III (case study evaluation), regardless of the numerical value of the mean score. If the minority opinions refer to important site-specific conditions, it is likely that their utility will become evident in one or more of the case study examples. These non-consensus issues will be included in EPAM as deemed appropriate following Phase III evaluation.

General Development Questions (Section A)

The five IMPACT AGENT classes defined in EPAM (impingement, entrainment, thermal loading, chemical change and physical change) remain as originally defined. The panel felt that all other agents (of which ten were suggested) were subclasses of one of the original five or represented special effects that were likely to exist at few sites, and/or for brief time periods (e.g., during construction phase). The panel did note that the redistribution of either organisms or water mass from one water body to another represented a special case of physical change that had a large potential for adverse effects. The consensus opinion was that the five major IMPACT AGENTS should guide the development of the operational monitoring program (i.e., the program that will be used during the period of plant life), and that all other suggestions should be considered within the context of the original five.

The panel was traditional in its definition of what constitutes an appropriate TARGET (i.e., an ecosystem component likely to change as a result of power plant operation or design). Only six choices (out of fifteen, see Appendix B.5) were given high scores by all panel members. The following six target types, in order or preference, will be the primary TARGETS in EPAM.

1. A distinct species (all life stages)
2. A distinct species (one life stage)
3. Measure of a physical phenomenon (e.g., temperature, discharge)
4. Concentration or mass of a given chemical (e.g., copper)
5. Population indicator (e.g., fecundity, change in population dynamics)
6. A taxonomically defined group (e.g., phytoplankton)

Two other choices received low mean scores but more than 10 percent of the panel felt they were very important. These two types of data, which provide surrogate information about the ecological condition of some of the target types listed above, will be considered as potential target categories in the case study evaluations. These choices were:

1. Records of social use of the area, such as swimming, boating, or fishing, that serve to document a particular habitat quality or the presence of particular biota (such records may be available from local or state park departments, state fish and game agencies, etc.)
2. Records that document the type and/or quantity of commercial or sport harvest removed from the area of interest.

The questionnaire also presented a number of potential target selections representing trophic level or community- or ecosystem-wide activity. These choices all received little support. Because the monitor team can conceive of situations in which such targets may be of importance, the Phase III Delphi round will reconsider the issue.

The Delphi panel agreed upon acceptable criteria for choosing test variables. The criteria to be incorporated in EPAM include the four original criteria suggested by the University of Washington team and six suggested by the Delphi panel. They are, in order of importance:

1. Potential error of the measurement
2. Utility of the variable for meeting the study objectives (Delphi suggestion)
3. Ability to interpret the variable (Delphi suggestion)
4. Importance of the variable in the local system (Delphi suggestion)
5.* Availability of comparative data (Delphi suggestion)

6.* Cost of measurement

7.* Technical requirements of measurement

8. Utility as judged by historical precedent

9. High quality control (Delphi suggestion)

10. Probability that the variable will change (Delphi suggestion)

* Criteria 5-7 ranked equal in importance.

Two other criteria, suggested by the Delphi panel, received low mean scores but were rated as very important by 11 and 15 percent of the panel, respectively. These criteria, theoretical significance of the variable and the potential that the variable measures a phenomenon with a high potential for changing something with social value (e.g., weight of commercial catch, availability of swimming beach, etc.) were judged by the monitor team to be subsets of criteria 4 and 8. They will, therefore, not be included as separate criteria.

A strong consensus supported all suggested criteria for judging the importance of the spatial and temporal extent of the hypothesized effect. These criteria can be found in Appendix B.5. In both instances the most important criterion was judged to be the extent to which the hypothesized effect overlapped with critical habitats (spatial) and critical life stages (temporal).

The panel felt that evaluation of a hypothesis having a strong likelihood of identifying an ecological change should not be included automatically in the monitoring program. They believed that the existence of alternative explanations (competing causes of change), measurement problems, or other difficulties might justify exclusion of such a hypothesis. This opinion lends strong support to the inclusion of the HYPOTHESIS CHOICE subroutine in EPAM. HYPOTHESIS CHOICE allows a users to choose from all suggested HYPOTHESES those which have the greatest potential for occurring and, at the same time, can be measured effectively.

EPAM Subroutine (Sections B through G)

In Section B of the questionnaire panelists were asked to consider the concept of defining the ecological value of a target by assigning numerical values to different functional roles. Ninety-three percent of the respondents thought that such an assignment was possible. The consensus scores for the four categories of ecological value that were included in this question are:

1. Target provides more that 20 percent of the total biomass at a given trophic level
2. Target provides a primary food supply (more than 30 percent) to an organism with economic or social value
3. Target provides a secondary food supply (more than 20 percent) to an organism with economic or social value
4. Target is a primary producer or primary producing group which provides more than 20 percent of the food base for primary consumption.

Other categories that may have potential for inclusion in the ecological value definition will be considered in the next Delphi round. Example candidates include habitat formers (e.g., kelp beds) and major predators.

Section C asked panelists to consider the concept of overlap between targets and impact agents in greater detail. The panel was supportive of the concept of geographic overlap and some measure of its variability. However, panelists suggested that the extent of overlap between an impact agent and the more sensitive life stages and habitat spaces of a target should also be considered explicitly in defining the amount of overlap. The panel also felt that overlap into refuge areas should be evaluated. Such measures will be included in the TARGET OVERLAP subroutine.

Section D asked panelists to consider various problems caused by impingement and entrainment. Two questions asked panelists to evaluate the kinds of conditions that would be likely to increase the severity of these problems. Eight conditions were considered to be important and are listed below in order of importance. The conditions suggested by the Delphi panel are noted.

1. Small fish abundant in the intake area
2. Intake design provides no potential for egress, no return devices, etc. (Delphi suggestion)
3. Intake location is coincident with critical habitat space (Delphi suggestion)

4. Shoaling fish present in intake area
5. Targets have physiological or behavioral preference for being in intake area (Delphi suggestion)
6. Ichthyoplankton, zooplankton, and phytoplankton are abundant in intake area
7. Plant operation coincides with natural catastrophic and/or other anthropogenically induced stresses (Delphi suggestion)
8. Screen mesh size larger than the size of organisms in the intake area (entrainment consideration) (Delphi suggestion)

One further condition, the presence of organisms that are very active in terms of daily activity, was not judged to be very important by the panel as a whole but was felt to be extremely important by 10 percent of the respondents. It will be considered with the other eight conditions in evaluating the importance of impingement and entrainment in the initial consideration of potential impact agents.

Panelists were asked to consider several other concepts in this same context. These included consideration of the relative dilution of cooling water in the water body, the intake velocity, and the presence of multiple plants. A prominent minority (25 percent of respondents) believed that the dilution volume was less important than the rate of dilution. The others felt that anything greater than 12 percent (volume of cooling water to volume of the water body) would increase problems. A majority of respondents (73 percent) believed that an intake velocity greater than 1 ft/sec would exacerbate impingement/entrainment problems, while 20 percent felt that the critical intake velocity was species-dependent. Seven percent of respondents argued that such a concept has not been useful in the past and should not be considered. The panel was split concerning the importance of multiple plants. Forty percent felt that no generalizations could be made, while almost 60 percent thought that the presence of one or more additional plants within a 25-mile radius would increase the potential for impingement/entrainment problems. Because of the lack of consensus on these three issues, all three will be evaluated with the case study data sets before a decision about inclusion in EPAM is made. Further evaluation by the Delphi panel will also be made during Phase III.

A strong consensus supported the statement that reproductive rate is important to consider when evaluating the potential for population declines following entrainment. However the panel was unable to specify what reproductive rates would result in population declines. This information will be pursued further by the monitor team at the University of Washington and in the next Delphi round prior to inclusion in EPAM.

Section E evaluated problems due to chemical or physical alterations. The panel agreed on several conditions that would be likely to increase the severity of such problems. These are noted below in order of importance. The conditions suggested by the Delphi panel for inclusion are so designated.

1. Biota in area are already close to tolerance levels and/or existing water quality is poor (Delphi suggestion).
2. Targets are immobile.
3. Flushing rate or mixing in discharge area is low (Delphi suggestion).
4. Organisms have reduced free mobility (e.g., juvenile fish).
5. Discharge concentrates in preferred depth strata (Delphi suggestion).
6. Organisms are attracted to discharge area as a result of habitat or physiological preference.
7. Bioaccumulation occurs in sediments (Delphi suggestion).

The panel also agreed that at least crude estimates of the intake or discharge envelope (e.g., uniform distributions) would be useful in initial estimates of impact.

The panel felt that predictions of thermal problems were most dependent upon the ambient temperature of the waterbody in question and the tolerances of the targets under consideration. Given this belief, they also reached a consensus that temporal duration and spatial extent of the maximum temperature was much more useful in predicting thermal problems than either a consideration of maximum temperature alone or the concept of a temperature change greater than or equal to 10 degrees C.

The questions in Section G about sampling costs reconfirmed that this issue is difficult or sensitive for many of the panelists to discuss in precise terms. As shown in Q2, only about 25 percent of the panel felt proficient in analyzing costs. While no one panelist's responsibilities explicitly exclude cost considerations, it appears that many organizations have specific persons assigned for cost evaluations. In addition, many members of the panel reported expertise in one or two specialty areas but claimed that the type of information requested (all trophic levels) was too comprehensive. Several members make "first-cut" cost estimates in the process of program design but do not make the final decisions. Other panel members are not directly involved in field monitoring programs and hence do not have relevant numbers.

No consensus was reached on identifying the relative cost of a replicate versus an independent sample for fish, benthic invertebrates, water quality or plankton. Panelists reported that fish and water quality replicates represented anywhere from 10 to 75 percent of the cost of an independent sample. There was no consensus within this range. The cost of a replicate for benthic invertebrate samples was reported to be between 25 and 75 percent, with 60 percent of the panelists reporting 75 percent. Plankton replicate costs were reported to range between 10 and 75 percent, again with 60 percent of the panel reporting 75 percent.

CONCLUSIONS OF THE DELPHI REVIEW

The Delphi process has performed a valuable role in providing a widespread (in terms of disciplines) basis of support for the philosophical stance and underlying structure of EPAM. The panel also helped to define, and in some cases identify, many of the detailed points that needed to be considered in the full development of the model. A more secondary, but no less useful, function of the panel has been as a source of various data necessary for the development of the statistical and optimization routines that are an integral part of EPAM. Panel members have also been cooperative in agreeing to participate in the Phase III case study evaluations.

The Delphi results strongly support the identification and hypothesis testing routines that are a central element of EPAM. The panel agreed that initial identification of impact agents is mandatory prior to the development of testable hypotheses. They also believed that the reduction of program objectives into formal hypotheses provides a strong basis for the development of a sound monitoring program. The incorporation of falsification routines, which identify alternative explanations for a hypothesized change and judge primary hypotheses on the basis of their relative freedom from alternative explanations, was strongly supported by the Delphi panel. This hypothesis testing structure will help to effect improvement in one area identified by the Delphi panel as unsatisfactory in past monitoring programs: the ability of a program to distinquish change due to power generation from other sources of variability (see Q1 results).

Another area identified by the Delphi panel as inadequate in past monitoring programs was the ability of a program to establish whether observed changes have ecological significance. The emphasis of EPAM on the development of hypotheses that have an increased potential to identify long-term change, and that represent extensive impact agent/target overlap in time and space (increased severity), will help to alleviate this problem. The Delphi panel also was instrumental in defining the kinds of conditions and ecosystem components that have an increased potential for change in general and for change likely to continue for a long time, or, alternatively, conditions and ecosystem components that are likely to be resilient to adversity. The panel will consider these questions again in the Phase III Delphi round.

Several of the subroutines in LEVELS 2 and 3 require a user to identify different components of the monitoring program and then rank these aspects on the basis of established criteria. As shown in the Results section for Q3 above, the Delphi panel contributed to the initial identification of program components (e.g., what is an appropriate target?) and to the definition of appropriate choice criteria (e.g., how should targets be selected?). As a result of the scoring system used in the Delphi questionnaire (1-5), default value ranks for the various criteria also were identified.

Regarding application of the Delphi technique itself, several conclusions are warranted. The relatively large size of the panel used in this application (a final membership of 62 persons) may not be necessary if the diversity of expertise can be represented in a smaller panel. In Q1 areas of consensus were apparent and did not change after the results of 15 to 20 panelists were collated. In Q3, 60 of the questionnaires were scored in two randomly chosen groups of 30 each. Weighted mean scores and importance rankings were not different between the two groups. It should also be noted that there was no apparent pattern or disciplinary bias associated with the persons who chose not to continue in the Delphi panel following initial agreement.

It appears that a strong consensus, once formed, is not shaken even when additional options occur in a repeated question. Responses to the two PDS questions repeated in Q1 were markedly similar, even though the reiteration presented many more options. The rating of alternatives for hypothesis evaluation did not change between Q2 and Q3. In some instances however, a reduced number of options led to a stronger consensus in a second iteration. For example, in Q2, 72 percent of the panel agreed to a given definition of ecological value. In Q3, the other options were removed and panel members were asked if they agreed with the majority opinion. Ninety-six percent of all members did.

Cogent definition of the question(s) being asked is mandatory for successful application of the Delphi technique. This is particularly true if the issue is complex or has no widely accepted attributes. The confusion that arose in discussion of long-term change in Q1 and Q2 was largely alleviated by the use of more clearly defined nomenclature in Q3. Perhaps related to this problem is the finding that the Delphi technique was most useful in this application for considering issues that had been identified by the monitor team before the first questionnaire was prepared.

Implementation of the Delphi process is time-consuming. The development of questionnaires and the processing of responses can require up to a person-month, depending on the length and complexity of the issues to be addressed. Clearly a small panel considering only one or two issues would be more efficient in terms of implementation. The role of the monitor team appears to be critical regardless of the size of the panel. The most important functions of the monitor team are in defining the issues to be considered and in analyzing the consensus opinion from the diversity of responses. The actual logistics of the application could be accomplished by any responsible person.

The Delphi process represents a very useful method for increasing the breadth and validity of coverage. In the current application, options missing from the original questionnaires but suggested by the panelists became important elements in EPAM. The basic structure of EPAM and functions of the various subroutines did not change as a result of the Delphi process, although this possibility certainly existed and could occur in another application. The Delphi panel was in strong support of the initial premise and format. While this may have been a self-selecting result, the diversity of the panel and the fact that panel members wrote extensive comments about areas on which they disagreed argue against this conclusion. Furthermore, many of the invitees who did not join the panel were unable to participate due to schedule conflicts or excessive work loads, and not because of lack of interest or sympathy.

In conclusion the Delphi technique appears to be a method that has potential utility for the process of monitoring program design, either in the development of a standardized set of general methods as has been demonstrated here, or in the development of a specific program at one site by a group of local experts working in conjunction with utility staff.

CHAPTER 4

COMPUTATIONAL ALGORITHMS EMPLOYED IN EPAM

This chapter details the computational algorithms employed primarily in LEVELS 0 and 1. LEVELS 2 and 3 are also discussed in the case of CHOICE.

METHODS OF ALTERNATIVE EVALUATION

The design of an ecological monitoring program requires the careful consideration of multiple objectives and a variety of physical, economic, and temporal constraints. The individuals charged with designing the monitoring program may be required to select from a set of conflicting, competing activities. If the monitoring network is to address multiple objectives, the decision maker must choose between activities that contribute to these objectives and select the activities that, when viewed over all objectives and constraints, represent a best compromise solution to the problem.

This section addresses two issues: (1) choosing between conflicting activities and selecting those that best achieve a given set of objectives, and (2) applying a solution methodology to aquatic biological monitoring. The methodology presented is based upon the use of pairwise comparisons between objectives and activities and subsequent application of an eigenvalue analysis. This procedure is shown to result in computationally efficient solutions and to generate useful, easily interpreted results. A computerized, interactive program that incorporates this technique into EPAM (denoted as the subroutine CHOICE) is described. Examples of the procedure are also presented.

Multiple Solution Techniques

The application of multiobjective analysis has grown considerably in recent years in both the public and private sector (Keeney and Raiffa, 1976). A majority of the work in this field has been quantitative in nature, and has emphasized mathematical solution techniques and optimization procedures. Cohon (1978) has characterized these procedures as techniques that generate noninferior solutions, incorporate preferences, and involve multiple decision makers. The techniques differ not only in mathematical form but also in the extent to which the analyst and decision maker interact in selecting among alternative solutions. Four of the more frequently used approaches include the weighting method (Goicoechea, et al., 1982), the surrogate worth trade-off method (Haimes and Hall, 1974), utility approaches (Keeney and Raiffa, 1976), and goal programming (Ignizio, 1981).

Two important objectives in the design of monitoring programs are minimizing the network's cost and maximizing the network's effectiveness or ability to detect change if it occurs. These conflicting objectives are exemplified by the following situations:

1. Given a hypothesis to be tested, select a monitoring design from among the possible allocations of sampling effort in time, space, and replicates to provide the best test of the hypothesis.
2. Given a set of hypotheses of interest, each concerned with a different target in the site area, list them in order from the most important to the least important.
3. Given a set of sampling programs to test hypotheses, select those that best achieve a given set of overall monitoring objectives.

Expert Involvement

Because multiobjective analysis involves choices between competing activities, a well informed decision-maker should be involved. An important requirement in solving these problems is a detailed understanding of the problem, its inherent conflicts, and the relative importance of its potential outcomes. As in the case of monitoring design, the complexity of the problem makes it unlikely that anyone other than an experienced decision maker will understand or appreciate the problem's nuances. The experience gained by an individual responsible for monitoring program design will be influenced not only by the design in general but by specific knowledge involving the location and environmental issues of importance.

Incorporating the decision-maker into the selection of the best compromise solution can, however, create difficulties. Decision-makers may not respond consistently at all times. In developing the relationships and comparisons that must be made between objectives, constraints, and criteria, the decision-maker should be rational, consistent, and certain of his or her judgments.

Research in the field of psychology (Fischhoff, et al., 1981) suggests that individuals may not demonstrate all of these traits when making decisions. It is often difficult for people to make clear distinctions between objectives and constraints. In addition, people are inconsistent for a variety of reasons and may not choose rationally between alternatives. These features should not necessarily be considered faults, but rather characteristics of individuals who may view a problem differently as more and different information becomes available.

Computer Program CHOICE

The design of monitoring networks occurs amidst constantly changing information about and knowledge of the aquatic environment. Because of this situation, a multiobjective technique was incorporated into the EPAM methodology that would allow users to evaluate carefully targets and criteria for their individual and specific settings and to select the monitoring design based on their unique needs. The method developed also would allow users to incorporate their objective knowledge and their judgment when it would provide the best information available.

This task was accomplished through the development of an interactive computer program, denoted CHOICE. The goal of this program is to allow users to describe objectives and potential alternatives throughout the EPAM program by assigning relative weights or values and then deriving rankings of importance. CHOICE is used at three levels: LEVEL 3 employs CHOICE to rank IMPACT AGENTS and their TARGETS, LEVEL 2 uses CHOICE to rank PRIMARY HYPOTHESES, and LEVEL 0 exercises CHOICE to rank individual SAMPLING PROGRAMS comprising a full MONITORING PROGRAM.

CHOICE was developed to allow decision-makers to demonstrate inconsistencies in their judgments, as will be described in detail in the following discussion, and to indicate the degree of the inconsistencies. CHOICE was based on a mathematical procedure suggested by Saaty (1977) for deriving weights for a set of activities based upon their relative ability to meet specified goals. The approach, unlike many others in the multiobjective literature, emphasizes the incorporation of human judgment into the decision-making process and the quantification of inconsistencies or uncertainties when they exist. The procedure, described in detail below, requires a pairwise comparison to be made between activities and allows a hierarchical structure to be developed for activities and objectives.

The following pages present an explanation of the procedure with a description of the program CHOICE. In addition, the technique is applied in two examples to illustrate its power as a decision-making tool.

Solution Technique

The user's fundamental requirement in multiobjective analysis is to choose between a variety of activities, each of which contributes in varying extents to one or more objectives. The challenge is to determine relative weights or priorities of each of the activities relative to each of the objectives and to combine the objectives to generate an overall ranking. In the case of aquatic sampling, there may be several objectives and several levels or hierarchies of objectives that arise. When this is the case, the user needs to rank the objectives in one level relative to those in the next higher level.

The technique presented here includes a method for scaling the weights in each level of a hierarchy with respect to an element of the next higher level. The user accomplishes this by constructing a matrix of pairwise comparisons of the activities whose entries indicate the degree to which each element dominates the others with respect to the given criteria.

This process can be posed as an eigenvalue problem (Saaty, 1977). Suppose that n activities are to be compared relative to their importance or worth in achieving an objective. A matrix can be generated that indicates the relative value or weight each activity has in achieving the objective. For instance,

$$
A = \begin{array}{c|cccc}
 & A_1 & A_2 & \cdots & A_n \\
\hline
A_1 & w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\
A_2 & w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\
A_3 & w_3/w_1 & w_3/w_2 & \cdots & w_3/w_n \\
\vdots & \vdots & \vdots & & \vdots \\
A_n & w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n
\end{array}
$$

where w_i is the weight assigned to an activity.

All elements in this matrix will be positive, and each will have the reciprocal property that $a_{ij} = 1/a_{ji}$. If this matrix is multiplied by the transpose of the weight matrix, w^T, the problem becomes:

$$\underline{A}\ \underline{w} = n\ \underline{w} \tag{4.1}$$

Thus far, it has been assumed that $\underline{w}$ is known. However, it is much more common that $\underline{w}$ is unknown and that what is available is the $\underline{A}$ matrix, the relative weighting derived by a decision-maker. Such matrices can be developed by questioning those familiar with the specific monitoring program. If $\underline{A}$ is known, then $\underline{w}$ can be recovered by solving:

$$(\ \underline{A} - n\ \underline{I}\ )\ \underline{w} = 0 \tag{4.2}$$

This equation represents the familiar eigenvector problem with $\underline{w}$ as unknown and $\underline{n}$ representing an eigenvalue of $\underline{A}$. It can be shown that since every row of $\underline{A}$ is a multiple of another, it has unit rank and thus all eigenvalues except one are zero. It follows that the value of the one non-zero eigenvalue, and therefore the maximum eigenvalue, is equal to n.

It has been implicitly assumed that the $\underline{A}$ matrix is cardinally consistent (i.e., element dominance is maintained). This may not be true because of human inconsistencies in judgment. The new matrix can be characterized as the original matrix with small perturbations that will result in small perturbations in the eigenvalues. The problem now becomes:

$$\underline{A}'\ \underline{w}' = \lambda'\ \underline{w}' \tag{4.3}$$

with the new associated eigenvalue problem of

$$(\ \underline{A'} - \lambda'\ \underline{I}\ )\ \underline{w}' = 0 \tag{4.4}$$

Saaty (1977) suggested that a determination of the inconsistency of the $\underline{A}'$ set of relative weights can be evaluated by comparing the eigenvalue obtained in this problem with the eigenvalue associated with a consistent matrix (which is equal to n). In addition, the eigenvector obtained represents the best compromise estimate of the true weights derived from the matrix of relative weightings.

Solution Procedure

Initiation of this solution procedure requires that the user develop a relative weight matrix for each objective to be considered. In the procedure, weights are assigned on the following scale:

Scale Value	Contribution to the Composite Set
1	All sets are of equal importance or this is the least important.
3	This set appears to be slightly more important.
5	Judgment suggests this set is more important.
7	This set has "demonstrated importance" over all others.
9	This set dominates all others.

(Even numbers are used to interpolate between the levels.)

Next the user constructs the positive pairwise correlation matrix for the weights of the components as described previously. Saaty (1977) suggested that the first row be completed and then the first column generated by using the reciprocals of the row entries. The second row then can be completed and used to complete the second column; this process continues through each row and column. When objective data are available, the weights are computed on a scale of one to ten. When objective data are unavailable, the weights are generated using a Delphi technique or subjective judgment.

Examples

The following examples illustrate the utility of the CHOICE program. The first example is extremely simple in construction and illustrates the application of the methodology to a single hierarchy problem. The second example expands the setting to a multiple hierarchy problem, a more realistic setting for aquatic monitoring design.

Example 1: Suppose that the user wants to compare the weights of five aquatic species. These weights may represent actual weights in terms of their mass or perhaps the economic value of the species. Suppose initial values are:

	Starfish	Salmon	Perch	Trout	Clam
Starfish	1/1	1/5	1/3	1/4	4/1

If the weights were derived in a rational and consistent manner, there would be no need to make any other comparisons. In this case the relative ranking would be:

Salmon	-	5
Trout	-	4
Perch	-	3
Starfish	-	1
Clam	-	0.25

However, as described previously, this may not be the case. Assuming that such comparisons are inconsistent, the entire comparison matrix might look like:

	Starfish	Salmon	Perch	Trout	Clam
Starfish	1	1/5	1/3	1/4	4
Salmon		1	2	2	8
Perch			1	1/2	4
Trout				1	7
Clam					1

(Note that the lower triangular matrix values would be the inverse of the given upper triangular values.)

Because the relative rankings of these elements are inconsistent, their relative values cannot be determined from this matrix. To do so the user could initiate CHOICE to determine the relative weights that are the most consistent with the weights reported. Calculation of the eigenvector for this inconsistent matrix using CHOICE determined the relative rankings to be:

Salmon	-	5.00
Trout	-	3.50
Perch	-	2.25
Starfish	-	1.10
Clam	-	0.45

The eigenvalue determined for this matrix is 5.15, indicating that it is approximately consistent. These relative weights could be used in the next level of objectives as indicated in the next example.

Example 2: The second example illustrates how multiple hierarchies can be incorporated into the methodology. The example considers a user wanting to select a monitoring network from three potential candidates. These networks are denoted as Design A, Design B, and Design C. Two organisms are evaluated by the networks, Fish and Benthic organisms. Suppose that two criteria are important in the evaluation of the network, the cost of the network and its power at detecting change. These criteria for the potential designs are summarized as:

Design A - Best power for Fish, most expensive
Design B - Best power for Benthic, second most costly
Design C - Poorest for power for both Benthic and Fish, least costly

Suppose that after examination, the relative weights derived for the criteria are determined to be equal:

	Power for Fish	Power for Benthic	Cost
Power for Fish	1	1	1
Power for Benthic		1	1
Cost			1

The relative weights for each design for each objective are determined:

Power for Fish

Design	A	B	C
A	1	4	6
B		1	2
C			1

Power for Benthic

Design	A	B	C
A	1	.8	0.6
B		1	1.5
C			1

Cost

Design	A	B	C
A	1	2	0.1
B		1	0.2
C			1

Rankings of the potential designs relative to the objectives, and the objectives relative to one another, are needed so that the best design, based upon the objectives, can be chosen.

From CHOICE the relative ranking of each of the alternatives is:

	Power For Fish	Power For Benthic	Least Cost
Alternative A	1.000	0.635	0.158
Alternative B	0.275	1.000	0.128
Alternative C	0.151	0.840	1.000
Eigenvalue	3.01	3.05	3.23

Based on the rankings of the three objectives, the final rankings of the alternatives calculated by CHOICE are:

Alternative A: 0.36 Alternative B: 0.23 Alternative C: 0.41

It is interesting to note the effect on the ranking of the alternatives if the weighting of the criteria is altered. Consider for instance, the following weighting of criteria:

	Power for Fish	Power for Benthic	Cost
Power for Fish	1	2	2
Power for Benthic		1	3
Cost			1

The rankings of the alternatives in this case are as follows and indicate the selection of a different alternative:

Alternative A: 0.45 Alternative B: 0.25 Alternative C: 0.30

Summary of Technique

The technique described above was incorporated into EPAM to supply a decision criterion to choose among sampling alternatives in the development of a sampling program. The required inputs for this type of analysis are matrices relating each of the alternatives to each criterion and relating the relative importance of each of the objectives to be achieved. These inputs are direct outputs of other subroutines of the EPAM model. With these inputs CHOICE can be used to allow decision-makers to determine the existence and extent of inconsistencies and to select among alternatives that best achieve expressed objectives.

DESIGN AND OPTIMIZATION OF SAMPLING PROGRAMS

The four questions the user must address in designing any sampling program are the type, location, frequency, and number of replicates to collect. The first issue, what to sample, is addressed in LEVELS 2 and 3, and is essentially not a statistical question. The second problem of where to sample can be considered from a statistical viewpoint only insofar as sampling locations are not constrained by the factors incorporated into the design. For instance, in the sampling of benthic organisms, substrate type may have a strong influence on organism abundance, and it simply will not make sense to sample in areas having substrates that are inhospitable to the organisms of interest. In some cases, other factors, such as water depth, circulation patterns, or salinity may constrain station location to the point that the designer is left with very limited flexibility in placing stations. On the other hand, sampling of some variables, such as water quality indicators, may be almost totally unconstrained, and then statistical considerations may be brought to bear on the problem of station location. The problems of how often to sample and how many replicates to collect usually may be addressed statistically, as these choices will directly influence the detectability of change. The station selection modules of EPAM (see Appendices D.5 and D.6) rely heavily on expert judgment to select candidate station locations. The statistical algorithms discussed in this section are then invoked to determine the number of stations, and their specific locations.

The type of statistical test to be used also bears directly on the sampling design. Factorial treatment designs, making use of the analysis of variance (ANOVA), are by far the most widely used technique for the detection of short-term ecological change (McKenzie, et al., 1977). They have been widely accepted because ANOVA is well understood and explicitly accounts for multiple sources of variation (depth, transect location, time of year, substrate type, etc.) that might otherwise confound the detection of change. In some cases other tests such as sequential comparison, regression analysis, pattern recognition, etc. might be more appropriate; however, factorial designs provide by far the most general basis for sampling design. It should be emphasized that any sampling design is, of necessity, dependent on the assessment methodology that will be used, and the work reported here is not immune from this restriction. Likewise, arguments could be made against ANOVA on the basis of its frequent misuse, which is perhaps a result of its familiarity to most analysts. Nonetheless, if properly used it is a versatile method, and is the sampling design approach outlined in this section.

As reported in the Phase I completion report, the majority of sampling designs for aquatic ecological impact detection have addressed the problem of short-term change. In this context, factorial designs appear to be the best approach. However, for the detection of long-term change time series analysis is perhaps the most applicable method. As will be described later in this section, time series methods are essential when temporal correlation of the model residuals becomes significant.

The most important consideration in determining space-time sampling tradeoffs, and the desired amount of replication, is the structure of the data. Specifically, in a factorial treatment design with three factors, for instance, it is assumed that

$$Y_{ijkl} = \mu + A_i + B_j + C_k + AB_{ij} + AC_{ik} + BC_{jk} + ABC_{ijk} + \varepsilon_{ijkl} \quad (4.5)$$

i.e., the observation is the sum of various effécts plus a random, or unexplained, error term. The correlation in the errors ε_{ijkl} determines the redundancy of observations in space and time, which can be exploited to address problems of sampling location, frequency of sampling, and replication. It should be emphasized that the ANOVA model is exclusively concerned with the variance explained by the various factors, assuming that the residual, or error term, ε_{ijkl}, is independent with respect to all factors. Such a condition appears to be the exception rather than the rule; moreover, the correlation between the residuals must be a key consideration in achieving efficient sampling designs.

Therefore, in the remainder of this chapter six topics related to sampling program design are addressed: (1) evidence of, and appropriate models for, correlation between the residuals; (2) the costs associated with various aspects of aquatic sampling, compiled from a number of sources; (3) optimization of sampling programs designed to detect short-term aquatic change when the residuals are uncorrelated and continuous in space and time; (4) optimization of sampling programs for short-term aquatic change when the error term is continuous but correlated in time and/or space; (5) development of a methodology for balancing allocation of sampling resources between monitoring programs designed to detect short- and long-term aquatic change, respectively; and (6) an approach for incorporating deterministic ecosystem models into monitoring designs.

Much of the information presented in this section is technical in nature, and a detailed understanding of the algorithms presented requires that the reader be comfortable with basic mathematical and statistical concepts. Therefore, a brief overview is included at the beginning of each section, and a summary at the end, so that the reader, if so inclined, can skip the intermediate derivations without loss of continuity.

Exploratory Data Analysis

Three sets of typical aquatic monitoring data were examined for evidence of correlation in the residuals at each station over time (temporal correlation), and correlation in the residuals between locations at a fixed time (spatial correlation). Then two different models of the decay of spatial correlation with distance were fitted to the data. The three sets of aquatic monitoring data included (1) extensive benthic data collected around the Palos Verdes Peninsula of the Southern California coastal area by the Los Angeles County Sanitation District (Garrison, 1981); (2) benthic and phytoplankton data from the Zion nuclear generating station, located on the western shore of Lake Michigan immediately south of the Wisconsin-Illinois state line; and (3) benthic, phytoplankton and zooplankton data from the Nine Mile Point nuclear generating station, located on Lake Ontario near Oswego, New York. The three sites are reasonably diverse with respect to water body type, test species, and climate. Brief descriptions of the principal characteristics of each data base follow.

Los Angeles County Sanitation District

The Los Angeles County Sanitation District (LACSD) is responsible for wastewater treatment for much of the greater Los Angeles area. A major LACSD outfall discharges primary treated wastewater approximately 1.75 miles (2.8 kilometers) offshore from Whites Point on the Palos Verdes Peninsula. Discharge of wastewaters to the marine shelf in this area has been occurring since 1937, and is believed to have greatly impacted the native benthic populations (Garrison, 1981).

The waters off the Palos Verdes Peninsula have been sampled extensively for soft-bottom benthic invertebrates (biomass measurements) since 1974. The monitoring network utilized is shown in Figure 4.1. It consisted of forty-four stations on four depth contours (eleven stations per depth contour). The depth contours were at 100, 200, 500, and 900 feet. The stations were located on an approximately rectangular grid, and spaced one mile apart. The analyses were undertaken on the assumption that the station coordinates were described by the nominal grid spacing.

The variable analyzed in this project was the average total benthic biomass (e.g., gm/m^2) per station. Samples were taken semi-annually from 1972 to 1978. Four replicates were taken at each station from 1972 to the first semester of 1977, while only three were taken from then on. The number of observations per station was not uniform; some stations were not sampled during 1972 and 1973, and a few other values are missing. The average number of observations per station was fourteen.

Nine Mile Point Nuclear Power Station Site

The Nine Mile Point station began commercial operation in December 1969 and draws once-through cooling water from Lake Ontario. The variables analyzed for this site were total benthic population density [number/m^2], and total phytoplankton density [cells/ml]. The monitoring network, shown in Figure 4.2, consisted of twenty benthic organism sampling sites distributed over five depth contours (10, 20, 30, 40, and 60-feet), and sixteen phytoplankton sampling sites, which coincided with most of the benthic stations. The 30-foot depth contour was not sampled for phytoplankton (Murarka, 1976).

Benthic organisms were sampled in the months of April, June, August, and October from 1975 to 1978. All benthic stations were sampled at this frequency to yield a total of sixteen observations per station. Phytoplankton was sampled monthly from April to December from 1976 to 1978, yielding a total of twenty-seven observations per station (no phytoplankton values were available for January, February or March).

Zion Nuclear Power Station Site

The data and information from the Zion site were taken from a report by Murarka et al. (1976). Like Nine Mile Point, this plant is located on one of the Great Lakes (Lake Michigan). The plant began commercial operation in 1974. The variahbles analyzed here were total benthos [number/m^2], total phytoplankton [number/ml], and total zooplankton [number/m^3].

Figures 4.3 and 4.4 show the field sampling locations for benthos and plankton, respectively. The monitoring network for benthic data included 20 stations and five contours at 10, 20, 30, 40, and 60 feet. The plankton sampling stations were included among the benthic sites; there were twelve stations (six treatment and six control) distributed on three depth contours at 10, 30 and 60 feet.

Benthos were sampled bimonthly for the years 1972 to 1975, giving a total of 24 observations per sampling location. Phytoplankton and zooplankton were sampled monthly for the same period of time yielding 48 observations per sampling station. Very few data were missing.

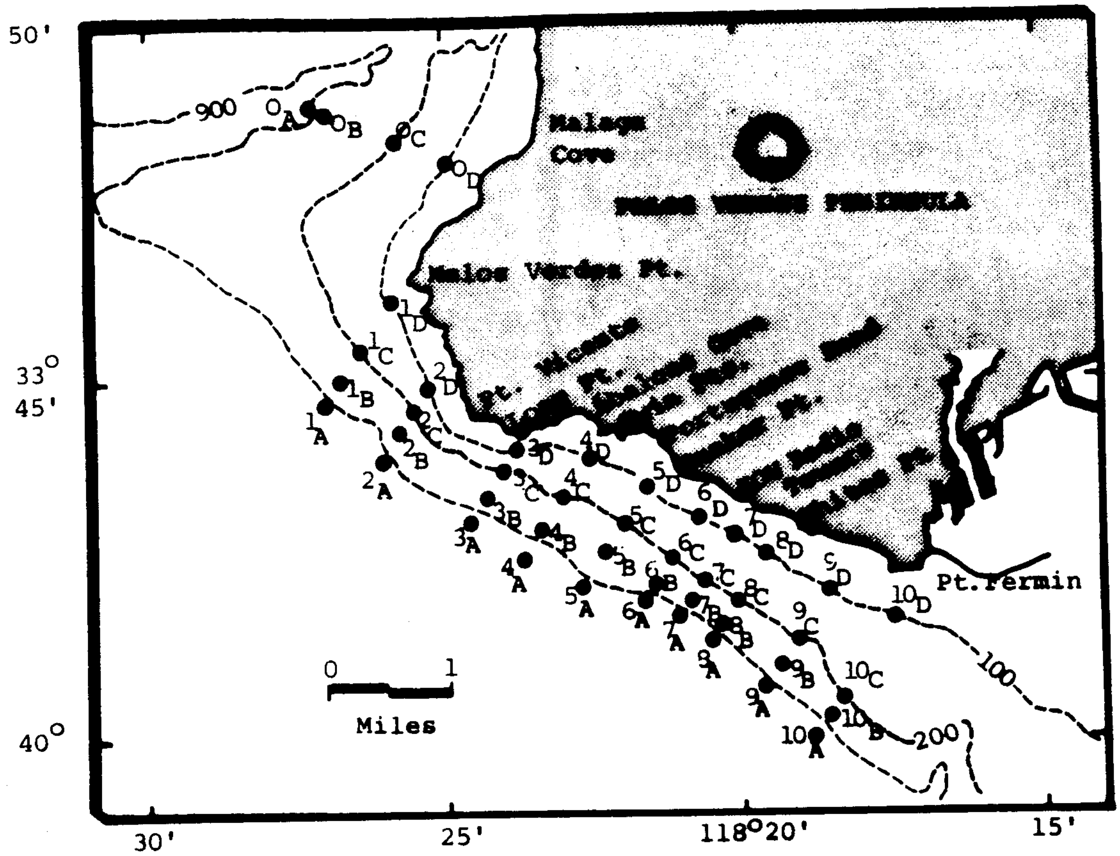

Figure 4.1. LACSD Sampling Network

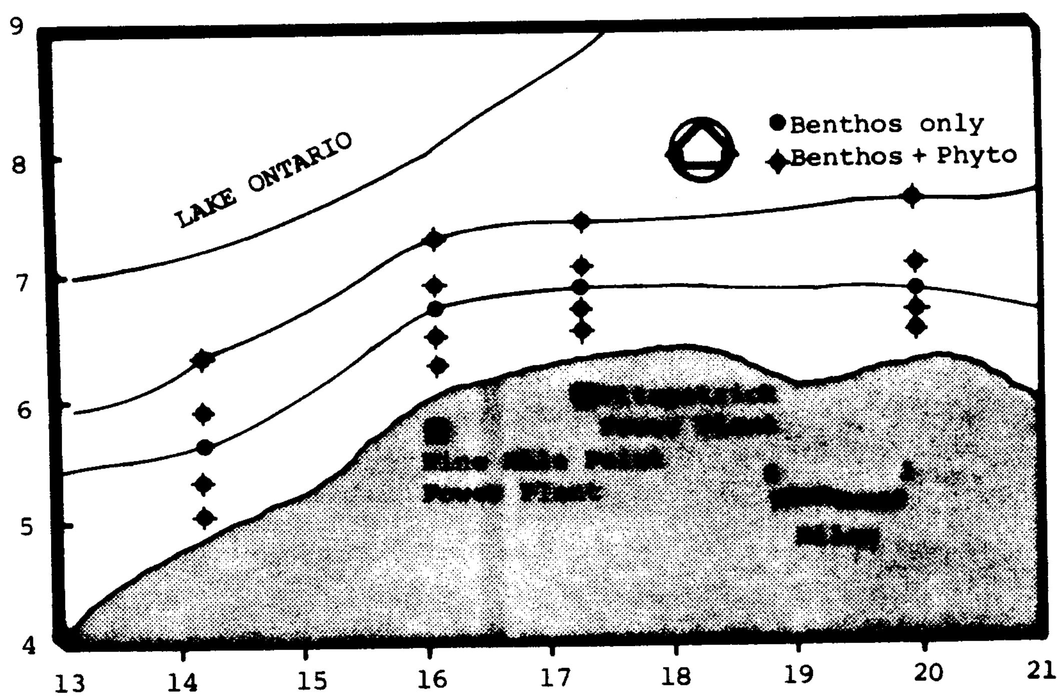

Figure 4.2. Nine Mile Point Sampling Network

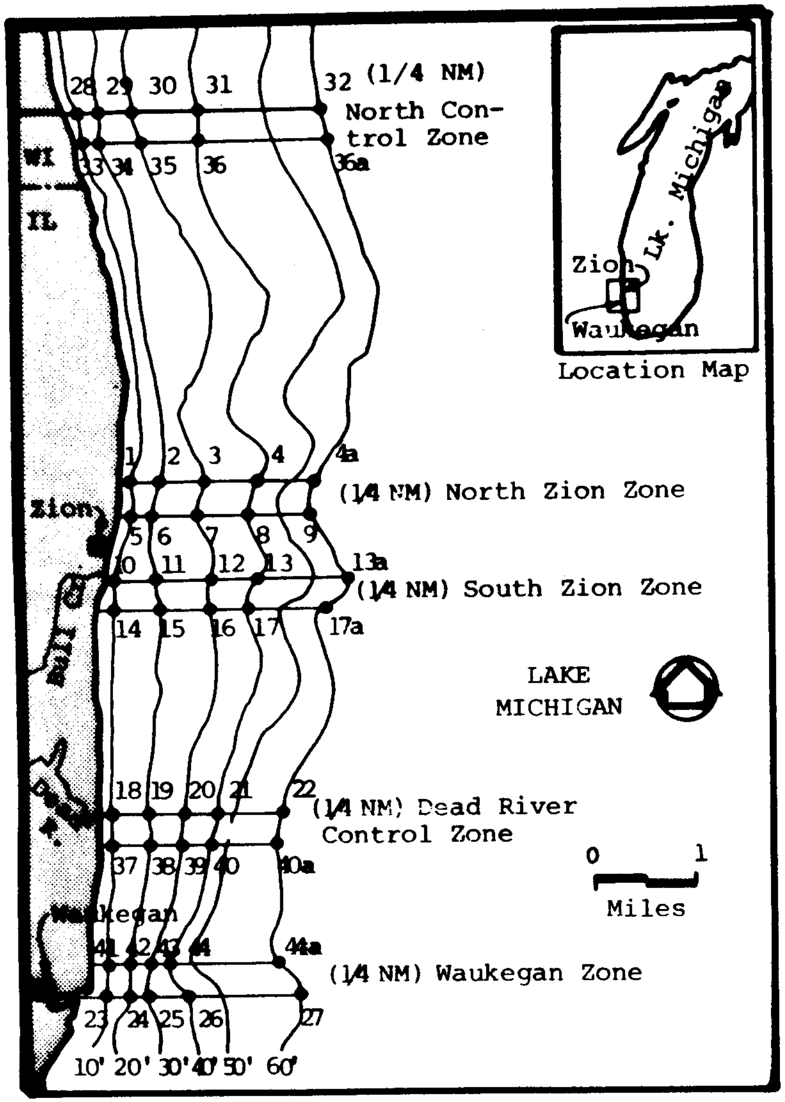

Figure 4.3. Zion Benthos Sampling Network

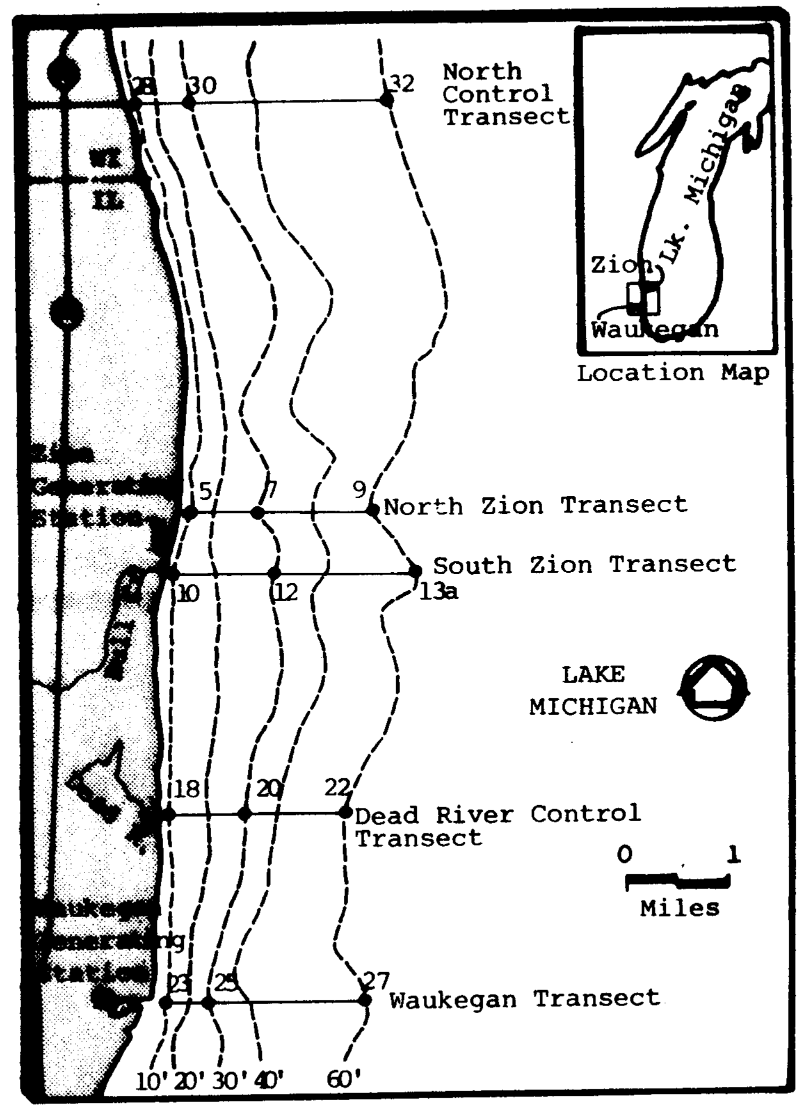

Figure 4.4. Zion Plankton Sampling Network

Estimating Temporal Correlation

The estimates of temporal correlation for phytoplankton at Zion are given in Table 4.1 and were derived as follows. The data were grouped into six seasons, with each season consisting of two months, so January-February was one season, March-April another, and so on. Two-month periods were combined into seasons to provide a reasonable sample size to estimate the means and standard deviations. Seasonal means and standard deviations were then estimated for each station (see Table 4.2). For each station, observations were standardized by subtracting the appropriate seasonal mean and then dividing by the seasonal standard deviation. Based on these standardized observations, maximum likelihood autocorrelation estimates were then calculated according to Box and Jenkins (1976). It should be noted that observations taken before and after plant operation were combined when estimates of seasonal means and standard deviations were computed. This procedure is justified by the fact that Murarka, et al. (1976) found no plant effect for either phytoplankton or zooplankton, although McKenzie, et al. (1977) found ambiguous results on this issue.

A problem in testing for spatial correlation is the effect of autocorrelation at each site on the estimator of cross-site correlation (Jenkins and Watts, 1968). Therefore, when autocorrelation is present, the data should be pre-whitened (i.e., the autocorrelation should be removed) before the cross correlation between stations is computed. The estimated autocorrelations were quite small at all sites, and in fact when an AR(1) process was fitted to the standardized series at each station, none of the AR(1) parameters was significantly different from 0 at the .05 significance level, or even at the .10 level (two-sided test). Therefore, pre-whitening was not necessary, although in other cases it may be. Estimates of temporal correlation at the other sites investigated are summarized in Table 4.3. These estimates were computed using the same approach as was used to compute the Zion estimates.

Estimating Spatial Correlation

Spatial correlations were determined by applying the following procedure to all sites. Given two sampling sites with their respective data sequences, the correlation of the residuals was calculated and plotted versus the separation distance between the sites. The scattergram obtained in this manner was

Table 4.1. Estimates of Autocorrelations at Zion,
Variable = ln (total number phytoplankton/ml)

Station Lag	5	7	9	28	30	32
1	-0.06	-0.22	-0.13	0.11	-0.06	-0.05
2	0.17	0.11	0.21	0.20	0.10	0.04
3	0.11	-0.01	0.03	0.06	-0.17	0.02

Table 4.2. Seasonal Means and Standard Deviations at Zion
Variable ln (total number of phytoplankton/ml) at Zion

SEASONAL MEANS

Station Season	5	7	9	28	30	32
1	7.68	7.57	7.43	7.78	7.57	7.30
2	8.22	7.97	7.66	8.04	7.72	7.08
3	7.72	7.60	7.35	7.89	7.52	7.34
4	7.25	6.92	6.70	7.33	6.79	6.37
5	6.96	6.81	6.49	7.04	6.75	6.46
6	7.44	7.25	6.94	7.60	7.32	6.67

SEASONAL STANDARD DEVIATIONS

Station Season	5	7	9	28	30	32
1	0.51	0.48	0.30	0.49	0.37	0.37
2	0.67	0.46	0.43	1.20	1.09	1.37
3	0.81	0.63	0.62	0.80	0.71	0.86
4	0.67	0.67	0.89	0.57	0.74	0.84
5	0.32	0.37	0.63	0.36	0.42	0.47
6	0.58	0.71	0.75	0.61	0.58	0.56

Table 4.3. Estimates of Autocorrelation

Site	Variable	Min	Max	Average	N	n
LACSD	Benthos	-0.53	0.79	0.14	44	12
NMP	Benthos	-0.18	0.65	0.18	20	15
	Phytopln	0.02	0.54	0.31	16	26
ZION	Benthos	-0.39	0.42	-0.07	20	21
	Phytopln	-0.22	0.08	-0.08	12	46
	Zoopln	-0.16	0.11	0.01	12	46

N = Number of sampling stations,
n = Number of observations at each sampling station.

then analyzed in search of a pattern of spatial correlation (a spatial correlation structure) that could be associated with the organism or environment in question.

Table 4.4 gives the estimates of the correlations between stations for the phytoplankton at Zion. Treatment stations appear to be just as highly correlated with each other as with their respective paired control stations. The variances of these estimates of cross correlation are on the order of $(1 - \rho^2)^2/n$, where ρ is the true correlation and n is the number of observations used to compute the estimate (Anderson, 1958). The smallest sample correlation is 0.596 from a sample size of $n = 45$, so a rough upper boundary on the variances of these estimates is 0.01. All estimates are significantly different from 0 ($p < .0005$, one-sided test).

Given a value for the separation distance, the mean of the various correlation coefficient estimates corresponding to it were averaged to yield Cor(h), the estimate of the correlation between two stations that are separated by a distance h. The estimates of the correlation coefficients were assumed to be distributed normally around the population value, Cor(h), with a variance of $(1 - \mathrm{Cor}(h)^2)^2/n$ (Anderson, 1958). A scattergram of Cor(h) versus h was then plotted. A better view of the general behavior hidden by the variability of the scattergrams was obtained through the application of smoothing techniques. The technique utilized is called the three-repeated running-medians technique and was suggested by Tukey (1977). The medians of each set of three points sliding one at a time are taken as the new smoothed sequence. This technique was applied twice (hence the name repeated running-medians) to each data base with the exception of LACSD, since in this case the first smoothing gave values that were very close together.

Once the smoothed correlation estimates had been obtained as a function of distance, two correlation functions were fitted to the data. The correlation functions considered were an exponential decay,

$$\mathrm{Cor}(h) = \exp(-\, b_e h) \qquad 0 < b_e < \infty \tag{4.6}$$

and a modified Bessel function, K_1,

$$\mathrm{Cor}(h) = b_b h K_1(b_b h) \qquad 0 < b_b < \infty \tag{4.7}$$

where $K_1(.)$ is the modified Bessel function of the first kind.

Table 4.4. Estimates of Correlations Between Stations at Zion, Variable = ln (total number phytoplankton/ml)

Station	5	7	9	28	30
7	.748				
9	.859	.857			
28	$\boxed{.852}$	.596	.797		
30	.715	$\boxed{.794}$	.848	.781	
32	.763	.696	$\boxed{.859}$	.761	.815

Based on observations from April 1972 to December 1975. Boxed values are correlations between control-treatment pairs.

The fits of the two functions were achieved by weighted least squares, where the weights were the inverse of the variance of the correlation at each separation distance. The results are shown in Figures 4.5-4.10. With the exception of the LACSD benthic data, all sites and variables exhibited quite high correlations at close distances.

Although temporal correlation in the residuals was not significant in most cases (for monthly and bimonthly sampling schedules), spatial correlation frequently was found to be present. This result has important implications for sampling design, and is discussed in the following section.

Effects of Correlated Errors on ANOVA

The following sections present: (1) an example of a monitoring program design in which ANOVA is used to test for plant impact, (2) an explanation of the assumptions of ANOVA and their relationship to temporal and spatial correlation in the error terms, and (3) the results of a Monte Carlo study of the performance of an ANOVA procedure when temporal or spatial correlation is present in the error terms.

A state-of-the-art design to test for short-term, local impacts in the aquatic environment adjacent to an electrical power plant was presented by Skalski and McKenzie (1982). An example they gave of such a design is shown in Table 4.5. This design can be used to investigate whether the presence and operation of a power plant is affecting the mean of some indicator variable (e.g., the abundance of a particular species or the concentration of a

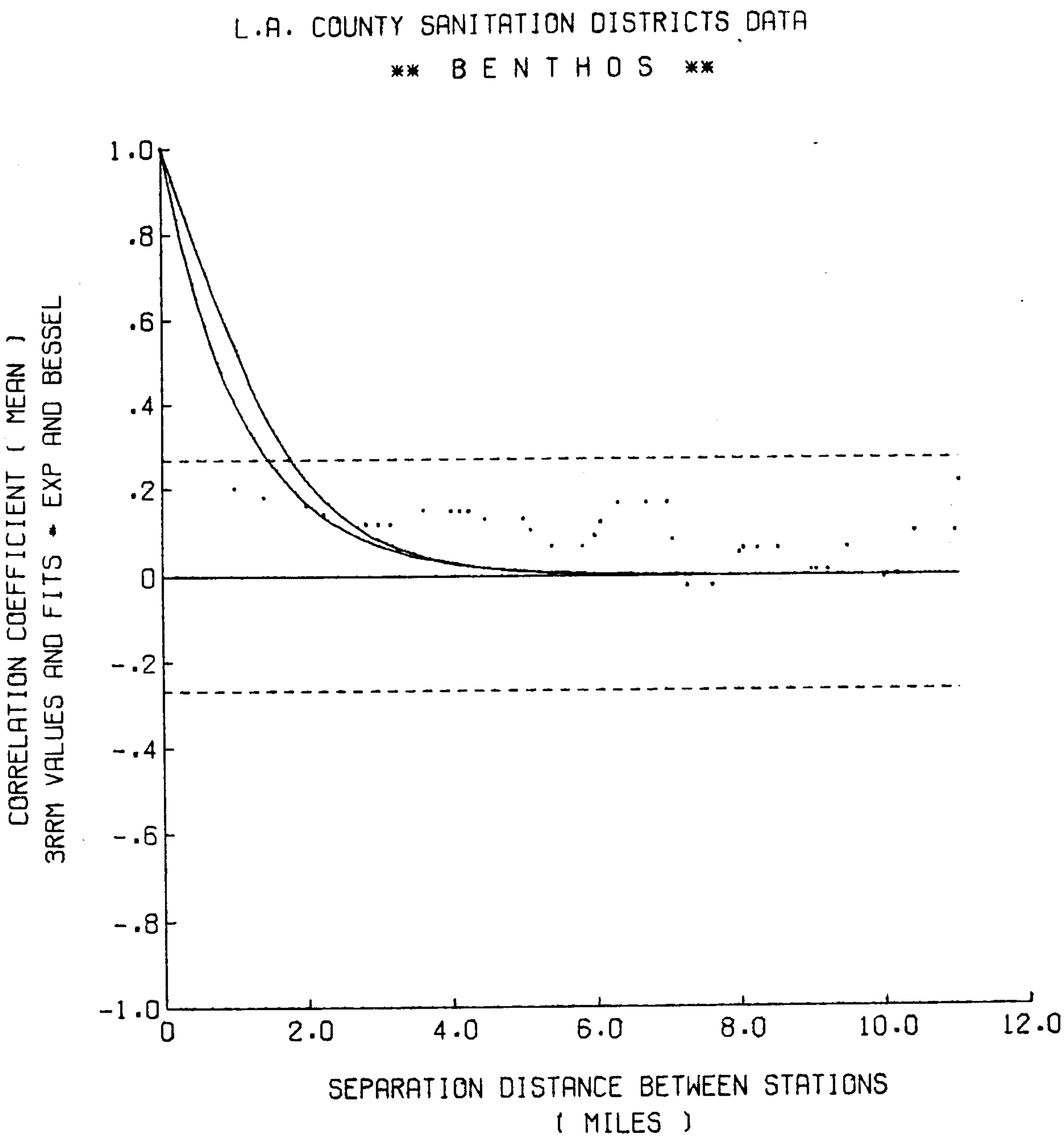

Figure 4.5. Error Correlation Versus Separation Distance: LACSD Benthos

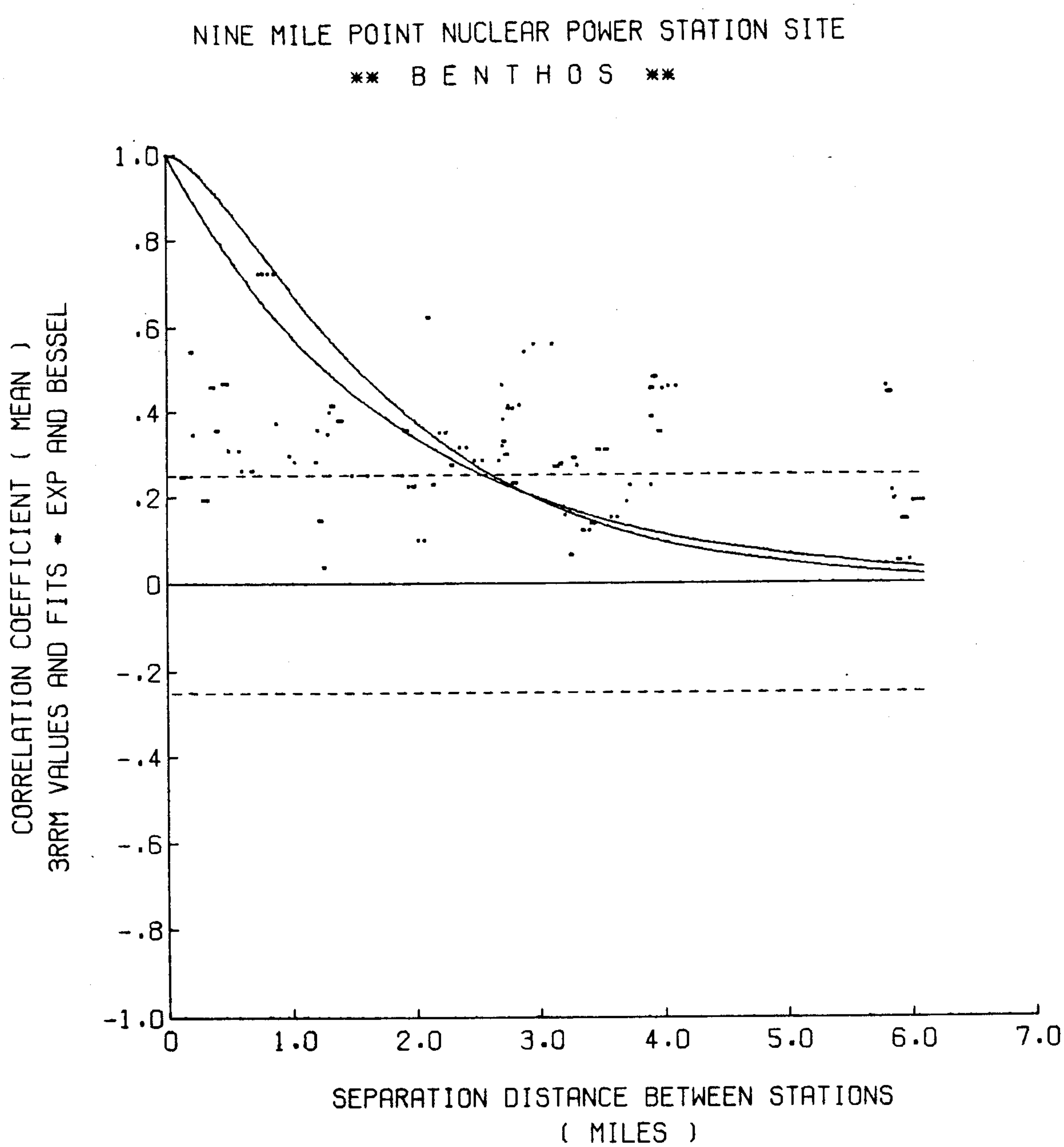

Figure 4.6. Error Correlation Versus Separation Distance: Nine Mile Point Benthos

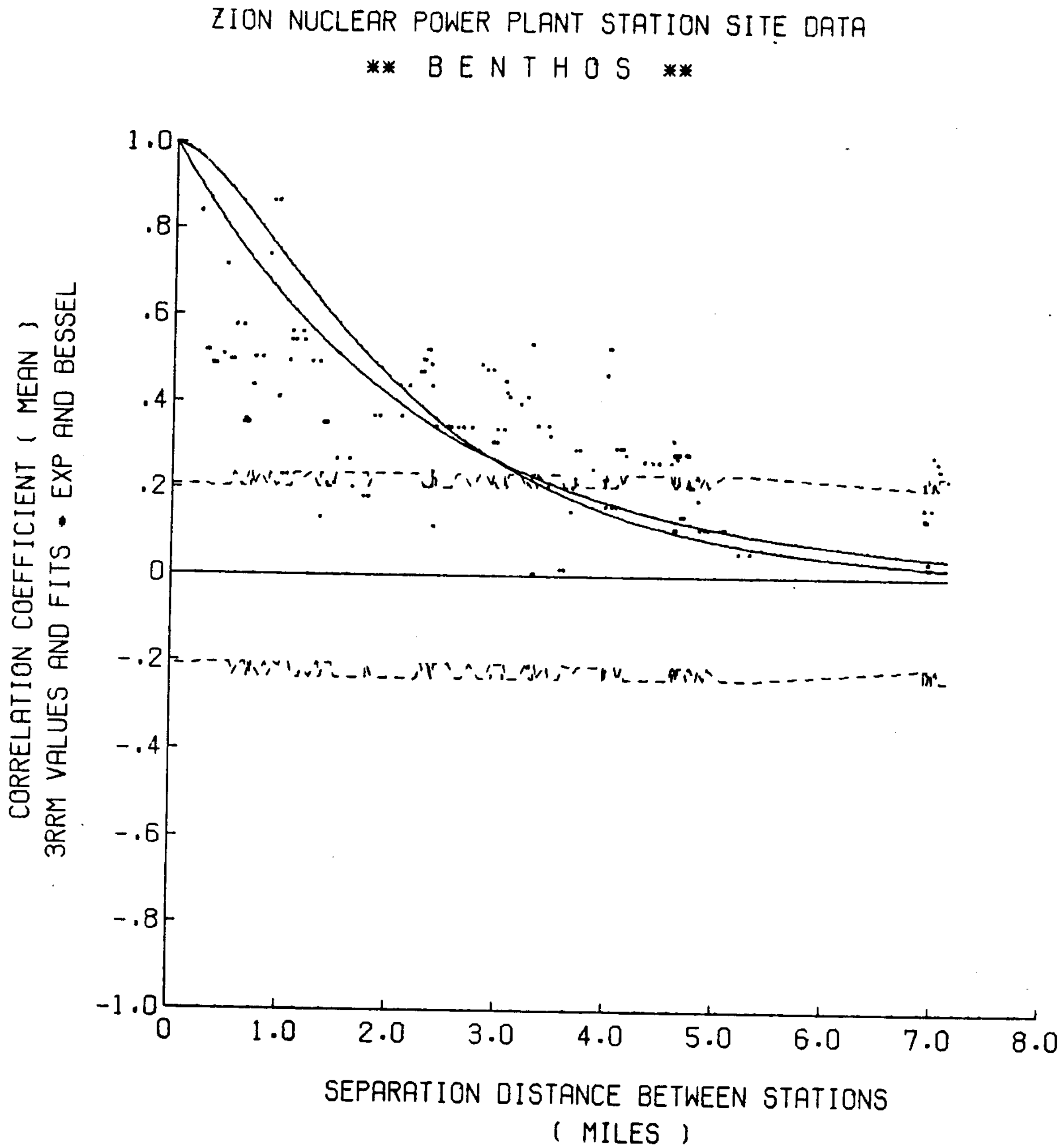

Figure 4.7. Error Correlation Versus Separation Distance: Zion Benthos

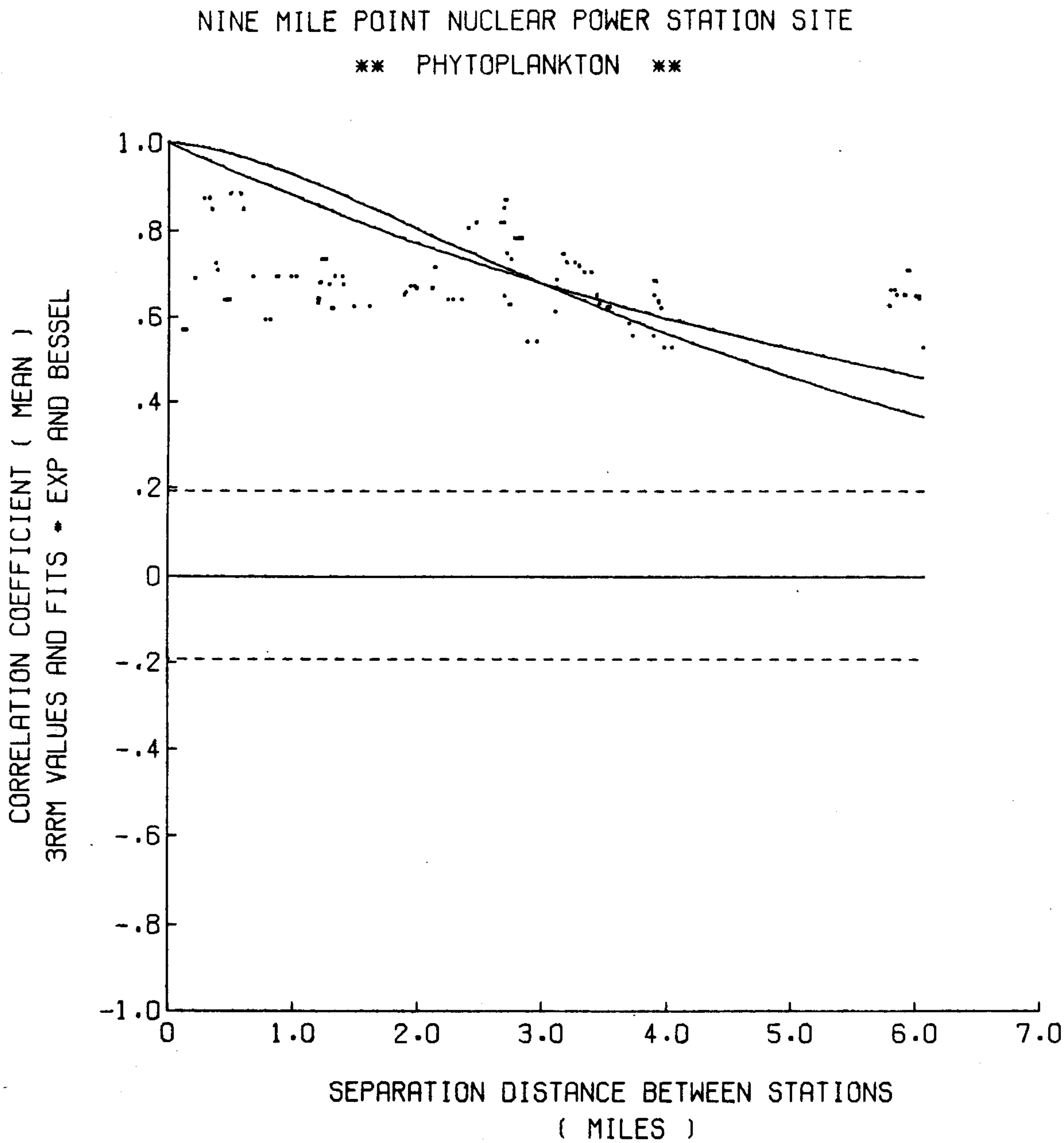

Figure 4.8. Error Correlation Versus Separation Distance: Nine Mile Point Phytoplankton

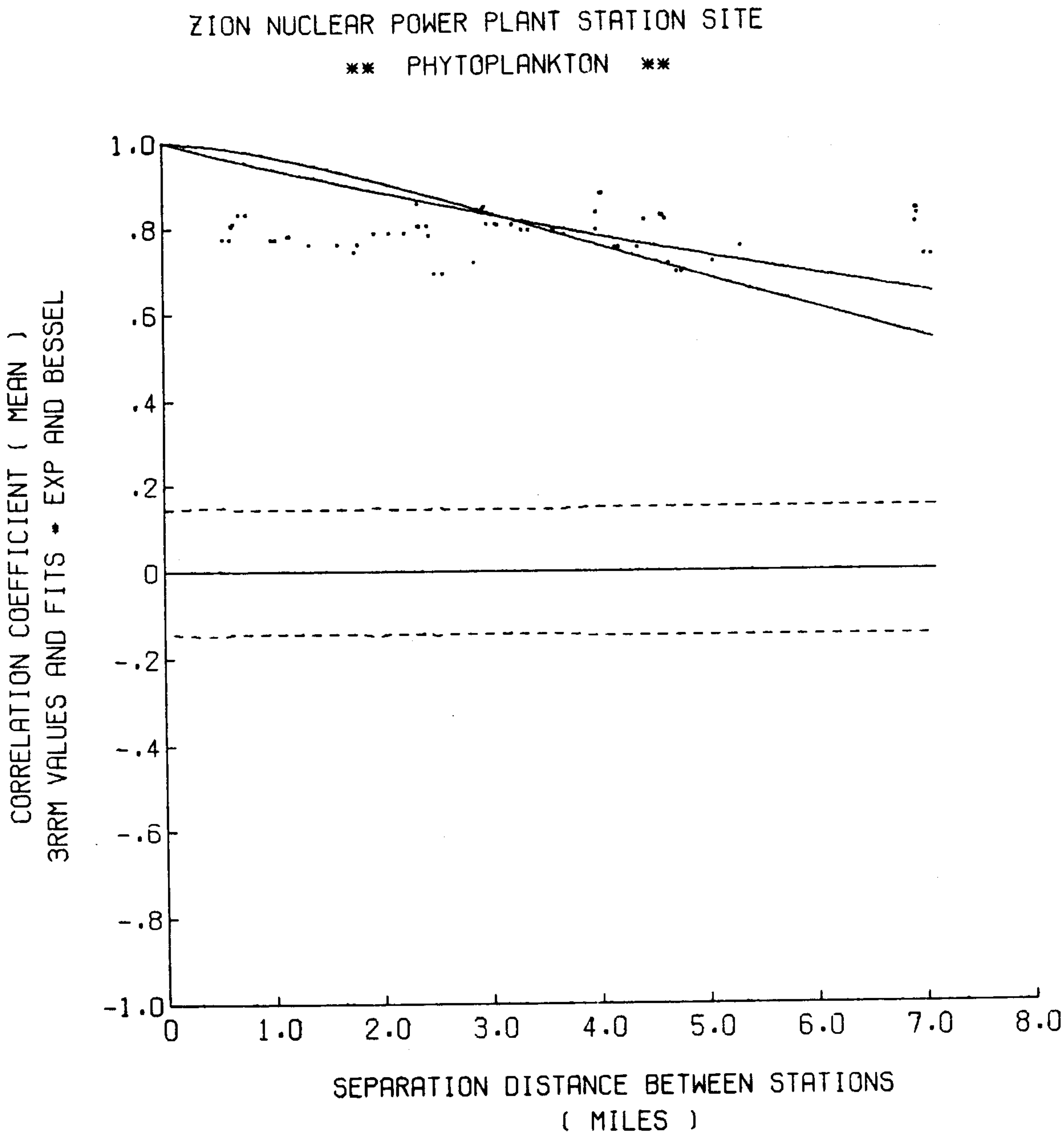

Figure 4.9. Error Correlation Versus Separation Distance: Zion Phytoplankton

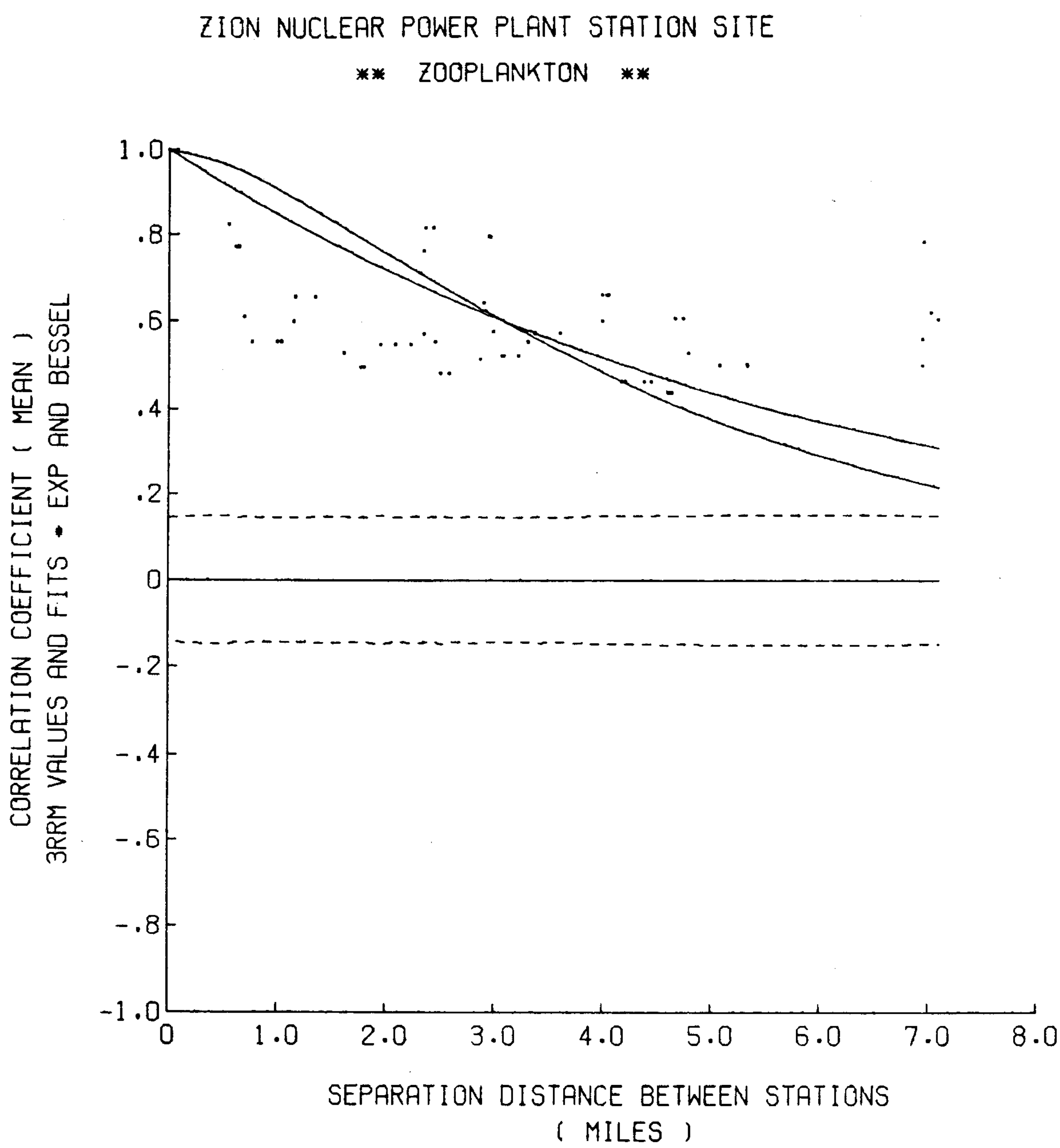

Figure 4.10. Error Correlation Versus Separation Distance: Zion Zooplankton

Table 4.5. Example of ANOVA Design (after Skalski and McKenzie, 1982)
Network Design: As in Figure 4.4.

Model: $y_{abcde} = \mu + A_a + B_b + C_c + D_d + AB_{ab} + AC_{ac} + AD_{ad}$

$+ BC_{bc} + BD_{bd} + CD_{cd} + \varepsilon_{abcde}$,

where y_{abcde} = the difference in the natural logarithms of organism abundance between control and treatment stations for station pair (c,d),

ε_{abcde} are i.i.d. $N(0,\sigma^2)$ random variables,

$e = 1, \ldots, n$,

n = number of years of sampling for each level of plant status

FACTOR	DESCRIPTION	NUMBER OF LEVELS
A	Plant Status	2 (1=pre-operational, 2=operational)
B	Time of Sampling	12 (monthly samples)
C	Location of Station Pair	3 (1=10 feet, 2=30 feet, 3=60 feet)
D	Position of Station Pair	2 (1=north, 2=south)

chemical substance). In this example it is assumed that the indicator variable is organism abundance. Six control-treatment pairs of sampling stations have been formed on the basis of depth contours (10, 30, and 60 feet) and location (north vs. south). The stations of each control-treatment pair are assumed to track each other, i.e., maintain a constant proportionality, or a constant difference, in the indicator variable over time (Skalski and McKenzie, 1982). Two stations that maintain a constant proportionality necessarily maintain constant differences between the logarithms. Two hypothetical examples of station pair tracking are shown in Figures 4.11a and 4.11b. Impact is defined as a statistically significant change in the proportional abundance (or change in the difference) of the indicator variable at control and treatment stations between pre-operational and operational periods of the power plant, (Skalski and McKenzie, 1982).

A statistical analysis of data that are gathered based on this design would involve an ANOVA procedure. In the example shown in Table 4.5, the factors included in the ANOVA are depth contour, station location, time of sampling (season), and plant status. If control and treatment stations truly track each other over time as in Figure 4.11a, season need not be included as a factor in the ANOVA. If, however, it is believed that differencing between treatment and control pairs does not remove all the seasonal fluctuations in the indicator variable, then season should be included as a factor. In this case, it is assumed that the control and treatment stations track each other in a manner similar to that shown in Figure 4.11b. Inclusion of season as a factor, moreover, will allow detection of differential changes across seasons, regardless of whether station pairs track each other as in Figure 4.11a or 4.11b. ANOVA procedures using models similar to the one given in Table 4.5 were used by McKenzie, et al. (1977, 1979) to evaluate data gathered at the Prairie Island, Zion, Haddam Neck, San Onofre, Calvert, Cliffs, and Pilgrim nuclear power plants.

Assumptions of ANOVA

Carrying out an ANOVA procedure requires that three assumptions be made about the error terms in the model (e.g., the ε_{abcde} in the model in Table 4.5), aside from assuming they all have mean 0 (Eisenhart, 1947). These

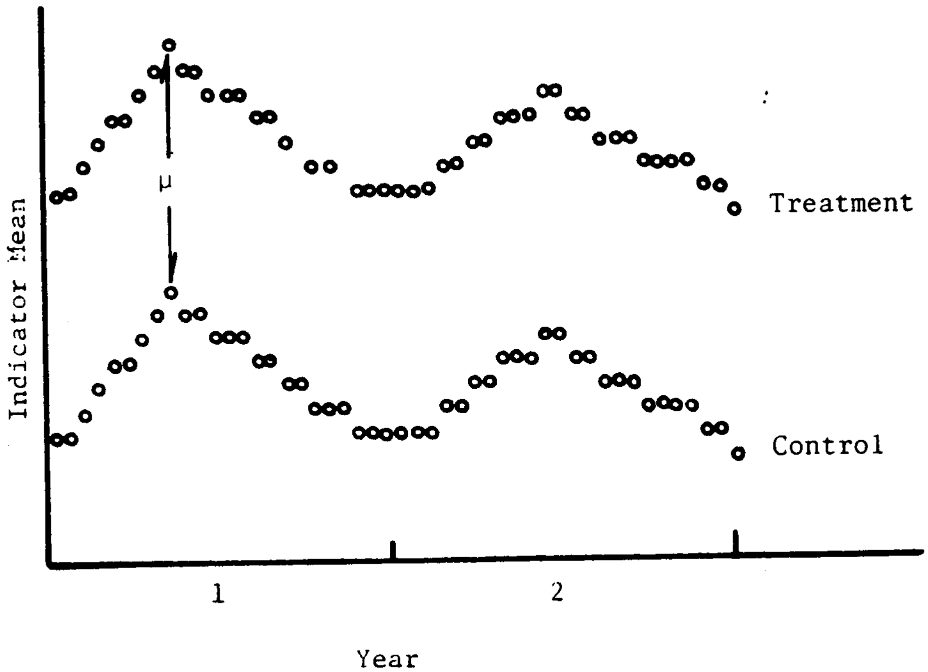

Figure 4.11a. Example of Perfect Station Pair Tracking

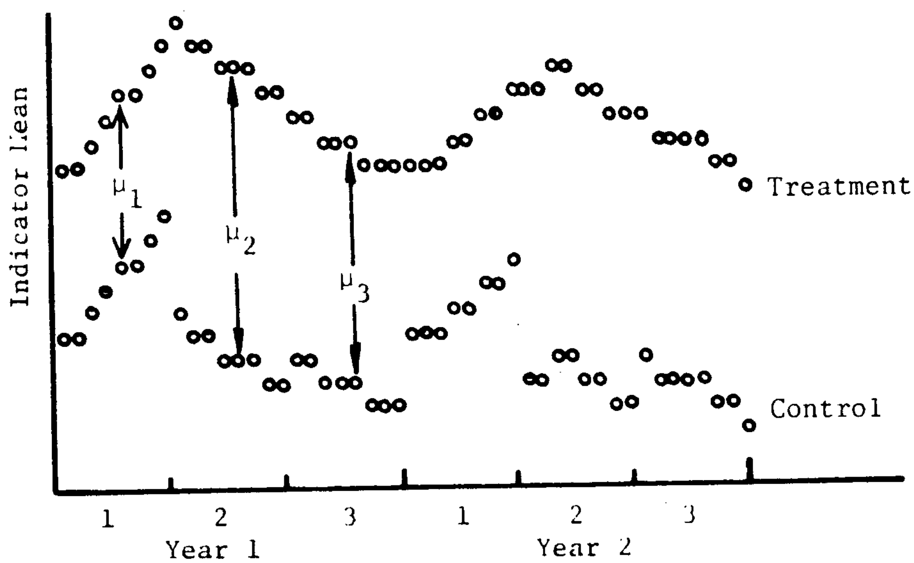

Figure 4.11b. Example of Seasonal Station Pair Tracking

assumptions, listed in order of decreasing importance (Glass, et al., 1977; Box, 1954) are:

1. The errors are independent (hence uncorrelated).
2. The errors all have the same variance.
3. The errors are all distributed as normal random variables.

The existence of serial (temporal) correlation among the observations at a treatment station and/or among the observations at its paired control station can induce temporal correlation among the differences in the observations for this station pair. Skalski and McKenzie (1982) stated, "use of proportional abundance [i.e., differencing the logarithms of abundance at paired treatment and control stations] reduces the serial correlation that can exist among the successive observations of abundance [or successive observations of the logarithm of abundance] at sampling stations ...". This proposition is not necessarily true. The relationship between temporal correlation among the observations at each station and temporal correlation among the differences between the observations for control-treatment pairs is given in Appendix D.1.

In addition, for each observation period, correlation among the N treatment and control stations (spatial correlation) can induce correlation among the N/2 station pair differences. The relationship between spatial correlation among the observations at each station and correlation among the differences is derived in Appendix D.2.

If spatial or temporal correlation among the observations at the stations induces correlation in the differences, these data (the differences) will violate assumption 1 of ANOVA given above, the assumption of independent errors. The qualitative effects of the violation of this assumption on the F-tests of the ANOVA have been known for some time (Cochran, 1947; Box, 1954).

Explanation of the Monte Carlo Study

To quantify the effects of spatially and temporally correlated errors on ANOVA set ups like the ones proposed by Skalski and McKenzie (1982), a Monte Carlo Study was carried out. The design used for this study was based on the one used at Zion Nuclear Power Plant (see Figure 4.4), and is given in Table 4.6. Observations were generated at each sampling station and the errors (deviations from the mean) were allowed to be correlated in space and time.

Table 4.6. Network Design and Model Used in Monte Carlo Study

Network Design: Stations 5, 7, 9, 28, 30, and 32 of Figure 4.4.

Station Number	Coordinates
5	0, 0
7	0.75, 0
9	1.88, 0
28	-0.48, 4.03
30	0.02, 4.03
32	1.92, 4.03

Model: $y_{ijk} = \mu + A_i + C_j + AC_{ij} + \varepsilon_{ijk}$,

where y_{ijk} = the difference between the observations at the treatment station and the control station for station pair j, for the kth sampling occasion at plant status level i

ε_{ijk} are i.i.,. $N(0,\sigma^2)$,

$k = 1, \ldots, n$

n = number of sampling occasions for each level of plant status

FACTOR	DESCRIPTION	NUMBER OF LEVELS
A	Plant Status	2 (1=pre-operational, 2=operational)
C	Location of Station Pair	3 (1=10 feet, 2=30 feet, 3=60 feet)

Temporal correlation was incorporated through the generation of an AR(1) process (Box and Jenkins, 1976) at each sampling station. Spatial correlation was taken to be an isotropic function of distance and took the form of the exponential decay model (4.6). The assumption of isotropy is open to question due to the influence of such factors as depth and current, which could lead to an anisotropic correlation function. The simple correlation model (4.6) was favored, however, because it requires the estimation of only one parameter, and because it is sufficient to indicate the general effects of spatially correlated errors on ANOVA.

The multivariate time series generating algorithm given in Appendix D.3 was used to simulate the observations at each station. It should be noted that the purpose of simulating observations at the stations was not to model the Zion phytoplankton data (although these data were used to provide an idea of the general levels of some of the parameters), but rather to provide a means of studying the effects of spatial and temporal correlation on ANOVA under controlled conditions.

The effects of spatial and temporal correlation on type I error only were considered in this study. The type I errors with regard to testing the following two hypotheses were examined:

1. The null hypothesis of no interaction between plant status effect and depth contour effect, given by:

   $$H_{0_1}: \ AC_{ij} = 0 \ \forall i,j$$

2. The null hypothesis of no plant status effect (no plant impact), given by:

   $$H_{0_2} : A_1 = A_2 = 0.$$

The interest in H_{0_1} lies in the fact that the presence of non-zero interaction terms complicates the interpretation of the results for the test of H_{0_2} (Seber, 1977). In practice then, it is often taken for granted that the usual test for main effects cannot be carried out if the null hypothesis of no interaction has been rejected.

For each set of parameters that determined a specific level of spatial and temporal correlation, 100 and 200 trials were carried out. Each trial consisted of:

1. The generation of 48 observations at each station (24 observations in the pre-operational phase, and 24 in the operational phase). The standard deviation of the observations was set at 0.6 because this is approximately the standard deviation of the observations at each station in the phytoplankton data from Zion (see Table 4.2). In some cases, the standard deviation of the observations was also set at 50 to test the invariance of the results.

2. The calculation of the differences between the treatment and control stations. All differences were distributed as normal random variables with mean 0. Hence, there was no seasonality in the differences (station pair tracking as in Figure 4.11a) in the generated data, and this factor was not included in the ANOVA model. Also, both H_{0_1} and H_{0_2} were always true.

3. Application of the test of H_{0_1}: no interaction.

4. If H_{0_1} was not rejected, application of the test of H_{0_2}: no plant impact.

In all the trials, the F-statistic was compared to the 95th percentile of its true distribution (in the absence of spatial and temporal correlation) under H_0. Thus, in the absence of spatial and temporal correlation, H_{0_1} and H_{0_2} should have been rejected about 5 percent of the time.

The results of the Monte Carlo study are shown in Tables 4.7 and 4.8, and in Figures 4.12 and 4.13. It appears that the probability of a type I error with regard to testing H_{0_1} (no interaction) decreases with increasing spatial correlation. On the other hand, the probability of a type I error with regard to testing H_{0_2} (no plant impact) increases with increasing spatial correlation, and with increasing temporal correlation.

Table 4.7. Effect of Spatial Correlation on ANOVA

1. $\emptyset = 0$, $\sigma = 0.6$, Number of trials = 200

b_e	$\hat{\alpha}_{int}$	$\hat{\alpha}_{ps}$
2	.08 (.02)*	.0625 (.0714)
1	.015 (.06)	.0761 (.0851)
0.5	.02 (.01)	.1173 (.1616)
0.1	0. (.01)	.1950 (.1111)
0.05	0. (0)	.1750 (.2700)

* Numbers in parentheses indicate results for 100 trials.

2. $\emptyset = 0$, $\sigma = 50$, Number of trials = 100

b_e	$\hat{\alpha}_{int}$	$\hat{\alpha}_{ps}$
2	.06	.0957
1	.01 (.045)*	.0505 (.0995)
0.5	.03	.1649
0.1	0 (.005)	.1700 (.1960)
0.05	0 (.005)	.2200 (.2261)

* Numbers in parentheses indicate results for 200 trials.

$\emptyset$ = the lag one correlation at each station,

σ = the standard deviation of the observations at each station,

b_e = the parameter of the spatial correlation function (4.1),

$\hat{\alpha}_{int}$ = the estimate of the probability of rejecting H_{0_1}: $AC_{ij} = 0 \ \forall i,j$

$\hat{\alpha}_{ps}$ = the estimate of the probability of rejecting H_{0_2}: $A_1 = A_2 = 0$.

Table 4.8. Effect of Temporal Correlation on ANOVA.

$b_e = 5$, $\sigma = 0.6$, Number of trials = 100

$\emptyset$	$\hat{\alpha}_{int}$	$\hat{\alpha}_{ps}$
0	0	.06
.05	.03	.03
.10	.04	.15
.30	.18	.16
.50	.40	.20
.70	.52	.31

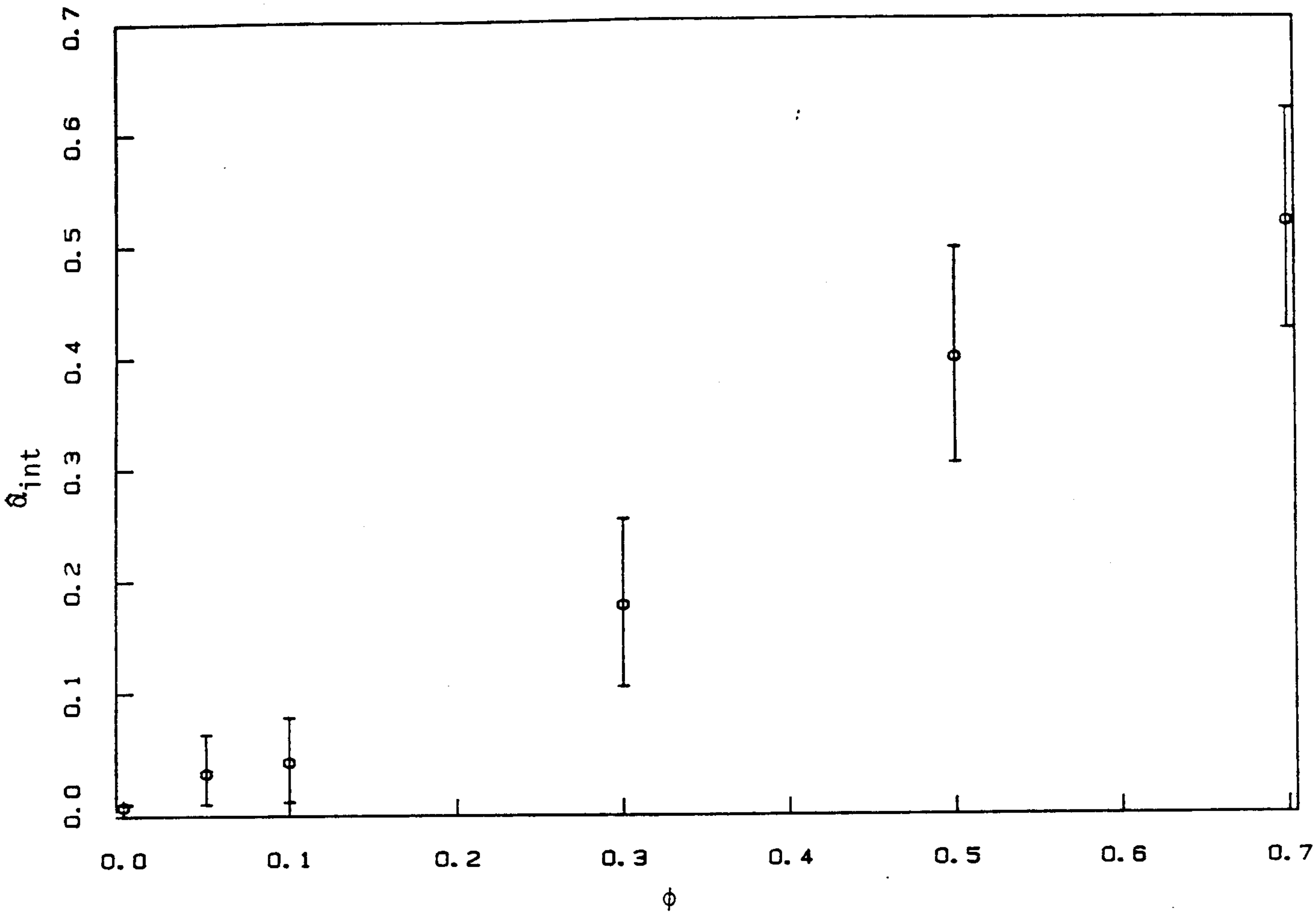

Figure 4.12. 95 Percent Confidence Intervals for $\hat{a}_{int}$.

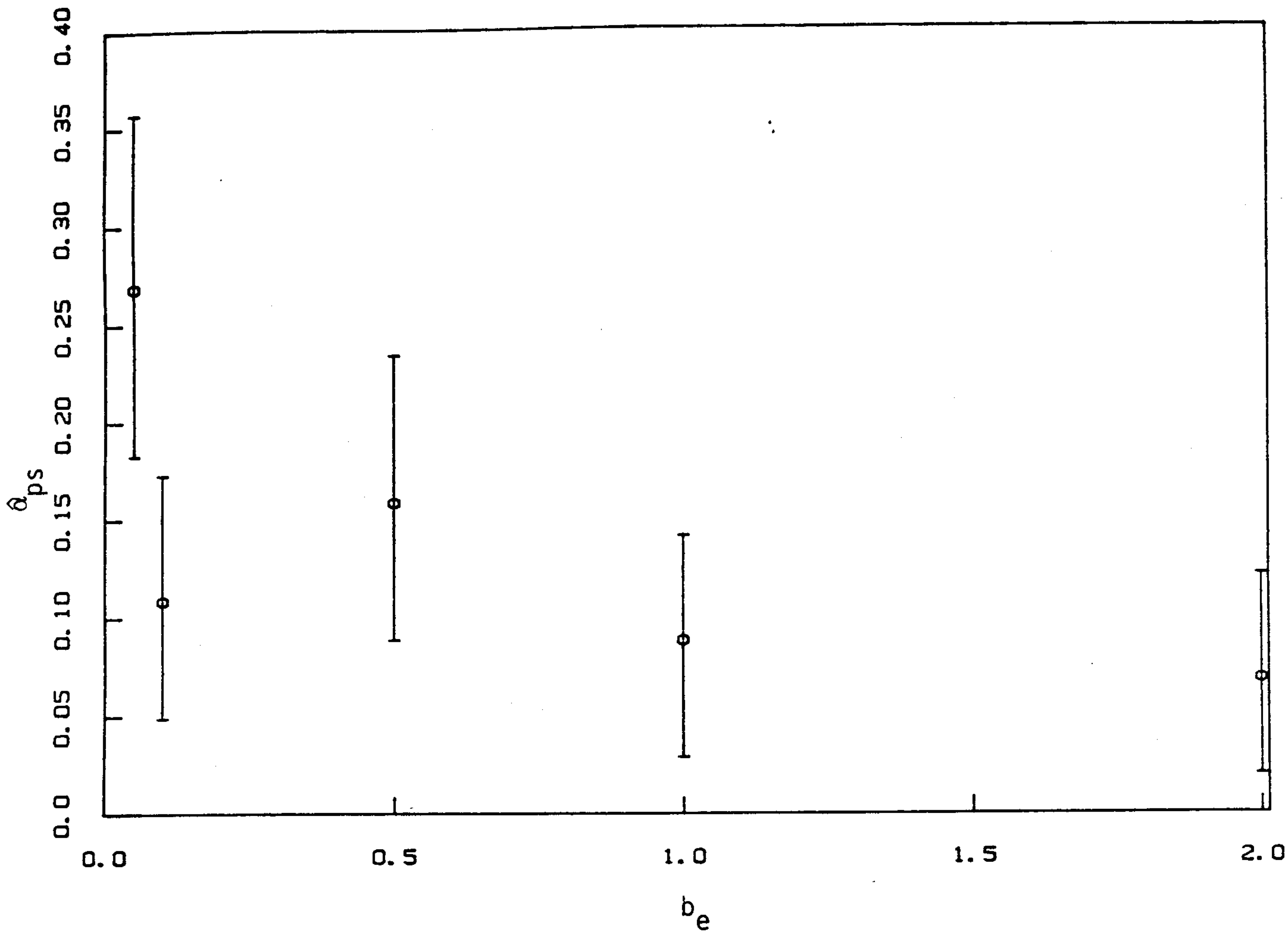

Figure 4.13. 95 Percent Confidence Intervals for $\hat{\alpha}_{ps}$.

The results for testing H_{0_2} agree with the findings of Box (1954) and Cochran (1947). They showed that positive correlation within groups can increase the type I error probability when means between groups are compared. For each station pair, the presence of temporal correlation induces serial correlation, within each level of plant status (see Appendix D.1). Hence, positive temporal correlation should increase the type I error probability for the test of plant status effect. The presence of positive spatial correlation induces positive correlation among the station pair differences for each sampling occasion (see Appendix D.2). Hence, positive spatial correlation also induces positive correlation within each level of plant status.

In the kinds of ANOVA models suggested for analyzing aquatic monitoring data, temporal correlation in the errors can increase the probability of a type I error when interaction or plant status effects are tested, and spatial correlation can increase the probability of a type I error when plant status effect is tested. An investigator analyzing aquatic monitoring data should test for spatial and temporal correlation in the errors (see Appendix D.5) before applying standard models and tests that assume uncorrelated errors. If temporal or spatial correlation appears to be present, modified techniques such as those introduced in the following sections should be used.

The next objective will be to develop optimization algorithms for the cases of correlated and uncorrelated errors. Because this optimization development incorporates cost considerations, an analysis of sampling costs, conducted during Phase II of the research, is discussed before the optimization sections.

Analysis of Sampling Cost

Sampling costs discussed here are based on a survey of data from five sources. A cost equation is described, and cost coefficients based on two different data bases are derived. This cost equation is used in the sampling program optimization algorithm described in the following sections. Numerous additional sources were contacted to provide such data, and an inquiry was placed in a Delphi questionnaire to elicit other potential sources. However, additional data sets were not found, either because cost data had not been assembled or could not be released.

An adequate determination of cost-effective monitoring requires accurate data describing the cost of sampling. A major problem in obtaining these data is a proper definition of what constitutes a sample. Confounding this problem is the introduction of continuous sampling technology into the field of aquatic monitoring. The cost to obtain a single observation using high technology (continuous sampling) is orders of magnitude less than the cost of traditional methods of discrete sampling and may provide quite different information. In addition, determining the marginal cost per sample can be an extremely difficult exercise because of the number of overlapping and dependent cost relationships that exist.

Evaluation of cost data was performed to generate parameters appropriate for the optimization techniques developed. Aquatic sampling data were obtained from five different organizations (Mattson, 1982). The general characteristics of these cost data are summarized in Table 4.9. All costs were adjusted to 1982 U.S. dollars, but differences in overhead rates, subsidies, profit margins, and salary differences could not be resolved. Major cost differences between sources may be due in part to the level of detail of laboratory or on-site analysis performed on each sample. Costs of specific levels of taxonomic identification or condition can vary greatly. Also the experience of the staff can influence sampling costs if costs of learning to sample properly are included. Some sources did not report net sizes, volumes of water sampled, or areas of substrate observed. Therefore, normalization of data was not possible.

Tables 4.10 and 4.11 present disaggregated cost data reported from sources B and E. While there are differences in boat and equipment costs, both these data sets suggest that laboratory and data management costs are major components of sampling costs. Proper allocation and definition of such costs will be essential in standardizing reported costs. Table 4.11 presents data that suggest labor and overhead costs also may require standardization before cost comparisons can be made. Because of the high level of variability of absolute sampling cost data demonstrated in these tables, relative costs of sampling were used in subsequent analyses, since relative cost data are less variable and easier for most individuals to estimate.

Table 4.9. Cost Data in 1982 Dollars for Five Sources

	Source A	Source B	Source C	Source D	Source E
Costs included:					
Collection	X	-	-	X	X
Boat/supplies	-	partial	partial	X	X
Equipment	X	-	X	X	X
Amortization	-	-	X	X	X
Supplies	X	X	X	X	X
Personnel Costs	X	X	X	X	X
Data Acquired:					
Density or Numbers	X	X	X	X	X
Species, Size, Weight Age, etc.	-	X	X	X	X
Sample Cost:					
Water (nutrient, conductivity, etc.)	19	95	35	--	85
Temp. and D.O. only	6	--	2	--	--
Phytoplankton	12	145	13	615	235
Zooplankton	18	--	130	230	380
Ichthyoplankton	--	660	110	155	--
Fish:					
Otter Trawl	28	460	--	120	295
Beach Seine	--	--	190	130	255
Trap	--	--	190	--	--
Gill Net	--	--	190	375	--
Benthic:					
Grab	23	--	--	--	185
Core	23	200	--	395	--
Quadrat	11	35	90	850	--
Acoustic Surveys ($/hr)	300	--	900	--	800

Source A - A 1982 research proposal submitted by a university for a variety of aquatic research sampling efforts. These are minimum costs due to low overhead (34%), low salaries, and exclusion of boat costs. The analysis, conducted only for density, excludes detailed species identification.

Source B - A 1982 operational phase environmental study for a West Coast power plant. Field sampling and laboratory equipment costs are not included in this set, but regular costs such as boat rental, transportation, and fuel costs are included.

Table 4.9 - Continued

Source C - A 1979 budget for a baseline study on a large lake. Only a portion of boat and motor costs are included since capital costs for equipment used but purchased for previous studies are not included. In the cost information submitted, the management, accommodations, data analysis, reports, and contingency costs were not included in the cost/sample values; they were allocated according to crew days to obtain overall cost/sample values.

Source D - A 1981 environmental laboratory study for a New England coastal power plant. Costs include overhead, laboratory O and M, boat O and M plus carrying charges or costs, miscellaneous expenses, and management/clerical salaries.

Source E - A 1982 budget for a baseline study on the West Coast for several sampling schemes. Equipment costs were estimated when the budget indicated that these items were being provided by the client or borrowed from another project. Overall management and direct costs that could not be directly associated with a sampling scheme were not included.

Table 4.10. Disaggregation of Sampling Costs

Samples	Percent of Total Cost: Field		Laboratory		Data Management		Total Cost $/Sample	
	B	E	B	E	B	E	B	E
Chemical/Physical	22	50	39	25	29	25	96	85
Phytoplankton	17	36	27	31	56	33	195	235
Bottom Fish	30	34	-	26	70	40	460	295
Benthos	16	17	34	47	50	36	197	185

Table 4.11. Percentage Contribution of Cost Categories to Cost Discrete Sample Values from Source E

Cost Category	Water Chemistry	Phytoplankton	Fish	Benthic
Boat, equipment and supplies	52	37	29	13
Labor	30	40	52	65
Indirect*	18	14	19	16
Labor	30	49	52	65

* Indirect cost on labor and supplies

For the purposes of this research sampling costs were divided into four components:

1. C_o, overhead costs, including program management, that are fixed and do not vary with the size of the program
2. C_t, costs associated with a sampling occasion, including travel to and from the site, boat rental and all other costs related to the number of sampling occasions
3. C_s, costs per sampling station, including travel between stations and incremental time spent for each station
4. C_r, costs per replicate, including sample collection and laboratory analysis

Using these components, the total cost equation becomes:

$$C = C_o + TC_t + STC_s s + RSTC_r \tag{4.8}$$

where T is the number of sampling occasions, S is the number of stations, and R is the number of replicates.

Data from benthic sampling studies from two separate sampling cost data bases were used to derive estimates of the cost coefficients C_o, C_t, C_s, and C_r. The first set of coefficients, reported in 1982 dollars, is:

C_t = 2,160
C_s = 79
C_r = 460

The second set of coefficients is:

C_t = 986
C_s = 37
C_r = 211

The cost coefficients of the second set are roughly one-half those of the first set; hence, the ratios of $C_t/C_s/C_r$ are about the same for both sets (27.3/1.0/5.8 vs. 26.6/1.0/5.7). It should be noted that benthic sampling costs are expected to be quite different from costs of sampling other aquatic biological populations, such as finfish. Benthic sample collection is not expensive, but laboratory costs are high, because a trained biologist may spend many hours classifying all of the benthic organisms in a single sample. Therefore, one would expect higher C_s values and lower C_r values in other aquatic monitoring programs.

Although the values of the cost coefficients associated with sampling a specific taxon may vary considerably from program to program, it appears that the ratios of these coefficients may remain fairly constant. The relationship of these ratios will determine the nature of the space-time tradeoff in a cost-effective network design.

Short-Term Sampling Network Optimization - Independent Residuals

The purpose of this section is to describe an ANOVA model to apply to aquatic monitoring data that allows for more than one replicate at each sampling occasion. The usual assumption of independent errors is used here; however, this is extended to the dependent case in subsequent sections. A procedure to optimize a sampling network based on this model is then described.

In this report, a sampling network design will be termed "optimal" if it satisfies one of the following two conditions:

1. For a fixed amount of resources, C', that can be allocated to sampling, the network design maximizes the power of detecting a specified impact over the set of all possible network designs whose cost is no more than C'.

2. For a fixed minimum power, B', of detecting a specified impact, the network design minimizes the cost of the design over the set of all possible network designs that yield at least power B'.

In other words, an optimal sampling network maximizes the power for a fixed cost or minimizes the cost for a fixed power.

This report considers only one basic design (Skalski and McKenzie, 1982): two levels of plant status (before operation vs. after), p station pairs, n_1 sampling occasions before plant operation, n_2 sampling occasions after plant operation, and r replicates at each sampling station and occasion.

Also, only two kinds of intervention response functions (IRF) are considered: a step IRF and a linear trend IRF (see Figure 4.14). Thus the set of design networks over which optimization takes place consists of this one design with p, n_1, n_2 and r allowed to vary. This framework is sufficiently flexible to encompass the objectives of most aquatic monitoring programs.

Sampling Design Framework

The models considered by Skalski and McKenzie (1982) assume only one replicate per station and sampling occasion. A simple model that allows for more than one replicate is:

$$Y_{ijkl} = \mu + A_i + B_j + AB_{ij} + C/AB_{(ij)k} + \varepsilon_{ijkl}, \tag{4.9}$$

where

Y_{ijkl} = the lth observation on sampling occasion k at station pair j for plant status at level i

A_i = effect due to plant status at level i (fixed)

B_j = effect due to station pair j (fixed)

AB_{ij} = effect due to plant status x station pair interaction at level ij (fixed)

$C/AB_{(ij)k}$ = effect of sampling occasion k at station pair j for plant status at level i (random)

ε_{ijkl} = effect of replicate l in sampling occasion k at station pair j for plant status at level i (random),

$i = 1,2$

$j = 1,\ldots,p$

$k = 1,\ldots,n$

$n = n_1 = n_2$

$l = 1,\ldots,r$

This model is basically the same as the one given in Table 4.6, except it allows for more than one replicate at each sampling station and occasion.

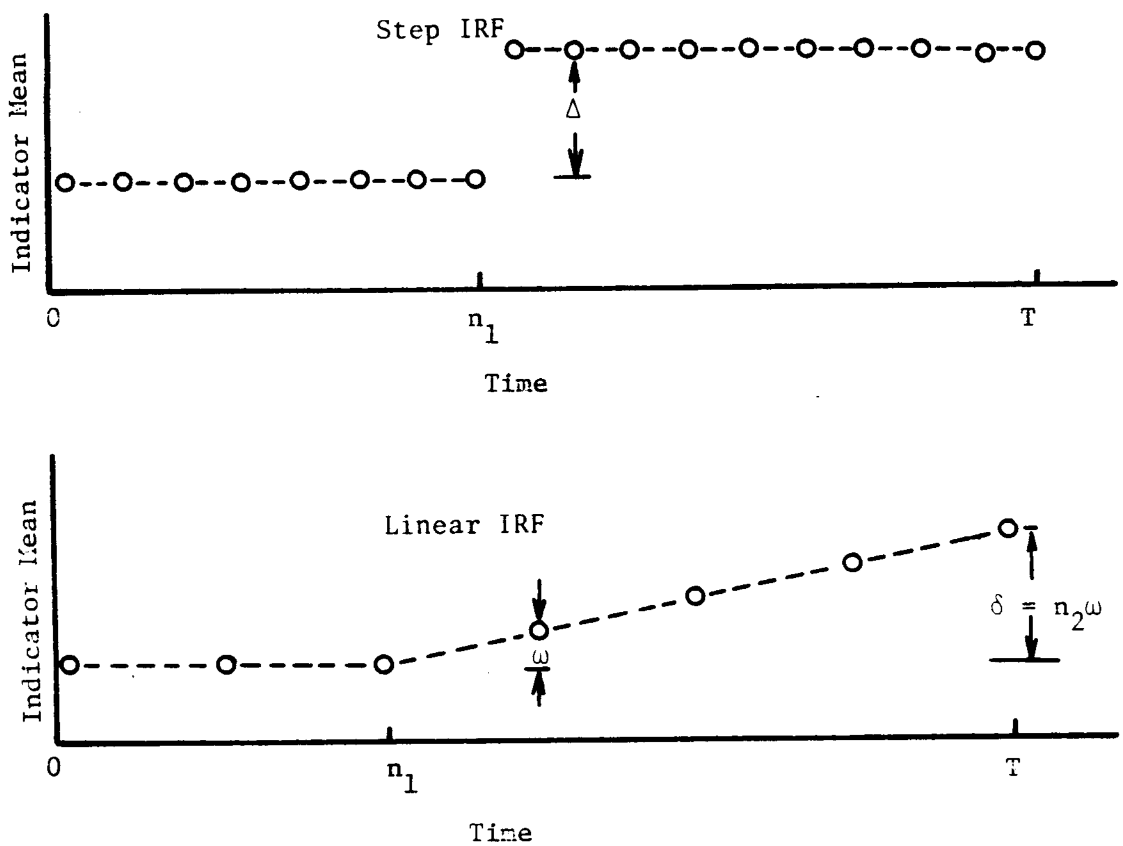

Figure 4.14. Two Kinds of Intervention Response Functions

Furthermore, a balanced design is assumed, that is, an equal number of sampling occasions before and after plant operation ($n_1 = n_2 = n$). Although it is possible to perform ANOVA on unbalanced data sets, balanced designs are highly preferable for two reasons. First, the F-tests are more powerful under a balanced design. Second, under an unbalanced design, the F-tests are not as robust to departures from normality or homoscedasticity (Seber, 1977).

In model (4.9), station pair (factor B) could be considered as a random factor. McKenzie et al. (1977) posit, however, that sampling station locations are chosen in a nonrandom, systematic fashion and therefore, following Winer (1971), station pair should be considered a fixed factor. This convention will be adopted in this report. The expected mean square for each source of variation in the model and the associated degrees of freedom are given in Table 4.12.

Table 4.12. Expected Mean Square and Degrees of Freedom for Factors in Model (4.9)*

Source of Variation	df	EMS
A	1	$\sigma_\varepsilon^2 + r\sigma_{C/AB}^2 + pnr\emptyset(A)$
B	(p-1)	$\sigma_\varepsilon^2 + r\sigma_{C/AB}^2 + 2nr\emptyset(B)$
AB	(p-1)	$\sigma_\varepsilon^2 + r\sigma_{C/AB}^2 + nr\emptyset(AB)$
C/AB	2p(n-1)	$\sigma_\varepsilon^2 + r\sigma_{C/AB}^2$
ε	2pn(r-1)	σ_ε^2

* The notation follows that of Cochran (1951).

In most aquatic sampling programs true replicates will not exist and what is called a replicate here, for a given sampling occasion and station, is actually a measurement taken at a time and location very close to that of the first measurement. For instance, if the program is sampling benthic organisms with an Ekman grab, it would not be wise to take a grab in exactly the same location where the first grab occurred and then call the second grab a replicate. Thus, the component of variance σ_ε^2 is itself composed of two sources of variability: the variance in the observations caused by measurement error and the variance in the observations caused by the small-scale spatial variability inherent in the variable of interest.

The power of any test of effects in an ANOVA model has the general form:

$$\text{Power} = P(F_{\nu_1,\nu_2}(\text{NCP}) > F^{\alpha}_{\nu_1,\nu_2}) \tag{4.10}$$

where

$F_{\nu_1,\nu_2}(\text{NCP})$ is a random variable distributed as a non-central F with ν_1 and ν_2 degrees of freedom, and non-centrality parameter NCP,

and

$F^{\alpha}_{\nu_1,\nu_2}$ is the $(1-\alpha)^{th}$ percent point of the central F distribution with ν_1 and ν_2 degrees of freedom.

For $\nu_1(\nu_2)$ held fixed, the power increases monotonically in $\nu_2(\nu_1)$ and NCP.

For model (4.9), the power of the test of no plant impact (H_0: $A_1 = A_2 = 0$) is given by Eq. 4.10, with ν_1, ν_2 and NCP suitably determined. The form of these three parameters can be calculated from Table 4.14. Thus,

$$\nu_1 = 1,$$
$$\nu_2 = 2p(n-1),$$

$$\text{NCP} = \left[\frac{r}{\sigma_\varepsilon^2 + r\sigma_{C/AB}^2} \cdot \frac{pn}{2}\,\Delta^2\right]^{1/2}$$

where Δ = the overall change in the mean of the variable after plant operation.

The above discussion shows that the power of the test of no plant impact is affected by two sources of variability in the data: σ_ε^2 and $\sigma_{C/AB}^2$. As indicated previously, σ_ε^2 is variability due to measurement error and small-scale spatial variability. $\sigma_{C/AB}^2$, on the other hand, can be described as follows. Model (4.9) assumes that each station pair maintains a constant difference over time within a certain margin of error unrelated to measurement error. $\sigma_{C/AB}^2$ is the variance of this error.

The power of the test of no plant impact is also affected by the magnitude of Δ and the values of n, p, and r. For all other quantities held fixed, NCP is proportional to the absolute value of Δ. Thus, the power increases as the size of the change increases.

The relationship between the power and the values of n, p, and r is more complex. With all other quantities fixed, NCP is proportional to n, but ν_2 is also proportional to n. Thus, increasing n increases the power by increasing both NCP and ν_2. This same argument applies to p. Conversely, the quantity $r/(\sigma_\varepsilon^2 + r\sigma_{C/AB}^2)$ is bounded below by $1/\sigma_{C/AB}^2$. Hence, unless σ_ε^2 is much larger than $\sigma_{C/AB}^2$ (i.e., the measurement error is much larger than the natural variability) increasing r will not greatly affect the value of NCP and thus will not greatly affect the power.

Let $G(\nu_1,\nu_2,NCP,x)$ denote the cumulative distribution function of a non-central F distribution with ν_1 and ν_2 degrees of freedom and non-centrality parameter NCP:

$$G(\nu_1,\nu_2,NCP,x) = P(F\nu_1,\nu_2(NCP) \leq x),$$

and let $G'(\nu_1,\nu_2,NCP,x) = 1 - G(\nu_1,\nu_2\ NCP,x)$. Then the power of the test of no plant impact in this case is given by:

$$\text{Power} = G'(\nu_1,\nu_2,\ NCP,F^{\alpha}_{\nu_1,\nu_2}) \tag{4.11}$$

with ν_1 and ν_2, and NCP as above.

The total number of sampling occasions is 2n, the number of sampling stations is 2p, and the number of replicates per station and sampling occasion is r. Thus, based on the cost equation (4.8) derived in the previous section, the cost of the sampling program is:

$$C = C_0 + 2nC_t + 4npC_s + 4nprC_r. \tag{4.12}$$

The cost function can be written as:

$$C(n,p,r) = 2nC_t + 4npC_s + 4nprC_r. \tag{4.13}$$

For a fixed cost C', a reasonable objective for the monitoring design is to maximize the function:

$$G'(\nu_1,\nu_2,NCP,F^{\alpha}_{\nu_1,\nu_2})$$

subject to the constraint $C(n,p,r) \leq C''$,

where $C'' = C' - C_0$.

For a fixed power B', the objective is to minimize the function:

$$C(n,p,r,)$$

subject to the constraint $G'(\nu_1,\nu_2,NCP,F^{\alpha}_{\nu_1,\nu_2}) \geq B'$.

The details of the optimization algorithm used are described below.

The statistical procedures described above were incorporated into a sampling optimization subroutine in EPAM. As in the other subroutines in EPAM, the optimization procedure is designed to be user friendly and interactive. To execute the optimization in EPAM the user must supply the following inputs: the cost coefficients (C_0, C_t, C_s, and C_r), σ^2_ε, $\sigma^2_{C/AB}$, Δ, and α. The cost coefficients and the variance must be estimated from preoperational or pilot study data. The user also is asked to supply reasonable upper limits on p, n, and r. For example, the user may desire no more than 20 station pairs, sampled at no more than 120 times for each level of plant status, with no more than 5 replicates for each station and sampling occasion. The lower bound on p and r is always 1 and the lower bound on n is always 2.

In the case of a fixed cost design, two additional required inputs are C' (the fixed cost) and the minimum power against Δ a sampling network design must possess in order to be considered as a candidate in the search for the optimal design. The constraint $C'' - C(n,p,r) = 0$ forces the range of r to be:

$$1 \leq r \leq (C'' - 4C_t - 8C_s)/(8C_r).$$

For r fixed at r', the range of p is:

$$1 \leq p \leq (C'' - 4C_t)/(8C_s + 8r'C_r).$$

If r is fixed at r', and p is fixed at p', then n is determined by:

$$n = C''/(2C_t + 4p'C_s + 4p'r'C_r). \tag{4.14}$$

The optimal network design is identified by allowing n, p, and r to range over their limits given above. Designs in which n, p, or r exceed their user-supplied upper bound and designs with power less than the user-supplied minimum power requirement are not considered as candidates in the search.

In the case of a fixed power design, two additional required inputs are B' (the fixed power) and the maximum cost a sampling network design may possess in order to be considered as a candidate in the search for the optimal design. The optimal design is identified by allowing n, p, and r to range over their limits (from 1 to the user-supplied upper bounds in the case of r and p, and from 2 to the user-supplied upper bound in the case of n). Any design in which the cost exceeds the user supplied maximum cost is not considered as a candidate in the search.

Example of Application of Optimization Algorithm

With the use of the second set of cost coefficients derived in the section Analysis of Sampling Cost (C_t = \$986, C_s = \$37, C_r = \$211), the optimization algorithm was applied to a hypothetical case study in which the variance parameters (σ^2_ε and $\sigma^2_{C/AB}$) were estimated from the LACSC benthic data discussed in the Exploratory Data analysis section. Data from eight stations for the years 1978-1979 were employed using as the variable total number of organisms per sample. In order to induce normality and homoscedasticity, the raw data were transformed by taking the natural logarithm (McKenzie, et al., 1977). σ^2_ε and $\sigma^2_{C/AB}$ were estimated from the data to be 0.24 and 0.06, respectively (see Appendix D.4). These estimates, however, are very approximate because there was no specific pairing of control and treatment stations (see Appendix D.4). The parameter Δ was taken to be -0.3567, corresponding to a 30 percent reduction in organism density (Skalski and McKenzie, 1982). The type I error, α, was set to 0.10. The current LACSD benthic sampling program uses three replicates per station and occasion, but four were used in the past, so the upper bound on r was set at four in all cases. Currently, eighteen stations are sampled at each occasion, but as many as forty-four were sampled in the past; therefore, the upper bound on p was varied between nine and twenty-two. Because each station is sampled twice per year, the upper bound on n was varied from four to twelve in steps of two (corresponding to two-to-six years of sampling for each level of plant status).

Table 4.13 presents the results of the optimization for a fixed cost, and Table 4.14 shows those of the optimization for a fixed power. In both, the optimal n increases and the optimal r decreases with an increasing upper bound on n. Similarly, the optimal p increases and the optimal r decreases with an increasing upper bound on p. These results are in accord with those expected, based upon the earlier discussion of the relationship between the power of the test of no plant impact and the quantities n, p, and r.

The effect of the ratio $\sigma_\varepsilon^2/\sigma_{C/AB}^2$ on the optimal number of replicates is summarized in Table 4.15. It is evident that as the ratio of measurement error to natural variability increases, more replicates are required to achieve a given power.

Another variable that affects the optimal number of replicates is obviously C_r, the cost of each replicate. As replicates decrease in cost, the optimal number increases toward the user-supplied upper bound on r. The effects of C_r and the ratio $\sigma_\varepsilon^2/\sigma_{C/AB}^2$ on the optimal number of replicates are summarized in Table 4.16.

The optimal sampling program design remains invariant when the ratio $C'/C_t/C_s/C_r$ stays constant. This relationship may be illustrated by comparing the results in Table 4.17 to those in Table 4.13b. In Table 4.17, C', C_t, C_s and C_r were all doubled compared to their values in Table 4.13b, but nothing else was changed. The results in the two tables are identical except for the values of the total costs of the programs. As already pointed out in the Analysis of Sampling Cost section, the ratios $C_t/C_s/C_r$ were almost exactly the same for the two sets of cost coefficients derived in that section.

These results show that there are four ways to decrease the required number of replicates in an optimal design: (1) increase the number of sampling

Table 4.13. Results of Short-Term Sampling Optimization, Fixed Cost (minimum power required = 0.50)

a. C' = $135,000

Optimal Design

n Upper Bound*	n	p	r	Cost ($)	Power
4	4	9	4	134,752	0.66
6	5	9	3	130,460	0.68
8	8	8	2	133,280	0.70
10	8	8	2	133,280	0.70
12	12	9	1	130,800	0.70

b. C' = $205,000

Optimal Design

n Upper Bound*	n	p	r	Cost ($)	Power
4	4	9	4	134,752	0.66
6	6	9	4	202,128	0.74
8	8	8	3	187,296	0.75
10	10	9	2	184,960	0.76
12	11	9	2	203,456	0.78

c. C' = $205,000

Optimal Design

p Upper Bound**	n	p	r	Cost ($)	Power
9	6	9	4	202,128	0.74
12	6	12	3	204,792	0.77
15	6	12	3	204,792	0.77
18	6	17	2	199,104	0.79
22	5	21	2	202,640	0.79

* p Upper Bound = 9
** n Upper Bound = 6

Table 4.14. Results of Short-Term Sampling Optimization, Fixed Power (maximum cost allowed = $205,000)

a. B' = 0.70

Optimal Design

p Upper Bound**	n	p	r	Cost ($)	Power
4	No network design candidate achieved power 0.70				
6	6	9	3	156,552	0.71
8	8	9	2	147,968	0.72
10	8	9	2	127,968	0.72
12	12	9	1	130,800	0.70

b. B' = 0.70

Optimal Design

p Upper Bound**	n	p	r	Cost ($)	Power
9	6	9	3	156,552	0.71
12	6	11	2	133,008	0.71
15	5	13	2	120,200	0.70
18	6	18	1	118,968	0.70

* p Upper Bound = 9
** n Upper Bound = 6

Table 4.15. Optimal Number of Replicates for Fixed Power Design as a Function of n Upper Bound, p Upper Bound, and $\sigma_\varepsilon^2/\sigma_{C/AB}^2$

a. B' = 0.70

n Upper Bound*	1.0	1.5	2.0	2.5	3.0	3.5	4.0
12	1	1	1	1	1	1	1
10	1	1	1	1	1	2	2
8	1	1	1	2	2	2	2
6	1	1	2	2	2	3	3

$\sigma_\varepsilon^2/\sigma_{C/AB}^2$ +

b. B' = 0.70

p Upper Bound**	1.0	1.5	2.0	2.5	3.0	3.5	4.0
18	1	1	1	1	1	1	1
15	1	1	1	1	1	2	2
12	1	1	1	2	2	2	2
9	1	1	2	2	2	3	3

$\sigma_\varepsilon^2/\sigma_{C/AB}^2$

* p Upper Bound = 9

** n Upper Bound = 6

+ $\sigma^2_{C/AB}$ = 0.06

Table 4.16. Optimal Number of Replicates for Fixed Cost Design as a Function of C_r and $\sigma^2_\varepsilon/\sigma^2_{C/AB}$.*

$C' = 135000$

C_r	$\sigma^2_\varepsilon/\sigma^2_{C/AB}$ = 0.5	1.0	2.0	3.0	4.0
\$218	1	1	1	1	1
\$185	1	1	1	2	2
\$158	1	1	2	2	2
\$105	1	2	2	2	2
\$ 79	2	2	2	2	2
\$ 52	2	3	3	3	3
\$ 26	4	4	4	4	4

* n upper bound = 12

p upper bound = 9

r upper bound = 4

$\sigma^2_\varepsilon/\sigma^2_{C/AB} = 0.06$

Table 4.17. Effects of Doubling Costs on the Fixed Cost Optimization (minimum power required = 0.50)

C' = $410,000

Optimal Design

n Upper Bound*	n	p	r	Cost ($)	Power
4	4	9	4	269,504	0.66
6	6	9	4	404,256	0.74
8	8	8	3	374,592	0.75
10	10	9	2	369,920	0.76
12	11	9	2	406,912	0.78

Compare to Table 4.13.b.

* p Upper Bound = 9

+ C_t = $1972, C_s, = $74, C_r = $422 occasions before and after plant operation, (2) increase the number of station pairs, (3) decrease the variability due to measurement error by using more accurate sampling techniques, or (4) any combination of (1)-(3).

Short-Term Sampling Network Optimization - Correlated Residuals

Sampling Design Framework

The algorithm used for design of a network for detecting short-term change when the residuals are correlated in space and/or time is based on the protocol described in Appendix D.5. It is a general form of the simplified ANOVA model discussed in the previous section, and allows for inclusion of season as a factor, and spatial and temporal correlation in the residuals. The model is similar to that described by Skalski and McKenzie (1982), except that it is couched in a multivariate framework, allowing differential change between station pairs. In addition to seasonal variation, and plant status, factors such as depth, transect location, etc. may be incorporated. As with the use of independent residuals, pilot study data or pre-operational data should be available to determine the accuracy of the control-treatment station pairing scheme and to estimate the necessary variances and correlations.

The non-centrality parameter NCP and the degrees of freedom ν_1 and ν_2 associated with the power of the applicable ANOVA test against the null hypothesis of no plant impact depend on whether season is incorporated as a factor in the ANOVA and whether spatial correlation is present. The values of ν_1, ν_2, and NCP for each of the four possible situations are given in Table 4.18 below. Table 4.19 may be referenced to for definitions of the variables used. A balanced design is assumed, so that $n_1 = n_2 = Sn$. Also, only one replicate per station and sampling occasion is assumed. Therefore, the cost equation (4.8) can be rewritten as:

$$C = C_0 + 2SnC_t + 4pSnC_{sr}. \tag{4.15}$$

where

$$C_{sr} = C_s + C_r \tag{4.16}$$

Table 4.18. Degrees of Freedom and Noncentrality Parameter for ANOVA with Correlated Residuals

Spatial Correlation Absent	1	2	$(NCP)^2$
Seasons Present	1	2pS(n-1)	$\frac{pSn}{2}\frac{\Delta^2}{\sigma^2}$
Seasons Absent	1	2p(Sn-1)	"
Spatial Correlation Present			
Seasons Present	p	2S(n-1)-p+1	$\frac{Sn}{2}\Delta^2(\sum_1^p \sum_1^p \theta_{ij})$
Seasons Absent	p	2(Sn-1)-p+1	"

where θ_{ij} = the $(ij)^{th}$ element of Σ^{-1}.

Table 4.19. Nomenclature Used in Factorial Treatment Design Algorithms

$$C'' = C' - C_0$$

C' = Fixed cost of sampling network
p = Number of station pairs
S = Number of sampling periods per year (number of seasons)
n_1 = Number of samples taken before plant operation
n_2 = Number of samples taken after plant operation
n = Number of years of sampling before (or after) plant operation (assumed equal)
θ_{ij} = ij'th element of inverse residual covariance matrix
NCP = Noncentrality parameter in noncentral F distribution
ν_1 = First degree of freedom in F (or noncentral F) distribution
ν_2 = Second degree of freedom in F (or noncentral F) distribution
n_0 = Number of sampling periods in nonseasonal design before and after plant operation (assumed equal)

If season is not included as a factor in the ANOVA, then n_0 will be used to denote the number of sampling periods for each level of plant status, (n_0 = Sn) so that (4.15) becomes:

$$C = C_0 + 2n_0C_t + 4pn_0C_{sr}. \tag{4.17}$$

The cost function can be written as

$$C_1(p,S,n) = 2SnC_t + 4pSnC_{sr}. \tag{4.18}$$

For a fixed cost C', a reasonable objective for the monitoring design is to maximize the function

$$G'(\nu_1,\nu_2,\text{NCP},F^{\alpha}_{\nu_1,\nu_2}),$$

subject to the constraint $C_1(p,S,n) \leq C''$, where $C'' = C' - C_0$.

For a fixed power B', the objective is to minimize the function

$$C_1(p,S,n),$$

subject to the constraint $G'(\nu_1,\nu_2,NCP,F^{\alpha}_{\nu_1,\nu_2}) \geq B'$. The details of the optimization algorithm used are described for several special cases below.

In the first part of this discussion of the optimization algorithm, it will be assumed that the cost of the sampling program is fixed. In this case, the constraint $C'' - C(p,S,n) = 0$ forces the equality

$$Sn = C''/(2C_t + p4C_{sr}). \tag{4.19}$$

If p, S, and n are all bounded below by 1, then the range of p is given by:

$$1 \leq p \leq (C'' - 2C_t)/4C_{sr}). \tag{4.20}$$

For the case of no spatial correlation, the algorithm requires that the magnitude of the impact and the variance of the differences between control and treatment stations be provided. When season is not included as a factor, the optimal design is identified by allowing p to vary over the limits given by equation (4.20). The results of the algorithm are then the optimal p and n_0.

When season is included as a factor, the optimization routine consists of the following steps:

1. Set the value of p.

2. Vary S and n over all possible combinations that satisfy equation (4.19).

3. Follow steps 1 and 2 for all possible values of p that satisfy equation (4.20).

The outputs of the design algorithm are then the optimal p, S, and n.

When spatial correlation is present, the required inputs are the change magnitude, the total set of 2p' station pair coordinates from which a subset will be chosen, and Σ, the p'xp' variance-covariance matrix of the differences between control-treatment station pairs.

In the subcase when season is not included as a factor, the sampling design algorithm consists of the following steps:

1. Set the value of p.
2. Choose the set of p station pairs (from the given set of candidate p' pairs) that yields the largest value of $(\sum_{1}^{p} \sum_{1}^{p} \theta_{ij})$.
3. Follow steps 1 and 2 for all possible values of p that satisfy equation (4.20).

The outputs from the algorithm are then the optimal p and n_0.

In the case when spatial correlation is present and season is used as a factor, the optimization routine proceeds as follows:

1. Set the value of p.
2. Choose the set of p station pairs (from the given set of candidate p' station pairs) that yields the largest value of $(\sum_{1}^{p} \sum_{1}^{p} \theta_{ij})$.
3. Vary S and n over all possible combinations that satisfy the constraint equation (4.19).
4. Follow steps 1-3 for all possible values of p that satisfy equation (4.20).

The output of the algorithm is the location of the p optimal station pairs, as well as the optimal S and n.

In the second part of this discussion of the optimization algorithm, it will be assumed that the power of the sampling program is fixed. When power, rather than cost, is fixed, the algorithm requires the user to supply reasonable limits on p, S, and n. For example, the user may desire no more than fifty station pairs, sampled at no more than twenty-four times per year, for no more than ten years for each level of plant status.

If spatial correlation is absent, the same inputs are required as for fixed cost. In the special case when season is not used as a factor, the procedure involves the following steps:

1. Set the value of p.

2. Find the smallest value of n_0 that yields at least power B'. If n_0 exceeds its user-prescribed limit, the value of p set in step 1 is discarded as a possibility.

3. Follow steps 1-2 for all possible values of p.

As in the case of fixed cost, the algorithm yields the optimal p and n_0.

When season is used as a factor, the optimization consists of the following steps:

1. Set the value of p.

2. Find the smallest value of Sn that yields at least power B'. If Sn exceeds its user-prescribed limit, the value of p set in step 1 is discarded as a possibility.

3. For a set value of Sn vary S and n over all possible combinations.

4. Follow steps 1-3 for all possible values of p.

Again, as in the case of fixed costs, the outputs are p, S, and n.

Finally, for the fixed power case with spatial correlation, the inputs are identical to the fixed cost case. When season is not used as a factor, the optimization consists of the following steps:

1. Set the value of p.

2. Choose the set of p station pairs from the given set of p' station pairs that yields the largest value of $(\sum_{1}^{p} \sum_{1}^{p} \theta_{ij})$.

3. Find the smallest value of n_0 that yields at least power B'. If n_0 exceeds its user-prescribed limit, the value of p set in step 1 is discarded as a possibility.

4. Follow steps 1-2 for all possible values of p.

The outputs are the optimal p, the locations of the optimal set of station pairs and the optimal n_0.

When seasons are used as factors, the optimization consists of the following steps:

1. Set the value of p.

2. Choose the set of p station pairs from the given set of p' station pairs that yields the largest value of $(\sum_{1}^{p} \sum_{1}^{p} \theta_{ij})$.

3. Find the smallest value of Sn that yields at least power B'. If Sn exceeds its user-prescribed limit, the value of p set in step 1 is discarded as a possibility.

4. For a set value of Sn (n_0), vary S and n over all possible combinations.

5. Follow steps 1-4 for all possible values of p.

The procedure which has been described allows the planner of an aquatic monitoring program to determine an optimal sampling network to detect a given short-term change. If cost is fixed, the optimal design yields maximum power (probability) of detecting the change. If power is fixed, the optimal design yields the minimum cost.

Resource Allocation to Short- Versus Long-Term Sampling

There has been much discussion among specialists in aquatic monitoring design concerning the relative merits of programs designed for short- as opposed to long-term detection of aquatic ecological impacts. Some plant effects are evidenced relatively quickly after first plant operation, and are susceptible to detection by short-term, intensive sampling. Other effects may occur at a relatively low level, or may be delayed in time, so that detection can only be accomplished (if at all) over a long period of time. From the st°andpoint of an electric utility, it is desirable to have sampling programs that can identify both types of change, and, where possible, the same sampling program should be used for both short- and long-term change in the interest of cost-effectiveness. In this section, an algorithm that allows such tradeoffs is described. The methodology employs a factorial design with analysis of variance for short-term change detection, and a modification of Intervention Analysis, a form of Box-Jenkins transfer function modeling (Box and Tiao, 1975) for long-term change detection (See Appendices D.6-D.7). The intervention analysis model allows for multiple control-treatment station pairs, and uses differencing procedures to form a single time series. Table 4.20 may be referenced for definitions of the variables used.

Cost Equation

The total cost of a sampling network is given by:

$$C = C_O + C_A + C_I, \tag{4.21}$$

where

C_O = overhead cost,

C_A = cost due to the design to detect short-term impact (ANOVA design),

C_I = additional cost due to the design to detect long-term impact (IA design).

Table 4.20. Nomenclature Used in Short- Versus Long-Term Tradeoff Design Algorithms

C' = Fixed cost of sampling network

$C'' = C' - C_0$

p_A = Number of station pairs for ANOVA design

p_I = Number of station pairs for IA design

S_A = Number of sampling periods (seasons) per year for ANOVA design

S_I = Number of sampling periods (seasons) per year for IA design

n = Number of years of sampling before (or after) plant operation for ANOVA design

n_a = Number of years of sampling after plant operation for IA design

n_1 = Total number of sampling periods before plant operation

n_2 = Total number of sampling periods after plant operation

p' = Number of station pairs in the set of station pairs from which the network will be constructed

Δ = Short-term step change

δ = Long-term total change

σ^2 = Variance of the differences between observations at control-treatment station pairs

Σ = $p' x p'$ variance-covariance matrix of the differences between control-treatment station pairs

$\emptyset$ = Lag-one auto-correlation for the differences between observations at control-treatment station pairs

B' = Minimum power required to detect a short-term step change

The cost associated with the ANOVA design (ignoring overhead) is given by:

$$C_A = 2S_A n C_t + 4p_A S_A n C_{sr}. \tag{4.22}$$

The additional cost of the IA design is given by:

$$C_I = S_I(n_a - n)C_t + 2p_I S_I(n_a - n)C_{sr}, \tag{4.23}$$

assuming $n_a > n$.

If season is not a factor, equation (4.22) can be rewritten as

$$C_A = 2n_1 C_t + 4p_A n_1 C_{sr}, \tag{4.24}$$

and equation (4.23) can be rewritten as

$$C_I = (n_2 - n_1)C_t + 2p_I(n_2 - n_1)C_{sr}. \tag{4.25}$$

Optimization Algorithm

The optimization algorithm described below assumes a fixed cost, C', for the sampling network.

When spatial correlation is absent, the user must supply Δ (the short-term step change), δ (the long-term total change), σ^2, $\emptyset$, and B' (the minimum power required to detect a short-term step change Δ). The first step taken is to find the minimum cost ANOVA design that yields power B' with respect to the step change Δ. This step is accomplished via the algorithm already described in the section on optimizing sampling networks for short-term change. The outcome of this step is the determination of the optimal C_A, p_A, S_A, and n (or n_1 in place of $S_A n$ if season is not a factor). The amount of funds left to spend on the IA design is ($C'' - C_A$). The second step of the algorithm finds the IA design that maximizes the power of detecting a long-term change given the constraints $C_I = (C' - C_A)$, $p_I \leq p_A$, and n_1 fixed at the value determined in the first step. The output for this step is the optimal p_I, S_I, and n_a (or n_2 in place of $S_I n_a$ if season is not a factor).

When spatial correlation is present, the user must supply the set of 2p' station pair coordinates from which a subset will be chosen, as well as Δ, δ, Σ, $\emptyset$, and B'. The optimization algorithm then proceeds in the same manner as when spatial correlation is absent. The output is the optimal $(A, p_A$, the optimal set of p_A station pairs, the optimal S_A, n (or n_1 in place of $S_A n$ if season is not a factor), the optimal p_I, S_I, and n_a (or n_2 in place of $S_I n_a$ if season is not a factor).

A procedure has been described that allows the planner of an aquatic monitoring program to determine optimal space-time tradeoffs in the sampling network for given levels of short-term and long-term change he or she wishes to detect with a given level of probability.

Use of Simulation Models in Sampling Design

One of the issues addressed in this project is the degree to which modeling can act as a substitute for field monitoring. While this investigation was conducted at a relatively low level in terms of project resources, some important progress has been made and is reported here. The approach taken was to apply the methods of optimal estimation to determine how much can be gained by substituting increasingly complex models in place of space/time resolution of a sampling program. The problem can be viewed as having three dimensions: model complexity, and spatial and temporal distribution of sampling.

It should be noted that the preceding three sections do not consider the model complexity dimension, i.e., they effectively assume that all information about the system must be gained through sampling, as opposed to modeling. This assumption is relaxed in this section. As in the field-oriented designs discussed earlier, the efficacy of monitoring program design is to be evaluated in terms of how well the design detects the impact resulting from some disturbance.

The objective of this section, therefore, is to develop a methodology that is appropriate for a wide variety of ecosystems. For the purposes of this development, however, the emphasis will be on benthic and nektonic communities affected by thermal power plants. The reasons for this choice are the sophistication of available models, and the importance of the species selected in environmental assessment programs.

Because a great deal of work has been devoted to the analysis of these ecosystem types, state-of-the-art models for these systems are fairly complex. One such model was developed for EPRI by Tetra Tech (1979), and was chosen for the purposes of the experiments reported here. It has been assumed that this model represents the real world after the variability of forcing functions and residual error have been included. Monte Carlo simulations, using this complex model, provided the data with which to evaluate the importance of model complexity. Methods developed for the analysis of other large-scale, complex systems, such as power grids and economic systems, were applied to the prototype to find levels of aggregation which describe appropriate levels of system response. Hypothesis testing methods, based on the likelihood function, and which make use of optimal estimation, evaluated the models of different complexity given a range of reasonable space/time sampling schemes.

State Estimation

The physical, chemical and biological dynamics of aquatic ecosystems can be described in terms of a conservation equation:

$$\frac{\delta \underline{C}}{\delta t} + u.\nabla\underline{C} = \nabla.\varepsilon\nabla\underline{C} + [\kappa]\ \underline{C} + \theta \qquad (4.26)$$

where

$\underline{C}$ = the nx1 vector of constituent values,

u = the velocity vector for the x,y,z, directions,

ε = the diffusion coefficient,

$[\kappa]$ = an nxn matrix of coefficients describing the system kinetics,

θ = a source (sink) term.

Finite difference or finite element methods allow transformation of this equation into state-space form. Incorporation of process model uncertainty and measurement model uncertainty into the state-space formulation makes it possible to design filtering schemes for weighting state estimates produced by the model and state estimates derived from the data. The process model becomes

$$X(j+1) = \underline{\emptyset X}(j) + \Lambda\underline{U}(j) + \Gamma\underline{W}(j) \qquad (4.27)$$

and the measurement model

$$\underline{Z}(j) = \underline{HX}(j) + \underline{V}(j) \qquad (4.28)$$

where

$Z(j)$ = the measured value of the state variable at the jth time increment,

$\underline{H}$ = a matrix for the relationship between the state variables and the measured value,

$U(j)$ = a forcing function

$V(j)$ = a white noise process, distributed as $N(0,[\underline{R}_j])$.

Λ = a coefficient matrix for the forcing function

Γ = a coefficient matrix for the white noise process

The filter design is accomplished by specifying a cost function associated with errors in state estimates, where errors result from differences between the estimated value of the state and the true value. In the case of the widely used Kalman filter, this cost function is the sum of the squared error and optimal estimates obtained by minimizing this function. In addition, the error is constrained to be unbiased, leading to the well-known relations for the filter gain, $[\kappa]$; the estimates of the state, $[X(n+1|n+1)]$; and the corresponding covariance matrix, $[\Sigma(n+1\pi n+1)]$ (Gelb, 1974).

Hypothesis Testing

The process and measurement models (Eqs. 4.27 and 4.28) lead to estimates of the state and the covariances of the state; however, the evaluation of impact requires a formalism for incorporating this information into hypothesis testing. Typically, hypothesis testing is done by establishing a set of hypotheses and evaluating some decision function in light of a pre-established criterion or decision rule.

For assessing environmental impact, the null hypothesis is

H_0: there is no environmental impact

as opposed to the hypothesis

H_1: there is an environmental impact.

The null hypothesis H_0 states that the model for no impact

$$\underline{\dot{X}}_1 = \underline{\theta X}_1 + \underline{\Gamma W} \tag{4.29}$$

is true, while the hypothesis H_1 states that the model of the observable impact is true

$$X_2 = \emptyset \underline{X}_2 + \Lambda \underline{U} + \underline{\Gamma W} \tag{4.30}$$

where the matrices $[\underline{\theta}]$ and $[\underline{\Gamma}]$ have the same elements in both equations (4.29) and (4.30). When either H_0 or H_1 are true, the likelihood functions ζ_1 and ζ_2, according to Schweppe (1973), are normally distributed:

$$\zeta\ell(n,j) \sim N(\mu_{k\ell}(n,j),\ \mu_{\ell\ell}n,j) \tag{4.31}$$

$$\mu_{k\ell}(n,j) = \sum_{n=1}^{N} \sum_{j=1}^{M} d_K(n,j)[\Sigma(n|n-1) + R(n)]^{-1}.d_{\ell}(n,j) \tag{4.32}$$

where the d_k are developed from the knowledge of the deterministic mean value of the state and by using the optimal filter equations appropriate to 4.27 and 4.28. Specifying decision criteria requires the use of the following decision rules for this model:

1. Choose H_0 if $\zeta_1 > \Delta^*$
   and $\zeta_2 < \Delta^*$

2. Choose H_1 if $\zeta_1 < \Delta^*$
   and $\zeta_2 > \Delta^*$

3. No decision if $\zeta_1 < \Delta^*$
   and $\zeta_2 > \Delta^*$

4. No decision if $\zeta_1 < \Delta^*$
   and $\zeta_2 > \Delta^*$

where

Δ^* = a decision criterion chosen to obtain the proper balance between the probability of detecting an impact when there is none and the probability of finding no impact when there really is one.

Implementation of this algorithm allows for the reduction in estimation uncertainty achievable by use of a model. A number of studies have examined this issue with the time and space distribution of samples as the dimensions of the problem, but only Moss (1980) has added the dimension of model accuracy. His analysis was one of hydrological parameters in a system of river basins.

Experiments currently in progress are testing this algorithm on a hypothetical power plant cooling lake, using a modification of the Tetra Tech lake ecosystem model as the prototype. Also being employed are three simplifications of this model, using several techniques for aggregating with respect to number of species and temporal and spatial resolution.

CHAPTER 5
CONCLUSION

The complexity of aquatic monitoring program design for electric power plant impact assessment makes development of a comprehensive framework a challenging task. Previous efforts, such as those of States et al. (1978) and Fritz et al. (1980), have made notable progress in this direction, but have not been fully successful. In this research, an algorithmic framework, termed Electric Power Aquatic Monitoring (EPAM) has been developed to provide a comprehensive set of guidelines, making use of interactive computer programming technology, to guide aquatic monitoring design. EPAM moves beyond the previous formats in several important respects:

1. It involves expert judgment in a unique manner and with greater efficiency than its predecessors.

2. EPAM represents a hierarchical framework aimed at monitoring program designers having different levels of knowledge about the system under consideration. This approach is well-suited to interactive computer manipulation, and the EPAM framework and many of the special-purpose algorithms have been programmed for this application. Versions are being developed for both main frame and microcomputer systems in FORTRAN and BASIC languages.

3. EPAM gives specific and detailed attention to certain statistical problems that have limited the success of aquatic monitoring programs, particularly the questions of spatial and temporal correlation in the residuals when factorial treatment designs are applied. This effort has extended to a consideration of monitoring program optimization for cost-effectiveness.

4. EPAM presents specific algorithms to aid the monitoring program designer in identifying IMPACT AGENTS, TARGETS, and HYPOTHESES OF IMPACT.

5. EPAM has been evaluated using a case history from aquatic monitoring records. Further case history evaluation will continue in Phase III.

The major findings of the Phase II research in these five areas and the conclusions drawn are discussed in the remainder of this section.

The complexity of aquatic ecosystems affected by thermal electric power generation necessitates the use of experts in many aspects of monitoring program design. This research has demonstrated that, to an extent, expert judgment can be built into a generalized framework to consider monitoring issues and to arrive at definite program elements. Use of the Delphi process for the purpose of exploring general consensus is feasible, although time-consuming, but convening multidisciplinary panels for the design of each major monitoring program, as suggested by Fritz et al. (1980) is probably not feasible.

Even with a general conceptual framework incorporating expert judgment, such as EPAM, expert opinion may still be required to make site-specific decisions with respect to individual monitoring programs. EPAM facilitates incorporating expert judgment at this level through use of a mathematical procedure (CHOICE) that prioritizes selections relative to the achievement of multiple monitoring objectives.

A hierarchical structure to guide monitoring program design is more appropriate, and potentially useful in more situations, than a single-tier approach, because electric utilities face the need to conduct monitoring in widely differing circumstances. These situations range, on the one hand, from a new plant on a poorly studied water body to an existing plant with a monitoring program of long duration where optimizing just one of the sampling efforts is desired. Since many utilities have large data bases that can be employed in planning future monitoring at the same site, it is advantageous to have a planning framework that can be entered at the level equivalent to the existing knowledge. Others may wish to design a monitoring program on paper at low cost using very minimal data. EPAM will accommodate such a user and provide the opportunity for upgrading as additional knowledge becomes available.

Relative to statistical design, standard procedures in the analysis of aquatic monitoring data assume uncorrelated errors. This assumption is critical and violation of it requires modified procedures. This research has explored this point and has concluded the following:

1. Temporal correlation in the errors in aquatic data appears to be negligible for monthly and bi-monthly sampling periods.

2. Spatial correlation in the errors in aquatic data is significant for space scales important to aquatic monitoring and must be considered in most cases, particularly for suspended and mobile organisms. Fixed organisms, such as benthos, tend to exhibit less spatial correlation and may be treated adequately using traditional methods that employ an assumption of independent residuals.

3. In the application of an aquatic monitoring optimization program that assumes uncorrelated sampling residuals, one replicate for each sampling station and sampling occasion resulted in the most cost-effective program. This conclusion was shown to be true for a wide range of sample variability and sampling cost ratios.

4. For those situations in which spatially correlated errors must be considered, a multivariate framework has been developed. The application of this framework to actual monitoring data will be evaluated in Phase III.

5. The effectiveness of an aquatic monitoring program in detecting both short and long-term effects should be considered in the design stage. Studies of short-term effects usually require spatially intensive sampling, whereas detection of long-term effects requires a record of extended length. An algorithm has been proposed for performing this tradeoff, using a factorial treatment design for short-term assessment and a control-treatment station pairing appropriate to time series analysis for long-term assessment.

6. The modeling of biological interactions and dynamics may substantially reduce the magnitude of the monitoring effort by providing insights into optimal location and frequency of sampling.

A partial limitation to the application of the procedures developed for sampling design at this time is that the structure of the residuals, which reflect both natural variability and measurement error, must be known. Measurement errors depend largely on the sampling techniques employed. Because this area has been poorly documented and is beyond the scope of this research, maximum effectiveness in using the knowledge gained must await research on this question.

LEVELS 2 and 3 of EPAM contain a number of rather specific procedures to aid a monitoring program designer in identifying and ranking basic elements of the program. Most of these factors have been considered in preceding frameworks, but the present effort is distinguished in the power plant aquatic monitoring field by its comprehensive outlook. States et al. (1978) discussed impact agents and recommended ranking them in importance. These investigators also gave attention to the vulnerability of ecosystem components to impact, hypothesis development and ranking, and selection of variables. Fritz et al. (1980) also defined and ranked impact agents as a primary underlying element of monitoring program design. They went on to consider interaction between affected species and the power plant and the central nature of hypotheses to a monitoring program. The present research has adopted and augmented these concepts and developed specific procedures for conducting analyses and arriving at decisions.

Although evaluation of EPAM by case study has barely begun at this stage, it will be the central component of the next phase of the project. Experience to date has indicated that verification of performance will make a major contribution to the conceptual model and provide an important incentive for its eventual use. There is no substitute for applying a process in the arena for which it has been developed to determine its utility. While the major components were set, the San Onofre case study revealed many specific features that required modification to close gaps in logic, to permit useful data to be employed, or to provide a means of proceeding where data were not as complete as may have been desired. Phase III case studies are expected to serve in the same fashion to evaluate the model's performance in different circumstances and to aid in 'tuning up' the algorithms.

REFERENCES

Anderson, T.W. 1958. An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York, NY.

Bakus, G.J., W.G. Stillwell, S.M. Latter, and M.C. Wallerstein. 1982. Decision making: with applications for environmental management. Environmental Management 6(6): 493-504.

Bickel, P.J. and K.A. Doksum. 1977. Mathematical Statistics. Holden-Day, Inc., San Francisco, CA.

Box, G.E.P. 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems. Annals of Mathematical Statistics 25: 290-302, 484-498.

Box, G.E.P., and G.M. Jenkins. 1976. Time Series Analysis: Forecasting and Control. Revised Edition, Holden-Day, Inc., San Francisco, CA.

Box, G.E.P. and Tiao, G.C. 1975. Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association 70: 70-79.

Cochran, W.G. 1947. Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics 3: 22-38.

Cochran, W.G. 1951. Testing a linear relation among variances. Biometrics 7: 17-32.

Cohon, J.L. 1978. Multiobjective Programming and Planning. Academic Press, New York, NY.

Comiskey, C.E. and C. Brandt. 1982. Appendix A. Quantitative Impact Assessment. In: J.D. Allen (ed.), Marine Ecosystem Monitoring Task Group, Ecology Committee, Science Advisory Board, U.S. Environmental Protection Agency, Washington, D.C.

Delbecq, A.L., A.H. Van de Ven, and D.H. Gustafson. 1975. Group Techniques for Programs Planning. Scott Foresman and Co. Glenview IL.

DeMartini, E.E. 1980. Report of the Marine Review Committee to the California Coastal Commission: Predictions of the Effects of the San Onofre Nuclear Generating Station and Recommendations. Part I: Recommendations, Predictions, and Rationale (including Technical Appendices), MRC Document 80-04 (I). Marine Review Committee, Solana Beach, CA.

DeMartini, E.E. 1981. January 1 - December 31, 1980, UCSB Fish Program Annual Report, Volume I, Summary of Findings: Preoperational Monitoring Program (Fish Studies), San Onofre Nuclear Generating Station. Marine Science Institute, University of California, Santa Barbara, CA.

Dietz, J. 1982. Personal communication. San Diego Gas and Electric Company, San Diego, CA.

Eisenhart, B. 1947. The assumptions underlying the analysis of variance. Biometrics 3: 1-21.

Feder, H.M., and C.H. Turner. 1974. Observations on Fishes Associated with Kelp Beds in Southern California, Fish Bulletin 160. California Department of Fish and Game, Sacramento, CA.

Fishhoff, B.S. Lichtenstein, P. Slovic, S. Derby, and R. Keeney. 1981. Acceptable Risk. Cambridge University Press, New York, NY.

Fritz, E.S., P.J. Rago, and I.P. Muraka. 1980. Strategy for Assessing Impacts of Power Plants on Fish and Shellfish Populations. Power Plant Project, Office of Biological Services, Fish and Wildlife Service. FWS/OBS-80/34.

Garrison, W.E. 1981. Ocean Monitoring and Research: Annual Report 1980-1981. Los Angeles County Sanitation District, Whittier, CA.

Gelb, A. (ed.) 1974. Applied Optimal Estimation. MIT Press, Cambridge, MA.

Gilliland, M.W. and P.G. Risser. 1977. The use of systems diagrams for environmental impact assessment: procedures and an application. Ecological Modeling 3:183-209

Glass, G.V. P.D. Peckham, and J.R. Sanders. 1972. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. Reviews in Educational Research 42: 237-288.

Goicoechea, A., D.R. Hansen, and L. Duckstein. 1982. Multiobjective Decision Analysis with Engineering and Business Applications. John Wiley and Sons, New York, NY.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. Wiley-Interscience, New York, NY.

Haimes, Y.Y., and W.A. Hall. 1974. Multiobjectives of water resources system analysis, the surrogate worth trade-off method. Water Resources Research 10(4): 615-624.

Holling, C.S. (ed.) 1978. Adaptive Environmental Assessment and Management. Wiley International Institute for Applied Systems Analysis, International Series on Applied Systems Analysis, No. 3. John Wiley and Sons, New York, NY.

Ignizio, J.P. 1981. The determination of a subset of efficient solutions via goal programming. Computers and Operations Research 3: 9-16.

Jeffers, J.N.R. 1978. Design of Experiments. Institute of Terrestrial Ecology. National Environment Research Council. England.

Jenkins, G.M., and D.G. Watts. 1968. Spectral Analysis and Its Applications. Holden-Day, Inc., San Francisco, CA.

Keeny, R.L., and J. Raiffa. 1976. Decisions with Multiple Objectives: References and Value Tradeoffs. John Wiley and Sons, New York, NY.

Knuth, D.E. 1981. The Art of Computer Programming, Vol. 2: Seminumerical Algorithms. Addison-Wesley Publishing Company, Inc., Reading, MA.

Lettenmaier, D.P. and L.C. Murray. 1977. Design of Nonradiological Aquatic Sampling Programs for Nuclear Power Plant Impact Assessment Using Intervention Analysis. U.S. Nuclear Regulatory Commission, Technical Report UW-NRC-6.

Linstone, H.A. and M. Turoff (eds.). 1975. The Delphi Method. Techniques and Applications. Addison-Wesley Publishing Company, Inc., Reading, MA.

Lockheed Center for Marine Research. 1977. Marine Biofouling Control Studies, San Onofre Nuclear Generating Station, Final Report. Lockheed Center for Marine Research, Carlsbad, CA.

Lockheed Center for Marine Research. 1978. San Onofre Nuclear Generating Station Units 2 and 3 Prototype Study for Heat Treatment Procedures, Final Report. Lockheed Center for Marine Research, Carlsbad, CA.

Los Angeles Department of Water and Power. 1980. Power Supply for Los Angeles, DWP 20M 8-80. Los Angeles Department of Water and Power, Los Angeles, CA.

Mattson, G.G. 1982. Analysis of Aquatic Monitoring Costs. Non-thesis Master of Civil Engineering paper. Department of Civil Engineering, University of Washington, Seattle, WA.

McGroddy, P.M., L.E. Larson, and D.R. Deneen. 1979. Physical and Hydraulic Description of Generating Station Intakes for Southern California Edison Company, 79-RD-63. Southern California Edison Company, Rosemead, CA.

McKenzie, D.H., Arnold, E.M., Skalski, J.R., Fickeisen, D.H., and Baker, K.S. 1979. A Quantitative Assessment of Aquatic Impacts of Power Plants. PNL-2891. Battelle, Pacific Northwest Laboratories, Richland, WA.

McKenzie, D.H., L.D. Kanberg, K.L. Gore, E.M. Arnold, and D.G. Watson. 1977. Design and Analysis of Aquatic Monitoring Programs at Nuclear Power Plants. PNL-2423. Battelle Pacific Northwest Laboratories, Richland, WA.

Meyers, C.D. and K.E. Bremer (eds.). 1975. A Statement of Concerns and Suggested Ecological Research. Report No. 1 of the Lake Michigan Cooling Water Studies Panel. U.S. Environmental Protection Agency, Region V, Chicago, IL, EPA-905/3-75-001.

Morrison, D.F. 1976 Multivariate Statistical Methods. McGraw-Hill Inc., New York, NY.

Moss, M.E. 1979. Space, time, and the third dimension (model error). Water Resources Research 15(6): 1797-1800.

Munn, R.E. (ed.). 1975. Environmental Impact Assessment: Principles and Procedures. SCOPE Workshop on Impact Studies in the Environment (WISE), United Nations Environmental Program (UNEP), Environment Canada and UNESCO; SCOPE Report No. 5, Toronto, Canada.

Murarka, I.P. 1976. An Evaluation of Environmental Data Relating to Selected Nuclear Power Plant Sites: the Nine Mile Point Nuclear Power Station Site. Argonne National Laboratory ANL/EIS-7. Argonne, IL.

Murarka, I.P., A. Policastro, E. Daniels, J. Ferante, and F. Vaslow. 1976. An Evaluation of Environmental Data Relating to Selected Nuclear Power Plant Sites: the Zion Nuclear Power Plant Site. Argonne National Laboratory ANL/EIS-5. Argonne, IL.

Palmer, J.B. 1983. Personal communication. Southern California Edison Company, Rosemead, CA.

Platt, J.R. 1964. Strong inference. Science 146(3642): 347-353.

Proctor, C.M., J.C. Garcia, D.V. Galvin, T. Joyner, G.B. Lewis, L.C. Loehr, A.M. Massa, and Ryckman, Edgerley, Tomlinson and Associates, Inc. 1980. An Ecological Characterization of the Pacific Northwest Coastal Region. Volume 1: Conceptual Model. U.S. Fish and Wildlife Service, Washington, D.C., FWS/OBS-79/11.

Rago, P.J., E.S. Fritz and I.P. Murarka. 1983. Assessing Impacts of Power Plants on Fish Populations: A General Strategy. Environmental Monitoring and Assessment 3(2): 185-201.

Saaty, T. 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15: 234-281.

Schweppe, F.C. 1973. Uncertain Dynamic Systems. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Seber, G.A.F. 1977. Linear Regression Analysis. John Wiley and Sons, New York, NY.

Silvey, S.D. 1975. Statistical Inference. Chapman and Hall, London, England.

Skalski, J.R. and D.H. McKenzie. 1982. A design for aquatic monitoring programs. Journal of Environmental Management 14: 237-251.

Southern California Edison Company. 1982. 1981 Annual Report, Marine Environmental Analysis and Interpretation, San Onofre Nuclear Generating Station. Southern California Edison Company, Rosemead, CA.

Southern California Edison Company. 1983a. Report on 1982 Data, Marine Environmental Analysis and Interpretation, San Onofre Nuclear Generating Station. Southern California Edison Company, Rosemead, CA.

Southern California Edison Company. 1983b. San Onofre Nuclear Generating Station Unit 1, 316(b) Demonstration. Southern California Edison Company, Rosemead, CA.

States, J.B. 1983. Personal communication. Battelle Pacific Northwest Laboratory. Anchorage, AK.

States, J.B., P.T. Huag, T.G. Shoemaker, L.W. Reed, and E.B. Reed. 1978. A Systems Approach to Ecological Baseline Studies. U.S. Fish and Wildlife Service, Washington, D.C., FWS/OBS-78/21.

Stock, J.N. and R.A. De La Parra. Undated. Use of a Thermal Backwash to Control Marine Biofouling at San Onofre Nuclear Generating Station. Internal Report. Southern California Edison Company, Rosemead, CA.

Tetra Tech, Inc. 1979. Methodology for Evaluation of Multiple Power Plant Cooling System Effects, Vol. 1-3 Report to Electric Power Research Institute, Tetra Tech, Inc., Lafayette Hills, CA.

Thomas, G.L., L. Johnson, R.E. Thorne, and W.C. Acker. 1980a. A Comparison of Fish Entrapment at Four Southern California Edison Company Cooling Water Intake Systems, FRI-UW-8023. Fisheries Research Institute, University of Washington, Seattle, WA.

Thomas, G.L., R.E, Thorne, W.C. Acker, T.B. Stables, and A.S. Kolok. 1980b. The Effectiveness of a Velocity Cap and Decreased Flow in Reducing Fish Entrapment, FRI-UW-8027. Fisheries Research Institute, University of Washington, Seattle, WA.

Tukey, J.W. 1977. Exploratory Data Analysis. Addison-Wesley Publishing Company, Inc., Reading, MA.

U.S. Environmental Protection Agency. 1977a. Interagency 316(a) Technical Guidance Manual and Guide for Thermal Effects Sections of Nuclear Facilities Environmental Impact Statements. U.S.E.P.A., Office of Water Enforcement, Permits Division, Washington, D.C.

U.S. Environmental Protection Agency. 1977b. Guidance for Evaluating the Adverse Impact of Cooling Water Intake Structures on the Aquatic Environment: Section 316(b) P.L. 92-500. U.S.E.P.A., Office of Water Enforcement, Permits Division, Washington, D.C.

U.S. Nuclear Regulatory Commission. 1976. Preparation of Environmental Reports for Nuclear Power Stations. Regulatory Guide 4.2, Revision 2. NUREG-0099. U.S.N.R.C., Office of Standards and Development, Washington, D.C.

Winer, B.J. 1971. Statistical Principles in Experimental Design, 2nd Ed., McGraw-Hill Book Company, New York, NY.