STATISTICAL AND ANALYTICAL
SUPPORT CONTRACT:
FINAL REPORT
(DELIVERABLE 4)

VOLUME I

Prepared for

Water Quality Data Analysis
Working Group
(Chesapeake Bay Program)
410 Severn Avenue
Annapolis, MD 21405

Prepared by

Martin Marietta Environmental Systems
9200 Rumsey Road
Columbia, MD 21045

February 1987

ACKNOWLEDGEMENTS

This report was prepared by Kenneth A. Rose, A. Fred Holland, Harold T. Wilson, and Richard A. Cummins. The authors wish to thank the members of the Analytical Working Group for their support and comments throughout this project.

EXECUTIVE SUMMARY

Under contract X-003321-02 with the Chesapeake Research Consortium (CRC), Martin Marietta Environmental Systems was contracted to provide statistical and analytical support to the Monitoring Subcommittee of the Chesapeake Bay Program (CBP). All tasks performed focused on data processing and analysis using data collected at mainstem stations between June 1984 (initiation of program) and September 1985 as part of the CBP water quality monitoring program and historical Chesapeake Bay data. The results from all tasks performed under this contract were grouped into three interim (or working) documents (referred to as Deliverables 1-3), and a final report (Deliverable 4). This report is Deliverable 4. Deliverables 1-3 are included with this final report as Appendices A-C in a companion volume.

The primary focus of this contract was the development of a statistical analysis framework for detection of trends in Chesapeake Bay water quality attributable to pollution-control management actions. In Chapter II, a procedure for selecting among the many possible statistical methods appropriate for trend analysis of water qualty data is presented. These methods are reviewed and summarized in Appendices C and D of Volume II of this report. The analysis selection procedure is based on the characteristics of the data being analyzed and is applicable to both historical and CBP water quality monitoring data. Based on the proposed analysis framework and graphical and tabular analyses of the CBP monitoring data (Appendices A and B), a preliminary evaluation of the sampling design of the CBP monitoring program is provided in Chapter III. In Chapter IV, the quality assurance/quality control (QA/QC) procedures we applied to CBP monitoring data prior to analyses are described.

Throughout Chapters II-IV, recommendations on various aspects of collection and analysis of Chesapeake Bay water quality monitoring program data are made. These recommendations are designed to improve the ability of the program to detect and understand changes in Chesapeake Bay water quality. Many of these recommendations propose additional work to address outstanding issues on data collection and analysis. Many of the proposed approaches to address these issues involve small to moderate expenditures of effort and, as much as possible, utilize existing data. These recommendations, described in detail in Chapters II-IV, are briefly summarized below, grouped together under the general topics of data analysis and sampling design.

Data Analysis

- Additional refinements of the proposed analysis framework should be performed. These refinements should include the fine-tuning of the analysis framework using three or more years of CBP monitoring data, further evaluation of multivariate statistical techniques for their applicability for characterizing spatial and temporal patterns, the assembly and statistical analysis of selected sampling stations with long-term time series of water quality data, and the development of methods for characterizing within summer variability of the duration and extent of hypoxia and stratification intensity

- Main Bay stations located in vicinity of tributaries should not be included in analyses characterizing mainstem water quality but should be analyzed independently with upstream tributary data to determine the effects of tributary inputs on mainstem water quality

- QA/QC procedures, including estimation of error rates, should be applied to data (both new and historical) before they are incorporated into the CBP database to ensure data are acceptable for public dissimination. Specific problems with the existing CBP historical database (e.g., incorrect depths associated with water quality variables) should be corrected to allow extension of historical time series with data from the present CBP water quality monitoring program and permit rigorous evaluation of the utility of correlational analysis on pooled data. The historical database should also be updated to include existing data not presently in the database. Emphasis should be placed on DO and salinity data collected at multiple depths during multiple cruises within summers to allow historical within summer variability in the duration and extent of low DO water to be determined

- Information on exchange rates among dissolved and particulate nutrients, chlorophyll-a, and DO for both the water column and the sediments should be incorporated into analyses of CBP water quality monitoring data.

Sampling Design

- Rigorous evaluation of the CBP water quality monitoring program design, including the conduct of power and sensitivity analyses, should be accomplished as soon as estimates of year-to-year variability in water quality parameters for the present program are available (i.e., 3-4 years of data). This evaluation should include comparison of data from the CBP monitoring program to data collected with alternative sampling strategies (e.g., random selection of station locations; integrated pump samples; continuous monitoring). These comparisons will allow assessment of the degree and magnitude of the biases associated with CBP collected data and confirmation of the representativeness of these data for characterizing Bay water quality in regions, seasons, and depth layers.

- As is presently being done, measurements of temperature, salinity, conductivity, DO, and pH should be taken at 2 m depth intervals for the above and below pycnocline layers, and DO and salinity measurements should be taken at 1 m depth intervals within the pycnocline for the central region of the Bay to as accurately as possible define DO and salinity isopleths. Existing data should be used to determine whether one grab sample is sufficient for characterization of nutrients and chlorophyll-a below the pycnocline at lateral stations

- Data generators, to the degree possible, should be required to use similar data collection and measurement techniques. Identical measurement methods should be required for nutrients and chlorophyll-a, and additional QA/QC comparisons involving all three data generators using split sample techniques should be performed to ensure similar data quality among data generators. This will reduce artificial longitudinal gradients due to differences in sampling methods and data quality from obscuring true north-south differences in water quality.

In conclusion, while there are aspects of the CBP main Bay water quality monitoring program that can be improved, the overall approach of the program is sound and will provide the empirical information needed to characterize and detect trends in Chesapeake Bay water quality and to evaluate the effectiveness of management actions. Continuation of this coordinated monitoring effort provides the best opportunity for generation of rigorous statements concerning the State-of-the-Bay and for the development of an ecologically sound water quality management strategy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. OBJECTIVES

In January 1986, the Chesapeake Research Consortium (CRC) contracted Martin Marietta Environmental Systems (contract X-003321-02) on behalf of the U.S. EPA Chesapeake Bay Liaison Office to provide statistical and analysis support to the Monitoring Subcommittee of the Chesapeake Bay Program (CBP). Specific analysis tasks were defined by the Water Quality Data Analysis Working Group of the Monitoring Subcommittee (the "Analytical Working Group") and focused on analysis and evaluation of the first 18 months of mainstem water quality monitoring data collected by the mainstem monitoring program. The work conducted for this contract represents the first time that the CBP mainstem water quality monitoring data have been examined from a Baywide perspective, and the results of this work will assist the CBP Liaison Office in:

- Characterizing variability in space and time among the measured water quality parameters
- Evaluating the adequacy of the existing monitoring program for characterizing spatial/temporal structure in water quality and for characterizing existing conditions
- Identifying statistical methods that are most applicable to evaluation of trends in and measuring responses of water quality parameters to management actions.

A wide range of data processing, graphics preparation, and statistical analyses were conducted. Results of these analyses were presented in three interim reports (Deliverables 1-3). Contents of these interim documents include:

- Graphical and tabular characterization of the 1984/1985 data (Deliverable 1 - Task A)
- Identification of a spatial and temporal aggregation schemes (Deliverable 1 - Task B and C)
- Statistical analysis to determine if information within depth layer samples was redundant (Deliverable 1 - Task D)

- Graphical and tabular analyses of recent and historical water quality data (Deliverable 2)

- Review and application of univariate statistical methods appropriate for trend detection in CBP water quality data (Deliverable 3).

This document is the final report (Deliverable 4) and synthesizes the information in Deliverables 1-3 into an analysis framework for applying the trend detection methods identified and reviewed in Deliverable 3. We also make recommendations for improving the quality and usefulness of the information being collected by the water quality monitoring program. Deliverables 1-3 are included with this report as Appendices A-C, respectively, in a companion volume.

The CBP mainstem water quality monitoring data are stored in a centralized database in Annapolis, MD. Prior to this project, these data were assumed to require only minor corrections before they could be used in statistical analyses. Unanticipated quality assurance/quality control (QA/QC) problems were, however, identified when the data for this project were transferred to Martin Marietta Environmental Systems. Most of the QA/QC problems resulted because of the complexity of the water quality monitoring program (e.g., 50 stations and up to 24 parameters sampled by three institutions at various depths) and the difficulty in initiation of a centralized database for such a large, complex monitoring program. In order to ensure acceptable data quality for statistical analysis, Martin Marietta Environmental Systems conducted a variety of QA/QC procedures on the data. The result of QA/QC efforts was production of a data set in which scientists and water quality managers could have a high degree of confidence for preparation of the "State-of-the-Bay" report and associated statistical analysis. This work was not presented in Deliverables 1-3 and is briefly described in Chapter IV of this report.

B. HISTORICAL PERSPECTIVE

The Chesapeake Bay is the nation's largest estuary and one of its most valuable natural resources. It is renowned for its fishery and shellfish harvests and its value as wildlife habitat. Chesapeake Bay fish and shellfish harvests as well as water and sediment quality declined as its watersheds and shorelines were developed. As a result of these declines, the U.S. Congress authorized the U.S. EPA to conduct a study of the Bay's water quality and its relationship to declines in living resources in 1975. This authorization established the EPA's Chesapeake Bay Program (CBP) with a Liaison Office located in Annapolis, MD. Results of the first five years of CBP study were published in 1983 and suggested that inputs of nutrients, suspended sediments, and toxic and hazardous substances were adversely affecting the "health" and productivity of the Bay. Perhaps the most important finding of initial CBP efforts, however, was that hydrodynamic and biological processes were linked and jointly control the "health" and water quality of Chesapeake Bay. As a result, the effects of man's activities on water quality and living resources were seldom localized but rather affect water quality and biological productivity on a regional basis.

In response to the threat of declining water and sediment quality to the Bay's living resources, the federal government and the states of the Chesapeake Bay region pledged to restore the environmental quality of the Chesapeake Bay and protect its living resources. The Chesapeake Bay Restoration and Protection Plan was finalized in September 1985. Participants in the plan included the U.S. EPA, State of Maryland, Commonwealth of Virginia, District of Columbia, and Commonwealth of Pennsylvania. Management actions adopted by the participants included:

- Decrease inputs of nitrogen and phosphorus
- Decrease inputs of toxic and hazardous substances
- Decrease inputs of sediments
- Improve and restore habitat quality for living resources.

A key element in the Restoration and Protection Plan was the establishment of a comprehensive monitoring program to collect the information required:

- To characterize existing conditions (i.e., define the "State-of-the-Bay") including separation of variation due to natural phenomena from changes due to pollutant inputs
- To track the responses (hopefully improvement) of the environmental quality and living resources to management actions taken by federal, state, and local governments
- To direct research (i.e., formulate hypotheses for testing) and provide data for modeling efforts aimed at identifying and understanding processes and mechanisms controlling water quality and biological productivity.

A network of 50 mainstem Chesapeake Bay water quality monitoring stations (28 in Virginia and 22 in Maryland) as well as an additional 77 tributary monitoring stations were established (Fig. I-1). The locations of monitoring stations were selected to provide a baywide characterization of existing water quality

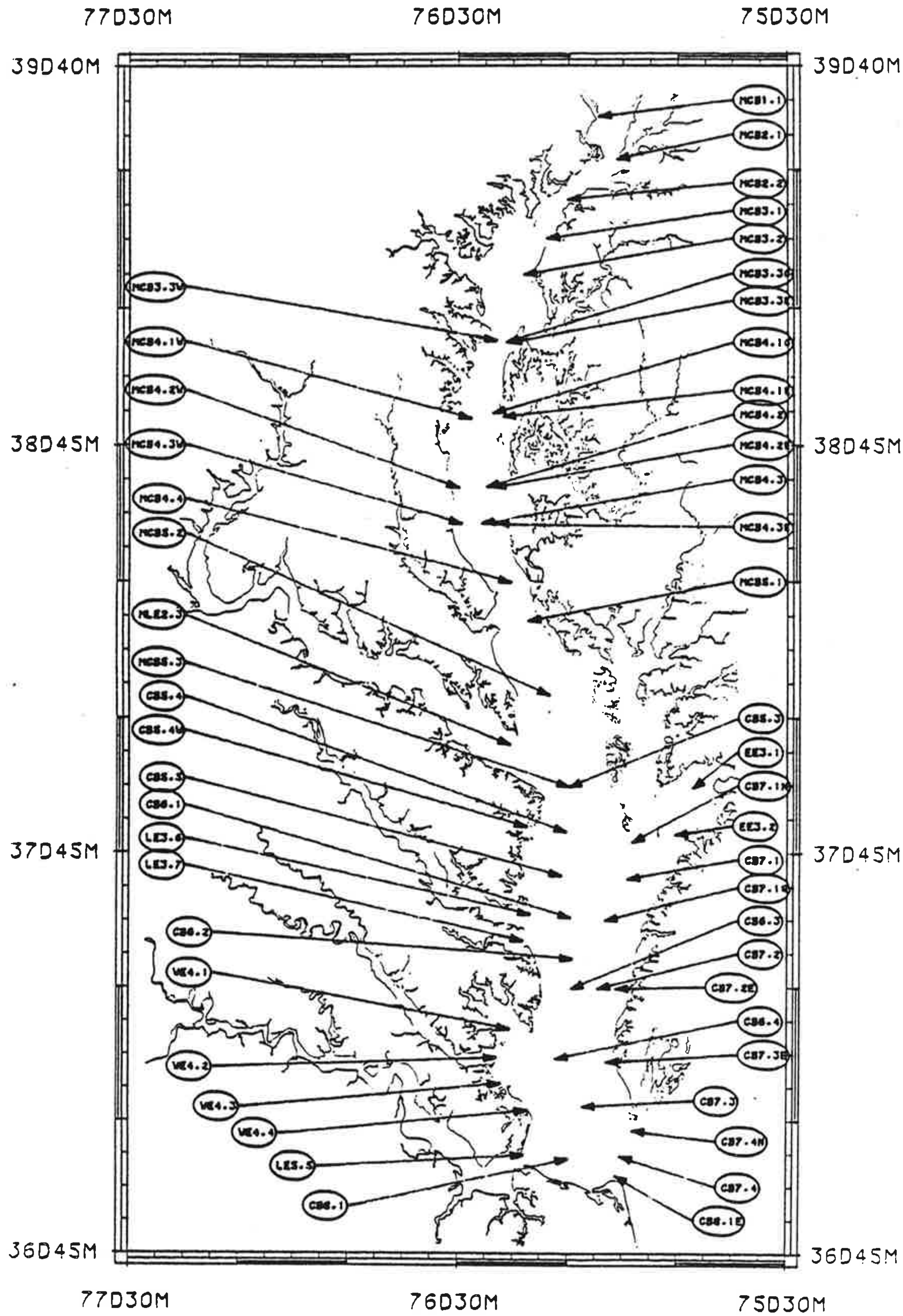


Figure I-1. Map of Chesapeake Bay showing the locations of the CBP mainstem water quality monitoring stations conditions. Virginia stations are sampled by the Virginia Institute of Marine Sciences (VIMS) and Old Dominion University (ODU). All Maryland stations are sampled by the Maryland Department of Mental Health and Hygiene, Office of Environmental Programs (OEP). VIMS, ODU, and OEP are referred to in later parts of this document as the data generators.

Sampling for the monitoring program is conducted twice monthly from March through October and once monthly from November through February, resulting in a total of 20 cruises per year. Each cruise covers, to the extent possible, the entire station array and takes about three days. Table I-1 lists the variables measured at each station.

The depths at which measurements are made varies with parameter and data generator. From June 1984 through April 1986, OEP measured temperature, conductivity, salinity, dissolved oxygen concentration (DO), and pH at 0.5 (surface), 1.0, 2.0, and 3.0 m below the air-water interface. Below 3 m depth, OEP took measurements at 3.0 m intervals, however, if DO varied more than 1.0 mg/l, or if conductivity varied more than 1,000 micromhos/cm, over any of the 3.0 m intervals, measurements were taken every 1.0 m within that 3.0 m interval. As of May 1986, OEP modified their sampling protocol to take measurements at 2 m intervals, with samples taken at 1 m intervals within any 2 m interval for which the change in conductivity exceeded 1,000 micromhos/cm or DO exceeded 1.0 mg/l. ODU and VIMS measured temperature, conductivity, dissolved DO and pH at 1.0 m (surface), and then at 2.0 m intervals thereafter.

Samples of nutrient variables, chlorophyll-a, and total suspended solids were taken near the surface and 1.0 m above the bottom by all data generators. In addition, OEP took samples for nutrients, chlorophyll-a, and total suspended solids just above and just below the pycnocline. A calculation made at the time of sampling was used to identify the maximum rate of vertical change of conductivity through the water column (i.e., the pycnocline). If a pycnocline was not detected, OEP took samples at 1/3 and 2/3 of the distance between the surface and bottom sample depths. At nine Virginia stations, ODU and VIMS took samples for nutrients, chlorophyll-a, and total suspended solids just above and just below the pycnocline if a pycnocline was detected. Detailed descriptions of the sample processing protocols and chemical analysis methods used by each data generator are available from the Chesapeake Bay Liaison Office.

Table I-1. Water quality variables reported at various depths at each of the CBP mainstem stations. Some of these variables are directly measured, while other are computed from measured variables

Temperature
Dissolved oxygen
Specific conductance (or salinity)
pH
Secchi depth
Total Kjeldahl nitrogen (filtered and unfiltered)
Nitrite plus nitrate concentration
Nitrite
Nitrate
Ammonia (filtered)
Particulate organic nitrogen
Total nitrogen
Total dissolved nitrogen
Total organic carbon
Dissolved organic carbon
Particulate organic carbon
Silicate (filtered)
Chlorophyll-a
Phaeophytin
Total suspended solids
Total phosphorus
Total dissolved phosphorus
Dissolved orthophosphorus
Particulate phosphorus

II. DATA ANALYSIS APPROACH AND RATIONALE

The major objective of the CBP Monitoring Program is to measure water quality responses to management actions taken to reduce nutrient and sediment loadings. Attainment of this objective requires hypothesis testing, and in Deliverable 3, univariate statistical methods applicable for detecting trends in water quality and testing hypotheses about water quality responses to management action were reviewed. In this Chapter, we present a decision network for application of the methods identified in Deliverable 3 and describe how multivariate statistical analyses fit into the overall analysis approach. Results from the implementation of this analysis approach should be useful for:

- Characterizing spatial and temporal variation in Chesapeake Bay water quality (i.e., determining the "State-of-the-Bay")
- Partitioning variation due to natural phenomena from changes due to nutrient and sediment loadings and identifying potential processes and mechanisms affecting water quality
- Tracking the response of Bay water quality to management actions that reduce nutrient and sediment inputs
- Developing hypotheses for evaluation and testing by the research and modeling programs
- Developing a water quality management strategy for Chesapeake Bay that is based on an understanding of the processes and mechanisms affecting Bay water quality.

Figure II-1 shows the central position of the analysis of water quality monitoring data in Chesapeake Bay Program.

A. RECOMMENDED ANALYSIS APPROACH

Figure II-2 shows the steps in the analysis approach we recommend the CBP use. The first step is QA/QC and summarization of the data transmitted from data generators to the database. QA/QC identifies erroneous data and ensures acceptable data quality for conduct of statistical analyses.

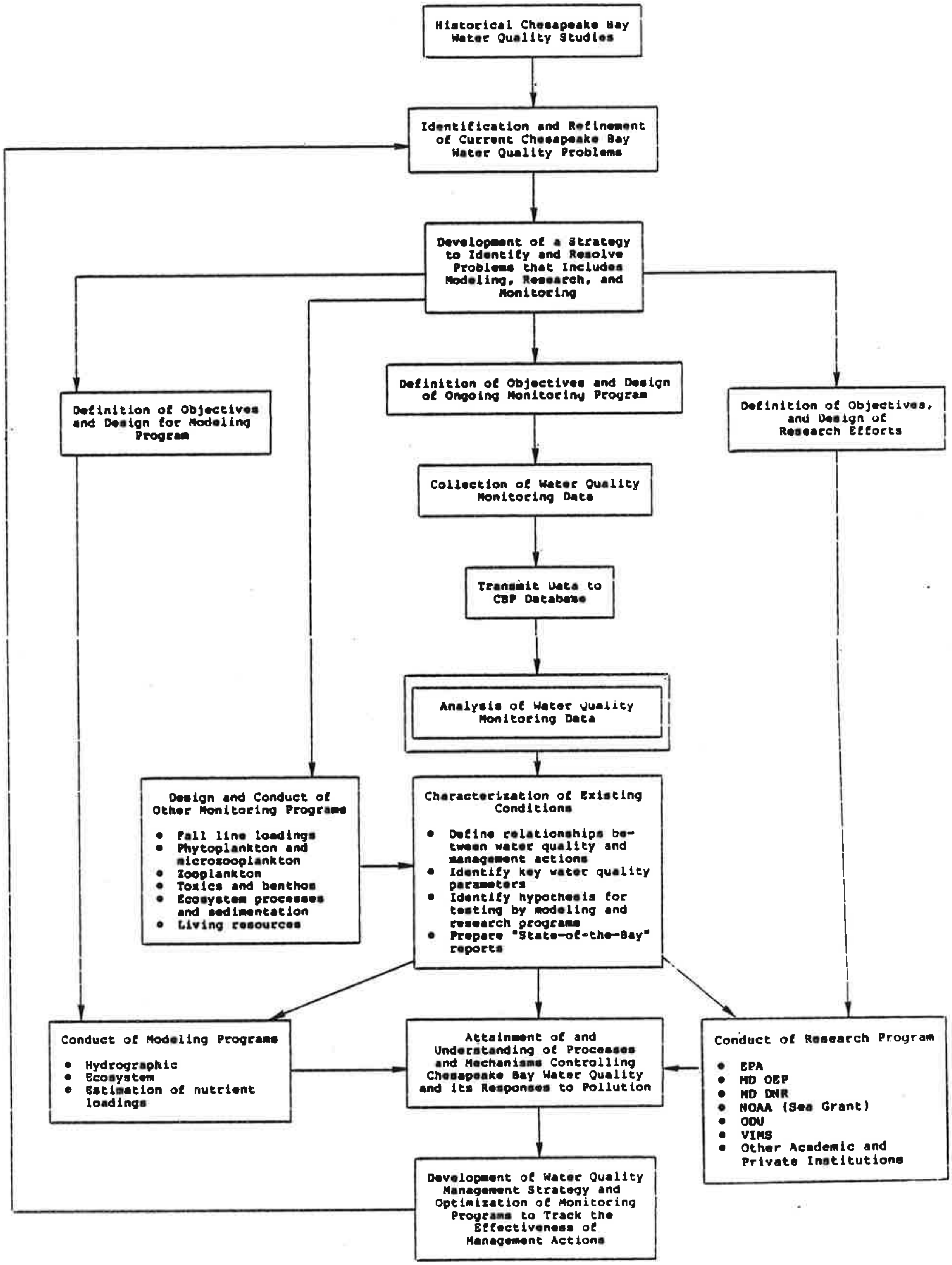


Figure II-1. Flow chart showing the relationship of water quality monitoring program to modeling and research programs and attainment of the overall goal of development of a Chesapeake Bay water quality management strategy. Note the key role of the analysis step in the overall program.

QA/QC - Data Summary

- Prepare graphical summaries for
  - -- Parameters (by depth) vs. station/latitude (Appendix A, Task A)
  - -- Isopleth maps of selected parameters (Appendix A, Task A)
  - -- Parameters (by depth) vs. cruise/date (Appendix A, Task A)
- Prepare tabular summaries of statistical properties (Appendix A, Task A)

↓

Spatial/Temporal Aggregation of the Data

- Define spatial/temporal aggregation scheme (multivariate statistical methods) -- e.g., Principal components analysis (Appendix A, Tasks B and C)
- Use recent and historical data to refine spatial/temporal aggregation scheme using other ordination/classification techniques
- Identify stations that are representative of regions (i.e., stations that are consistently classified in particular regions)
- Identify transitional stations (i.e., stations that are classified in two or more regions)
- Identify time periods when water quality is relatively stable from cruise to cruise (i.e., stable seasons)
- Identify time periods when water quality is in a transition from one state to another
- Identify water quality parameters making major contributions to similarities and differences among regions and cruises

↓

Select Analysis Variables (Univariate and Multivariate)

- Response variables
  - -- Specific water quality parameters (Table I-1)
  - -- Integrated water quality measures (e.g., volume of hypoxic water)
  - -- Index values (e.g., Estuarine Index of Enrichment - see Reed and McErlean 1979; first component score of a PCA)
- Explanatory variables
  - -- Functional (e.g., degree of stratification)
  - -- Empirical (e.g., station, year, cruise)
- Subset data based on spatial/temporal analysis results and depth layer
- Confirm homogeneity of trend
- Identify appropriate trend detection methods

↓

Conduct Trend Analysis

Figure II-2. Major steps in the proposed analysis approach

Summarizing and display of the data is an essential part of most QA/QC programs. Data summaries that should be accomplished include graphs of parameters (regional and seasonal averages) vs depth and latitude, isopleth maps of selected parameters (DO, salinity) by cruise, graphs of parameters (regional and seasonal averages) vs cruise/date, and tabular summaries by season and region, including information on central tendencies, frequency distributions, frequency of censored data, and variance. Data summaries and graphs will be used to guide selection of statistical analyses during application of the analysis decision network. Data summaries and graphs also provide a first order description of the spatial and temporal structure in the data and provide water quality managers with real time information on existing conditions.

The second major step in the analysis process is aggregation of the data in a manner which captures the spatial and temporal structure in Chesapeake Bay water quality. Because water quality is the net result of variation of many parameters (DO, chlorophyll-a, salinity, nutrients), a multivariate analysis technique is required to appropriately aggregate stations and cruises into groups. In Deliverable 1 (Appendix A, Tasks B and C), we used principal components analysis (PCA) for this purpose. PCA was selected because it was a method for parsimoniously reducing the large amounts of complex and correlated water quality data collected into a smaller set of uncorrelated variables (i.e., principal components). PCA is relatively straight forward to conduct, and analysis results are relatively easy to interpret. The aggregations of stations and seasons that resulted from PCA were consistent with existing hydrodynamic, chemical, and biological processes controlling Bay water quality (see Appendix A, Fig. II-2). Other multivariate analysis techniques may also be appropriate for describing spatial and temporal structure in CBP water quality monitoring data (see Appendix D). Techniques that identified transitional stations or seasons would be particularly useful for refining the spatial/temporal aggregation scheme. We did not evaluate the applicability of other multivariate statistical methods for making spatial/temporal aggregations because the method we selected was very successful and most of the remaining effort for this contract was directed toward evaluating the trend analysis techniques that will be discussed later.

RECOMMENDATION: Multivariate statistical techniques, particularly those identified in Appendix D, should be evaluated to determine their applicability for characterizing spatial and temporal variation in CBP water quality monitoring data. Both recent and historical data should be used in these analyses.

The third step in the analysis approach is to select analysis variables based on the working hypothesis that increased inputs of nutrients and sediments have enhanced algal productivity and indirectly increased the extent and duration of hypoxic/anoxic conditions due to decomposition of excess production. Response variables include measures of water quality that are likely to be affected by changes in nutrient loadings including individual water quality parameters that are directly measured by the monitoring program (DO, nutrient concentrations, chlorophyll-a concentration), integrated water quality measures that can be calculated using the data (volume of hypoxic/anoxic water), and combinations of the values of multiple water quality parameters into predetermined indices (e.g., see Reed and McErlean 1979), or data derived indices (e.g., first component of a PCA) of water quality. Multivariate statistical techniques may be particularly useful for defining univariate indices because they provide an objective means of capturing the complex spatial/temporal structure of water quality into a single measure. Explanatory variables are included in analyses to partition variation due to natural phenomena from that associated with inputs of nutrients or pollutants. They are either empirical (station, year, cruise) and are included in analyses to account for known sources of spatial and temporal variation, or they may be functional (e.g., a measure of salinity stratification) and are included in analyses to account for postulated relationships among water quality parameters. Following selection of response and explanatory variables, the data should be subset into regional and seasonal groups for trend analysis based on results of PCA or other aggregation analyses, homogeneity of trend within groups, practical considerations (likelihood of detecting trends, interpretability of results), and hypotheses to be tested. A detailed discussion of the spatial and temporal scale appropriate for trend analysis is presented in Deliverable 3 (Appendix C).

Conduct of trend analysis is the final step within the analysis process. In Deliverable 3 (Appendix C) we presented a list and review of univariate statistical methods appropriate for trend detection, and applied these methods to historical and recent Chesapeake Bay water quality data. Figure II-3 is a decision network for selecting among the various statistical methods presented in Deliverable 3 depending upon the characteristics of the response and explanatory variables to be analyzed. This decision network is based on the answers to the following questions:

- Is the response variable censored, and if so, what percentage of the observations are at detection concentrations?

- Are functional explanatory variables to be included in the analysis (e.g., intensity of stratification as measured by salinity gradient with depth for analysis of trends in DO) and if so, are they censored?

- How many empirical explanatory variables (i.e., stations, cruises, depths) are to be included, and are they consistently defined over time?

Figure II-3 shows how the answers to these questions directs a analyst to the appropriate analysis method (boxes labeled A-M in Table II-1). As discussed in Deliverable 3, the implementation of each method depends on the specifics of the available data and the hypotheses to be tested. To illustrate, if we use the decision network to identify which analysis method is appropriate to test if year-to-year variation in summer DO in the mid Bay region is increasing or decreasing for the CBP monitoring data without accounting for year-to-year differences in the degree and intensity of stratification, we end up at box E. This is because:

- DO is as a continuous variable that is not censored.
- No functional explanatory variables are included in the analysis.
- Explanatory variables include station, year, and cruise, and are measured consistently over time.

The specific ANOVA model for mainstem monitoring data would include terms for station, year, and cruise nested in year. Continuing in the decision network (Fig. II-3a), an assessment of whether the assumptions underlying ANOVA are satisfactorily met would next be performed using statistical tests (e.g., Durbin-Watson statistic) and residuals analysis. If deemed necessary, any adjustments to data in order to better meet these assumptions would then be identified and applied. Provided the assumptions of ANOVA can be satisfactorily met, the last step in the decision network involves interpretation of results. Figure II-4 (reproduced from Deliverable 3) shows another decision network for the interpretation of the results from the application of such an ANOVA model. Figure II-4 demonstrates an important consideration. That is, once the appropriate statistical model has been defined it may be necessary to conduct more than one test to refine results and formulate conclusions. Other examples of the implementation of the methods listed in the decision network are provided in Deliverable 3 using both historical and recent CBP monitoring water quality data. These examples are summarized in Table II-2, including their correspondence to the general methods listed in Fig. II-3.

Caution must be used when interpreting the results from the proposed analysis approach. This is because many of the proposed analyses involve the grouping of multiple stations into "regions", multiple cruises into "seasons", and multiple depth-specific measurements into "depth layers." For both

Figure II-3a and II-3b appear on following two pages.

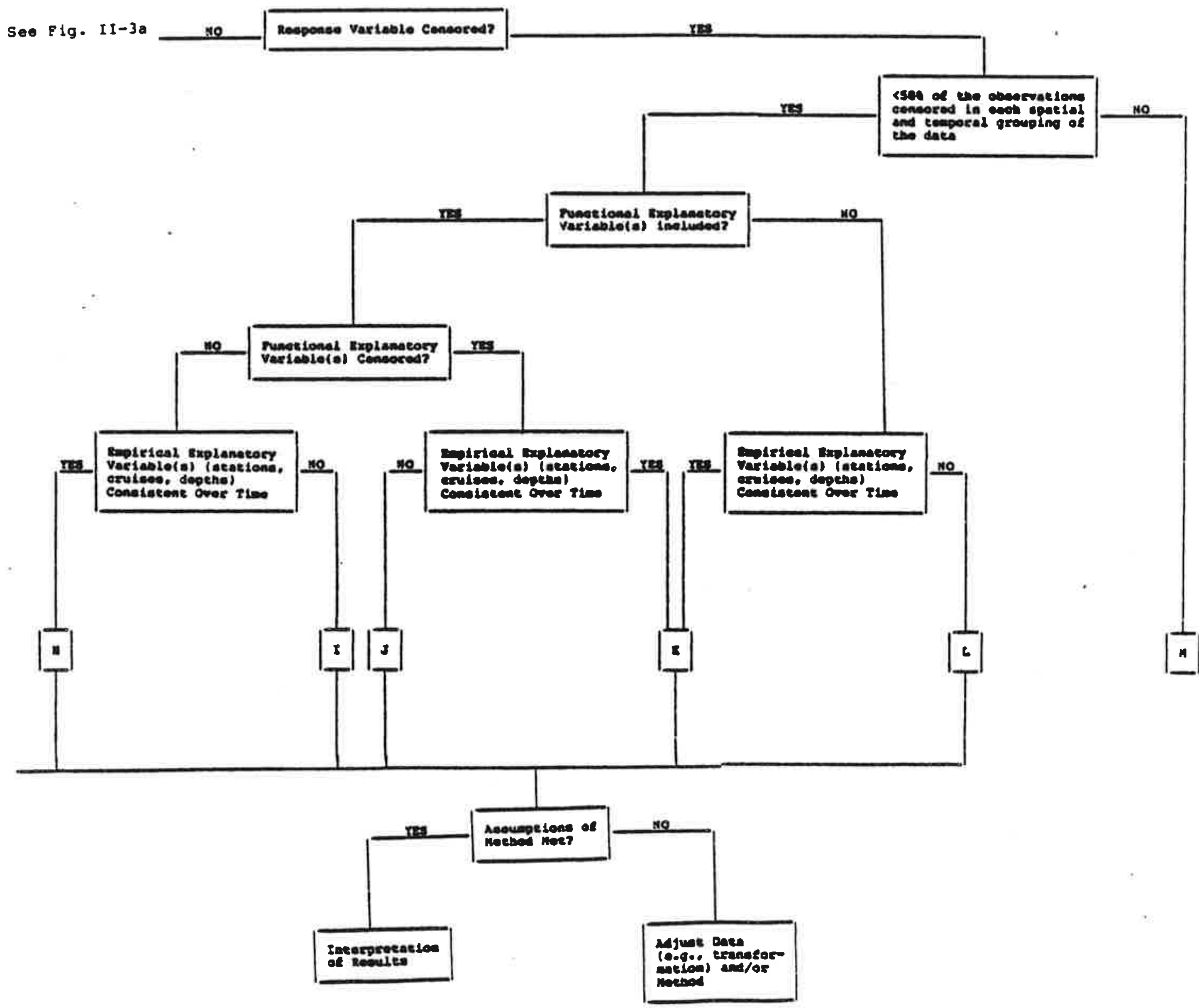
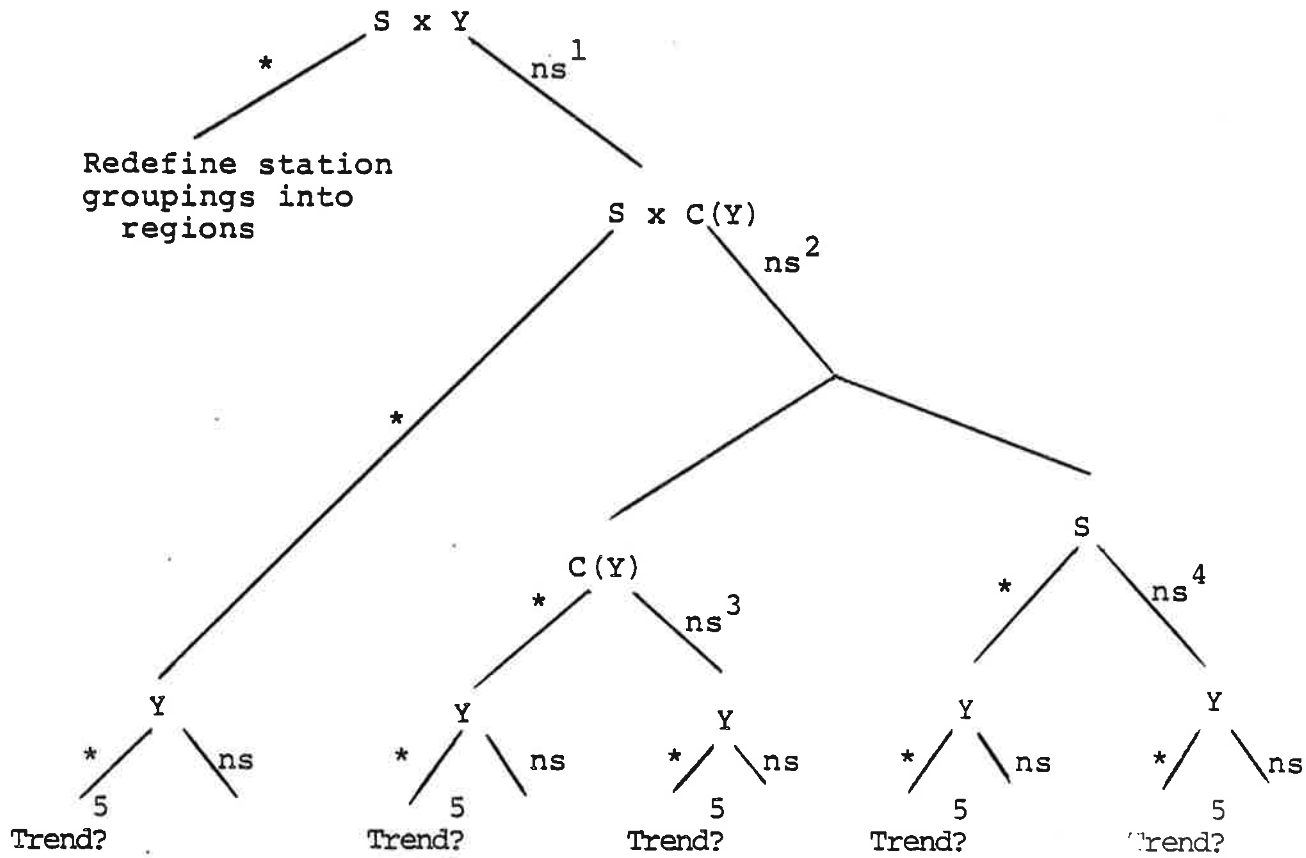

See Fig. II-3a
NO
Response Variable Censored?
YES
<50% of the observations censored in each spatial and temporal grouping of the data
YES
NO
Functional Explanatory Variable(s) included?
YES
NO
Functional Explanatory Variable(s) Censored?
NO
YES
Empirical Explanatory Variable(s) (stations, cruises, depths) Consistent Over Time
YES
NO
Empirical Explanatory Variable(s) (stations, cruises, depths) Consistent Over Time
NO
YES
Empirical Explanatory Variable(s) (stations, cruises, depths) Consistent Over Time
YES
NO
H
I
J
K
L
M
Assumptions of Method Met?
YES
NO
Interpretation of Results
Adjust Data (e.g., transformation) and/or Method

Figure II-3b.

Table II-1. Key to the statistical methods denoted in boxes labeled A-M in Figure II-3 (KT = Kendall's Tau; SRC = Spearman Rank Correlation)

Box Label	Method
A	ANCOVA
B	KT and SRC on residuals obtained from a regression of mean response variable on functional explanatory variable
C	ANOVA or ANCOVA with categorized functional explanatory variable
D	KT and SRC on residuals obtained from data alignment procedure using mean response variable and categorized functional explanatory variable
E	ANOVA
F	Friedman's two-way ANOVA and Hirsch's modification to KT
G	KT and SRC on mean response variable
H	Logistic regression
I	KT and SRC on residuals obtained from a regression of median response variable on functional explanatory variable
J	Linear logit models
K	KT and SRC on residuals obtained from data alignment procedure using median response variable and categorized functional explanatory variable
L	KT and SRC on median response variable
M	Data are too censored for analysis of trends using these methods*

* Possible approach is an analysis of the frequency of observations at detection concentration (e.g., binomial model).



1. Option of eliminating S x Y from model.
2. Option of eliminating S x C(Y) from model.
3. Option of eliminating C(Y) from model. (If options 2 and 3 are invoked, cruises in a season are treated as "replicates.")
4. Option of eliminating S from model. (If options 1, 2, and 4 are invoked, stations are treated as replicates.")
5. Trend detection using pairwise or multiple comparison tests on year-specific parameters.

Figure II-4. Decision network for interpretation of the results from an ANOVA model applied to CBP water quality monitoring data to detect trends in summer DO concentrations. The ANOVA model is based on station (S), year (Y) and cruise nested in year (C(Y)) factors with appropriate two-way interactions. Each step involves the evaluation of the significance of main effects or interaction terms (* = significant, ns = not significant)

Table II-2. Summary of trend analyses demonstrated in Deliverable 3, and their correspondence to the methods proposed in the decision tree (Figure II-3)

Data	Method	Response Variable	Empirical Explanatory Variables	Functional Explanatory Variables
Calvert Cliffs	G	Mean DO over stations and cruises (months)	Year	None
Calvert Cliffs	(F) Friedman's two-way ANOVA	Mean DO over cruises (months)	Year; station as blocking variable	None
	(F) Hirsch's modification to Kendall's Tau	Mean DO over stations	Year	None
Calvert Cliffs	B & D	Mean DO over stations and cruises (months)	Year	Salinity difference (Bottom-Surface)
Calvert Cliffs	A	Mean DO over stations	Year; cruise (or month)	Salinity difference
Calvert Cliffs	E	Mean DO over stations	Year; cruise (or month)	None
1984/1985 CBP	E	Chlorophyll-a	Year; station; cruise	None
1984/1985 CBP	J	Ammonia	Year; cruise	None

historical and recent CBP monitoring data stations, cruises, and depths are not randomly sampled. Therefore, conclusions from these analyses are appropriate only for the specific stations, cruises, and depth measurements used in the analysis. Additional assumptions about how representative stations are of the "region", cruises are of the "season", and depth specific measurements are of the "depth layer" are required to extrapolate conclusions to regions, seasons, and depth layers.

The decision network is a reasonable approach for selecting among the appropriate analysis methods based on characteristics of the data. In many cases, however, it is possible that several different analysis approaches could be used for a given data set. For this reason, the analysis approach and decision network should be refined based on further application to monitoring data collected by the CBP water quality monitoring program.

RECOMMENDATION: Additional refinement of the proposed decision network for statistical trend detection should be performed using water quality monitoring data after three or more years of data covering a range of conditions are available.

B. LINKAGE WITH HISTORICAL DATA

Detection of changes in water quality requires a sufficiently long time series to allow for management actions to have an effect and to permit the removal of natural variation for a range of conditions (e.g., high flow and low flow years). The linkage of CBP monitoring data to historical water quality data provides a means for extending the temporal coverage of data collected as part of the present monitoring program to include a range of conditions. The farther back in time present data can be extended, the more likely and quickly trends in water quality attributable to management actions can be detected.

An obvious approach to linking the recently collected water quality data to historical data is to continue analyses of trends performed as part of the 1976-1981 CBP (Flemer et al. 1983a and b) and related studies (Officer et al. 1984). These analyses included:

- Pooling historical data over space and time and performing correlation analyses to determine if water quality variables have changed systematically over years. For these analyses, data were averaged by regions on an annual and seasonal basis for the entire water column, top 10 m of the water column and for the portion of the water column greater than 10 m

- Graphical comparison of recent and historical DO data for years with similar freshwater inflows and salinity distributions including visual comparison of isopleths along the north-south axis of the mainstem, vertical profiles at selected stations, estimates of the volume of low DO water, and projections of the areal extent of low DO water for selected years.

Below general conceptual problems with this analysis approach and specific deficiencies in the CBP historical database that precluded use of this analysis approach for this contract are discussed, and recommendations to correct these problems are presented.

Correlation Analyses

When data from a number of studies with different objectives are pooled, the resulting data are frequently biased due to differences in study design, sampling methods, and measurement techniques. If data are pooled over a large number of studies, some of these biases may be "averaged out." In some cases, correction factors can be used to correct for differences in sampling and measurement methods (e.g., artificial censoring of data with lower detection limits). For the most part, however, the biases introduced by pooling data from a number of studies are unknown and not quantifiable. One method to reduce the bias in long-term data sets and check on the reliability of analysis results for pooled data are to conduct analyses for selected stations with long historical records that used relatively consistent sampling methods and had the same monitoring objectives over time. Results from these station-specific analyses can then be compared to results for analyses conducted on the pooled data. An example of this approach was provided in Deliverable 3 (Appendix C), in which historical data on four main Bay stations located near Calvert Cliffs were analyzed for DO trends using general linear models (ANOVA, ANCOVA) as well as distribution free (nonparametric) correlational techniques. Data were analyzed by specific dates (months) and stations, and pooled across stations and months.

When results of analyses based on pooled data and station-specific data agree then there is higher confidence that the observed trends are real. If the results of the two analysis approaches disagree with each other, it is generally not possible to determine which one is correct. Trends are variable within regions and seasons and station locations in the region and cruise dates in the season were not randomly selected and thus may not be representative of regional and seasonal water quality. Furthermore, statistical analyses applied to environmental data,

Table II-3. Average values of total nitrogen concentration (mg/L) from recent (1984/1985) and historical mainbay data. Results are presented by year, season, and depth strata for CBP segment CB3.

Variable: Total nitrogen concentration   Segment: CB3

	10 Meters or Less*						More Than 10 Meters*					
	Spring**		Summer**		Fall**		Spring**		Summer**		Fall**	
Year	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
1969	0.890	16	0.641	69	0.973	29	--	--	--	--	--	--
1970	1.455	89	0.723	161	0.891	72	1.215	13	1.002	20	0.872	11
1971	1.129	94	0.976	169	--	--	0.967	13	0.901	25	--	--
1972	--	--	--	--	--	--	--	--	--	--	--	--
1973	1.429	57	1.035	58	--	--	1.110	6	0.818	6	--	--
1974	1.283	56	0.797	64	0.772	18	--	--	--	--	--	--
1975	1.353	97	0.974	84	1.045	64	1.006	19	0.710	9	--	--
1976	1.234	42	0.765	56	1.216	33	--	--	0.668	7	--	--
1977	1.072	90	0.448	97	0.821	61	--	--	--	--	--	--
1978	2.231	36	1.076	43	0.671	50	--	--	--	--	--	--
1979	1.403	23	0.914	13	1.253	18	--	--	--	--	--	--
1980	1.593	26	1.469	60	2.317	16	--	--	--	--	--	--
1981	--	--	--	--	--	--	--	--	--	--	--	--
1984	--	--	--	--	0.582	81	--	--	--	--	0.530	19
1985	0.920	96	0.786	97	--	--	0.681	16	0.833	23	--	--

* Depth Strata

** Season (Winter: D, J, F; Spring: M, A, M; Summer: J, J, A; Fall: S, O, N) which rarely satisfy all of the assumptions of various methods, should be viewed as indicative, rather than as conclusive evidence, of trends in water quality. Determining which of the analysis results is correct therefore becomes difficult.

RECOMMENDATION: Identify and assemble long-term time series of water quality data at selected individual or clusters of stations within major regions of the Bay and major tributaries. The review of historical data performed by Heinle et al. (1980) should be used as one of the basis for identifying appropriate data sets and water quality variables for which long-term data exist. DO, secchi depth, and salinity are the likely variables for which reliable time series can probably be assembled. The water quality data that are eventually compiled should be analyzed for trends using the methods described and illustrated in the previous section by applying the trend analysis decision network.

As part of Deliverable 2, we attempted to extend the correlational analyses used previously by the CBP to include the 1984/1985 monitoring data (see Appendix B, Task F). Table II-3 is an example of total nitrogen historical time series (including 1984/1985 data) for the depth layers defined as above 10 m and below 10 m. Correlation analysis (e.g., Pearson correlation as used by the CBP) and distribution-free methods (as outlined in Deliverable 3 and Fig. II-3), can be used to determine if these data vary in a systematic manner over years. However, after these time series were compiled, QA/QC problems (i.e., incorrect depths associated with water quality data) in the CBP historical database were discovered by CBP personnel thus questioning any results from trend analyses of these data. At this time, the implications of these data problems on the results of depth layer specific analyses are unknown. Once errors have been identified and corrected, these analyses can be completed to evaluate trends in Bay water quality.

RECOMMENDATION: Correct the CBP historical database to permit correlational analyses of historical data for major depth layers (e.g., <10m; >10m). Once the database problems are rectified, historical time series can be updated with data from the present monitoring program and the analysis approach of correlation methods to evaluate trends in pooled data can be empirically evaluated for its utility in trend detection.

Graphical Comparisons

We also had difficulties extending graphical comparisons of historical and recent data used previously by other researchers to quantify long-term changes in the volume of hypoxic/anoxic

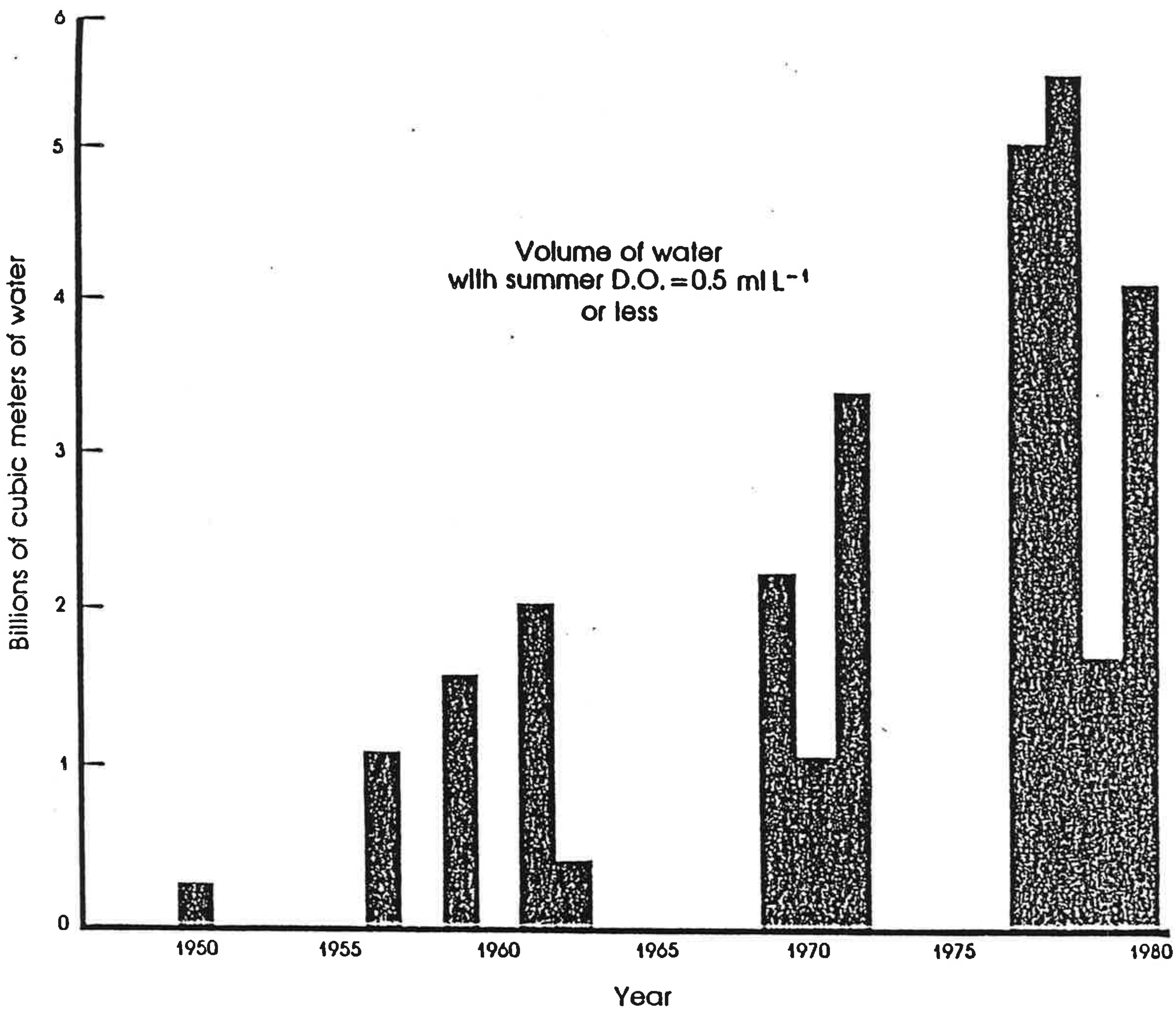


Figure II-5. Historical time series of the estimated volume of Chesapeake Bay water with DO $\leq$ 0.5 ml/L during the summer. (Reproduced from Flemer et al. 1983a) waters. To illustrate these difficulties, consider our attempt to extend the time series showing summertime estimates of the volume of low DO water for selected years (Fig. II-5). We were unable to satisfactorily extend this time series within the scope of this contract because detailed documentation on the methodology used to estimate the volume of low DO water was not available. We therefore had to request documentation from the original investigator (Mr. Ned Berger) who was overseas and unable to respond for several months. Furthermore, upon receipt of Mr. Berger's documentation, we found that application of the methodology required estimation of quantities that could vary from investigator to investigator because the methodology require subjective judgement to be exercised. We thus had to ensure that our implementation of the method on recent data was consistent with the previous implementation procedures used by others. However, the Chesapeake Bay Institute (CBI) data used to estimate the historical time series were not available in the CBP historical database. Thus, to extend the time series, the data from original CBI reports would have to be keypunched and verified. This additional work was beyond the scope of this contract.

Examination of DO isopleths for 1984 and 1985 data (Appendix B, Task A) showed that cruise to cruise variation in the extent of hypoxic/anoxic water could be as large as that which occurred from year to year. This high intra-summer variation in hypoxia/anoxia makes comparisons of recent and historical data difficult because most historical estimates of the volume of low DO water were based on single cruises within a summer and do not include estimates of within summer variability. Estimation of within summer variability for the volume of low DO water for historical data is required before 1984/1985 conditions can be compared to historical conditions. Unfortunately, the historical data in the CBP database do not include all data that encompass multiple cruises in each summer. It will, therefore, not be possible to make meaningful comparisons of the extent and duration of hypoxia/anoxia between recent and historical data until additional data for multiple cruises have been added to the database.

RECOMMENDATION: Update the CBP historical water quality database to include post-1981 Chesapeake Bay Institute data and other historical data (e.g., PROVIDER studies) not presently included. Studies that include measurements of DO and salinity at multiple depths during multiple cruises within a summer should be incorporated first.

Methods for characterizing the variability of the location, extent, and duration of hypoxia, should be developed. An initial approach would be application of the methodology used previously by the CBP for estimation of the volume of hypoxic water to both historical and CBP monitoring data from individual cruises, within a summer, and examination of the mean and range of these estimates for each summer over years. This approach would also allow comparison to the historical time series based on CBI data, and a preliminary assessment of the magnitude of intra-summer variability relative to historical trends. As part of the application of these methods, data from lateral stations may be useful for qualitatively characterizing the across-Bay variability in hypoxia. Due to only three stations being on any given lateral transect and the location of many of these lateral stations in relatively shallow water that experience intermittent stratification, quantitative assessment of across-Bay variation in hypoxia is probably not possible with these data.

RECOMMENDATION: Investigate methods for characterizing within summer variability of the duration and extent of hypoxia and stratification intensity. These methods should be applied to available historical DO and salinity data that consist of multiple cruises within a summer (from an updated CBP historical database) and data from the present monitoring program.

We want to emphasize that the above recommendations and discussion does not invalidate previous graphical comparisons of historical and recent data. Such comparisons are the first step in determining if a problem exists. What we are advocating is that the analysis process should now proceed toward inclusion of within year variability into historical comparisons.

III. PRELIMINARY EVALUATION OF THE SAMPLING PROGRAM

Most of the effort and expense for monitoring programs occur during the collection and processing of samples (Downing 1979; Millard and Lettenmaier 1986). It is therefore important that the sampling design and approach for a long-term study like the CBP water quality monitoring program be rigorously evaluated to ensure that it is compatible with the planned analysis approach. Such an evaluation will ensure the data that are necessary to accomplish program objectives and conduct planned analyses are collected. When a rigorous evaluation of the sampling program has been completed and power and sensitivity analyses have been conducted, allocation of sampling effort can be adjusted to obtain the desired percision and accuracy without collecting excessive amounts of information.

The major goals of this study were to identify statistical analysis techniques for determining the magnitude and direction of trends in Bay water quality and to develop an analysis framework for applying the trend detection methods identified to assess the effectiveness of management actions to improve Bay water quality. Rigorous evaluation of the sampling program was not a part of the scope of work for this contract nor was it possible to accomplish this evaluation with the information available at this time. It is important, however, to determine at this time that the information required to apply the analysis framework are being collected and are of a quality appropriate to conduct the statistical analyses that are planned. The goal of this chapter is to accomplish this a preliminary evaluation for the CBP water quality monitoring program.

Power and sensitivity analyses are important steps in a rigorous evaluation of the sampling program. Although initially called for as part of this contract, power analyses were not performed because these analyses require the estimation of specific parameters in statistical models of interest (e.g., ANOVA). Power analyses based on only 18 months of data would likely result in parameter estimates (and therefore conclusions) concerning the probabilities of detecting changes that were specific to conditions in the Bay occurring during the 18 months covered by the data. More general (and thus more meaningful) conclusions about the power of specific analyses to detect differences can only be made when power and sensitivity analyses include 3-4 years of data that represent a wide range of conditions. Conclusions based on power analyses using historical data would be specific to the sampling design that generated the historical data, and would not be relevant to the existing design of the CBP water quality monitoring program. When power analyses are accomplished, they should focus on the probability of detecting year-to-year changes in season-specific water quality parameters (e.g., summer DO), based on data generated from the present design of the CBP monitoring program (i.e., stations, cruises, depths sampled).

RECOMMENDATION: A rigorous evaluation of the sampling program including the conduct of power and sensitivity analyses should be accomplished as soon as estimates of year-to-year variability for the current monitoring program are available. Data spanning 3-4 years are likely sufficient for such an evaluation. This evaluation can be used to improve allocation of sampling effort and to determine the likelihood of the present program for addressing its objectives.

A. SYSTEMATIC VS RANDOM COLLECTION OF SAMPLES

The ongoing sampling program takes samples systematically in space (i.e., at predetermined, fixed station locations) and time (i.e., on sampling dates at approximately evenly spaced sampling intervals). A random allocation of sampling effort in space and time (e.g., a random stratified sample design) is, however, the preferred sampling strategy for many of the analyses identified in the previously discussed analysis framework (Cochran 1977; Green 1979). A random sample design was not used because:

- A random stratified design was much more costly to implement
- Anoxia, and most other water quality problems in Chesapeake Bay, are most severe and of greatest concern in the central regions of the Bay. It was therefore decided to concentrate sampling effort in the region of concern
- Research vessels were available for only limited blocks of time on consecutive days. Thus, it was not possible to randomly collect samples in time
- Data from a random design are difficult to compare to historical data collected at known stations and times.

The major advantages of a systematic sampling program were thus broad geographical coverage with the limited funds and research vessel time available. The major disadvantage of the systematic sampling program is that the measurements taken may not be representative of all conditions of concern. For instance, variance estimates for water quality variables within a region or season for which few observations are taken include variation attributable to known but uncontrolled (and therefore unestimable) sources and may also be biased in an unknown fashion. For example, the lack of coordination between the timing of cruises and tidal conditions results in variance estimates of water quality parameters for a season (group of cruises) that include a component due to differences in actual tidal condition from cruise to cruise. Similarly, because station locations are fixed, variance estimates of water quality parameters for a region (group of stations) can be biased if water quality at these stations differs consistently from other (unsampled) locations in the same region. As a result, interpretation of analysis results and formulation of conclusions about water quality on regional and seasonal scales are speculative unless the magnitude and direction of the bias can be assessed. However, to rigorously define the magnitude and direction of biases for a systematic sampling design will likely be excessively time consuming and costly.

RECOMMENDATION: Estimates of the means and variances for the water quality parameters sampled should be compared for randomly and systematically collected data for the regions and seasons which have severe water quality problems (central region of the Bay during summer). These comparisons will permit a determination of the approximate degree and magnitude of bias associated with the ongoing program and confirmation of the representativeness of the systematically collected data. Without this comparative information, it may not be possible to determine if the water quality responses (or lack of them) represent true responses or whether they are due to unknown sources of bias. This study is also necessary for estimating the "real" power of the ongoing program to detect responses to management actions because standard power analyses assume randomly collected samples. The data required to make these preliminary comparisons may be available from a pilot study conducted by OEP when they were designing their sampling program (report in preparation).

B. SPATIAL ALLOCATION OF SAMPLING EFFORT

As noted in the Introduction, a network of 50 stations (28 in Virginia and 22 in Maryland) that geographically cover the entire main Bay are currently sampled by the CBP. Our spatial/temporal analysis of the 1984-1985 data showed that this array of stations was adequate to capture the major regional structure in water quality. We, in fact, concluded that the stations located in the vicinity of tributary mouths (LE2.3, EE3.1, EE3.2, LE3.6, LE3.7, WE4.1, WE4.2, WE4.3, WE4.4, LE5.5) had a distinctly different spatial/temporal structure than nearby mainstem stations. [Note that although located near the mouth of the Great Wicomico River, station 5.4W exhibited similiar water quality dynamics as nearby mainstem stations.] This finding suggests that the water quality in the vicinity of tributary mouths is strongly influenced by conditions occurring in the tributaries. When comparable water quality data for tributaries are available analyses should be conducted to confirm this speculation. Until these analyses are accomplished, data from stations located in the vicinity of tributary mouths should not be included in analyses characterizing trends in water quality for the adjacent portions of the main Bay.

Data from stations located in the vicinity of tributary mouths are, however, a special concern to water quality managers because one of the long-term patterns that has been suggested for these habitats is that they are more frequently affected by hypoxia/anoxia in recent times than they have been historically (Seliger et al. 1985). In addition, habitats in the vicinity of tributary mouths frequently support productive populations of oysters and clams and are utilized by fish population as staging areas for seasonal movement and migrations (Lippson et al. 1979). Analyses evaluating trends in water quality for tributary mouths should therefore be conducted. However, before trend analyses can be conducted for stations located in tributary mouths, spatial/temporal analyses must be accomplished to determine if these stations are all unique or if they can be aggregated in some manner (e.g., upper Bay, mid Bay, lower Bay systems; low flow systems, high flow systems; highly perturbed and developed systems, relatively unperturbed and undeveloped systems).

RECOMMENDATION: Main Bay stations located in the vicinity of tributaries should not be included in analyses characterizing mainstem bay water quality or evaluating trends in mainstem bay water quality parameters. Rather, these data should be analyzed with upstream tributary data as a separate group and used to characterize and possibly to categorize tributary inputs.

It is possible that when the sampling program is rigorously evaluated a reallocation of sampling effort will be appropriate. For example, our preliminary summarization of the first 18 months of data suggests that spatial variation in some water quality parameters of concern (e.g., nutrient concentrations) is relatively small in Virginia waters compared to that which occurs in Maryland waters. This difference in variability suggests that variation in collected data for some water quality parameters is more homogeneous in the lower Bay than in the upper Bay, and that it may be possible to characterize water quality there with less spatial coverage. However, the homogeneous water quality data for the lower Bay also suggest it will be more difficult to measure temporal trends or responses to management actions there. Determination of which if any stations should be eliminated or reallocated to different habitats should be based on the sensitivity of the previously identified analysis methods to deletion of data associated with individual stations. One method for accomplishing sensitivity analysis would be to apply the appropriate analysis techniques repetitively eliminating one station at a time and observing the effects on results. Stations which are candidates for elimination or reallocation are those for which elimination of data have little influence on analysis results. The correspondence of stations to historical water quality sampling locations and sampling locations for living resources must also be considered as part of any reallocation/elimination process. It would be unwise to modify sampling effort until an analysis of the first three years and perhaps first five years of data had been accomplished.

RECOMMENDATION: Any reallocation or reduction in sampling effort should only be considered after the first several years of data have been analyzed to determine the sensitivity of analyses in the proposed decision network to a range of possible modifications. To the degree possible, data covering a range of conditions (e.g., periods of normal freshwater inflow and periods of high and low freshwater inflow) should be included in sensitivity analyses.

C. TEMPORAL ALLOCATION OF SAMPLING EFFORT

The monitoring program consists of 20 cruises annually with two cruises per month from March through October and one cruise per month between November and February. With this frequency of cruises, the finest time scale of water quality variation which can be characterized is seasonal patterns. We used principal components analyses to aggregate 1984/1985 cruises into seasons (Appendix A, Tasks B and C). This analysis showed that fall and spring are transition and dynamic periods between the extreme conditions which occur in winter and summer. A major objective of some of the trend analyses to be conducted will be to partition predictable seasonal periodicities before testing hypotheses about spatial patterns, year-to-year variation, and long-term responses of water quality parameters to decreases in nutrient and sediment loadings. Therefore, average annual patterns must be defined. Accurate characterization of annual cycles requires intensive sampling during the appropriate seasonal periods and a data collection phase that samples a range of natural conditions. The seasonal averages computed for the ongoing program are judged to be adequate and appropriate because sufficient samples are collected to describe and quantify major aspects of seasonal variation and partition it from other sources of variation, particularly responses to management actions.

Results of some of the ongoing research programs indicate the time scale of variation in the duration and extent of hypoxic water and algal blooms may be on time scales of hours to days (Dr. Mary Tyler, Martin Marietta Environmental Systems, personal communication). This scale of variation is not captured by the ongoing monitoring program because it is impractical to sample more frequent than biweekly from research vessels given the broad geographical coverage required and the associated costs. However, data on a finer time scale would be useful for characterizing the extent of hypoxia/anoxia and the frequency of algal blooms.

RECOMMENDATION: Water quality data on temporal variation that occurs from hours to days should be obtained to confirm the accuracy of seasonal characterizations of water quality based on bimonthly cruise data. Possible approaches for collecting these data are to use continuous monitoring at selected locations and remote sensing techniques. Initial emphasis with continuous monitoring should be placed on time and depth measurements of DO and salinity during the summer in the central regions of the mainstem Bay. Emphasis with remote sensing should be placed on characterizing day-to-day variability of chlorophyll-a concentrations and turbidity. In addition, special studies on short-term variability in DO conditions have been conducted as part of the overall CBP program and funded under the Sea Grant program (Dr. Thomas Malone, University of Maryland). Data from these studies should be evaluated to determine the magnitude of short-term variation in the duration and extent of hypoxic/anoxic water.

D. VERTICAL SAMPLING ALLOCATION

The water quality of the Chesapeake Bay varies substantially with depth. Much of this variation is associated with the characteristic two layered estuarine circulation and the degree of stratification that results from it. The steepest vertical variation occurs between 7 and 10 m below the water surface in the pycnocline and is generally greatest in spring and summer and least in fall and winter. As noted in the previous section, vertical variation in water quality is, however, variable in both time and space within seasons. The strongest stratification generally occurs in the central regions of the Bay resulting in summer anoxia. Stratification in headwaters and at the Bay mouth are relatively small. The degree of vertical stratification is also influenced by aperiodic events such as storms.

Because of vertical stratification, measurement of water quality variation with depth is fundamental to any characterization of Bay water quality and is a crucial part of the CBP monitoring program. In fact, analyses of long-term trends and responses of many water quality variables to management actions will likely be limited to specific depth layers (e.g., below pycnocline DO). It is unlikely that responses of water quality to management actions within the pycnocline can be determined using conventional measurement techniques. Water quality in the pycnocline varies on time scales not easily measured by existing instrumentation (i.e., minutes to days) and is sufficient that even if the measurements required could be collected, it is unlikely that responses to management actions would be detected.

The CBP monitoring program consists of two general depth sampling strategies: direct in situ measurements of DO, salinity, conductivity, temperature, and pH and discrete grab samples above and below the pycnocline for chlorophyll-a, nutrients, and total suspended solids.

DO and salinity measurements are used to define the duration and extent of hypoxic/anoxic water and are important variables for evaluation of water quality responses to management actions. DO is primarily a response variable, and salinity is primarily a functional explanatory variable. Isopleth maps and trend analysis are the primary methods used to evaluate long-term trends in DO and salinity. Figure III-1 shows DO and salinity data for 22-24 July 1985, a representative cruise. Additional isopleth maps are presented in Appendix B, Task A. As is visually apparent from these maps, the additional vertical DO and salinity measurements collected by OEP in the pycnocline region of the central Bay allowed for more accurate construction of isopleths, and thus likely more precise estimates of the volume of hypoxic water. Additional measurements within the pycnocline did not appear to be necessary in the lower Bay or lower salinity regions of the upper Bay.

RECOMMENDATION: OEP should continue to take DO and salinity measurements at 1 m depth intervals within the pycnocline for the deep central regions of the Bay. When a pycnocline is present lower Bay data generators should also collect DO and salinity at 1 m depth intervals data within the pycnocline.

OEP has historically taken some of their DO and salinity measurements a 3 m depth intervals. This sampling strategy could result in collection of relatively few samples within the above and below pycnocline layers when the position of the pycnocline is broad and variable. Application of the trend detection methods for continuous variables, like DO, salinity, and uncensored nutrient variables, discussed in the previous chapter require at least two observations for the above and below pycnocline layers at each station. Collection of additional samples within the above and below pycnocline layers

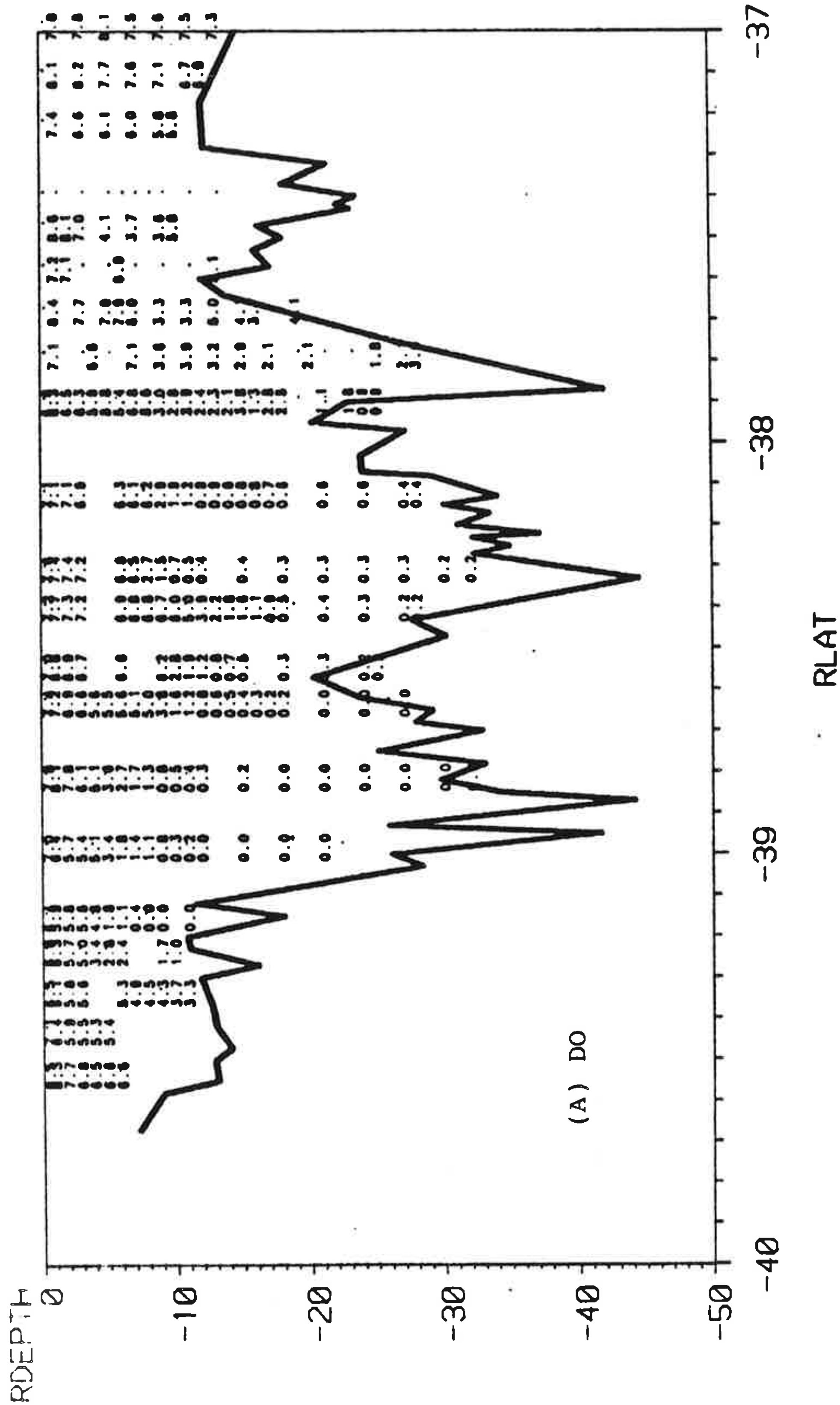


Figure III-1. DO and salinity with depth along the north-south main axis of Chesapeake Bay for cruise 23 (22-24 July 1985)

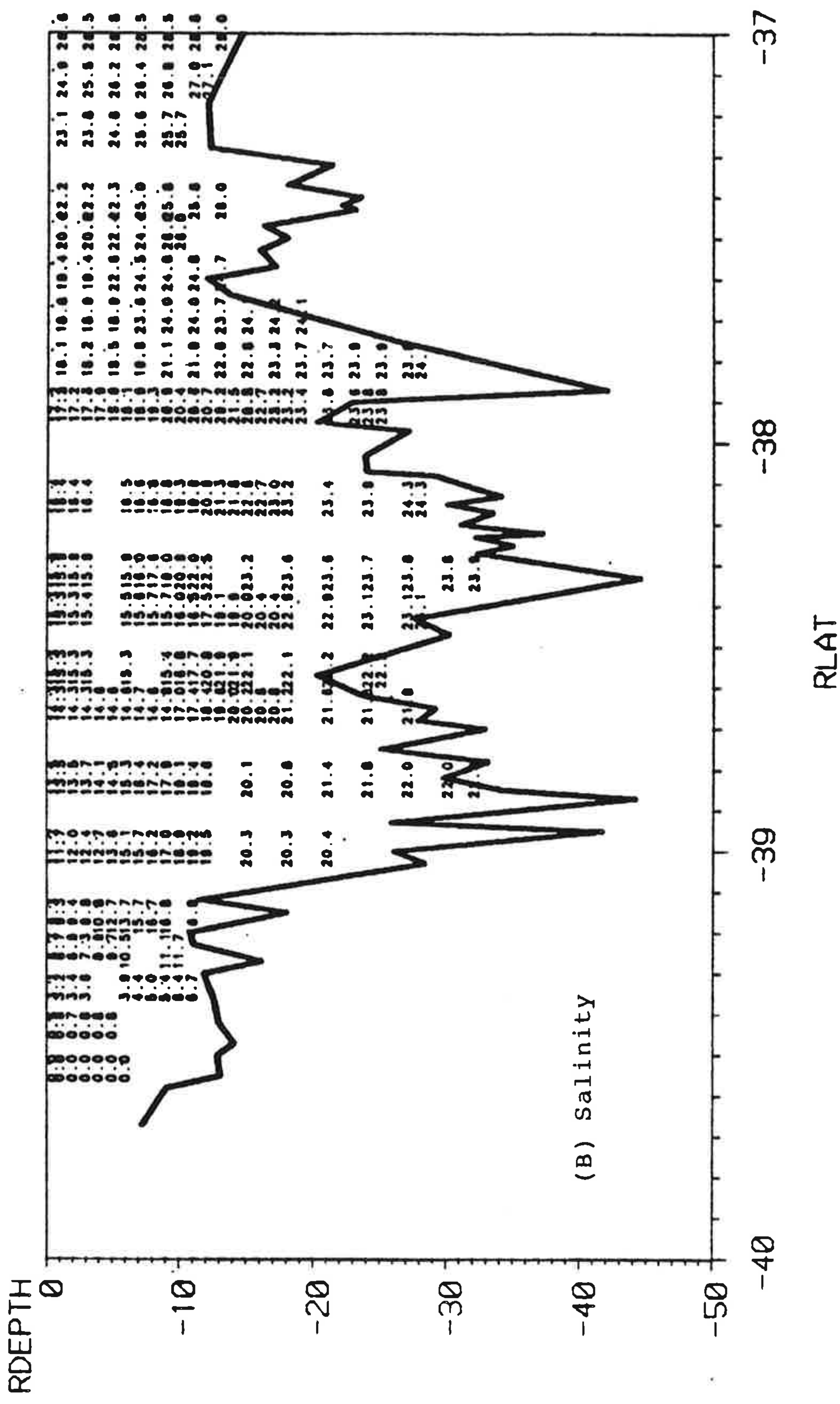


Figure III-1. Continued

would increase the power of most of the trend analysis proposed in the previous chapter and would better describe vertical stratification within the water column.

RECOMMENDATION: As is presently being done, direct measurements of temperature, salinity, conductivity, DO, and pH should be taken at 2 m depth intervals for the above and below pycnocline layers.

Measurement of chlorophyll-a and nutrient concentrations is relatively expensive compared to measurements of DO, salinity, conductivity, temperature, and pH. Because of cost constraints, measurement of these parameters is limited to two depths above the pycnocline (near the surface and 1 m above the upper depth of the pycnocline) and two depths below the pycnocline (near the bottom and 1 m below the lower depth of the pycnocline). The location of the pycnocline is determined by the DO and salinity vertical profiles.

Analysis of the 1984/1985 nutrient and chlorophyll-a data indicated that within a depth layer, nutrients and chlorophyll-a observation varied systematically (Appendix A, Task D) and the two measurements taken within each layer were significantly different. Although the differences were small, this finding suggests the two within-layer samples are not true replicates. If nutrients and chlorophyll-a vary linearly with depth within a layer, combining the two samples likely results in an unbiased estimate of the mean and an over estimate of the variance. If the pattern of variation of nutrients within depth layers is not linear then the direction and magnitude of bias is not known. Pooling the within layer samples for analyses may result in biased estimates of regional conditions and uninterpretable results or erroneous conclusions.

RECOMMENDATION: Alternative sampling strategies for collecting nutrient and chlorophyll-a data for above and below the pycnocline layers of the water column should be identified and compared to data collected by the presently used grab sample method. The goal of this comparison would be to confirm that the sampling strategy presently used results in collection of samples representative of depth layers. An alternative sampling strategy that may be appropriate would be to collect two integrated pump samples for both the above and below pycnocline layers. This sampling strategy would collect a similar total number of samples to that presently collected and could potentially result in better representation of within layer variation. A disadvantage of a sampling strategy based on integrated pump samples is that information on vertical gradients within a depth layer is lost. Evaluation of whether two grab samples per depth layer are sufficient to characterize within-layer water quality requires data that are not collected by the existing mainstem water quality monitoring program. A simple approach for using historical data for verifying the adequacy of taking two samples per depth layer would consist of superimposing the present sampling scheme on historical data with detailed vertical resolution and comparing the estimates of the within-layer means and variances for the complete data set to those resulting from taking two samples per depth layer scheme at the depths sampled by the existing program. The historical data required for this evaluation may be available from a pilot study conducted by OEP during the design phase for the upper Bay monitoring program, other historical studies, or the phytoplankton component of the CBP monitoring program.

At several latitudes in the Bay, data collection is accomplished in the deep central portion of the Bay as well as at lateral stations located east and west of the main axis stations (see Fig. I-1). Total depth for some of the lateral stations is similar to the pycnocline depth (approximately 9 m). As a result, water quality at these lateral stations was more similar to stations located upstream than they were to adjacent main axis stations (see regional groupings of stations, Appendix A, Tasks B and C). When a pycnocline was detected at lateral stations, its close proximity to the bottom resulted in below pycnocline and bottom grab samples occurring at similar depths. Therefore, it may be sufficient to take only one grab sample below the pycnocline at lateral stations.

RECOMMENDATION: Use graphical inspection with existing data to determine whether one grab sample is sufficient for characterization of water quality below the pycnocline at lateral stations.

E. VARIABLES MEASURED AND MEASUREMENT TECHNIQUES

The station array monitored by the three data generators form a north-south gradient in the Bay (Fig. I-1), with OEP stations located between the Susquehanna and Potomac Rivers, VIMS stations located between the Potomac and York Rivers, and ODU stations located between the York River and the mouth of the Bay. Thus, differences in data collection methods or data quality among the data generators may be confounded with naturally occurring north-south differences in Bay water quality and adversely affect inter-regional (spatial) comparisons and measurement of the responses to management actions (see Appendix C). A major reason why it was difficult to use historical data to rigorously assess water quality responses to previous water quality management actions was because measurements were not collected using consistent methods on a Baywide basis.

RECOMMENDATION: Data generators, to the degree possible, should be required to use similar data collection and measurement techniques. Furthermore, to ensure data determined by data generators using the same methods are of similar quality, QA/QC comparisons among data generators using split samples from individual grab samples should be performed regularly. These type of QA/QC comparisons should be expanded from those presently being performed between VIMS and ODU, and should also include OEP. This will greatly assist interpretation of analysis results and will reduce the likelihood of "real" inter-regional responses to management actions being confounded with north-south differences in measurement techniques.

Variables Measured

Major objectives of Bay cleanup efforts are to reduce the extent and duration of anoxia, improve water clarity, and reduce the frequency of blooms by nuisance algae species on the presumption that improved habitat will aid in restoring living resources in the Chesapeake Bay. The major management actions taken to achieve these actions are reductions in nutrient and sediment loadings. The empirical measurements collected by the CBP water quality monitoring program include the important physical variables (salinity, temperature) and water quality parameters (dissolved and particulate nutrients and carbon, DO, chlorophyll-a) required to address these objectives (Table I-1). However, all of the variables measured by the water quality monitoring program are concentration-based measurements. In order to understand the mechanisms and processes affecting Bay water quality leading to anoxia and controlling algal blooms, estimates of the exchange rates among nutrient and chlorophyll-a pools, and between the sediments and overlying water, are required. Process rate data can also be used to check the realism of Bay water quality models and are required to making informed selections among the many possible alternative management actions. For example, a major decision facing Bay water quality managers is whether they should take actions to reduce nitrogen, phosphorus, or both to reduce primary productivity. The decision as to whether to control nitrogen or phosphorus is frequently based on the ratios of nutrient concentrations and experimental and modeling results. For the reasons described above, direct measurement of process rates would provide information useful for making a more informed decision about appropriate nutrient control actions.

RECOMMENDATION: Analyses of the CBP monitoring program data should include information on exchange rates among dissolved and particulate nutrients, chlorophyll-a, and DO for both the water column and the sediments. Only a portion of the required data is presently being collected by other components of the

CBP monitoring program. This information will assist in verifying water quality models, in explaining responses to management actions, in defining research questions, and in making informed selections among the many alternative actions that could be taken to improve Bay water quality. Because of the spatial and temporal variability in water column process rates and the expense associated with collecting these measurements (e.g., Taft et al. 1975), it is not practical to include direct measurement of water column process rates as part of the regular monitoring activities. Some information on water column process rates are also available from the literature (e.g., McCarthy et al. 1975), from the ongoing field studies (Dr. Thomas Fisher, Horn Point Environmental Laboratory, University of Maryland, personal communication), and from water quality modeling activities. Additional process rates can be estimated from concentration-based measurements collected by other components of the CBP monitoring program (e.g., nutrient recycling and algal grazing due to zooplankton). If additional measurements are needed, these measurements need only be made at "key" locations and times. The ongoing ecosystem processes component of the CBP monitoring program should be used to define the location and frequency of measurements.

Measurement Techniques

While the major water quality variables of interest were all ultimately reported by each data generator, there were some differences in how these values were obtained. Of particular importance were differences in measurement methods for determining dissolved and particulate nutrient forms including:

- Differences in how the dissolved and particulate forms were determined. For example, OEP directly measured particulate organic nitrogen, whereas VIMS and ODU calculated particulate organic nitrogen as the difference between total Kjeldahl nitrogen and ammonia.
- Differences in analytical measurement techniques of the same variable. For example, all three data generators directly measured dissolved inorganic phosphorus concentrations. However, because of different analytical techniques, minimum detection concentrations differed between OEP (0.0016 mg P/L) and ODU/VIMS (0.01 mg P/L).
- Differences in computational methods used for estimating detection concentrations.

Among-data generator differences in nutrient analytical methods introduce artificial biases into the data complicating interpretation of analysis results. For example, during periods of low nutrient concentration (e.g., summer and fall), N:P ratios along the main axis of the Bay become confounded with differences in detection limits among the data generators. As a result, N:P ratios become the ratios of detection limits and are not reflective of actual regional differences in nitrogen or phosphorus levels. In Deliverable 3, we reviewed and evaluated analysis approaches for trend detection. This review indicated that the most appropriate way to conduct trend analysis was after aggregating the data into regional and seasonal groups. If this analysis approach is used erroneous conclusions that are partially, or wholly, due to differences in analytical methods, rather than on "real" differences in how regions of the Bay are responding to management action, may result.

RECOMMENDATION: All data generators should be required to use identical measurement methods for nutrients and chlorophyll-a for both main Bay and tributary monitoring and the required measurement methods should be those that result in most accurate determinations of important nutrient forms. Because of the interrelated manner in which some nutrient forms are compiled, partial resolution of differences in measurement methods among data generators would not adequately rectify this deficiency in the current monitoring program.

IV. DATABASE PROCESSING

Maintenance of and access to a centralized database is the most effective means of dissiminating the data collected for the CBP monitoring program to the many agencies, organizations, and institutions that require access to it. Because a major use of the water quality monitoring data is as inputs for analyses and models used for developing water quality management strategies and making other resource management decisions, it is critical that the data in the database be of a high quality. As noted in previous sections, however, establishment and maintenance of a centralized database is a difficult task mainly because the data are collected by different institutions using slightly different sampling designs and methods. The historical data were also collected by many institutions, and many instances the detailed sampling specifications and therefore the quality of these data are unknown.

The CBP centralized database is maintained by the Computer Sciences Corporation (CSC) in Annapolis, MD at the Chesapeake Bay Liaison Office. CSC is responsible for entry of data into the database and for ensuring the data is available to authorized users. One of CSC's initial efforts has been to develop a rigorous QA/QC program for assuring the quality of the CBP water quality monitoring data in the database. This QA/QC program was in early stages of development when data were first transferred from CSC to Martin Marietta Environmental Systems. Although CSC now has a formal QA/QC program, it is still being adapted and modified to address specific problems as they are identified. The objective of this chapter is to make recommendations relevant to the QA/QC program based on problems encountered during analysis of the data. These recommendations are not intended to constitute the complete array of QA/QC procedures that should be applied to new (or old) data in the CBP database. Rather, they are presented to document procedures that we applied successfully to identify erroneous data and other problems before the conduct of analyses. CSC and the CBP can use this information to determine if any of these procedures should be incorporated into the general QA/QC protocol for the CBP database.

The first step in the Martin Marietta Environmental Systems QA/QC procedures was to determine if all stations were properly labeled. This was followed by a check to determine if the numbers of observations reported for each station was equal to the expected number. The expected number of observations was calculated for each sampling location based on sampling protocols and data provided by CSC and the CBP Liaison Office. These QA/QC checks resulted in the production of a series of tables that listed:

- Station labels that were not anticipated
- The number of non-missing observations for each variable for each station by data generator (Table IV-1 is an example for Cruise 5 for stations sampled by VIMS)
- Stations that had two or more records for the same depth (Table IV-2 is an example for OEP stations)
- Stations and cruises which had a greater number of observations than expected (Table IV-3 is an example for ODU station CB7.4 for cruise 23).

Once problems with the numbers of observations were identified and corrected by discussions with data generators, the CBP Liaison Office, and CSC, the data were screened for inappropriate values. The first part of the screening process checked the data for inappropriate variable types. For instance, numeric variables incorrectly defined as character variables, or inconsistent symbols used to denote values of variables reported at detection limits (e.g., the letter "D" instead of the symbol "<").

Next, the data were checked for potentially erroneous values (e.g., values that exceeded normal ranges). Because the 1984/1985 data were the initial entries into the CBP database, historical data were used to define normal ranges for each station-date combination. As additional data are collected and entered into the CBP database, it should be pooled (stations into regions, cruises into seasons, depth specific data into above pycnocline/below pycnocline values) and used to compute frequency distributions for each season, region, and depth layer combination. These frequency distributions should then be used to define normal ranges, and extreme values for each variable (e.g., those in upper and lower 5% of recorded values) should be identified and verified to determine if they are unrealistic.

Detection limits in the CBP water quality monitoring data vary spatially and temporally among data generators. This situation will likely persist in the future as new analytical methods are developed and implemented into the monitoring program because it is unlikely that all data generators will implement new procedures at the same time. Table IV-4 is an example of a simple procedure used by Martin Marietta Environmental Systems to ensure that detection limits for each data generator are consistent with previous data. Table IV-4 shows all of the detection values reported by ODU for each water quality variable.

Table IV-1. An example of a table produced as part of Martin Marietta Environmental Systems QA/QC procedures applied to the CBP mainbay monitoring data showing the number of nonmissing observations for cruises and stations monitored by VIMS

CRUISE NUMBER : BAY005
DATE OF CRUISE: AUG 29 84

STATION	DATE	COND	SI	TOC	DOC	TSS	TKNW	TKNF	NH4	NO23	NO2	NO3	TDP	TP	CHLA	DON	PON	TDN	TN	DOP	POC	SECCHI	DISOXY	PH	SALIN	WTEMP
CB5.5	12	11	2	2	2	2	2	1	2	2	2	2	2	2	1	1	1	1	2	2	2	1	11	2	11	11
CB6.1	8	7	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	7	2	7	7
CB6.2	7	6	0	0	0	2	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	6	2	6	6
CB6.3	7	7	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	0	7	2	7	7
CB7.1	15	14	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	14	2	14	14
CB7.1S	9	8	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	8	2	8	8
CB7.2	12	11	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	11	2	11	11
CB7.2E	8	7	1	1	1	2	1	0	1	1	1	1	1	1	2	0	0	0	1	1	1	1	7	2	7	7
EE3.1	3	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	0	2
EE3.2	17	15	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	16	2	15	16
LE3.6	6	5	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	5	2	5	5
LE3.7	5	4	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	4	2	4	4
WE4.1	4	3	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	3	2	3	3
WE4.2	8	7	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	7	2	7	7
WE4.3	4	3	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	3	2	3	3
WE4.4	5	4	2	2	2	2	2	1	2	2	2	2	2	2	2	1	1	1	2	2	2	1	4	2	4	4

STATIONS NOT FOUND:

CB5.3, CB5.4W, CB5.4, CB7.1N

Table IV-2. An example of a table produced as part of Martin Marietta Environmental Systems QA/QC procedures applied to the CBP mainbay monitoring data. OEP stations that had two or more observations at the same depth are identified by the label A

SOURCE: MARYLAND OEP
DATE : 3-31-86

CRUISE

STATION	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
CB3.3C			A	A	A	A			A	A					A		A			A		A	A	A	A	A	
CB4.1C										A																A	
CB4.1E			A						A					A	A		A										
CB4.2E																											
CB4.3C					A							A			A					A			A	A	A	A	A
CB4.4						A				A							A	A					A	A	A	A	A
CB5.2				A		A				A	A						A	A		A		A	A	A	A		A
CB5.3																											
LE2.3																			A		A						
CB5.1																A											
CB3.1											A				A						A						
CB4.3E																											

LEGEND: A = 2 records occur for the same depth with different values for nutrient variables
Possible lab rep??

NOTE : All stations had water quality data sampled at 1 meter intervals. Sampling of nutrients occurred at 0.5 m, bottom and at 2 points between even values (e.g. 8.5m & 24.5m).

Table IV-3. An example of a table produced as part of Martin Marietta Environmental Systems QA/QC procedures applied to the CBP mainbay monitoring data showing observed values at station CB7.4N on cruise 23, which had a greater number of observations than expected

---CRUISE=BAY023 STATION=CB7.4N-----------------------------------

SDEPTH	SEGMENT	SOURCE	CHLA_D	PHEA_D	PH	TKNF	NO2	NO3	TOC	DOC
0	CB7	VA/ODU			.	.	.	.	.	.
1	CB7	VA/ODU	.	<	8.25000	0.430000	0.000500000	0.0100000	1.74263	1.09869
1	CB7	VA/ODU		<	8.25000	0.280000	0.000500000	0.0100000	1.25382	1.15122
3	CB7	VA/ODU			.	.	.	.	.	.
3	CB7	VA/ODU			.	.	.	.	.	.
5	CB7	VA/ODU			.	.	.	.	.	.
5	CB7	VA/ODU			.	.	.	.	.	.
7	CB7	VA/ODU			.	.	.	.	.	.
7	CB7	VA/ODU			.	.	.	.	.	.
9	CB7	VA/ODU			.	.	.	.	.	.
9	CB7	VA/ODU			.	.	.	.	.	.
11	CB7	VA/ODU			.	.	.	.		.
11	CB7	VA/ODU				.	.		.	
12	CB7	VA/ODU		<	8.20000	.	0.000500000	0.0100000	1.23287	1.15974
12	CB7	VA/ODU		<	8.25000	.	0.000500000	0.0100000	1.29819	0.50000

SDEPTH	POC	TSS	COND	DATE	TP	TDP	TKNF_D	NO2_D	TOC_D	DOC_D	TN	DISOXY
0	.	.	.	22JUL85	.	.					.	.
1	1	4	44670.0	22JUL85	0.0500000	0.0410000		<			0.385500	7.80000
1	1	4	.	22JUL85	0.0649999	0.0510000		<			0.355500	.
3	.	.	44950.0	22JUL85	.	.						8.00000
3	.	.	.	22JUL85	.	.					.	.
5	.	.	44990.0	22JUL85	.	.					.	7.80000
5	.	.	.	22JUL85	.	.					.	.
7	.	.	44970.0	22JUL85	.	.					.	7.80000
7	.	.	.	22JUL85	.	.					.	.
9	.	.	44970.0	22JUL85	.	.					.	7.80000
9	.	.	.	22JUL85	.	.					.	.
11	.	.	44990.0	22JUL85	.	.					.	7.70000
11	.	.	.	22JUL85		.						.
12	1	.	45000.0	22JUL85	0.0420000	.		<			0.445500	7.60000
12	1	.	.	22JUL85	0.0709999	.		<		.	0.385500	

Table IV-4. An example of a table produced as part of Martin Marietta Environmental Systems QA/QC procedures applied to the CBP mainbay monitoring data showing the number of distinct detection concentrations and their values for all variables measured at ODU stations

	COND	SI	TOC	DOC	TKNW	TKNF	NH4	NO23	NO2	NO3	TDP
# of VALUES:	0	1	1	1	0	1	2	163	1	1	1
VALUES:											
		0.0280	1.0000	1.0000		0.1000	0.0100	0.0344	0.0010	0.0100	0.0100
							0.0056	0.0327			
								0.0544			
								0.0311			
								0.0394			
								0.0361			
								0.0411			
								0.0294			
								0.0244			
								0.0461			
								0.0377			
								0.0752			
								0.0110			
								0.0116			
								0.0129			
								0.0673			
								0.0153			
								0.0211			
								0.0168			
								0.0440			

	TP	CHLA	DON	PON	TDN	TN	DOP	POC	DISOXY	PHEA	PO4F	PHOSP
# of VALUES:	0	5	86	29	160	328	95	410	0	0	1	20
VALUES:												
		-5.407	2.3000	-.0200	1.0744	1.1244	-.0100	0.7500			0.0100	-.0050
		-.0000	2.5900	-.1100	1.1544	1.1427	-.0040	-.0500				0.0100
		-9.767	2.9800	-.0500	0.9794	1.1344	0.0380	0.5900				-.0030
		-30.42	3.2700	-.1900	0.9811	1.0711	0.0490	-.0600				-.0380
		-.5874	3.0200	0.3200	1.1194	1.1194	0.1200	-.2200				-.0050
			2.4900	0.3700	1.0744	1.1061	0.0550	0.4300				0.0650
			2.7400	0.5100	1.0361	1.0711	0.0980	0.0800				0.0380
			2.7800	0.4200	1.0661	1.0594	0.0210	-.4600				-.0070
			0.2200	0.2700	2.3852	1.0094	0.0010	0.4300				-.0150
			0.3000	0.6600	2.6116	0.8827	0.0340	0.2700				0.0530
			0.0880	-.0100	3.0573	1.2544	0.0120	0.4300				-.0070
			0.0740	-.0400	3.2916	1.1161	0.0140	0.0200				-.0040
			0.0810	-.0200	3.0410	0.9861	0.0080	-.5000				.0030
			0.0720	-.0100	2.5110	1.0677	0.0290	0.6400				-.0020
			0.0860	-.0700	2.7610	1.9061	0.0180	0.4000				-.0050
			0.2800	-.0100	0.3653	0.9161	0.0160	0.4400				-.0010
			0.2400	-.0300	0.5668	3.5452	0.0200	0.2600				-.0040
			0.3200	-.0300	0.2769	4.3110	0.0100	0.2300				-.0040
			0.2500	-.0300	0.2839	4.1816	0.0230	0.4400				-.0010
			0.2900	-.0700	0.3912	5.0629	-.0060	0.4100				-.0010

The final step in Martin Marietta Environmental Systems QA/QC checks was random comparison of selected data in the datafiles against original data in reports. CSC should follow a similar but even more rigorous procedure and systematically contrast a subsample (randomly or arbitrarily selected) of values in the database against original data reports or field data sheets. Industrial Quality Control techniques should then be used to estimate error rates (e.g., see Chapter 17 in Duncan 1974) and corrections should be made to datafiles until the actual error rate is within pre-established limits of acceptability. We are aware that data generators are required to verify their data entries before transmitting data tapes to CBP and recommend that this procedure continue in the future. The above recommended checking by CSC against original data sheets is in addition to the line-by-line checking required of data generators. We feel these additional checks are necessary because our experience with SAS databases has shown us that the efficiency of line-by-line checks vary among individuals and agencies depending upon the level of sophistication of the computer software used to assist with this checking, the design of data sheets, and the amount of coding/data transfer that occurs before keypunching. If the data in the database are not shown to be reliable then they will not be used by decision makers, scientists, or the public regardless of their accessiblity.

RECOMMENDATION: Before data (new or historical) are incorporated into the CBP database QA/QC procedures must be accomplished that identify:

- Improperly labeled stations and variables
- Stations/variables where expected observations differ from the actual number that were taken
- Inappropriate variable types
- Values that are in a range that suggest they may be erroneous
- Stations, variables, or data for which detection limits are not consistent with previous data
- Data that do not agree with values in original reports.

V. CONCLUDING REMARKS

In this final report, we have presented recommendations on aspects of data collection, database processing, and data analysis of the present CBP main Bay water quality monitoring program. While there are aspects of the program that can be improved, we also want to emphasize that the overall approach of the program is sound and will provide the data needed to characterize and detect trends in Chesapeake Bay water quality. Continuation of this coordinated monitoring approach provides the best opportunity for generation of statements concerning trends in Bay water quality and for providing an ecologically sound basis for the implementation of pollution-control management actions.

VI. LITERATURE CITED

Cochran, W.G. 1977. Sampling Techniques. John Wiley and Sons. NY, 428 p.

Downing, J.A. 1979. Aggregation, transformation, and the design of benthic sampling programs. J. Fish. Res. Bd. Can. 36:1454-1463.

Duncan, A.J. 1974. Quality Control and Industrial Statistics. R. Irwin: Homewood, IL.

Flemer, D.A., and others. 1983a. Chesapeake Bay: A profile of environmental change. U.S. Environmental Protection Agency.

Flemer, D.A., and others. 1983b. Chesapeake Bay: A profile of environmental change, Appendices. U.S. Environmental Protection Agency.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley and Sons: New York.

Heinle, D.R., C.F. D'Elia, J.L. Taft, J.S. Wilson, M. Cole-Jones, A.B. Caplins, and E. Cronin. 1980. Historical review of water quality and climatic data from Chesapeake Bay with emphasis on effects of enrichment. Chesapeake Research Consortium Publication No. 84.

Lippson, A.J., M.S. Haire, A.F. Holland, F. Jacobs, J. Jensen, R.L. Moran-Johnson, T.T. Polgar, and W.R. Richkus. 1979. Environmental Atlas of the Potomac Estuary. Prepared for Maryland Department of Natural Resources, Power Plant Siting Program by Martin Marietta Corporation, Baltimore, MD. 280 p.

McCarthy, J.J., W.R. Taylor, and J.L. Taft. 1975. The dynamics of nitrogen and phosphorus cycling in the open waters of the Chesapeake Bay. pp. 664-681. In: Marine Chemistry in the Coastal Environment. T.M. Church (ed.). American Chemical Society Symposium Series 18, Washington, DC.

Millard, S.P., and D.P. Lettenmaier. 1986. Optimal design of biological sampling programs using analysis of variance. Estuarine, Coastal and Shelf Science 22:637-656.

Officer, C.B., R.B. Biggs, J.L. Taft, L.E. Cronin, M.A. Tyler, and W.R. Boynton. 1984. Chesapeake Bay anoxia: origin, development, and significance. Science 223:22-27.

Reed, G.J., and A.J. McErlean. 1979. A preliminary report on selected and developed indices for use in the detection, measurement, and assessment of estuarine nutrient enrichment. Chesapeake Research Consortium Publication No. 70.

Seliger, H.H., J.A. Boggs, and W.H. Biggley. 1985. Catastrophic anoxia in the Chesapeake Bay in 1984. Science 228:70-73.

Taft, J.L., W.R. Taylor, and J.J. McCarthy. 1975. Uptake and release of phosphorus by phytoplankton in the Chesapeake Bay Estuary, USA. Marine Biology 33:21-32.