Data Management Strategy
for the Tampa Bay National Estuary Program:

Recommendations and Implementation Plan

Prepared for:

Tampa Bay National Estuary Program
111 7th Ave. South
St. Petersburg, FL 33701

Prepared by:

Coastal Environmental Services, Inc.
1099 Winterson Rd., Suite 130
Linthicum, MD 21090

August 1992

Printed on Recycled Paper
FOREWORD

This report dated August 1992 and titled "Data Management Strategy for the Tampa Bay National Estuary Program: Recommendations and Implementation Plan" was prepared by Sara J. Cairns and Anthony J. Janicki of Coastal Environmental Services, Inc. for the Tampa Bay National Estuary Program, as part of the Data Management Strategy project. All work was prepared under a contract entered into on 15 November 1991 by and between the Tampa Bay Regional Planning Council in behalf of Tampa Bay National Estuary Program and Coastal Environmental Services, Inc. This is Technical Publication #08-92 of the Tampa Bay National Estuary Program.
ACKNOWLEDGMENTS

The authors of this report would like to thank Richard Eckenrod and Holly Greening of the Tampa Bay National Estuary Program for their assistance in completing this project. We are also grateful to the 49 people who responded to our Survey of Data User Needs and Goals: their willingness to take the time and effort to reply to our inquiries was critical to our being able to develop recommendations based on actual rather than hypothetical needs of data users in the Tampa Bay area. We are similarly indebted to the participants of the Data Management Strategy Workshop held on 3-4 February 1992, who contributed their comments and suggestions regarding effective data management practices. David Stage (Office of the Governor) was extremely helpful in providing information on the work of the Growth Management Data Network Coordinating Council and the Florida Spatial Data Directory.

We would also like to acknowledge the able assistance of David Wade, Peggy Derrick, Douglas Heimbuch, and Susan Janicki (Coastal Environmental Services, Inc.) in conducting the survey and workshop.
ABSTRACT

One goal of the Tampa Bay National Estuary Program (TBNEP) is to identify and develop the plans for a data management system that meets the needs of potential users who can benefit from improved access to information that is pertinent to resource management. To meet this goal the TBNEP contracted with Coastal Environmental Services, Inc. to develop a Data Management Strategy (DMS) for Tampa Bay. As part of this project the needs and goals of data users in the Tampa Bay area were investigated by means of a survey, followed by a workshop at which participants discussed how the DMS could improve on existing information resources. The recommendations resulting from the survey and workshop discussions focused on (1) developing a Central Subject Directory (CSD) to existing data, (2) standardizing data management forms and protocols used by different agencies, and (3) providing data management services for the TBNEP throughout its tenure. The purpose of the CSD will be to permit users to locate existing data rapidly, easily, and thoroughly. It will be created using information from an existing computer database, and will be designed to be an element of a statewide directory being developed by the Florida Growth Management Data Network Coordinating Council (FGMDNCC). The standardized forms and protocols for data documentation will be based on practices already in use wherever possible, and will be developed in close consultation with local agencies. Many local agencies have signed a Memorandum of Understanding (April 1992) as participants in the Tampa Bay Regional Coordinating Council, as part of the FGMDNCC initiative, in which they agree to: submit entries to a CSD describing their in-house databases; meet minimum standards for in-house documentation of data; and routinely provide this documentation when transferring data to other agencies. Those agencies which have not signed the Memorandum of Understanding will be urged to participate in the initiative. The data management needs of the TBNEP will be met by (1) providing technical support to TBNEP contractors when they submit their data to the Environmental Protection Agency's Ocean Data Evaluation System, (2) providing suitable forms and protocols for TBNEP contractors to use in documenting the data they collect, and (3) providing data reduction services, including data summaries, maps, and presentation materials, on an as-needed basis.
TABLE OF CONTENTS

FOREWORD ............................................................. ii
ACKNOWLEDGEMENTS ........................................... iii
ABSTRACT ............................................................. iv

LIST OF TABLES ....................................................... ix
LIST OF FIGURES ..................................................... x

1. INTRODUCTION ...................................................... 1 - 1

2. DATA USER NEEDS IN THE TAMPA BAY AREA .................. 2 - 1
   2.1 SURVEY OF DATA USER NEEDS AND GOALS ............... 2 - 1
       2.1.1 Methods .................................................. 2 - 1
       2.1.2 Results .................................................. 2 - 2
   2.2 DATA MANAGEMENT STRATEGY WORKSHOP .................. 2 - 4
       2.2.1 Workshop Goals and Charges to the Participants .... 2 - 5
       2.2.2 Summary of Information Needs and Resources in
               the Tampa Bay Area .................................. 2 - 5
               First charge: Identify and assign priorities
               to user needs ............................................ 2 - 6
               Second charge: Features of existing
database
               management systems .................................. 2 - 7
               Third charge: Realistic modifications and
               additions to existing systems ....................... 2 - 9
       2.2.3 Presentations and Demonstrations .................... 2 - 10
       2.2.4 Conclusions ............................................ 2 - 10
TABLE OF CONTENTS (cont.)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. DATA MANAGEMENT RECOMMENDATIONS</td>
<td>3 - 1</td>
</tr>
<tr>
<td>3.1. CENTRAL SUBJECT DIRECTORY</td>
<td>3 - 2</td>
</tr>
<tr>
<td>3.1.1 Justification</td>
<td>3 - 2</td>
</tr>
<tr>
<td>3.1.2 Characteristics</td>
<td>3 - 4</td>
</tr>
<tr>
<td>3.1.3 Recommended Directory Contents</td>
<td>3 - 5</td>
</tr>
<tr>
<td>3.1.4 Evaluation of Existing Directories</td>
<td>3 - 6</td>
</tr>
<tr>
<td>Florida Spatial Data Directory (FSDD)</td>
<td>3 - 6</td>
</tr>
<tr>
<td>Contents</td>
<td>3 - 7</td>
</tr>
<tr>
<td>Search and subset capabilities</td>
<td>3 - 7</td>
</tr>
<tr>
<td>User time investment</td>
<td>3 - 8</td>
</tr>
<tr>
<td>Required facilities</td>
<td>3 - 8</td>
</tr>
<tr>
<td>Bibliographic Database (BDB)</td>
<td>3 - 8</td>
</tr>
<tr>
<td>Contents</td>
<td>3 - 9</td>
</tr>
<tr>
<td>Search and subset capabilities</td>
<td>3 - 9</td>
</tr>
<tr>
<td>User time investment</td>
<td>3 - 10</td>
</tr>
<tr>
<td>Required facilities</td>
<td>3 - 10</td>
</tr>
<tr>
<td>3.1.5 Recommended Structure of the Central Subject Directory</td>
<td>3 - 10</td>
</tr>
<tr>
<td>3.2 DECENTRALIZED COMPONENTS: DATA MANAGEMENT PROTOCOLS AND RECORDS</td>
<td>3 - 11</td>
</tr>
<tr>
<td>3.2.1 Data Collection</td>
<td>3 - 12</td>
</tr>
<tr>
<td>Review of data sheets</td>
<td>3 - 13</td>
</tr>
<tr>
<td>Quality assurance and quality control</td>
<td>3 - 13</td>
</tr>
<tr>
<td>Storing of data sheets</td>
<td>3 - 13</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (cont.)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2 Computer File Construction and Maintenance</td>
<td>3-14</td>
</tr>
<tr>
<td>Data entry and verification</td>
<td>3-14</td>
</tr>
<tr>
<td>Dataset description (documentation)</td>
<td>3-15</td>
</tr>
<tr>
<td>Study overview documentation</td>
<td>3-15</td>
</tr>
<tr>
<td>Methods and sampling design documentation</td>
<td>3-16</td>
</tr>
<tr>
<td>Dataset contents documentation</td>
<td>3-18</td>
</tr>
<tr>
<td>Dataset quality documentation</td>
<td>3-19</td>
</tr>
<tr>
<td>Maintenance</td>
<td>3-20</td>
</tr>
<tr>
<td>3.2.3 Data Transfers</td>
<td>3-20</td>
</tr>
<tr>
<td>Basic transfer</td>
<td>3-21</td>
</tr>
<tr>
<td>Describing the request</td>
<td>3-21</td>
</tr>
<tr>
<td>Transfer format</td>
<td>3-21</td>
</tr>
<tr>
<td>Dataset description</td>
<td>3-21</td>
</tr>
<tr>
<td>Data reduction transfers</td>
<td>3-23</td>
</tr>
<tr>
<td>3.2.4 Summary</td>
<td>3-24</td>
</tr>
<tr>
<td>3.3 TBNEP PROGRAMMATIC DATA MANAGEMENT NEEDS</td>
<td>3-25</td>
</tr>
<tr>
<td>3.3.1 Short-term Data Needs</td>
<td>3-26</td>
</tr>
<tr>
<td>3.3.2 Data Reduction Services</td>
<td>3-27</td>
</tr>
<tr>
<td>3.3.3 Public Education Needs</td>
<td>3-27</td>
</tr>
<tr>
<td>3.3.4 ODES Data Submission Guidelines</td>
<td>3-28</td>
</tr>
<tr>
<td>3.4 SUMMARY</td>
<td>3-28</td>
</tr>
<tr>
<td>4 IMPLEMENTATION PLAN</td>
<td>4-1</td>
</tr>
<tr>
<td>4.1 SHORT-TERM (FIRST YEAR) IMPLEMENTATION</td>
<td>4-1</td>
</tr>
<tr>
<td>4.1.1 Central Subject Directory (CSD)</td>
<td>4-1</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (cont.)

Page

4.1.2 Standard Data Management Forms and Protocols ........ 4 - 3
4.1.3 TBNEP Support ........................................ 4 - 4

4.2 LONG-TERM IMPLEMENTATION .............................. 4 - 5
4.2.1 Central Subject Directory .............................. 4 - 5
4.2.2 Standard Data Management Forms and Protocols ........ 4 - 6

4.3 SUMMARY .................................................... 4 - 6

5 REFERENCES ..................................................... 5 - 1

APPENDICES

A Survey of Data User Needs and Goals - Survey Instrument
B Survey of Data User Needs and Goals - Summary of Responses
C Sample Data Documentation Forms and Protocols
D TBNEP Dataset Documentation Package
E ODES Data Submission Guidelines
LIST OF TABLES

Table 2.1  Agencies from which responses to the Survey of User Needs and Goals were received. ................. 2 - 12

Table 3.1  Recommended characteristics of the central subject directory .............................................. 3 - 30

Table 3.2  Recommended contents of the central subject directory ......................................................... 3 - 32

Table 3.3.  Match in variables between the Bibliographic Database (BDB), the central subject directory (CSD), and the Florida Spatial Data Directory (FSDD) ........................................ 3 - 33

Table 4.1.  Activities required for long-term maintenance of the Central Subject Directory ...................... 4 - 8

Table 4.2.  Chronology for the first year of the implementation of the Data Management Strategy .............. 4 - 9
LIST OF FIGURES

Figure 1.1. Summary of events and final products for the Tampa Bay National Estuary Program's Data Management Strategy project ............... 1 - 3

Figure 2.1. Inter-agency transfers of data reported in the Survey of Data Users Needs and Goals ..................... 2 - 1
CHAPTER 1. INTRODUCTION

The Tampa Bay National Estuary Program (TBNEP) is intended to address a number of critical issues related to the natural resources of the Tampa Bay area. Its goals include assessing trends in natural resources, identifying causes of environmental problems, assessing pollutant loadings, preparing a Comprehensive Conservation and Management Plan (CCMP), coordinating activities among a variety of participants, and monitoring the effectiveness of actions taken under the CCMP. Each of these activities will depend heavily on access to data on natural resources and environmental conditions in the Bay area.

The TBNEP is sponsoring data collection projects designed to determine current conditions in the Bay. Federal, state, regional, and local agencies are involved in many ongoing research and monitoring activities that generate information critical to resource management in Tampa Bay. Extensive historical data also exist that have been collected by projects ranging from local special-interest studies to Bay-wide surveys: these data will be critical to evaluating how conditions in the Bay have changed in recent times.

The usefulness of any data in contributing to resource management decisions will depend on how well they are or have been documented. A number *per se* is useless no matter how accurately it measures some quantity (e.g., phosphate levels in the Alafia River) unless that number is

- identified accurately (defined, with units of measurement),
- linked to documentation describing the sampling methods and quality control procedures used in collecting and processing the data, and
- stored in a manner and location that makes it accessible.

Proper data management is the process that ensures that data are defined, documented, and kept accessible for future use.

The large quantity of existing information on the natural resources of Tampa Bay provides both an opportunity and a challenge. The opportunity is to use existing data to extend our understanding of the Bay’s ecology and to answer specific questions without wasteful (or impractical) duplicate data collection efforts. The challenge is to do so without wasting an equivalent amount of effort in locating and accessing data that are stored in a bewildering variety of sites and formats.

The goal of the TBNEP’s Data Management Strategy (DMS) project is to
identify and develop the plans for a data management system that meets the needs of potential users who can benefit from improve access to information that is pertinent to resource management. Coastal Environmental Services, Inc., was contracted by TBNEP to develop the Data Management Strategy plans.

We used a sequential approach, soliciting input from data producers and users in the Tampa Bay area, and submitting preliminary results to TBNEP for review before proceeding with the next phase (Figure 1.1). We collected information on the needs of data users and existing database management systems through a survey of resource managers in the Tampa Bay area. We also conducted a workshop at which data producers and users discussed their information needs and resources, and were asked to contribute input to the TBNEP Data Management Strategy.

In this report we present our recommendations regarding specific steps that can be taken to create and maintain an effective data management strategy for Tampa Bay. The recommendations are based on our investigation of the needs and goals of data users in the Tampa Bay area, on the current state of information resources and data transfer practices in the area, and on basic tenets of data management.
Figure 1.1. Summary of events and final products for the Tampa Bay National Estuary Program's Data Management Strategy project.
CHAPTER 2. DATA USER NEEDS IN THE TAMPA BAY AREA

In order to make sound recommendations regarding a Data Management Strategy (DMS) for the Tampa Bay National Estuary Program (TBNEP) we first needed extensive information on the data needs of resource managers. We conducted a survey of data users in the Tampa Bay area that was designed to determine the information needs and current practices of resource managers, professionals, and educators. We then conducted a workshop at which the participants discussed the results of the survey and provided additional input. The following sections describe how the survey was conducted and what the results were, as well as what issues were raised at the Data Management Strategy workshop and what conclusions were reached.

2.1 SURVEY OF DATA USER NEEDS AND GOALS

In December 1991 and January 1992, Coastal Environmental Services, Inc., conducted a Survey of Data User Needs and Goals for the TBNEP. The goal of the Survey was to determine the information needs and practices of resource managers, professionals, and educators in the Tampa Bay area. Specific topics of interest were: what natural resources information is needed by each survey respondent; how he or she currently accesses this information; and what services that are not now available would improve access to information. Survey responses would be used to help determine specific recommendations for the TBNEP’s Data Management Strategy.

2.1.1 Methods

Our target population for the survey was individuals in the Tampa Bay area whose jobs require access to natural resources information: particularly resource managers. We used an inventory prepared by the Tampa Bay Regional Planning Council ("The Future of Tampa Bay", 1984) to prepare a list of target organizations. The inventory was of those federal, state, regional, and local government agencies that had jurisdiction over activities associated with Tampa Bay. We identified individuals to contact within each organization (and updated the agency list when necessary) using (a) conversations with TBNEP staff, (b) the TBNEP Technical Advisory Committee membership list, and (c) calls to organizations and individuals.

The survey instrument that we developed (Appendix A) had 16 questions in four general categories:
• Respondent background (e.g., type of job, types of projects worked on)

• Type of natural resources information needed by the respondent (e.g., dates, subject matter)

• Current data transfer methods used (e.g., sources, hardware and software)

• Suggestions for future enhancements to data access (e.g., centralization of information, access to information processing services).

All questions were multiple choice or short-answer, and a table of existing data bases was included to help respondents respond to questions regarding their recent data transfer experiences.

We mailed survey instruments to a total of 71 people. Of these, 24 had been contacted in advance and agreed to respond to the survey via a telephone interview: the mailing provided them with a copy of the instrument to look over during the interview. The other 47 people received a survey with a cover letter explaining the purpose of the survey and asking that they fill out the form and return it to Coastal in an enclosed, stamped envelope. The telephone interviews were conducted primarily during the week of 6-10 January, 1992. The last mailed response we received arrived on 30 January 1992.

2.1.2 Results

The response rate to the survey was 92% (22 of 24) for those who agreed in advance to a phone interview, and 57% (27 of 47) for those responding by mail. We received surveys from 26 government agencies and 5 professional organizations (Table 2.1).

The survey was targeted primarily at management-level users, and 54% of respondents had worked as a project manager for projects requiring information on the natural resources of Tampa Bay. (A complete compilation of responses is provided in Appendix B.) The type of projects that respondents had worked on in the last year were primarily planning, regulatory, or advisory, although 37% also reported working on research or educational projects. The types of information that respondents had needed to locate in the last year fell into the general subject areas of surface and ground water, ecology and biology, human population
impacts, and meteorology. Specific examples were highly varied, both in subject matter and the geographic area for which data were needed.

In locating data during the last year, survey respondents reported that they had used a total of 37 datasets from more than a dozen agencies (Figure 2.1). Inter-agency data access was clearly a common activity, involving many different agencies. The variety of subject areas for which information was needed, and the variety of data sources, is probably reflected in the nearly unanimous opinion of survey respondents that a guide to existing information should be made available at a central location. Only 36% of respondents thought that the actual data should be centralized, whereas 91% thought that a guide to existing data by topic, place and date, along with instructions on how to obtain copies, "should definitely be made available at a central location".

When asked to comment on the dates and types of data that should be covered by the TBNEP Data Management Strategy, survey respondents indicated that the strategy should encompass a wide range of data. The general topic areas suggested on the survey form (Geology, Surface and Ground Waters, Meteorology, Biology, and Human Population) were each selected by 85-96% of the respondents as being areas that should definitely be included in the DMS. Sub-topics that were widely agreed upon as being important included:

**Surface/Ground Waters:** Nutrients (94% of respondents); Water chemistry (90%); Flow rates (88%); Tributary watersheds (85%); Circulation/flushing (85%); Tidal patterns (85%)

**Biology:** Seagrasses (83%); Wetlands (83%)

**Meteorology:** Rainfall (81%)

**Geology:** Bathymetry of Tampa Bay (81%)

Specific topics suggested by respondents for inclusion in the DMS included subtidal bottom types, rate and direction of changes in water table, service area boundaries for sewer systems, radiation levels, and red tide occurrences. With regard to dates, recent data, even from ongoing projects, was considered a high priority for inclusion in the DMS by most (80%) of the respondents. Early (pre-1980) data was considered a low priority by only 20% of respondents, and 40% gave it a high priority rating.

Geographical Information Systems (GIS) programs are increasingly popular computer tools for presenting and analyzing spatial data. The survey results indicate that GIS applications are well-known (84% of respondents had seen a GIS demo) and widely used (55% of respondents worked for an organization that
owned a GIS, and another 14% expected to acquire one in the next 1-2 years). In rating the importance to their work of access to a GIS, 52% of respondents rated access as "extremely important" currently, and 77% expected it to be extremely important in the next five years.

The accessibility of data that are stored in computer files can be strongly influenced by what hardware and software are used. Survey responses indicate that IBM-compatible personal computers are almost universally in use in accessing and manipulating information. Only 11% of respondents indicated that Macintosh computers were in use by their agency. Among the larger and more powerful computers, 50% of respondents reported that their agency uses a mainframe and 41% reported use of a workstation. A large majority (80%) reported use of a modem, indicating that online access to information stored at a remote location is feasible.

Computer files can present a challenge to data transfers due to incompatibility between software used to store and move files. When asked about what software programs were commonly used by their agency for different tasks, respondents reported 21 different communications programs in use, 10 spreadsheets, 16 database programs, 12 statistics programs, and 28 graphics programs. The programs used by the largest proportion of respondents were the Lotus 123 spreadsheet program (69%) and the dBASE database (60%). The only communications package used by more than 10% of all respondents was ProCom Plus (22%). GIS software can be particularly problematic with regard to data transfers due to the size and complexity of the files. Survey respondents reported that at least nine different GIS programs are in use in the Tampa Bay area, with ARC/INFO being the most widely used.

2.2 DATA MANAGEMENT STRATEGY WORKSHOP

The Tampa Bay National Estuary Program (TBNEP), with the assistance of its contractor Coastal Environmental Services, Inc., held a two-day workshop to review and discuss critical issues and components of the Data Management Strategy project. The Data Management Strategy workshop, held at the University of South Florida's St. Petersburg Campus on 3-4 February, 1992, had as its focus the results of the Survey of Data User Needs and Goals and discussions of current data handling/transfer protocols and their inherent advantages and disadvantages. The specific goals and accomplishments of the workshop are summarized below.
2.2.1 Workshop Goals and Charges to the Participants

Two major objectives were stated at the Data Management Strategy workshop:

- to discuss information needs and resources in the Tampa Bay area, and
- to provide a forum for input to the Data Management Strategy.

To these ends, the workshop participants were given three charges:

- identify and assign priorities to user needs in terms of types of information and modes of access to that information;
- identify features of existing database management systems that promote or interfere with effective access to information; and
- propose realistic alternatives to existing systems that will help to meet user needs.

On the first day, the results of the survey of the needs of data users in the Tampa Bay area were presented. The results were discussed to obtain feedback from the workshop participants on whether the responses to the survey accurately and completely identified areas of concern that should be given high priority in developing a data management strategy. Current practices in data transfer between agencies were also discussed, to identify features of data management systems that promote or interfere with effective information exchange. Two presentations by invited speakers focused on how other National Estuary Programs have dealt with the problems of data management.

On the second day, the focus of the discussion was on what steps could and should be taken to improve access to information on natural resources in the Bay. Suggestions were made for modifications or additions to existing data management systems. Demonstrations were then provided for several existing database management systems in the Tampa Bay area.

2.2.2 Summary of Information Needs and Resources in the Tampa Bay Area

Dick Eckenrod, TBNEP Program Director, summarized the goals of the TBNEP, which include developing an accessible information management system. The implementation of the Comprehensive Characterization and Management Plan
(CCMP) for Tampa Bay will be supported by the Data Management Strategy. The Data Management Strategy project has the immediate goal of supporting TBNEP data management needs and the long-term goal of developing a strategy for improving access to information on natural resources and environmental conditions.

On the first day of the workshop, Coastal presented the results of the survey of resource managers in the Tampa Bay area described in the previous section. The presentation of workshop objectives and survey results was followed by discussions focused on each of the three charges given to the workshop participants.

First charge: Identify and assign priorities to user needs

The TBNEP has specific data needs that focus on the development of the CCMP. TBNEP will maintain only limited hardware and software capabilities, yet will need access to statistical analyses, GIS products, sophisticated graphics, and models of natural systems. The United States Environmental Protection Agency (EPA) requires that all newly generated data for the TBNEP must be entered into the EPA's Ocean Data Evaluation System (ODES), and protocols must be developed for dealing with types of data that ODES is not designed to handle. TBNEP also has an important focus on public education, and needs access to particular types of data, clearly presented, to perform this function.

The needs of resource managers in Tampa Bay agencies include access to data from many different subject areas. The point was made that it is encouraging that managers do recognize the range of topics that are relevant to any one resource issue, and that data management to improve information access should not be limited to a few priority items. The need of many users for access to the most recent available information must be balanced against the importance of checking and validating incoming data. A set of levels of validation could be defined that would be used to describe the extent of checking that had been performed to date on a set of data. A standard form included in every data transfer would then allow data users to know exactly how much - or little - validation had been performed on the data.

An issue raised related to validation was how data users can learn about errors and corrections to datasets. A common form of data use is to obtain a copy of a dataset and then use that copy in-house. It is not feasible for data producers to notify everyone who has obtained a copy of the data of each correction made. Most corrections to well-validated data will be minor, but one suggested user need was for a standard format, and possibly a single location (e.g., a computer bulletin board), for reporting such corrections.
The form of data required by data users definitely includes maps as well as numeric and text information. A particular challenge is posed by the increasing use of GIS services to store, manipulate, and present spatial data. As with numeric data, transfers of GIS data are complicated by variations in hardware and software. Transfer protocols for GIS data are relatively new, however. Those that have already been developed need to be disseminated, and others need to be developed. The cost and complexity of GIS systems mean that they are not as ubiquitous as PC-based programs, and there is considerable need for protocols for transfers into and out of, as well as between, GIS programs.

For all types of data, the strain on individual agencies of sharing data with other users could be reduced by developing (a) standard protocols for data transfers, and (b) translation tables for the different codes used by different data producers. By developing such transfer aids TBNEP could serve an important role in integrating existing data management systems.

Different users need different levels of reduction of primary data. Raw data can be reduced to summary tables, reduced to the results of specific analyses, or used as input to models where the main interest is in the data generated by the model (e.g., predicted trends in abundance). Some users can perform all these functions themselves, and only need access to the raw data. Others need only the reduced or transformed data, and could benefit from data management services that included data reduction or analysis.

There was general agreement that a directory to what data exist, and where they are located, would be very useful. It would reduce the search time for locating data, decrease the risk of duplicating effort, and allow the identification of data gaps.

Second charge: Features of existing database management systems

Based on personal experiences with databases in the Tampa Bay area, workshop participants identified features that promoted or interfered with access to information. The points brought up during the discussion were based on the experiences of both data users and data producers in dealing with specific information requests.

One point raised was that people looking for information often need assistance in defining their specific data needs. One-on-one contact with a person who is familiar with the data available from a given agency may be essential to translating a need for information into a specific data request.
Several participants felt that an important feature of a database is the ability to estimate in advance the cost of accessing data. One component of this issue is the quantity of available data that meets a specific request. The sheer volume of data may be too much for the time and monetary budget of a given task. Another component is the actual charges (e.g., for online time, for technical advice, for the hardware and software needed to transfer and then access the data).

Large or complex data management systems may require a major investment in staff training if they are to be used effectively. The initial time investment to learn the system can act as a powerful deterrent to using it. Manuals are rarely well-written and do not necessarily improve access to a system. The point was made that some simple systems (e.g., the Lotus 123 spreadsheet program) may be unsophisticated, but they work. No matter how powerful the capabilities of a given data management system, some users may simply want to get a copy of some of the data in the form of a spreadsheet file.

User-friendly "front ends" to databases may improve access to complex systems, but were considered a hindrance by some data users, if they restricted access to a few menu options. Programs that allow preliminary searching and subsetting of the data before choosing a portion to download can be very useful, however. An important feature of such a system is the ability to name and store subsets as they are created, so that the user can return to previously created subsets.

Data users may need help in understanding what they need in the way of hardware, software, and/or personnel to obtain a copy of the data in a format that they can use. Standardized forms listing transfer format options (e.g., low- or high-density floppy disks, ASCII files or SAS datasets) could serve this function. Inexperienced users could learn from the forms exactly what they need to know before making a transfer request, and experienced users could quickly indicate their needs.

A strong link between a dataset and documentation describing it is essential when data are transferred between agencies. The most basic type of documentation is a data dictionary that defines variables and codes used in storing the data. Also critical is information on sampling design and methods used in collecting the data, and on quality control procedures used in analyzing and processing the data. If this information (or directions on where to locate it) is routinely provided with each data request, the strain on data providers of responding to queries for clarification is reduced. The risk that data will be misused due to a lack of familiarity with how and why they were collected is also reduced.
Third charge: Realistic modifications and additions to existing systems

One consensus of the previous discussions was that the TBNEP could play a valuable integrative role by coordinating and facilitating data transfers between existing agencies. A major issue, however, was whether the TBNEP should also promote the creation of a centralized directory or repository of information on natural resources.

Four major types of information were identified as candidates for placing in a centralized location:

- a subject directory, containing information on where data collected by different agencies is located, organized by subject matter;
- study information, containing information on sampling design, methods, variables collected, and quality control for specific field studies;
- data dictionaries, containing variable definitions, labels, units of measure, and codes for computer files; and
- the actual data, in the form of data sheets, reports, maps, or computer files.

The consensus of the workshop was that the first (a subject directory) should be centralized and computerized, whereas the other three could be located at the individual agencies that produce or control different datasets. The role of the data management strategy could be to develop standard protocols for producing and maintaining the information, and to work on inter-agency agreements on a commitment to sharing information.

A major effort is already underway at the state level to promote the effective sharing of spatial databases among agencies. The Growth Management Data Network Coordinating Council has been developing the Florida Digital Spatial Database System. Steps already completed include creating a formal procedure for developing data standards, the signing of interagency agreements on data sharing, and the development of an online directory to independently operated spatial databases. Possible problems with relying exclusively on this program for TBNEP data management are: (1) The state program is still under development as of June 1992, and may not be fully functional in time for the more short-term needs of the TBNEP and local resource managers. (2) Non-geographically referenced data have not been included in the ongoing pilot study in the state program, but the TBNEP Data Management Strategy should include the full range of natural resources data available. (3) The turnaround time between submittal of information to the state
system and accessibility to users will probably be longer than for a local system. Clearly, however, any components of the Florida Digital Spatial Database System that fulfill the needs of the TBNEP Data Management Strategy will be incorporated as appropriate. High priority will be given to avoiding duplication of effort and to maintaining complete compatibility with the state program.

2.2.3 Presentations and Demonstrations

The workshop benefitted from presentations and demonstrations provided by invited guests who had experience with other National Estuary Programs and with specific databases containing data for the Tampa Bay area. Thomas Gulbransen presented an overview of data and information management systems (DIMS) developed by other National Estuary Programs. Dennis Suszkowski described the approaches and some of the products of data management for the New York-New Jersey Harbor Estuary Program. The second afternoon of the workshop was devoted to demonstrations of specific databases. These were:

- ODES (Environmental Protection Agency),
- STORET (USGS),
- a GIS (Florida DNR),
- COMPAS and ELMER (NOAA and FL DNR), and
- Bibliographic Database (BDB) (SFWMDSWIM).

2.2.4 Conclusions

The input provided by workshop participants confirmed and expanded on the results of the survey of user needs. Inter-agency use of data on natural resources in the Tampa Bay area is already extensive, but many specific areas were identified where access to data should be improved. High-priority user needs include access to:

- diverse subject areas;
- up-to-date data, subject to well-specified validation checks; and
- GIS and other hardware/software capabilities that may not be available within individual agencies.

The workshop consensus was that actual data should remain decentralized, in the control of knowledgeable producers. Access could be improved, however, by:
• creating a centralized subject directory to existing information;
• developing and disseminating standard protocols for data transfers between specific hardware and software configurations;
• creating translation tables for codes used by different data producers; and
• developing inter-agency agreements to standardize data management procedures and to ensure that adequate levels of service to external users are maintained.

In developing the Data Management Strategy full use should be made of procedures already developed by local, state, and federal agencies. This will avoid duplication of effort and promote the development of an effective, cooperative approach to data management among Tampa Bay resource managers.
Table 2.1. Agencies from which responses to the Survey of User Needs and Goals were received.

<table>
<thead>
<tr>
<th>Federal</th>
<th>State</th>
<th>Regional</th>
<th>County</th>
<th>City</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOAA - National Marine Fisheries Service</td>
<td>Florida Department of Environmental Regulation (FDER)</td>
<td>Southwest Florida Water Management District (SWFWMD)</td>
<td>Hillsborough Co. Engineering Services Department</td>
<td>City of Clearwater Pollution Control</td>
<td>Camp Dresser &amp; McKee, Inc.</td>
</tr>
<tr>
<td>Tampa Bay National Estuary Program</td>
<td>Florida Department of Community Affairs</td>
<td>Tampa Bay Regional Planning Council (TBRPC)</td>
<td>Hillsborough Co. Environmental Protection Commission (HC-EPC)</td>
<td>City of Clearwater Public Works Department</td>
<td>Dames and Moore</td>
</tr>
<tr>
<td>U.S. Fish and Wildlife Service (USFS)</td>
<td>Florida Department of Natural Resources (FDNR)</td>
<td>West Coast Regional Water Supply Authority (WCRWSA)</td>
<td>Hillsborough Co. Planning and Development Department</td>
<td>City of Tampa Planning Department</td>
<td>Lewis Environmental</td>
</tr>
<tr>
<td>U.S. Geological Survey (Water Resources Division - Tampa) (USGS)</td>
<td>Florida Game and Freshwater Fish Commission (FGFWFC)</td>
<td></td>
<td>Pinellas Co. Department of Communications and Information Systems</td>
<td>City of Tampa Parks Department</td>
<td>Mote Marine Laboratory</td>
</tr>
<tr>
<td></td>
<td>Florida Sea Grant</td>
<td></td>
<td>Pinellas Co. Department of Environmental Management (PC-DEM)</td>
<td>City of Tampa Sanitary Sewers Department</td>
<td>National Audubon Society of Tampa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pinellas Co. Planning Department</td>
<td>City of St. Petersburg Planning Department</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pinellas Co. Public Works</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pinellas Co. Sewers Department</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1. Inter-agency transfers of data reported in the Survey of Data Users Needs and Goals. Arrows point to the agency that requested the data. (FL NAI = Florida Natural Areas Inventory. For other abbreviations, see Table 2.1.)
CHAPTER 3. DATA MANAGEMENT RECOMMENDATIONS

The goal of the Data Management Strategy (DMS) developed by the Tampa Bay National Estuary Program (TBNEP) is to improve access to information on natural resources and environmental conditions in Tampa Bay. As described in the previous chapter, a survey was conducted to identify the needs of data users in the area. A subsequent DMS workshop provided a forum for resource managers in the area to offer input to the development of the strategy. We have developed a set of recommendations for the DMS that incorporates the results of the survey and workshop.

The proposed strategy is aimed at encouraging data sharing among decentralized databases. In a centralized system, data producers would submit their data, with accompanying documentation, to a central repository. Central repositories typically require that data be submitted in a standard format that is compatible with the hardware and software used by the repository. All data requests would be directed to this repository, allowing for "one-stop shopping" by data users. In a decentralized system, data producers maintain their own databases, typically in a variety of physical locations and using a variety of software and hardware configurations. Centralized systems, while exciting in their concept, too often fail in their implementation. This is due in large part to the fiscal requirements not only for start-up, but also for routine maintenance and updating.

Keeping the data in the control of individual agencies has the advantage of not requiring the development and maintenance of a central repository. It also keeps the datasets closely linked to people who are experts on their contents and structure. It has the disadvantage, however, of making it difficult to locate data for a given information need. The DMS therefore includes the development of a central subject directory, designed for the specific purpose of directing users to data stored in individual agencies. We recommend that this directory be kept narrowly focused on this one purpose, to ensure that it remains easy to access and update.

In addition to the creation of a central subject directory, we propose that as a part of the DMS certain protocols for data management be adopted by all agencies in the Tampa Bay area. These protocols will cover the entire range of records that are needed by data users in order to make efficient, well-informed use of existing data. The records that we recommend are those that should be maintained for any database, whether shared among different agencies or not, and whether stored in a centralized or decentralized system. Since most data producers do maintain these records, the protocols will be adopted from those
already developed by local, state, or federal agencies whenever possible. In some cases, our recommendation will be that a specific form be used by all agencies. In others, we will recommend that a general template be followed and adapted as necessary to the needs of individual agencies and datasets.

In considering the usefulness of existing protocols, we have focused on protocols developed by two data management systems. One is the U.S. EPA’s Ocean Data Evaluation System (ODES). All primary data collected for TBNEP must be submitted to ODES, making compatibility with ODES formats a high priority. The second system is being developed at a state level by the Florida Growth Management Data Network Coordinating Council (FGMDNCC). The FGMDNCC is developing a strategy for sharing data among agencies that includes inter-agency agreements on data sharing and the development of a central directory to spatial data. We evaluate the feasibility of using this system for Tampa Bay data management, and give high priority to ensuring that any products developed by TBNEP are compatible with the state system.

In the following sections we describe in detail our Data Management Strategy recommendations for the central subject directory and the decentralized protocols and records. In the next chapter, we provide an implementation plan for carrying out these recommendations.

3.1 CENTRAL SUBJECT DIRECTORY

As indicated by the survey results and workshop discussions, there is a need for a centrally-located directory to existing information on Tampa Bay. The purpose of such a directory would be to improve access to existing information by allowing users to rapidly identify what data have already been collected, and to find out where further information on the data is located. In the following sections we provide a detailed justification for why a directory should be developed, and specify what the capabilities of the directory should be. Based on these specifications, we evaluate two existing directories. To meet the immediate needs of Tampa Bay data users, we propose that a directory be developed that is based on one and consistent with the other existing directory.

3.1.1 Justification

The workshop and the survey of data user needs identified two major features of data use in the Tampa Bay area that indicate the need for a central directory to existing information. The first feature is the diversity of topics, formats, and dates that were identified as high-priority data needs. The second feature is current patterns of inter-agency access to data: many different datasets
located in a variety of different agencies are already being used. Since access is needed to many different data sources, a central directory would greatly reduce the search time required to locate all the data that are relevant to any one data need.

To serve the needs of data users in the Tampa Bay area, the DMS should cover a diversity of types of information. When asked what types of data should definitely be included in the DMS, most survey respondents selected multiple items from all five of the major categories offered (geology, meteorology, surface and ground waters, biology, and human population: see Appendix A). Workshop participants agreed that resource managers need access to a wide range of subjects, and that data management should not be restricted to a narrow range of subject matter.

The need of users for a variety of data formats also broadens the scope of data management for Tampa Bay. In addition to numeric data, users need access to maps (both as hard copy and as computer files) and text (particularly documentation and reports describing data). In giving examples of specific information needs they had in the last year, survey respondents described questions that required access to maps in 39% of the examples given. The importance of map data is also indicated by the importance of Geographic Information Systems (GIS) in the Tampa Bay area: 78% of survey respondents had worked on a project using GIS in the last year, and 77% indicated that they expected access to a GIS to be "extremely important" to their work in the next five years.

Access to data from recent (1980+) and ongoing projects was considered a high priority by 80% of the survey respondents. However, earlier data were needed by a substantial minority: 40% of the respondents gave a high priority to having access to pre-1980 data. Early data may be needed to assess trends and changes in natural resources, or may be the most recent available information.

Extensive data already exist for the Tampa Bay area for a variety of subject areas and in a variety of formats. Survey respondents indicated that during the last year they had made use of data from a total of 39 different datasets (other than those in their own agency), from more than 10 different agencies. The fact that many datasets exist, in many different locations, poses a considerable challenge to data users trying to locate information. Users that are aware of one source may not realize that additional data are also available. Users trying to locate all available data need to make extensive queries.

The situation described above is clearly one where a central directory to existing information would be useful to data users. A directory would reduce the search time required to locate data. It would increase the use of historical study results, decrease duplication of effort, and allow data gaps to be identified. It
could also reduce the strain on data producers by providing a preliminary screening of queries. Those needing information could first search the directory, and approach agencies only after identifying specific datasets of interest.

3.1.2 Characteristics

The exact form of the central subject directory should be dictated by its purpose. We propose the following statement of purpose and describe the specific characteristics that the directory should have in order to meet that purpose.

**Purpose:** To permit users to locate existing data rapidly, easily, and thoroughly.

The ability of the directory to achieve this purpose will be determined by its characteristics. We have grouped these characteristics into four categories: the directory's contents, its search and subset capabilities, the time investment required to use it, and the facilities required to maintain and access it.

- **Contents.** The directory entries should include descriptions of text, numeric, and map data. Ideally the time frame will include all existing historical data. The subject matter covered should include geology, meteorology, surface and ground waters, biology, and human population impacts.

  Offering a thorough search of existing data is critical to the function of the central directory. The directory does not serve its purpose if a user needs to check many other locations. The contents should therefore include most existing datasets.

- **Search and subset capabilities.** To locate data, users must be able to specify conditions that the data must meet. At a minimum, these should include the specific subject matter, the geographic area, and the time span of the data. Additional search conditions could include items such as sampling or analytical methods used. These should only be included, however, if doing so does not interfere with the other specifications for the directory.

- **User time investment.** The time required to specify each information need, process the request, and deliver the results should be minimal. Very little time should be required for a typical user to learn how to use the directory effectively. Requests should be easy to modify, and it should be possible to combine and compare the results of different searches.
• **Required facilities.** Users should be able to access the directory from their desks, e.g., by modem or in a format (e.g., floppy disk) readable by most personal computers. The agency responsible for maintaining the directory should not have to invest extensively in hardware or software, or in staff training.

Table 3.1 lists the recommended characteristics of the central subject directory. It also provides examples of each characteristic.

### 3.1.3 Recommended Directory Contents

The central subject directory should direct users to existing datasets. A dataset may be a document (e.g., report or map) or a computerized file. We recommend that there be a one-to-one correspondence, whenever possible, between an entry in the directory and a particular data collection effort (e.g., field study, monitoring program, or mapping effort). Individual datasets that have been stored as part of a larger database should appear as separate entries in the directory.

Each entry in the directory should contain:

- a concise description of the contents of the dataset (subject, location, dates), and

- a guide to where to locate the data.

More detailed information (e.g., on what sampling methods or quality control procedures were used during data collection) should be available at the location where the data are stored, but should not be included in the directory. Restricting the information contained in the directory to the minimum needed to help users locate the data serves two main functions. It increases the speed of searches by decreasing the size of the database, and it promotes updating of the directory by making it easy to add information for new datasets.

The proposed contents of the directory are listed in Table 3.2. The subject of the dataset is described in the title, the study objective, the keywords, and a list of parameters collected. Fields describing the location and date(s) of the study are included to allow searches of the directory using these variables. Further information on each dataset in the directory could be found in the citation(s) provided, or by contacting the agency holding the dataset.
3.1.4 Evaluation of Existing Directories

Two existing directories could be used for the Tampa Bay subject directory. One is the Florida Spatial Data Directory (FSDD) under development by the Florida Growth Management Data Network Coordinating Council (FGMDNCC). The other is the Bibliographic Database (BDB) developed by the Southwest Florida Water Management District (SWFWMD). Using an existing system decreases the cost of developing the central directory, if it can be done without compromising the purpose of the central subject directory.

A third possibility for use as a directory would be ODES, the EPA’s Ocean Data Evaluation System. ODES is actually a data repository, and thus requires more effort to submit information to it than would be required by a system that was designed only to act as a directory (the actual data, fully documented and with accompanying quality control reports, must be submitted to ODES). At this time ODES does not contain data from Tampa Bay. Only National Estuary Program contractors collecting primary data will be required to submit data to ODES, so it will probably never contain the full range of subject material recommended for the central subject directory. For these reasons we do not recommend that ODES be considered for use as the central subject directory. ODES would, however, be one of the data storage locations that users of the directory could be referred to.

Florida Spatial Data Directory (FSDD)

The FSDD is being designed to serve as a state-wide central directory to geographically referenced data. It is not yet operational, but considerable background work and development have already been completed. Each entry will describe a geographically referenced dataset. The basic contents of the directory, as described in the FGMDNCC’s Draft Florida Spatial Data Directory Users Manual (dated 31 August 1991) are:

- Subject descriptions (category, subcategory, subject and keywords)
- Area of coverage, by political boundary
- Data source: political entity, date range of data, resolution and scale
- Derived data description: accuracy, format, media
- Data information and transfer contacts: agency, name, address, phone number
- Reference map title, description, and file name
• Description of new data development projects

• File name of data dictionary and quality and accuracy report

The FSDD will be accessible online (with an associated bulletin board), and copies will also be distributed on floppy disks ("Distributed Directory"). Both systems (online and disks) have menu-driven front ends that allow users to search, view, and record information in the directory. Both systems can also be used to add new entries to the FSDD.

Our evaluation of the FSDD for use as the central subject directory for Tampa Bay is organized by the characteristics described above: contents, search and subset capabilities, the time investment required of users, and the required facilities. The evaluation is based on documents provided to Coastal in December, 1991. It is important to note that the FSDD is still under development and the items noted here may change in its final form. We note cases where we are aware of specific plans that will affect its future development.

Contents. At its current stage of development the design of the FSDD does not meet our requirement for diverse subject matter and format. It has been designed for geographically referenced data, and does not, for example, provide for citations of text reports. There is an intention, however, to expand the FSDD in the future to include entries that cover all types of data needed for growth management.

The FSDD does not at this time contain enough entries to allow a thorough search of existing datasets. The actual contents of the Distributed Directory version of the FSDD as of December 1991 were limited to a few sample entries. However, interagency agreements being negotiated for the FSDD may ensure that it will provide thorough coverage of data developed from its implementation date on.

Search and subset capabilities. The online version of the FSDD meets our specifications for allowing searches by subject matter and geographic area. The Distributed Directory menu system allows only a limited type of subject search, by presenting a list of contents sorted by category or subject. (Since the file is stored in dBASE, experienced dBASE users could perform more powerful searches by accessing the file directly.) Neither the online or distributed system menu includes an option for searches by date.

The result of a search consists of the subject, category, subcategory fields and ID number of all entries that meet the search conditions. Individual entries can then be viewed by specifying the ID number (one entry contains nine screens of information). There is no provision for combining two or
more search conditions into one search or for comparing the results of two searches.

**User time investment.** Very little learning is required to search the FSDD using the menu-driven front end. The search options available can be specified quickly and easily, in part because they are limited in number and complexity. Conducting searches with more than one condition could be laborious, requiring that the user manually record and compare the files located by each search.

The time required to process requests is difficult to judge given the lack of entries in the dataset. We would predict that the more successful the project is (i.e. the more dataset information that is added), the more processing time could become a problem. If searches continue to be limited to a few of the fields, searches will be rapid even in a very large dataset. There will be considerable demand for more flexible searches, however. Speed of processing will then be more of a problem with a state-wide directory than with a local one, due to the difference in the sheer volume of entries.

The time investment required to add entries to the directory is a critical feature of the central directory. The FSDD allows direct entry into the computer file, eliminating the need for transfers from paper to computer media. It asks for more extensive information about the dataset, however, than is needed simply to locate potentially interesting data. (There are 110 fields per entry in the Distributed Directory.) As noted above, including detailed information allows for more precise searches, but may discourage data producers from submitting information to the directory.

**Required facilities.** The FSDD is designed to be accessible from users' desks, either by modem or by using the distributed disks. The facilities required to store and maintain the FSDD will include a hardware platform that is accessible by modem and that is large and powerful enough to accommodate the state-wide data that will be incorporated into the FSDD.

**Bibliographic Database (BDB)**

The BDB provides a detailed summary of reports describing data collected in the Tampa Bay area. It was completed in 1989 and has not been updated since. The basic contents of the BDB are detailed descriptions of the contents, and data sources of text reports (Brooks and Doyle, 1989):

- Citation: authors, date, title, source
• Subject descriptions: keywords, statement of purpose

• Location of data collection (Tampa Bay subdivision or specific site)

• Sample methods summary: dates, sampling interval, number of sites, number of observations, parameters recorded

• Status of data: location, storage media (hardware and software), availability of documentation

• Data information contact: name, address, phone

The BDB is available at SWFWMD through the Surface Water Improvement and Management Data Management and Support System (SWIM DMSS), residing on the SWFWMD's IBM 3090. It is also available on distributed floppy disks, as a REF11 file. It will soon be available as a dBASE file from the TBNEP office. Some familiarity with REF11 or dBASE is currently required to use the files distributed on floppy disks.

As with the FSDD, our evaluation of the BDB for use as the central subject directory for Tampa Bay is organized by the categories described above: contents, search and subset capabilities, the time investment required of users, and the required facilities.

Contents. The BDB does meet our requirement for diverse subject matter and format. The scope of work of the BDB project was to "identify, locate, briefly describe, and evaluate the quality of all sources of data pertinent to the ecology of Tampa Bay" (Brooks and Doyle, 1989). The major data types included fell under the headings of biology, chemistry, geology, physics, data bases (including maps and GIS), and meteorology. Human population data was included only for direct impacts to the Bay (e.g., dredge and fill operations, fisheries, point source discharges).

The BDB contains 1,356 entries, and is the most comprehensive source available for historical data on Tampa Bay. It has not, however, been updated since its completion in 1989.

Search and subset capabilities. Searches by subject, geographic area, and date are possible using the BDB. The dBASE version can be searched either by using the menu-driven ASSIST option or by writing dBASE programs. Both approaches allow multiple conditions to be set. Multiple options also exist for printing or storing the outcome of a search.
User time investment. REF11 and dBASE are both powerful database software packages, offering extensive options for searching and reporting, as noted above. The usual tradeoff exists, however, between power and ease of use. For the dBASE version of the BDB, a moderate time investment in learning the menu options of dBASE would allow a user to apply a wide variety of search strategies. A larger time investment, however, would be required to learn and apply the full range of options available through the programming language.

The time required to process requests depends on the computer the BDB is stored on, the complexity of the data request, and the software version used (REF11 or dBASE). A simple search of the dBASE version of the BDB (searching for a match between one search string and the contents of one field) will require a few seconds or minutes if the entire database can be searched from the ASSIST menu option. Up to 20 minutes could be required for the same search if the dBASE files must be kept compacted due to space problems (the entire BDB in dBASE requires more than 5 megabytes of storage).

Search time would be decreased by limiting the fields in the database to the few required to direct users to the data. Limiting the number of fields would also limit the time investment required to add entries to the BDB.

Required facilities. To use the BDB at their desks, users currently need an IBM-compatible personal computer and a suitable software package: REF11, dBASE, or one of the many packages that can read dBASE files directly. The BDB is currently stored on a mainframe computer at SWFWMD, but the REF11 and dBASE versions can reside on a personal computer.

3.1.5 Recommended Structure of the Central Subject Directory

We propose that the central subject directory for Tampa Bay should use the existing BDB as a base of information, but should be designed to allow easy transfer of information from or into the FSDD. Extensive records for historical data can quickly and easily be added to the central subject directory by copying information from the BDB. Compatibility with the FSDD can be maintained by ensuring that matching variables have the same format and length limits. Such compatibility is highly desirable to ensure that information submitted to one directory can be made available in the other, without requiring double entry by the agency submitting the information.

Our recommendations for the central subject directory contents include some variables that are not in the BDB. Data for these variables will have to be entered
during the initialization of the directory. The proposed central subject directory also includes some variables that are not currently part of the FSDD. The FSDD would have to be modified before all the information in the central subject directory could be transferred to the FSDD. Table 3.3 summarizes the match between variables in central subject directory, the BDB, and the FSDD.

The recommended characteristics of the central subject directory given in Table 3.1 are achieved in part by limiting the contents of the directory to only those variables that are essential to locating data. The rest of the recommendations will be met during the implementation phase, e.g., by developing programs that will help users to search the database efficiently. The steps required to implement our recommendations regarding the central subject directory are described in detail in Chapter 4. Our implementation recommendations will include the development of programs for transferring information from the central subject directory into the FSDD.

3.2 DECENTRALIZED COMPONENTS: DATA MANAGEMENT PROTOCOLS AND RECORDS

In accordance with the responses to the survey of data user needs and the workshop discussions, we recommend that datasets in the Tampa Bay area remain decentralized, in the control of individual agencies. The Data Management Strategy to improve access to these data involves obtaining commitments from agencies to agree to follow standard procedures in documenting how data are collected, processed, and stored within the agency, as well as developing standard procedures for transfers between agencies.

Our overall goal in making these recommendations is to improve access to existing information. For the data user, this includes ensuring that data being transferred are well documented. For the data producer, it includes reducing the effort required to meet data requests by standardizing the process.

Datasets should be well-documented within agencies. We present guidelines for what the components of this documentation should include, and how detailed they should be. We further recommend that agencies agree to follow these general guidelines, and in some cases use standard forms. Such agreements would ensure that, no matter what the source of the data, a user would know what information is available to describe the data, and what information would be provided along with a data transfer.

Agencies that are responsible for major datasets will typically have developed procedures for allowing outside users to request and receive data. Our proposal is that these procedures be standardized across agencies. We provide a
justification for each recommendation that spells out the benefits to both data users and producers of standardizing data transfer procedures.

This section is organized chronologically, following data from the collection process, through in-house processing, and finally to transfers to other data users. Records should be kept throughout this process that document what has been done. It is important to note that we expect that the records maintained by different agencies will vary in their level of detail. For example, one agency might record each correction made to a dataset, while another agency might simply keep a record of the date when the most recent correction was made. In the following section we describe what a complete set of records describing a dataset would consist of. For each type of record, we describe what it should contain and propose minimum criteria that can be used to judge whether a given form or protocol is adequate to achieve its purpose.

We recommend that records describing the datasets be stored in the same location as the decentralized databases, i.e. by the agency that creates them. Descriptive documentation would then be available at the same place as the actual data. This is the current pattern of data transfers between Tampa Bay agencies. The changes that we are recommending are that:

- agencies make a commitment to maintaining the minimum records needed for effective data sharing,
- these records be standardized among agencies, wherever feasible, and
- agencies agree to provide these records with each data transfer, as a standard practice.

These changes should make the current system work more efficiently. Any extra effort required from data producers should be minimal, and should benefit in-house users as well as those requesting data from other agencies.

During the implementation phase, existing protocols for maintaining database records will be compiled and evaluated for use in the DMS, with new protocols being developed only if necessary. The evaluation of existing protocols will be based on the criteria given in the following sections.

### 3.2.1 Data Collection

Complete and careful documentation of the data collection process is critical to the success of a data management system. Checks on the accuracy of data at the time of data collection are essential. Early identification of errors benefits the
data producer by preventing wasted time and reducing the risk of drawing erroneous conclusions. Documenting what quality control procedures were followed is particularly important to secondary users of the data. No matter how accurate historical data are, they may be of no practical value if there is no record of how they were collected.

**Review of data sheets**

The majority of all data errors occur when transcribing data to and from data sheets. We recommend that all data sheets be reviewed, signed and dated by the crew chief or chief scientist before the data are transferred to electronic format. Data sheets should be checked for the following features:

- readability (are the recorded values legible?)
- completeness (are there any missing values?)
- plausibility (are the results within reasonable limits?)

**Quality assurance and quality control**

Quality assurance (QA) steps are taken to ensure that a study is adequately implemented to provide data of high quality. Quality control (QC) samples are used during a study to check the immediate instrument calibration or response. The measurement obtained from a quality control sample is expected to fall within specific acceptance criteria or control limits (EPA 1978).

Guidelines for QA/QC procedures are available from the U.S. Environmental Protection Agency (EPA) for most fields of study (e.g., EPA 1978 and EPA 1988). Whatever QA/QC procedures are followed, we recommend that QA/QC activities be documented for all sampling programs. Checks on data quality must not only be performed, they must also be reported so that secondary users of the data can confirm that they were conducted.

**Storing of data sheets**

Each agency should have a system for filing (archiving) data sheets for data that have been transferred to electronic format. Archiving raw data provides a back-up if the electronic format becomes damaged or misplaced. In addition, if the validity of the data is questioned, data sheets may be consulted or reviewed to identify technical difficulties or measurement error at the time of collection.
3.2.2 **Computer File Construction and Maintenance**

Creating a digitized computer file typically includes the following steps:

- data entry and verification: digitizing the data and checking the accuracy of the data entry, and
- describing the dataset: creating records of how the data are organized and what the numbers mean.

After this process is completed the maintenance of the dataset may involve further changes. These changes could include corrections based on questions raised during use of the data, or the creation and storage of subsets or summaries of the data. Records must be kept of what happens at each of these steps, if the data are to be useful in future analyses.

**Data entry and verification**

Data entry and verification is the process of changing data from one format to another and verifying that the new version exactly matches the original. For most numeric and text data, data entry is the keypunching of data and data verification is the process of eliminating keypunch errors. In addition to checking the accuracy with which the original data were copied, verification may also include checks on the reasonableness of the data.

The most accurate method of data verification for keypunched data is double entry followed by electronic logical consistency and completeness checks. Double entry is the process in which the same data are keyed into two separate data files by two different key operators. The files are then compared electronically. Data can also be checked visually against the original. With either visual or electronic checks, subsets can be verified rather than the entire dataset. Verification of the entire dataset is preferable, but methods exist for verifying selected subsets in such a way that the maximum error rate in the final dataset can be estimated and kept below a fixed limit (e.g., Duncan 1974). After data verification, logical consistency checks should be performed. These include range checks on continuous variables, code checks on categorical variables, and cartographic and topological checks on map data. Completeness checks should also be performed, which include confirming that all variables, and all values within variables, have been entered.

Quality control during data entry and verification is of particular concern with GIS data. The processing involved in entering map data into a GIS includes many potential sources for errors. We strongly recommend that data producers report
the lineage, positional accuracy, attribute accuracy, logical consistency, and completeness of their GIS data. A template for such a report has been developed by the FGMDNCC, and we recommend its adoption as the standard format to be used by Tampa Bay agencies.

Verifying data and resolving errors before further processing or before releasing data to other users saves time and money for all parties involved. Data producers minimize the number of corrections that must be made by catching errors early. Data users are less likely to have to repeat analyses and will have more confidence in their results if they know that the data were properly entered and verified.

We recommend that the data entry and verification protocol used by individual agencies should, at a minimum, include the following:

- verification of entered data, e.g., double entry or visual checks,
- logical consistency and completeness checks, and
- maintenance of a record of what verification has been performed.

Dataset description (documentation)

Documentation describing individual datasets is an essential resource for database users, whether data are archived within an agency or transferred between agencies. These records should be developed and maintained by data producers, and stored at the same location as the actual data. The minimum documentation needed to allow the efficient transfer of data between agencies includes:

- an overview of the study for which the data were collected,
- a description of the sampling design and the collection methods,
- a description of the dataset contents, and
- a report on the level of data quality review.

The rationale and criteria for each of these types of documentation are described in the following sections. Examples of protocols for each type of documentation are presented in Appendix C.

**Study overview documentation.** An overview of the study for which the data were collected should be available to the users of the data. A study overview
is an efficient method of describing the objectives and methods of the study. The establishment of a standard format for reporting a study overview would increase the ease with which data sharing could be accomplished.

The criteria for a evaluating a useful study overview are the typical criteria used to evaluate an abstract of a report or technical paper. The study overview should summarize:

- Who collected the data
  - the official title of the study
  - the name of the agency responsible for data collection

- What data were collected
  - a list of the parameters measured

- Why the data were collected
  - the objective(s) of the study

- Where the data were collected
  - the geographic coordinates bounding the study area

- When the data were collected
  - the temporal range over which data were collected
  - the sampling interval

- How the data were collected
  - a brief description of methods used to collect the data

When appropriate, a brief summary of study results could also be included. Most of this information will be included in the central subject directory described earlier (see Table 3.2). We therefore recommend that the central subject directory format be used by all agencies in preparing study overviews for each of their datasets. Recording the information in this format will have the added benefit of allowing the dataset to easily be added to the central directory.

**Methods and sampling design documentation.** A description of the methods and sampling design with which the data were collected should be available to the users of the data. Knowledge of the methods and sampling design is needed to make valid inferences from the data, and to evaluate the level of confidence in those inferences. This documentation is especially critical for data sets that are being archived or transferred to other agencies.

The criteria used to evaluate methods and sampling design documentation are the typical criteria used to evaluate the methods section of a report or technical
paper. The methods documentation should include a sufficient level of information so that the experiment or measurement could be repeated

Sampling methodologies will differ between and within agencies, and according to type of study. Each study will require data sheets and variables specific to the purpose of the study. To meet the needs of inter-agency dataset sharing, we recommend that, at a minimum, agencies record the following information for every study:

- **Sampling design**
  - whether sample sites (e.g., trawl locations) were chosen randomly or systematically
  - whether sample units (e.g., individual fish measured for length) were chosen randomly or systematically

- **Temporal and spatial measurements**
  - x-y coordinates for each sampling site
  - water depth (if applicable)
  - date
  - time of day

- **In-situ (field) measurements**
  - type of equipment or instrument used for collection
  - instrument model number (if applicable)
  - parameters measured
  - mesh or gear size (if applicable)
  - units of measurement

- **Analytical or laboratory measurements**
  - analytical method number (e.g., EPA or APHA)
  - type of instrument used
  - instrument model number (if applicable)
  - detection limits
  - units of measurement

The diversity in methods used in data collection preclude the adoption of a particular standard format for recording methods. However, widely accepted standard method numbers provide a convenient method of documenting methods for many measurements (i.e. U.S. EPA and American Public Health Association standard method numbers).

We recommend that a methods report be written for each study that includes at least the above information. The level of detail should be sufficient to allow a competent researcher to repeat the study.
Dataset contents documentation. Clearly, a description of the contents of each data set should be available to the users of the data. This documentation is often referred to as a "Data Dictionary" or "Annotated Variable List", and it provides a guide to the interpretation and technical use of the data. Dataset contents documentation is essential for all phases of data management including using the data within an agency, archiving data, and transferring data between agencies. The ability to use a dataset is always dependent on the documentation of its contents.

A useful dataset contents description has two components. The first component describes the overall dataset, and the second component describes each individual variable in that dataset. The documentation for the overall dataset should specify:

- The technical information on hardware and software that must be known in order to read the data. This includes:
  - a media description (e.g., DSDD, 5.25" floppy disk),
  - an operating system description (e.g., MS DOS, version 3.2), and
  - an encoding format description (e.g., ASCII file delimited by commas).

The media and operating system descriptions are straightforward. Data encoding formats, however, are diverse and are often proprietary in nature, and an efficient standard form of documenting formats is difficult to define. However, the encoding format of most data can be identified by the specification of the software package used to store it (e.g., this is a PC SAS version 6.04 dataset).

The documentation for each individual variable in the dataset should specify:

- The technical information that must be known in order to read the data (e.g., variable names, column positions, whether the data are character or numeric).
- The technical information that must be known in order to interpret the data (e.g., variable definitions, coded value definitions, units of measurement, number of significant digits).

We recommend that dataset contents be completely described, using a standard format. The template for recording this information that was developed by the FGMDNCC is presented in Appendix C.
Dataset quality documentation. A description of the steps taken to ensure the quality of each data set should be available to the users of the data. Dataset quality can be defined as a measure of the amount of critical review a dataset has undergone. Documentation of dataset quality is especially critical when data are being archived or transferred between agencies.

Dataset quality documentation should communicate to a data user what critical review steps have been applied to a dataset. Our recommended strategy for communicating this information is to classify review steps into discrete levels, and to label each data set with an identifier based on these discrete levels. The proposed quality review code identifies the level of scrutiny that has been applied to a particular data set, and it includes the following values:

- **Level 0 = No Scrutiny**
  This level describes raw data as sampled in the field or experimental setting (e.g., data sheets, aerial photos, measuring instrument printouts and datasets)

- **Level 1 = Reviewed**
  This level describes Level 0 data that have been visually inspected for readability and completeness (e.g., are handwritten data legible, or are computer files complete?)

- **Level 2 = Verified**
  This level describes Level 1 data that have been scrutinized for data encryption errors and data measurement errors (e.g., checked for keypunch errors by double entry, checked for instrument failure and plausibility of values by range checks, checked to see if all samples are in data)

- **Level 3 = Validated**
  This level describes Level 2 data that have been scrutinized by a person proficient in the field of interest (e.g., checked for unusual values, and evaluated for systematic biases).

Detailed standards for each of the above levels, and forms used to report quality review steps and their outcome, will be developed during the implementation phase of the DMS (Chapter 4). We recommend that quality review forms be completed for each dataset.
Maintenance

Maintaining data files includes keeping a record of transactions to the data file as well as protecting the data file from corruption. Keeping a log or audit-trail of transactions provides the data producer with a method of recreating data if necessary. It allows a data user to determine the original source of data, and to find out what changes have been made to it since the data were collected. The transactions to be recorded include any that involve reviewing, changing, or transferring data in the dataset. At a minimum, the information recorded would indicate that an (unspecified) change or correction had been made on a given date. The exact nature of the change could be noted if feasible.

Maintenance of a dataset is a dynamic rather than static process, since changes can be made at any time. One of the concerns raised during the DMS workshop was how data users could be made aware of corrections to a dataset that occur after the user has obtained a copy of the data. We recommend that all agencies in the Tampa Bay area agree to maintain data tracking forms that record corrections made to the data. Users will then know that this information exists, and can obtain copies of the form to identify changes made to the dataset. The log of transactions should also include a record of data transfers. If a major error is discovered in the dataset, each person that had received a copy of the data could be notified, preventing further misuse of the data.

We recommend that agencies maintain a transaction log for each dataset. A transaction log should be as simple as possible while containing adequate detail to describe each transaction affecting the data accurately. An example of a log sheet is provided in Appendix C.

3.2.3 Data Transfers

A data transfer must occur in order to share data between agencies or individuals. Data sharing is critical to effective resource management in the Tampa Bay area. If data are not shared, unnecessary and wasteful duplication of effort may result. When data are shared, limited research resources can be focused on data gaps where they will be most effective.

We define two major types of data sharing. One is a basic transfer of data from an existing dataset, with no changes or modifications. The other is a transfer of data that have first been manipulated to meet a specific request. An example of the first would be copying a numeric file to a floppy disk and mailing the disk. An example of the second would be taking the same numeric file, running a statistical analysis, and mailing the results of the analysis. Each type of transfer will be considered in the next two sections.
Basic Transfer

Even basic data transfers can pose major practical problems. The requester must be able to describe clearly what data are needed. The data must be transferred in a format that the requester can understand and use. Background information that will allow the requester to evaluate the data received and then use them appropriately must accompany any data transfer. Each of these steps must be accomplished with as little cost as possible to both parties.

Describing the request. Data users must first identify existing datasets that contain information that they need. This is the purpose of the central subject directory described in the previous chapter. As was mentioned at the DMS workshop, users often need assistance in translating their information needs into specific data requests. This requires one-on-one contact with someone who is familiar with the dataset(s) that the user is interested in. An important component of the proposed central subject directory is the name, address, and phone number of a person who will act as a contact for each dataset.

A data request must be specified in terms of:

- the dataset containing the information and
- the subset(s) of the dataset needed: which variables, and which records.

Subsetting may be requested if the data user only needs a portion of the dataset, consisting of particular items (variables), or data from a particular location or range of dates.

Transfer format. The data may be available only in the format in which they are stored. Alternatively, the data producer may be willing and able to convert the data to a different format. In either case, the mode of transfer must be specified in terms of:

- the media (e.g., tape, disk, or hard copy),
- the file format of computer files (e.g., ASCII, dBASE, or ARC/INFO files), and
- the version of the software and operating system used in creating the file.

Dataset description. In order for a data file to be useful to a data requester, the data provider must supply information about the data file being
transmitted. The information includes the study overview, methods and sampling design report, dataset contents documentation, and dataset quality documentation described earlier. Supplying this information at the time of transmittal provides the requester with the tools necessary to use the data appropriately and reduces the burden on the provider of having to address questions that the documentation can answer.

Accurately defining the request at the beginning of the process is essential: if a data transfer takes place and the data are not usable by the requester, time and money have been wasted by both parties. A standard data request form is useful for ensuring that the requester provides all the necessary information. A form can also alert the requester to what technical details he or she must specify, sparing the data producer from repeatedly answering the same questions or providing the same verbal instructions. A data request form should allow data users to specify clearly each of the items described above under "Describing the request" and "Transfer format".

A critical part of the request process is providing the requester with an estimate of the cost and turnaround time for each request. Both will depend on the specifics of the request given above. A standard form prepared by the data producer, however, can provide data users with a preliminary estimate of the cost and turnaround time. This information should be part of a form prepared by data producers that describes their policies and procedures regarding data transfers.

There are numerous software and hardware options for storing and manipulating data. Moving information from one location to another is often theoretically possible, but practically difficult and time-consuming. From the point of view of the data user, data stored in a software package to which he or she does not have access may be completely unusable. From the point of view of the data producer, attempting to provide data in whatever form it may be requested would require an excessive amount of time and computer expertise.

We recommend that a set of software-specific translation procedures be developed. Many such procedures can simply direct users to the translation options provided by most software packages. Others will require complex instructions. In these cases, once a translation is worked out for one data transfer, an easy method of making it available to other users should be provided.

Translation procedures (moving data from one environment to another) are potentially problematical with GIS data. Map data in a GIS are typically associated with 'attribute' files that describe different characteristics of the points, lines, or polygons in the map. Some software packages can easily transfer the entire GIS database, whereas others cannot keep the attribute files linked to the map data. In addition, some software manufacturers have not been willing to cooperate in
developing programs that will transfer their files to another company’s GIS. We recommend that a summary of existing GIS transfer protocols be developed and made available to local agencies. As additional protocols are developed by individuals, agencies, or software vendors, these can be added to the summary.

Most datasets use codes to store data. Codes uniquely describe information that is too long or complex to include in its entirety (e.g., species names or gear type and dimensions). Agencies typically develop standard codes for in-house use, but different codes will be used by different agencies. Codes also tend to change over time, even within agencies. Differences in codes can impede data sharing. Even when code definitions are provided with each transfer, translating from one set of codes to another can be time-consuming and error-prone.

The problems associated with variations in codes were raised at the DMS workshop. One solution is to promote the use of standard codes by all agencies. Another is to develop translation tables for existing codes, that can be used repeatedly by different agencies. We recommend that the DMS include an assessment of the size of the problem, and address the problem as part of the implementation phase. Data users and producers should be queried to identify major datasets with matching variables but different coding schemes. If major coding problems are identified, computer files of translation tables should be developed and provided to those agencies whose codes have been used in developing the tables.

Data reduction transfers

Data users are not necessarily interested in, or able to make use of, the raw data stored in a dataset. They may need an altered form of the data: a summary statistic, a table, a specially prepared map, or the predictions of a model that uses the data as input. Data users may not have the facilities or expertise to prepare such summaries of the data. Alternatively, they may be able to, but only at a much larger investment of time and money than would be needed if the agency holding the data prepared the summary.

Public education provides an important example of a need for data reduction services. Educators and people acting as public information contacts may not have the technical expertise or the facilities to prepare graphics or extract summary information. If they were available, however, such summaries would be extremely useful in performing their jobs. As another example, students who are interested in the natural resources of the Bay will also usually need summary rather than raw data, without having the resources to prepare the summaries themselves.
We recommend that agencies managing data in the Tampa Bay area develop a policy for offering data reduction services. Alternative levels of service could consist of:

- no data reduction services provided
- services provided only after negotiating a contract with the individual or agency requesting the service
- a standard set of services available, at specified cost, to anyone requesting data.

The latter option could be offered only for selected datasets within an agency. To reduce the effort required to provide the service, brief summaries of what services are available should be prepared and made available to data users. Such a summary would include:

- the dataset for which manipulation and reduction services are available
- a brief list of the types of services available
- a list of the types of deliverable products available (e.g., aerial photographs, maps, data tables)
- a description of charges for the services

It would be important to include, with each product delivered to a client, sources for background information on the data being delivered, including:

- citation(s) for reports documenting the dataset, and
- citations for technical reports and literature describing how the data were manipulated.

3.2.4 Summary

The protocols described in this chapter will be an essential part of the DMS for improving access to information by promoting the sharing of existing data between agencies. The recommendations given above are based on basic tenets of good data management. Most agencies follow protocols similar to the ones given here as part of their routine processing of data. Agencies should not only follow these basic protocols, but also keep a detailed record of their actions. The success
of the DMS will depend on a strong and continuing commitment from data producing agencies to follow good data management practices.

For each protocol, we have explained why these records are essential to the effective use of data both within and outside of the agency maintaining the records. To summarize, the recommendations made in this chapter are that agencies should keep the following records for each dataset:

- a study overview
- a description of study methods and sampling design
- a description of dataset contents
- a record of dataset quality review, including
data sheet review and archiving procedures,
data entry procedures for digital data, and
verification and validation procedures; and
- a transaction log.

We also recommend that agencies have available:

- a summary of transfer policies and procedures,
- a data request form,
- software-specific translation protocols,
- code translation tables, and
- a summary of data reduction services offered.

3.3 TBNEP PROGRAMMATIC DATA MANAGEMENT NEEDS

The data management needs of the TBNEP include many of the same needs described above for all data users. In addition to these, there are some specific data management needs which must be considered in developing the DMS. These can be grouped as follows:

- short-term data needs
- data reduction services
• public education needs

• ODES data submission guidelines.

3.3.1 Short-term Data Needs

The focus of the TBNEP is the development of the Comprehensive Conservation and Management Plan (CCMP). The CCMP development will require data which will vary widely in type, sources, and formats. Successful development will therefore depend on timely and efficient management of data. One necessary component of TBNEP data management will be a library system for tracking data and reports generated or received by the TBNEP office. The library system should be designed at a minimum to allow entries to be located by the date, source, and topic of the item sent or received.

One particular source of data will include results from contractors who are either generating primary data, which must be submitted to ODES (see below), or who are providing data analysis results (e.g., maps). To ensure that data submitted by TBNEP contractors can be used with the minimum of effort and error, we recommend that criteria for data submittal be developed for those data not submitted to ODES. These criteria should include, at a minimum:

• data medium - compatible with hardware available to TBNEP,

• data format - compatible with hardware and software available to TBNEP,

• data documentation - as described in detail earlier in this chapter, and

• hard copy format, especially for map outputs.

TBNEP, during its tenure, will also be providing data to a number of users. To facilitate these data requests and to maintain control of intermediate and final data products, we recommend that standard protocols be developed. These protocols would include:

• a log of data requests and data transfers - including requester, files or products transferred, dates of request and transfer; and

• a data transmittal form - a check list to ensure transfer of the following: data medium, data format, and data documentation.
We have prepared a preliminary set of forms that can be provided to TBNEP contractors to ensure adequate, standardized documentation of data submitted to TBNEP (Appendix D). These forms should be evaluated and modified as necessary during the implementation phase of the Data Management Strategy project.

3.3.2 Data Reduction Services

TBNEP will require a wide variety of data reduction services to support the CCMP development. Most of these services will be provided by contractors on a project-by-project basis. However, other services may be needed, especially for program presentation materials or CCMP-related products. These services will likely be out-of-scope for particular contracts, or needed on a more ad hoc basis. These materials will include data summaries, in the form of tables or graphs, and GIS or other map products. They may also include statistical analyses or model outputs which TBNEP cannot develop given their hardware and software capabilities. We recommend that TBNEP consider how these services could be provided in a timely and cost-effective manner.

3.3.3 Public Education Needs

Public education is an important component of the TBNEP. It will not need the technical data management services being planned for data users in general. There are at least two areas, however, where the DMS could include components that will assist with public education.

One area is data reduction services, similar to those described above for the program as a whole. A strategy for providing TBNEP staff with presentation materials and data summaries would be particularly useful to the public education effort.

The second area is related to the question of what data have been and are being collected for Tampa Bay. We have recommended the development of a computerized central subject directory for technical data users. A more descriptive summary of the research and data management focus of agencies in the Tampa Bay area could be useful to the public education component of the TBNEP. The summary could serve two functions. It could direct the public to information sources without requiring any technical expertise or detailed knowledge of existing datasets. Equally important, such a summary could be used to educate the public regarding ongoing efforts to understand and protect natural resources in Tampa Bay. A summary suitable for these purposes will be developed as part of the Base Program Analysis to be conducted for TBNEP in funding year 1992.
3.3.4 ODES Data Submission Guidelines

The U.S. EPA requires that primary data collected for the TBNEP be submitted to its Ocean Data Evaluation System (ODES). Extensive documentation and support services exist to help investigators submit their data to ODES. We recommend that TBNEP contractors be given a brief package of information that summarizes their ODES data submission responsibilities and directs them to further technical support services. A draft of this package is presented in Appendix E.

3.4 SUMMARY

Given guidance by TBNEP and the community of data providers and users in the Tampa Bay area, we have prepared recommendations for a Data Management Strategy that is designed to improve access to information while taking advantage of existing database management protocols. Our strategy focuses on integrating and coordinating existing systems, with the ultimate goal of improving data transfers for both the providers and users of these data. The basic components of the proposed DMS are:

- a central subject directory,
- a set of standard protocols and report forms to be used by agencies with dataset holdings, and
- short-term data management services for TBNEP.

The central subject directory is needed to direct data users to existing data, which will remain decentralized. It will make use of information already entered into an existing directory (the Bibliographic Database), and will be designed to be compatible with the Florida Spatial Data Directory currently being developed by the state.

Datasets will remain decentralized, in the control of individual agencies. We recommend that these agencies follow standard protocols for data management. In particular, we propose that the agencies should agree to maintain at least those records that are needed for effective secondary use of data. We describe these records in detail, and provide specific statements of the advantages to data producers of maintaining and standardizing these records.

The TBNEP has data management needs that have a shorter time frame or, in a few cases, that are different from those of other data users. These needs include tracking data requests and transfers among TBNEP contractors, obtaining data reduction services, supporting public education data needs, and providing
TBNEP contractors with guidelines for ODES data submission. We have prepared a package notifying TBNEP contractors of their ODES data submission responsibilities, and have made recommendations for meeting the other TBNEP needs as part of the Data Management Strategy.
Table 3.1. Recommended characteristics of the central subject directory

| CONTENTS |
|----------|----------|
| **Characteristic** | **Examples:** |
| Diverse subject matter, format of data | Include Descriptions of: Reports Numeric computer files GIS computer files Maps For: Geology, meteorology, surface and ground waters, biology, and human populations |
| Thorough search of existing data | Initially, add entries for as much existing data as possible For maintenance, make new data entry quick and easy to encourage updating |

<table>
<thead>
<tr>
<th>SEARCH AND SUBSET CAPABILITIES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristic:</strong></td>
</tr>
<tr>
<td>Search conditions include:</td>
</tr>
<tr>
<td>• Subject matter</td>
</tr>
<tr>
<td>• Geographic area</td>
</tr>
<tr>
<td>• Date(s)</td>
</tr>
<tr>
<td>Subset capabilities include:</td>
</tr>
<tr>
<td>• Combining search conditions</td>
</tr>
<tr>
<td>• Comparing searches</td>
</tr>
</tbody>
</table>
Table 3.1 (continued).

### USER TIME INVESTMENT

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Example:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal &quot;learning curve&quot;</td>
<td>Keep it simple, use commonly used software, provide a menu-driven front-end</td>
</tr>
</tbody>
</table>

**Allow rapid**
- Specification of a request
- Processing of a request
- Entry of new data

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Provide a checklist of options</td>
<td></td>
</tr>
<tr>
<td>Minimize the size and complexity of the directory</td>
<td></td>
</tr>
<tr>
<td>Include in the directory only the minimum information needed to locate data</td>
<td></td>
</tr>
</tbody>
</table>

### REQUIRED FACILITIES

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Example:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessible from a user's own desk</td>
<td>Online access by modem, and/or</td>
</tr>
<tr>
<td></td>
<td>Floppy disks, tape cartridges, etc.</td>
</tr>
</tbody>
</table>

**Minimal**
- hardware
- software
- staff training required for storage and maintenance

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimize file size, e.g., to allow storage of the directory on a personal computer</td>
<td></td>
</tr>
<tr>
<td>Use commonly used, widely distributed software.</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.2. Recommended contents of the central subject directory

<table>
<thead>
<tr>
<th>Dataset Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contents:</strong> Brief statement of objective</td>
</tr>
<tr>
<td>Keywords</td>
</tr>
<tr>
<td>Parameters collected</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location: Geographic area</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Date(s) of study (source data)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Citation (report describing the study)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Dataset Location:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agency</td>
</tr>
<tr>
<td>Name to contact</td>
</tr>
<tr>
<td>Address</td>
</tr>
<tr>
<td>Phone number</td>
</tr>
</tbody>
</table>
Table 3.3. Match in variables between the Bibliographic Database (BDB), the central subject directory (CSD), and the Florida Spatial Data Directory (FSDD).

<table>
<thead>
<tr>
<th>Variables</th>
<th>CSD</th>
<th>BDB</th>
<th>FSDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset Title</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Study Objective(s)</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Keywords</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>List of Parameters</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
</tr>
<tr>
<td>Location</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Date(s)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Citation (report describing the study)</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Data Location</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

* A data dictionary for the dataset, including variable names and descriptions, may be available as a separate file.
CHAPTER 4 - IMPLEMENTATION PLAN

The purpose of the implementation plan for the Tampa Bay National Estuary Program's (TBNEP) Data Management Strategy (DMS) is to provide guidelines for carrying out the recommendations made in Chapter 3. During the next three years the TBNEP will be involved in initiating and maintaining the DMS. It is important to note, however, that continued data management after the tenure of the TBNEP is a major concern that will influence the final design and implementation of the DMS.

We describe in this chapter how TBNEP and its contractors can initiate the recommended DMS and encourage its use among local agencies. We also specify the activities that will be critical to the long-term continuation of the DMS. At the end of the first year of the implementation phase of the DMS project, data management products and practices that allow rapid, easy access to natural resources information should be available for use by agencies in the Tampa Bay area.

4.1 SHORT-TERM (FIRST YEAR) IMPLEMENTATION

The main components of the DMS to be initiated in the first year of implementation are a (1) a Central Subject Directory that will direct data users to existing data sets, (2) standard data management forms and protocols that will be adopted by Tampa Bay area agencies, and (3) data management support for the TBNEP office.

4.1.1 Central Subject Directory (CSD)

To initialize the CSD, we recommend that the TBNEP contractor for the implementation phase carry out the following steps (Table 4.1).

1. Conduct final preparations for creating the CSD.

   • In consultation with the TBNEP office, sketch out a long-range maintenance plan for the CSD. Topics to consider include the location at which the CSD will be stored and maintained, the activities needed to ensure that the CSD remains a useful tool for data users, and funding sources for maintenance and updating of the CSD. The purpose of preparing such a plan early in the implementation phase is to ensure that long-term issues are considered during the initialization of the CSD and to
identify areas that will need further planning or research and should therefore be begun early in the program.

- Finalize the structure of the CSD database. This requires constructing a dBASE file with specific variables and variable attributes. Before finalizing the structure, the current status of the Florida State Data Directory (FSDD) should be checked, to ensure that the CSD design is as compatible as possible with the FSDD design.

- In consultation with TBNEP staff, determine priorities for updating the information from the SWFWMD Bibliographic Database (BDB) that will be transferred to the CSD. Datasets should be identified that have been developed since 1989 and that should definitely be included in the first version of the CSD.

2. Create a working copy of the CSD that contains information, transferred from the BDB, on historical data.

- Transfer entries from the BDB to the CSD. Depending on the final design of the CSD, this may require combining information from multiple fields into one field as well as simple field-to-field transfers. Quality control checks should be conducted to ensure that the transfer is complete and accurate.

- Create programs that search the CSD by subject matter, geographic area, and date. The programs should allow users to specify multiple search conditions and to easily store and retrieve their search results. The programs should be made available through a user-friendly shell that provides users of the directory with menu-driven access to the search programs.

- In preparation for updating of the CSD, obtain documentation describing the high-priority recent datasets identified in step (1) above. The documentation should be detailed enough to provide all the information needed to add an entry to the CSD database.

3. Develop programs for adding entries to the CSD and update the directory with recent, high-priority data sets.

- Create programs that can be used to add entries to the CSD. The programs should include checks on entered values wherever feasible. A user-friendly shell should be developed that prompts users for all necessary information, and provides lists of allowable values if appropriate.
• Update the CSD by adding entries for the high-priority datasets identified in step (1) above.

4. Prepare complete documentation of the CSD, including instruction in its use.

• Document the structure and contents of the CSD. The most basic documentation will be a data dictionary that defines each of the variables in the CSD database. In addition, a record should be maintained of what steps were taken to add entries to the CSD and to verify their accuracy.

• Develop a set of clearly-written instructions for using the CSD. These instructions should be used in conducting a training workshop for Tampa Bay data users that will introduce them to and instruct them in the use of the CSD.

4.1.2 Standard Data Management Forms and Protocols

To activate the DMS recommendations regarding standard data management practices among decentralized agencies, we recommend that the TBNEP contractor for the implementation phase carry out the following steps (Table 4.1).

1. Take steps to ensure that Tampa Bay agencies participate in both the planning and implementation of the DMS.

• In consultation with the TBNEP staff, identify key agencies and agency contacts. Key agencies will be those most likely to act as data suppliers, e.g., those with extensive or particularly relevant dataset holdings.

• Meet with agency contacts and discuss agency participation in the DMS. During these meetings, explain the DMS recommendations regarding agency agreements related to the CSD and in-house data management practices. The main goal of the meetings should be to involve the agencies as active partners in the development of the DMS. In particular, input from agency contacts should be solicited on exactly what the agreements should include, and how they should be initiated.

• Draft a Memorandum of Understanding in which each agency will agree to
  - meet minimum standards for in-house documentation of datasets,
  - submit entries to the CSD on a regular basis for datasets produced by the agency, and
  - routinely provide agreed-upon data documentation forms with each data transfer.
2. Finalize the recommended forms and protocols for data management.

- Based on the recommendations made in Chapter 3 of this report, review the forms and protocols provided in Appendix C. Each form should be adopted only after evaluating it with regard to (1) whether it meets the minimum standards specified in Chapter 3 and (2) whether it is acceptable to local agencies. Agency contacts should be consulted during this process, and existing forms and protocols should be used wherever possible.

- Submit the final collection of forms and protocols to TBNEP for review. Final modifications to the recommended forms and protocols should then be made as necessary based on the results of the review process.

3. Support the Tampa Bay Regional Coordinating Council Memorandum of Understanding.

- Encourage as many smaller agencies as is feasible to enter into the existing Tampa Bay Regional Planning Council Memorandum of Understanding.

- Prepare a report containing a complete set of forms and protocols, suitable for distribution to participating agencies. The report should describe the purpose of each type of data management record, and should include the minimum criteria from Chapter 3 for whether alternative forms or protocols are acceptable. There should be a clear distinction between cases where specific recommended forms are provided, and those where the form or protocol is intended simply as a guideline, to be modified by individual agencies.

4.1.3 TBNEP Support

The contractor for the implementation phase of the DMS will need to carry out the following activities (Table 4.1) to provide the TBNEP office with the data management services it needs.

1. Provide support to TBNEP contractors who need assistance in submitting their data to ODES.

   - Act as a liaison between TBNEP contractors and ODES, if so requested. The DMS contractor should help other TBNEP contractors to obtain ODES technical support, and should act as a central clearing house for any problems encountered by the contractors.
2. Provide data tracking protocols to the TBNEP office and TBNEP contractors.
   • As early as possible, the DMS contractor should provide the TBNEP office
     with the data documentation forms and protocols that are relevant to
     data transfers to and by the TBNEP office. TBNEP contractors should use
     these protocols in documenting the data that they will submit to TBNEP,
     and the TBNEP office can use the protocols for in-house data
     management. These protocols should be custom-designed, if necessary,
     to meet TBNEP needs (including developing a library system for tracking
     data requests and transfers).

3. Provide data reduction services to the TBNEP office, on an as-needed basis.
   • On request, provide data summaries, maps, or presentation materials that
     are needed to further TBNEP activities.

4.2 LONG-TERM IMPLEMENTATION

   The design and initial implementation of the DMS is intended to provide a
   solid base for long-term data management in the Tampa Bay area. Serious
   consideration must be given to how the strategy that is implemented by TBNEP will
   be maintained at the conclusion of the program. This question must be considered
   at each step of implementing the DMS, but the answer cannot be spelled out in
   detail at this time since it will be highly dependent on the types of inter-agency
   agreements that are acceptable to Tampa Bay agencies. We provide here a brief
   summary of the main issues that will affect the long-term effectiveness of the
   DMS. These issues should be discussed early in the implementation phase at
   meetings with agency contacts.

4.2.1 Central Subject Directory

   Maintenance of the CSD will require storing, distributing, updating, and
   promoting the use of (marketing) the directory (Table 4.2). Storage and distribution
   are basic to making the directory accessible to data users. Updating will be critical
   to maintaining the usefulness of the directory. Marketing to increase the number
   of users of the directory will increase its effectiveness and will make it easier to
   obtain the funding required to maintain the CSD.

   TBNEP funds will be used to initiate the CSD. Alternative sources must be
   identified as soon as possible, as part of the process of building support for the
   CSD. Possible strategies for funding include:
• User-supported. Participating agencies contribute funds, and user fees are charged.

• Government-supported. Federal or state funding.

In addition to basic maintenance of the CSD, funding to compensate data producers for the time and materials required to respond to data requests will be very important in ensuring widespread agency participation in the DMS.

4.2.2 Standard Data Management Forms and Protocols

Proper data documentation is generally recognized as being desirable but is not necessarily fully carried out by data producers. Initial agency agreements are one method of ensuring that data transfers include all the information needed by data users. Adherence to the agreements should be monitored, however, to provide an objective basis for applying methods of motivating/enforcing compliance and providing grounds for extending the agreements. Thorough documentation of data benefits the data producer as well as outside users, and methods of encouraging the use of standardized DMS protocols should be based on this fact. The DMS should allow for and encourage improvements and modifications of the initial recommended forms and protocols. Flexibility will encourage participation, and incorporating feedback will increase the effectiveness of the DMS.

4.3 SUMMARY

Recommendations regarding the TBNEP Data Management Strategy are provided in Chapter 3 of this report. Their implementation will require an initial phase of creating a Central Subject Directory and developing inter-agency agreements on data management and transfer protocols. Their goal is to provide long-term data management that will improve access to information on natural resources and environmental conditions.

Implementing the CSD requires transferring information on historical data from an existing bibliography, updating the information to include recently collected data, and developing user-friendly shells to allow general access to the Directory. In creating the CSD, careful consideration must be given to its long-term maintenance, e.g., to how it will be stored, distributed, updated, and marketed.

Implementing the decentralized components of the DMS requires developing inter-agency agreements related to data management. These agreements will be designed to improve access to information by standardizing data documentation and instructions on requesting copies of the data. The adoption of the agreements
by local agencies will depend on close communication between the TBNEP Data Management Strategy contractor and agency representatives during the implementation phase. Monitoring compliance with the agreements will be essential to ensuring their long-term effectiveness. Flexibility in modifying the recommended forms and protocols based on the practical experience of agencies using them will also contribute to their widespread adoption and long-term usefulness.

Data management support for the TBNEP is essential to meeting the Program's goals. The TBNEP requires extensive access to historical data in order to set living resource targets, identify critical habitats, and assess impacts. The TBNEP must also coordinate data collection by its contractors, and needs to provide contractors with instructions on how data should be submitted, with what accompanying documentation. Public education is another important TBNEP activity that can benefit from access to data management services, in the form of access to data summaries and effective graphics for public presentations.

Efficient sharing of information among agencies provides multiple benefits to data users. The recommended Data Management Strategy described in this report is intended to enhance access to information for data users, while minimizing the cost to data producers. We anticipate that the results of improved sharing of information among Tampa Bay data producers and users will be improved management of natural resources, to the benefit of all Tampa Bay residents.
Table 4.1. Activities required for long-term maintenance of the Central Subject Directory.

<table>
<thead>
<tr>
<th>Type of Activity</th>
<th>Specific Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage</td>
<td>'Base' copy of the database containing the directory</td>
</tr>
<tr>
<td></td>
<td>Documentation of the directory</td>
</tr>
<tr>
<td></td>
<td>Instructions, guidelines for use of the directory</td>
</tr>
<tr>
<td>Distribution</td>
<td>Maintain a bulletin board for remote access.</td>
</tr>
<tr>
<td></td>
<td>Process requests (send floppy disks with copies of the CSD and accompanying documentation and instructions).</td>
</tr>
<tr>
<td>Updating</td>
<td>Process new entries submitted by participating agencies.</td>
</tr>
<tr>
<td></td>
<td>Monitor compliance with inter-agency agreements.</td>
</tr>
<tr>
<td></td>
<td>Transfer entries from and to the FSDD.</td>
</tr>
<tr>
<td></td>
<td>Make any corrections to the CSD that are submitted by users.</td>
</tr>
<tr>
<td>Marketing</td>
<td>To users: Publicize the availability of the directory to potential users.</td>
</tr>
<tr>
<td></td>
<td>To producers: Make modifications that will enhance directory usefulness to contributing agencies.</td>
</tr>
</tbody>
</table>
Table 4.2. Chronology for the first year of the implementation of the Data Management Strategy.

<table>
<thead>
<tr>
<th><strong>Central Subject Directory (CSD)</strong></th>
<th><strong>Decentralized Components</strong></th>
<th><strong>TBNEP Support</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Long-range maintenance plan for CSD: activities, location, funding</td>
<td>Identify key DMS agencies and agency contacts</td>
<td>Offer ODES support</td>
</tr>
<tr>
<td>Finalize CSD structure (check FSDD structure with state)</td>
<td>Meet with agency contacts: prepare draft agreement</td>
<td>Design in-house TBNEP library</td>
</tr>
<tr>
<td>Identify high-priority datasets for updating the CSD</td>
<td></td>
<td>Discuss procedure for providing data reduction services</td>
</tr>
<tr>
<td>Transfer BDB records to CSD</td>
<td>Adopt/modify forms and protocols recommended in Chapter 3</td>
<td>Provide TBNEP-specific data documentation forms and protocols</td>
</tr>
<tr>
<td>Develop user-friendly shell for CSD searches</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obtain documentation for high-priority CSD datasets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Update CSD with high-priority data</td>
<td>Submit report for TBNEP and agency approval</td>
<td></td>
</tr>
<tr>
<td>Develop user-friendly data entry shell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compile complete set of documentation for the CSD database</td>
<td>Finalize agency agreements</td>
<td></td>
</tr>
<tr>
<td>Hold CSD training workshop</td>
<td>Distribute form/protocol report</td>
<td></td>
</tr>
</tbody>
</table>
Appendix A

Survey of Data User Needs and Goals - Survey Instrument
Data Management Strategy

Survey of User Needs and Goals

Goal of the Data Management Strategy

The Tampa Bay National Estuary Program is developing a Data Management Strategy (DMS). The goal of the DMS is to improve the ability of resource managers to locate, store, and access information regarding natural resources (including environmental conditions) in and around Tampa Bay. This could be achieved by developing procedures for improving data access among existing agencies, or by developing a central repository that contains the actual information.

Goals of the Survey of User Needs and Goals

This survey is intended to find out what types of information you need and how you currently obtain them. In addition, we are interested in your opinions about desirable characteristics of the DMS. For example, these might include the ability to provide such products as statistical summaries or maps.

Examples of "Information":

Types of information to be considered in the DMS include:

- **Field Data** (e.g. Hillsborough County water quality data, DNR fisheries-related data files)
- **Maps** (e.g. the National Wetlands Inventory, land use maps)
- **Documentation** (e.g. articles and reports describing data collection methods)
- **Bibliographic Data** (e.g. DIALOG or an online library catalog)
Survey of User Needs and Goals

General User Information

1. In what capacity(s) do you make use of information on the natural resources (including environmental conditions) of Tampa Bay?
   ____ Director (e.g. Agency Director, Department Head)
   ____ Project Manager
   ____ Data Analyst
   ____ Other: ________________________________

2. Which of the following categories best describe the projects that you have worked on in the last year (approximate percent of projects falling into each category):
   ____ Planning (establish goals, set guidelines, develop implementation strategies)
   ____ Regulatory (issue permits or veto a project or activity)
   ____ Advisory (make recommendations or comments)
   ____ Research (contribute research and/or education information)
   ____ Other: ________________________________

3. What final products do you develop using information on natural resources?
   ____ Reports or presentations containing
     ____ Tables
     ____ Graphics
     ____ Statistical analyses
   ____ Maps
   ____ Models
   ____ Other: ________________________________

4. Are there ways of using (processing) information that would be very useful to your job that are not currently available to you? Please list or describe services other than simple data access (e.g. statistical analysis, map overlays...) that would be particularly useful to you if offered by the Data Management Strategy.
5. Please give two (representative) examples of specific information on natural resources or environmental conditions in the Tampa Bay area that you have needed to find in the last year:

1) 

2) 

Evaluation of Existing Data Bases

Please answer the next two questions by marking the list of data bases and Geographic Information Systems (GIS) on the following pages. (Please add data bases to the list if you use ones that are not listed here.)

6. What data bases have you obtained information from in the last year?

In the 'Frequency' column note the approximate number of times you have used the data base in the last year. If the frequency is more than 10, just mark as "10+".

Note: For this question, answer for your agency or department if you are familiar with data transfers in general (fill in A below). Otherwise, answer only for projects on which you have worked (skip A and check B below).

A. Responding for the following agency/department (give the name):

B. Responding only for projects I have worked on: ___

7. On the list of data bases, there is one column for each of 10 features that may affect your interest in using the data base. For each data base you have used or considered using, please indicate which features tend to increase (+), decrease (-), or have no effect (0) on your interest in the data base.

Note: A survey of data base managers will be used to objectively describe each data base. Here we are interested in your impressions and opinions.

If you consider features not listed in the table to be important, please note them here:
### Survey of User Needs and Goals

<table>
<thead>
<tr>
<th>Data Base</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effect on Interest in the Data Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ = Increases Interest, - = Decreases Interest</td>
</tr>
<tr>
<td>0 = Little if any effect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Contents¹</th>
<th>Access Tools²</th>
<th>Transfers³</th>
<th>Info on Data⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td>Detail</td>
<td>Searches</td>
<td>Subsets</td>
</tr>
</tbody>
</table>

1. Contents relative to your needs:
   - Type of data
   - Detail available (dates, map scale...)

2. Tools for data access and manipulation:
   - Searches (locating information)
   - Subsets (extracting)
   - Analysis (tables, statistics)

3. Data transfers:
   - Ease
   - Cost
   - Time for turnaround

4. Information on the data:
   - Methods used to collect the data
     (where, when, how)
   - Quality assurance

---

<table>
<thead>
<tr>
<th>Environmental Protection Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODES (Ocean Data Evaluation System)</td>
</tr>
<tr>
<td>STORET (STOrage and RETrieval)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>U.S. Geological Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAWDEX (National Water Data Exchange)</td>
</tr>
<tr>
<td>WATSTOR (Water Data Storage and Retrieval System)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>National Oceanic and Atmospheric Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPAS (Coastal Ocean Management, Planning and Assessment System)</td>
</tr>
<tr>
<td>NEDRES (National Environmental Data Referral Service)</td>
</tr>
</tbody>
</table>
# Survey of User Needs and Goals

<table>
<thead>
<tr>
<th>Data Base</th>
<th>Frequency</th>
<th>Effect on Interest in the Database</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ = Increases Interest, - = Decreases Interest 0 = Little if any effect</td>
</tr>
<tr>
<td></td>
<td>Contents(^1)</td>
<td>Access Tools(^2)</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Type</td>
<td>Detail</td>
<td>Searches</td>
</tr>
</tbody>
</table>

**Southwest Florida Water Management District**

- SWIM Bibliographic Data Base
- SWIM Water Quality Data
- Little Manatee River Water Quality Data
- Land Use/Land Cover GIS
- Storm Water Permits GIS
- Surface Water Permitting Data Base
- Water Use Permitting Data Base

**Environmental Protection Commission of Hillsborough County**

- Surface Water Quality Data Base
- Mitigation of Wetlands Data Base

---

1. Contents relative to your needs
   - Type of data
   - Detail available (dates, map scale...)

2. Tools for data access and manipulation:
   - Searches (locating information)
   - Subsets (extracting)
   - Analysis (tables, statistics)

3. Data transfers:
   - Ease
   - Cost
   - Time for turnaround

4. Information on the data:
   - Methods used to collect the data (where, when, how)
   - Quality assurance
<table>
<thead>
<tr>
<th>Data Base</th>
<th>Frequency</th>
<th>Effect on Interest in the Database</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ = Increases Interest, - = Decreases Interest 0 = Little if any effect</td>
</tr>
<tr>
<td></td>
<td>Contents¹</td>
<td>Access Tools²</td>
</tr>
<tr>
<td></td>
<td>Type</td>
<td>Detail</td>
</tr>
<tr>
<td>City of Tampa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water Quality Monitoring Program Data Base</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Natural Resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land Cover GIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant Communities GIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCS Soils Data GIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisheries-Independent Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game and Fresh Water Fish Commission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wildlife Habitat Mapping GIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wading Bird Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wildlife Observations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Florida Natural Areas Inventory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rare/Endangered Species</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rare and Exemplary Natural Communities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Contents relative to your needs
   Type of data
   Detail available (dates, map scale...)

2 Tools for data access and manipulation:
   Searches (locating information)
   Subsets (extracting)
   Analysis (tables, statistics)

3 Data transfers:
   Ease
   Cost
   Time for turnaround

4 Information on the data:
   Methods used to collect the data
   (where, when, how)
   Quality assurance
### Survey of User Needs and Goals

<table>
<thead>
<tr>
<th>Data Base</th>
<th>Frequency</th>
<th>Effect on Interest in the Data Base</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ = Increases Interest, - = Decreases Interest</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 = Little if any effect</td>
</tr>
<tr>
<td>Contents¹</td>
<td>Access Tools²</td>
<td>Transfers³</td>
</tr>
<tr>
<td>Type</td>
<td>Detail</td>
<td>Searches</td>
</tr>
</tbody>
</table>

#### West Coast Regional Water Supply Authority

- Hydrologic Data Base
- Water Quality Data Base
- Vegetation Data

#### Others:

- DIALOG online information retrieval system

---

1. Contents relative to your needs
   - Type of data
   - Detail available (dates, map scale...)

2. Tools for data access and manipulation:
   - Searches (locating information)
   - Subsets (extracting)
   - Analysis (tables, statistics)

3. Data transfers:
   - Ease
   - Cost
   - Time for turnaround

4. Information on the data:
   - Methods used to collect the data (where, when, how)
   - Quality assurance
## Survey of User Needs and Goals

<table>
<thead>
<tr>
<th>Data Base</th>
<th>Frequency</th>
<th>Effect on Interest in the Data Base</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ = Increases Interest, - = Decreases Interest</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contents¹</th>
<th>Access Tools²</th>
<th>Transfers³</th>
<th>Info on Data⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Detail</td>
<td>Searches</td>
<td>Subsets</td>
</tr>
</tbody>
</table>

1 Contents relative to your needs
- Type of data
  - Detail available (dates, map scale...)

2 Tools for data access and manipulation:
  - Searches (locating information)
  - Subsets (extracting)
  - Analysis (tables, statistics)

3 Data transfers:
  - Ease
  - Cost
  - Time for turnaround

4 Information on the data:
  - Methods used to collect the data
  - (where, when, how)
  - Quality assurance
Survey of User Needs and Goals

Geographic Information System Needs

Geographic Information System (GIS) software provides tools for storing, analyzing, and presenting spatial data.

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Does your agency own a GIS?</td>
<td>Yes / No</td>
</tr>
<tr>
<td>8a. If so, which software package?</td>
<td></td>
</tr>
<tr>
<td>8b. On what type of computer? (e.g. PC, workstation, mainframe)</td>
<td></td>
</tr>
<tr>
<td>8c. If not, do you expect to acquire one in the next 1-2 years?</td>
<td>Yes / No / Don’t Know</td>
</tr>
<tr>
<td>9. Have you ever seen a demo of a GIS?</td>
<td>Yes / No</td>
</tr>
<tr>
<td>10. Has a GIS been used in any project on which you have worked?</td>
<td>Yes / No</td>
</tr>
<tr>
<td>11. Have you personally used a GIS?</td>
<td>Yes / No</td>
</tr>
<tr>
<td>12. How important do you consider access to a GIS to be for your work?</td>
<td>Currently? ____</td>
</tr>
<tr>
<td>(On a scale of 0-5 where 0 = Don’t know, 1 = Little or no importance, 5 = Extremely important)</td>
<td>In the next 5 years? ____</td>
</tr>
</tbody>
</table>
Survey of User Needs and Goals

Hardware and Software

13. What hardware and software does your agency commonly use to access and manipulate information for projects with which you are familiar?

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Yes / No</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM-compatible PC</td>
<td></td>
</tr>
<tr>
<td>MacIntosh</td>
<td></td>
</tr>
<tr>
<td>Workstation</td>
<td></td>
</tr>
<tr>
<td>Mainframe</td>
<td></td>
</tr>
<tr>
<td>Other:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Yes / No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modem</td>
<td></td>
</tr>
<tr>
<td>Floppy disks</td>
<td>Yes / No (for transfers)</td>
</tr>
<tr>
<td>Scanner</td>
<td>Yes / No (to input data)</td>
</tr>
<tr>
<td>Removable cartridges</td>
<td></td>
</tr>
<tr>
<td>Tapes</td>
<td></td>
</tr>
<tr>
<td>Other:</td>
<td></td>
</tr>
</tbody>
</table>

Software
(please give the name of commonly used programs, e.g. dBASE or SAS)

<table>
<thead>
<tr>
<th>Communications</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfers:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Database:</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Statistics:</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Graphics</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Other:</th>
<th></th>
</tr>
</thead>
</table>
Recommendations for the Data Management Strategy

14. Implementing the DMS will depend on funding and other constraints. In your opinion, which of the following types of data should definitely be included in the DMS? Please circle your choices (individual words within lines, entire lines, or major categories).

Geology
- Geological formations
- Glacial deposits, sediments
- Soil type
- Bathymetry (for Tampa Bay)
- Other: ______________________

Climate
- Rainfall, temperature, wind speed
- Severe weather events
- Other: ______________________

Surface and Ground Waters
- Tributary watersheds, flow rates
- Nutrients, water chemistry
- Ground water flow, aquifer characteristics
- Circulation, flushing, tidal patterns (in Tampa Bay)
- Other: ______________________

Biology
- Phytoplankton, zooplankton
- Seagrasses, wetlands
- Terrestrial land cover (habitat maps)
- Benthic invertebrates, shellfish, fish
- Reptiles, amphibians, birds
- Marine mammals, terrestrial mammals
- Other: ______________________

Human Population
- Census data
- Land use, political boundaries, property maps
- Permitting: air / water / construction / dredging
- Dredge and fill areas, marinas and docks
- Archeological/Historical sites
- Other: ______________________

Other: ______________________
15. In terms of the dates when data were collected, how important is it to include data from the following periods in the Data Management Strategy? (H = High Priority, M = Medium Priority, L = Low Priority)

___ Early (pre-1980)

___ Recent (1980+)

___ Very current (including ongoing projects)

16. Data are currently scattered among many different agencies. A central directory or storage location could be created as part of the Data Management Strategy but would require a considerable investment to develop and maintain. If created it could contain any or all of the following types of information. Which (if any) do you think should definitely be made available at a central location?

___ A guide to what information exists for the Tampa Bay area by topic, place, date.

___ Information on where data is stored and instructions on how to obtain copies.

___ Documentation describing sampling methods used in collecting the data.

___ Information on steps taken to ensure the accuracy of the data in each data base.

___ The actual data.

Thank you!
Appendix B

Survey of Data User Needs and Goals - Summary of Responses
Survey of Data User Needs and Goals

Summary of Responses (49 Returns)

1. In what capacity(s) do you make use of information on the natural resources (including environmental conditions) of Tampa Bay?

   26% Director (e.g. Agency Director, Department Head)
   54% Project Manager
   33% Data Analyst
   41% Other (Stormwater planning; SeaGrant extension agent; land acquisition; planning administrator; chemist; sanctuary manager; chief biologist; consultant; permit review; engineer; modeling; extensive studies; education; environmental scientist)

2. Which of the following categories best describe the projects that you have worked on in the last year (approximate percent of projects falling into each category):

   67% Planning (establish goals, set guidelines, develop implementation strategies)
   65% Regulatory (issue permits or veto a project or activity)
   59% Advisory (make recommendations or comments)
   37% Research (contribute research and/or education information)
   30% Other

3. What final products do you develop using information on natural resources?

   Reports or presentations containing
   80% Tables
   78% Graphics
   71% Statistical analyses
   75% Maps
   37% Models
   18% Other:
4. Are there ways of using (processing) information that would be very useful to your job that are not currently available to you? Please list or describe services other than simple data access (e.g. statistical analysis, map overlays...) that would be particularly useful to you if offered by the Data Management Strategy.

* mapping of habitats and vegetative communities
* one-stop shopping for Tampa Bay resource information
* infrastructure mapping (channels, piping)
* "user-friendly" front-ends for the local ARC/INFO GIS Systems
* hydrologic soil classifications (A, B, C, D soils)
* hydrologic land uses
* land use translations (depends upon specific types of monitoring)
* models for determining impacts during low-flow periods for tidally influenced tributaries in watershed basins.
* access to STORET and opportunity to learn STORET methods
* regulation updates (for stormwater, wetlands, etc.)
* interfacing of software (system that allows everything for one report to be stored on one disk)
* GIS
* LANDSAT imagery
* mapping
* aerial photography
* ownership (property) maps

* time series aerial photographs
* GIS
* detailed habitat mapping
* hard copies of GIS and overlaps
* digitized older data
* economic analysis (useful for regulation justification) e.g. total contribution per acre of wetland communities to region (including all values)
* GIS overlaps
* descriptive statistics (e.g. sort, means, averages, ranges, etc.)
* A DMS that would inter-relate with models for particular basin projects (channel improvements, SWIM surface water quality enhancement projects) would help review and prioritize such projects. We would like to see how each channel project affects Tampa Bay and, if applicable, any rivers.
4 (cont.). Are there ways of using (processing) information that would be very useful to your job that are not currently available to you?

* water quality analysis/cumulative data (maps of some info.)

* wetland impacts - (maps also)

* environmental/urban ecology modeling

* wildlife/habitat resources modeling

* hydrology/flow statistics for riverine systems

* documented positive management results showing clear cause and effect relationship.

* the consolidation and conversion (along with proper quality assurance checks) of disparate data sets on natural resources from the bay

* data reduction services, and to a lesser degree, statistical analysis

* graphics program compatible with all the other graphics users: FDER, SWFWMD, FGFWFC Tallahassee

* time series analysis

* spectral analysis

* bay information on a CD that can be searched at the local libraries (USF, SPJC, etc.)

* GIS based maps which are compatible at all levels.

* accurate digital maps of a wide variety of environmental overlap data that can be processed and is managed and with accurate data documentation accessible on a real-time basis at no cost.

* central location for obtaining maps

* complete statistical analysis of existing water quality data

* automated mapping/data management system (GIS)

* easy access to summary information on a Tampa Bay region basis

* ambient monitoring summaries for water and sediment quality

* inventory or map of unique marine and shoreline habitats

* regulatory tracking system that combines federal, state and local government permit and mitigation actions

* qualitative indexes for wetland and upland habitats

* trend analysis and models that assess cumulative impact on the bay systems (i.e., the impacts of pollutant loading and habitat loss on living marine resources and other related wildlife)
5. Please give two (representative) examples of specific information on natural resources or environmental conditions in the Tampa Bay area that you have needed to find in the last year:

* bathymetry

* updated marine habitat maps

* colonial water-bird breeding areas in Tampa Bay

* habitat map of MacDill airforce base

* trend maps for natural resources

* water quality conditions

* what kinds of water quality impacts are flowing into the Bay?

* How much freshwater is flowing into the bay during low-flow periods?

* water quality for tributaries flowing into Tampa Bay

* nutrient data for Tampa Bay

* Delaney Creek water quality data

* seagrass distribution

* areas of public ownership

* restoration projects

* hydrological information

* flow data for tributaries in Boca Ciega Bay

* rainfall data for Indian Rock Beach for past 20 years

* bacteriological data on Lake Chatoqua (sp?)

* research on Allen’s Creek (modeling)

* aerial photography of Egmont Key to determine erosion

* nitrogen to phosphorous ratio for all of Tampa Bay

* seagrass mapping

* water quality data from the HCEPC database

* hydrologic and meteorological data from SWFWMD’s Hydrologic database.

* NOAA rainfall (climate data)

* meteorological data

* water quality/light penetration

* Where are sludge disposal areas located in the watershed?

* When a permittee sends in compliance information, where does that data go?

* water quality chemical data

* biological data

* commercial fishery landings
Please give two (representative) examples of specific information on natural resources or environmental conditions in the Tampa Bay area that you have needed to find in the last year:

* vegetative community mapping
* % coverage seagrass/mangroves in Hillsborough, Manatee, and Pinellas counties over time
* fisheries data
* drainage areas and flows
* economic value of marine systems (e.g. salt marsh, mangrove community) measured as commercial/recreational fish landings
* contribution of stormwater loads from various land uses around Tampa Bay
* locations of environmentally sensitive areas in the bay and in the watershed
* ambient water quality measurements
* How will replacement of Mullet Key Bridge affect marine water quality & adjacent shellfish?
* What is water quality in Tampa Bay at mouth of Delaney Creek & how will Hillsborough County’s proposed channel improvement affect it?
* hard bottom habitat offshore of sandy beaches (barrier islands)
* infauna of dredge & nourishment sites
* Water quality for all receiving water bodies discharging into Bay
* flow capacities (low, high, natural) for segment of creek or riverine systems
* estimated pollutant load in stormwater
* seagrass bed scarring frequency & recovery rates
* locations of species concentrations
* habitat maps
* available and potential habitat for select shellfish populations
* delineation of specific benthic habitats and description of their dominant components
* location of breeding bird colonies
* location of sea grass beds
* distribution of wetland vegetation in tributaries (current and historical)
* general water quality data
* long-term water quality trends & inter-relationships
Please give two (representative) examples of specific information on natural resources or environmental conditions in the Tampa Bay area that you have needed to find in the last year:

* circulation, residence time, transport rates for Tampa Bay
* seagrass beds/endangered habitat
* historic information of environmental quality
* zooplankton species composition & abundance
* juvenile fish distribution within rivers that feed to Tampa Bay
* wetlands, habitat, wildlife
* water quality at small area level
* seagrasses
* circulation
* distribution & composition of submerged aquatic vegetation
* water quality trends and existing conditions
* Tampa Bay water quality data - comprehensive long-term data
* USGS gaging station flow data from tributaries into Tampa Bay
* water quality associated with specific land uses (e.g. agriculture)
* stage duration data on wetlands water levels
* current, tidal, freshwater inflow data
* distribution of salinity in Hillsborough River
* location of publicly owned land in Tampa Bay watershed
* existing channel and adjacent resources information for Port of Tampa and Port Manatee
* quality of wetland and upland habitats in NW Tampa and their importance to wildlife
* Historical and current seagrass distribution maps
* water quality in Allen’s Creek
* water quality parameters for tributaries to Tampa Bay
* climate data – air quality, rainfall, temperature
* water quality
* seagrass distribution
* shoreline habitat type and coverage for permit review purposes
* recreational/commercial fisheries harvest data
Please give two (representative) examples of specific information on natural resources or environmental conditions in the Tampa Bay area that you have needed to find in the last year

* meteorological data (rainfall, wind)
* water quality data (EPC)
* population growth
* registered vehicles

**Evaluation of Existing Data Bases**

Please answer the next two questions by marking the list of data bases and Geographic Information Systems (GIS) on the following pages. (Please add data bases to the list if you use ones that are not listed here.)

6. What data bases have you obtained information from in the last year?

   In the ’Frequency’ column note the approximate number of times you have used the data base in the last year. If the frequency is more than 10, just mark as "10+".

Note: The 'Total' column includes a separate tally (in parentheses) of respondents reporting use of in-house data (datasets stored at the agency where they work).
<table>
<thead>
<tr>
<th>Data Base</th>
<th>Number of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency of Use</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
</tr>
<tr>
<td><strong>Environmental Protection Agency</strong></td>
<td></td>
</tr>
<tr>
<td>ODES (Ocean Data Evaluation System)</td>
<td>1</td>
</tr>
<tr>
<td>STORET (STOrage and RETrieval)</td>
<td>5</td>
</tr>
<tr>
<td><strong>U.S. Geological Survey</strong></td>
<td></td>
</tr>
<tr>
<td>NAWDEX (National Water Data Exchange)</td>
<td>1</td>
</tr>
<tr>
<td>WATSTOR (Water Data Storage and Retrieval System)</td>
<td>3</td>
</tr>
<tr>
<td><strong>National Oceanic and Atmospheric Administration</strong></td>
<td></td>
</tr>
<tr>
<td>COMPAS (Coastal Ocean Management, Planning and Assessment System)</td>
<td>1</td>
</tr>
<tr>
<td>NEDRES (National Environmental Data Referral Service)</td>
<td>1</td>
</tr>
<tr>
<td><strong>Southwest Florida Water Management District</strong></td>
<td></td>
</tr>
<tr>
<td>SWIM Bibliographic Data Base</td>
<td>7</td>
</tr>
<tr>
<td>SWIM Water Quality Data</td>
<td>2</td>
</tr>
<tr>
<td>Little Manatee River Water Quality Data</td>
<td>3</td>
</tr>
<tr>
<td>Land Use/Land Cover GIS</td>
<td>5</td>
</tr>
<tr>
<td>Storm Water Permits GIS</td>
<td>2</td>
</tr>
<tr>
<td>Surface Water Permitting Data Base</td>
<td>4</td>
</tr>
<tr>
<td>Water Use Permitting Data Base</td>
<td>2</td>
</tr>
<tr>
<td>SWIM Seagrass Survey</td>
<td>1</td>
</tr>
<tr>
<td>Hydrologic Data Base</td>
<td></td>
</tr>
<tr>
<td>Data Base</td>
<td>Number of Respondents</td>
</tr>
<tr>
<td>----------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>Frequency of Use</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
</tr>
<tr>
<td><strong>Environmental Protection Commission of Hillsborough County</strong></td>
<td></td>
</tr>
<tr>
<td>Surface Water Quality Data Base</td>
<td>8</td>
</tr>
<tr>
<td>Mitigation of Wetlands Data Base</td>
<td>1</td>
</tr>
<tr>
<td><strong>City of Tampa</strong></td>
<td></td>
</tr>
<tr>
<td>Water Quality Monitoring Program Data Base</td>
<td>4</td>
</tr>
<tr>
<td><strong>Department of Natural Resources</strong></td>
<td></td>
</tr>
<tr>
<td>Land Cover GIS</td>
<td>4</td>
</tr>
<tr>
<td>Plant Communities GIS</td>
<td>2</td>
</tr>
<tr>
<td>SCS Soils Data GIS</td>
<td>2</td>
</tr>
<tr>
<td>Fisheries-Independent Data</td>
<td>2</td>
</tr>
<tr>
<td><strong>Game and Fresh Water Fish Commission</strong></td>
<td></td>
</tr>
<tr>
<td>Wildlife Habitat Mapping GIS</td>
<td>4</td>
</tr>
<tr>
<td>Wading Bird Data</td>
<td>2</td>
</tr>
<tr>
<td>Wildlife Observations</td>
<td></td>
</tr>
<tr>
<td><strong>Florida Natural Areas Inventory</strong></td>
<td></td>
</tr>
<tr>
<td>Rare/Endangered Species</td>
<td>4</td>
</tr>
<tr>
<td>Rare and Exemplary Natural Communities</td>
<td>4</td>
</tr>
<tr>
<td>Data Base</td>
<td>Number of Respondents</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>Frequency of Use</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
</tr>
<tr>
<td>West Coast Regional Water Supply Authority</td>
<td></td>
</tr>
<tr>
<td>Water Quality Data Base</td>
<td></td>
</tr>
<tr>
<td>Department of Environmental Regulation</td>
<td></td>
</tr>
<tr>
<td>Point Source Discharges</td>
<td>1</td>
</tr>
<tr>
<td>Permit Information</td>
<td>1</td>
</tr>
<tr>
<td>Groundwater Management System</td>
<td>1</td>
</tr>
<tr>
<td>Nonpoint Pollution Assessment Data</td>
<td>1</td>
</tr>
<tr>
<td>Water Resources Bibliography</td>
<td>1</td>
</tr>
<tr>
<td>Others:</td>
<td></td>
</tr>
<tr>
<td>DIALOG online information retrieval system</td>
<td></td>
</tr>
<tr>
<td>Florida Breeding Bird Atlas</td>
<td></td>
</tr>
<tr>
<td>NMFS Commercial Fish Landings</td>
<td></td>
</tr>
<tr>
<td>Wetland Status Tracking System</td>
<td></td>
</tr>
</tbody>
</table>

7. On the list of data bases, there is one column for each of 10 features that may affect your interest in using the data base. For each data base you have used or considered using, please indicate which features tend to increase (+), decrease (-), or have no effect (0) on your interest in the data base.

(Responses to Question #7 indicated that it was not interpreted consistently. We did not attempt, therefore, to summarize the responses quantitatively.)
8. Does your agency own a GIS? Yes - 55%

8a. If so, which software package?
ARC/INFO 52%
GEOVISION 22%
OTHER 26%
(1 response each for: INTEGRAPH; MAPINFO; AUTOCAD; ERDAS; GENESIS; EPPL7; PROPRIETARY)

8b. On what type of computer? PC 44%
(e.g. PC, workstation, Workstation 30%
mainframe) Mainframe 37%
Other 7%

8c. If not, do you expect to acquire one in the next 1-2 years? Yes - 14%

9. Have you ever seen a demo of a GIS? Yes - 84%

10. Has a GIS been used in any project on which you have worked? Yes - 82%

11. Have you personally used a GIS? Yes - 33%

12. How important do you consider access to a GIS to be for your work? (On a scale of 0-5 where
0 = Don't know, 1 = Little or no importance, 5 = Extremely important)

Currently: Next 5 years:
0 - 9% 2%
1 - 11% 0%
2 - 9% 6%
3 - 23% 4%
4 - 5% 11%
5 - 52% 77%
Hardware and Software

13. What hardware and software does your agency commonly use to access and manipulate information for projects with which you are familiar?

Hardware

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM-compatible PC</td>
<td>94%</td>
</tr>
<tr>
<td>MacIntosh</td>
<td>11%</td>
</tr>
<tr>
<td>Workstation</td>
<td>41%</td>
</tr>
<tr>
<td>Mainframe</td>
<td>50%</td>
</tr>
<tr>
<td>Other: (VAX cluster, platforms)</td>
<td>10%</td>
</tr>
<tr>
<td>Modem</td>
<td>80%</td>
</tr>
<tr>
<td>Floppy disks</td>
<td>95% (for transfers)</td>
</tr>
<tr>
<td>Scanner</td>
<td>25% (to input data)</td>
</tr>
<tr>
<td>Removable cartridges</td>
<td>25%</td>
</tr>
<tr>
<td>Tapes</td>
<td>45%</td>
</tr>
<tr>
<td>Other: (Hard drive, optical disks)</td>
<td>5%</td>
</tr>
</tbody>
</table>

Software

(please give the name of commonly used programs, e.g. dBASE or SAS)

<table>
<thead>
<tr>
<th>Type of Software</th>
<th>Total Number of Programs Reported</th>
<th>Programs Cited by &gt;= 10% of Survey Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Program</td>
</tr>
<tr>
<td>Communications</td>
<td>21</td>
<td>ProCom Plus</td>
</tr>
<tr>
<td>Spreadsheet</td>
<td>10</td>
<td>Lotus 123</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quattro Pro</td>
</tr>
<tr>
<td>Database</td>
<td>16</td>
<td>dBASE III or IV</td>
</tr>
<tr>
<td>Statistics</td>
<td>12</td>
<td>SAS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statgraph</td>
</tr>
<tr>
<td>Graphics</td>
<td>28</td>
<td>Harvard Graphics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lotus 123</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quattro Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUTOCAD</td>
</tr>
</tbody>
</table>
Recommendations for the Data Management Strategy

14. Implementing the DMS will depend on funding and other constraints. In your opinion, which of the following types of data should definitely be included in the DMS? Please circle your choices (individual words within lines, entire lines, or major categories).

Note: ‘#’ = the number of respondents selecting each topic. Items in plain typeface were offered as choices on the survey instrument. Items in italics were written in by respondents.

<table>
<thead>
<tr>
<th>#</th>
<th>Item</th>
<th>#</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>Geology</td>
<td>39</td>
<td>Climate</td>
</tr>
<tr>
<td>10</td>
<td>Geological formations</td>
<td>34</td>
<td>Rainfall</td>
</tr>
<tr>
<td>27</td>
<td>Glacial deposits</td>
<td>30</td>
<td>Temperature</td>
</tr>
<tr>
<td>37</td>
<td>Sediments</td>
<td>28</td>
<td>Wind speed</td>
</tr>
<tr>
<td>39</td>
<td>Soil type</td>
<td></td>
<td>Severe weather events</td>
</tr>
<tr>
<td>3</td>
<td>Bathymetry (for Tampa Bay)</td>
<td>3</td>
<td>Evapotranspiration</td>
</tr>
<tr>
<td>2</td>
<td>Topography</td>
<td>2</td>
<td>Wind direction (by season)</td>
</tr>
<tr>
<td>2</td>
<td>Sediment type</td>
<td>1</td>
<td>Rainfall quality</td>
</tr>
<tr>
<td>1</td>
<td>Soil chemistry</td>
<td>1</td>
<td>Atmospheric deposition of pollutants (dryfall/wetfall)</td>
</tr>
<tr>
<td>1</td>
<td>Phosphate deposits</td>
<td>1</td>
<td>Fire frequency</td>
</tr>
<tr>
<td>1</td>
<td>Mining locations (past and present)</td>
<td>1</td>
<td>Radiation</td>
</tr>
<tr>
<td>1</td>
<td>Coral/hard bottom locations</td>
<td>1</td>
<td>Air quality</td>
</tr>
<tr>
<td>1</td>
<td>Island and mangrove locations</td>
<td>1</td>
<td>PAR</td>
</tr>
<tr>
<td>1</td>
<td>Bottom sediment composition</td>
<td>1</td>
<td>Frontal passages</td>
</tr>
<tr>
<td>1</td>
<td>Subtidal bottom types</td>
<td></td>
<td>Human Population</td>
</tr>
<tr>
<td>41</td>
<td>Surface and Ground Waters</td>
<td>31</td>
<td>Census data</td>
</tr>
<tr>
<td>42</td>
<td>Tributary watersheds</td>
<td>34</td>
<td>Land use</td>
</tr>
<tr>
<td>45</td>
<td>Flow rates</td>
<td>28</td>
<td>Political boundaries</td>
</tr>
<tr>
<td>43</td>
<td>Nutrients</td>
<td>28</td>
<td>Property maps</td>
</tr>
<tr>
<td>32</td>
<td>Water chemistry</td>
<td>25</td>
<td>Permitting: air</td>
</tr>
<tr>
<td>33</td>
<td>Ground water flow</td>
<td>35</td>
<td>Permitting: water</td>
</tr>
<tr>
<td>41</td>
<td>Aquifer Characteristics</td>
<td>28</td>
<td>Permitting: construction</td>
</tr>
<tr>
<td>41</td>
<td>Circulation</td>
<td>34</td>
<td>Permitting: dredging</td>
</tr>
<tr>
<td>41</td>
<td>Flushing</td>
<td>34</td>
<td>Dredge and fill areas</td>
</tr>
<tr>
<td>41</td>
<td>Tidal patterns</td>
<td>26</td>
<td>Marinas and docks</td>
</tr>
<tr>
<td>2</td>
<td>Water levels</td>
<td>27</td>
<td>Archeological/historical sites</td>
</tr>
<tr>
<td>1</td>
<td>Saltwater/freshwater wedges</td>
<td>1</td>
<td>Service area boundaries (for sewer systems, etc.)</td>
</tr>
<tr>
<td>1</td>
<td>Water quality loading rates</td>
<td></td>
<td>Habitat restoration efforts</td>
</tr>
<tr>
<td>1</td>
<td>Water table changes (rate &amp; direction)</td>
<td>1</td>
<td>Local government jurisdictions (with reference to data they collect)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>In-water infrastructure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>Adjacent land uses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>Areas with special management designations</td>
</tr>
</tbody>
</table>
15. In terms of the dates when data were collected, how important is it to include data from the following periods in the Data Management Strategy? (H = High Priority, M = Medium Priority, L = Low Priority)

L = 20% M = 40% H = 40% Early (pre-1980)
L = 0% M = 20% H = 80% Recent (1980+)
L = 4% M = 15% H = 80% Very current (including ongoing projects)

16. Data are currently scattered among many different agencies. A central directory or storage location could be created as part of the Data Management Strategy but would require a considerable investment to develop and maintain. If created it could contain any or all of the following types of information. Which (if any) do you think should definitely be made available at a central location?

91% A guide to what information exists for the Tampa Bay area by topic, place, date.
91% Information on where data is stored and instructions on how to obtain copies.
40% Documentation describing sampling methods used in collecting the data.
45% Information on steps taken to ensure the accuracy of the data in each data base.
36% The actual data.
Comments:

* Standardized format for data collection method for GIS information needs to be developed.

* Is it important to have all of this data in DMS, or just to know where to get it? Need: simple format, one central location.
  - where and how can I get it?
  - make sure data is actually there and available
  - UPDATING is extremely important!

* Don't attempt to categorize unpublished data; don't attempt to link existing databases into one central location (it's a waste of time).

* Maintenance and updating are extremely important.

* There is a definite need for coordination of existing databases.
  - State of FL has GIS network, but access and compatibility are not always easy
  - difficult for government agencies to coordinate w/other agencies.
  - The simpler the approach, the better off we'll be.
  - Need for PC based software that is compatible to mainframes.

* Hardware limitations will be the major problem for those agencies who don't have advanced computer systems.

* Nothing should ever have to be key-punched twice! Everything should be compatible (in format) for easy access.

* Who should be the central custodian of environmental data?

* There is a definite need for a central data repository
  - issues: data uniformity for various computer software; updating; maintenance

* There is a need for improved access to information on public ownership data through tax collections office.

* There is a need for plant community data vegetation surveys.
Comments (cont.)

* Many informational sources are unavailable to our office from ignorance of their existence, or a cheap way to access them.

* It's difficult to imagine a Tampa Bay system delivering processed information based on the current status of the availability of information. Data access that is "simple," in an established format, and that has been through some sort of QA/QC protocol would be a major accomplishment. My major problem is that most of the organizations have not yet established a standardized method for data management, and thus, asking the organizations themselves is often a laborious and difficult task for the organization personnel. Once data is finally obtained, it is rarely in a usable format!

* The actual data in a central repository would be extremely difficult to accomplish based on my opinion of the adequacy of data management by organizations at the present time.

* DMS should only contain actual data if funding permits. However, agencies with existing centralized systems will continue to use their own and duplicate storage will undoubtedly occur.

* Water quality data in databases is hard to understand in terms of quality of the data. There is a lot of sophistication on my part (and of others) in accessing most databases. Much data is not retrievable accept by starting from scratch--finding data contacts and visiting to copy paper copies. One realistic product from the DMS could be a DOC list for types of data.
Appendix C

Sample Data Documentation Forms and Protocols

C1 - Study Overview
New York-New Jersey Harbor Estuary Program Data Documentation Form
(Instructions, blank form, and sample)

and Analysis: Data Documentation and Statistical Support. Prepared
for the New York/New Jersey Harbor Estuary Program.

C2 - Description of Study Methods and Sampling Design
1991 Aerial Survey of Sea Mammals in Tampa Bay, Methods Description
from Report.

Weigle, B.L., J.E. Reynolds III, B.B. Ackerman, I.E. Beeler, and P.L. Boland.
1991. Distribution and abundance of bottlenose dolphins (Tursiops
truncatus) in Tampa Bay. In Proceedings, Tampa Bay Area Scientific
Information Symposium 2. Available from TEXT Technical Writing
Service, Tampa, Florida.

C3 - Description of Dataset Contents
Florida Growth Management Data Network Coordinating Council, Standards
for the Data Dictionary and Quality and Accuracy Report.

Data Dictionary and Quality and Accuracy Report.

C4 - Record of Dataset Quality Review
Florida Growth Management Data Network Coordinating Council, Standards
for the Data Dictionary and Quality and Accuracy Report.

Data Dictionary and Quality and Accuracy Report.
C5 - Transaction Log
Modified From Maryland Power Plant Research Program, Data Tracking Form.


C6 - Summary of Data Transfer Policies and Procedures
Southwest Florida Water Management District, GIS Data Distribution Procedures.


C7 - Data Request Form
Southwest Florida Water Management District, Geographic Information System Data Order Form.


C8 - Code Translation Tables
Virginia Marine Resources Commission (VMRC) Fisheries Statistics Gear Code Translation Table.


C9 - Data Reduction Services Offered
Virginia Marine Fisheries Commission Interactive Data Summarization Programs.

New York-New Jersey Harbor Estuary Program Data Documentation Form
(Instructions, blank form, and sample)

and Analysis: Data Documentation and Statistical Support. Prepared
for the New York/New Jersey Harbor Estuary Program.
New York-New Jersey Harbor Estuary Program
Data Documentation Form - Instructions

Purpose: To help make existing data available for further research and management use.

In particular, to provide information about the data that will allow other investigators to make a preliminary judgement regarding whether or not the data is of interest to them.

One form has been prepared for use with a wide variety of data. Not all categories (i.e. fields) will be applicable to all data sets. Please use n/a (not applicable) or similar comments to indicate that a given field is deliberately left blank.

A sample is provided to help clarify how the form should be filled out. When in doubt, consider the purpose of the form and try to provide the information that will help to meet that purpose.

Brief description of each item:
Dataset Title Identify the study for which the data was collected
Date(s) Year(s) for which data exists
Taxonomic group(s) Species (or more general categories) for which data exists
Publications/reports Citations for literature reporting the results of the study, or describing the methods used.
Collected by Principal investigator for the study. A contact for questions regarding the hows or whys of data collection.
Sponsor(s) Name of organization funding/sponsoring the study.
Available from Who to contact to request a copy of the data
Available as Format in which copies of the data are available.
Restrictions on access Limits on who can obtain copies of the data, or on what parts of the data set are available
Abstract 1-paragraph description of the study. Be sure to include goals (original purpose of the study).
<table>
<thead>
<tr>
<th><strong>Area sampled</strong></th>
<th>Total region for which data is available.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method of choosing sites</strong></td>
<td>Whether sampling sites were chosen at random (or stratified random, etc. designs), vs. fixed sites chosen using criteria such as ease of access or known interest relative to the goals of the study.</td>
</tr>
<tr>
<td><strong>Total number of sites</strong></td>
<td>Distinct locations at which sampling took place.</td>
</tr>
<tr>
<td><strong>Number of obs. per-site:</strong></td>
<td>Number of times (or range of such values) that individual sites were sampled.</td>
</tr>
<tr>
<td><strong>Dates</strong></td>
<td>Month and year</td>
</tr>
<tr>
<td><strong>Sampling frequency</strong></td>
<td>Continuous, daily, monthly, seasonal, etc.</td>
</tr>
<tr>
<td><strong>Variations/Modifications</strong></td>
<td>Major deviations from sampling design (e.g. locations sampled, number of samples, gear) during data collection, or changes in the design during the course of the study.</td>
</tr>
<tr>
<td><strong>Variables in the data set</strong></td>
<td>List of variables for which data exists. (e.g., if looking for salinity values, would you find them in this data set?).</td>
</tr>
</tbody>
</table>

Do not literally include all variables. For example, a trawl data set may have variables for gear code, mesh size, duration of tow, etc. It would be sufficient to report that counts of fish were made (variable) using otter trawls (equipment). The idea is to help someone decide whether this data set may be of interest to them (worth further investigation), vs. completely describe it. When in doubt, though, feel free to provide more details rather than fewer.
<table>
<thead>
<tr>
<th>Dataset Title</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Date(s)</td>
<td></td>
</tr>
<tr>
<td>Taxonomic group(s)</td>
<td></td>
</tr>
<tr>
<td>Keywords</td>
<td></td>
</tr>
<tr>
<td>Publicatons/Reports</td>
<td></td>
</tr>
<tr>
<td>Data collected by</td>
<td></td>
</tr>
<tr>
<td>Name:</td>
<td>Address:</td>
</tr>
<tr>
<td>Phone Number:</td>
<td>Sponsor(s)</td>
</tr>
<tr>
<td>Data available from:</td>
<td></td>
</tr>
<tr>
<td>Name:</td>
<td>Address:</td>
</tr>
<tr>
<td>Phone:</td>
<td></td>
</tr>
<tr>
<td>Available as (circle all that apply):</td>
<td>Print: Raw data / Summary tables</td>
</tr>
<tr>
<td>Computer files: ASCII</td>
<td>dBASE / Lotus 123 / SAS</td>
</tr>
<tr>
<td>Other:</td>
<td></td>
</tr>
<tr>
<td>Restrictions on access:</td>
<td></td>
</tr>
</tbody>
</table>
Abstract: Goals, Methods (when & where), Results (what recorded)

Spatial extent of study

<table>
<thead>
<tr>
<th>Area over which samples collected: (describe and/or attach map)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method of choosing sampling sites (random vs. fixed):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of sites:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of observations per site:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

Temporal extent of study

<table>
<thead>
<tr>
<th>Beginning and ending dates:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sampling frequency (e.g. weekly, seasonal):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

Variations/Modifications in sample design during data collection: (describe separately for field and lab, if appropriate)

|                                                             |
|                                                             |

C1 - 4
### Variables in the data set

<table>
<thead>
<tr>
<th>Variable description</th>
<th>Equipment and/or Methodology used (where applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
New York-New Jersey Harbor Estuary Program Data Documentation Form

SAMPLE

Dataset Title | U.S. Fish and Wildlife Service / U.S. Corps of Engineers Survey of Fishery Resources in the Kill Van Kull and Hudson River Estuary
---|---
Study Date(s) | 1984-1985

<table>
<thead>
<tr>
<th>Taxonomic group(s)</th>
<th>Fish (all taxa collected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keywords</td>
<td>Hudson River Estuary, Kill Van Kull, fish populations, water quality</td>
</tr>
</tbody>
</table>

Data collected by

<table>
<thead>
<tr>
<th>Name</th>
<th>Marine Science Research Center of SUNY at Stony Brook, US Army Corps of Engineers, and US Fish and Wildlife Service.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td>xxxxx (Hypothetical example)</td>
</tr>
<tr>
<td>Phone Number</td>
<td>xxxxx</td>
</tr>
<tr>
<td>Sponsor(s)</td>
<td>U.S. Army Corps of Engineers, and U.S. Fish and Wildlife Service.</td>
</tr>
</tbody>
</table>

Data available from:

<table>
<thead>
<tr>
<th>Name</th>
<th>Environmental Analysis Branch, U.S. Army Corps of Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td>xxxxx (Hypothetical example)</td>
</tr>
<tr>
<td>Phone</td>
<td>xxxxx</td>
</tr>
</tbody>
</table>

Available as (circle all that apply):

- Print: Raw data / Summary tables
- Computer files: ASCII
dBASE / Lotus 123 / SAS
- Other: 

Restrictions on access: Not known
Abstract: Goals, Methods (when & where), Results (what recorded)

The purpose of the study was to determine what fish are present in Newark Bay, Kill Van Kull and the Hudson River Estuary, and (combined with existing information) to determine why they are there. The fishery resources of the study area were surveyed on a monthly basis from July 1984 to June 1985, using a variety of gear types. Water quality data were recorded concurrently with counts and lengths of fish by species.

Spatial extent of study

Area over which samples collected: (describe and/or attach map) Portions of Newark Bay, Arthur Kill, Kill Van Kull, Hudson River, East River and the Upper Bay portion of New York Harbor. Boundaries are the New Jersey Turnpike extension bridge over Newark Bay at Bayonne, NJ, the Goethals Bridge in the Arthur Kill, a Line from Fort Wadsworth to Fort Hamilton just south of the Verrazano-Narrows Bridge, the Manhattan Bridge in the East River, the Henry Hudson Parkway Bridge over the Harlem River and the Tappan Zee Bridge over the Hudson River.

Method of choosing sampling sites (random vs. fixed): not known

Total number of sites: 57

Number of observations per site: 9-12

Temporal extent of study

Beginning and ending dates: July 1984 - June 1985

Sampling frequency (e.g. weekly, seasonal): monthly

Variations/Modifications in sample design during data collection: (describe separately for field and lab, if appropriate)

Detailed sampling occurred from July to October 1984, and was considerably reduced thereafter. Ice restrictions and vessel down time led to some stations not being sampled as scheduled. Over 70% of the stations were not sampled at least one of the 12 sampling months.
**Variables in the data set**

<table>
<thead>
<tr>
<th>Variable description</th>
<th>Equipment and/or Methodology used (where applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salinity</td>
<td>Hydrolab Surveyor II</td>
</tr>
<tr>
<td>Conductivity</td>
<td>&quot;</td>
</tr>
<tr>
<td>pH</td>
<td>&quot;</td>
</tr>
<tr>
<td>Water temperature</td>
<td>&quot;</td>
</tr>
<tr>
<td>Sample depth</td>
<td>not known</td>
</tr>
<tr>
<td>Total count (by species)</td>
<td>Gill nets (large, small, and experimental), Fyke nets, Otter trawls, and Beach Seines</td>
</tr>
<tr>
<td>Length (of individuals)</td>
<td>? (to nearest mm)</td>
</tr>
</tbody>
</table>
C2 - Description of Study Methods and Sampling Design

1991 Aerial Survey of Sea Mammals in Tampa Bay, Methods Description from Report.

METHODOLOGY

Aerial surveys of manatees and dolphins in Tampa Bay were conducted twice monthly from November 1987 through February 1989 and once per month thereafter. No flight was made in November 1989 or November 1990. Data through December 1990 are reported here. Each survey involved two Cessna 172 aircraft flying simultaneously, one on each side of the bay. Flights were made at an altitude of 150 m and an air speed of approximately 80-85 knots. J. Reynolds or P. Kouyoulius was primary observer in the aircraft that surveyed the western half of Tampa Bay; B. Weigle, E. Beeler, or B. Ackerman was primary observer in the aircraft that covered the eastern half.

The aerial surveys concentrated on nearshore areas of the bay from the shoreline to about 1.0 km offshore (Figure 1). This area was calculated to contain 568.5 km² of water surface area. Approximately 440 km² in the middle of the bay were not surveyed because the flights were designed to maximize manatee counts (Reynolds et al. 1991) by concentrating on shallower waters where manatees and their primary food source, seagrasses, are located. Flight paths were parallel to the shoreline, periodically circling to provide observers with adequate time to spot dolphins and manatees. When dolphins or manatees were observed or suspected to be present, additional circles were made for as long as necessary to obtain an accurate count. Data recorded for each group included total number of dolphins present; number of calves present (defined as dolphins less than half the length of a closely associated animal); the animals' behavior, including feeding, resting, or direction of travel; and their precise location.

Sightings of dolphins were recorded on photocopies of NOAA navigation charts (1:40,000 scale). Other data including time of day, air temperature, water clarity, and sea-surface conditions were also recorded on the maps. Following each flight, data were entered into a Geographic Information System (GIS) at the Florida Marine Research Institute (Florida Department of Natural Resources). Observations were categorized into one of 21 geographic zones of the bay.

Statistical tests of counts between seasonal time periods were made using analysis of variance (ANOVA) procedures, using Duncan's multiple range test for multiple comparison of means (SAS 1985). Dolphin density was calculated for each zone, by season and for all surveys, by normalization of the counts using the number of aerial surveys and the area of the zone. The seasons were defined as follows: winter—December through February; spring—March through May; summer—June through August; fall—September through November.
C3 - Description of Dataset Contents


II.B. DATA DICTIONARY: Template

A. Entity Template

1. Label
2. Entity Authority
3. Definition
4. Quantity of Data

B. Attribute Template

1. Label
2. Attribute Authority
3. Definition
   a. Description
   b. Measurement/Determination
4. Domain value
   a. Value format
      1. Domain
         a. Character type
         b. Allowable values (domain enumeration)
            1. Length
            2. Number of significant digits
            3. Units of measure
      b. Categorical
         1. Value
         2. Meaning
      c. Continuous
         1. Range of values
            a. Minimum
               1. value
               2. inclusive/exclusive
            b. Maximum
               1. value
               2. inclusive/exclusive
         2. Typical value
      5. Other editing information
II.C. DATA DICTIONARY: Template Field Descriptions

A. Entity Template

1. Label
2. Entity Authority
3. Definition
4. Quantity of Data

B. Attribute Template

1. Label
2. Attribute Authority
3. Definition

Aliases need to be included in the definition if they are important.

a. Description

b. Measurement/Determination

In dealing with scientific data, an area of major concern is how that data was measured or determined.

4. Domain value

a. Value format

Describes the format that the attribute value can take.

1. Domain

a. Character type

There are six major specifications (see type in glossary).
b. Allowable values (domain enumeration)

1. Length
2. Number of significant digits
3. Units of measure

b. Categorical

1. Value
2. Meaning

c. Continuous

1. Range of values
   a. Minimum
      1. value
      2. inclusive/exclusive
   b. Maximum
      1. value
      2. inclusive/exclusive

2. Typical value

   *Give some indication as to what a typical value would be. This may be described as a mean, median or mode, if appropriate. It is not necessary to calculate these values. The purpose is to provide a "general understanding of what is to be expected." Textual description is also appropriate with support for the derived number.*

5. Other editing information

   *This would include programmatic edits from the source of data entry. Examples of edits would be upper or lower case, values = A through G, values less than 0, etc.*

   *If editing features such as date fields, dollar marks, etc. are included with the data, this information should be included here.*
II.D. DATA DICTIONARY: Definitions and Use of Terms

aliases:

Other words for the same variable. These normally are not relevant to the transfer of data but if so, they should be included in the definition.

attribute:

A defined characteristic of an entity, for example, composition is a possible attribute for a bridge.

attribute authority:

The organization and/or document through which a meaning is assigned to the attribute label.

attribute value:

A specific quality or quantity assigned to an attribute (where entity is "bridge" and attribute is "composition," an attribute value might be "steel").

authority:

The organization and/or document through which a meaning is assigned to the entity label.

bitfield (unsigned binary, per agreement):

A sequence of on or off states to be represented by bitfield data--unsigned.

character mode bitfield:

A sequence of on or off states to be represented by bitfield data using the binary characters "0" and "1".
categorical values:

Data elements which only take up certain values, i.e., a department number which can take on the values 06, 20 and 33, but no other values.

continuous:

Data elements, which for all practical purposes, can take any value within a range, i.e., a dollar amount from zero to $999,999,999.99 to the nearest cent.

domain:

The set in which a variable is expressed, i.e., alpha, alphanumeric, graphic character, integer, etc.

entity:

A real world phenomenon that is not subdivided into phenomena of the same kind (i.e., a bridge).

entity authority:

The identification of the organization and/or document through which meaning is assigned to an entity label.

exclusive:

The value is not included (if a lower limit of a range is 2.0, and it is exclusive, 2.0 is not a member of the range).

inclusive:

The value is included (if a lower limit of a range is 2.0, and it is inclusive, 2.0 is a member of the range and the smallest value of that range).

integer:

A positive, negative, or unsigned whole number.
label:

A descriptive or identifying word.

length:

The maximum number of digits a number can have. This is field-specific information.

number of decimal places:

The number of places allowed to the right of the decimal place. A statement about the accuracy (significant digits) of the number of decimal places should be included.

number of significant digits:

The number of decimal places that are meaningful. For example, in dealing with dollars and cents there are two significant digits. If you have a value such as $1.53 multiplied by .18, you will have an answer of .1754, but the answer will only be valid (and sensible) to the second decimal. Thus the correct answer, rounding to the nearest 100th, is .18.

real:

A positive or negative number with a fraction. A rational or irrational number.

template:

An outline to be followed when recording information.
type:

A data type indicates the manner in which the field or subfield will be encoded. This is relevant to the data transfer and not to a data dictionary.

A Graphics characters, alphanumeric characters, or alphabetic characters
I Implicit-point (integer)
R Explicit-point unscaled (real)
S Explicit-point scaled (real with exponent)
B Bitfield data (unsigned binary, per agreement)
C Character mode bitfield (binary in zero and one characters)

units of measure:

Identifies what measurement was used for a value, i.e. dollars, francs, feet, inches, meters, pounds, kilograms, etc.

value:

The number or code stored.
C4 - Record of Dataset Quality Review


III.A. QUALITY AND ACCURACY REPORT: Procedures for Completing

The Templates for the Quality and Accuracy Report are divided into two sections, vector and raster. The appropriate template is to be filled out according to the type of data.

The templates are simply an outline to be followed and filled out in text format. The following rules apply:

- Each outline should be addressed and filled out with the appropriate information, notation that it is not applicable or that the information is unknown.
- Outline levels should never be changed or deleted.
- The Quality and Accuracy report should be made available in an ASCII format.

The template outlines follow. For additional information in filling out the individual sections, see the annotated templates beginning on page 26.
III.B. QUALITY AND ACCURACY REPORT: Truth In Labelling

The purpose of the template for the Quality and Accuracy Report is to provide as much information as possible. The intent is to follow the "truth-in-labelling" practices proposed by the workgroup that developed the Spatial Data Transfer Specifications proposed by the U.S. Geological Survey. The following quotes come from An Interim Proposed Standard for Digital Cartographic Data Quality; Supporting Documentation by N. Chrisman.

We find "quality" to be a wide-ranging concern which can cover any issue affecting the use of cartographic data. The potential uses of digital cartographic data are so diverse that a fixed set of numerical thresholds could not adjust to the potential uses. In more circumscribed application areas (for example, a multipurpose cadastre or a forest inventory), a set of thresholds might be fruitful. Because these standards must serve the whole profession, we foresee a truth-in-labelling standard instead. The idea is to communicate actual numerical properties of the data in a way that potential users can make their own informed decisions on fitness.

The truth-in-labelling concept may seem less rigorous in that it blesses the status quo. Any imprecise, inaccurate database could meet the standard in the formal sense by proclaiming those imprecisions and inaccuracies. These standards place a substantial responsibility on the user to evaluate the quality report to ensure fitness for the particular application.

It is with this thought in mind, that the templates have been created. Formalizing the structure, but still allowing the basis to be textual in content, the report is better suited for user evaluation of the data.
III.C. QUALITY AND ACCURACY REPORT: Template - Vector Data

A. Lineage

1. Description of source material(s)
   a. Name
   b. Scale (specify ratio)
   c. Datum
   d. Map Projection check list
      ___ No projection
      ___ Lambert Conformal Conic (ie. State Plane-Florida North)
      ___ Transverse Mercator (ie. State Plane-Florida East/West)
      ___ Albers Equal-Area Conic
      ___ Polyconic
      ___ Equidistance Cylindrical
      ___ Miller Oblated Stereographic
      ___ Stereographic (nonpolar/polar)
      ___ Regulator Mercator
      ___ Modified Transverse Mercator
      ___ Bipolar Oblique Conic Conformal
      ___ Other (name/explain)
   e. Media of source
   f. Condition of Media
   g. Creator organization/individual

      Name _________________________________
      Address ________________________________
      Phone (_ _) ____________________

h. Date of source material

   1. Time interval covered
   2. Update schedule
2. Derivation methods for data
   a. Method of derivation
      1. preautomation compilation
      2. digitizing/scanning/transformations
      3. equipment
         a. model
         b. resolution
         c. tolerance of digitizer - (if digitized)
   b. Date of automation
      1. Initial date
      2. Update schedule
   c. Control Points
      — No projection
      — State Plane-Florida North (Lambert Conformal Conic)
      — State Plane-Florida East/West (Transverse Mercator)
      — Other (name)
   d. Explanation of procedures used to digitize/scan/transform etc. the data.
      1. Name of transformation methodology
      2. Description of algorithm
      3. Mathematics used in the transformation
      4. Set of sample computations
   e. Software system and version used.

B. Positional Accuracy
   1. Line Work Replication Check
      a. positional
      b. completeness
2. Absolute measure of error reference in units of the coordinate system.
   a. Value
   b. Method of derivation of that number

3. Test Reports (use one or more of the following)
   a. Deductive estimate
      1. date of tests
      2. report
   b. Internal Evidence (geodesy)
      1. date of tests
      2. report
   c. Comparison to Source (overlay)
      1. date of tests
      2. report
   d. Independent source of higher accuracy
      1. date of tests
      2. report

C. Attribute Accuracy

1. Continuous (define what documents can be referenced here)
   a. Absolute measure of error reference in units of the measurement system.
      1. Value
      2. Method of derivation of that number
   b. Test Reports (use one or more of the following)
      1. Deductive estimate
         a. date of tests
         b. report
2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical
   a. Test Reports (if appropriate use one of the following)
      1. Deductive Estimate
         a. date
            1. date of tests
            2. date of materials used
            3. rates of change
         b. report
            1. basis for the deduction
            2. deduction

2. Independent Samples
   a. date
      1. date of tests
      2. date of materials used
      3. rates of change
   b. map scale
   c. dates of materials used
   d. report (misclassification matrix)
1. sampling procedures  
2. sampling locations  
3. results  

3. Tests based on Polygon Overlay  
   a. date  
      1. date of tests  
      2. date of materials used  
      3. rates of change  
   b. map scale  
   c. dates of materials used  
   d. report  
      1. misclassification matrix  
      2. explain relationship between maps  

D. Logical consistency  

1. Tests of valid values  
   a. test performed  
   b. date  
   c. results  

2. Cartographic Tests  

Does the provider feel the map is cartographically clean?  

_____ Yes  
_____ No  

a. Answer the following  

1. Do lines intersect only where intended?  
   (yes/no/unknown)  
2. Were duplicate lines eliminated? (yes/no/unknown)  
3. Are all areas completely described? (yes/no/unknown)  
4. Have overshoots and undershoots been eliminated?  
   (yes/no/unknown)  
5. Have all slivers been eliminated? (yes/no/unknown)
6. Does documentation in the lineage section (d-5) describe why the data is cartographically clean? (yes/no/unknown)

b. Tests

1. tests performed
2. date
3. reference to documentation

Conclusions:

3. Topological Tests

a. test performed
b. date
c. software

1. name
2. version

d. results

1. Test for polygon coverage

   a. How many polygons are represented on the digital map product?
   b. Has a polygon closure been verified?
   c. Are Polygon-Ids assigned to each polygon on the digital map?

   1. Do polygons have more than one Polygon-Id? (Yes/No)
   2. Are the Polygons-Ids unique: (Yes/No)

2. Test for line coverage

   a. How many lines are represented on the digital map product?
   b. Do the line segments have unique line segment values? (Yes/No)
   c. Is the digital map topologically clean? (Yes/No)
3. Test for point coverage
   a. How many points are represented on the digital map product?
   b. Are the Point-Ids unique: (Yes/No)

E. Completeness of Source Materials

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
III. D. QUALITY AND ACCURACY REPORT: Template - Raster Data

* - Represents differences from the Vector Data Quality and Accuracy Report

A. Lineage

1. Description of source material

a. Name
b. Scale (specify ratio)
c. Datum
d. Map Projection check list*

___ No projection
___ Lambert Conformal Conic (ie. State Plane-Florida North)
___ Transverse Mercator (ie. State Plane-Florida East/West)
___ (UTM) Universal Transverse Mercator
___ Lambert Azimuthal Equal-Area (sphere)
___ Albers Equal-Area Conic
___ Polyconic
___ Equidistance Cylindrical
___ Miller Oblated Stereographic
___ Stereographic (nonpolar/polar)
___ Regulator Mercator
___ Modified Transverse Mercator
___ Bipolar Oblique Conic Conformal
___ Other (name/explain)

e. Media of source*

1. media
2. Type of imagery

___ Landsat MSS (Multispectral Scanner)
___ Landsat TM (Thematic Mapper)
___ SPOT Panochromatic
___ SPOT Multispectral
___ Other satellite imagery
___ Aerial Photography
___ Other
f. Condition of Media

g. Creator organization/individual

Address of creating entity
Name ________________________________
Address ________________________________
Phone (___) ______

h. Date of source material

1. Scene number(s) and dates of source material
2. Update schedule

i. Ancillary Source materials
j. Imported data

2. Derivation methods for data

a. Method of derivation

1. digitizing/scanning/transformations etc.
2. equipment

   a. model
   b. resolution
   c. tolerance of digitizer - if digitized

b. Date of automation

1. Initial date
2. Update schedule

b. Control Points

1. Coordinate reference system check list*

   ___ No projection
   ___ State Plane-Florida North (Lambert Conformal Conic)
   ___ State Plane-Florida East/West (Transverse Mercator)
   ___ Other (name)
2. List of coordinate pairs

d. Explanation of procedures used to digitize/scan/transform etc.

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used

B. Positional Accuracy

1. Absolute measure of error reference in units of the coordinate system.
   a. Value
   b. Method of derivation of the number

2. Test Reports *
   a. Deductive estimate
      1. date of tests
      2. report

C. Attribute Accuracy

1. Continuous
   a. Absolute measure of error reference in units of the coordinate system.
      1. Value
      2. Method of derivation of that number

   b. Test Reports (use one or more of the following)
      1. Deductive estimate
         a. date of tests
         b. report

23
2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical
   a. Test

   1. Deductive Estimate
      a. date
         1. date of tests
         2. date of materials used
         3. rates of change
      b. report
         1. basis for the deduction
         2. deduction

   2. Ground Truthing*
      a. Were ground truthing tests performed (yes/no)
      b. date range of tests
      c. How many points were used
      d. Description of the test
D. Logical consistency *

1. Tests of valid values
   a. test performed
   b. date
   c. results

E. Completeness of Source Materials

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
III.E. QUALITY AND ACCURACY REPORT: Annotated Template - Vector Data

A. LINEAGE

1. Description of source material(s)

   a. Name

   b. Scale (specify ratio)

      scale: ratio between the distance on a map, chart, photograph or image and the corresponding distance on the surface of the Earth.

   c. Datum

      Geodetic Datum: A set of constants specifying the coordinate system used for geodetic control, i.e., for calculating coordinates of points on the earth.

   d. Map Projection check list

      map projection: systematic drawing of lines of a plane surface to represent the parallels of latitude and the meridians of longitude of the Earth.

      __ No projection
      __ Lambert Conformal Conic (ie. State Plane-Florida North)
      __ Transverse Mercator (ie. State Plane-Florida East/West)
      __ Albers Equal-Area Conic
      __ Polyconic
      __ Equidistance Cylindrical
      __ Miller Oblated Stereographic
      __ Stereographic (nonpolar/polar)
      __ Regulator Mercator
      __ Modified Transverse Mercator
      __ Bipolar Oblique Conic Conformal
      __ Other (name/explain)
Annotated Template - Vector Data  
Page Two

(A.1.e.)

e. Media of source

source: origins of the data.

media: physical substance the data was taken from (mylar, paper, electronic tape, etc.).

f. Condition of Media

A qualitative statement of the condition of the media. Statements such as "good" or "bad" must be given a basis.

g. Creator organization/individual

Address of creating entity
Name __________________________________________
Address ______________________________________
Phone (___) ________

h. Date of source material

A history of the development of the source material. (There may be multiples of this entry.)

1. Time interval covered

A history of the development of the source material.

2. Update schedule

Schedule on which the data is being updated. If updating is continuous some indication of the turnover should be given.
2. Derivation methods for data

The purpose of this step is to describe how the data was brought into the system. By knowing how the data was created and the technology used, limits on the accuracy may be deduced.

a. Method of derivation

1. preautomation compilation
2. digitizing/scanning/transformations etc.
3. equipment
   a. model
   b. resolution
   c. tolerance of digitizer - (if digitized)

b. Date of automation

1. Initial date
2. Update schedule

c. Control Points

control point: any station in a horizontal or vertical control network that is identified in a data set or photograph and used for correlating the data shown in that data set or photograph.

coordinate system: a particular kind of reference frame or system, such as plane rectangular coordinates or spherical coordinates, that uses linear or angular quantities to designate the position of points within that particular reference frame or system.
Annotated Template - Vector Data
Page Four

(A.2.c.1)

1. Coordinate reference system check list
   __ No projection
   __ State Plane-Florida North (Lambert Conformal Conic)
   __ State Plane-Florida East/West (Transverse Mercator)
   __ Other (name)

2. List of control point coordinate pairs

d. Explanation of procedures used to
digitize/scan/transform etc. the data

This is a description of procedures that would indicate the
quality/accuracy of the data captured. Information that would not
provide insights should not be included. Transformation routines
that are supplied by vendors should include the name of the
transformation module. User created transformations should
include the following:

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used

B. POSITIONAL ACCURACY

Tests of accuracy after all transformations have been performed on a
particular layer.

1. Line Work Replication Check
   a. Positional
   b. Completeness
2. Absolute measure of error reference in units of the coordinate system.

(provides a numerical estimate of expected discrepancies)

"quality of the final product after all transformations."

a. Value
b. Method of derivation of that number

3. Test Reports (use one or more of the following)

a. Deductive estimate
deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

1. date of tests
2. report

b. Internal Evidence (geodesy)

1. date of tests
2. report

c. Comparison to Source (overlay)

1. date of tests
2. report

d. Independent source of higher accuracy

1. date of tests
2. report
C. ATTRIBUTE ACCURACY

Accuracy assessment for measures on a continuous scale shall be performed using procedures similar to those used for positional accuracy (providing a numerical estimate of expected discrepancies).

There has been considerable discussion on how much detail is required at this step. It is the view of the developers of the report that as much information be provided as possible. This does not mean that a test must be performed that wouldn't normally be performed, but it does mean that all tests that were performed should be reported. The level of reporting should be at such a level as to be useful to the recipient. The exact format is up to the discretion of the developer.

1. Continuous or numerical (See Data Dictionary Template)
   a. Absolute measure of error referenced in units of the measurement system.

   (provides a numerical estimate of expected discrepancies)
   "quality of the final product after all transformations."

   1. Value
   2. Method of derivation of that number

   b. Test Reports (use one or more of the following)

   1. Deductive estimate

   deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

   a. date of tests
   b. report
Annotated Template - Vector Data
Page Seven

(C.1.b.2.)

2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical (See Data Dictionary Template)

The intent is for a statistical sampling of an attribute that would
give a level of assurance as to the accuracy of the data. This was
not intended for a item by item data check.

   a. Tests (if appropriate use one of the following)

   1. Deductive Estimate
      a. date
         1. date of tests
         2. date of materials used
         3. rates of change

   In the case of discrepancies between the date of the tests and the
dates of the materials used, the rates of change in the phenomena
classified will be reported. It is up to the creator to identify
dates as being discrepant. The intent is to make note of temporal
items.
(C.2.a.1.b.)

b. report

1. basis for the deduction
2. deduction

2. Independent Samples

a. date
   1. date of tests
   2. date of materials used
   3. rates of change

In the case of discrepancies between the dates of the tests and the dates of the materials used the rates of change in the phenomena classified will be reported.

b. map scale

c. dates of materials used

d. report (misclassification matrix)

1. sampling procedures
2. sampling locations
3. results
3. Tests based on Polygon Overlay

These are for polygons that represent a category, such as wetlands, clay soils, etc.

a. date
   1. date of tests
   2. date of materials used
   3. rates of change

   In the case of discrepancies between the dates of the tests and the dates of the materials used, the rates of change in the phenomena classified will be reported.

b. map scale

b. dates of materials used
d. report
   1. misclassification matrix
   2. explain relationship between maps

D. LOGICAL CONSISTENCY

1. Tests of valid values

   Test for attribute data.

   a. test performed
   b. date
   c. results

2. Cartographic Tests

   Does the provider feel the map is cartographically clean?

   Yes
   No
(D.2.a.)

a. Answer the following

1. Do lines intersect only where intended? (yes/no/unknown)
2. Were duplicate lines eliminated? (yes/no/unknown)
3. Are all areas completely described? (yes/no/unknown)
4. Have overshoots and undershoots been eliminated? (yes/no/unknown)
5. Have all slivers been eliminated? (yes/no/unknown)
6. Does documentation in the lineage section (d-5) describe why the data is cartographically clean? (yes/no/unknown)

b. Tests

1. tests performed
2. date
3. reference to documentation

Conclusions:

3. Topological Tests

topology: a branch of geometrical mathematics which is concerned with order, contiguity, and relative position, rather than actual linear dimensions.

topologic error checking: process of ensuring the logical consistency of the data is intact; all polygons are closed, all arcs are connected to nodes, etc.

a. test performed
b. date
c. software

1. name
2. version
(D.3.d.)

d. results

1. Test for polygon coverage
   a. How many polygons are represented on the digital map product?
   b. Has a polygon closure been verified?
   c. Are Polygon-Ids assigned to each polygon on the digital map?
      1. Do polygons have more than one Polygon-Id? (Yes/No)
      2. Are the Polygon-Ids unique? (Yes/No)

2. Test for Line coverage
   a. How many lines are represented on the digital map product?
   b. Do the line segments have unique line segment values? (Yes/No)
   c. Is the digital map topologically clean? (Yes/No)

3. Test for point coverage
   a. How many points are represented on the digital map product?
   b. Are the Point-Ids unique? (Yes/No)
E. COMPLETENESS OF SOURCE MATERIALS

The purpose of the completeness section is to describe the set of information collected in comparison to a larger set. For example, a set called "Well Data Points" may be all manmade wells in the area described or it may be only private wells used for homes.

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
III.E. QUALITY AND ACCURACY REPORT: Annotated Template -
Raster Data

A. LINEAGE

1. Description of source material

The purpose is to describe the origins of the data from which the "map" was created allowing the receiver to evaluate the utility of
the data.

a. Name
b. Scale (specify ratio)

c. Datum

Geodetic Datum: A set of constants specifying the coordinate
system used for geodetic control, i.e., for calculating
coordinates of points on the earth.

d. Map Projection check list*

map projection: systematic drawing of lines of a plane surface to
represent the parallels of latitude and the meridians of longitude
of the Earth.

__ No projection
__ Lambert Conformal Conic (i.e. State Plane-Florida
  North)
__ Transverse Mercator (i.e. State Plane-Florida
  East/West)
__ (UTM) Universal Transverse Mercator
__ Lambert Azimuthal Equal-Area (sphere)
__ Albers Equal-Area Conic
__ Polyconic
__ Equidistance Cylindrical
__ Miller Oblated Stereographic
__ Stereographic (nonpolar/polar)
Annotated Template - Raster Data
Page Two

(A.1.d.)

___ Regulator Mercator
___ Modified Transverse Mercator
___ Bipolar Oblique Conic Conformal
___ Other (name/explain)

e. Media of source*

1. media

    source: origins of the data

    Media that the data was on (mylar, paper, electronic tape, etc.).

2. Type of imagery

___ Landsat MSS (Multispectral Scanner)
___ Landsat TM (Thematic Mapper)
___ SPOT Panochromatic
___ SPOT Multispectral
___ Other satellite imagery
___ Aerial Photography
___ Other

f. Condition of Media

A qualitative statement of the condition of the media. Statements such as "good", or "bad" must be substantiated.

g. Creator organization/individual

Address of creating entity
Name __________________________________________
Address _______________________________________
Phone (___) ______
Annotated Template - Raster Data
Page Three

(A.1.h.)

h. Date of source material

A history of the development of the source material.
There may be multiples of this entry.

1. Scene numbers(s) and dates of source material*

Include scene numbers if appropriate.
There may be multiples of this entry.

2. Update schedule

Schedule on which the data is being updated. If updating is
continuous some indication of the turnover should be provided.

i. Ancillary Source materials

ancillary data: auxiliary or supplementary data.

j. Imported data

What was the format and what was the software.

If graphic data was not created in the system that it is being
used in, what structure was it imported in. For example was it a
DLG III format, SDTS or was it a vendor created importation
package.

2. Derivation methods for data

The purpose is to describe how the data was brought into the system.
By knowing how the data was created and the technology used, limits
on the accuracy may be deduced.
(A.2.a.)

a. Method of derivation

1. digitizing/scanning/transformations etc.
2. equipment
   a. model
   b. resolution
   c. tolerance of digitizer - if digitized

b. Date of automation

1. Initial date
2. Update schedule

c. Control Points

control point - any station in a horizontal or vertical control network that is identified in a data set or photograph and used for correlating the data shown in that data set or photograph.

coordinate systems - a particular kind of reference frame or system, such as plane rectangular coordinates or spherical coordinates, that use linear or angular quantities to designate the position of points within that particular reference frame or system.

1. Coordinate reference system check list*

   - No projection
   - State Plane-Florida North (Lambert Conformal Conic)
   - State Plane-Florida East/West (Transverse Mercator)
   - Other (name)

2. List of coordinate pairs
(A.2.d.)

d. Explanation of procedures used to digitize/scan/transform etc. the data

This is a description of procedures that would indicate the quality/accuracy of the data captured. Information that would not provide insights should not be included. This is particularly important for transformations that did not use canned software packages. If there were transformations of this nature, the following should be included:

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used

B. POSITIONAL ACCURACY

Tests of accuracy after all transformations have been performed on a particular layer.

1. Absolute measure of error reference in units of the coordinate system.

   (provides a numerical estimate of expected discrepancies)
   "quality of the final product after all transformations."

   a. Value.
   b. Method of derivation of that number

2. Test Reports (use one or more of the following)*
(B.2.a.)

a. Deductive estimate

deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

1. date of tests
2. report

C. ATTRIBUTE ACCURACY

Accuracy assessment for measures on a continuous scale shall be performed using procedures similar to those used for positional accuracy (providing a numerical estimate of expected discrepancies).

1. Continuous

a. Absolute measure of error reference in units of the measurement system.

(provides a numerical estimate of expected discrepancies) "quality of the final product after all transformations."

1. Value
2. Method of derivation of that number

b. Test Reports (use one or more of the following)

1. Deductive estimate

deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

a. date of tests
b. report
Annotated Template - Raster Data
Page Seven

(C.1.b.2.)

2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical

The intent is for a statistical sampling of an attribute that would give a level of assurance as to the accuracy of the data. This was not intended for a item by item data check.

   a. Test (if appropriate use one of the following)

1. Deductive Estimate
   a. date
      1. date of tests
      2. date of materials used
      3. rates of change

In the case of discrepancies between the date of the tests and the dates of the materials used the rates of change in the phenomena classified will be reported. It is up to the creator to identify dates as being discrepant. The intent is to make note of temporal items.
Annotated Template - Raster Data
Page Eight

(C.2.a.1.b.)

b. map scale

c. report
   1. basis for the deduction
   2. deduction

2. Ground Truthing*

   a. Were ground truthing tests performed?
      (yes/no)
   b. date range of tests
   c. How many points were used
   d. Description of the test

D. LOGICAL CONSISTENCY *

1. Tests of valid values

   Test for attribute data.

   a. test performed
   b. date
   c. results

E. COMPLETENESS OF SOURCE MATERIALS

The purpose of the completeness section is to describe the set of information collected in comparison to a larger set. For example a set called "Well Data Points" may be all manmade wells in the area described or it may be only private wells used for homes.

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
C5 - Transaction Log

Modified From Maryland Power Plant Research Program, Data Tracking Form.

DATA TRACKING INFORMATION FORM

PROJECT CODE: ___________________________

CHARGE NUMBER: ________________________

PRINCIPAL INVESTIGATOR: ________________________________

TITLE: _____________________________________________

DATA COLLECTED BY: ________________________________

DATA START DATE: ________________________________

DATA END DATE: ________________________________

STUDY LOCATION: __________________________________

REPORT REFERENCE: __________________________________

PUBLICATION DATE: ________________________________

Please give a brief description of what is to be done with this data. Include a due date.

<table>
<thead>
<tr>
<th>DATE RECEIVED</th>
<th>BY:</th>
<th>PI:</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATE VERIFIED</td>
<td>BY:</td>
<td>PI:</td>
</tr>
<tr>
<td>DATE PUT UP</td>
<td>BY:</td>
<td>PI:</td>
</tr>
<tr>
<td>DATE VALIDATED</td>
<td>BY:</td>
<td>PI:</td>
</tr>
<tr>
<td>(ready in data base)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATE(S) REVISED</td>
<td>BY:</td>
<td>PI:</td>
</tr>
<tr>
<td></td>
<td>BY:</td>
<td>PI:</td>
</tr>
</tbody>
</table>

SERIAL NUMBER: _______________________________________

KEYWORDS: ____________________________________________

STORAGE LOCATION: ___________________________________

VOLSER: _____________________________________________

C5 - 1
C6 - Summary of Data Transfer Policies and Procedures

Southwest Florida Water Management District, GIS Data Distribution Procedures.

GIS SYSTEM DOCUMENTATION

GEOGRAPHIC INFORMATION SYSTEM DATA DISTRIBUTION PROCEDURES

Mapping and GIS Section
Southwest Florida Water Management District
1. INTRODUCTION

The District has developed a geographic information system (GIS) database to support many of its planning, environmental and regulatory activities. This database includes a considerable amount of information that is potentially of value to federal, state, regional and local governmental agencies as well as to private businesses. The purpose of this document is to describe the procedures for distributing GIS data to interested public and private individuals or groups.

2. DATA AVAILABILITY

Data available for distribution include those for which all quality control procedures have been completed and that are loaded into the District's ARC/INFO Map Library. Currently available data layers are listed in APPENDIX A. Detailed information about specific layers may be obtained by contacting the District's Mapping and GIS Section.

3. DISTRIBUTION FORMAT

Data will be distributed in ARC/INFO EXPORT format on nine track computer compatible magnetic tapes. EXPORT coverages will be available only in non-compressed ASCII format as described in the ARC/INFO Users Guide. Data densities of 1600 or 6250 Bits Per Inch (BPI) may be specified. Translation of data to other digital or hardcopy formats is the sole responsibility of the purchaser of the data.

Data are only available in the spatial format that they are stored in the District GIS database. These formats are a function of data density and correspond to the United States Geological Survey's (USGS) 1:24,000 and 1:100,000 quadrangles or in a single District-wide coverage. Any subsetting of data is the sole responsibility of the recipient.

4. PRIORITIZATION OF DATA DELIVERIES

The District's highest priority for the use of its GIS is the support of in-house activities. Though the District will attempt to meet requests for data in a timely manner, requests by non-District staff for data will be completed only after all in-house GIS requirements are met.

5. DISCLAIMER OF WARRANTY

These data were developed to assist the District in its geographic modelling efforts. The data are being provided on an 'as is' basis. The District specifically disclaims any warranty, either expressed or implied, including, but not limited to, the implied warranties or merchantability and fitness for a particular use. The entire risk as to quality and performance is with the user. In no event will the District or its staff be liable for any direct, indirect, incidental, special, consequential, or other damages, including loss of profit, arising out of the use of this data even if the District has been advised of the possibility of such damages.

GEOGRAPHIC INFORMATION SYSTEM DATA DISTRIBUTION PROCEDURES (REV. 3: 11/91) 2
B. ORDERING PROCEDURES/CHARGES

All requests for data must be made in writing. Data are to be ordered using the following procedures.

A. Complete the attached Geographic Information System Data Order Form.

B. On the attached Data Layer List (Appendix A), check off the requested data layers. Specific information on individual layers can be obtained by contacting the Mapping and GIS Section at (904) 756-7211, Ext. 4278.

C. On the attached Quad List (Appendix B), check off the USGS 7.5 minute quadrangles for which data are requested. Data orders must be referenced to the USGS quadrangles. Orders referencing only county or municipal boundaries will not be processed.

D. All prices listed below reflect current District operating expenses and will be adjusted quarterly as necessary. These charges represent the cost to the District for the distribution of the digital files. Charges are calculated as follows:

- Tape Charge: $12.00/tape
- Computer Charge: $17.45/hour (Minimum charge of 1 hour)
- Personnel: $13.00/hour (Minimum charge of 1 hour)
- Documentation: $0.15/page
- Shipping: Varies

Tape charges represent the cost of nine-track magnetic tapes. Computer charges are based on connect time and run times required to process an order. Personnel time represents operator time for processing data. Shipping charges will vary depending on the method of shipment requested.

Upon receipt of the GIS Data Order Form, the requestor will be informed as to the availability of the requested data and cost estimates will be given at the time of the order. Based upon the cost estimate, the requestor must affirm that he understands and will pay the amount of estimated cost prior to the beginning of data processing work by the District.

Final data charges may vary from the estimated amount because of the difficulties involved in pre-determining run times prior to creating EXPORT files. Upon creation of the computer tapes, the requestor will be notified of the final costs. Data will be delivered upon receipt of payment.

E. The District reserves the right to waive all or part of the cost associated with the distribution of GIS data where reciprocal exchange-of-information cooperative agreements exist with other governmental, educational or non-profit organizations.
C7 - Data Request Form

Southwest Florida Water Management District, Geographic Information System Data Order Form.

GEOGRAPHIC INFORMATION SYSTEM DATA ORDER FORM

MAPPING AND GIS SECTION
SOUTHWEST FLORIDA WATER MANAGEMENT DISTRICT
2379 BROAD STREET
BROOKSVILLE, FLORIDA 34609-6899
(904) 796-7211 ext. 4278

PLEASE PRINT OR TYPE

NAME: _____________________________ PHONE: _______________________

ORGANIZATION NAME: ____________________________________________

ORGANIZATION TYPE (check one):

_ Federal  _ State  _ County  _ Local  _ Private

_ Educational  Other: _____________________________

ADDRESS:

______________________________________________________________

______________________________________________________________

DATE OF REQUEST: ____________ DATA SEARCH ONLY* ____________ DATA REQUEST

DATA USE (Optional): _________________________________________

MEDIA FORMAT:

Tape Density: __ 1600 bpi  __ 6250 bpi

FOR DISTRICT USE ONLY - DO NOT WRITE BELOW THIS LINE

==================================================================================================

CHARGES

DATE ORDER RECEIVED: ____________ LABOR ____________

ESTIMATED CHARGE: ____________ TAPE(S) ____________

REQUESTOR APPROVAL OF
ESTIMATED CHARGE: ____________ COMPUTER ____________

ORDER FILLED BY: ____________ SHIPPING ____________

________________________________

DOCUMENTATION ____________

________________________________

AMOUNT DUE ____________

*WRITTEN REQUEST REQUIRED
APPENDIX A - DATA LAYER LIST

The following layers are currently available for distribution. The availability of these data for specific areas may vary depending on the current status of District's data collection program.

- **District-owned Lands** - Boundaries of all District-owned lands.
- **Drainage Basin Boundaries** - 1:24,000 drainage basin boundaries as delineated by the USGS Water Resources Division, Tallahassee Office.
- **DRASIC** - DRASIC model and component layers for each aquifer.
- **Land use/cover** - Land use/cover mapped from color infrared photography flown between December 1989 and December 1990.
- **Political Boundaries** - County and Basin Board Boundaries.
- **Public Land Survey** - Section, Township, Ranges from 1:24,000 USGS maps.
- **Seagrasses** - Seagrass maps for Tampa Bay.
- **Soils, Detailed** - USDA/SCS detailed soils maps from county soils atlases (Hernando, Pasco, Sarasota, Hardee, Desoto, Citrus and Polk Counties).
- **TOPO5** - Five foot contour data from USGS quadrangles.
- **TOPO2** - Two foot data from aerial mapping program.
- **Regulatory** - Permit points and boundaries.
- **Recharge** - Recharge to the Floridan aquifer from the USGS.
- **1:24,000 Roads** - Transportation network digitized from USGS 1:24,000 quads.
APPENDIX B - USGS Quadrangles within the Southwest Florida Water Management District

Please indicate with a check the USGS 1:24,000 quadrangles for which you are requesting data. If you require all 1:24,000 quads within a 1:100,000 quadrangle, please check the 1:100,000 quad name (in boldface).

---

**GAINESVILLE:**
- ARCHER

**OCALA:**
- BRONSON
- BRONSON NE
- BRONSON SW
- BRONSON SE
- WILLISTON
- FLEMINGTON
- McINTOSH
- REDDICK
- TIDEWATER
- WACCASASSA BAY
- ROMEO
- LEBANON STATION
- LEVANT
- WITHLACOOCHEE BAY
- COTTON PLANT
- OCALA WEST
- DUNELLON SE
- SHADY

**INVERNESS:**
- RED LEVEL
- RED LEVEL
- CRYSTAL RIVER
- HOLDER
- STOKES FERRY
- LAKE PANASOFFKEE NW
- OXFORD
- OZELLO WEST
- OZELLO
- HOMOSASSA
- LECANTO
- INVERNESS
- RUTLAND
- WILDWOOD
- CHASSAHOWITZKA BAY
- CHASSAHOWITZKA
- WAHOO
- BUSHNELL
- BROOKSVILLE NW
- CHADWICK
- NOBLETON
- BROOKSVILLE
- BAYPORT
- WEEKIWACHEE
- BROOKSVILLE
- BROOKSVILLE SE
- ST CATHERINE
- WEBSTER

**TARPON SPRINGS:**
- ARIPEKA
- PORT RICHEY NE
- MASARYKTOWN
- SPRING LAKE
- LACOCHEE
- CLAY SINK
- PORT RICHEY
- FIVAY
- EHREN
- DADE CITY
- BRANCHBOROUGH
- TARPON SPRINGS
- ELFERS
- ODESSA
- LUTZ
- SAN ANTONIO
- ZEPHYRHILLS
- SOCRUM
- DUNEDIN
- WESLEY CHAPEL
- CITRUS PARK
- SULPHUR SPRINGS
- THONOTOSASSA
- PLANT CITY EAST
- PLANT CITY WEST
- OLDSMAR

**ST PETERSBURG:**
- CLEARWATER
- SAFETY HARBOR
- GANDY BRIDGE
- TAMPA
- BRANDON
- DOVER
- NICHOLS
- ST PETERSBURG
- PORT TAMPA
- GIBSONTON
- RIVERVIEW
- LITHIA
- KEYSVILLE
- ST PETERSBURG BEACH
- PASS-A-GRILLE BEACH
- WIMAUMA
- COCKROACH BAY
- RUSKIN
- PORT LONESOME
- DUEETE NE
- EGGMONT KEY
- ANNA MARIA
- PALMETTO
- PARRISH
- KEENTOWN
- DUEETE
- RYE

---

C7 - 3

GEOGRAPHIC INFORMATION SYSTEM DATA DISTRIBUTION PROCEDURES (REV. 3: 11/91) 5
__SARASOTA:
  _BRADENTON BEACH_ _BRADENTON_ _LORRAINE_ _VERNA_
  _MYAKKA CITY NW_ _MYAKKA HEAD_ _SARASOTA WEST_ _VERNA_
  _SARASOTA_ _BEE RIDGE_ _OLD MYAKKA_ _MYAKKA CITY_
  _EDGEVILLE_ _BIRD KEYS_ _LAUREL_ _LOWER MYAKKA LAKE_
  _MYRDCK NW_ _MURDOCK NE_ _VENICE_ _MYAKKA RIVER_
  _MURDOCK_ _MURDOCK SE_

__CHARLOTTE HARBOR:
  _ENGLEWOOD NW_ _ENGLEWOOD_ _EL JOBEAN_ _PUNTA GORDA_
  _PLACIDA_ _PUNTA GORDA SW_ _PUNTA GORDA SE_
  _PORT BOCA GRANDE_

__ORLANDO:
  _LADY LAKE_ _LEESBURG WEST_ _CENTER HILL_
  _MASCOTTE_ _CLERMONT WEST_

__KISSIMEE:
  _BAY LAKE_ _LAKE NELLIE_ _LAKE LOUISA_ _ROCK RIDGE_
  _POYNER_ _LAKE LOUISA SW_ _INTERCESSION CITY_
  _PROVIDENCE_ _POLK CITY_ _GUN LAKE_ _DAVENPORT_
  _LAKELAND_ _AUBURNDALE_ _WINTER HAVEN_ _DUNDEE_
  _LAKE HATCHINEHA_

__BARTOW:
  _MULBERRY_ _BARTOW_ _ELOISE_ _LAKE WALES_
  _HESPERIDES_ _BRADLEY JUNCTION_ _HOMELAND_
  _ALTURAS_ _BABSON PARK_ _LAKE WOHYAKAPKA_
  _BAIRD_ _BOWLING GREEN_ _BEREAH_ _FROSTPROOF_
  _LAKE ARBUCKLE_ _LAKE ARBUCKLE NE_ _FORT GREEN_
  _WAUCHULA_ _GRIFFINS CORNER_ _AVON PARK_
  _LAKE ARBUCKLE SW_ _LAKE ARBUCKLE SE_

__ARCADIA:
  _ONA_ _ZOLFO SPRINGS_ _SWEETWATER_ _CREWSVILLE_
  _SEBRING_ _LORIDA_ _LIMESTONE_ _GARDNER_
  _CREWSVILLE SW_ _CREWSVILLE SE_ _LAKE PLACID_ _NOCATEE_
  _LAKE JUNE-IN-WINTER_ _LONG ISLAND MARSH NW_ _VENUS NW_ _CHILDs_
  _ARCADIA_ _LONG ISLAND MARSH SE_ _LONG ISLAND MARSH SW_
  _FORT OGDEN_ _ARCADIA SE_ _VENUS SW_ _VENUS_

__FORT MYERS:
  _CLEVELAND_ _BERMONT_ _TELEGRAPH SWAMP NW_ _TUCKERS CORNER_
  _TELEGRAPH SWAMP NE_ _GILCHRIST_ _TELEGRAPH SWAMP SE_
Virginia Marine Resources Commission (VMRC) Fisheries Statistics Gear Code Translation Table.

<table>
<thead>
<tr>
<th>NEMFIS Code</th>
<th>Gear Name</th>
<th>Old Wash., DC Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Trawl, Line</td>
<td>675</td>
</tr>
<tr>
<td>20</td>
<td>Hand Line, Other</td>
<td>610</td>
</tr>
<tr>
<td>21</td>
<td>Handline, Auto Jig</td>
<td>621</td>
</tr>
<tr>
<td>30</td>
<td>Harpoon, Other</td>
<td>754</td>
</tr>
<tr>
<td>31</td>
<td>Harpoon, Swordfish</td>
<td>751</td>
</tr>
<tr>
<td>32</td>
<td>Harpoon, Turtle</td>
<td>752</td>
</tr>
<tr>
<td>34</td>
<td>Spears</td>
<td>760</td>
</tr>
<tr>
<td>40</td>
<td>Longline, Setline</td>
<td>675</td>
</tr>
<tr>
<td>41</td>
<td>Longline, Drift W/ Hooks</td>
<td>678</td>
</tr>
<tr>
<td>50</td>
<td>Trawl, Otter, Bottom, Fish</td>
<td>210</td>
</tr>
<tr>
<td>51</td>
<td>Trawl, Otter, Bottom, Crab</td>
<td>205</td>
</tr>
<tr>
<td>52</td>
<td>Trawl, Otter, Bottom, Scallop</td>
<td>214</td>
</tr>
<tr>
<td>55</td>
<td>Trawl, Otter, Bottom, Lobster</td>
<td>212</td>
</tr>
<tr>
<td>56</td>
<td>Trawl, Otter, Bottom, Paired</td>
<td>235</td>
</tr>
<tr>
<td>58</td>
<td>Trawl, Otter, Bottom, Shrimp</td>
<td>215</td>
</tr>
<tr>
<td>59</td>
<td>Trawl, Otter, Bottom, Other</td>
<td>220</td>
</tr>
<tr>
<td>60</td>
<td>Troll Line, Other</td>
<td>660</td>
</tr>
<tr>
<td>61</td>
<td>Troll Line, Tuna</td>
<td>655</td>
</tr>
<tr>
<td>62</td>
<td>Trot Line with Bait</td>
<td>680</td>
</tr>
<tr>
<td>63</td>
<td>Power Troll, Salmon</td>
<td>651</td>
</tr>
<tr>
<td>64</td>
<td>Power Troll, Tuna</td>
<td>656</td>
</tr>
<tr>
<td>65</td>
<td>Power Troll, Other</td>
<td>661</td>
</tr>
<tr>
<td>70</td>
<td>Seine, Haul, Common</td>
<td>20</td>
</tr>
<tr>
<td>71</td>
<td>Seine, Long</td>
<td>30</td>
</tr>
<tr>
<td>80</td>
<td>Trap, Floating</td>
<td>295</td>
</tr>
<tr>
<td>90</td>
<td>Dip Net, Common</td>
<td>703</td>
</tr>
<tr>
<td>91</td>
<td>Dip Net, Drop Net</td>
<td>705</td>
</tr>
<tr>
<td>92</td>
<td>Dip Net, Push Net</td>
<td>725</td>
</tr>
<tr>
<td>100</td>
<td>Gill Net, Sink/Anchor, Other</td>
<td>430</td>
</tr>
<tr>
<td>101</td>
<td>Gill Net, Set/Stake, Sea Ra</td>
<td>420</td>
</tr>
<tr>
<td>102</td>
<td>Gill Net, Stake, Other</td>
<td>425</td>
</tr>
<tr>
<td>110</td>
<td>Gill Net, Drift, Other</td>
<td>470</td>
</tr>
<tr>
<td>111</td>
<td>Gill Net, Drift, Sea Bass</td>
<td>460</td>
</tr>
<tr>
<td>112</td>
<td>Gill Net, Drift, Shad</td>
<td>465</td>
</tr>
<tr>
<td>113</td>
<td>Gill Net, Drift, Salmon</td>
<td>455</td>
</tr>
<tr>
<td>114</td>
<td>Gill Net, Drift, Barracuda</td>
<td>450</td>
</tr>
<tr>
<td>115</td>
<td>Gill Net, Crabs</td>
<td>410</td>
</tr>
<tr>
<td>120</td>
<td>Seine, Purse, Other</td>
<td>145</td>
</tr>
<tr>
<td>121</td>
<td>Seine, Purse, Herring</td>
<td>110</td>
</tr>
<tr>
<td>122</td>
<td>Seine, Purse, Mackerel</td>
<td>120</td>
</tr>
<tr>
<td>123</td>
<td>Seine, Purse, Menhaden</td>
<td>125</td>
</tr>
<tr>
<td>124</td>
<td>Seine, Purse, Tuna</td>
<td>140</td>
</tr>
<tr>
<td>131</td>
<td>Dredge, Scallop, Bay</td>
<td>823</td>
</tr>
<tr>
<td>132</td>
<td>Dredge, Scallop, Sea</td>
<td>825</td>
</tr>
<tr>
<td>140</td>
<td>Pound Net, Other</td>
<td>290</td>
</tr>
<tr>
<td>141</td>
<td>Pound Net, Crab</td>
<td>280</td>
</tr>
<tr>
<td>142</td>
<td>Pound Net, Fish</td>
<td>275</td>
</tr>
</tbody>
</table>
C9 - Data Reduction Services Offered

Virginia Marine Fisheries Commission Interactive Data Summarization Programs.

VMRC Harvest Statistics, 1929-1988

The following program is an interactive system designed to ease summarization of historical commercial harvest data. The data has come from several sources and has not been fully verified, as such results should be double checked with existing published sources.

Press Enter Key to Start
ESC Key to Return to DOS

Hit <ESC> key to return to DOS. Hit <RETURN> to begin program.

SCREEN 2

Enter Filename (Default VANMFS.SRT):

Enter Starting Date (Default = 1929): 1983

Enter Ending Date (Default = 1988):

Enter Search Title: Virginia Alewife Landings 1982-1988

Press ESC then Return to go to start screen
SCREEN 4

TABLE OPTIONS

1. Area Table...catch and revenue by area
2. Gear Table...catch and revenue by gear
3. Month Table...catch and revenue by month
4. ASCII Data...puts subsets of raw data in an ASCII data file

CHOOSE A TABLE

To select the desired task, type either 1, 2, 3 or 4. For the current program version the user must use the numeric keys above the letters on the main portion of the keyboard.

SCREEN 5

Please Wait

Now computing an Area Table for the data file vanmfs.
Virginia Alewife Landings - 1983 to 1988
Data from 1983 to 1988.
For the species codes: / 10/
For all gear codes.
For all county codes.
For all area codes
After selection of the desired task screen 5 appears with a description of the selection criteria. When the search is complete, timing information of the search is printed at the bottom of screen 5. The user must then strike the <RETURN> key to have the summarized data printed to the screen. Examples of tables generated are as follows:

### Virginia Alewife Landings - 1983 to 1988

**File ALEWIFE1.5YR**

01-22-1990 16:11:42

Data from 1983 to 1988.

Species / 10/, All Gears, All Counties, All Water Bodies,

#### Catch by area

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Atlantic Ocean</th>
<th>Eastern Shore</th>
<th>Chesapeake Bay</th>
<th>James River</th>
<th>York River</th>
<th>Rappahannock River</th>
<th>Potomac River</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>0</td>
<td>0</td>
<td>36</td>
<td>186</td>
<td>0</td>
<td>192</td>
<td>1,424</td>
<td>1,838</td>
</tr>
<tr>
<td>1984</td>
<td>62</td>
<td>0</td>
<td>3</td>
<td>412</td>
<td>0</td>
<td>49</td>
<td>750</td>
<td>1,257</td>
</tr>
<tr>
<td>1985</td>
<td>25</td>
<td>0</td>
<td>32</td>
<td>122</td>
<td>0</td>
<td>11</td>
<td>241</td>
<td>432</td>
</tr>
<tr>
<td>1986</td>
<td>70</td>
<td>0</td>
<td>5</td>
<td>310</td>
<td>0</td>
<td>0</td>
<td>372</td>
<td>758</td>
</tr>
<tr>
<td>1987</td>
<td>85</td>
<td>0</td>
<td>0</td>
<td>237</td>
<td>0</td>
<td>0</td>
<td>460</td>
<td>783</td>
</tr>
<tr>
<td>1988</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>508</td>
<td>19</td>
<td>8</td>
<td>158</td>
<td>715</td>
</tr>
</tbody>
</table>

**SUM** 243 0 98 1,775 20 260 3,385 5,782

**AVG** 40 0 16 296 3 43 564 964

**PCT** 4.197 0.000 1.695 30.695 0.343 4.494 58.548 100.000

**Note -- All values are in thousands of pounds.**

### Virginia Alewife Landings - 1983 to 1988

Data from 1983 to 1988.

Species / 10/, All Gears, All Counties, All Water Bodies,

#### Revenue by area

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Atlantic Ocean</th>
<th>Eastern Shore</th>
<th>Chesapeake Bay</th>
<th>James River</th>
<th>York River</th>
<th>Rappahannock River</th>
<th>Potomac River</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>33</td>
<td>0</td>
<td>102</td>
<td>88</td>
<td>236</td>
</tr>
<tr>
<td>1984</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>3</td>
<td>42</td>
<td>72</td>
</tr>
<tr>
<td>1985</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>1986</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td>1987</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>47</td>
<td>62</td>
</tr>
<tr>
<td>1988</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>36</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>52</td>
</tr>
</tbody>
</table>

**SUM** 10 0 16 128 1 106 217 480

**AVG** 2 0 3 21 0 18 36 80

**PCT** 2.055 0.000 3.436 26.763 0.272 22.148 45.270 100.000

**Note -- All values are in thousands of dollars.**
Appendix D

TBNEP Dataset Documentation Package

Study Overview

Dataset Contents

Quality Review Summary
Vector and Raster Data
Attribute Data

Transaction Log

Data Request Form
## Study Overview Form

**Tampa Bay National Estuary Program**  
111 7th Avenue South  
St. Petersburg, FL 33701  
(813) 893-2765

<table>
<thead>
<tr>
<th>Project Name:</th>
<th>Submitted by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Number:</td>
<td>Date:</td>
</tr>
<tr>
<td>Dataset Title:</td>
<td></td>
</tr>
<tr>
<td>Filename(s):</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keywords:</td>
</tr>
</tbody>
</table>

| Report Citation(s): |

### Data Collected By:

<table>
<thead>
<tr>
<th>Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
</tr>
</tbody>
</table>

| Sponsoring Agency: |

### Data Available From:

<table>
<thead>
<tr>
<th>Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
</tr>
</tbody>
</table>

| Phone: |
Tampa Bay National Estuary Program

<table>
<thead>
<tr>
<th>Available as:</th>
<th>Hard copy:</th>
<th>Raw data / Summary tables / Maps</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Circle all that apply)</td>
<td>Computer files:</td>
<td>ASCII / Lotus 123 / dBASE / SAS / ARCINFO</td>
</tr>
<tr>
<td>Restrictions on access:</td>
<td>Other:</td>
<td>________________________________</td>
</tr>
</tbody>
</table>

Abstract (attach separate page if necessary):

Spatial Extent of Data:

- Geographic Area:

- Resolution:

Temporal Extent of Data Collection:

- Range of Dates Included:

- Sampling Frequency:
### Study Overview

**Variations and Modifications in Sample Design During Data Collection:**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Method/Equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Form Filled in By:**

<table>
<thead>
<tr>
<th>Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
</tr>
<tr>
<td>Date:</td>
</tr>
</tbody>
</table>
Description of Items

<table>
<thead>
<tr>
<th>Dataset Title</th>
<th>A descriptive title identifying the dataset.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filename(s)</td>
<td>Name of the computer file(s) containing the dataset.</td>
</tr>
<tr>
<td>Subject</td>
<td>A brief summary of the contents of the dataset.</td>
</tr>
<tr>
<td>Keywords</td>
<td>A list of words or phrases identifying topics related to the dataset.</td>
</tr>
<tr>
<td>Report Citations</td>
<td>Citations for literature reporting the results of the study, or describing the methods used.</td>
</tr>
<tr>
<td>Data Collected by</td>
<td>Principal investigator for the study. A contact for questions regarding the hows or whys of data collection.</td>
</tr>
<tr>
<td>Sponsoring Agency</td>
<td>Name of agency or organization funding/sponsoring the study.</td>
</tr>
<tr>
<td>Available from</td>
<td>Who to contact to request a copy of the data</td>
</tr>
<tr>
<td>Available as</td>
<td>Format in which copies of the data are available.</td>
</tr>
<tr>
<td>Restrictions on access</td>
<td>Limits on who can obtain copies of the data, or on what parts of the data set are available</td>
</tr>
<tr>
<td>Abstract</td>
<td>Summary of the study goals, methods, and results.</td>
</tr>
<tr>
<td>Geographic Area</td>
<td>Region for which data are available (e.g. county, USGS quadrangle(s), watershed).</td>
</tr>
<tr>
<td>Resolution</td>
<td>Level of spatial detail, e.g. smallest resolvable area, or number of sampling sites within the geographic area.</td>
</tr>
<tr>
<td>Range of Dates Included</td>
<td>Start and end dates of data collection.</td>
</tr>
<tr>
<td>Sampling Frequency</td>
<td>Continuous, daily, monthly, seasonal, etc.</td>
</tr>
<tr>
<td>Variations and Modifications</td>
<td>Major deviations from sampling design (e.g. locations sampled, number of samples, gear) during data collection, or changes in the design during the course of the study.</td>
</tr>
<tr>
<td>Variables and Methods/Equipment</td>
<td>List of variables included in the dataset, and methods or equipment used in data collection.</td>
</tr>
</tbody>
</table>
Dataset Contents Form

Project Name: ____________________________  Submitted by: ____________________________
Project Number: ____________________________  Date: ____________________________

Dataset Title: ____________________________
Filename(s): ____________________________

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Units</th>
<th>Definition</th>
<th>Description: Range (continuous) or Value Definitions (categorical)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tampa Bay National Estuary Program
111 7th Avenue South
St. Petersburg, FL 33701
(813) 893-2765
Quality Review Summary
Vector and Raster Data

Project Name: ____________________________  Submitted by: ________________
Project Number: _________________________  Date: ________________

Dataset Title: _______________________________________________________
Filename(s): __________________________________________________________

The Florida Growth Management Data Network Coordinating Council (FGMDNCC) has developed the following templates for Quality and Accuracy Reports. Indicate which of these templates have been used to produce a Quality and Accuracy Report for this dataset:

____ Vector Data Template
____ Raster Data Template

Attach the completed report(s) to this form.
III.A. QUALITY AND ACCURACY REPORT: Procedures for Completing

The Templates for the Quality and Accuracy Report are divided into two sections, vector and raster. The appropriate template is to be filled out according to the type of data.

The templates are simply an outline to be followed and filled out in text format. The following rules apply:

- Each outline should be addressed and filled out with the appropriate information, notation that it is not applicable or that the information is unknown.

- Outline levels should never be changed or deleted.

- The Quality and Accuracy report should be made available in an ASCII format.

The template outlines follow. For additional information in filling out the individual sections, see the annotated templates beginning on page 26.
III.B. QUALITY AND ACCURACY REPORT: Truth In Labelling

The purpose of the template for the Quality and Accuracy Report is to provide as much information as possible. The intent is to follow the "truth-in-labelling" practices proposed by the workgroup that developed the Spatial Data Transfer Specifications proposed by the U.S. Geological Survey. The following quotes come from An Interim Proposed Standard for Digital Cartographic Data Quality; Supporting Documentation by N. Chrisman.

We find "quality" to be a wide-ranging concern which can cover any issue affecting the use of cartographic data. The potential uses of digital cartographic data are so diverse that a fixed set of numerical thresholds could not adjust to the potential uses. In more circumscribed application areas (for example, a multipurpose cadastre or a forest inventory), a set of thresholds might be fruitful. Because these standards must serve the whole profession, we foresee a truth-in-labelling standard instead. The idea is to communicate actual numerical properties of the data in a way that potential users can make their own informed decisions on fitness.

The truth-in-labelling concept may seem less rigorous in that it blesses the status quo. Any imprecise, inaccurate database could meet the standard in the formal sense by proclaiming those imprecisions and inaccuracies. These standards place a substantial responsibility on the user to evaluate the quality report to ensure fitness for the particular application.

It is with this thought in mind, that the templates have been created. Formalizing the structure, but still allowing the basis to be textual in content, the report is better suited for user evaluation of the data.
III.C. QUALITY AND ACCURACY REPORT: Template - Vector Data

A. Lineage

1. Description of source material(s)
   a. Name
   b. Scale (specify ratio)
   c. Datum
   d. Map Projection check list

   __ No projection
   __ Lambert Conformal Conic (ie. State Plane-Florida North)
   __ Transverse Mercator (ie. State Plane-Florida East/West)
   __ Albers Equal-Area Conic
   __ Polyconic
   __ Equidistant Cylindrical
   __ Miller Oblated Stereographic
   __ Stereographic (nonpolar/polar)
   __ Regulator Mercator
   __ Modified Transverse Mercator
   __ Bipolar Oblique Conic Conformal
   __ Other (name/explain)

   e. Media of source
   f. Condition of Media
   g. Creator organization/individual

   Name
   ______________________________

   Address
   ______________________________

   Phone (__) ______

h. Date of source material

1. Time interval covered
2. Update schedule
2. Derivation methods for data

a. Method of derivation

1. preautomation compilation
2. digitizing/scanning/transformations
3. equipment
   a. model
   b. resolution
   c. tolerance of digitizer - (if digitized)

b. Date of automation

1. Initial date
2. Update schedule

c. Control Points

___ No projection
___ State Plane-Florida North (Lambert Conformal Conic)
___ State Plane-Florida East-West (Transverse Mercator)
___ Other (name)

d. Explanation of procedures used to digitize/scan/transform etc. the data.

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used.

B. Positional Accuracy

1. Line Work Replication Check

   a. positional
   b. completeness
2. Absolute measure of error reference in units of the coordinate system.
   a. Value
   b. Method of derivation of that number

3. Test Reports (use one or more of the following)
   a. Deductive estimate
      1. date of tests
      2. report
   b. Internal Evidence (geodesy)
      1. date of tests
      2. report
   c. Comparison to Source (overlay)
      1. date of tests
      2. report
   d. Independent source of higher accuracy
      1. date of tests
      2. report

C. Attribute Accuracy

1. Continuous (define what documents can be referenced here)
   a. Absolute measure of error reference in units of the measurement system.
      1. Value
      2. Method of derivation of that number
   b. Test Reports (use one or more of the following)
      1. Deductive estimate
         a. date of tests
         b. report
2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical
   a. Test Reports (if appropriate use one of the following)

1. Deductive Estimate
   a. date
      1. date of tests
      2. date of materials used
      3. rates of change
   b. report
      1. basis for the deduction
      2. deduction

2. Independent Samples
   a. date
      1. date of tests
      2. date of materials used
      3. rates of change
   b. map scale
   c. dates of materials used
   d. report (misclassification matrix)
1. sampling procedures
2. sampling locations
3. results

3. Tests based on Polygon Overlay
   a. date
      1. date of tests
      2. date of materials used
      3. rates of change
   b. map scale
   c. dates of materials used
   d. report
      1. misclassification matrix
      2. explain relationship between maps

D. Logical consistency

1. Tests of valid values
   a. test performed
   b. date
   c. results

2. Cartographic Tests

   Does the provider feel the map is cartographically clean?

   ___ Yes
   ___ No

   a. Answer the following

   1. Do lines intersect only where intended? (yes/no/unknown)
   2. Were duplicate lines eliminated? (yes/no/unknown)
   3. Are all areas completely described? (yes/no/unknown)
   4. Have overshoots and undershoots been eliminated? (yes/no/unknown)
   5. Have all slivers been eliminated? (yes/no/unknown)
6. Does documentation in the lineage section (d-5) describe why the data is cartographically clean? (yes/no/unknown)

b. Tests

1. tests performed
2. date
3. reference to documentation

Conclusions:

3. Topological Tests

a. test performed
b. date
c. software

1. name
2. version
d. results

1. Test for polygon coverage

a. How many polygons are represented on the digital map product?
b. Has a polygon closure been verified?
c. Are Polygon-Ids assigned to each polygon on the digital map?

1. Do polygons have more than one Polygon-Id? (Yes/No)
2. Are the Polygons-Ids unique: (Yes/No)

2. Test for line coverage

a. How many lines are represented on the digital map product?
b. Do the line segments have unique line segment values? (Yes/No)
c. Is the digital map topologically clean? (Yes/No)
3. Test for point coverage
   a. How many points are represented on the digital map product?
   b. Are the Point-Ids unique: (Yes/No)

E. Completeness of Source Materials

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
III. D. QUALITY AND ACCURACY REPORT: Template - Raster Data

* - Represents differences from the Vector Data Quality and Accuracy Report

A. Lineage

1. Description of source material

   a. Name
   b. Scale (specify ratio)
   c. Datum
   d. Map Projection check list*

   __ No projection
   __ Lambert Conformal Conic (ie. State Plane-Florida North)
   __ Transverse Mercator (ie. State Plane-Florida East/West)
   __ (UTM) Universal Transverse Mercator
   __ Lambert Azimuthal Equal-Area (sphere)
   __ Albers Equal-Area Conic
   __ Polyconic
   __ Equidistant Cylindrical
   __ Miller Oblated Stereographic
   __ Stereographic (nonpolar/polar)
   __ Regulator Mercator
   __ Modified Transverse Mercator
   __ Bipolar Oblique Conic Conformal
   __ Other (name/explain)

   e. Media of source*

      1. media
      2. Type of imagery

         __ Landsat MSS (Multispectral Scanner)
         __ Landsat TM (Thematic Mapper)
         __ SPOT Panochromatic
         __ SPOT Multispectral
         __ Other satellite imagery
         __ Aerial Photography
         __ Other
f. Condition of Media

g. Creator organization/individual

Address of creating entity
Name ________________________________
Address ________________________________
Phone (____) ______

h. Date of source material

1. Scene number(s) and dates of source material
2. Update schedule

i. Ancillary Source materials
j. Imported data

2. Derivation methods for data

a. Method of derivation

1. digitizing/scanning/transformations etc.
2. equipment

   a. model
   b. resolution
   c. tolerance of digitizer - if digitized

b. Date of automation

1. Initial date
2. Update schedule

c. Control Points

1. Coordinate reference system check list:

   __ No projection
   __ State Plane-Florida North (Lambert Conformal Conic)
   __ State Plane-Florida East/West (Transverse Mercator)
   __ Other (name)
2. List of coordinate pairs

d. Explanation of procedures used to digitize/scan/transform etc.

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used

B. Positional Accuracy

1. Absolute measure of error reference in units of the coordinate system.

   a. Value
   b. Method of derivation of the number

2. Test Reports *

   a. Deductive estimate

      1. date of tests
      2. report

C. Attribute Accuracy

1. Continuous

   a. Absolute measure of error reference in units of the coordinate system.

      1. Value
      2. Method of derivation of that number

   b. Test Reports (use one or more of the following)

      1. Deductive estimate

         a. date of tests
         b. report
2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical
   a. Test

   1. Deductive Estimate
      a. date
         1. date of tests
         2. date of materials used
         3. rates of change
      b. report
         1. basis for the deduction
         2. deduction

2. Ground Truthing*
   a. Were ground truthing tests performed (yes/no)
   b. date range of tests
   c. How many points were used
   d. Description of the test
D. Logical consistency *

1. Tests of valid values
   a. test performed
   b. date
   c. results

E. Completeness of Source Materials

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
III.E. QUALITY AND ACCURACY REPORT: Annotated Template - Vector Data

A. LINEAGE

1. Description of source material(s)

   a. Name

   b. Scale (specify ratio)

      scale: ratio between the distance on a map, chart, photograph or image and the corresponding distance on the surface of the Earth.

   c. Datum

      Geodetic Datum: A set of constants specifying the coordinate system used for geodetic control, i.e., for calculating coordinates of points on the earth.

   d. Map Projection check list

      map projection: systematic drawing of lines of a plane surface to represent the parallels of latitude and the meridians of longitude of the Earth.

      _ No projection
      _ Lambert Conformal Conic (ie. State Plane-Florida North)
      _ Transverse Mercator (ie. State Plane-Florida East/West)
      _ Albers Equal-Area Conic
      _ Polyconic
      _ Equidistance Cylindrical
      _ Miller Oblated Stereographic
      _ Stereographic (nonpolar/polar)
      _ Regulator Mercator
      _ Modified Transverse Mercator
      _ Bipolar Oblique Conic Conformal
      _ Other (name/explain)
Annotated Template - Vector Data
Page Two

(A.1.e.)

(e. Media of source

    source: origins of the data.

    media: physical substance the data was taken from (mylar, paper, electronic tape, etc.).

(f. Condition of Media

A qualitative statement of the condition of the media. Statements such as "good" or "bad" must be given a basis.

(g. Creator organization/individual

    Address of creating entity
    Name
    Address
    Phone (__) _____

(h. Date of source material

A history of the development of the source material. (There may be multiples of this entry.)

1. Time interval covered

A history of the development of the source material.

2. Update schedule

Schedule on which the data is being updated. If updating is continuous some indication of the turnover should be given.
2. Derivation methods for data

The purpose of this step is to describe how the data was brought into the system. By knowing how the data was created and the technology used, limits on the accuracy may be deduced.

a. Method of derivation

1. preautomation compilation
2. digitizing/scanning/transformations etc.
3. equipment
   a. model
   b. resolution
   c. tolerance of digitizer - (if digitized)

b. Date of automation

1. Initial date
2. Update schedule

c. Control Points

c: any station in a horizontal or vertical control network that is identified in a data set or photograph and used for correlating the data shown in that data set or photograph.

c: a particular kind of reference frame or system, such as plane rectangular coordinates or spherical coordinates, that uses linear or angular quantities to designate the position of points within that particular reference frame or system.
(A.2.c.1)

1. Coordinate reference system check list

- No projection
- State Plane-Florida North (Lambert Conformal Conic)
- State Plane-Florida East/West (Transverse Mercator)
- Other (name)

2. List of control point coordinate pairs

d. Explanation of procedures used to
digitize/scan/transform etc. the data

This is a description of procedures that would indicate the
good quality/accuracy of the data captured. Information that would not
provide insights should not be included. Transformation routines
that are supplied by vendors should include the name of the
transformation module. User created transformations should
include the following:

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used

B. POSITIONAL ACCURACY

Tests of accuracy after all transformations have been performed on a
particular layer.

1. Line Work Replication Check

   a. Positional
   b. Completeness
(B.2.)

2. Absolute measure of error reference in units of the coordinate system.

(provides a numerical estimate of expected discrepancies)

"quality of the final product after all transformations."

a. Value
b. Method of derivation of that number

3. Test Reports (use one or more of the following)

a. Deductive estimate

deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

1. date of tests
2. report

b. Internal Evidence (geodesy)

1. date of tests
2. report

c. Comparison to Source (overlay)

1. date of tests
2. report

d. Independent source of higher accuracy

1. date of tests
2. report
C. ATTRIBUTE ACCURACY

Accuracy assessment for measures on a continuous scale shall be performed using procedures similar to those used for positional accuracy (providing a numerical estimate of expected discrepancies).

There has been considerable discussion on how much detail is required at this step. It is the view of the developers of the report that as much information be provided as possible. This does not mean that a test must be performed that wouldn't normally be performed, but it does mean that all tests that were performed should be reported. The level of reporting should be at such a level as to be useful to the recipient. The exact format is up to the discretion of the developer.

1. Continuous or numerical (See Data Dictionary Template)
   a. Absolute measure of error referenced in units of the measurement system.

   (provides a numerical estimate of expected discrepancies)
   "quality of the final product after all transformations."

   1. Value
   2. Method of derivation of that number

   b. Test Reports (use one or more of the following)

   1. Deductive estimate

   deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

   a. date of tests
   b. report
Annotated Template - Vector Data
Page Seven

(C.1.b.2.)

2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical (See Data Dictionary Template)

The intent is for a statistical sampling of an attribute that would
give a level of assurance as to the accuracy of the data. This was
not intended for a item by item data check.

   a. Tests (if appropriate use one of the following)

       1. Deductive Estimate

           a. date

           1. date of tests
           2. date of materials used
           3. rates of change

In the case of discrepancies between the date of the tests and the
dates of the materials used, the rates of change in the phenomena
classified will be reported. It is up to the creator to identify
dates as being discrepant. The intent is to make note of temporal
items.
Annotated Template - Vector Data
Page Eight

(C.2.a.1.b.)

b. report

1. basis for the deduction
2. deduction

2. Independent Samples

a. date
1. date of tests
2. date of materials used
3. rates of change

*In the case of discrepancies between the dates of the tests and the dates of the materials used the rates of change in the phenomena classified will be reported.*

b. map scale
c. dates of materials used
d. report (misclassification matrix)

1. sampling procedures
2. sampling locations
3. results
Annotated Template - Vector Data
Page Nine

(C.2.a.3.)

3. Tests based on Polygon Overlay

These are for polygons that represent a category, such as wetlands, clay soils, etc.

a. date

1. date of tests
2. date of materials used
3. rates of change

In the case of discrepancies between the dates of the tests and the dates of the materials used, the rates of change in the phenomena classified will be reported.

b. map scale
c. dates of materials used
d. report

1. misclassification matrix
2. explain relationship between maps

D. LOGICAL CONSISTENCY

1. Tests of valid values

Test for attribute data.

a. test performed
b. date
c. results

2. Cartographic Tests

Does the provider feel the map is cartographically clean?

_ Yes
_ No
(D.2.a.)

a. Answer the following

1. Do lines intersect only where intended? (yes/no/unknown)
2. Were duplicate lines eliminated? (yes/no/unknown)
3. Are all areas completely described? (yes/no/unknown)
4. Have overshoots and undershoots been eliminated? (yes/no/unknown)
5. Have all slivers been eliminated? (yes/no/unknown)
6. Does documentation in the lineage section (d-5) describe why the data is cartographically clean? (yes/no/unknown)

b. Tests

1. tests performed
2. date
3. reference to documentation

Conclusions:

3. Topological Tests

topology: a branch of geometrical mathematics which is concerned with order, contiguity, and relative position, rather than actual linear dimensions.

topologic error checking: process of ensuring the logical consistency of the data is intact; all polygons are closed, all arcs are connected to nodes, etc.

a. test performed
b. date
c. software

1. name
2. version
d. results

1. Test for polygon coverage
   a. How many polygons are represented on the digital map product?
   b. Has a polygon closure been verified?
   c. Are Polygon-Ids assigned to each polygon on the digital map?

   1. Do polygons have more than one Polygon-Id?  (Yes/No)
   2. Are the Polygon-Ids unique?  (Yes/No)

2. Test for Line coverage
   a. How many lines are represented on the digital map product?
   b. Do the line segments have unique line segment values?  (Yes/No)
   c. Is the digital map topologically clean?  (Yes/No)

3. Test for point coverage
   a. How many points are represented on the digital map product?
   b. Are the Point-Ids unique?  (Yes/No)
E. COMPLETENESS OF SOURCE MATERIALS

The purpose of the completeness section is to describe the set of information collected in comparison to a larger set. For example, a set called "Well Data Points" may be all manmade wells in the area described or it may be only private wells used for homes.

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
III.E. QUALITY AND ACCURACY REPORT: Annotated Template -
Raster Data

A. LINEAGE

1. Description of source material

The purpose is to describe the origins of the data from which the "map" was created allowing the receiver to evaluate the utility of the data.

   a. Name
   b. Scale (specify ratio)

scale: ratio or fraction between the distance on a map, chart, photograph or image and the corresponding distance on the surface of the Earth.

   c. Datum

Geodetic Datum: A set of constants specifying the coordinate system used for geodetic control, i.e., for calculating coordinates of points on the earth.

   d. Map Projection check list*

map projection: systematic drawing of lines of a plane surface to represent the parallels of latitude and the meridians of longitude of the Earth.

   __ No projection
   __ Lambert Conformal Conic (ie. State Plane-Florida North)
   __ Transverse Mercator (ie. State Plane-Florida East/West)
   __ (UTM) Universal Transverse Mercator
   __ Lambert Azimuthal Equal-Area (sphere)
   __ Albers Equal-Area Conic
   __ Polyconic
   __ Equidistance Cylindrical
   __ Miller Oblated Stereographic
   __ Stereographic (nonpolar/polar)
Annotated Template - Raster Data
Page Two

(A.1.d.)

___ Regulator Mercator
___ Modified Transverse Mercator
___ Bipolar Oblique Conic Conformal
___ Other (name/explain)

e. Media of source*

1. media

source: origins of the data

Media that the data was on (mylar, paper, electronic tape, etc.).

2. Type of imagery

___ Landsat MSS (Multispectral Scanner)
___ Landsat TM (Thematic Mapper)
___ SPOT Panochromatic
___ SPOT Multispectral
___ Other satellite imagery
___ Aerial Photography
___ Other

f. Condition of Media

A qualitative statement of the condition of the media. Statements such as "good", or "bad" must be substantiated.

g. Creator organization/individual

Address of creating entity
Name
Address
Phone  (___) _____
h. Date of source material

A history of the development of the source material.
There may be multiples of this entry.

1. Scene numbers(s) and dates of source material

Include scene numbers if appropriate.
There may be multiples of this entry.

2. Update schedule

Schedule on which the data is being updated. If updating is continuous some indication of the turnover should be provided.

i. Ancillary Source materials

ancillary data: auxiliary or supplementary data.

j. Imported data

What was the format and what was the software.

If graphic data was not created in the system that it is being used in, what structure was it imported in. For example was it a DLG III format, SDTS or was it a vendor created importation package.

2. Derivation methods for data

The purpose is to describe how the data was brought into the system. By knowing how the data was created and the technology used, limits on the accuracy may be deduced.
Annotated Template - Raster Data
Page Four

(A.2.a.)

a. Method of derivation

1. digitizing/scanning/transformations etc.
2. equipment

   a. model
   b. resolution
   c. tolerance of digitizer - if digitized

b. Date of automation

1. Initial date
2. Update schedule

c. Control Points

control point - any station in a horizontal or vertical control network that is identified in a data set or photograph and used for correlating the data shown in that data set or photograph.

coordinate systems - a particular kind of reference frame or system, such as plane rectangular coordinates or spherical coordinates, that use linear or angular quantities to designate the position of points within that particular reference frame or system.

1. Coordinate reference system check list*

   __ No projection
   __ State Plane-Florida North (Lambert Conformal Conic)
   __ State Plane-Florida East/West (Transverse Mercator)
   __ Other (name)

2. List of coordinate pairs

41
(A.2.d.)

d. Explanation of procedures used to digitize/scan/transform etc. the data

This is a description of procedures that would indicate the quality/accuracy of the data captured. Information that would not provide insights should not be included. This is particularly important for transformations that did not use canned software packages. If there were transformations of this nature, the following should be included:

1. Name of transformation methodology
2. Description of algorithm
3. Mathematics used in the transformation
4. Set of sample computations

e. Software system and version used

B. POSITIONAL ACCURACY

Tests of accuracy after all transformations have been performed on a particular layer.

1. Absolute measure of error reference in units of the coordinate system.

   (provides a numerical estimate of expected discrepancies)
   "quality of the final product after all transformations."

   a. Value.
   b. Method of derivation of that number

2. Test Reports (use one or more of the following)*
(B.2.a.)

a. Deductive estimate

deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

1. date of tests
2. report

C. ATTRIBUTE ACCURACY

Accuracy assessment for measures on a continuous scale shall be performed using procedures similar to those used for positional accuracy (providing a numerical estimate of expected discrepancies).

1. Continuous

a. Absolute measure of error reference in units of the measurement system.

(provides a numerical estimate of expected discrepancies) "quality of the final product after all transformations."

1. Value
2. Method of derivation of that number

b. Test Reports (use one or more of the following)

1. Deductive estimate

deduction: The deriving of a conclusion by reasoning. It may be necessary that a best guess is given. Any assumptions that were made to derive the conclusion should be described.

a. date of tests
b. report
2. Internal Evidence (geodesy)
   a. date of tests
   b. report

3. Comparison to Source (overlay)
   a. date of tests
   b. report

4. Independent source of higher accuracy
   a. date of tests
   b. report

2. Categorical

The intent is for a statistical sampling of an attribute that would give a level of assurance as to the accuracy of the data. This was not intended for a item by item data check.

   a. Test (if appropriate use one of the following)

   1. Deductive Estimate
      a. date
         1. date of tests
         2. date of materials used
         3. rates of change

In the case of discrepancies between the date of the tests and the dates of the materials used the rates of change in the phenomena classified will be reported. It is up to the creator to identify dates as being discrepant. The intent is to make note of temporal items.
(C.2.a.1.b.)

b. map scale
c. report
   1. basis for the deduction
   2. deduction

2. Ground Truthing*

   a. Were ground truthing tests performed? (yes/no)
   b. date range of tests
   c. How many points were used
   d. Description of the test

D. LOGICAL CONSISTENCY *

1. Tests of valid values

   Test for attribute data.

   a. test performed
   b. date
   c. results

E. COMPLETENESS OF SOURCE MATERIALS

The purpose of the completeness section is to describe the set of information collected in comparison to a larger set. For example a set called "Well Data Points" may be all manmade wells in the area described or it may be only private wells used for homes.

1. Selection Criteria
2. Definitions used
3. Other relevant mapping rules
4. Deviation from standard definitions and interpretations
5. Description of the relationship between the objects
6. Tests for taxonomic completeness
   a. procedures
   b. result
Indicate which level of quality review has been completed for this dataset:

_____ Level 0 = No Scrutiny

Raw data as sampled in the field or experimental setting (e.g.,
data sheets, aerial photos, measuring instrument printouts and
datasets).

_____ Level 1 = Reviewed

Level 0 data that have been visually inspected for readability
and completeness (e.g., are handwritten data legible, or are
computer files complete?).

_____ Level 2 = Verified*

Level 1 data that have been scrutinized for data encryption
errors and data measurement errors (e.g., checked for keypunch
errors by double entry, checked for instrument failure and
plausibility of values by range checks, checked to see if all
samples are in data).

_____ Level 3 = Validated*

Level 2 data that have been scrutinized by a person proficient in
the field of interest (e.g., checked for unusual values, and
evaluated for systematic biases).

* For Levels 2 and 3, attach a description of the review process, including the
dates when specific checks on this dataset were completed.
Guidelines

The purpose of the transaction log is to maintain a record of the origin and in-house maintenance of a dataset. At a minimum, the log should include:

- A title and/or identification number that unambiguously identifies the dataset.

- The date the dataset was received or created, and the source of the data.

- A record of all changes made in the contents of the data (including corrections, additions, and deletions). The log may either describe the change in detail, or give the date and a brief description that can be used to locate a complete record of the change(s) stored elsewhere.

Additional transactions that could be recorded include:

- What quality review steps have been conducted, and the date they were completed.

- A record of who has received a copy of the data, on what date.
<table>
<thead>
<tr>
<th>Date</th>
<th>Transaction</th>
<th>By</th>
<th>Supervisor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Request Form

Tampa Bay National Estuary Program
111 7th Avenue South
St. Petersburg, FL 33701
(813) 893-2765

PLEASE PRINT OR TYPE

Name: ____________________________ Phone: ______________

Organization Name: ____________________________

Organization Type (check one):

Federal State County Local Private
Educational Other: ____________________________

Address: ______________________________________

______________________________________________

Date of Request: ____________________________

Dataset Requested: ____________________________

Medium Requested (choose one):

____ DD 720K 3-1/2" floppy disk ______ HD 1.2M 3-1/2" floppy disk
____ DD 720K 5-1/4" floppy disk ______ HD 1.2M 5-1/4" floppy disk
Appendix E

Ocean Data Evaluation System (ODES) Data Submission Guidelines
Ocean Data Evaluation System (ODES) Data Submission Guidelines

There is an EPA requirement that all primary data collected by Tampa Bay National Estuary Program must be entered into ODES. ODES is EPA's marine and environmental monitoring database. Data submitted to ODES has computerized error checking performed on it and is reviewed technically. An overview of the data submission process is shown in Figure 1. Data submitter responsibilities are summarized in Table 1.

ODES Contents

Types of data contained in ODES are listed in Table 2. In order to enter data into ODES format, the data submitter must use:

• ODES standard data formats
• ODES data codes
• a standard file format defined by ODES

These data formats and data codes were established by the National Oceanographic Data Center (NODC). Information on ODES data formats and codes is contained in the ODES Data Submissions Manual.

ODES2 Data Entry System

A PC data entry system ODES2 can be used to help enter data into ODES file formats. ODES2 uses screens for data entry and creates data files in the standard file format defined by ODES. ODES2 provides pop-up dictionaries of ODES codes that eliminate the need to look up codes in the ODES Data Submission Manual.

Descriptive Overview

The data submitter must complete a 3-4 page form that briefly describes study methods and quality control. A copy of the form for submitting this descriptive overview for water quality data is attached. Similar forms are available for each type of ODES data.

The ODES Data Submission Manual and the ODES2 Data Entry System are available upon request from:

Mr. Robert King
ODES Work Assignment Manager
United States Environmental Protection Agency
Office of Wetlands, Oceans and Watersheds Assessment and Watershed Division
401 M Street, S.W. (WH-553)
Washington, D.C. 20460
(202) 260-7028
Figure 1. Data Submission Process from ODES Data Submission Manual.
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Obtain EPA approval to submit data.</td>
</tr>
<tr>
<td>2</td>
<td>Perform quality control checks during data entry</td>
</tr>
<tr>
<td>3</td>
<td>Enter data into ODES format</td>
</tr>
<tr>
<td>4</td>
<td>Prepare a description of the data set, include in a discussion of all quality control procedures followed during data collection, analysis, and entry.</td>
</tr>
<tr>
<td>5</td>
<td>Certify that data are accurate.</td>
</tr>
<tr>
<td>6</td>
<td>Send data, data set description and data certification statement to the ODES Staff.</td>
</tr>
<tr>
<td>7</td>
<td>Review hard copy of data and listing of errors detected during machine checks of information.</td>
</tr>
<tr>
<td>8</td>
<td>Provide verification statement regarding any necessary data revisions or corrections to the ODES Technical Reviews.</td>
</tr>
<tr>
<td>9</td>
<td>Provide the ODES Technical Reviewers with additional information, as needed, to evaluate the quality of the data.</td>
</tr>
<tr>
<td>Table 2. ODES Data Types</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td></td>
</tr>
<tr>
<td>Bacterial/Viral</td>
<td></td>
</tr>
<tr>
<td>Benthic Infauna Survey</td>
<td></td>
</tr>
<tr>
<td>Bioaccumulation</td>
<td></td>
</tr>
<tr>
<td>Bioassay</td>
<td></td>
</tr>
<tr>
<td>Current Meter</td>
<td></td>
</tr>
<tr>
<td>Fish Pathology</td>
<td></td>
</tr>
<tr>
<td>Influent/Effluent</td>
<td></td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
</tr>
<tr>
<td>Plankton Abundance</td>
<td></td>
</tr>
<tr>
<td>Receiving Water Quality</td>
<td></td>
</tr>
<tr>
<td>Sediment Grain Size</td>
<td></td>
</tr>
<tr>
<td>Sediment Pollutant</td>
<td></td>
</tr>
<tr>
<td>Trawl/Seine Survey</td>
<td></td>
</tr>
</tbody>
</table>
Attachment 1

Sample ODES Data Submission Formats


Description of Receiving Water Data Submitted to ODES
VI. ODES File Type 144 for Receiving Water Quality Data

A. Introduction

ODES File Type 144 for Receiving Water Quality Data can be used to report concentrations of organic and inorganic chemicals, or measurements of selected non-chemical water quality variables (e.g., temperature, BOD). This chapter describes how to compile and submit to the ODES Staff a data set containing receiving water quality data. The data set will consist of a set of different record types arranged in a hierarchical order. Section B of this chapter provides an illustration of each of the different record types and explains the hierarchical relationships among them. Section C contains detailed instructions for compiling your data set. The instructions for each of the data elements include information such as the length of the field, the type of data (e.g., numeric, character), and a description of the data element. In most cases where a data element is a code (e.g., a one-character code for detection limits), a list of valid codes will accompany the data element description. In some cases, where the list of data entry codes is long (e.g., codes for chemical analysis methods), you will be referred to Appendix C. A list of ten-character ODES Chemical Codes for identifying chemicals can be found in Appendix A of this manual.

B. Hierarchical Relationships

This version of File Type 144 is composed of seven record types:

- Record Type "A" is the Survey Header Record. It is used to report information common to the entire data set (e.g., survey dates, investigator's name).

- Record Type "Z" is the Header Record for Quality Assurance Samples. It is used to identify data for analytical blanks or other quality assurance samples referred to within the data set.

- Record Type "C" is the Station Header Record. It is used to report information about the station where the sample was collected (e.g., station location, water depth).

- Record Type "D" is the Station Environment Record. It is used to report information about the environmental conditions at the station where the sample was collected (e.g., current speed, wave height; air temperature).

- Record Type "E" is the Sample Record. It is used to report information about each whole water sample (e.g., sample depth, sphere).

- Record Type "F" is Data Record 1. It is used to report measurements for selected non-chemical water quality variables (e.g., temperature, BOD).

- Record Type "G" is Data Record 2. It is used to report concentrations of inorganic and organic chemicals (e.g., Cadmium, DDT).
Exhibit VI-B-1 provides an illustration of the structure of each of these record types. These record images, in conjunction with the detailed instructions provided in Section C, will enable you to quickly and accurately complete individual data entry sheets (or enter this data onto tape or disk) for each of the record types.

In addition it is also necessary to understand the relationships among the different record types so that you can compile a comprehensive data set in the proper hierarchical format. The hierarchical order described herein is designed to minimize the duplication of data entry tasks and is in accordance with standards adopted by the National Oceanographic Data Center (NODC). The relationships are straightforward, and taking just a few minutes now to understand them could save you considerable time and effort later.

As shown above, the different record types report information at different logical levels. For example, the Survey Header Record (Record Type "A") reports information common to the entire survey data set, whereas the Sample Record (Record Type "C") reports information about each of the samples taken during the survey. Thus, there will be only one Survey Header Record per data set, but there will typically be multiple Sample Records. Accordingly, for each Sample Record, there will be multiple Data Records.

Exhibit VI-B-2 shows how the different record types would be arranged in a data set reporting data from a survey conducted at two stations, where one water sample was taken at each station. For each sample, measurements were recorded for both non-chemicals (i.e., Data Record 1) and chemicals (i.e., Data Record 2). Data for quality assurance samples (e.g., analytical blanks) are reported first, followed by the sampling data for the first and second station.

If you have any questions about how to complete your data entry sheets or how to compile your data, please contact the ODES Staff. (See the Preface for address and telephone information.)

C. Detailed Data Element Descriptions

Detailed descriptions for all of the data elements in ODES File Type 144 for Receiving Water Quality Data are provided below. All record types for a given data set will share the same nine-character file identifier. This file ID consists of three characters to identify the file type, i.e., "144", followed by a two-character ID for the municipality/institution/agency conducting the survey, the last two digits of the year in which the survey was taken, a series number assigned by the investigator and a scan code to indicate sampling frequency and collection period.

The contents of all numeric fields should be right-justified, preceded by either blanks, zeroes, or where applicable, positive (+) or negative (-) signs. The contents of all decimal points are implied rather than physically included (see Exhibit VI-B-1). All character fields should be left justified, blank-filled.

*Survey Header Record*

This record is mandatory and should be submitted only once for each data set. See Exhibit VI-C-1 for a description of all data elements in this record.
EXHIBIT VI-B-1 (Cont'd)

Station Header Record

<table>
<thead>
<tr>
<th>FILE TYPE</th>
<th>SOURCE ID</th>
<th>SERIES NUMBER</th>
<th>RECORD TYPE</th>
<th>STN. ID</th>
<th>STATION LATITUDE</th>
<th>STATION LONGITUDE</th>
<th>DATE</th>
<th>TIME</th>
<th>RELATION TO ZID</th>
<th>LOCATION CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1</td>
<td>0000000001</td>
<td>1</td>
<td>N</td>
<td>1222222222</td>
<td>3333333333</td>
<td>4444444444</td>
<td>5555555555</td>
<td>6666666666</td>
<td>7777777777</td>
</tr>
<tr>
<td>144</td>
<td>1</td>
<td>0111111111</td>
<td>1</td>
<td>N</td>
<td>1234567890</td>
<td>1234567890</td>
<td>0123456789</td>
<td>0123456789</td>
<td>0123456789</td>
<td>0123456789</td>
</tr>
</tbody>
</table>

Station Environmental Record

<table>
<thead>
<tr>
<th>FILE TYPE</th>
<th>SOURCE ID</th>
<th>SERIES NUMBER</th>
<th>RECORD TYPE</th>
<th>STN. ID</th>
<th>BOTTOM TYPE</th>
<th>CURRENT DEP.</th>
<th>WIND DEP.</th>
<th>TIDE DEPTH</th>
<th>DEPTH OF THERMAL</th>
<th>AIR TEMPERATURE</th>
<th>SALINITY</th>
<th>WAVE PERIOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1</td>
<td>000000001</td>
<td>1</td>
<td>D</td>
<td>1222222222</td>
<td>3333333333</td>
<td>4444444444</td>
<td>5555555555</td>
<td>6666666666</td>
<td>7777777777</td>
<td>8888888888</td>
<td></td>
</tr>
<tr>
<td>144</td>
<td>1</td>
<td>0111111111</td>
<td>1</td>
<td>D</td>
<td>1234567890</td>
<td>1234567890</td>
<td>0123456789</td>
<td>0123456789</td>
<td>0123456789</td>
<td>0123456789</td>
<td>0123456789</td>
<td></td>
</tr>
</tbody>
</table>
EXHIBIT VI-B-2
Receiving Water Quality Data
Order of Record Types/Reason for Occurrence

Survey Header Record
- Describes the survey - first record in the data set

Station Header Record
- Describes first station where samples were collected

Station Environment Record
- Describes environmental conditions at the first station

Sample Record
- Describes first water sample

Data Record 1
- Reports measurements for selected non-chemical water quality variables for first replicate observation

Data Record 2’s
- Report concentrations for inorganic and organic chemicals for first replicate observation

Data Record 2’s

Header Record for Analytical Blanks
- Begins the reporting of data for all analytical blanks in this survey

Sample Record
- Reports identifying information for first analytical blank

Data Record 2’s
- Report chemical concentrations for first chemical blank

See record arrangement shown for the first station where samples were collected
Header Record for Quality Assurance Samples

This record is used to identify data for all quality assurance (QA) samples recorded in the data set. All QA sample measurements (e.g., field, laboratory, or transport blanks or spiked matrix samples) should be reported within this level of the hierarchy. See Exhibit VI-C-2 for a description of all data elements in this record. (See Exhibit VI-B-2 for a description of how data records for QA samples should be organized within the data set.)

Station Header Record

This record is mandatory; one record should be generated for each sampling station. Stations should be uniquely identified by the Station ID Prefix and Station ID fields (columns 11-15). See Exhibit VI-C-3 for a description of all data elements in this record.

Station Environment Record

This record is mandatory; one Station Environment Record must accompany each Station Header Record included in the data set. Station Environment Records are "linked" to Station Header Records by the Station Prefix and Station ID fields (columns 11-15). See Exhibit VI-C-4 for a description of all data elements in this record.

Sample Record

This record reports information common to each sample or analytical blank. This is a mandatory record; one record should be generated for each sample or quality assurance samples collected. Columns 18-80 are relevant only for Sample Records which identify actual field samples. When the Sample Record is used to identify analytical blanks or other QA samples, columns 18-80 should be left blank. See Exhibit VI-C-5 for a description of all data elements in this record.

Data Record 1

Data Record 1 reports data on selected water quality variables for receiving water quality samples. This record is "linked" to the Sample Record by the Sample Number. Data Record 1 is not used when reporting data for analytical blanks or other QA samples. See Exhibit VI-C-6 for a description of all data elements in this record.

Data Record 2

This record reports concentrations for inorganic and organic chemicals; data for as many as three chemicals can be reported on each Data Record 2. Each compound reported in a single record should have analytical methods in common and be associated with a single blank sample. These records are "linked" to Sample Records by the Sample Number. See Exhibit VI-C-7 for a description of all data elements in this record.
D. Data Entry Codes

This section contains a list of all the code types used in ODES File Type 144 for Water Quality Data. The codes are listed by their identifiers and the fields in which they are to be used. The elements of these code types are listed in the appendices: Appendix A contains ODES Chemical and Water Quality Codes, Appendix B contains NODC/ODES Taxonomic Codes, and Appendix C contains all other ODES code types.

<table>
<thead>
<tr>
<th>Code Identifier</th>
<th>Field Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPA-1</td>
<td>Relation to ZID</td>
</tr>
<tr>
<td>0346</td>
<td>Station Location code</td>
</tr>
<tr>
<td>0077</td>
<td>Bottom Type</td>
</tr>
<tr>
<td>0110</td>
<td>Current Direction</td>
</tr>
<tr>
<td>0109</td>
<td>Sea State</td>
</tr>
<tr>
<td>0154</td>
<td>Tide Stage</td>
</tr>
<tr>
<td>0362</td>
<td>Wave Height</td>
</tr>
<tr>
<td>0376</td>
<td>Gear Type</td>
</tr>
<tr>
<td>0093</td>
<td>Sphere</td>
</tr>
<tr>
<td>EPA-4</td>
<td>Salinity Equipment</td>
</tr>
<tr>
<td>0094</td>
<td>Transmissivity Equipment</td>
</tr>
<tr>
<td>EPA-5</td>
<td>Dissolved Oxygen Equipment</td>
</tr>
<tr>
<td>EPA-2</td>
<td>Limit code</td>
</tr>
<tr>
<td>EPA-3</td>
<td>Qualifier code</td>
</tr>
<tr>
<td>0377</td>
<td>Measurement code</td>
</tr>
<tr>
<td>ODES</td>
<td>Chemical Code</td>
</tr>
<tr>
<td>0350</td>
<td>Chemical Analysis Method</td>
</tr>
<tr>
<td>EPA-10</td>
<td>Scan ID</td>
</tr>
<tr>
<td>Starting Column</td>
<td>Length of Field</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>39</td>
<td>6</td>
</tr>
<tr>
<td>45</td>
<td>15</td>
</tr>
<tr>
<td>60</td>
<td>17</td>
</tr>
</tbody>
</table>
EXHIBIT VI-C-1 (Cont'd)
Survey Header Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
<td>4</td>
<td>Blank</td>
<td>BLANK</td>
</tr>
</tbody>
</table>
EXHIBIT VI-C-2
Header Record for Quality Assurance Samples

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This number appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description forms which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No. EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;Z&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>70</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
</tbody>
</table>
EXHIBIT VI-C-3
Station Header Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This number appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description forms which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No. EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;C&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>Alphanumeric</td>
<td>STATION IDENTIFICATION PREFIX - a two-character code to indicate that a station has been revisited during a survey. Leave this field blank for the first occupation of the station in a given survey. Valid codes for reoccupations of the station are R1-R9.</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>Alphanumeric</td>
<td>STATION IDENTIFIER - three characters assigned by the investigator to identify the sampling station.</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>DDMMSS</td>
<td>STATION LATITUDE - latitude of the station. DD is degrees. MM is minutes. SS is seconds.</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>Code</td>
<td>HEMISPHERE - set to &quot;N&quot;.</td>
</tr>
</tbody>
</table>

VI-13
<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>7</td>
<td>DDDMMSS</td>
<td>STATION LONGITUDE - longitude of the station. DDD is degrees. MM is minutes. SS is seconds.</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>Code</td>
<td>HEMISPHERE - set to &quot;W&quot;.</td>
</tr>
<tr>
<td>31</td>
<td>6</td>
<td>YYMMDD</td>
<td>DATE - date of sample collection. YY is the last two digits of the year. MM is the number of the month (01-12). DD is the day (01-31).</td>
</tr>
<tr>
<td>37</td>
<td>4</td>
<td>HHMM</td>
<td>TIME - starting time of sample collection in 24-hour format. HH is the hour (00-23). MM is the number of minutes (00-59).</td>
</tr>
<tr>
<td>41</td>
<td>2</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>43</td>
<td>1</td>
<td>Code</td>
<td>RELATION TO ZID - a one-character code to describe the classification of the station with respect to the zone of initial dilution (ZID). This field is applicable only to EPA's 301(h) program. Leave this field blank for all stations not classified as part of a 301(h) monitoring program. Where applicable, use Code EPA-1:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>W - Within ZID</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B - ZID Boundary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N - Near Field</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R - Reference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F - Far Field</td>
</tr>
<tr>
<td>44</td>
<td>6</td>
<td>Numeric</td>
<td>DISTANCE TO ZID - six digits for distance in meters from the station to the edge of the ZID. This field is only applicable for EPA's 301(h) program.</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>Numeric</td>
<td>WATER DEPTH - depth at station in meters with one decimal place.</td>
</tr>
</tbody>
</table>
EXHIBIT VI-C-3 (Cont'd)
Station Header Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>1</td>
<td>Code</td>
<td>STATION LOCATION CODE - a one-character code to describe the location of the station. Use Code No. 0346:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A -- Stream</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B -- Estuary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C -- Lake</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D -- Ocean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E -- Well</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F -- Other</td>
</tr>
<tr>
<td>56</td>
<td>25</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
</tbody>
</table>
## Exhibit VI-C-4
Station Environment Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description forms which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No. EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;D&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>Alphanumeric</td>
<td>STATION IDENTIFICATION PREFIX - a two-character code to indicate that a station has been revisited during the survey. Leave this field blank for the first occupation of the station in a given survey. Valid codes for reoccupations of the station are R1-R9.</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>Alphanumeric</td>
<td>STATION IDENTIFIER - three characters assigned by the investigator to identify the sampling station.</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>Blank</td>
<td>BLANKS.</td>
</tr>
</tbody>
</table>
### EXHIBIT VI-C-4 (Cont'd)
Station Environment Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2</td>
<td>Code</td>
<td>BOTTOM TYPE - a two-character code to describe the specific bottom type. Use Code No. 0077. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>Numeric</td>
<td>CURRENT SPEED - three digits for the speed of the surface current in meters per second with two decimal places.</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>Code</td>
<td>CURRENT DIRECTION - a two-character code for the direction of the current. Use Code No. 0110. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>27</td>
<td>2</td>
<td>Numeric</td>
<td>WIND SPEED - two digits for surface wind speed in meters per second with one decimal place.</td>
</tr>
<tr>
<td>29</td>
<td>2</td>
<td>Code</td>
<td>WIND DIRECTION - a two-character code for the wind direction. Use Code No. 0110. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>Code</td>
<td>SEA STATE - a one-character code to describe sea conditions. Use Code No. 0109:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 -- Calm-Glassy (0 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 -- Calm-Rippled (0-.1 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 -- Smooth-Wavelet (.1-.5 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 -- Slight (.5 -1.25 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 -- Moderate (1.25-2.5 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 -- Rough (2.5 -4.0 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 -- Very Rough (4-6 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7 -- High (6-9 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 -- Very High (9-14 meters)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9 -- Phenomenal (&gt;14 meters)</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>Numeric</td>
<td>TIDE HEIGHT - three digits for tide height in meters with one decimal place.</td>
</tr>
</tbody>
</table>
# EXHIBIT VI-C-4 (Cont'd)
Station Environment Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>1</td>
<td>Code</td>
<td>TIDE STAGE - a one-character code for tide stage. Use Code No. 0154:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blank -- No Information</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 -- Ebb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 -- Ebb Slack</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 -- Flood</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 -- Flood Slack</td>
</tr>
<tr>
<td>36</td>
<td>3</td>
<td>Numeric</td>
<td>DEPTH OF THERMOCLINE - three digits to record depth of the thermocline or mixing layer in meters.</td>
</tr>
<tr>
<td>39</td>
<td>3</td>
<td>Numeric</td>
<td>TRANSPARENCY - three digits for secchi disk depth in meters with one decimal place.</td>
</tr>
<tr>
<td>42</td>
<td>4</td>
<td>Symbolic/ Numeric</td>
<td>AIR TEMPERATURE - four digits for the temperature of the air at the surface. Negative temperatures should be preceded with a minus sign. Temperature is recorded in degrees Celsius with one decimal place.</td>
</tr>
<tr>
<td>46</td>
<td>4</td>
<td>Symbolic/ Numeric</td>
<td>WATER TEMPERATURE - four digits for the temperature of the water at the surface. Negative temperatures should be preceded with a minus sign. Temperature is recorded in degrees Celsius with one decimal place.</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>Numeric</td>
<td>SALINITY - four digits for salinity of the water at the surface in parts per thousand (PPT) with two decimal places.</td>
</tr>
<tr>
<td>54</td>
<td>2</td>
<td>Code</td>
<td>WAVE HEIGHT - a two-character code for wave height. Use Code No. 0362. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>56</td>
<td>2</td>
<td>Numeric</td>
<td>WAVE PERIOD - two digits for the average wave period in seconds.</td>
</tr>
<tr>
<td>58</td>
<td>23</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
</tbody>
</table>

VI-18
### EXHIBIT VI-C-5
#### Sample Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This number appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description forms which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No. EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;E&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>Alphanumeric</td>
<td>SAMPLE NUMBER - three characters assigned by the investigator to identify the sample. Samples should be numbered sequentially. Quality control samples such as field blanks, lab blanks or spiked matrix samples are identified by a sequential number prefixed by a letter as indicated below:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F -- field blank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L -- lab blank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T -- transport blank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S -- spiked matrix sample</td>
</tr>
</tbody>
</table>
## EXHIBIT VI-C-5 (Cont'd)
Sample Record

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>6</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>24</td>
<td>6</td>
<td>Numeric</td>
<td>SAMPLE DEPTH - depth in meters where the whole water sample was collected, with two decimal places. Surface microlayer samples should be coded as &quot;000000&quot;.</td>
</tr>
<tr>
<td>30</td>
<td>12</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>42</td>
<td>2</td>
<td>Code</td>
<td>GEAR TYPE - a two-character code to identify the general gear type used to collect the sample. Use Code No. 0376:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>05 -- Bottle (Niskin, Rosette, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>09 -- Pump (Plankton, Midwater, Airlift, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>99 -- Misc (Hand-Gathered)</td>
</tr>
<tr>
<td>44</td>
<td>37</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>Starting Column</td>
<td>Length of Field</td>
<td>Field Format</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This number appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description forms which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No. EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;F&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>Alphanumeric</td>
<td>SAMPLE NUMBER - three characters assigned by the investigator to identify the sample. Samples should be numbered sequentially. Quality control samples such as field blanks, lab blanks or spiked matrix samples are identified by a sequential number prefixed by a letter as indicated below:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F -- field blank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L -- lab blank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T -- transport blank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S -- spiked matrix sample</td>
</tr>
<tr>
<td>Starting Column</td>
<td>Length of Field</td>
<td>Field Format</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This number appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description forms which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;F&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>BLANK</td>
<td>BLANK.</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>Alphanumeric</td>
<td>SAMPLE NUMBER - three characters assigned by the investigator to identify the sample. Samples should be numbered sequentially. Quality control samples such as field blanks, lab blanks or spiked matrix samples are identified by a sequential number prefixed by a letter as indicated below:</td>
</tr>
</tbody>
</table>

F - field blank  
L - lab blank  
T - transport blank  
S - spiked matrix sample
EXHIBIT VI-C-6 (Cont'd)
Data Record 1

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>1</td>
<td>Code</td>
<td>TRANSMISSIVITY EQUIPMENT CODE - a one-character code to describe the type of equipment to measure transmissivity. Use Code No. 0094:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 -- Turbidometer, in JTU</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 -- Transmissometer, in Percent of Light Transmission over a 10 cm Path</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 -- Fluorometer; Suspended Solids Calibration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 -- Nephelometer, in NTU</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 -- Turbidometer in FTU</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 -- Transmissometer</td>
</tr>
<tr>
<td>38</td>
<td>3</td>
<td>Numeric</td>
<td>pH - three digits for pH, with two decimal places.</td>
</tr>
<tr>
<td>41</td>
<td>4</td>
<td>Numeric</td>
<td>DISSOLVED OXYGEN - four digits for dissolved oxygen (DO), in milligrams per liter, with two decimal places.</td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td>Code</td>
<td>DISSOLVED OXYGEN EQUIPMENT CODE - a one-character code to describe the type of equipment used to measure dissolved oxygen (DO). Use Code No. EPA-5:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 -- Winkler Titration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 -- Probe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 -- CTD</td>
</tr>
<tr>
<td>46</td>
<td>5</td>
<td>Numeric</td>
<td>TOTAL SUSPENDED SOLIDS - five digits for the concentration of Total Suspended Solids in milligrams per liter (MGL) with two decimal places.</td>
</tr>
<tr>
<td>51</td>
<td>4</td>
<td>Numeric</td>
<td>LIGHT EXTINCTION COEFFICIENT - four digits for the extinction coefficient with two decimal places.</td>
</tr>
<tr>
<td>55</td>
<td>14</td>
<td>Blank</td>
<td>BLANK.</td>
</tr>
<tr>
<td>69</td>
<td>4</td>
<td>Numeric</td>
<td>BIOLOGICAL OXYGEN DEMAND (BOD) - four digits for the BOD concentration (5-day) in milligrams per liter (MGL).</td>
</tr>
</tbody>
</table>
EXHIBIT VI-C-6 (Cont'd)
Data Record 1

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>4</td>
<td>Numeric</td>
<td>OIL AND GREASE - four digits for the concentration of oil and grease in milligrams per liter with one decimal place.</td>
</tr>
<tr>
<td>77</td>
<td>4</td>
<td>Numeric</td>
<td>SETTLABLE SOLIDS - four digits for the settling rate, in milliliters per liter per hour (ml/l/hour).</td>
</tr>
<tr>
<td>81</td>
<td>4</td>
<td>Numeric</td>
<td>VOLATILE SUSPENDED SOLIDS - four digits for concentration of Volatile Suspended Solids, in milligrams per liter (MGL), with one decimal place.</td>
</tr>
<tr>
<td>Starting Column</td>
<td>Length of Field</td>
<td>Field Format</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>Code</td>
<td>FILE TYPE - set to &quot;144&quot;.</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Alpha</td>
<td>SOURCE IDENTIFIER - a unique identifier, assigned by EPA, which appears on all records within the data set.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>YY</td>
<td>YEAR - the last two digits of the year in which the sample was taken. This number appears on all records within the data set.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Alphanumeric</td>
<td>SERIES NUMBER - An alpha (A-Z) or numeric (1-9) code assigned by the investigator. Used to identify subgroups of data within a survey (e.g., analytical groups, or data collected at different frequencies). This code is defined by the investigator in the data description form which accompany the data sets. If there is no series number, enter &quot;0&quot; (zero) in the field.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Code</td>
<td>SCAN ID - A one-character code used to indicate sampling frequency and collection period. Use a numeric code to identify the quarter when semi-annual or quarterly data were collected. Use an alpha code to indicate when monthly data were collected. Use Code No. EPA-10 (See Appendix C).</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>Code</td>
<td>RECORD TYPE - set to &quot;G&quot;.</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>BLANK</td>
<td>BLANK.</td>
</tr>
</tbody>
</table>
| 14              | 3               | Alphanumeric | SAMPLE NUMBER - three characters assigned by the investigator to identify the sample. Samples should be numbered sequentially. Quality control samples such as field blanks, lab blanks or spiked matrix samples are identified by a sequential number prefixed by a letter as indicated below: F -- field blank L -- lab blank T -- transport blank S -- spiked matrix sample
<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>1</td>
<td>Code</td>
<td>SPHERE - a one-character code to identify the sphere from which the data came. Use Code No. 0093. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>Numeric</td>
<td>REPLICATE NUMBER - a unique identifier for each replicate observation taken from the whole water sample. This number is assigned by the investigator.</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>Code</td>
<td>CHEMICAL CODE (1) - a ten-character code to identify the chemical measured; entries must be left-justified. (See Appendix A for list of ODES Chemical Codes.) For unlisted codes, contact the ODES Staff. Do not assign codes independently.</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>Code</td>
<td>MEASUREMENT CODE (1) - code to describe the units of the reported concentration (e.g., ppb, ppm, activity). Use Code No. 0377. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>Code</td>
<td>QUALIFIER CODE (1) - a one-character code to provide additional qualifying information about the measurement. Use Code EPA-3. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>Numeric</td>
<td>CONCENTRATION (1) - concentration of the chemical measured, in units indicated by the measurement code, up to four significant digits. Use the EXPONENT field to indicate desired decimal point position. Do not enter a concentration value of zero. Leave blank if necessary. (See Appendix A for an example of how to enter chemical concentration data.)</td>
</tr>
<tr>
<td>36</td>
<td>1</td>
<td>Symbolic</td>
<td>EXPONENT SIGN (1) - sign (+/-) of the exponent for the concentration entry. If exponent is 0, leave this blank.</td>
</tr>
<tr>
<td>37</td>
<td>1</td>
<td>Numeric</td>
<td>EXPONENT (1) - exponent of the concentration value reported. Use exponent to indicate decimal places.</td>
</tr>
<tr>
<td>Starting Column</td>
<td>Length of Field</td>
<td>Field Format</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>38</td>
<td>10</td>
<td>Code</td>
<td>CHEMICAL CODE (2) - a ten-character code to identify the chemical measured; entries must be left-justified. (See Appendix A for a list of ODES Chemical Codes.) For unlisted codes, contact the ODES Staff. Do not assign codes independently.</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>Code</td>
<td>MEASUREMENT CODE (2) - code to describe the units of the reported concentration (e.g., ppb, ppm, activity). Use Code No. 0377. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>49</td>
<td>1</td>
<td>Code</td>
<td>QUALIFIER CODE (2) - a one-character code to provide additional qualifying information about the measurement. Use Code EPA-3. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>Numeric</td>
<td>CONCENTRATION (2) - concentration of the chemical measured, in units indicated by measurement code, up to four significant digits. Use the EXPONENT field to indicate desired decimal point position. Do not enter a concentration value of zero. Leave blank if necessary. (See Appendix A for an example of how to enter chemical concentration data.)</td>
</tr>
<tr>
<td>54</td>
<td>1</td>
<td>Symbolic</td>
<td>EXPONENT SIGN (2) - sign (+/-) of the exponent for the concentration entry. If the exponent is 0, leave this blank.</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>Numeric</td>
<td>EXPONENT (2) - exponent of the concentration value reported. Use exponent to indicate decimal place.</td>
</tr>
<tr>
<td>56</td>
<td>10</td>
<td>Code</td>
<td>CHEMICAL CODE (3) - a ten-character code to identify the chemical measured; entries must be left-justified. (See Appendix A for a list of ODES Chemical Codes.) For unlisted codes, contact the ODES Staff. Do not assign codes independently.</td>
</tr>
<tr>
<td>Starting</td>
<td>Length</td>
<td>Field</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Column</td>
<td>of Field</td>
<td>Format</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>1</td>
<td>Code</td>
<td>MEASUREMENT CODE (3) - code to describe the units of the reported concentration (e.g., ppb, ppm, activity). Use Code No. 0377. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>67</td>
<td>1</td>
<td>Code</td>
<td>QUALIFIER CODE (3) - a one-character code to provide additional qualifying information about the measurement. Use Code EPA-3. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>68</td>
<td>4</td>
<td>Numeric</td>
<td>CONCENTRATION (3) - concentration of the chemical measured, in units indicated by measurement code, up to four significant digits. Use the EXPONENT field to indicate desired decimal point position. Do not enter a concentration value of zero. Leave blank if necessary. (See Appendix A for an example of how to enter chemical concentration data.)</td>
</tr>
<tr>
<td>72</td>
<td>1</td>
<td>Symbolic</td>
<td>EXPONENT SIGN (3) - sign (+/-) of the exponent for the concentration entry. If the exponent is 0, leave this blank.</td>
</tr>
<tr>
<td>73</td>
<td>1</td>
<td>Numeric</td>
<td>EXPONENT (3) - exponent of the concentration value reported. Use exponent to indicate decimal place.</td>
</tr>
<tr>
<td>74</td>
<td>2</td>
<td>Code</td>
<td>EXTRACTION METHOD CODE - a two-character code to indicate the method used to extract or digest the sample matrix and remove or isolate the chemical of concern. Use Code No. 0350. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>76</td>
<td>2</td>
<td>Code</td>
<td>SAMPLE CLEAN-UP CODE (1) - a two-character code used to indicate an additional step taken to further purify the sample extracts or digestates. Use Code No. 0350. (See Appendix C for a list of codes.)</td>
</tr>
</tbody>
</table>
EXHIBIT VI-C-7 (Cont'd)
Data Record 2

<table>
<thead>
<tr>
<th>Starting Column</th>
<th>Length of Field</th>
<th>Field Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>78</td>
<td>2</td>
<td>Code</td>
<td>SAMPLE CLEAN-UP CODE (2) - Clean-up Code (2) refers to the second cleanup procedure used during sample processing. Use Code No. 0350. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>80</td>
<td>2</td>
<td>Code</td>
<td>SAMPLE CLEAN-UP CODE (3) - Clean-up Code (3) refers to the third cleanup procedure used during sample processing. Use Code No. 0350. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>82</td>
<td>2</td>
<td>Code</td>
<td>INSTRUMENT CODE - a two-character code to identify the final chemical analysis method(s) used for analyzing the sample. The code should represent the final analysis method or combined methods as listed in Code No. 0350. (See Appendix C for a list of codes.)</td>
</tr>
<tr>
<td>84</td>
<td>3</td>
<td>Alphanumeric</td>
<td>SAMPLE NUMBER FOR ANALYTICAL BLANK - a three-character ID for the analytical blank associated with chemical concentrations reported on this record. This ID &quot;links&quot; this Data Record 2 to a particular analytical blank. This ID must match one of the Sample Numbers listed on the Sample Records for data from analytical blanks.</td>
</tr>
</tbody>
</table>
DESCRIPTION OF RECEIVING WATER DATA SUBMITTED TO ODES

Because of the many types and sources of data that may be added to ODES, it is of great value for users to know the goals and techniques of each sampling program. Much of this information, so valuable for interpretation of data, is not inherent in the data itself; it can only be supplied in narrative form. Each data submitter is therefore requested to provide a descriptive overview of the sampling program. The following questions are intended to indicate the important issues which affect a data set's use and its comparability to other data sets. The information supplied in response to these questions will help all ODES users choose data appropriate for their purposes. Answers the following questions are requested for each sediment pollution file submitted for addition to ODES. Please attach extra pages as necessary. If the requested information is available in annual or quarterly monitoring reports, these may also be submitted.

Data set ID#:__________ File Type: 144W
Submitter: ____________________________________________

Please give the name of an individual who can be contacted for additional information concerning this data set:

__________________________________________ (name)
__________________________________________ (organization)
__________________________________________ (address)
__________________________________________

Phone: ( ) _____ - ________________

1. If you have utilized the Series identification field to define subsets of your data, please provide a description of the code used and its definition. If more than one Series ID was used, provide information for all codes.

2. Please describe the goals of the sampling program.
3. Please describe the distribution of the sampling stations throughout the water body sampled. If stations were selected in order to characterize some particular feature (e.g., an effluent source), please identify this feature and present the rationale for the placement of stations in relation to this feature.

4. Please describe the gear, any special features of its application and techniques used for sample collection.

5. How were samples collected and handled during transportation? How were samples stored?

6. What component of the sample was analyzed (i.e., whole water, dissolved fraction, or suspended particulates)?

7. Please provide the following information on analytical techniques for each class of chemical compounds. (Please attach on a separate sheet.)

<table>
<thead>
<tr>
<th>Analyte Size</th>
<th>Sample Type and Preparation</th>
<th>Sample Preservation</th>
<th>Holding Time Range</th>
<th>Method</th>
<th>Instrumentation</th>
<th>Detection Limits</th>
</tr>
</thead>
</table>

8. Please provide the following information on the frequency of laboratory quality control checks and provide a copy of the results of any such analyses. (Please attach information on a separate sheet.)
9. If results submitted to ODES have been corrected for blanks or recovery response, please describe the manner in which this has been done.

10. Please describe any features of this data set which may affect its use to generally characterize environmental data.

11. Please describe the quality assurance/quality control procedures used to verify the correct coding and entry of data.

12. If any other types of data were collected concurrently which have been or will be added to ODES, please indicate the appropriate ODES file-type(s).

13. In what report or document can the raw data be found? How could an individual obtain a copy of the raw data?