SECTION 5: DATA MANAGEMENT SYSTEM

This section discusses essential data management system elements of the coastwide seagrass monitoring program including:

    1. Collection and compilation of the major datasets
        a. Habitat distribution and landscape features of beds.
        b. Species occurrence.
        c. Vegetation abundance and phenology parameters.
        d. Physiological parameters, tissue composition, etc.
        e. Abiotic site parameters (water, sediment, climatic, etc.)
        f. Non-seagrass biotic measurements (epiphytes, macroalgae).
    2. Custodial responsibility for databases
      (Storage and QA/QC maintenance of monitoring data; designated
        custodial agencies)
    3. Data clearinghouse and web-based linkage of databases
      (How datasets are distributed/accessed on a network or the Internet)
    4. Database format and software considerations (GIS, database programs,
        data standards, etc.)

Basic to this Plan is the design and establishment of a multifunctional data management system (DMS) for compiling, storing, maintenance and distribution of seagrass monitoring datasets. As described in the previous section, the datasets include seagrass distribution maps and landscape data (mainly vegetation patterns and areal coverage), quantitative vegetative and habitat ecological parameters, water and sediment quality data, and other ancillary coastal and hydrographic data. A basic prerequisite is that all monitoring data must be spatially referenced (georegistered) when collected, thus making it amenable to storage or analysis in a GIS database. Differential GPS should be the main method used if possible; otherwise sites must be located precisely on a USGS 7.5' map.

The environmental and biological data collected through surveys, monitoring programs, or by special studies need to be compiled into an information management system for ready access and analysis of data using a clearinghouse approach. In order to accomplish this, a specially maintained DMS will be developed and established by the three State resource agencies (TPWD, TGLO, TCEQ). Although monitoring data may be sampled and collected by other groups, particularly research scientists, these State entities are identified as the appropriate custodians for coordinating storage of the identified coastal monitoring information. This arrangement can best assure quality control and upkeep of relevant datasets. A clearinghouse system should be designed specifically to provide convenient access and integration of the data by state or federal agency managers and regulators, and research scientists.

Data Management Issues for the Seagrass Monitoring Program:

1.   The Monitoring Workgroup agreed that management of seagrass monitoring data is an appropriate function of the three state agencies with coastal resource management and conservation responsibilities: TPWD, TGLO, and TCEQ. While other entities such as university researchers, contractors for agencies (state or federal), estuary programs or nongovernmental organizations might collect such data, the function of compiling, storing, and distribution of such data should be centralized under the three state agencies to assure format, consistency, quality control, and certification of the datasets.   This ultimately meets the management objectives of these agencies for monitoring seagrass, and also provides the public and other government agencies access to the datasets.

2. The three agencies serve as custodians for specific types of datasets based on their respective missions.   TPWD will continue compiling and maintaining seagrass distribution maps and landscape data, including digital photography and imagery. To the extent possible, it will maintain the necessary biological database containing ancillary information on seagrass vegetation parameters, other ecosystem flora and fauna, and estuarine biological processes. GLO will continue to maintain coastal infrastructure and base map datasets as used for Coastal Management and Oil Spill Response Programs. GLO and TPWD will leverage State resources and coordinate in obtaining future coastal aerial phatography datasets to monitor seagrass distribution. TCEQ will maintain coastal water and sediment quality monitoring data, as part of their Surface Water Quality Monitoring Program.   USGS water data and various NOAA datasets could also be included.   Although not previously mentioned, TWDB maintains hydrologic and tide datasets for bays and estuaries.

3. All datasets will be in a georegistered format, such that they can be easily manipulated and analyzed with GIS software (assume ESRI/ArcInfo compatible). Other general guidelines may be developed concerning database format to be followed, parameter fields, etc.   The question of QA/QC requirements for compilation and maintenance of water quality/chemical parameters or seagrass biological datasets must be addressed. What criteria must be met for such data to be certified for inclusion in the TNRCC or TPWD databases? In the case of monitoring data collected by research scientists, a standardized quality assurance system must be followed prior to conducting sampling surveys.   TPWD and TCEQ will define criteria that must be met in order for data to be included in their databases, irrespective of source.

4. An efficient, user-friendly distributed clearinghouse system will be developed to allow remote access to these custodial datasets. This could necessitate one agency assuming major responsibility for the server system development, with Internet links to other sites (at a minimum TPWD, TGLO, TCEQ and TWDB) where data resides for downloading.

Development of Prototype DMS and Clearinghouse

Under this plan, the TGLO, TPWD, and TCEQ support developing an integrated clearinghouse and DMS. TPWD GIS staff have already begun preliminary evaluation of a coastal monitoring data clearinghouse to perform seagrass assessments, and linkage of databases through a web-based server system. The relevant seagrass databases are actually seen as parts of major existing databases in the two agencies, TPWD and TCEQ. The TCEQ already maintains the TRACS Database for surface water quality monitoring (SWQM), and TPWD is organizing its seagrass distribution maps and imagery datasets into a departmental Resource Information System (RIS).

The Steering Committee proposes the following distributed data model system to access seagrass monitoring data. In this model, water quality data housed at TCEQ, seagrass data from TPWD, and other coastal data from TGLO, will be combined at the application level. For such a system to be successful, data providers must agree on data formats, access protocols, documentation and quality. Implementation of technologies such as ESRI's ArcIMS will require that participants coordinate closely to insure maximum interoperability and compatibility of the data at the application level. Additionally the data management system will be multi-tiered, supporting a wide range of end user applications from desktop GIS and statistical analysis packages to Web browser based data query tools. It is anticipated that most users will access data through the Internet as well as local ESRI trusted clients. Database access may be provided via ArcIMS map services, data extraction applications, FTP access, data mirroring, or through ODBC access.

Following from above, how should specific datasets be housed and shared? For example, it may be best to provide seagrass coverages through ArcIMS map services rather than through FTP access. Alternatively, if users require raw data for spatial analysis, shape files or coverages could be provided through FTP or through data streaming (feature map services) or ArcIMS extract servers. Water quality data may require replication or access through proprietary TCEQ applications and-or ArcIMS map services.

Data providers will be required to provide FGDC compliant (Federal Geographic Data Commission) metadata for all included data sets (FGDC 2003). Such metadata must also be easily accessible. Accordingly, data custodians will consider providing metadata through establishment of metadata servers or by providing metadata to a central repository, such as a seagrass-specific metadata server hosted by TPWD or another data custodian.

As a proof of concept, TPWD GIS staff have developed a web-based implementation of a distributed data-sharing tool using ESRI's ArcIMS technology. A map service was created for internal use that combines TPWD's seagrass data with SPOT 10m imagery and digital orthoquad maps (DOQs) served from TNRIS, as well as a water quality map service from TCEQ. The application allows the user to create dynamic maps of seagrass while accessing water quality data from TCEQ that may be relevant to specific seagrass areas of interest. This pilot project clearly demonstrates both the

strengths and weaknesses of such a data distribution approach. Further discussion of operations development will be required by agency staff to address issues pertaining to data quality, security, and data access or downloading.

Future application development will focus on facilitating assessment and monitoring of actual or potential environmental impacts to seagrass. Accordingly, access to water quality information, both current and historical, will be critical to this effort. Specific end user needs must be addressed, as well as opportunities to leverage existing web-based applications such as TCEQ's Texas Surface Water Quality Viewer. Consideration of the end user needs will greatly determine the conceptualization, application development, and implementation of logical and physical data models for this project.