Interim Information Management System Plan

APPENDIX J:

INTERIM
Information System Management Plan
For SWAMP

December 2002

Prepared by:

Cassandra Roberts
Marine Pollution Studies Lab
Moss Landing, CA 95039
email: roberts@mlml.calstate.edu

Interim Information Management System Plan

Interim Information Management System Plan

I. PROGRAM DESCRIPTION

One major challenge in conducting a statewide monitoring effort is development of a unified data system. For instance, in many cases the participating SWAMP organizations have previously developed data management systems of their own, or for their own specific objectives. These systems vary in the types of data captured, the software systems in which they are stored, and the degree of data documentation. In order to meet the SWAMP Program goal of centralized data management, a cooperative information management system is necessary to ensure that the collected data can be shared effectively among participants.

Information management needs to occur on several levels. First, a process must be developed to ensure the quality, compatibility, and timeliness of the data each organization collects. Once collected and organized, it must be available in as timely a manner as possible to the Regional Board SWAMP staff and others for review, analysis and ultimately for interpretation. Ultimately, one of the major goals of the use of this information, once interpreted, is to make it available to other interested organizations and the general public. The SWRCB SWAMP Program is also in the process of creating and maintaining an official website for the SWAMP Program, upon which such reports and data and other information would become available.

Appendix J herein, the Interim SWAMP Information Management System (SIMS) Plan, documents and describes in detail the information management system (IMS) in that will support data capture and reporting during the initiation of SWAMP, although several elements are still being documented and finalized. The Interim SIMS Plan focuses on four major functions of the SIMS:

- The standard protocols each participating agency or laboratory will use to transfer data from their internal date generators to the SWAMP IMS.
- The process by which data will be submitted to the SWAMP data managers, including the path and quality control procedures the data will follow until it has been accepted.
- The technical specification (guidelines) of how the data will be organized in the SWAMP database.
- The milestones and mechanisms by which the data will be made accessible to project participants, other organizations, and the general public.

II. APPROACH TO INFORMATION MANAGEMENT

The Information Management System has several purposes, most importantly to provide a mechanism for sharing data among project participants. Data sharing is required if the SWAMP goal of producing an integrated hydrologic unit assessment of the State's surface waters is to be achieved. While this is the primary focus, the IMS has been developed in recognition that SWAMP represents an initial effort toward data standardization among regions, agencies and laboratories and that protocols adopted here may be later used for other data sharing purposes beyond this project. Thus, the system was designed to

Interim Information Management System Plan

be flexible to future adaptation. In addition, the system was constructed primarily to serve the RBS and technical committees, but it has also been designed to supply data to non-project scientists and the interested public.

The IMS will be based on a centralized data storage model. A centralized system was selected because SWAMP is an integrated project and the typical data user will be interested in obtaining synoptic data sets from discrete hydrologic units or large geographical regions of the state. The centralized system was also selected over the alternative of a distributed system linked through a server or series of FTP sites because sophisticated tools would need to be developed and implemented for users to access those sites. There is also valid concern over the difficulty of maintaining a linked-distributed system for an extended number of years. Current budget allocations make the centralized system a more achievable model for handling data in the SWAMP program

The centralized database will be developed using standardized data transfer protocols (SDTP) for data exchange and Data Entering/Editing Forms for field data and observations. The SDTP details the information to be submitted with each sample collection or processing element, the units and allowable values for each parameter, and the order in which that information will be submitted. They are necessary to ensure that data submitted by the participants are comparable and easily merged without significant effort or assumptions by the organization responsible for maintaining the centralized data system. Use of SDTP allows each participating organization to retain data they generate in their local data management system while providing a mechanism for data exchange among project participants and a means for populating a centralized database.

The SWAMP database will be organized through a relational structure. The central database will be called the Replicate Master and will contain a temporary and permanent side, which are further described in the Data Flow Section below. The relational structure involves use of multiple data tables linked through one or more common fields or primary keys. A relational structure allows data created at different times (e.g. lab data vs. field data) to be entered at the time of data production, minimizing the possibility of data loss. This relational structure also minimizes redundant data entry, by allowing data that are recorded only once (e.g. station location) to be entered into separate tables rather than to be repeated in every data record.

The data table structure of this database was designed around a sample driven model. One distinct feature of this database captures a “nominal” position of the station (lat/long) which is stored in the stations table while still capturing a “actual” position of each sample. This is important because many different organizations will be occupying a station at different times to collect different samples. An example would be one group collects water samples, another group would deploy and retrieve bivalves, while yet another would collect stream bioassessment information at a station. This database structure was also designed with surface water sampling in mind, however, it is also built to capture information collected at multiple depths in the water column more commonly observed in marine and freshwater lake sampling systems.

Interim Information Management System Plan

It is imperative that station failures are documented in the sample table of the database to insure that a value is not missing from the database but was indeed documented as not being sampled. An example would be a station not being sample because it was "dry". This will be further described in the Master Table Structure (Chapter V) to follow.

III. ROLES AND RESPONSIBILITIES

SWAMP is a cooperative effort among eleven organizations (SWRCB, nine RWQCBs, CDFG) plus numerous additional subcontractor labs which have limited experience working together. Effective implementation of the SWAMP Information Management System Plan requires clearly defined roles for participants. For the purpose of defining roles, there will be four types of participants in SWAMP:

- Data generators - Field crew leaders (Key Data Entry) and laboratory supervisors who will be responsible for compiling data their organization generates and entering the data into the Data Entering/Editing Forms or the SDTP tables.
- SWAMP IM Coordinator (SIMC)- Responsible for working with Data Generators and leading the Data Management Team (DMT) at Moss Landing Marine Laboratories to develop SDTP, and for creation and management of the centralized SWAMP database.
- SWAMP Data QA Coordinator (SIMQA) - Responsible for overseeing quality assurance during migration of completed datasets to permanent data in the SWAMP database.
- SWIM IM Coordinator- Responsible for accepting data from SWAMP, placing it in the SWRCB SWIM database, and transferring it to other EPA databases, such as STORET.

Data Flow (still under development)

Official submission to the database can occur in two ways, 1) by form entry and 2) by batch loading (Figure 3) using SDTP. These data will reside in the temporary side of the Replicate Master. Data will be considered draft as it is loaded into the database (meets statewide comparability criteria) and compared with the task orders and found to be complete. Completeness checks will be accomplished with coordination between the SIMC and the Regional Board Staff requesting the work to be done as they have the best understanding of the study design. Data is considered complete when all results are entered in the database for a specific sample. Any SWAMP participants can have access to this draft data by requesting a replicate from the SIMC.

Once the draft data is certified complete, it will be transferred to the permanent side of the Replicate Master and sent to the SIMQA (QA Officer replica) for quality assurance checks. Once the data is validated by the SIMQA it will be considered final data.

Following certification of all portions of the data by the SIMQA, the SIMC will submit the integrated across-state data set to be stored in a Manger's replica of the SWAMP database. The SIMC will be the point of contact for data requests about the integrated data set. The SIMC will also be responsible for

Interim Information Management System Plan

making the SWAMP data available to other data centralization functions such as the SWRCB SWIM database. The SWIM IM Coordinator (SWIMC) will be responsible for maintaining the current version of the SWAMP database within SWIM, and transferring it to other databases, such as STORET.

General Structure of Database

The SWAMP database currently contains 20 data tables (Figures 2 and 3). There are 10 entry level data tables and 10 permanent level data tables, both containing similar content. The main table is the Sample table, which includes a single data record for sample taken. Samples created can be 1) laboratory samples (lab generated), 2) analytical samples (field generated), 3) field observations or 4) field results. The Sample table includes all fields necessary to uniquely describe a sample. This sample is linked in a one:one or one:many relationship with all subsequent data tables. It is imperative that the *StationCode, SampleDate*, and *SampleTime* remain the same for all the field-generated samples, observations and results in order to link the information.

The combination of the fields *StationCode, EventType, SampleDate, SampleTime, SampleTypeCode, Duplicate, DepthSampleCollected, DistanceFromBank*, and *AgencyCode* will ensure that each record in the Sample table is unique. Sample records need to be linked with all results data and thus become the foundation of the database. The chemistry and toxicity results tables, all laboratory and analytical data are captured at the level of individual replicate, rather than in a summarized form. It is essential that the laboratories receiving samples be supplied with the information in this table for each sample.

Form Entry/Editing Protocols

Key Data Entry people (limited number per RWQCB) will enter field data into a replicate of the central SWAMP database data entry/editing forms provided to them by the DMT. Limited analytical data can also be entered through the form entry system. The DMT will provide training and support for use of these forms. The individual replicates will be synchronized with the central SWAMP database (Replicate Master). Recommended QA for form-entered data include double checking of data, or at minimum 20%, and range checks of the Field Results table. Data will next be submitted to the SIMC for synchronization to the Replicate Master and QA of data types.

Standardized Data Transfer Protocols

The data formats for the SDTP table submissions are detailed in App J Section C_SWAMPDataFormats. These data formats include Lookup lists that are required to use in order for the data to be loaded into the database. The DMT will work with analytical labs on an individual basis to make this process as seamless as possible. Fields for summary quality assurance information are also included. A detailed laboratory QA report will be required and addressed in detail in the SWAMP QAMP.

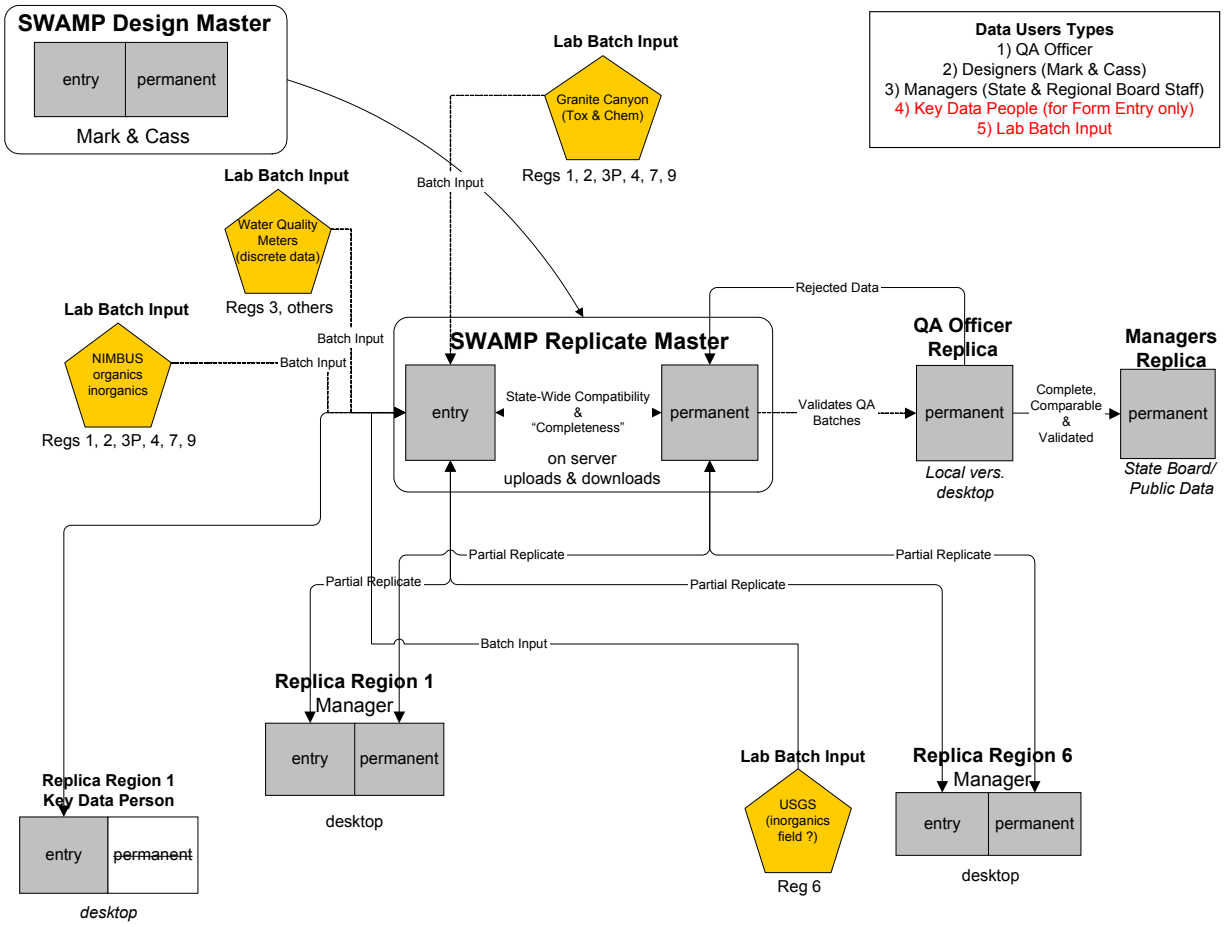
Upon receipt, the DMT will update a data submission log to document the data received from each submitting organization. The DMT will then initiate a series of error checks to ensure the data: 1) are within specified ranges appropriate to each parameter measured, 2) contain all required fields, 3) have encoded valid values from constrained look-up lists where specified, and 4) are in correct format (text in

Interim Information Management System Plan

text fields, values in numeric fields, etc.). If there are only a few, easily correctable errors, the DMT will make the changes. Changes will only be made with the consent of the data generator, with a list sent back to the data generator documenting the changes. If, there are numerous errors, or corrections difficult to implement, the DMT will send the data file back to the submitting organization with a list of necessary corrections. The submitting organization will make the corrections and resubmit the file to the DMT, who will subject the file to error checking once again. Each of these paths will be documented by the DMT as part of the submittal tracking process.

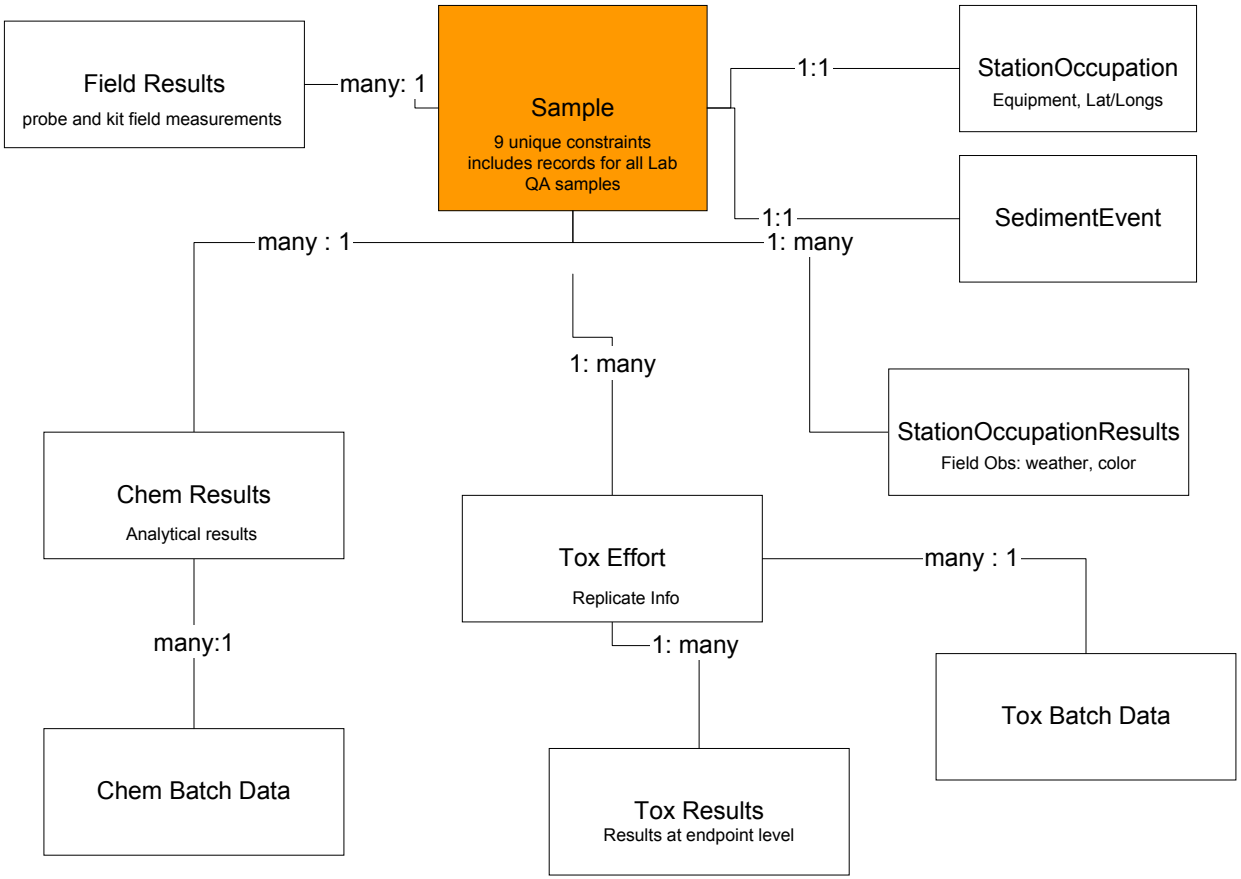
Interim Information Management System Plan

Figure 1: Flowchart of Statewide SWAMP Data Generators/Roles and Responsibilities



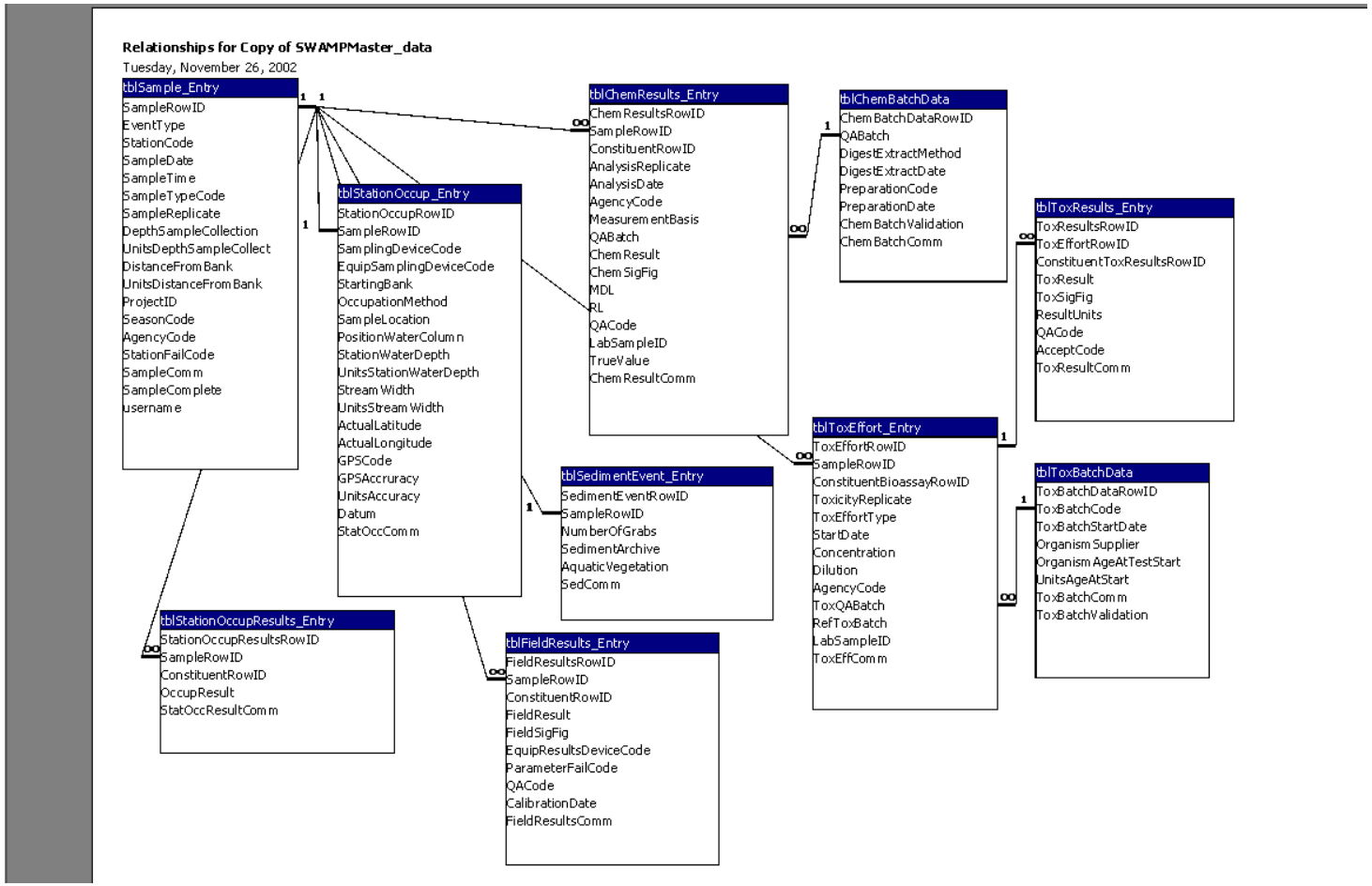
Interim Information Management System Plan

Figure 2: Outline of SWAMP Standardized Data Transfer Protocol Tables



Interim Information Management System Plan

Figure 3: Contents of SWAMP Standardized Data Transfer Protocol Tables



Interim Information Management System Plan

Data revisions (still under development)
Data can be revised in several ways depending on the stage of the data. When data is in the temporary side of the database, key data entry people will have the ability to revise data using the Data Entry/Editing Forms. When data is synchronized with the Replicate Master these edits will be committed to the database. It is important to note that the key data entry people or the DMT who make these edits bare the responsibility of making sure they are valid. Data deletions at this stage could have severe consequences to the database and should be used with care. Data being submitted using the SDTP can either be revised before or after it is submitted to the DMT.

Once the data is transferred to the permanent side of the Replicate Master, only the DMT, Designated Regional Board Staff and SIMQA will be able to edit it.

Schedule
The schedule for data submission varies by data type. Data collected in the field will be due first, while data produced through extensive laboratory analysis will be produced on a schedule consistent with nominal laboratory processing times. Key data entry people should provide their data to the DMT so that there is sufficient time for the DMT to resolve any data discrepancies and to ensure the data are in the proper format for the addition of the batch input data.

Data Sheets
To assist organizations in meeting the data entry forms and improve the efficiency of data input, the DMT has created a series of data sheets. These sheets follow closely with the data entry forms, however data gatherers are not required to use them.

IV. DATA ACCESS (still under development)

All measurement and supporting data gathered during SWAMP will be made available to all participating organizations and to the general public, though the schedule of availability and point of contact will vary by user. The different schedules reflect the differing levels of quality assurance and data documentation that will have been completed at various stages in the project.

The first location of data availability will be the SIMC, who will be responsible for the SWAMP database generated within the state. The SIMC will be free to distribute SWAMP data collected within the state, at any point after the data has approved as complete by the SIMQA and submitted to the final SWAMP database. Data released prior to having been transmitted and accepted by the SIMC and SIMQA should be identified as DRAFT data, not SWAMP data, because SWAMP quality assurance procedures will not yet have been performed. If Draft data is released, all filenames will include the word “DRAFT”. If hardcopies of Draft data are released, the pages must be stamped “Draft”. It is highly recommended that data released prior to its submittal to the SIMC be limited to organizations directly participating in the SWAMP project, rather than to outside agencies or the general public. Releases to the general public are not recommended until quality assurance has been performed by the SIMQA and metadata documentation is completed.

Nodes (planned for future implementation, if funding allows)

Interim Information Management System Plan

The second location of data availability will be the SWIMC, who will be responsible for integrating SWAMP and other state program data sets into the SWIM database. These data sets may be made available through other centralized or distributed databases, as coordinated by the SWIMC. It is the responsibility of the SWIMC to obtain express permission of the individual Program Managers prior to distribution of their respective program's data outside of the SWIM database.

Metadata

Each release of data to the public will include comprehensive documentation about SWAMP and the accompanying data sets. Referred to as metadata, this documentation will include database table structures (including table relationships) and lookup tables used to populate the fields in each table. It will also include quality assurance classifications of the data and documentation of the methodologies by which the data were collected.

A second type of metadata will document changes made to the data over time. As the data are used, we anticipate that errors will be found. As changes to the data are made, they will be documented in a file organized by date and data table. Including this file with each data download will allow users to reconcile potential differences in analysis output that result from using different versions of the data.

Metadata will follow guidelines from the Federal Geographic Data Committee, Content standard for digital geospatial metadata, version 2.0. FGDC-STD-001-1998 (FGDC 1998), including the Biological Data Profile and the Biological Names and Taxonomy Data Standards developed by the National Biological Information Infrastructure (NBII 1999). For tabular data, metadata that meet the FGDC content standard are contained by a combination of the SWAMP Data Directory and the SWAMP Data Catalog. For Arc/Info coverages, the metadata are in the .DOC file embedded in the coverage. This file stays with the coverage. When the coverage is moved to a public web site, it will be duplicated to an ASCII text file.

Interim Information Management System Plan

Table 1. Contact information for SWAMP information management.

SWAMP IM Coordinator (SIMC): Cassandra Roberts *Moss Landing Marine Laboratories* *Marine Pollution Studies Lab* 7544 Sandholdt Road Moss Landing, CA 95039 Ph. 831 771-4163 Fax 831 633-0128 roberts@mlml.calstate.edu Data Management Team (DMT) Mark Pranger *Moss Landing Marine Laboratories* *Marine Pollution Studies Lab* 7544 Sandholdt Road Moss Landing, CA 95039 Ph. 831 771-4176 Fax 831 633-0128 pranger@mlml.calstate.edu	SWAMP Data QA Coordinator (SIMQA) Not filled at this time--anticipated to be hired within 2003 if SWRCB funding allows.

For Future Implementation, if budget allows for development of nodes(these entities are currently providing extensive technical assistance into the database development):

Regional IM Coordinator (RIMC)- California Central Valley:
San Francisco Estuary Institute:Bruce Thompson
Department of Water Resources: Carl Jacobs

Regional IM Coordinator (RIMC)- - Southern California:
Larry Cooper
Southern California Coastal Water Research Project
7171 Fenwick Lane
Westminster, CA 92683
Ph. 714 894-2222
Fax 714 894-9699
larryc@sccwrp.org

Interim Information Management System Plan

V. MASTER TABLE STRUCTURES

For the following sections pertaining to Master Data Table Structures and Contents, the following document formats apply:

1. Consistent formats found throughout these sections:
- Table names will be underlined
- Field names will be *italicized*
- Values found within a field are **bolded**

2. Data will be given to SWAMP Information Management Team using standardized formats. It is the responsibility of the primary group submitting data (field crews or analytical labs) to format these data. The SWAMP database will use standardized look up lists where possible, per Section C below.

Interim Information Management System Plan

Section A. Relational Sample Data Table Structures

Samples

PURPOSE: The purpose of this table is to identify each sample generated in the SWAMP program. Samples created can be 1)laboratory samples (lab generated), 2) analytical samples (field generated), 3) field observations or 4) field results. It is imperative that the *StationCode, SampleDate*, and *SampleTime* remain the same for all the field-generated samples, observations and results in order to link the information.

The combination of the fields *StationCode, EventType, SampleDate, SampleTime, SampleTypeCode, Duplicate, DepthSampleCollected, DistanceFromBank*, and *AgencyCode* will ensure that each record in the Sample table is unique. Sample records need to be linked with all results data and thus become the foundation of the database. It is essential that the laboratories receiving samples be supplied with the information in this table for each sample.

BUSINESS RULES:

Field Generated Data:

EventType **FieldDescription** and *SampleType* **FieldObs** will be assigned to records that contain the visual observations of sample crews at site locations, i.e. water color, sediment composition, etc. These data will be recorded as results in the StationOccupResults table.

EventType **WaterChem** and *SampleType* **FieldMeasure** will be assigned to records that record all probe measurements and onsite/kit chemistry. These data will be recorded as results in the FieldResults table.

Field Generated Samples:

EventType will be assigned as **WaterTox_Chem**, **WaterChem, SedTox_Chem** or **Sed_Chem** as appropriate to the type of analysis being conducted. These data will be recorded as results in the ChemResults and ToxResults tables.

SampleType will be assigned as **grab** (used for single point of collection of water or sediment) or **integrated** (used for composited sample of either water or sediment).

Laboratory Generated Samples (QA samples):

Laboratory samples that are a modified field sample, such as matrix spike (**MSp**), will have all the same field information as the original sample except for the *SampleType,* which will be assigned the appropriate descriptor, such as **MSp** from luList_SampleType**.** The *AgencyCode, and ProjectID* for these modified field samples will remain as the original collecting agency and associated *ProjectID*.

Laboratory QA samples generated in the lab (**CRM, LCM, LabBlanks**, **Negative Controls**, etc.) will be assigned *EventType* **WaterChem**, **WaterTox**, **SedChem** or **SedTox** as appropriate for the laboratory and sample medium. The combination of the fields *SampleType* and a *StationCode* in the Sample Table will be used to distinguish

Interim Information Management System Plan

laboratory generated (QA data) from field generated results. Blanks and spikes do not have a *StationCode* associated with them so the Agency Code (from luList_AgencyCode) is used in its place. Certified Reference Material (**CRM)** or laboratory reference material (**LRM**) can have a *StationCode* with a pre-existing name that may help to identify the sample (ex. RMPMegaBass).

All of these QA samples will be associated with field samples by the same QABatch code. Formats for the *ProjectID* will be Fiscal Year + "SWLAB" (example being **01SWLAB**) if these QA samples are used in batches with combined fiscal years, use the later fiscal year in the *ProjectID*. *SeasonCode* should equal **Not Applicable**.

Sample Replicate:

Use this column only to identify true replicates created in the field. Default is zero. Laboratory replicates will be identified in the ChemResults table.

***Depth Sample Collected*:**

The depth the sample was taken is recorded in this field.

Water and sediment **grab** samples document the level from the surface in the water or sediment column at which the sample was taken. Units will also be recorded. *DepthSampleCollected* for water would be from the water's surface and recorded in meters (m) while depth collected for sediment would be from the sediment surface and recorded in cm. Samples created in the laboratory or during conditions where the depth was not recorded receive a value of –88.

For **integrated** samples collected from the same depth at different locations, i.e. points across a river, the actual sample depth should be recorded. This applies to both water and sediment samples. Integrated samples collected at multiple depths, i.e. samples integrated from the water column or sediment cores, should receive a depth of -88 and the actual depths of collection recorded in the *SampleComm* field.

***DistanceFromBank*:**

This field is used only when it is necessary to distinguish samples taken at different distances across a river, otherwise field samples receive a –88 for this field. The horizontal distance from the starting bank is captured in this field in meters. The starting bank, **LB** or **RB**, is recorded in the StationOccupation table. Samples created in the laboratory or during conditions where the distance was not recorded receive a value of –88.

SeasonCode:

Season descriptions (from luList_SeasonCodes) are provided by Regional Boards and can be unique to every region. This code will be used to distinguish sampling events where the same station is sampled multiple times throughout the year. *SampleDate* is not always a good sort field when multiple groups are visiting the same station on different dates. Data can be pooled together by the *SeasonCode*. The *SeasonCode* is a sort characteristic more than likely not applicable at the State level, but very helpful on a regional level. This field will also be critical when evaluating completeness of the data.

Interim Information Management System Plan

Laboratory-generated samples will receive a *SeasonCode* of **Not Applicable**.

Failures to obtain sample:

If a station was intended but not able to be sampled (for whatever reason), all appropriate sample types in the task order are still given a record with an appropriate *StationFailureCode* (see luList_StationFailureCode). This is critical to ensure completeness of a study design. This failure documentation needs to be filled out by the sampling crew. Subsequent laboratory result tables (ChemResults and ToxResults) will not be filled out.

TABLE SUBMISSION:

This table should be entered through the forms created in the SWAMP Access database by the Key Data People (KDP). The records can then be replicated up to the Master SWAMP database. The analytical labs will need to attach these sample records to each result before data can be uploaded to the database.

Field Name	Type	Required	Description
SampleRowID	Autonum	NA	Unique Row Identifier; generated from within Database
EventType	Text	Y	From Lulist_EventType
StationCode	Text	Y	SWAMP designated station label Xxx###### X=Regional Board, xx= Hydrologic Unit, ######= alpha-numeric code
SampleDate	Date	Y	Date of station visit expressed as dd/mmm/yyyy
SampleTime	Time	Y	Time of first sample taken; used for all samples collected at station (24 hour time)
SampleTypeCode	Text	Y	From LuList_SampleType
SampleReplicate	Number	Y	Used to identify the replicate number for a sample created in the field
DepthSampleCollection	Number	Y	Depth at which analytical sample was collected; units in meters; -88 if integrated sample
UnitsDepthSampleCollection	Text	Y	Units for water events in meters; sediment events in cm
DistanceFromBank	Number	Y	Required if multiple samples are taken along a transect; Horizontal distance from bank where sample was taken;
UnitsDistanceFromBank	Text	Y	Units in meters
ProjectID	Text	Y	Used to relate data to Task Orders; YYSWX### YY=Fiscal Year, SW=SWAMP, X=Regional Board, ###=Task Order
SeasonCode	Number	Y	From LuList_SeasonParent

Interim Information Management System Plan

Field Name	Type	Required	Description
AgencyCode	Text	Y	Agency creating the sample from luList_Agencies
StationFailCode	Text	N	Acceptable failure codes from luList_FailureCodes
SampleComm	Text	N	additional comments related to Sample
Username	Text	NA	Field used for Data Mangement

Section B1. Relational Field Results Table Structures

Station Occupation

PURPOSE: The purpose of the StationOccupation table is to document conditions under which each sample was collected. There can be a record for every **Grab**, **Integrated** or **FieldMeasurement** record generated in the Sample table.

BUSINESS RULES:
SamplingDeviceCode* vs. *EquipSamplingDeviceCode:

SamplingDeviceCode used to describe the method used for collecting the sample and *EquipSamplingDeviceCode* is used for identifying the specific equipment associated with the collection.

One exception is with the *SampleType* = **FieldMeasurement** where *SamplingDeviceCode* should equal **Field Equipment** and *EquipSamplingDeviceCode* should equal **Field Equipment described w/result**. The individual equipment associated with the each field measurement will be recorded in the FieldResults table.

TABLE SUBMISSION:
This table should be entered through the forms created in the SWAMP Access database. It will then be replicated up to the Master SWAMP database.

Field Name	Type	Required	Description
***StationOccupationRowID**	Text	NA	Unique Row Identifier; generated from within database
SampleRowID	Number	NA	Links record to the Sample Table
SamplingDeviceCode	Number	Y	Method used to collect sample from luList_ SamplingDevic
EquipSamplingDeviceCode	Number	Y	ID of the equipment used from luList_EquipSamplingDevi
StartingBank	Text	Y	Bank where distances are measured from; Left or Right Bank
OccupationMethod	Text	Y	Method of station occupation; "Walk In", "From Bridge",

Interim Information Management System Plan

Field Name	Type	Required	Description
			or report RV Name.
SampleLocation	Text	Y	Describe where sample was taken from Bank, Midchannel, Thalweg, OpenWater, Not Applicable, Not Recorded
PositionWaterColumn	Text	Y	Category of position in water column where sample was taken; subsurface, middepth, nearbottom or Not Applicable, Not Recorded
StationWaterDepth	Double	Y	Water Depth of the station where sample was taken, -88 if not recorded
UnitsStationWaterDepth	Text	Y	Units for Station Water Depth
StreamWidth	Double	Y	Stream Width at the Station where sample was taken, -88 if not recorded
UnitsStreamWidth	Text	Y	Units for Stream Width
ActualLatitude	Number	Y	Decimal Degrees 5 decimal places (NAD 83)
ActualLongitude	Number	Y	Decimal Degrees 5 decimal places (NAD 83); add negative sign
GPSCode	Number	Y	Id of the GPS equipment from luList_FieldEquipment
GPSAccuracy	Number	N	Record Accuracy of GPS
UnitsAccuracy	Text	N	Units associated with the GPS Accuracy
Datum	Text	Y	Datum used on GPS (NAD 83 recommended)
StatOccComm	Text	N	Additional comments regarding Station Occupation

Station Occupation Results

PURPOSE: The purpose of this table is to record all habitat descriptions (water, sediment, weather, etc.) associated with a station occupation. If both sediment and water are collected during a single station occupation, the information about both the sediment and water characteristics is recorded here.

TABLE SUBMISSION:
This table should be entered through the forms created in the SWAMP Access database. It will then be replicated up to the Master SWAMP database.

Field Name	Type	Required	Description
*** StationOccupResultsRowID**	AutoNumber	NA	Unique Row Identifier; generated from within database
SampleRowID	Number	NA	Links record to the Sample Table
ConstituentRowID	Number	Y	Unique ID to used to define sample
OccupationResult	Text	Y	Result of constituent
StatOccResultComm	Text	Y	Additional comments for Occupation Results

Interim Information Management System Plan

Sediment Event
PURPOSE: The SedimentEvent table documents all relevant information about each sediment sample collected. This table is only filled out if the *EventType* is **SedChem** or **SedTox_Chem,** and *SampleType* is **Grab** or **Integrated**.

TABLE SUBMISSION**:**
This table should be entered through the forms created in the SWAMP Access database. It will then be replicated up to the Master SWAMP database.

Field Name	Type	Required	Description
***SedimentEventID**	AutoNumber		Unique Row Identifier; generated from within database
SampleRowID	Number		Links record to the Sample Table
NumberofGrabs	Number	Y*	*Required if grab sampler was used, record number of grabs used to obtain sample, -88 if not applicable
SedimentArchive	Yes/No	Y	Was there an archive of sediment taken?
AquaticVegetation	Text	N	If yes, describe any aquatic vegetation
SedComm	Text	N	Additional comments

Section B2. Relational Laboratory Results Data Table Structures

Toxicity Effort (not yet completed)
PURPOSE: The purpose of the sediment toxicity results table is to document the results of all toxicity tests conducted in association with the project. Each record represents the results of an individual replicate for an individual species processed in a batch of replicates.

The concentration field is used only for reference toxicant test sample records.

Included in this table should be all replicates for reference or home sediments, negative controls or reference toxicants. The QACode field in this table refers to specific replicate in this table whereas, acceptability of the entire batch is documented in the ToxicityBatchInfo table.

Interim Information Management System Plan

TABLE SUBMISSION:

Field Name	Type	Required	Description
ToxEffortRowId	AutoNumber	NA	Unique Row Identifier; generated from within database
SampleRowID	Number	NA	Unique Row identifier linking to Sample Table
ConstituentBioassayRowID	Number	NA	Unique code
ToxicityReplicate	Number	Y	Sample Replicate for a specific test; WQ replicate is zero.
ToxEffortType	Text	Y	from luList ToxEffortType
StartDate	Date/Time	N	the date starting date of the test
Concentration	Double	Y	used for Reference Toxicity Tests
Dilution	Double	Y	The dilution factor expressed as a proportion
AgencyCode	Text	Y	Agency conducting the analysis - from luList Agencies
ToxQABatch	Text	Y	Batch number for samples processed as a batch; Xx##=X first letter of Genus; x first letter of species; ## sequential number for that species test for the year. If RefToxBatch= XxRT##
RefToxBatch	Text	N	Reference Toxicant Batch Number; XxRT##
LabSampleID	Text	Y	Unique Laboratory Sample Id assigned at Lab
ToxEffComm	Text	Y	Comments

Toxicity Results (not yet completed)

TABLE SUBMISSION:

Field Name	Type	Required	Description
ToxResultRowId	AutoNumber	NA	Unique Row Identifier; generated from within database
ToxEffortRowID	Number	NA	Unique Row identifier linking to Sample Table
ConstituentToxResultsRowID	Number	NA	
ToxicityResult	Number	Y	the numerical result of the test
ToxSigFig	Text	Y	Significant figures associated with the result
ToxQACode	Double	Y	the quality assurance code from luList_QA
AcceptCode	Text	Y	from luList_ToxTestAccept; refers to all replicates for the Test Endpoint
ToxResultComm	Double	N	additional comments

Interim Information Management System Plan

Toxicity Batch Data (not yet completed)

TABLE SUBMISSION:

Field Name	Type	Required	Description
ToxBatchDataRowId	AutoNumber	NA	Unique Row Identifier; generated from within database
ToxBatchCode	Number	NA	Unique Row identifier linking to Sample Table
ToxBatchSartDate	Date/Time	Y	Start of Tox Batch Test Date
OrganismSupplier	Text	N	Place or vendor of test organism
OrganismAgeAtStartTestg	Double	N	Units used to quantify AgeAtTestStart
UnitsAgeAtStart	Double	N	the quality assurance code from luList_QA
ToxBatchValidation	Text	Y	
ToxBatchComm	Double	N	additional comments

Chemistry Results

PURPOSE: The purpose of the ChemistryResults table is to document the analysis results for sediment chemistry, sediment grain size, and water column chemistry. Each record represents a result from a specific analysis for a particular parameter at a single station or a single QA sample. This table will also contain all supporting QA sample results.

BUSINESS RULES:

Constituent Codes:

The *ConstituentCode* in this table is made up of Medium, Method, Analyte, Fraction and Units for each result. Each field has an associated lulist for these values.

Results vs. TrueValue:

The reported result is the number gathered from the analytical instrument. The "True Value" is the actual concentration of the parameter in the reference sample. The purpose of the "True Value" is to facilitate the calculation of percent recovery.

AnalysisReplicates:

AnalysisReplicates are used to distinguish between replicates created in the laboratory. The result for each replicate will be numbered starting at zero, e.g. the result for the first replicate will have an *AnalysisReplicate* of **0** and the result for the second replicate will have an *AnalysisReplicate* of **1**, etc.

Interim Information Management System Plan

Matrix Spikes:

The reported *ChemResult* is the number gathered from the instrument and is the net amount recovered from the sample including the spike concentration. For spiked samples, the *TrueValue* is the total concentration of the parameter in the sample plus the spiked concentration. Percent recovery will be calculated by dividing the *ChemResult* by the *TrueValue* times 100. If the sample being used for the matrix spike requires dilution, the reported values for the *SampleType* **MSp** are the actual values from the instrument, not dilution corrected values. The dilution factor is reported in the *ChemResultComm* field.

Surrogate Corrected Data:

If surrogate-corrected results data are reported, the values will be qualified by a surrogate corrected value (**SCV**) in the *QACode* field. For QA samples, the surrogate compounds will be reported as the measured result and the true value for that compound.

Non-Detects (ND) and Detected not Quantified (DNQ):

When a result is below method detection level (MDL), a (-) MDL value for the analyte is reported in the ChemResults field and qualified as "ND" in the QACode field; likewise, when a result is between the MDL and the Reporting Limit (RL), the analytical result is reported and qualified as DNQ in the QACode field.

Lab Control Material (LCM**)& Certified Reference Materials (**CRM**)**

When reporting these samples the medium used is **labwater**, **tissue** or **sediment**.

QACode

QA codes can be found in the luList_QACodes. When multiple codes are needed, a comma should separate them.

TABLE SUBMISSION:

The combination of the fields SampleRowID, ConstituentRowID, AnalysisReplicate, AnalysisDate, AgencyCode, and MeasurementBasis will ensure that all records in the table are unique. Submission Formats: refer to attached Excel File worksheet ChemResults

Field Name	Type	Required	Description
***ChemistryResultsRowID**	AutoNumber	NA	Unique Row Identifier; generated from within database
SampleRowID	Number	NA	Unique Row identifier linking to Sample Table
ConstituentRowCode	Number	NA	Unique code used to describe a chemistry result

Interim Information Management System Plan

Field Name	Type	Required	Description
AnalysisReplicate	Integer	Y	Used to distinguish between splits created in the laboratory
AnalysisDate	Date/Time	Y	Date the sample was process on the analytical instrument
AgencyCode	Text	Y	Lab code from luList_AgencyCode
MeasurementBasis	Text		either "DW" for dry weight or "WW" for wet weight
QABatch	Text	Y	The code for all of the samples processed in the same batch
ChemResult	Double	Y	Numerical result; Non-detects are (–) actual detection limit; Detected but not Quantified (DNQ) results are reported as the actual value below the reporting limit
ChemSigFig	Text	Y	Significant figures associated with the result
MDL	Double	Y	Based on 40CFR136 definition
RL	Double	Y	Defined by individual laboratory;
ChemQACode	Text	Y	From luList_QACode
LabSampleID	Text	N	Unique sample identifier supplied by lab
True Value	Double	*	QA samples only. * Required for matrix spikes and certified reference materials, and surrogates
ChemResultComm	Text	N	Additional comments pertaining to chemistry result

Chemistry Batch Data

PURPOSE: This table contains supporting analysis data related to any particular batch of samples. Each record represents the conditions under which each batch was processed. The QABatch code should match the code reported in the ChemResults table.

TABLE SUBMISSION:

The QABatch fields will ensure that each record in the table is unique.
Submission Formats: refer to attached Excel File worksheet ChemistryBatchData

Field Name	Type	Required	Description
ChemBatchDataRowID	Autonumber	NA	Unique Row Identifier; generated from within database
QABatch	Text	Y	Code for all of the samples processed in

Interim Information Management System Plan

Field Name	Type	Required	Description
			the same batch
DigestExtractMethod	Text	Y	Digestion or extraction from lulist_Digestion
DigestExtractDate	Date	Y	The date the digestion or extraction was initiated; expressed as dd/mmm/yyyy
PreparationCode	Text	N	If applicable, supply Sample Preparation method used for samples
PreparationDate	Date	N	The date the preparation was started expressed as dd/mmm/yyyy.
ChemBatchComm	Text	N	Additional comments

FieldResults

PURPOSE: The purpose of this table is to store results for measurements obtained while in the field. These measurements can be taken from a water quality meter, probe, or kit.

BUSINESS RULES:

Missing Results:

All parameters for a given station will have a record in the database if one or more of the assigned parameters were taken. The *FieldResults* for the missing parameter will be assigned a **-88** in the database. If all parameters were unable to be sampled, the sample receives the appropriate StationFailure code in the Sample table.

***FieldResult* Non-Detects**

If a given parameter is below the detection limit of the instrument, a (–) detection limit of the parameter is reported instead and the QA code is non-detect (**nd**).

TABLE SUBMISSION:

Field Name	Type	Required	Description
FieldResultsRowID	Autonumber	NA	Unique Row identifier; Generated from within Database
SampleRowID	Autonumber	NA	Unique Row identifier linking to Sample table
ConstituentRowID	Autonumber	NA	Unique code used to describe a Field Measurement result
FieldResult	Double	Y	Result of probe or kit measurement;

Interim Information Management System Plan

Field Name	Type	Required	Description
FieldSigFig	Text	Y	Significant figures associated with the result
EquipResultsDeviceCode	Integer	N	ID of the equipment used to obtain the result
ParameterFailCode	Integer	*	Required if parameter was not collected; acceptable failure codes from luList FailureFieldResults
FieldQACode	Text	Y	Use codes from luList QACodes
CalibrationDate	Date	Y	Date of most recent calibration; expressed as dd/mmm/yyyy
FieldResultsComm	Text	N	Additional comments

Interim Information Management System Plan

Section C. Look Up Tables

Please refer to the Excel file within Appendix J with the filename:

"AppJSectionC_SWAMPDataFormats111802_1202",

This file is also located on the internet at:

http://www.swrcb.ca.gov/swamp/qapp.html

Interim Information Management System Plan

Section D. Field Data Sheets

The following pages contain the forms to be utilized by SWAMP Field personnel while conducting field activities such as sample collection of various media, and such as multiparameter probe measurements and other field data measurements.

Interim Information Management System Plan

SWAMP Station Occupation Results

***Station ID:**

***Project ID:**

***Sample Season:**

***Date:** M M D D Y Y Y Y

***Sample Time: (time of first sample)**

Arrival Time:

Departure Time:

PG: OF PGS

Entered Dbase

Event Type

Sample Type
FieldObs

SampleDepthCollection
-88

DistanceFromBank
-88

***Crew:**

Photos (RB & LB are assigned when facing downstream)
RB/LB/BB/ US/DS/##
RB/LB/BB/ US/DS/##
RB/LB/BB/ US/DS/##

***Precipitation**
dry
drizzle
rain
thunderstorm

Sea State (if applicable):
Calm
Rough
Choppy

***Sky**
clear
partly cloudy
overcast
fog

Wind Direction (from) / no wind = xx :
N W E S

***Habitat**
dry
non-wadeable stream
wadeable stream
wadeable concrete channel
standing water
other_____

Wind Speed (kts) :

***Water Color**
clear
green
yellow
brown
other

***Water Clarity**
clear
semi-clear
turbid

***Water Odor**
hydrogen sulfide
sewage
petroleum
mixed
none

***Sediment Color**
black
brown
gray
yellow
mixed
other

*** Sediment Composition**
course sand
fine sand
silt / clay
cobble
gravel
mixed
other

***Sediment Odor**
none
hydrogen sulfide
sewage
petroleum
mixed
other

Station Occupation Comments

Access key required Yes / No
Contact Info:

Gaging Station #:

***Elevation (ft or m) :**

*** required field; underlined fields used as primary keys in dbase**

SWAMP SOFDS 1/02/2003

Interim Information Management System Plan

SWAMP Shallow Water Sampling Event

Entered Dbase

*Station ID: | *Project ID: | *Sample Season: | PG: OF PGS

*Date: M M D D Y Y Y Y | *Sample Time: | Field Duplicates yes / no | SampleType= FieldBLDup

Event Type	Sample Type	*Sample Device:	*Occupation Method	*Sample Location
WaterTox_Chem WaterChem WaterTox	Grab Integrated	Indiv. Bottle by hand Indiv. Bottle by pole sampler Indiv. Bottle by bucket sampler Teflon Tubing Kemmer Sampler other_____	Walk In From Bridge R/V _____	Bank MidChannel Thalweg Open Water

*Starting Bank: LB / RB (facing downstream)

*GPS / DGPS	Lat Degrees	sec / hunds	Long Degrees	sec / hunds
Nominal			-	
*Actual dec degrees			-	

Accuracy (ft / m) | 5 decimals | *GPS Model: | 5 decimals | Datum: NAD 83 other _____

*Station Water Depth (m) : (point of sample) | *Stream Width (m) : (point of sample)

Samples Taken (# of containers filled)

	DepthCollect (m)	*Inorganics	*Bacteria	*Chl a/Boron	*TSS	*TOC /DOC	*Total Mercury	*Dissolved Mercury	*Dissolved Metals	*Total Metals	*Organics	*Toxicity	TIE	
SUBSURF/MID/ BOTTOM ABOVE/THERMO/ BELOW														
				Vol Filt: (ml)				Preservative time						
	Integrated; -88 in dbase; (describe depths in comments)							*Preserved	In lab	In lab				

Only enter if multiple distances are taken

Event Type: WaterChem | Sample Type: FieldMeasure

	*Depth Collect (m or f)	*Distance from Bank (m or f)	Velocity (fps / mps)	Air temp C	H2O temp C	pH	O2 mg/L	O2 %	Specific Conductivity (mS uS /cm)	Turbidity ntu	
SUBSURF/MID/BOTTOM ABOVE/THERMO/ BELOW											
SUBSURF/MID/BOTTOM ABOVE/THERMO/ BELOW											
SUBSURF/MID/BOTTOM ABOVE/THERMO/ BELOW											
		*Instrument:									
		*Calibration date									

Sample Comments: (failure of probe parameter should be marked as “probe failure”)

Meter Used:_____

Prop used: AA / Mini

____ rev. @ ____ (sec)

*** required field; underlined fields used as primary keys in dbase**

SWAMP WQFDS 1/02/2003

Interim Information Management System Plan

SWAMP Deep Water Sampling Event

*Station ID: | *Project ID: | *Sample Season: | PG: OF PGS | Entered Dbase

*Date: M M D D Y Y Y Y | *Sample Time: | Field Duplicates yes / no | SampleType= FieldBLDup

Event Type: WaterTox_Chem, WaterChem, WaterTox
Sample Type: Grab, Integrated
*Sample Device: Indiv. Bottle by hand; Indiv. Bottle by pole sampler; Indiv. Bottle by bucket sampler; Teflon Tubing; Kemmer Sampler; other_____
*Occupation Method: Walk In; From Bridge; R/V _____
*Sample Location: Bank; MidChannel; Thalweg; Open Water
*Starting Bank: looking downstream LB / RB

*GPS / DGPS	Lat Degrees	sec / hunds	Long Degrees	sec / hunds
Nominal			-	
*Actual dec degrees		5 decimals	-	5 decimals

*Station Water Depth (m) : point of sample | *Stream Width (m) : point of sample | Accuracy (ft / m) | *GPS Model: | Datum: NAD 83 other _____

Samples Taken (# of containers filled)

SUBSURF/MID/BOTTOM ABOVE/THERMO/BELOW	DepthCollect (m)	*Inorganics	*Bacteria	*Chl a/Boron Vol Filt (ml)	*TSS	*TOC /DOC	*Total Mercury	*Dissolved Mercury	*Dissolved Metals	*Total Metals	*Organics	*Toxicity	TIE	
								Preservative time						

Event Type: WaterChem
Sample Type: FieldMeasure
Only enter if multiple distances are taken

SUBSURF/MID/BOTTOM ABOVE/THERMO/ BELOW	*Depth Collect (m or f)	*Distance from Bank (m or f)	Velocity (fps / mps)	Air temp C	H2O temp C	pH	O2 mg/L	O2 %	Specific Conductivity (mS uS /cm)	Turbidity ntu	
	0.2										
	1.2										
	2.2										
	3.2										
	4.2										
	5.2										
	6.2										
	7.2										
	8.2										
	9.2										
	10										
	15										
	20										
	25										
		*Instrument:									
		*Calibration Date									

Sample Comments:

Prop used: AA / Mini

Meter Used:_____ _____ rev. @ _____ (sec)

Interim Information Management System Plan

SWAMP Sediment Sampling Event

PG: OF PGS

Entered in Dbase

***Station ID:**

***Project ID:**

***Sample Season:**

***Date:** M M D D Y Y Y Y

***Sample Time:**

Field Duplicates yes / no

SampleType= FieldBLDup

Event Type
SedTox_Chem
SedChem
SedTox

Sample Type
Integrated
Grab

***depth sample collection (cm)**

***SampleDevice:**
Core / Scoop / Other
PE / PC
_____
Sediment Grab type_____
(Include area of grab)

***Occupation Method**
Walk In
From Bridge
R/V _____

***Sample Location**
Bank
MidChannel
Thalweg
Open Water
Mixed

***Starting Bank:** LB / RB

***Station Water Depth (m) : point of sample**

***Stream Width (m) : point of sample**

Same As WQ: Yes

*GPS / DGPS	Lat Degrees	sec / hunds	Long Degrees	sec / hunds
Nominal			-	
*Actual dec degrees			-	

5 decimals 5 decimals

Accuracy (ft / m)

***GPS Model:**

Datum: NAD 83 other _____

Total Number of Grabs for Sample

(Required for Integrated Samples taken with a Grab sampler)

Aquatic Vegetation in sample area Yes No

Describe or ID if Yes

Samples Taken (# of containers filled)

*Organics	*Metals/ HgT	*Grain size TOC	*Toxicity	*Archive Chem	Selenium	

Comments

*** required field; underlined fields used as primary keys in dbase SWAMP SFDS 1/3/03**