Task 4: Data Management

Table of Contents

Task 4: Data Management

Introduction

This task involves the preparation of water quality monitoring data to ensure data quality and compatibility with TCEQ requirements.

Data Management Roles

Basin Planning Agency Data Management

The Planning Agency Data Manager is responsible for preparing data sets of quality-assured data in accordance with the Data Management Reference Guide (DMRG) for submittal to the TCEQ. Each data set submitted to the TCEQ should contain data collected under a **single** QAPP.

The Planning Agency Data Manager will review each data set using the Data Review Checklist (Exhibit 4A) and will prepare a Data Summary (Exhibit 4B) to be submitted with each data set. The Data Summary will contain basic identifying information about the data set, information regarding inconsistencies and errors identified during data verification and validation steps, and/or problems with data collection efforts. The Planning Agency will submit data to the TCEQ CRP Project Manager as a data package containing at least the *Data Loading Validator Report* generated from the SWQMIS User Test Environment (UTE), the *Data Summary* and the *data text files*.

TCEQ Data Management

TCEQ CRP Project Manager Role

The TCEQ CRP Project Manager is responsible for receiving and reviewing data submitted by CRP Planning Agencies to ensure consistency and conformance with Quality Assurance Project Plan (QAPP) specifications and contract requirements. The TCEQ CRP Project Manager reviews documents contained in the data package submitted by the Planning Agency. The TCEQ CRP Project Manager will conduct checks for data coding and the acceptability of reported values. Parameters codes and stations will be cross-referenced against the specifications as outlined in the QAPP.

The SWQMIS *Data Loading Validator Report* from the UTE will be compared to the appropriate QAPP, the *Data Summary*, and the Monitoring Activities section of the quarterly progress report (see Exhibit 1D), to reconcile expected versus actual results. One step in this review is assessing whether the required Limits of Quantitation (LOQ's) listed in the QAPP correspond to the minimum, or "less than", values submitted in the data set. When discrepancies exist without adequate explanation in the *Data Summary*, the TCEQ CRP Project Manager will ask the Planning Agency to reconcile the differences and make corrections as necessary. This may require additional review of the QAPP when methods, parameter codes, or required reporting limits are changed. When all errors and discrepancies have been reconciled, the revised dataset will be forwarded to and reviewed by the TCEQ Data Manager.

TCEQ Data Manager Role

The TCEQ Data Manager is responsible for receiving data sets from the CRP Project Manager, and tracking actions taken on each data set. The TCEQ Data Manager ensures that data are reported following instructions in the latest version of the Surface Water Quality Monitoring Data Management Reference Guide (TCEQ DMRG) and the SWQMIS Users Guide. The TCEQ Data Manager will run the SWQMIS Loader Verification Utility in the Production Environment of SWQMIS (SWQMIS PROD) and, if no errors are identified, will send the Data Loading Validator Report to the TCEQ CRP Project Manager for approval. If errors are found, the TCEQ Data Manager will notify the TCEQ CRP Project Manager. The TCEQ CRP Project Manager will then communicate with the Planning Agency to resolve these errors.

After the TCEQ Data Manager receives approval from the TCEQ CRP Project Manager, the dataset is then ready for loading into SWQMIS (PROD).

The TCEQ Data Manager also addresses requests for new codes, for new monitoring stations, or for corrections to existing data.

Preparing and Reporting Data

Formatting Data

To prepare data for upload into SWQMIS, two text files must be created. The text files, called the 'Events File' and the 'Results File,' are formatted using the data dictionaries defined in Chapter 7 of the DMRG. The text files are related by a unique identification number, the Tag ID. A Tag ID is assigned to each sampling event. The 'Events File' has a sampling event, with a unique date, time and place (depth & station). The 'Results File' contains each measurement that was collected and analyzed for that event; this leads to many results for each event.

Generic Sample/Events File Structure

Tag|Station Id|End Date|End Time|End Depth|Start Date|Start Time|Start Depth|Category|Type|Comment|Submitting Entity|Collecting Entity|Monitoring Type

Generic Results File Structure

Tag|End Date|Parameter|GT/LT|Value|LOD|LOQ|Qualifier Code|Verify Flag

When formatting the two files, the vertical bar is used to delimit the fields. This vertical bar is called a "pipe", and so the text files are commonly known as "pipe-delimited" text files. It is important to use the "pipe" because other commonly used delimiters, such as commas and slashes, may be used in other fields or may cause a problem when loading the fields into a database. Punctuation and hard returns are strongly discouraged as they can cause problems when exporting or aggregating data.The text files should only contain the data. Planning Agency's should omit header rows as these are problematic for the SWQMIS loader program.

Coding Data

(Tag_id, Tag Prefix, Submitting Entity, Collecting Entity, and Monitoring Type)

The *Tag ID* is a unique alphanumeric identifier assigned to each sampling event; this *Tag ID* links the event to the measurement values recorded in the Results file, in a one-to-many relationship. The *Tag Prefix* is the unique one or two letter code added to the beginning of the *Tag_ID* which identifies the Planning Agency that is reporting the data set.

A list of valid Tag Prefixes can be found in Chapter 5 of the DMRG.

Submitting Entity codes identify the agency listed in the QAPP as the entity responsible for submitting the data to the TCEQ (QAPP Entity).

Collecting Entity codes identify the organization responsible for collecting the data (field entity).

Monitoring Type codes identify the type of monitoring under which the reported data was collected.

Example:

> SR|LW|RT means the data was submitted under the Sabine River Authority (SR) QAPP and collected by the City of Longview (LW) without targeting any certain environmental condition (RT). The RT monitoring type code identifies that this data was collected during routine monitoring. The Tag prefix would be "J" which is assigned to the Sabine River Authority.

Lists of valid entity codes can be found in Chapter 4 of the DMRG.

Biological Electronic Data Reporting

There is no difference in format between biological data flat files and routine surface water quality monitoring data flat files. Partners and contracted monitoring entities must report biological data via pipe-delimited flat files.

A separate Tag ID should be assigned for each type of data collected during a biological sampling event. SWQMIS uses the Sample Event and Sample Set structure. This structure is a one-to-many relationship with one Sample Event (the entire biological monitoring event) containing multiple Sample Sets. Some Sample Sets are biological (Nekton Electrofishing, Nekton Seining, etc.) and some are non-biological (24 Hour Data and Routine Chemistry), but all are included in the Sample Event.

Each biological Sample Set must include the Parameter Code 89888 which identifies that Sample Set as containing biological data. The value selected for Parameter Code 89888 will be determined by the Sampling Category (see Table 12.1 in Chapter 12 of the DMRG). To see how parameters are grouped under each sampling category, refer to Chapter 6 of the DMRG Commonly Reported Parameter Codes for Biological Data.

BLOB Files

Reporting biological monitoring data also requires attaching Binary Large Object (BLOB) files to the SWQMIS Sample Events and/or Sample Sets. BLOB files must be named in a format that includes the station ID, water body name, sample end date, and type of file (e.g., 13486-GreensCreek-24May2013-HabitatTransectWorksheets). BLOB files can be attached at the sample event and sample set levels in SWQMIS by authorized TCEQ staff.

Typically, there are four BLOB file types for each SWQMIS Sample Event for biological data:

- monitoring summary information, including:

  - Aquatic Life Monitoring checklist
  - A site map of the sampling area of sufficient scale to annotate both the sampling reach and transect lines
- nekton voucher photos (unless actual voucher specimens are retained),
- Stream Physical Characteristics Worksheet with the transect data,
- Habitat transect photos,
- Any other additional file as discussed between the Planning Agency and the TCEQ CRP Project Manager.

For individual SWQMIS Sample Sets/Tag ID there can be more than one attachment. All photographs for one SWQMIS Sample Set/Tag ID should be combined into one document that includes descriptive information for each individual photograph (preferably a .pdf, but a Word document, or Power Point file will suffice).

Electronic data should include a README.txt file that includes a list of each BLOB file, a description for each BLOB file, and a designated place for the BLOB file to be attached (either the SWQMIS Sample Event ID, or specific Tag ID).

Please reference Chapter 12 of the DMRG or Exhibit 4D for additional information. An explanation of how biological data is tagged can be found in Exhibit 4C.

TCEQ CRP Project Managers will accept electronic files. Electronic files submitted should consist of the ASCII pipe-delimited flat files, plus any additional files specified by the TCEQ CRP Project Manager or contract.

Required Files for Biological Data Submissions:

- ASCII Pipe-Delimited EVENT Text File
- ASCII Pipe-Delimited RESULT Text File
- README.txt File
- BLOB Files

Event File and Results File Format

The generic format of both the Event and Result files is the same as any other non-biological dataset; the Results file will have one or more records associated with each Event record. In each biological sampling category (Nekton Electrofishing, Benthics, Habitat etc.), the Planning Agency must include one record for parameter 89888 which will categorize it as a biological dataset.

README File Format

The generic format of the README file is shown below:

File|Description|Tag ID

If the file will be attached to a Sample Event, please enter ‘Sample Event Level’ in the Tag ID field.

Composite Samples

Composite samples require entries in several additional fields in the Events file. These fields are Startdate, Starttime, Startdepth, Category, and Type.

Category must be one of four codes: T for time composites, S for space composites, B for both space and time composites, or F for flow-weighted composites.

The *Type* field must be a two-digit number (including leading zeros, if necessary) indicating the number of grabs, CN for continuous, or GB when the number of grabs is unknown.

Data Review and Validation

The data review and validation process combines the data validation and verification requirements defined in Task 2 with those outlined in this task. The major considerations for this process are evaluating data reasonableness and verifying that quality control limits were met. Data that do not conform to the specifications of the QAPP will not be submitted to the TCEQ. Each anomaly that causes a data point to not meet QAPP specifications will be described in the *Data Summary* (Exhibit 4B).

Data Review Checklist

The *Data Review Checklist* (Exhibit 4A) covers three main types of review: data format and structure, data quality review, and documentation review.

The *Data Format and Structure* section includes checks for required entries and formats. This section can be automated by developing a computer program that checks the database for outliers, other data anomalies, and some types of data transcription errors. These checks should address questions such as:

- Are there any duplicate Tag ID numbers and are Tag Prefixes correct?
- Do the sampling dates in the Results file match those in the Events file for each Tag ID?
- Are the codes for Submitting Entity, Collecting Entity, and Monitoring Type consistent with the entity and type of monitoring conducted?
- Are the sampling dates and times in the correct format with leading zeros (MM/DD/YYYY) and (HH:MM)?

See the Data Review Checklist, Exhibit 4A, for a complete list.

The *Data Quality Review* section includes checks specific to the acceptability of the data. This requires a more in-depth review of the data by personnel that understand the results of the laboratory analyses. This section includes checks that should address questions such as:

- Are the required reporting limits consistent with those in the QAPP?
- Have outliers been confirmed and a code entered into the *Verify_flg* field?
- Do the laboratory results appear reasonable and acceptable when compared to other corollary data and/or historical measurements?
- Are all sampling sites defined in the QAPP?
- Are all parameter codes listed in the QAPP?

See the Data Review Checklist, Exhibit 4A, for a complete list.

The *Documentation Review* section includes checks of the quality control information that is developed and provided by the laboratory or field staff.

This section includes checks that address questions such as:

- Are blank results acceptable as specified in the QAPP?
- Were control charts used to determine the acceptability of field duplicates?
- Were there any failures in field and laboratory measurement systems that were not resolvable and resulted in unreportable data?
- Have any anomalies been reported on the *Data Summary*?
- Was the laboratory's NELAP Accreditation current for the analysis conducted?

Validating Outliers

The TCEQ establishes a minimum and maximum value for many parameters above or below which a value is considered an "outlier." The min/max values represent a statistically derived range based on historical data (e.g., 1st and 99th percentile) except for in cases where the parameter is reported as either absent or present. Reported values that are found to be outliers should be checked against field and laboratory records to verify the correctness of the value as described in Task 2. The Planning Agency Data Manager should ensure that these outliers are flagged in the data set to show that they have been confirmed. If an outlier is not flagged, the SWQMIS data loader will find the anomaly and will not accept the data. All outliers must be flagged in the Results file by the inclusion of a "1" in the *Verify_flg* field.

A file containing all parameters and their min/max levels for outlier values in SWQMIS is available upon request.

SWQMIS Validations

Planning Agencies load datasets into the test environment of SWQMIS to ensure that data is formatted correctly and meets data reporting requirements as described in the DMRG and Table 3.1 of the *Surface Water Quality Monitoring Procedures, Volume 1: Physical and Chemical Monitoring Methods*. Instructions for loading data can be downloaded from the Texas Clean Rivers Program: Data, Forms and Map Resources webpage.

The SWQMIS Validator Report provides a list of errors in the data set relating to duplicate Tag IDs and records, data structure, outliers, significant figures, and rounding for certain parameters. Some errors can be over-written by the system and will not require a correction to the data files, while others are not. A full list of validations performed in SWQMIS is located in Exhibit 4E.

Once the data is free of errors, the SWQMIS Validator produces a report called the *Data Loading Validator Report* that contains the following:

- Date Range,
- Tag_ID Range,
- count of records in the data set,
- Parameter codes submitted with the data set and the number of times each was reported,

- the minimum and maximum values submitted for each Parameter code,
- stations for which data was submitted,
- number of events at each station,
- Submitting Entity / Collecting Entity / Monitoring Type codes in the data set,
- outliers in the data set,
- a comparison of data reported against the historical statistics at each station,
- the ten highest and lowest values reported for each parameter.

The report should be reviewed against the Monitoring Activities section of the quarterly progress report (see Exhibit 1D), the QAPP, and the *Data Summary* (Exhibit 4B) to ensure consistency in reported values and site locations.

Submitting Data to the TCEQ

Water quality data will be submitted to the TCEQ CRP Project Manager via email as a deliverable at least three times per year. The deliverable is due on March 1, August 1, and December 1 of each year. The March 1st deliverable will ensure that all the data collected through November 30th of the prior calendar year will be submitted to the TCEQ for use in the Texas Integrated Report (IR) in years when a "call for data" is issued. In these years, Planning Agencies should also coordinate with sub-participants to submit data by the March 1st deadline as well. The August 1st date will ensure data is submitted prior to the close of the fiscal year, and the December 1st date ensures the TCEQ has an updated water quality database through the close of the prior fiscal year.

Each data deliverables should include data following these guidelines:

- Data Deliverable, contains data through 11/30 – March 1st
- Data Submittal Progress Report – March 1, 2021
- Data Deliverable, contains data through 4/30 – August 1st
- Data Deliverable, contains data through 8/31 – December 1st

Along with the water quality data, the Planning Agency's Data Manager provides the SWQMIS *Loading Validator Report* (.pdf or .html file) with a *Data Summary* (Exhibit 4B) to the TCEQ CRP Project Manager. The *Data Summary* explains data discrepancies (e.g., missing measurements), describes field and lab issues, and indicates whether the Corrective Action Process has been initiated. *Corrective Action Status Reports* referencing Corrective Action Plans relating to the missing measurements will be submitted with the Progress Report after they are finalized. Additionally, when data cannot be submitted due to either sampling complications or QC failures, a running tally of this data should be maintained and reported on the Data Summary to determine overall percent completeness per data submittal period.

In the second year of the biennium, the Planning Agency will provide a Data Submittal Progress Report (Exhibit 1D) by March 1st in addition to the data deliverable. The Data Submittal Progress Report deliverable will summarize the extent to which data collected through November 30th, 2020 has been submitted to the TCEQ.

Other Data Management Considerations

Creating New Sampling Stations and New Codes

Requests for the creation of new monitoring stations will be handled via SWQMIS. Each Planning Agency has been given access to the Monitoring Stations Module and the Reports Module of SWQMIS. Specific instructions for requesting new sampling stations are included in Section 4.3 of the DMRG, Chapter 3. Additional helpful information for creating new stations or requesting changes to existing stations can be found in Chapter 3 of the DMRG. Planning Agencies should be sure to submit requests for new monitoring stations in the Production environment of SWQMIS. Requests for new stations submitted in the Test environment cannot be processed.

Requests for new codes relating to Submitting Entity, Tag Prefix, Collecting Entity and parameter codes should be coordinated with the TCEQ CRP Project Manager. The Monitoring Type codes have been set for specific data use purposes, and new codes are created only if there is a demonstrable need. Lists of existing codes are available in the DMRG. To request new codes, complete the appropriate Data Management form.

Data Correction Requests

If the Planning Agency finds that water quality monitoring data are in error in its database, this fact should be communicated to the TCEQ. A Data Correction Request is intended to correct erroneous data that has been uploaded to SWQMIS. If the data has not been loaded to Production, then a Data Correction Form is not required, and the Planning Agency can make revisions to the dataset and resubmit it. A SWQM Data Correction Request Form should be used to specify the applicable corrections. The forms should be submitted electronically to the TCEQ CRP Project Manager.

If a large number of errors or systematic errors are found which make use of the form unreasonable, contact your TCEQ CRP Project Manager for alternate electronic reporting methods.

Data Resubmittals

In certain cases, data that is identified as not conforming to the QAPP may still be submitted for inclusion to SWQMIS via a data resubmittal form. Data resubmittal forms are appropriate only in circumstances where the quality of the data is known and meets the data quality objectives of the project; the data resubmittal form is intended primarily to address minor deviations from the QAPP that otherwise do not negatively impact the acceptability, quality, or assurance of quality of the data. It is incumbent upon the appropriate Planning Agency personnel to complete the data resubmittal form, detailing the nonconforming data and describing the documentation which demonstrates the acceptability of the data—TCEQ will then evaluate the Data Resubmittal and associated documentation to determine if the data meet the program data quality objectives. Examples of documentation which may be required to accompany a Data Resubmittal include but are not limited to Corrective Action Forms and QAPP amendments.

Data on Basin Planning Agency Webpages

Public availability of surface water quality data is an important facet of the Clean Rivers Program. The Planning Agency should make current data accessible to the public and should include data collected by the TCEQ, if feasible. A disclaimer should be provided on the website if the complete dataset for the basin is not available. New data should be added to the Planning Agency webpage at least twice annually. The Planning Agency may choose to provide a link to the TCEQ water quality data viewer to satisfy this deliverable.

Distribution of Provisional Data

CRP data that have not been fully reviewed and validated by the TCEQ may be provided upon request at the Planning Agency's own discretion. If the Planning Agency opts to provide provisional data to interested parties a disclaimer **must** be included stating that the data are provisional. The disclaimer must indicate that the data have not undergone TCEQ review and validation and, therefore, may be subject to revision or exclusion from SWQMIS.

CRP Data Management Training

The TCEQ may conduct data management training workshops, as needed, in response to new data management procedures or requirements. These workshops will typically be held in conjunction with other CRP training.

Exhibit 4A
Data Review Checklist

EXHIBIT 4A

Data Review Checklist

This checklist is to be used by the Planning Agency and other entities handling the monitoring data in order to review data before submitting it to the TCEQ. This table may not contain all of the data review tasks being conducted.

Data Format and Structure	Yes, No, or N/A
Are there any duplicate Tag ID numbers in the Events file?	
Do the Tag prefixes correctly represent the entity providing the data?	
Have any Tag ID numbers been used in previous data submissions?	
Are TCEQ station location (SLOC) numbers assigned?	
Are sampling Dates in the correct format, MM/DD/YYYY with leading zeros?	
Are sampling Times based on the 24-hour clock (e.g. 09:04) with leading zeros?	
Is the Comments field filled in where appropriate (e.g. unusual occurrence, sampling problems, unrepresentative of ambient water quality)?	
Submitting Entity, Collecting Entity, and Monitoring Type codes used correctly?	
Do sampling dates in the Results file match those in the Events file for each Tag ID?	
Are values represented by a valid parameter code with the correct units?	
Are there any duplicate parameter codes for the same Tag ID?	
Are there any invalid symbols in the Greater Than/Less Than (GT/LT) field?	
Are there any Tag IDs in the Results file that are not in the Events file or vice versa?	
Data Quality Review	Yes, No, or N/A
Does the data conform to quality assurance specifications as outlined in the QAPP?	
Are “less-than” values reported at the LOQ? If no, explain in Data Summary.	
Have the outliers been verified and a "1" placed in the Verify_flg field?	
Have checks on correctness of analysis or data reasonableness been performed? Example: Is ortho-phosphorus less than total phosphorus? Are dissolved metal concentrations less than or equal to total metals? Is the minimum 24-hour DO less than the maximum 24-hour DO? Do the values appear to be consistent with what is expected for the site?	
Have at least 10% of the data in the data set been reviewed against the field and laboratory data sheets?	
Are all parameter codes in the data set listed in the QAPP?	
Are all stations in the data set listed in the QAPP?	
Documentation Review	Yes, No, or N/A
Are blank results acceptable as specified in the QAPP?	
Were control charts used to determine the acceptability of field duplicates?	
Was documentation of any unusual occurrences that may affect water quality included in the Event File Comments field?	

Were there any failures in sampling methods and/or deviations from sample design requirements that resulted in unreportable data? If yes, explain in Data Summary.	
Were there any failures in field and/or laboratory measurement systems that were not resolvable and resulted in unreportable data? If yes, explain in Data Summary.	
Was the laboratory's NELAP Accreditation current for analysis conducted?	

Exhibit 4B
Example Data Summary

EXHIBIT 4B

Data Summary

Dataset Information

Data Source: _____

Date Submitted: _____

Tag_id Range: _____

Date Range: _____

___ I certify that all data in this data set meets the requirements specified in Texas Water Code Chapter 5, Subchapter R (TWC §5.801 et seq) and Title 30 Texas Administrative Code Chapter 25, Subchapters A & B.

___ This data set has been reviewed using the criteria in the Data Review Checklist.

Planning Agency Data Manager:_____Date: _____

Comments

Please explain in the table below any data discrepancies discovered during data review including:

- Inconsistencies with LOQs
- Failures in sampling methods and/or laboratory procedures that resulted in data that could not be reported to the TCEQ (indicate items for which the Corrective Action Process has been initiated and send *Corrective Action Status Report* with the applicable Progress Report).

Dataset X contains data from FYXX QAPP Submitting Entity code XX and collecting entity XX. This is field and lab data that was collected by the collecting entity. Analyses were performed by the lab name. The tables on the following page explain discrepancies or missing data as well as calculated data loss.

Discrepancies or missing data for the listed Tag ID:

Tag ID	Station ID	Date	Parameters	Type of Problem	Comment/PreCAPs/CAPs
BP00339 BP00340 BP00341 BP00342	17000 17001 13230 13229	08/16/2011	00094	Calibration	Not Calibrating/Pre-Cap #P065, *Recent probe Troubleshooting did not deal with the problem, may need to be sent for repairs
BD00921	13272	09/20/2011	00094	Unknown	CE did not provide an explanation as to why (00094) was not reported. Pre-Cap #P066
BD00915	13272	06/21/2011	00300	Field	DO value illegible
BF00267	13103	05/18/2011	00400	Unknown	CE Did not provide an explanation as to why it was not reported.
BF00269	17596	07/26/2011	00300	Field	Value was reported in percent saturated instead of µS/cm.
BA00237 BA00238 BA00239	13560 13208 18792	07/27/2011	00400	Quality	Did not report pH because value was out of range. Pre-Cap #P062.
BA00240 BA00241 BA00242	13560 13208 18792	07/27/2011	Sediment Field Parameters	Field	No Field data reported. Pre-Cap# P071
BM00456	13664	07/26/11	31699 32211 32218	Shipping	No 48 hour parameters (E. coli, pheophytin/chlorophyll) due to late cooler

Data Loss

Parameter	Missing Data points out of Total	Percent Data Loss for this Dataset	Parameter	Missing Data points out of Total	Percent Data Loss for this Dataset
00094	5/27	19%	31699	1/27	4%
00300	2/27	7%	32211	1/22	5%
00400	4/27	15%	32218	1/17	6%

Note: There were a total of 35 events, which consisted of 32 water samples, 5 of which were dry, and 3 sediment samples. Also note inconsistencies in parameters reported due to the use of 2 labs for analysis. In addition, there are inconsistencies in parameters reported due to sites like X & X, where not all parameters are reported. The following are the calculations of how many values should be reported for these parameters:

*Field: 5 Dry events, 3 Sediment with no field, therefore total field parameters counted are 27 events. Why these weren't collected

*Lab events: 35 total – 10 Here – 4 field only = 21 – 3 sediment = 18 conventional – 1 partial conventional for ammonia, phos, Nit+Nit (13103) = 17 conventionals

*E.coli: 5 without HT 31704, Ecoli should be 18 (17+Station 13103)+4 Presidio=22 Lab1 + 5 Lab2= 27

*Chlorophyll: 17+5 Lab1= 22

*Pheophytin: 17 Lab2

Exhibit 4C
Understanding Biological Event Data Tagging

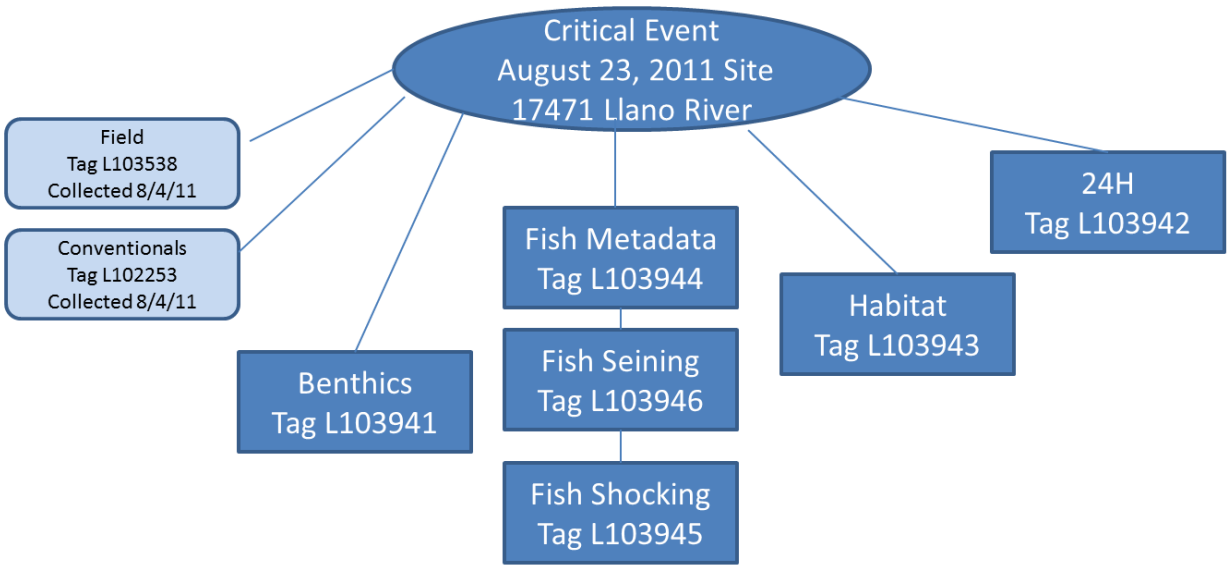
Biological Event Data

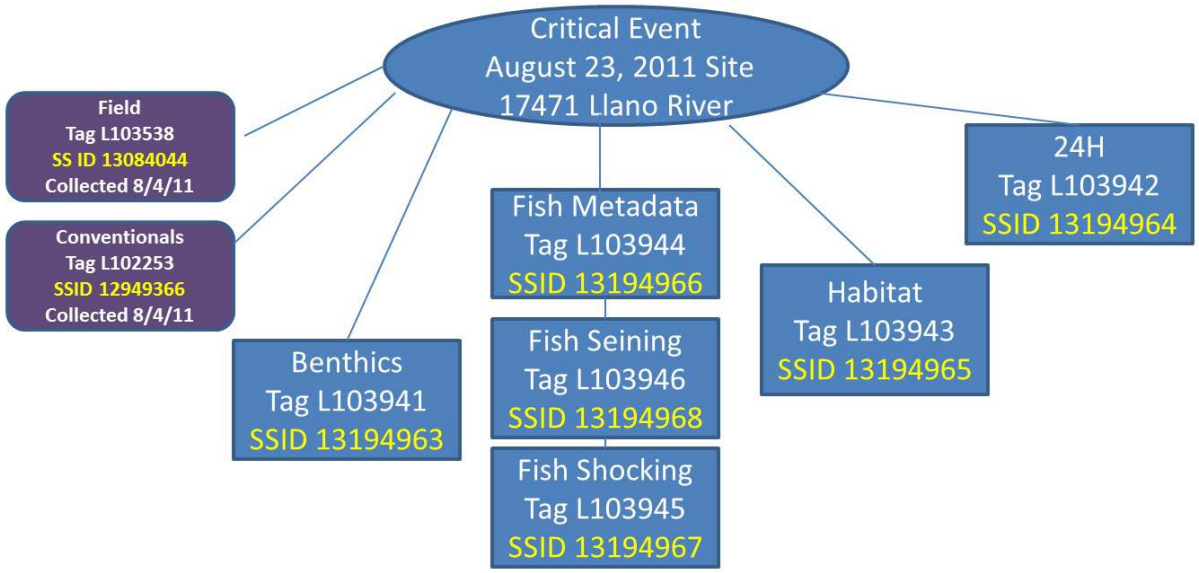
Details on how to submit biological data are in Chapter 12 of the Data Management Reference Guide

- Critical Event – August 23, 2011
- Site 17471 – Llano River near Junction
- Collection included: 24H, Benthic macroinvertebrates, fish, habitat
- The field and conventional data were collection 3 weeks earlier 8/4/11

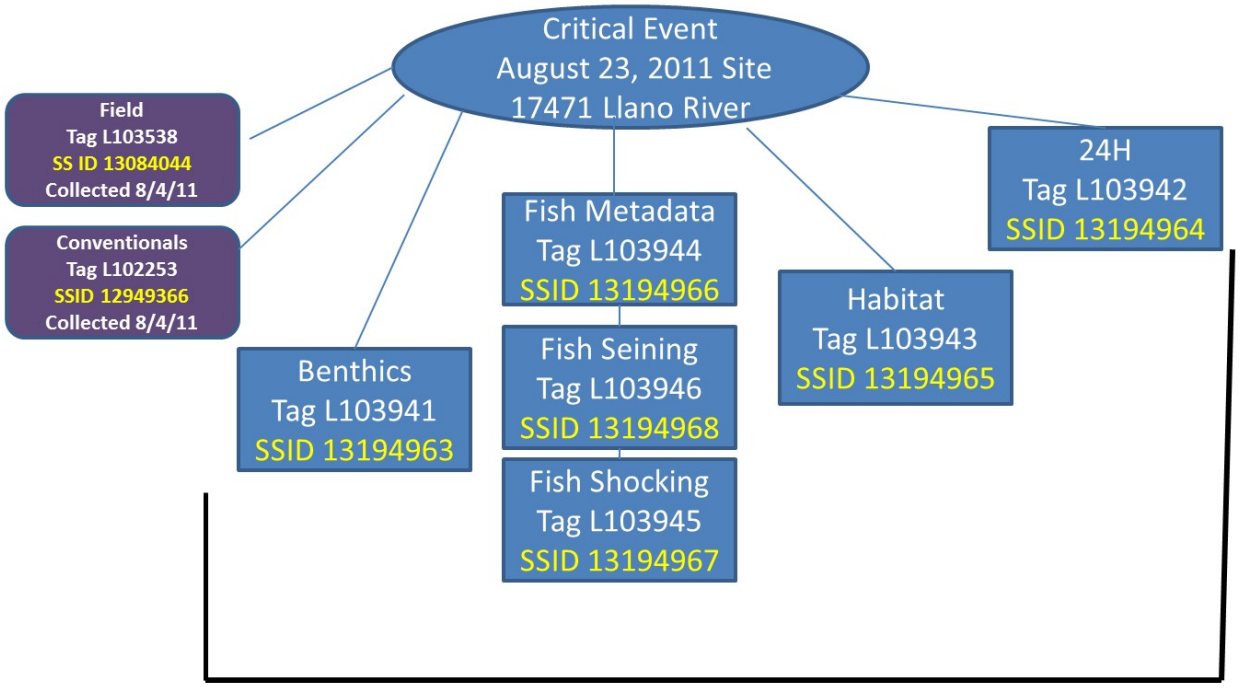
Terminology

- Tag
  - L103942
- In SWQMIS a Tag is called "RFA Tag #"
- Once the Tag gets loaded into SWQMIS, the system assigns a "Sample Set Id #"
  - Tag/RFA Tag # L103942 got assigned Sample Set Id 13194964





Once data are loaded to SWQMIS Production,
SWQMIS assigns a Sample Set Id # (**SSID**)to each Tag/RFA Tag Id #



Sample Event Id # 1351024

SWQMIS runs rules on data loaded into Production, and associates data collected by the same submitting entity, at the same site, and the same day +/- 2 days

The association results in a **Sample Event Id #**

ReadMe.txt

This file contains information about BLOB files. The ReadMe.txt file directs the TCEQ Data Manager where to attach each BLOB file.

	Filename	Description	Tag Id
1	17471-LlanoRiver ALM-2011 Summary Report 1.pdf	17471 Llano River Aquatic Life Monitoring Report for 2011 including sampling location map, checklist, and overall site information	**Sample Event Level**
2	17471-LlanoRiver ALM-23 Aug 2011 Fish Voucher Photos.pdf	17471 Llano River fish voucher photos from the August 2011 seining and shocking sampling efforts	L103944
3	17471-LlanoRiver ALM-23 Aug 2011-Habitat Transect Photos.pdf	17471 Llano River habitat transect photos from the August 2011 biological sampling event	L103943
4	17471-Llano River ALM-23 Aug 2011-Habitat Transect Worksheet.pdf	17471 Llano River habitat transect data and map from the August 2011 biological sampling event	L103943

* Sample Event Id # 1351024, assigned by SWQMIS after data loaded into Production

These include site maps, the Stream Physical Characteristics Worksheet with the transect data, fish voucher photos, or other biological data related images.
BLOB files must be named in a format that includes the station ID, water body name, and sample end date. BLOB files can be attached at the sample event and sample set levels in SWQMIS by authorized TCEQ staff.

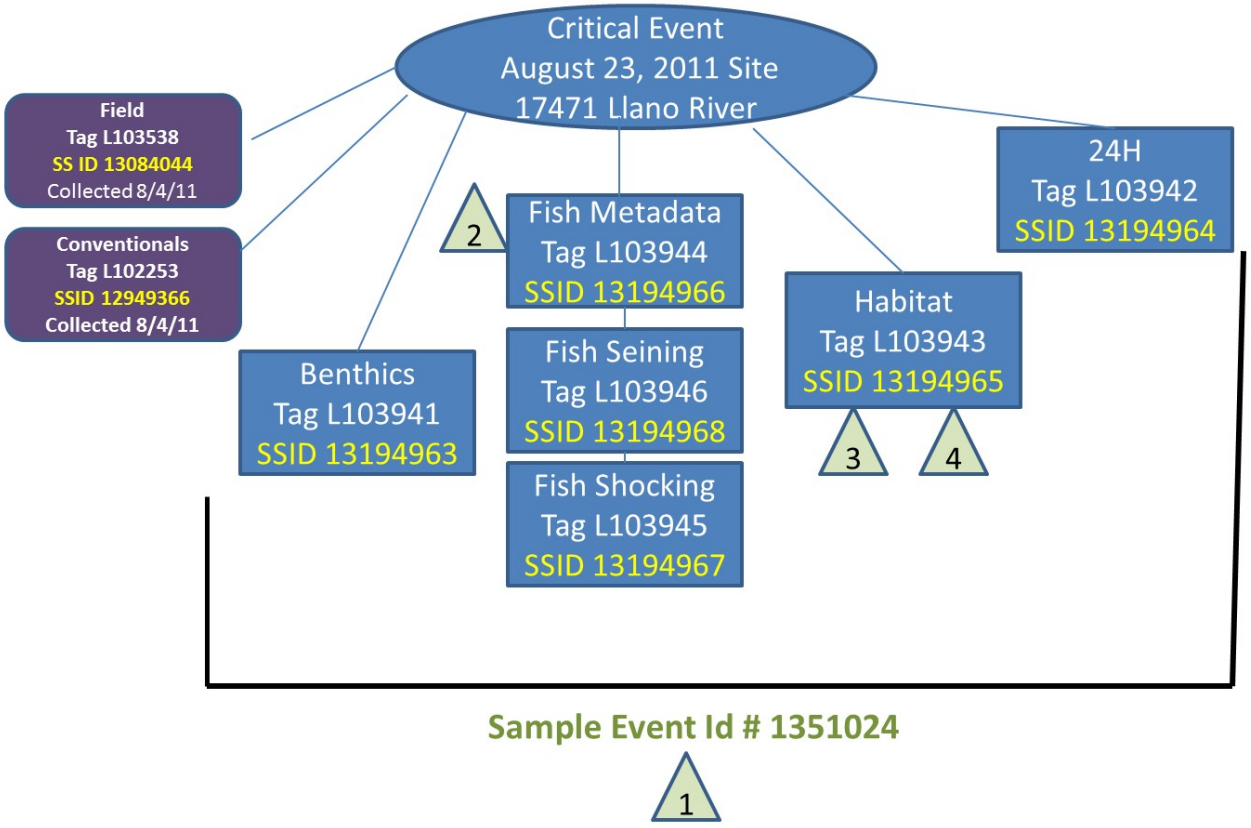

Critical Event
August 23, 2011 Site
17471 Llano River
Field
Tag L103538
SS ID 13084044
Collected 8/4/11
Conventionals
Tag L102253
SSID 12949366
Collected 8/4/11
Benthics
Tag L103941
SSID 13194963
2
Fish Metadata
Tag L103944
SSID 13194966
Fish Seining
Tag L103946
SSID 13194968
Fish Shocking
Tag L103945
SSID 13194967
Habitat
Tag L103943
SSID 13194965
3
4
24H
Tag L103942
SSID 13194964
Sample Event Id # 1351024
1

Exhibit 4D
Binary Large Object (BLOB) File Guidance and Examples

Exhibit 4D

Biological Data and BLOB File Guidance and Examples

This document provides guidance for **submitting** biological data. For guidance in the **collection** of the biological data consult the TCEQ *Surface Water Quality Monitoring Procedures, Volume 2: Methods for Collecting and Analyzing Biological Assemblage and Habitat Data* (RG-416) in conjunction with the TCEQ *Guidance for Assessing and Reporting Surface Water Quality in Texas*.

Guidance

General requirements and limitations for BLOB files:

- Files may be in a PDF, Microsoft Word, Microsoft Excel, or PowerPoint Format; PDF is preferred
- Each file should include in its heading monitoring station ID, station short description (when possible), and collection date
- Maximum size for each file is 15 MB
- Maximum number of files that can be attached per Event or per Sample Set is 5

Items below may be necessary as a part of biological data reporting. These items should be combined and submitted electronically to the TCEQ. See Appendix C of the TCEQ *Surface Water Quality Monitoring Procedures, Volume 2: Methods for Collecting and Analyzing Biological Assemblage and Habitat Data* (RG-416) for field forms and Chapter 12 (page 12-2) of the DMRG for additional information about BLOBs.

The following are required elements for submitting BLOB files:

ReadMe file (see page 4-4 and 4-5)

The type of information to include in a readme file

- File|Description|Tag ID
- Parameter code 89888 value
- The value description (see Chapter 12 of the DMRG for 89888 values)
- BLOB level (where should it be attached in SWQMIS? Event or Sample Set level?)

ALM Checklist:

- The ALM Checklist and Site Map can be combined into a single BLOB if needed to stay within the 5 file maximum attachments per Event or per Sample Set

Site Map(s):

- The map of the area where collection occurred

- Indication of the location of habitat transects

Voucher photos:

- All voucher photos should be combined into one document regardless of the method used to sample
- Each photo should be labeled with station short description, station id, species name, and collection date

Habitat transect photos

- All habitat photos should be combined into one document
- Photos should indicate the direction in which the photo was taken (i.e. left bank, right bank, upstream, downstream)
- Each photo should be labeled with station short description, station id, and collection date.

Habitat Transect Worksheet:

- The worksheet should be filled out clearly and completely. If not using the Habitat Transect Worksheet provided in the SWQM procedures, it should contain the same information.

ReadMe File

File Name	File Description	Tag ID
12830_MedinaRiver_Transects_3_18_14.pdf	12830 Medina River at Old English Crossing 3/18/14 Habitat Transect Data	SA15453
12830_MedinaRiver_HabitatPhotos_3_18_14.pdf	12830 Medina River at Old English Crossing 3/18/14 Habitat Transect	SA15453
12830_MedinaRiver_NektonPhotos_3_18_14.pdf	12830 Medina River at Old English Crossing 3/18/14 Seining and Shocking Voucher Photos	SA15452
12830_MedinaRiver_SiteMap_3_18_14.pdf	12830 Medina River at Old English Crossing 3/18/14 Site Map	
12830_MedinaRiver_ALM_3_18_14.pdf	12830 Medina River at Old English Crossing 3/18/14 Aquatic Life Use Information Data	
14929_SaldadoCreek_Transects_4_3_14.pdf	14929 Saldado Creek at Comanche Park 4/3/14 Habitat Transect Data	SA15487
14929_SaldadoCreek_HabitatPhotos_4_3_14.pdf	14929 Saldado Creek at Comanche Park 4/3/14 Habitat Transect Photos	SA15487
14929_SaldadoCreek_NektonPhotos_4_3_14.pdf	14929 Saldado Creek at Comanche Park 4/3/14 Seining and Shocking Voucher Photos	SA15485
14929_SaldadoCreek_SiteMap_4_3_14.pdf	14929 Saldado Creek at Comanche Park 4/3/14 Site Map	
14929_SaldadoCreek_ALM_4_3_14.pdf	14929 Saldado Creek at Comanche Park 4/3/14 Aquatic Life Use Information data	

Examples

ALM Checklist, 10210_DaysCreek_ALM_4_3_15.pdf

Nekton sampling event 1:

Minimum 15-minute (900 seconds) electrofishing: (Yes) No
Minimum 6 seine hauls (or equivalent effort to sample 60 meters): (Yes) No
Fish sampling conducted in all available habitat types: (Yes) No
If no, please describe why:

Benthic macroinvertebrate sampling event 1:

Indicate method(s) used:
Rapid bioassessment (5-minute kicknet or snags): X
Quantitative (Surber, snags, or dredge): ___

Habitat assessment event 1:

TCEQ habitat protocols: (Yes) No

Stream flow measurement event 1

Instantaneous flow measurement: (Yes) No
USGS gauge reading: Yes (No)

Nekton sampling event 2:

Minimum 15-minute (900 seconds) electrofishing: (Yes) No
Minimum 6 seine hauls (or equivalent effort to sample 60 meters): (Yes) No
Fish sampling conducted in all available habitat types: (Yes) No
If no, please describe why:

Benthic macroinvertebrate sampling event 2:

Indicate method(s) used:
Rapid bioassessment (5-minute kicknet or snags): X
Quantitative (Surber, snags, or dredge): ___.

Habitat assessment event 2:

TCEQ habitat protocols: (Yes) No
If no, flow,wetted channel width, photographs, description of bank conditions relative to first event, and description of canopy cover conditions relative to first event must be provided in this packet:

Stream flow measurement event 2

Instantaneous flow measurement: (Yes) No
USGS gauge reading: Yes (No)

Assessment Results (Optional)

Fish community index event 1:

Exceptional (High) Intermediate Limited

Fish community index event 2:

Exceptional (High) Intermediate Limited

Benthic macroinvertebrate community index event 1:

Exceptional (High) Intermediate Limited

Benthic macroinvertebrate community index event 2:

Exceptional (High) Intermediate Limited

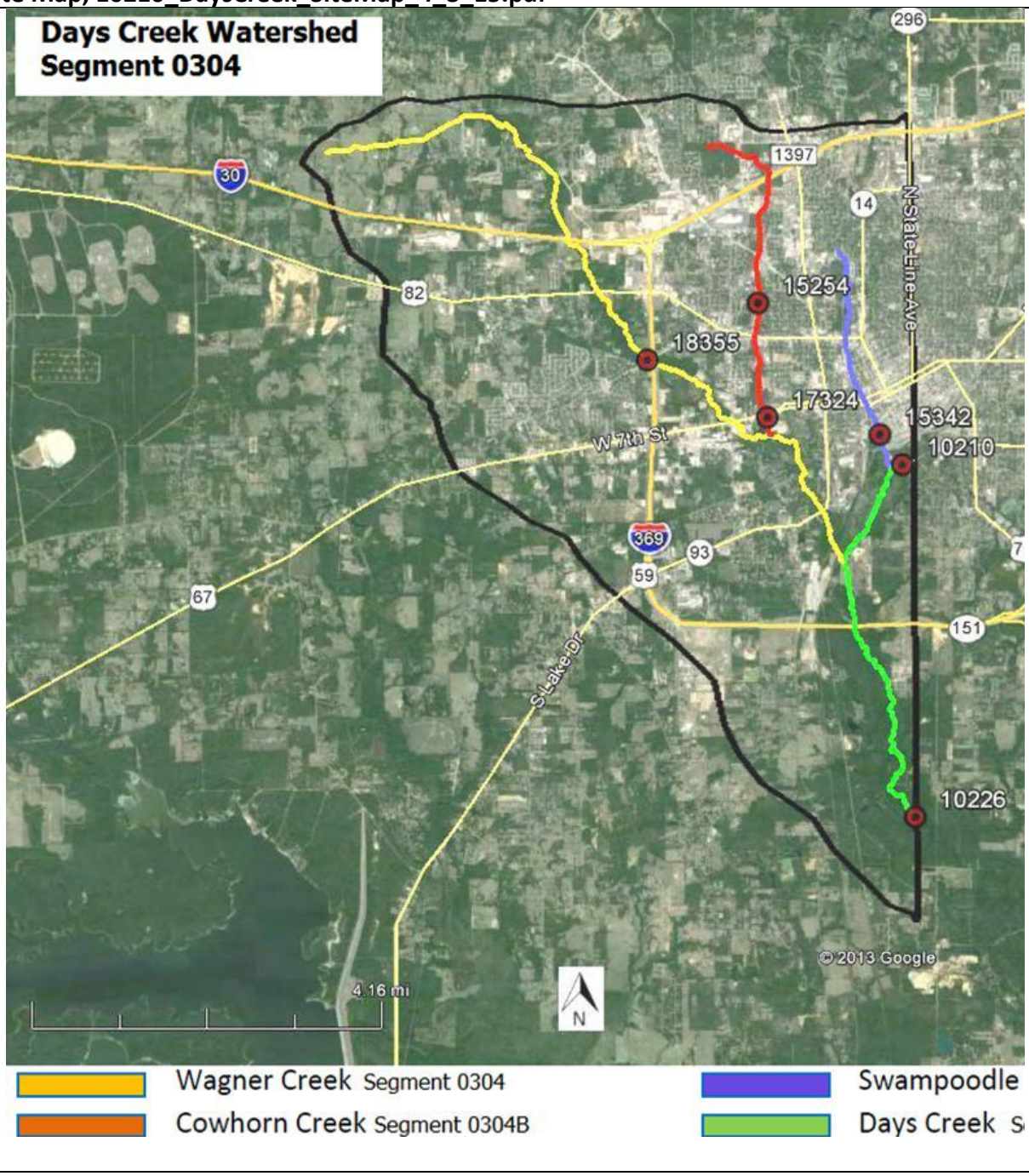
Habitat index event 1:

Exceptional (High) Intermediate Limited

Habitat index event 2:

Exceptional (High) Intermediate Limited

Site Map, 10210_DaysCreek_SiteMap_4_3_15.pdf



Voucher Photos, 20641_ColoradoRiver_NektonPhotos_5_15_13.pdf

Colorado River Above Lake Buchanan fish voucher photos
TCEQ ID 20641 05/15/2013

Pimephales vigilax

Lepomis cyanellus

Habitat transect photos, 20662_SanSabaRiver_HabitatPhotos_3_27_13.pdf

San Saba River at San Saba CR 340
TCEQ ID 20662

Habitat Transect Photos
03/27/2013

Upstream View

Left Bank View

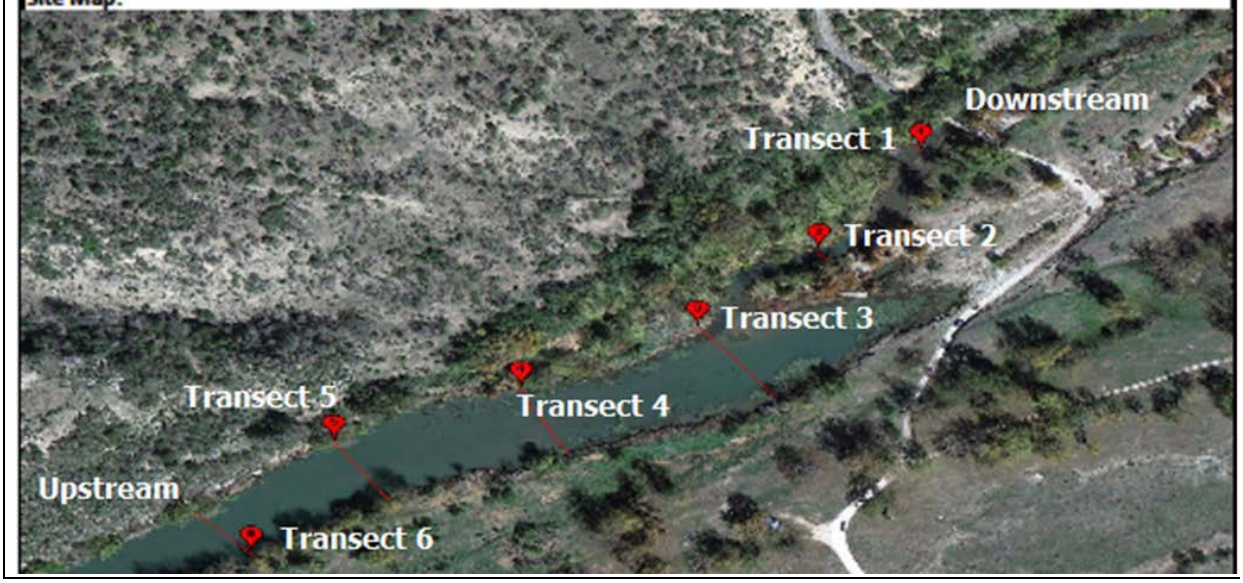
Transect 1

Right Bank View

Downstream View

Habitat Transect Worksheet, 20662_SanSabaRiver_Transects_3_27_13.pdf

Stream Physical Characteristics Worksheet								
Observers:	Djurecka, Dcowan, Jwoods,CPetri, FM			Date:	3/27/2013	Time:	11:00	
Weather Conditions:	Partly Cloudy, 45 deg F, N wind at 15 mph							
Stream:				Site ID:	20662	Segment:		
Location of site:	San Saba River at San Saba CR 340					Reach:	500 m	
Obs Stream Uses:	Recreation, Agriculture							
Stream Type:		perennial			intermittent with perennial pools			
Stream Bends:	4	Well Defined:	2	Mod Defined:	2	Poor Defined:	0	
Aesthetics	wilderness		natural		common		offensive	
Channel Obstr/Mods:	0					Number of Riffles:	1	
Channel Flow Status:	high		moderate		low		no flow	
Riparian Vegetation	Left	Right	Stream Flow:	31 cfs	Max Pool Depth:	3.5 m		
% Trees	15	8			Max Pool Width:	57 m		
% Shrubs	10	3	Transect placement relative to a fixed point:					
% Grassess&Forbs	27	36						
% Cult. Fields	0	0						
% Other	48	53						
Notes:								
Site Map:								



Habitat Transect Data

Date	3/27/2013					
Site	San Saba River at San Saba CR 340					
TCEQ ID	20662					
	Transect 1	**Transect 2**	**Transect 3**	**Transect 4**	**Transect 5**	**Transect 6**
Stream Type (Riffle,Run,Glide, or Pool)	RI	G	RU	P	P	P
Stream Width	23.8	12.4	24	57	57	50.5
Left Bank Slope	90	60	15	125	125	90
% Left Bank Erosion	15	25	40	75	80	80
Left Bank Width of Natural Buffer Vegetation	>20	>20	>20	>20	>20	>20
Right Bank Slope	75	25	3	35	90	25
% Right Bank Erosion	70	70	5	80	70	70
Right Bank Width of Natural Buffer Vegetation	>20	>20	>20	>20	>20	>20
% Tree Canopy	32.4	39.7	0.0	47.1	50.0	50.0
Dominant Substrate Type (1=clay,2=silt,3=sand, 4=gravel,5=cobble, 6=boulder,7=bedrock,8=other)	5	2	3	2	7	2
Stream Depth at Point 1	0.10	0.02	0.03	0.28	0.10	0.50
Stream Depth at Point 2	0.03	0.18	0.25	1.00	1.28	1.45
Stream Depth at Point 3	0.04	0.36	0.43	0.90	1.90	2.60
Stream Depth at Point 4	0.00	0.55	0.22	0.70	1.94	3.22
Stream Depth at Point 5	0.12	0.75	0.53	0.72	1.88	3.50
Stream Depth at Point 6	0.15	0.80	0.22	0.88	2.00	3.44
Stream Depth at Point 7	0.23	0.76	0.16	0.93	2.00	3.38
Stream Depth at Point 8	0.13	0.72	0.18	1.12	1.90	3.30
Stream Depth at Point 9	0.14	0.45	0.20	1.35	1.94	3.20
Stream Depth at Point 10	0.07	0.31	0.15	1.03	1.63	2.01
Stream Depth at Point 11	0.03	0.03	0.12	0.30	0.30	0.65
% Substrate Gravel or Larger	98	30	5	40	50	0
Maximum Thalweg Depth	0.23	0.8	0.53	2	2	3.5
% Instream Cover	85	15	30	25	10	10
Macrophyte Abundance 1=Abundant, 2=Common, 3=Rare,4=Absent	2	2	1	1	4	3
Algae Abundance 1=Abundant, 2=Common, 3=Rare,4=Absent	2	3	3	3	3	3
Left Bank % Trees	20	10	0	20	25	15
Left Bank % Shrubs	40	0	15	0	0	5
Left Bank % Grasses, Forbs	20	60	45	20	5	10
Left Bank % Cult Fields	0	0	0	0	0	0
Left Bank % Other	20	30	40	60	70	70
Right Bank % Trees	10	10	0	10	5	10
Right Bank % Shrubs	20	0	0	0	0	0
Right Bank % Grasses, Forbs	20	30	80	20	30	35
Right Bank % Cult Fields	0	0	0	0	0	0
Right Bank % Other	50	60	20	70	65	55
Trasect Latitude	31.19072	31.19003	31.18950	31.18908	31.18872	31.18794
Trasect Longitude	-98.90272	-98.90322	-98.90381	-98.90467	-98.90558	-98.90600

Exhibit 4E
SWQMIS Validations

Exhibit 4E

SWQMIS Validations

SWQMIS performs serval initial validations which includes:

- Checks for correct file formats as specified in the DMRG.
- Checks for validity of the following codes by cross-referencing them against the SWQMIS reference tables:
  - Submitting Entity,
  - Collecting Entity,
  - Monitoring Type Codes,
  - StationIDs.
  - TagID prefixes,
  - Parameter Codes.
- Checks the Result file for duplicate Parameter Codes on unique TagIDs.
- Checks the Event file for duplicate TagIDs that already exist in SWQMIS.

SWQMIS also checks records in the SWQMIS against the dataset for records with the same station, date, SE/CE/MT, depth, parameter codes, etc. to be sure that loading the current dataset would not create duplicate records in SWQMIS.

The following are the validation rules and SWQMIS's automated behavior:

Rule	SWQMIS Action & Examples	Notes
Water temperature (00010) must be reported to the nearest tenth of a degree. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 30.11 corrects to 30.1 28.55 corrects to 28.6 30 corrects to 30.0	
pH (00400) must be reported to the nearest tenth of a pH standard unit. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 7.22 corrects to 7.2 6.88 corrects to 6.9 7 corrects to 7.0	
Dissolved oxygen (00300) must be reported to the nearest tenth of a mg/L. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 6.33 corrects to 6.3 4.19 corrects to 4.2 6 corrects to 6.0	

Specific conductance (00094) must be reported to three significant figures when the value exceeds 100. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 1014 corrects to 1010 1267 corrects to 1270	
Salinity (00480) must be reported to the nearest tenth of a part/thousand when the reported value is above 2.0. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 3.12 corrects to 3.1 7.77 corrects to 7.8 3 corrects to 3.0	
If the Station is a freshwater or inland (brine) location, do not report salinity. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).		Parameter must be removed by submitting entity.

Rule	**SWQMIS Action & Examples**	**Notes**
Secchi disk (00078) must be reported to two significant figures. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 0.351 corrects to 0.35	
Days since last significant precipitation (72053) must be reported as a whole number. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
If sample collected when raining or has rained within the last 24 hours, report a value of <1 for Days since last significant precipitation (72053). . (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
E. coli (31699) must be reported as a whole number with two significant digits. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 854 corrects to 850	
E. coli (31699) must not be reported as zero. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	

Enterococcus (31701) must be reported as a whole number and with two significant figures. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 858 corrects to 860	
Enterococcus (31701) must not be reported as zero. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
Fecal coliform (31616) must be reported as a whole number with two significant figures. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 1214 corrects to 1200	
Fecal coliform (31616) must not be reported as zero. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	

Rule	**SWQMIS Action & Examples**	**Notes**
Fecal coliform (31616) must not be reported as TNTC. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
Flow (00061) values less than 10 and greater than 0.1 must be reported to the nearest tenth. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 8.62 corrects to 8.6 4 corrects to 4.0	
Flow (00061) values greater than 10 must be reported to the nearest whole number. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Auto-correction 15.6 corrects to 16	
Flow (00061) values less than 0.01 must be reported as <0.01. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
Flow severity (01351) must be a whole number in the range of 1 through 6. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	

If Flow (00061) is reported as zero, then Flow Severity (01351) must be reported as 1. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
If Flow Severity (01351) is reported as 6, then Flow (00061) must not be reported. (Rule is from SWQM Procedures Manual Volume 1 (10/2008), pages 3-14, Table 3.1).	Manual correction	
Composite samples must include all required fields.	Manual correction	
Each TagID in the Event file must have at least one reported Result.	Manual correction	
Each TagID in the Result file must have a TagID in the Event file or already in SWQMIS.	Manual correction	
Startdate must be before Enddate. If the Startdate and Enddate are the same, then Starttime must be prior to Endtime.	Manual correction	
Outliers must include a '1' in the Verify_flg field in the Result file.	Manual correction	