



EQuIS Data Processor (EDP)



**Users Manual
Version 2.0**



U.S. Environmental Protection Agency
Region 5, 77 West Jackson Boulevard
Chicago IL 60604

August 2008

DISCLAIMER OF ENDORSEMENT

Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government, and shall not be used for advertising or product endorsement purposes.

TABLE OF CONTENTS

1.0 INTRODUCTION 1

2.0 EDD DATA FILE CHECKS 2

3.0 INSTALLING EARTHISOFT'S EDP 3

4.0 USING THE EDP 11

 4.1 LOADING EDD FILES 13

 4.2 ERROR LOGS 16

 4.3 CORRECTING ERRORS 17

 4.4 USING FIND AND REPLACE 18

 4.6 SAVING CHANGES TO THE EDD FILE 18

 4.7 THE SIGN AND SUBMIT FUNCTION 19

5.0 SUBMITTING THE EDD DATA PACKAGE 21

6.0 UPDATING THE REFERENCE VALUE FILE 21

7.0 UPDATING THE FORMAT FILE 21

APPENDIX A 22

ACRONYMS

File Extensions	.ftp	File Transfer Protocol
	.rvf	EarthSoft Reference Value File
	.xse	Extensible Markup Language Schema Encrypted File
	.zip	WinZip Archive File
Acronyms	EDD	Electronic Data Deliverable
	EDP	Electronic Data Processor
	EFDC	Electronic Field Data Checker
	ELDC	Electronic Laboratory Data Checker
	EPA	Environmental Protection Agency
	EQiS	Environmental Quality Information System
	HTML	HyperText Markup Language
	TB	Trip Blank
	TIC	Tentatively Identified Compound

1.0 INTRODUCTION

The Environmental Quality Information System (EQuIS) Data Processor (EDP) has been made available to data providers in order to check their Electronic Data Deliverable (EDD) files prior to submittal to EPA Region 5. The EDP is used to check EDD files to ensure they are formatted as described in the ‘Comprehensive Electronic Data Deliverable Specification Manual’. If the EDP detects errors, the errors will be identified and can be viewed directly within the EDP or via an error log. After the errors are corrected by the data provider, the EDP should be re-run to assure that no errors remain. The EDD files can then be signed and submitted and posted to the Region 5 EDD .ftp site for incorporation into the regional database.

The EDP is a product of EarthSoft, Inc. and replaces the two previous Region 5 EDD checker applications, the Electronic Laboratory Data Checker (ELDC) and the Electronic Field Data Checker (EFDC). The EDP is a single application that checks all EDD files and provides a much easier user interface for identifying and correcting errors.

This document’s sole purpose is to assist EPA Region 5 data providers in the installation and use the EDP in conjunction with submitting EDD data files to Region 5. Therefore, this document only provides information pertaining to the specific requirements of the Region 5 EDD specifications and is not intended to be a comprehensive EDP Users Manual. For more detailed discussion of the functionality and technical specifications of EDP, please refer to EarthSoft’s web site at www.earthsoft.com.

2.0 EDD DATA FILE CHECKS

The EDP checks data for the following:

- Required Fields;
- Field Lengths;
- Data Types;
- Valid Dates;
- Reference Values;
- Duplicate Rows;
- Range Checking; and
- Record Parent-Child Relationships.

Appendix A describes the checks performed by the EDP.

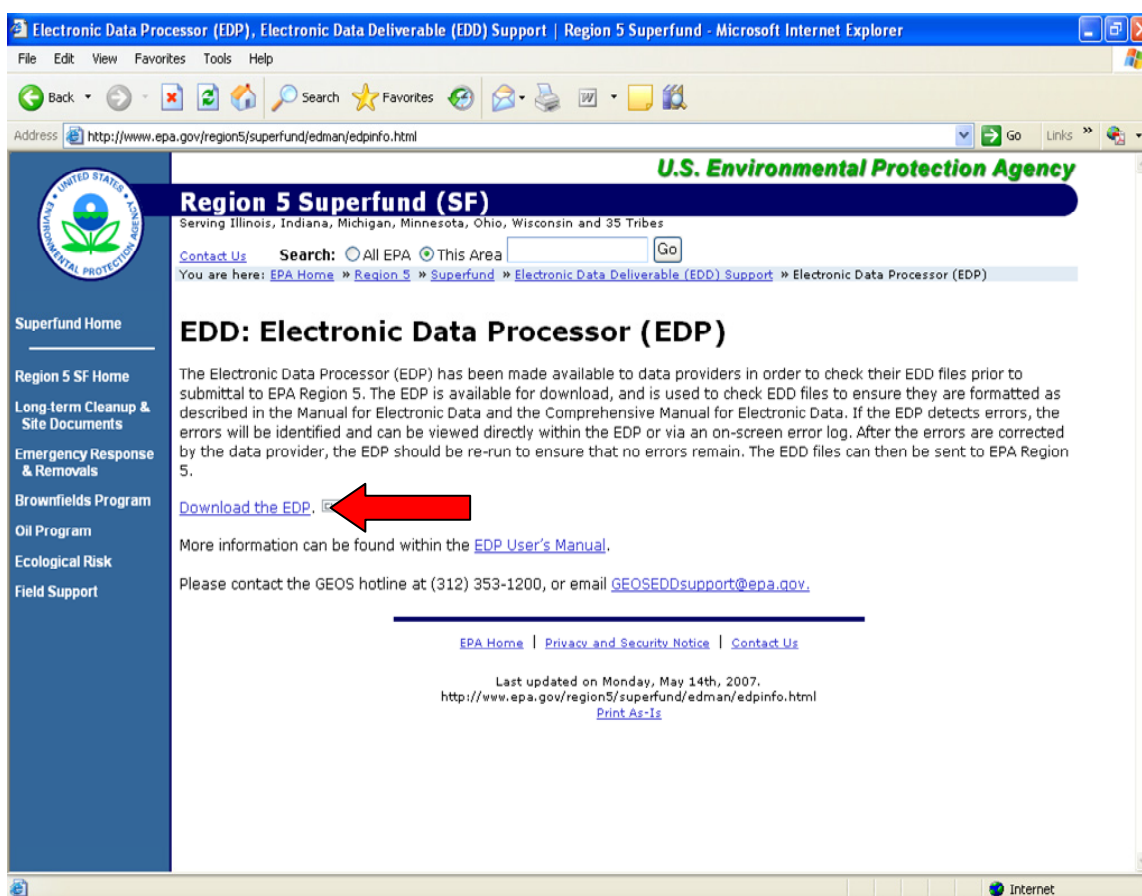
3.0 INSTALLING EARTHISOFT'S EDP

Installation of EDP involves the following three steps:

- 1) Downloading the EDP application from the EPA Region 5 website;
- 2) Installing the EDP; and
- 3) Registering the EDP.

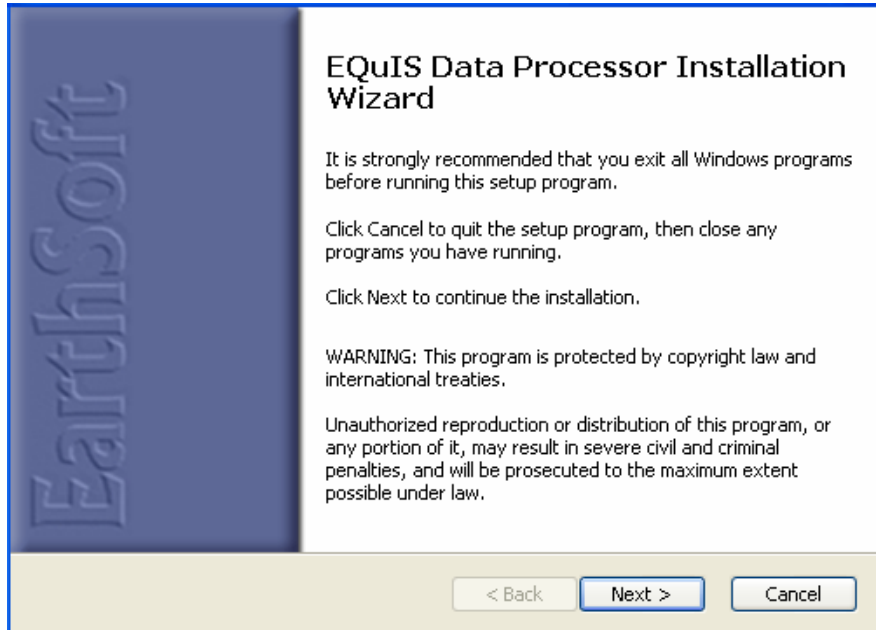
Downloading the EDP

The EDP installation application can be downloaded from the EPA Region 5 GEOS EDD Support website at <http://www.epa.gov/region5/superfund/edman/edpinfo.html>. Click on the 'Download the EDP' link and to begin the download process.



Installing the EDP

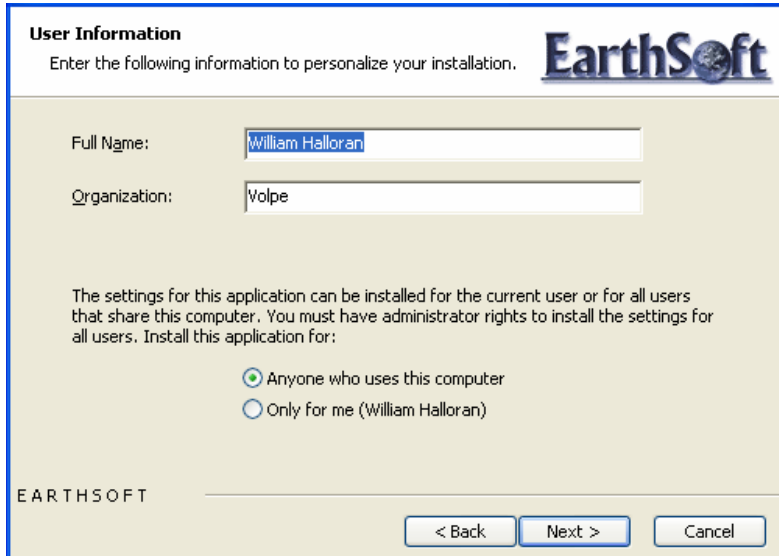
Open the directory where the EDP installation application was downloaded, and double-click the file. The installation wizard will launch. Follow the instructions which will lead step by step through the installation procedure. It is important to note that during installation you should have no other programs running.



Click the next button. The License Agreement screen will appear. Select 'I accept the license agreement' radio button and click the 'Next' button.



Enter full name and organization and select the desired setting. Click the 'Next' button.



User Information **EarthSoft**
Enter the following information to personalize your installation.

Full Name:

Organization:

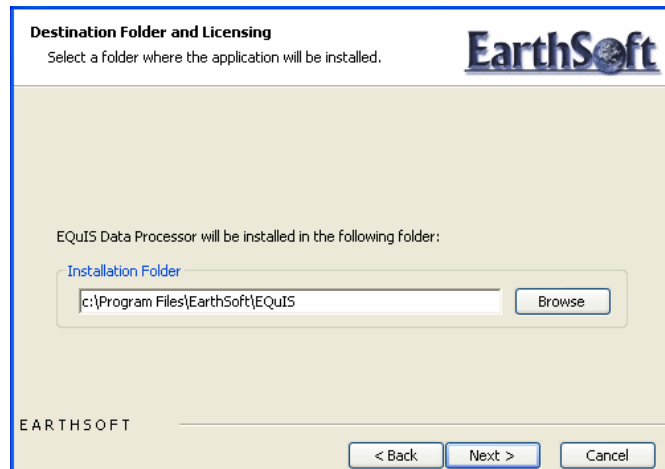
The settings for this application can be installed for the current user or for all users that share this computer. You must have administrator rights to install the settings for all users. Install this application for:

Anyone who uses this computer
 Only for me (William Halloran)

EARTH SOFT

< Back Next > Cancel

Select the destination folder for the application files. Click the 'Next' button.



Destination Folder and Licensing **EarthSoft**
Select a folder where the application will be installed.

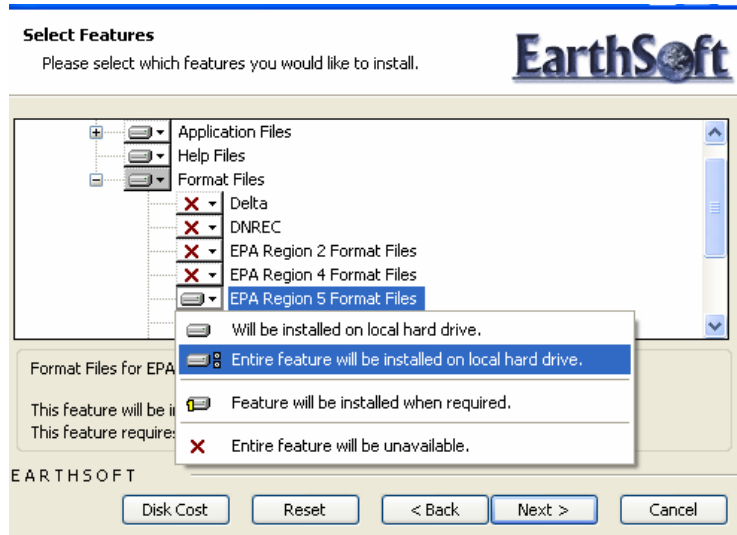
EQUS Data Processor will be installed in the following folder:

Installation Folder

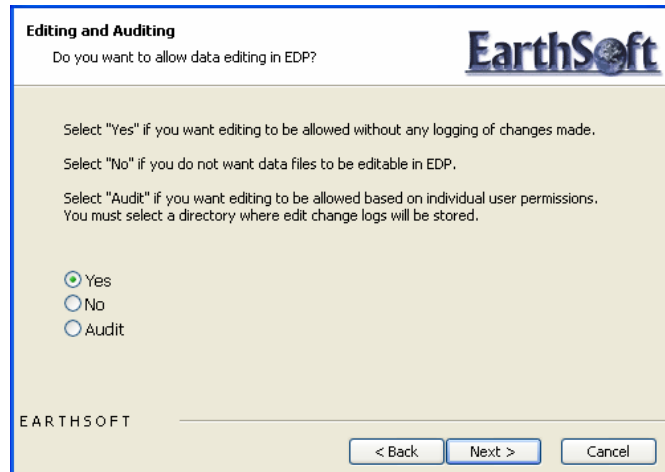
EARTH SOFT

< Back Next > Cancel

Click the icon next to EPA Region 5 Format Files and select 'Entire feature will be installed on local hard drive' and then click the 'Next' button.



Click the 'Next' button and you will be presented with the Editing and Auditing screen.



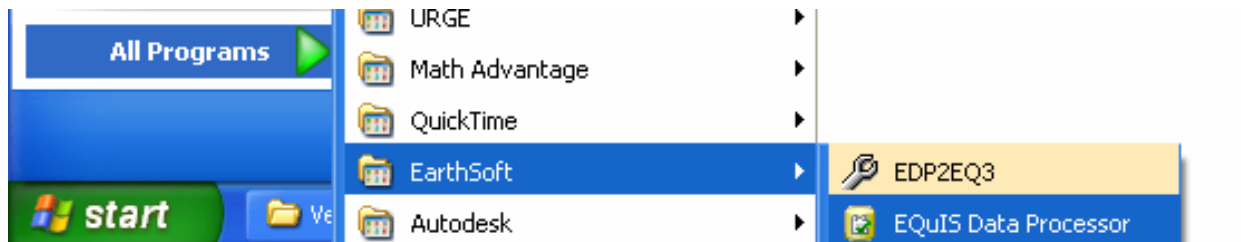
Selecting 'Yes' will allow you to make edits directly to the data in the EDD file via the EDP. Selecting 'No' will not permit any editing to the EDD file via the EDP and all edits to the file must be done directly to the EDD text file after exiting the EDP. You also have the option of auditing all changes that are made to the data files. A log is created that includes the date and time, the user, the original value, and the new value. If Audit is selected, you will also have to select a directory to which the auditing files will be stored. Once the selection is made, click the 'Next' button.



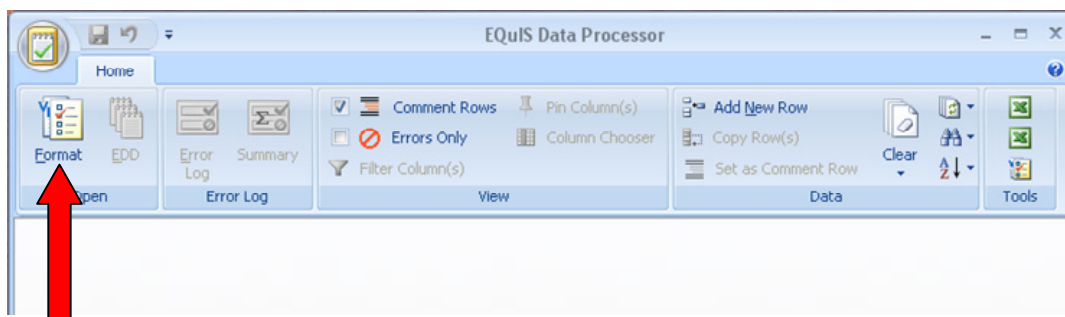
Click the 'Next' button to begin the installation. When the installation is done, you will be presented with a window that verifies that the EDP has been successfully installed. Click the 'Finish' button to exit the installation.

Registering the EDP

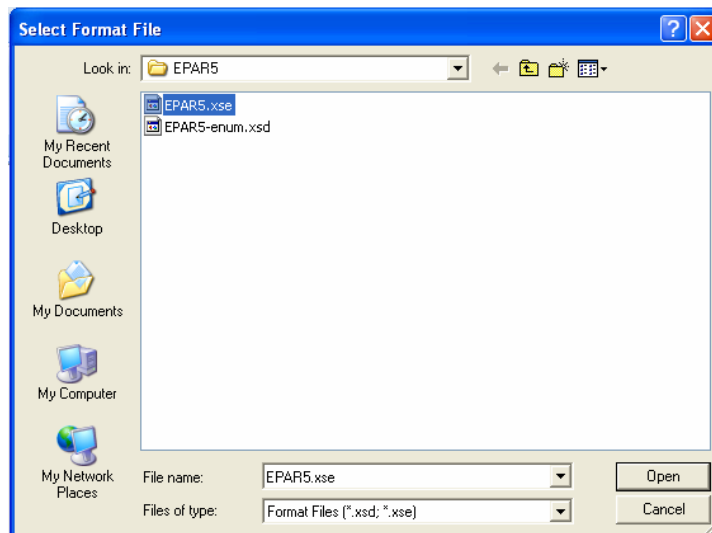
Once installed, the EDP must be registered. Start the EDP application by selecting Start>All Programs>EarthSoft>EQUS Data Processor.



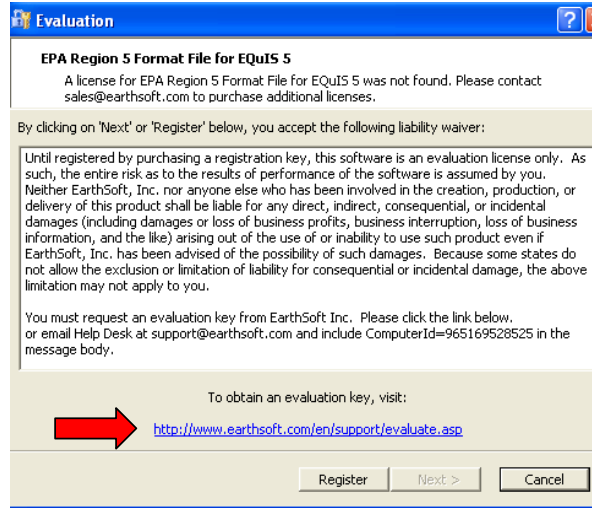
The EDP application will start and a blank screen appears. Select 'Format' from the upper menu.



Select the 'EPAR5.xse' file and click the 'Open' button.

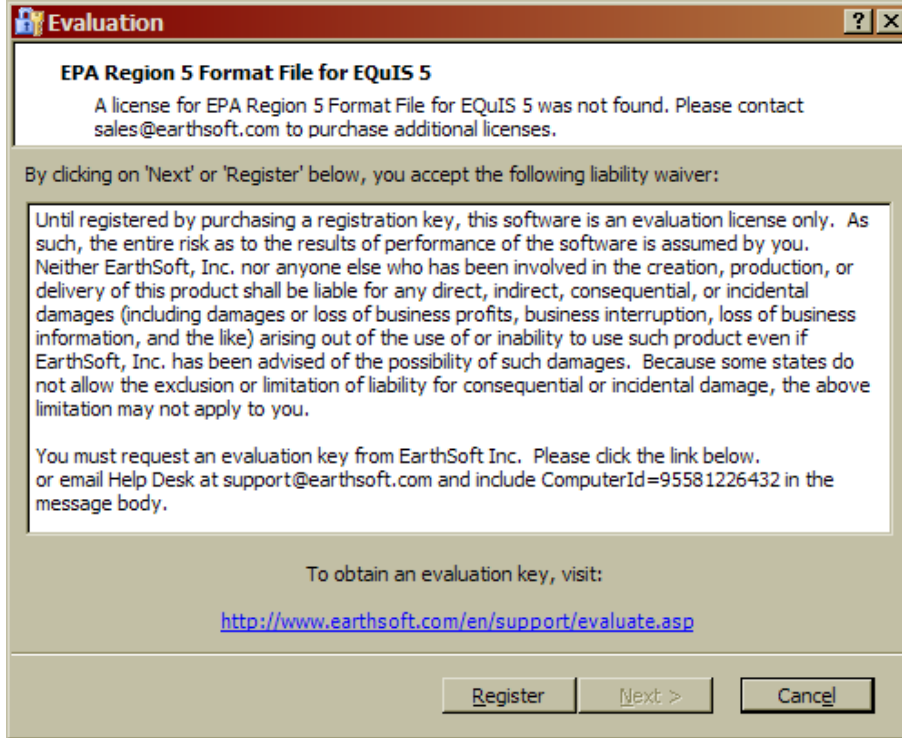


The EDP Registration Screen appears. Click the link to the EDP registration web page.

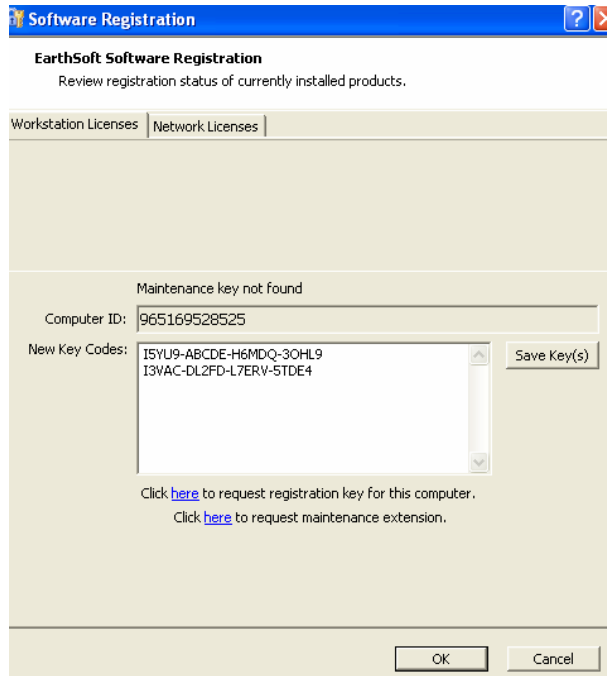


The 'Request Product Evaluation Key' screen will appear. Enter the requested information in the 'Key Request' form. When all information has been entered, select 'Submit'. After the registration form has been submitted, a registration key will be sent to the e-mail address provided in the registration form.





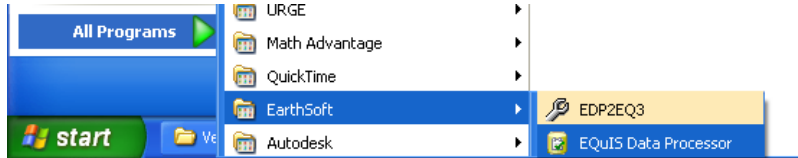
Once the registration key has been received, register the EDP by starting the application, Start>All Programs>EarthSoft>EQUIS Data Processor. The above screen appears. Click ‘Register’ button/link.



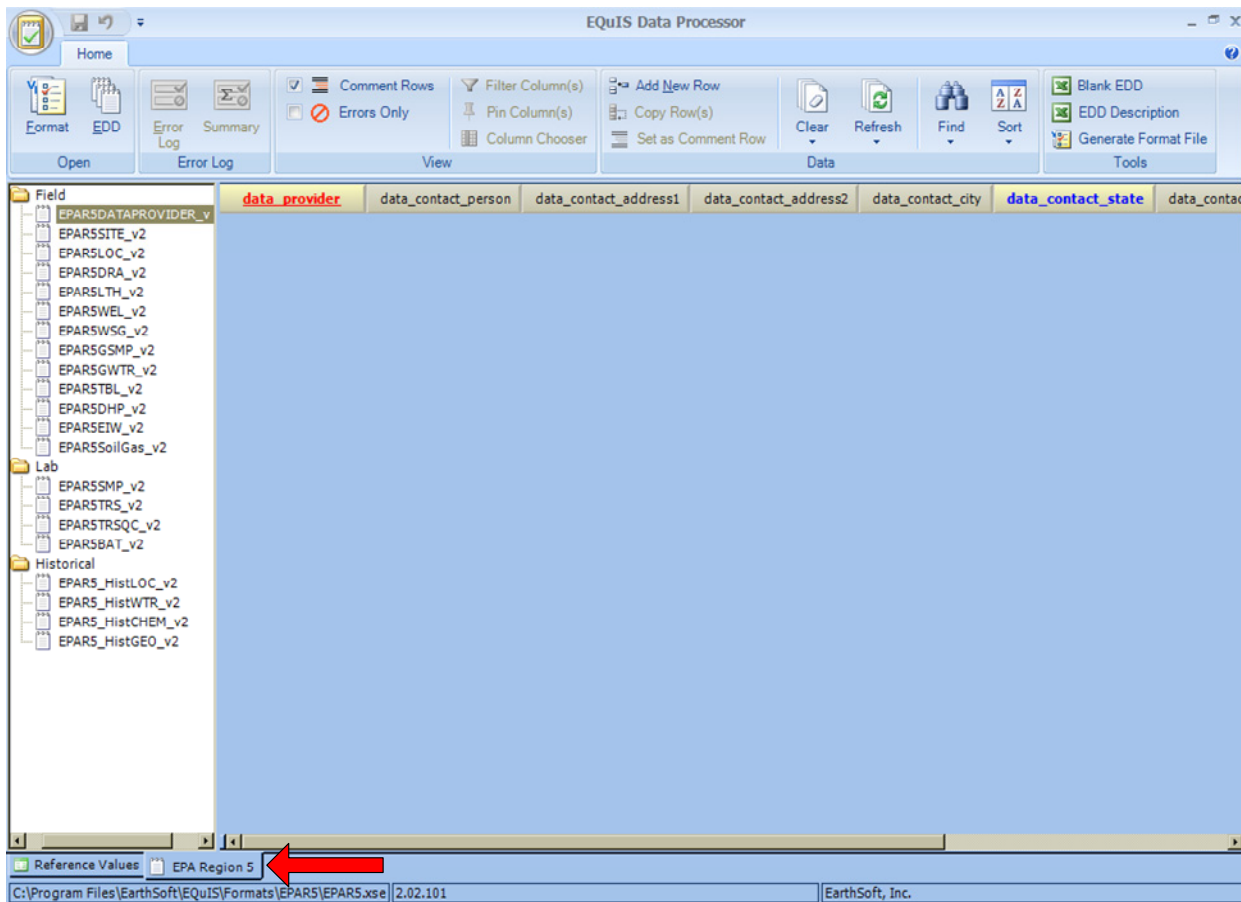
Enter the registration key sent via e-mail in the ‘New Key Codes’ field and click the ‘Save Key(s)’ button. A screen stating that the registration succeeded should appear. The EDP is now registered and ready for use.

4.0 USING THE EDP

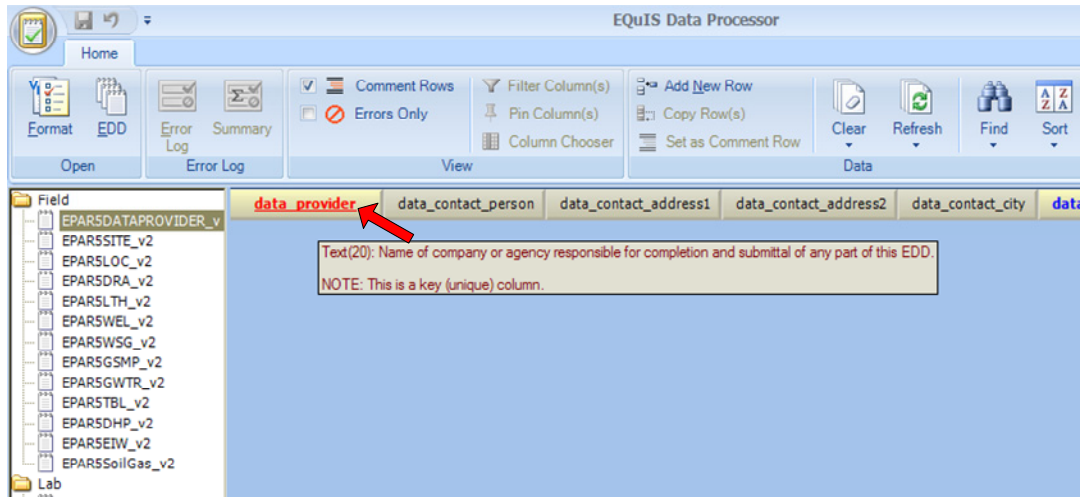
Start the application by selecting Start>All Programs>EarthSoft>EQiS Data Processor from the Windows ‘Start’ menu.



The EDP will open. Select the ‘EPA Region 5’ tab located at the bottom of the screen (as indicated by the arrow in the example below). The Region 5 file formats are displayed along the left side of the window. An empty table with the field names associated with the highlighted file type is displayed along the top. Each of the EDD file types listed in the EDP corresponds to the EDD files described in the Region 5 ‘Comprehensive Electronic Data Deliverable Specification Manual’ and ‘Basic Electronic Data Deliverable Specification Manual’. In the screen below, the ‘EPAR5DATAPROVIDER_v2’ format has been selected and its associated fields are displayed across the top.



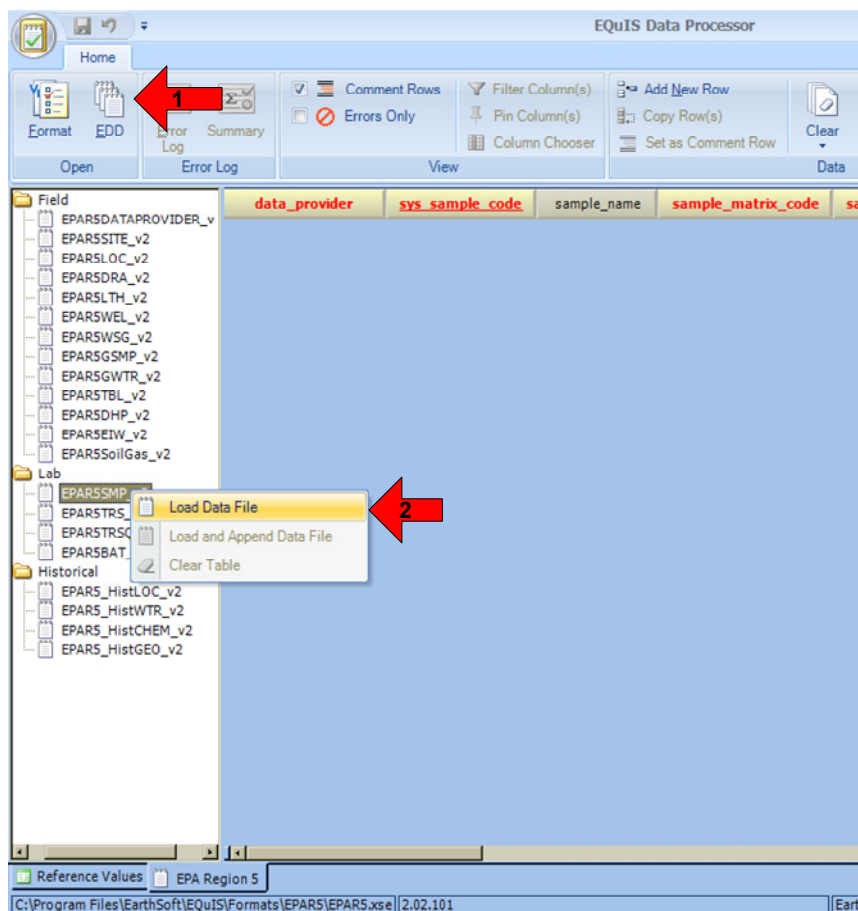
Fields with red text are 'Required' fields and cannot be left blank; they must be populated with data. Information about each field is provided when the cursor is placed over the field name (As indicated in the example below).



4.1 Loading EDD Files

Files are checked either by loading individually created EDD files into EDP or by loading a single Access database created with individual tables named according to the naming convention, or an Excel spreadsheet with tabs named according to the naming convention.

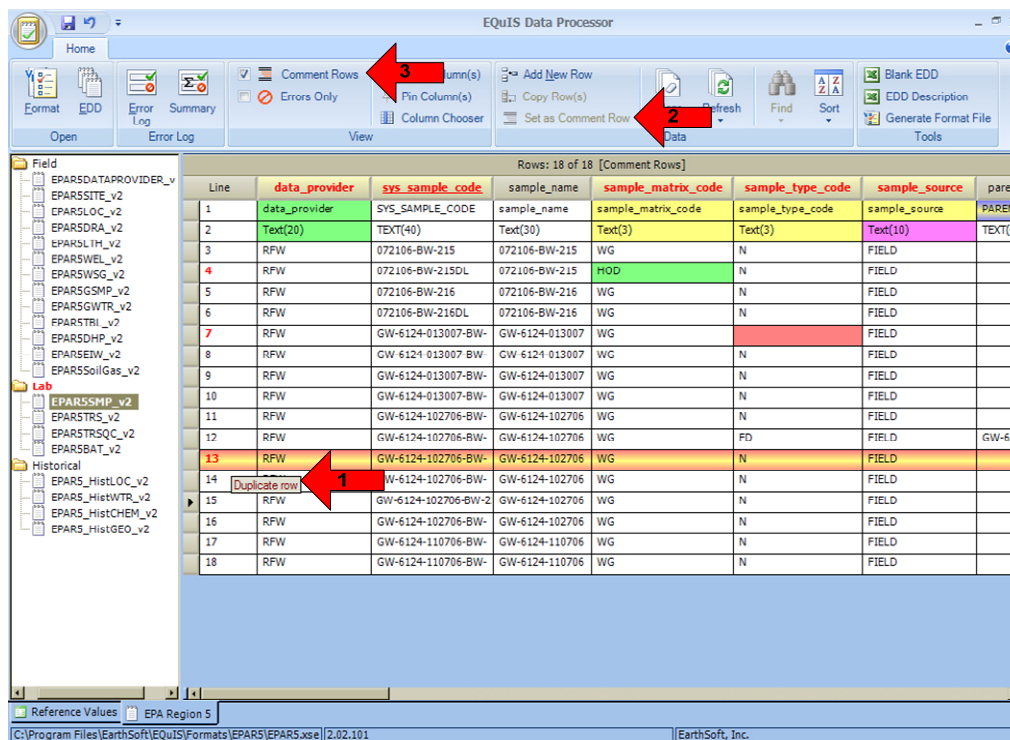
To load a single EDD file, first, select the format type of the EDD file to be checked from the format list at the left. In the example below, an EPAR5SMP_v2 file is going to be checked, therefore, the EPAR5SMP_v2 format has been selected. Next, load the EDD data file by clicking the EDD icon located in the top menu bar (as indicated by arrow 1 in the example below) or right-click on the format type and select 'Load Data File' (as indicated by arrow 2 in the example below).



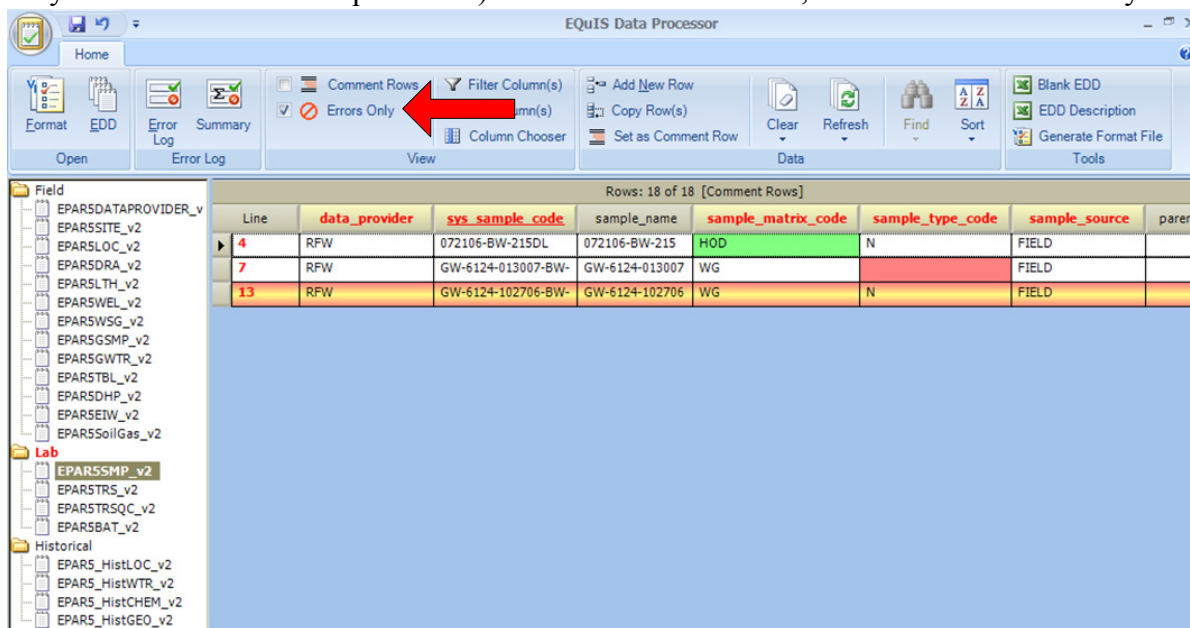
Use the Browse window to locate the EDD file and select 'Open'. The data file will load to the EDP and be checked during loading. Data will be displayed in the table and any detected errors will be shaded. **Note:** *If the data file contains header rows, EDP will identify fields in the header rows as errors unless each header row is preceded by a pound-sign character (#).*

To load a single EDD file containing multiple format sections, click the EDD button from the menu bar (Arrow 1 above) and use the browsing window to locate the EDD file and select 'Open'. The EDP will then load the constituent parts of the EDD into the appropriate locations and display any errors. **Note:** *This method may take several minutes.*

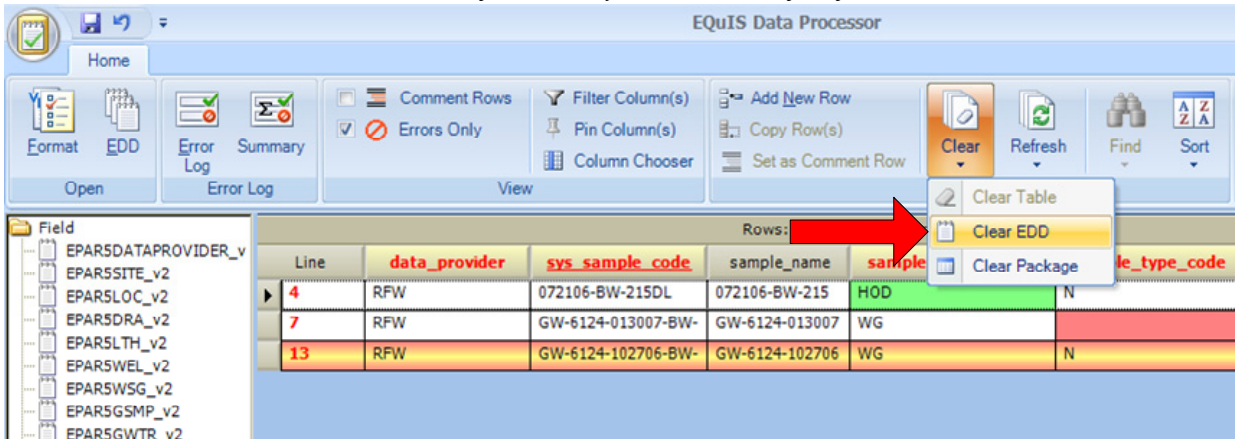
In the screen below, rows 4, 7, and 13 have errors. Each type of error is shaded differently. Place the cursor over the error to show the type of error (Arrow 1). To hide header rows that appear as errors, highlight the header rows, select ‘Set as Comment Row’ (Arrow 2) from the top menu, and then uncheck ‘Comment Rows’ also located in the top menu (Arrow 3). To unhide the header row, re-check the ‘Comment Rows’ box.



To view only the rows with errors, check the box next to ‘Errors Only’ located in the top menu bar (as indicated by the arrow in the example below). To restore all the rows, uncheck the ‘Errors Only’ box.

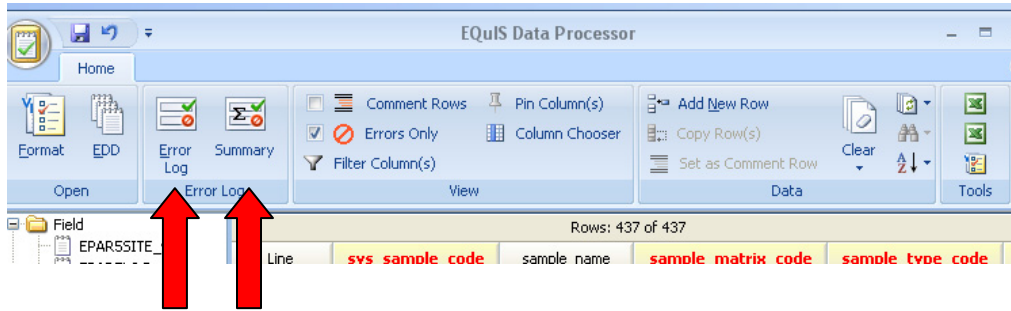


To clear the data from EDP, select 'Clear' from the top menu, then select 'Clear EDD' (as indicated by the arrow in the example below). The EDD file will be cleared from the EDP viewer. *Note: Clearing the data from the EDP will not delete the EDP file, it only removes the file from the viewer.*



4.2 Error Logs

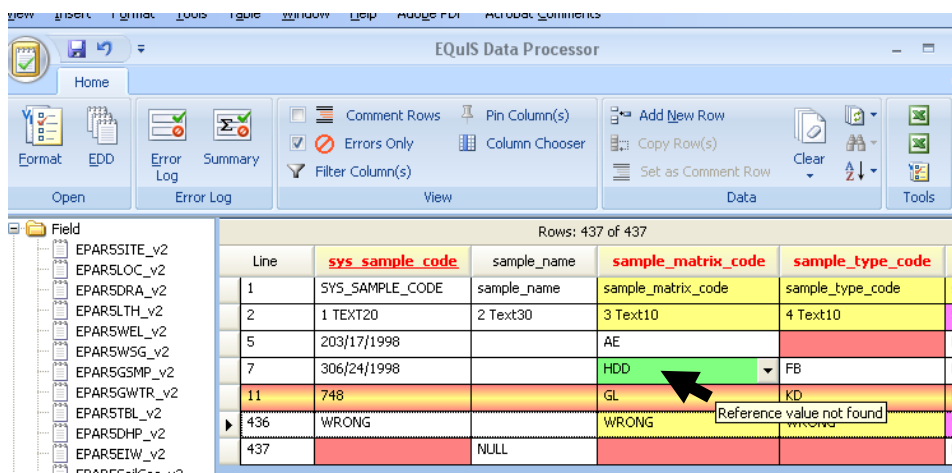
EDP produces an error log that can be saved as an HTML formatted file. In the top menu, select ‘Error Log’ to view and save the error details or ‘Summary’ to view and save a summary of the errors (as indicated by the arrows in the example below). Use the Browse window to locate the desired location and select ‘Save’. The error log will then be saved in the selected folder.



4.3 Correcting Errors

As stated above, the data are being checked for errors by the EDP as the EDD files are loading. The fields with errors will be shaded different colors depending on the type of error. The types of errors being checked for by the EDP are described in Appendix A.

A description of the error is provided when the cursor is placed over the field. In the example below, the sample_matrix_code value in line 7 is not a valid value.



Two methods can be used to correct this valid value error:

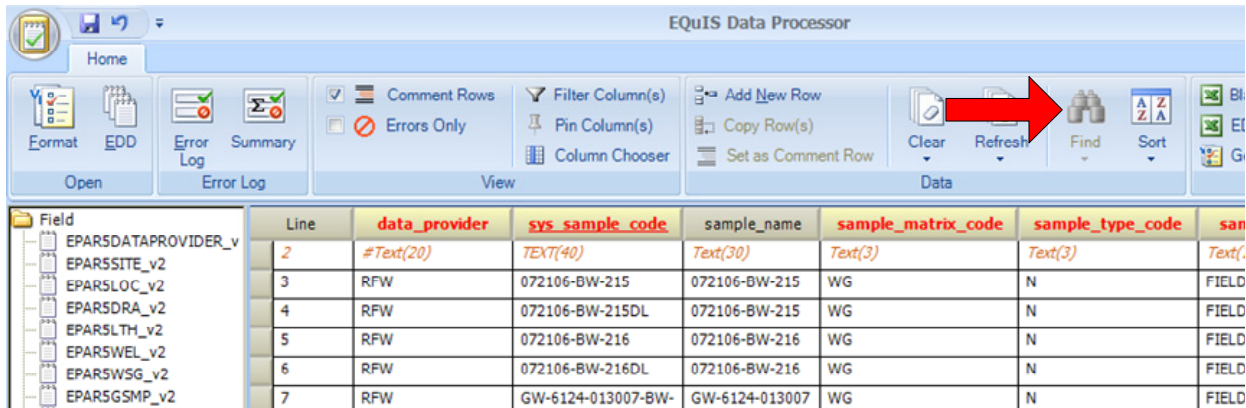
1. Exit the EDP and then open the EDD file using a text editor or spreadsheet application, correct the error, resave the file, and then re-load the EDD back into the EDP to ensure no further errors.
2. To correct errors directly within EDP, click in the error field and type the correct value. If the field is restricted to a list of valid values, the values will be provided from a drop down list by clicking on the down arrow located on the left side of the field. Once the error is corrected and the cursor is moved out of the field (i.e., user clicks on another field) the shading signifying an error will disappear.

Note: This method can only be used if 'Yes' or 'Audit' was selected in the 'Editing and Auditing' screen during EDP installation.


Note: If data providers believe that a new reference value is required, they should follow the process described in the 'Comprehensive Electronic Data Deliverable Specification Manual' to request that the value be added.

4.4 Using Find and Replace

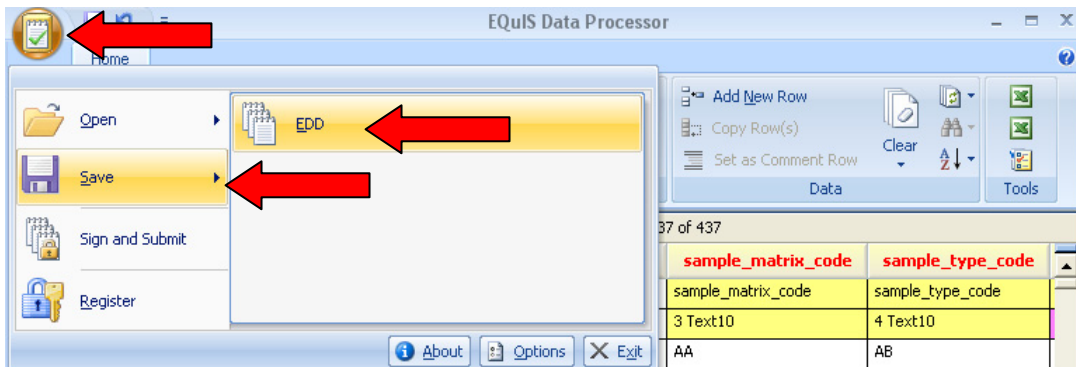
The ‘Find and Replace’ function allows searching the file for a specified value and then replacing it with another value. This function is useful when there are a number of similar values that need to be changed. The ‘Find and Replace’ function is activated by selecting the binocular icon located in the Data menu (as indicated by the arrow in the example below). The ‘Find and Replace’ dialog will appear. Type the value to be replaced in the ‘Find What’ field and the new value in the ‘Replace With’ field. Select ‘Find’ to view fields with the value and ‘Replace’ to replace the original value with the new value.



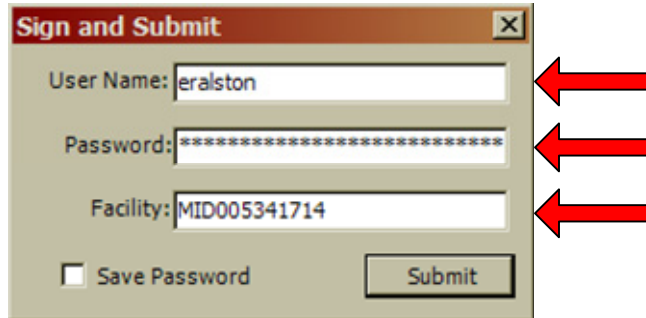
4.6 Saving Changes to the EDD File

To save the changes made to the EDD data file, click the  icon located in the upper right corner of the EDP and then select Save>EDD. Use the Browse window to locate the folder where the EDD file is to be saved and click ‘Save’. Any changes made to the EDD will now be saved.

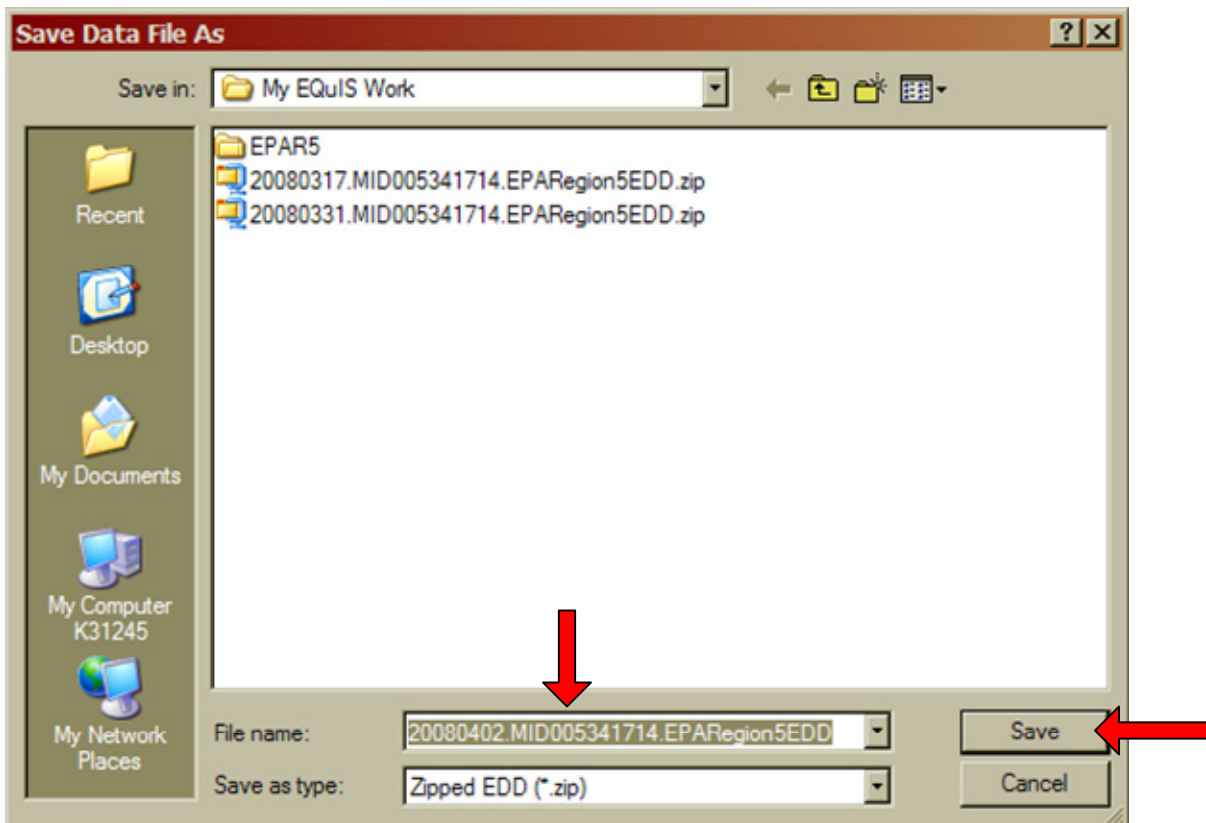
Note: Only the individual EDD file selected in the left hand format field and displayed in the main workspace will be saved. Therefore each EDD file should be saved separately, as appropriate.



This will open the ‘Sign and Submit’ authentication screen. This should be populated with your Region 5 provided user name, password, and the 12 character EPA ID for the site pertaining to the EDD files.



After entering the correct authentication information, click the ‘Submit’ button. Next select the location where you would like the .zip file saved and ensure that the file name of the ‘Data Package’ matches the required naming convention. Click the ‘Save’ button.



5.0 SUBMITTING THE EDD DATA PACKAGE

After the EDD files have been checked by the EDP and the ‘Sign and Submit’ process has been completed, the Data Package is ready for submittal to EPA Region 5. Please follow the procedures for submitting EDD files described in Region 5’s ‘Comprehensive Electronic Data Deliverable Specification Manual’ or ‘Basic Electronic Data Deliverable Specification Manual’ found on the EPA Region 5 GEOS EDD Support website located at <http://www.epa.gov/region5superfund/edman>.

6.0 UPDATING THE REFERENCE VALUE FILE

Periodically, EPA Region 5 will post an updated reference value file (.rvf) on the Region 5 GEOS EDD Support website located at <http://www.epa.gov/region5superfund/edman>. Follow the steps below to update the reference values in the EDP application:

1. Download the most recent reference value file from EPA Region 5 GEOS EDD Support website.
2. Replace the existing reference value file, ‘EPAR5.rvf’, located in the C:\Program Files\EarthSoft\EQuIS\Formats\EPAR5 folder, with the downloaded file.
3. The next time EDP is started, the new reference values will be loaded.

7.0 UPDATING THE FORMAT FILE

If Region 5 makes changes to the format file (.xse) of the existing EDD, the EDP application will need to be updated with a new format file. Follow the steps below to update the format file in the EDP application:

1. Download the most recent format file from EPA Region 5 GEOS EDD Support website, <http://www.epa.gov/region5superfund/edman>.
2. Replace the existing format file, ‘EPAR5.xse’, located in the C:\Program Files\EarthSoft\EQuIS\Formats\EPAR5 folder, with the downloaded file.
3. The next time EDP is started, the new format files will be loaded.

***Note:** Region 5 does not expect to make changes to the format in the near future; however, if changes are made, Region 5 will provide notification of the format changes to all data providers.*

APPENDIX A

This appendix describes the errors identified by the EDP.

1. Valid Value Not Found and Value Not Found on List

The value in the field does not match the values listed in the Valid Value Appendix of the Comprehensive Electronic Data Deliverable Specification Manual. Use the field descriptions in Section 3, 4, and 5 of the 'Comprehensive Electronic Data Deliverable Specification Manual' for details of values that are required. If an appropriate value does not exist, contact EPA Region 5 to request that a new valid value be added to the valid value table.

2. Value Exceeds Field Length

The number of characters of the value entered in the field exceeds the maximum allowed number of characters. See Section 3, 4, and 5 of the 'Comprehensive Electronic Data Deliverable Specification Manual' for the maximum field lengths.

3. Missing Required Field

The field must be populated with a value. The field cannot be left null (i.e., blank). See Section 3, 4, and 5 of the 'Comprehensive Electronic Data Deliverable Specification Manual' for information on required fields.

4. Invalid Data Type

The value is not the appropriate data type. Each field has a specific data type that must be used, such as text, date/time, or numeric. If the appropriate data type for a field is 'Date/Time', then the value must be a date in MM/DD/YYYY HH:MM format. See the Data Type description in Section 3, 4, and 5 of the 'Comprehensive Electronic Data Deliverable Specification Manual' for the appropriate data type.

5. Out of Range

The value is not within the allowable range of values. Most numeric fields will not allow a negative value to be entered. Only certain exceptions are made for field measurements.

6. Duplicate Row

Two or more records have the same values in the primary key fields. The primary key fields are the fields that make each record in the file unique. No two records can have the same values in the primary keys. For example, the EPAR5LOC file has the sys_loc_code field as the primary keys. Two records that both have 'MW-01' in the sys_loc_code fields would be considered duplicate records. To make each record unique, one record would have to be changed so that the sys_loc_code was something other than 'MW-01'. Refer to Section 2.6 of the 'Comprehensive Electronic Data Deliverable Specification Manual' for further discussion of duplicate records.

7. Orphan Row

The record is missing a required parent record. Records that depend on information (i.e., child records) from another record (i.e., parent record) must reference the parent record and the parent record must exist in the corresponding file. For example, each row in the EPAR5TRS file must include a sys_sample_code that corresponds to a sys_sample_code reported in the EPAR5SMP file.

If a record in the EPAR5TRS file has a sys_sample_code of GWSMP-006 then a record must also be included in the EAPR5SMP file with a sys_sample_code of GWSMP-006. If a record in the EPAR5TRS file has a sys_sample_code that is not included in the EPAR5SMP file, an ‘orphan record’ error will be identified. See Section 2.6 of the ‘Comprehensive Electronic Data Deliverable Specification Manual’ for further discussion of child/parent records.

Note: When checking the EPAR5BAT file, the EPAR5TRSQC file must be checked through the EDP prior to checking the EPAR5BAT file, otherwise all records in the EPAR5BAT file will be identified as ‘Orphan Row’ Errors. The EPAR5BAT file is only submitted if an EPAR5TRSQC file is submitted.

8. Reportable_result Cannot be ‘Yes’ When lab_qualifiers = E, G, P, or R

Identifies records as errors that have a laboratory qualifier that signifies the result should not be reportable, yet the record has ‘YES’ reported in its reportable_result (EPAR5TRS and EPAR5TRSQC, Column 33) field. If a record has a lab_qualifier (EPAR5TRS and EPAR5TRSQC, Column 35) value of ‘E’, ‘G’, ‘P’ or ‘R’ they are considered as not reportable and the reportable_result (EPAR5TRS and EPAR5TRSQC, Column 33) field must be populated with ‘NO’.

9. Result_value is Required when detect_flag = Y

Identifies records that have the detect_flag (EPAR5TRS and EPAR5TRSQC, Column 34) value of ‘Y’ yet there is no value reported in the result_value field (EPAR5TRS and EPAR5TRSQC, Column 30). This error applies only to records of target analytes (TRG) and tentatively identified compounds (TIC). If a record has a value of ‘TRG’ or ‘TIC’ in the result_type_code (EPAR5TRS and EPAR5TRSQC, Column 32) and the detect_flag (EPAR5TRS and EPAR5TRSQC, Column 34) has a value of ‘Y’, the result_value field (EPAR5TRS and EPAR5TRSQC, Column 30) must be populated with the test result value (i.e., it cannot be left blank).

10. Reporting_detection_limit Cannot be Null when detect_flag = N

Identifies records with detect_flag (EPAR5TRS and EPAR5TRSQC, Column 34) value of ‘N’ and the reporting_detection_limit (EPAR5TRS and EPAR5TRSQC, Column 40) field is null. All records that have a value of ‘N’ in the detect_flag field must have the reporting_detection_limit field populated with the appropriate detection limit value (i.e. it cannot be left null).

11. Parent_sample_code is Required Where sample_type_code = BD, FD, FR, FS, LR, MS, SD, or MSD

Identifies records that have a sample_type_code (EPAR5SMP, Column 5) of ‘BD’, ‘FD’, ‘FR’, ‘FS’, ‘LR’, ‘MS’, ‘SD’, or ‘MSD’ but are missing the appropriate parent_sample_code (EPAR5SMP, Column 7). The above sample_type_codes signify duplicates, and the sample identifier (i.e., sys_sample_code) of the original sample from which the duplicate was derived must be populated in the parent_sample_code (EPAR5SMP, Column 7) field. The parent_sample_code value must match the sys_sample_code of the original sample and the original sample must also be reported as a separate record in the EPAR5SMP file (i.e., there should be a record for the original sample and a separate record for the duplicate sample).

12. Dates Not in Sequence

Errors will be identified if certain dates precede other dates, such as the analysis date preceding the sampling date.

13. Qc_spike_measured Less Than qc_original_conc and qc_dup_spike_measured Less Than qc_dup_original_conc

The qc_original_conc and the qc_dup_original_conc must be greater than qc_spike_measured and qc_dup_spike_measured, respectively. This applies to the EPAR5TRSQC format.

14. Sys_loc_code is Required when sample_type_code = 'N'

A location identifier (i.e., sys_loc_code) must be provided for all samples that are normal environmental samples. Therefore, all records in the EPAR5SMP_v2 file that have a sample_type_code (Column 5) of 'N' must also have the sys_loc_code field (Column 10) populated (i.e., this field cannot be left blank).