EQuIS Data Processor Reference Manual

January 2020

USEPA Region 4 Superfund and Emergency Management Division

Updated by:

earthsoft

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.

Contacts

For questions and comments, contact:

Your Remedial Project Manager (RPM) or,

DART Coordinator
SEMD, 11th Floor East
United States Environmental Protection Agency, Region 4
Sam Nunn Atlanta Federal Center
61 Forsyth Street, SW
Atlanta, GA 30303-8960

(941)544-5428

R4DART@epa.gov

Acronyms

CAS RN – Chemical Abstracts Service Registry Number

CLP **–** Contract Laboratory Program

DART **–** Data Archival and ReTrieval

EDD **–** Electronic Data Deliverable

EDGE – EQuIS Data Gathering Engine

EDP **–** EQuIS Data Processor

EQuIS – Environmental Quality Information System

EPA **–** United States Environmental Protection Agency

PRP **–** Potentially Responsible Party

QC – Quality Control

RPM – Remedial Project Manager

SEMD – Superfund and Emergency Management Division

XML – Extended Markup Language

Table of Contents

1 Introduction

1.1 Purpose

The sole purpose of this document is to assist the EPA Region 4 Superfund and Emergency Management Division (SEMD) data providers with the installation and use of the EQuIS™ Data Processor (EDP) in conjunction with submitting Electronic Data Deliverable (EDD) files to the EPA Region 4. Therefore, this document only provides information pertaining to the specific requirements to validate EDDs for the EPA Region 4 and is not intended to be a comprehensive EDP Manual. For a more detailed discussion of the functionality and technical specifications of EDP, please refer to EarthSoft's website at: https://help.earthsoft.com.

For a more detailed discussion of the EPA Region 4 EDD specifications, please refer to the *EPA Region 4 EDD Format File Guide* found on the EPA Region 4 Superfund website at:

https://www.epa.gov/superfund/region-4-superfund-electronic-data-submission

1.2 Scope and Application

The methods described in this document are to be used by all data providers when preparing and submitting environmental data electronically to the EPA Region 4, regardless of the originating program.

Following these procedures will help to reduce errors in data submitted to the EPA Region 4 and will enforce consistency, maintaining the strength and integrity of the EPA Region 4 EQuIS database. This strength of data allows for more informed and cost-effective decision-making.

1.3 Definitions

Data Provider – "Data Provider" and "Sampling Company" are defined to be interchangeable with regard to EDD submittals. The data provider may be defined as the person or agency that organized, formatted, and submitted the electronic data from a sampling event. This may or may not be the sampling company, particularly when working with historic data. Unless otherwise noted by the RPM, the prime contractor or grantee is always entered as the Sampling Company for the Samples and the Data Provider for the Geology EDDs.

Electronic Data Deliverable – An Electronic Data Deliverable (EDD) is a tabular, flat file format for sharing, manipulating, and using data, and is commonly a text or Microsoft Excel file type. For the EPA Region 4, an EDD is an archive zip file saved with a ".dat" extension holding one or more tab-delimited text files containing the environmental data to be submitted. An EDD file follows a consistent design meant to organize information in a useful format and typically contains header row(s) that describe what information should be completed in each column and in what format that data should be entered (see image on next page).

	Column 1	Column 2	Column 3	Column 4
Header Row 1	**#sys_loc_code**	**x_coord**	**y_coord**	**coord_type_code**
Header Row 2	*#Text(20)*	*Numeric*	*Numeric*	*Text(20)*
Data Row 1	EPA-1S	-82.317493	28.057509	LAT LONG
Data Row 2	EPA-2I	-82.317659	28.057151	LAT LONG

EQuIS Data Processor – The standalone version of the EQuIS Data Processor (EDP) is a powerful desktop interface to check for data completion and referential integrity, identify and correct errors, and create compressed files containing multiple related EDDs in a single useable format for upload and storage in a relational data base system.

EDP Format File – The EDP format file is the essence of data checking with EDP. The EDP format file set contains the definitions and restrictions for each individual field in available data tables. The format files control data checks, such as range checking, formatting, and enumerations. Format files are used in the field data collection tools EQuIS Collect and EQuIS Data Gathering Engine (EDGE) as well as EDP. The EPA Region 4 has a custom EDP format file, *EPAR4.zip*. It is extremely important to ensure you are using the most current file. Check the EarthSoft website to see if your version of the EPA R4 format file is current before working on your data.

Scribe – Scribe is a software tool developed by the EPA to assist in the process of managing environmental data. Scribe captures soil, water, air, and biota sampling, observational, and monitoring field data. Scribe can import EDDs from analytical laboratories, location data from a global positioning system (GPS), and from exported EQuIS EDDs.

Sampling Company – "Data Provider" and "Sampling Company" are defined to be interchangeable with regard to EDD submittals. The data provider may be defined as the person or agency that organized, formatted, and submitted the electronic data from a sampling event. This may or may not be the sampling company, particularly when working with historic data. Unless otherwise noted by the RPM, the prime contractor or grantee is always entered as the Sampling Company for the Samples and the Data Provider for the Geology EDDs."

Reference Value File – The reference value file (RVF) is associated with the EDP format file and is denoted with an .rvf file extension. This file contains the valid values reference tables that EDP uses to populate the drop-down menus that control when a specific type of value is required in an EDD. Example reference values are "mg/kg" (milligrams per kilogram) for a unit code or "GW" (groundwater) for a media code. These fields limit the type of data permitted in certain columns of the EDD, and the most recent valid values are in the RVF file. Therefore, it is extremely important to ensure you are using the most current file. Check the EarthSoft website to see if your version is current before working on your data.

ZIP Archival File Format – The ZIP file format is a data compression and archival format that contains one or more files that have been compressed, to reduce their file size, or stored as-is. ZIP archival files typically use the file extensions ".zip" or ".ZIP".

EDP automatically creates an archival zip file for submittal to the EQuIS database, however, the EPA electronic mail system will reject files with a .zip file extension due to security concerns. Therefore, the EDD package created by EDP with the extension ".zip" should be renamed with the extension ".dat". Details of this process are discussed in Section 3.8.

2 Getting Started

The standalone version of the EQuIS Data Processor is a product of EarthSoft, Inc. and replaces all previous methods of EDD checking, whether electronic or manual. EDP is a single application that checks all EDD files currently used by the EPA Region 4 and provides an easy-to-use, straight-forward interface for identifying and correcting data errors.

EDP Standalone has been made available to data providers to check their EDD files prior to submittal to the EPA Region 4. EDP is used to ensure EDD files are formatted as described in the EPA Region 4 File Format Guide. If EDP detects errors, the errors can be viewed directly within EDP or via an error log. After the errors are corrected by the data provider, the revised EDD files should be processed by EDP to ensure that no errors remain. The EDD files can then be "signed and submitted," saved with the appropriate filename, and emailed to the EPA Region 4 DART EQuIS database at EPAR4@EQuISonline.com.

Getting started with EDP Standalone involves three steps:

1) Downloading the EDP Standalone application,
2) Installing EDP Standalone, and
3) Registering EDP Standalone with EarthSoft.

Note: You must be an administrator or user with "Power User" privileges on your computer to install EDP. Check with your IT support before downloading and installing any software and only download EDP directly from EarthSoft or the EPA Region 4.

2.1 Downloading EDP Standalone

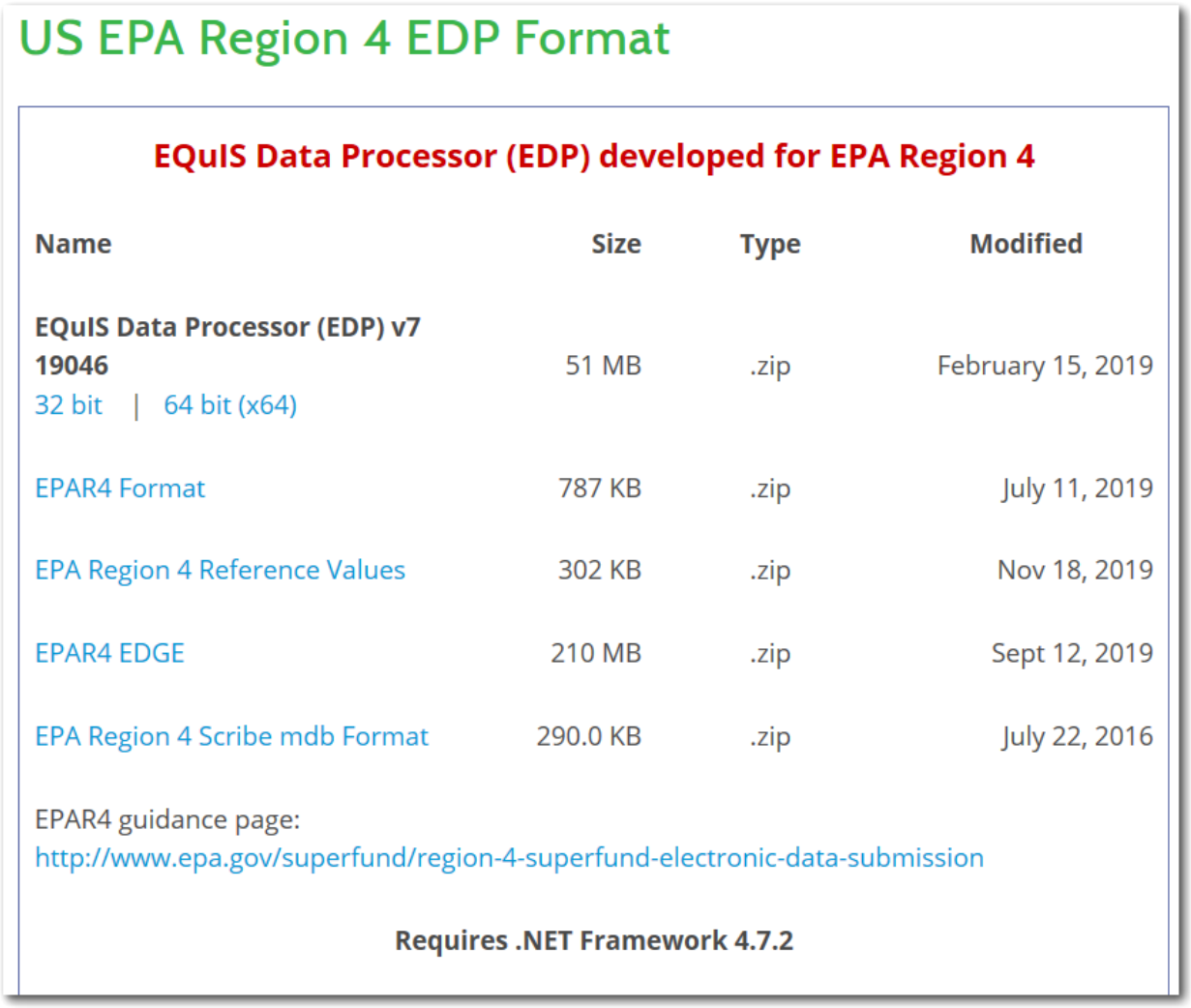
The most recent build of the EQuIS Data Processor 7 Standalone application can be downloaded directly from EarthSoft for no cost at:

https://earthsoft.com/products/edp/edp-format-for-epar4/.

The EPA Region 4 format file (*EPAR4.zip*) and reference value file (EPAR4rvfyyyymmdd.zip, where yyyymmdd is the date of the file creation) are also available for download from the EarthSoft website.

Note: The EPA Region 4 'EDP 7 Format File' requires the 'EDP 7 Version' application.

Ensure that you have the correct version for the Microsoft .NET Framework installed before installing EDP. The download page will provide the version needed for the most recent version of EDP. Information on checking your .NET version and obtaining the correct version can be found on the Microsoft website at http://www.microsoft.com/NET/, and additional information regarding Microsoft can be found below the installation instructions on the EarthSoft website.

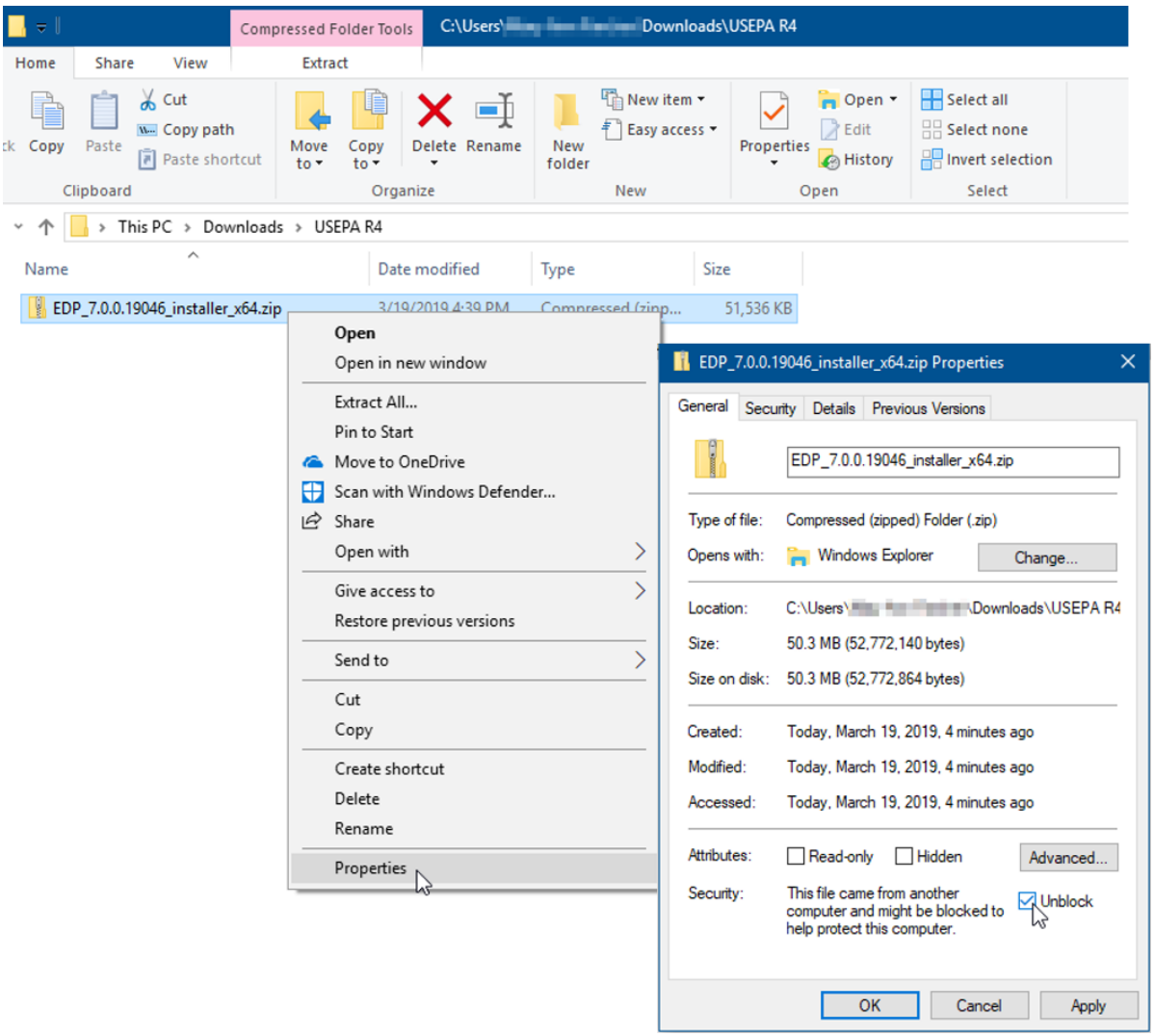


2.2 Installing EDP

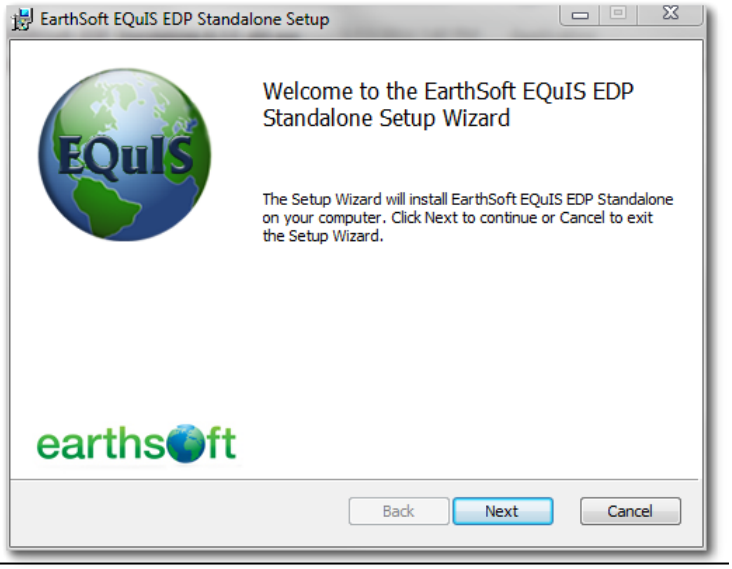
Open the directory where the EDP application was downloaded and unblock the file prior to unzipping the file.

Note: When downloading files from the Internet or other location, Windows may set an attribute on the file to "Blocked." When this happens, the file may not load properly. This behavior is the default for .NET 4 and is designed to help protect your computer from executing malicious files. Whenever you download a file, it is recommended that you check for the blocked attribute, and then "Unblock" the file so it will load properly. It is easier to unblock a .zip file rather than unblocking each of the individual files that are extracted from it.

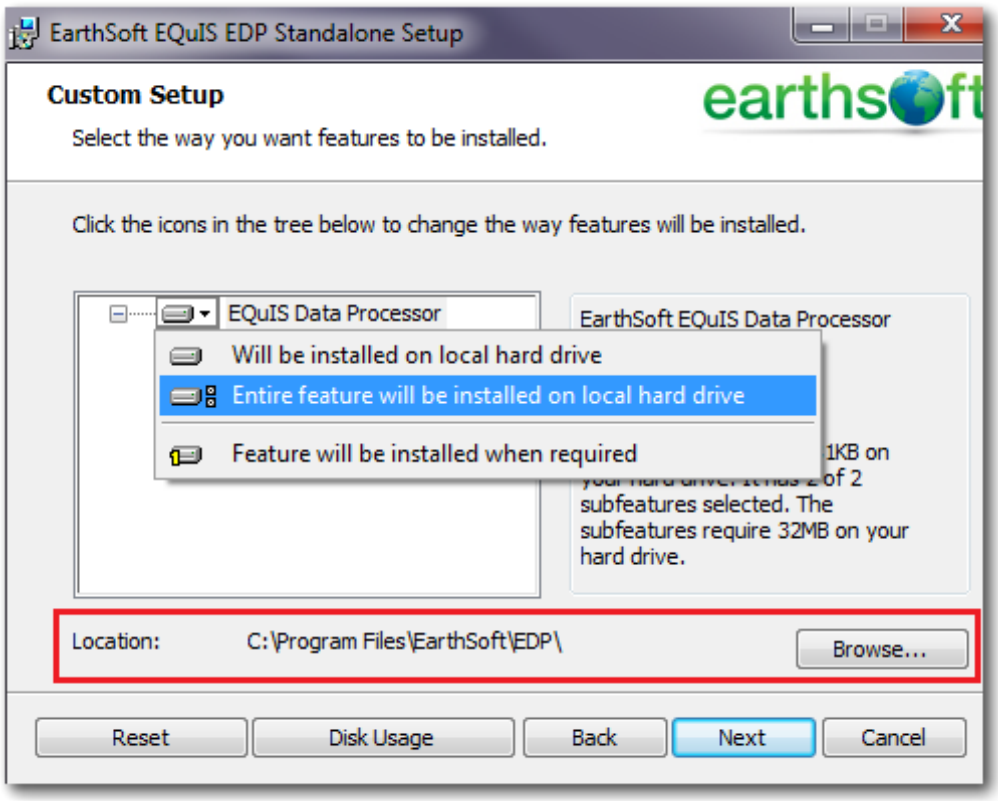
After downloading the file, check to see if it has been blocked by Windows by right-clicking on the downloaded file, and then selecting **Properties** from the drop-down menu. A file properties dialog will be displayed. On the **General** tab, click the **Unblock** checkbox or button and then select the **OK** button to remove this restriction on the file.



After the file has been unblocked, unzip the EDP installation folder. Double-click the EDP application installer to open the Installation Wizard. The install wizard will guide you step-by-step through the installation procedure. It is important to note that during installation you should have no other programs running.



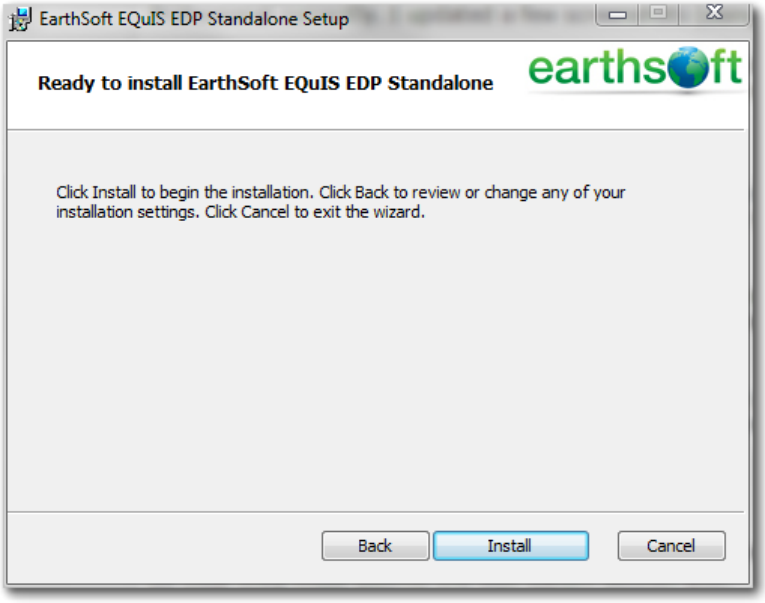
Click the **Next** button. The License Agreement screen will appear. Check the **I accept the terms in the License Agreement** radio button and click the **Next** button to open the Custom Setup screen.



Click the icon next to 'EQuIS Data Processor' and select '**Entire feature will be installed on local hard drive**'. At the bottom of the window, select the destination folder for the application files. Click the **Next** button. The "Ready to install" screen will open.

Click the **Install** button to begin the installation.

When the installation is complete, you will be presented with a window that verifies that EDP Standalone has been successfully installed. Click the **Finish** button to exit the installation.

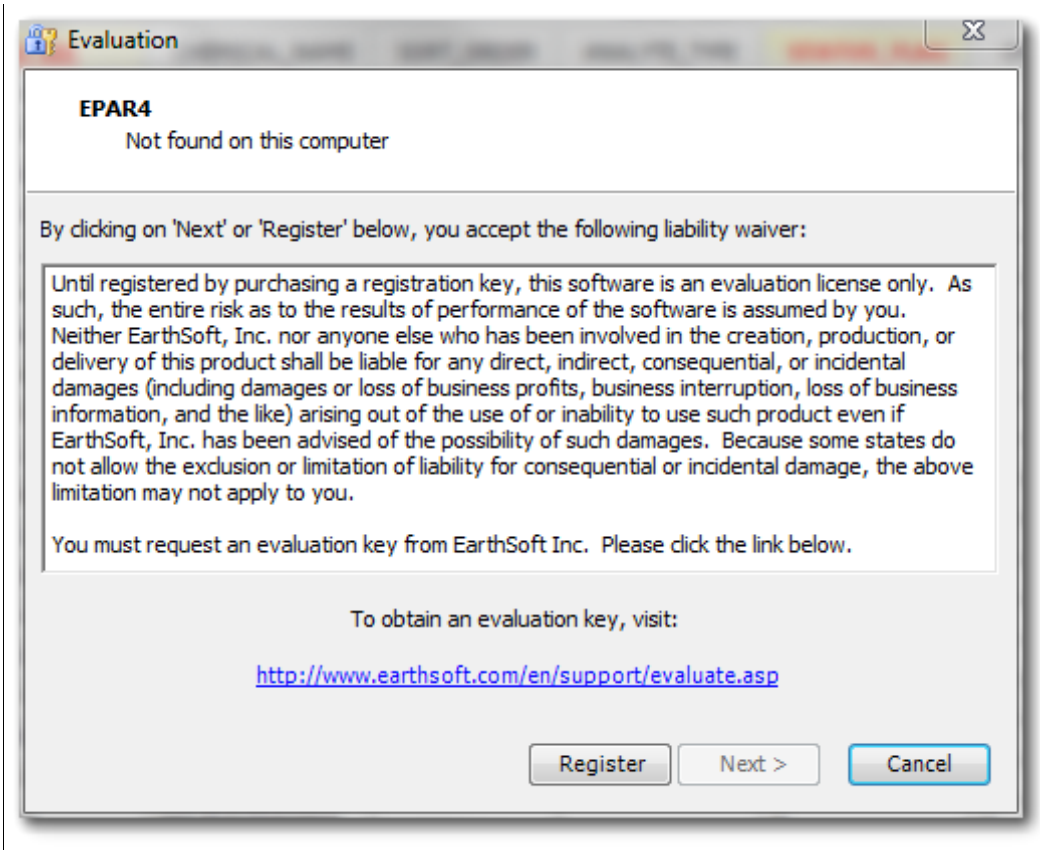


Next, return to the EarthSoft website to download the EPAR4 format file and reference values file. Select the "EPAR4 Format" link and save the *EPAR4.zip* file to the desired folder. Select the "EPA Region 4 Reference Values" link and save the *EPAR4rvfyyymmdd.zip* file to the desired folder. We recommend saving the files to the *C:\Program Files\EarthSoft\EDP\Formats* folder. [If you were required to install the 32-bit version of the software, this path will be *C:\Program Files (x86)\EarthSoft\EDP\Formats*. The remaining references within this document will assume the 64-bit version.] Unblock the zipped format file according to the instructions above, but do not unzip this file. Unblock the EPAR4yyyymmdd.zip file and extract the *EPAR4.rvf* file from the reference values zip file.

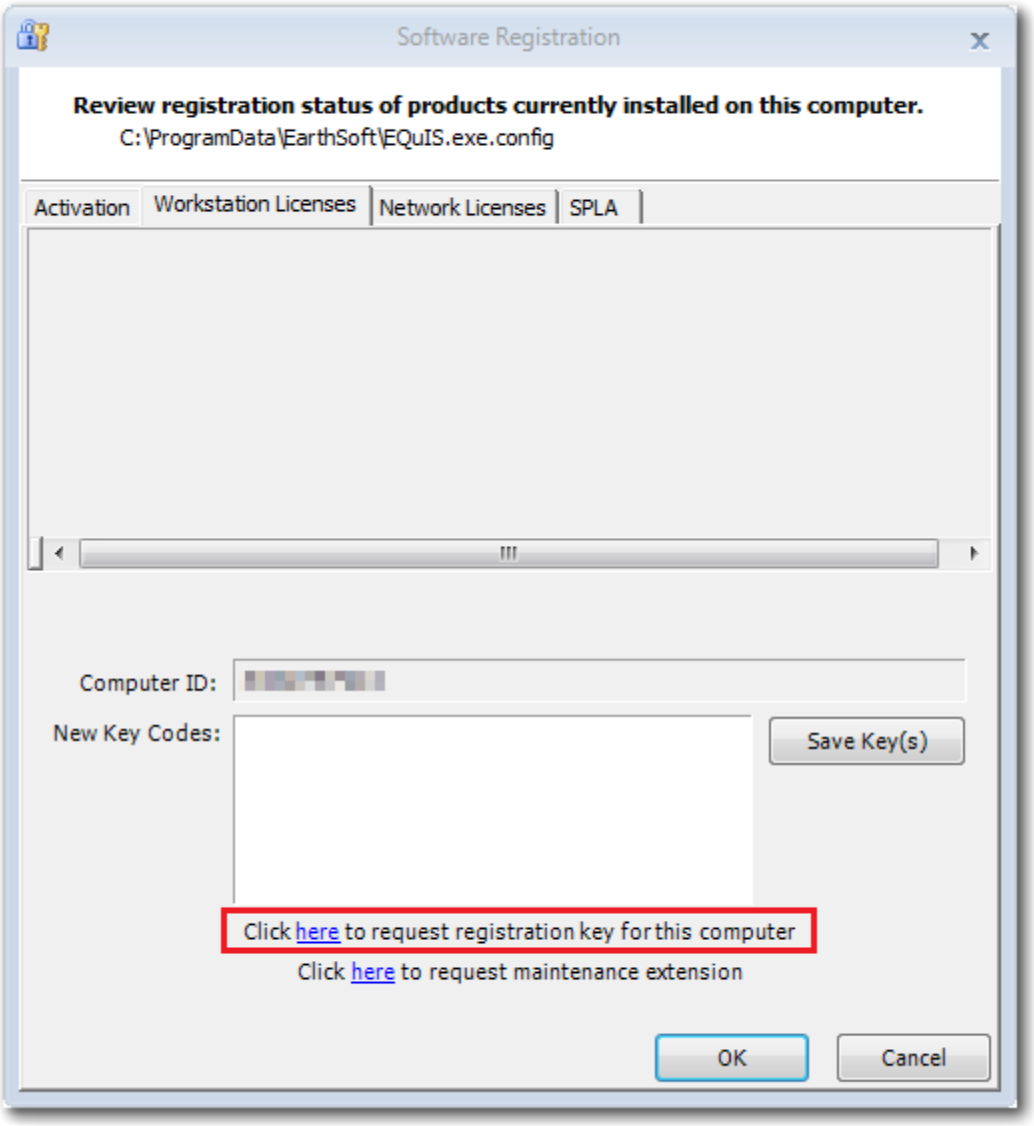
2.3 Registering EDP

Once installed, EDP must be registered. Start the EDP application by selecting **EDP Standalone** from the Windows Start menu. The EDP application will start with a blank workspace. Select **Format** from the Open group of the EDP Home ribbon.

Browse to select the *EPAR4.zip* format file. If you moved the format file to the suggested folder, the path you select will be the *C:\Program Files\EarthSoft\EDP\Formats* folder. Select the file, and then click the **Open** button. The "Evaluation" screen will appear. Click the **Register** button.

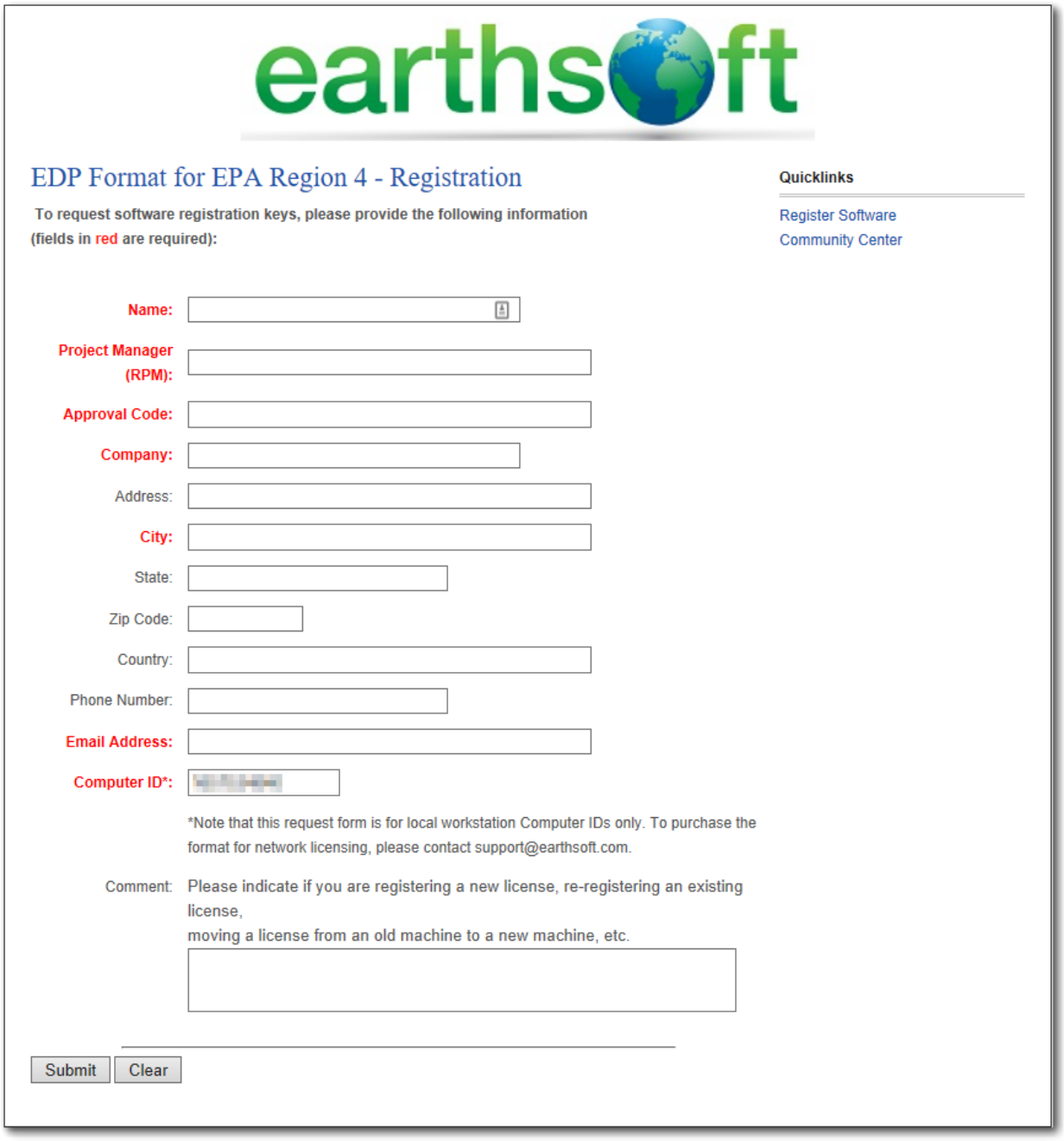


The 'Software Registration' screen will appear. Go to the **Workstation Licenses** tab. Click the first link to open the registration request page in your web browser.



Enter the requested information in the "EDP Format for EPA Region 4 – Registration" form (see image on next page). Your computer ID should be automatically populated. If you are working with several RPMs, enter the primary RPM or the RPM who provided your approval code in the appropriate field of the registration form. You will need to request the approval code from either your RPM or email the R4DART@epa.gov. When all information has been entered, click the **Submit** button. With the proper approval code, the software key codes will be automatically emailed to the email address entered in the registration form.

Once the new key codes have been received, register EDP by copying the registration key codes into the **New Key Codes** box on the **Workstation Licenses** tab of the Software Registration window. Click the **Save Key(s)** button. A screen stating that the "Registration succeeded" should appear. Click **OK**. EDP is now registered and ready for use. Registration is unique to a computer. You will need to repeat this process to use EDP on a different computer.


earthsoft
EDP Format for EPA Region 4 - Registration
To request software registration keys, please provide the following information (fields in red are required):
Quicklinks
Register Software
Community Center
Name:
Project Manager (RPM):
Approval Code:
Company:
Address:
City:
State:
Zip Code:
Country:
Phone Number:
Email Address:
Computer ID*:
*Note that this request form is for local workstation Computer IDs only. To purchase the format for network licensing, please contact support@earthsoft.com.
Comment:
Please indicate if you are registering a new license, re-registering an existing license,
moving a license from an old machine to a new machine, etc.
Submit
Clear

3 Using EDP

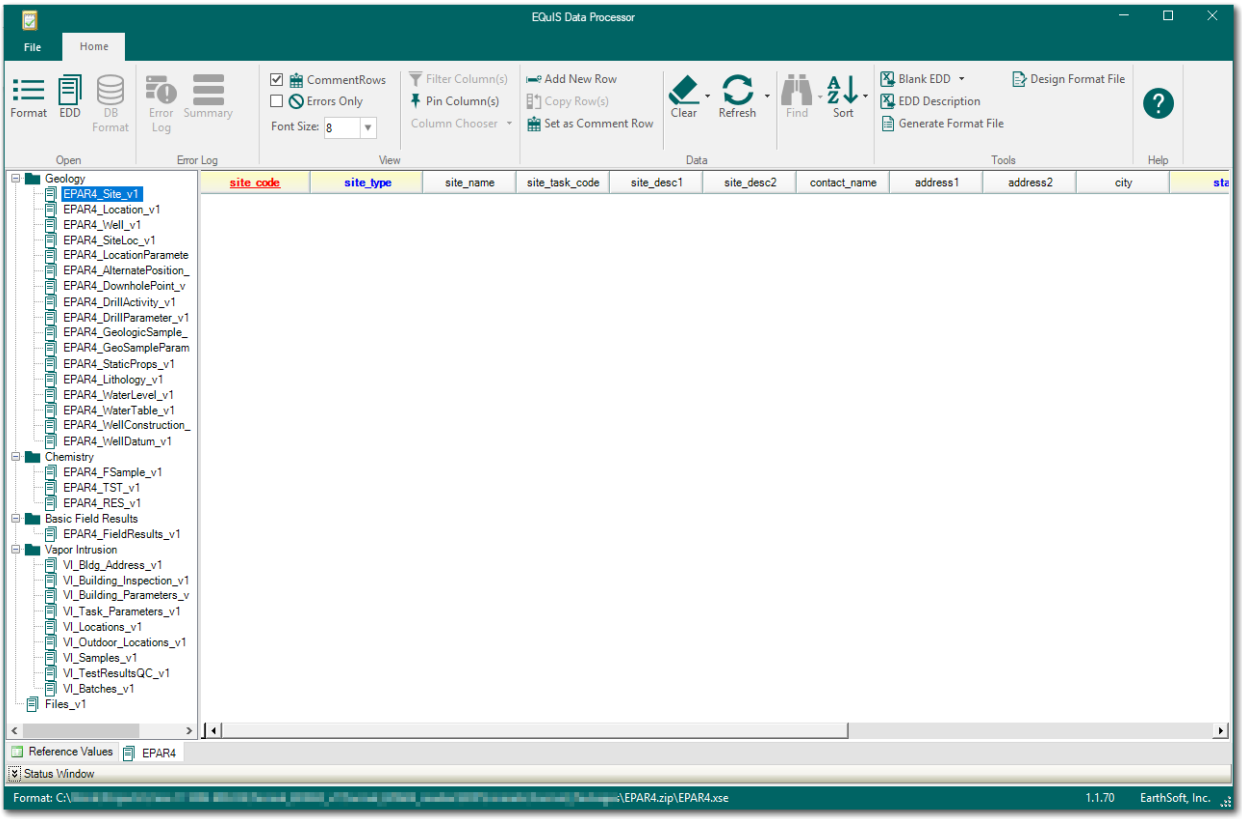
EDP is available as two desktop applications: (1) EDP Standalone, which is available via the download and registration process outlined in Section 2 above; and (2) Professional EDP, which is only available to users who have purchased and licensed EQuIS Professional. This manual only addresses the use of EDP Standalone for the EPA Region 4.

Review of EDDs via EDP is completed via the following workflow:

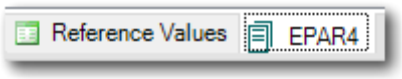
1. Launch EDP and load the EPAR4 Format
2. Load the EDDs
3. Resolve data errors (if any)
4. Sign and Submit the EDD to EPA Region 4

3.1 Start EDP and Load Format

Start the application by selecting **EDP Standalone** from the Windows Start menu. EDP will open. If the Region 4 format file does not load automatically, you will need to load it manually. Select **Format** from the Open group of the EDP Home ribbon. Select the *EPAR4.zip* format file from the *C:\Program Files\EarthSoft\EDP\Formats* folder and click the **Open** button. If requested, also browse to where you extracted the *EPAR4.rvf* file and select the this as your reference values file. The lower-left displays the elapsed time while the software is loading the format and respective reference values.

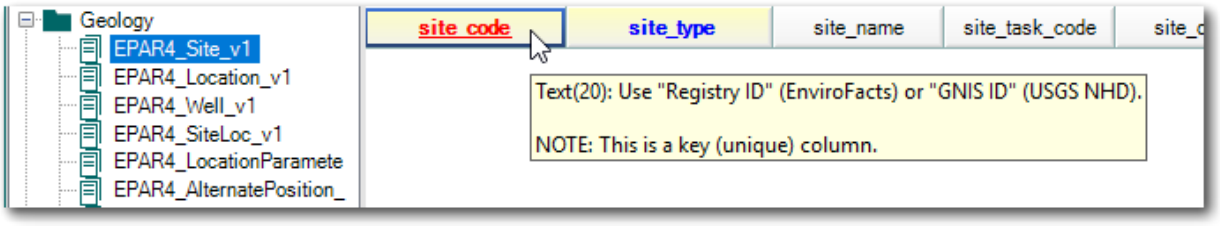


Two tabs are displayed near the bottom of the screen above the Status Window. Select the **EPAR4** tab to view the current EPA Region 4 EDD format specifications you will use to load and check your data.

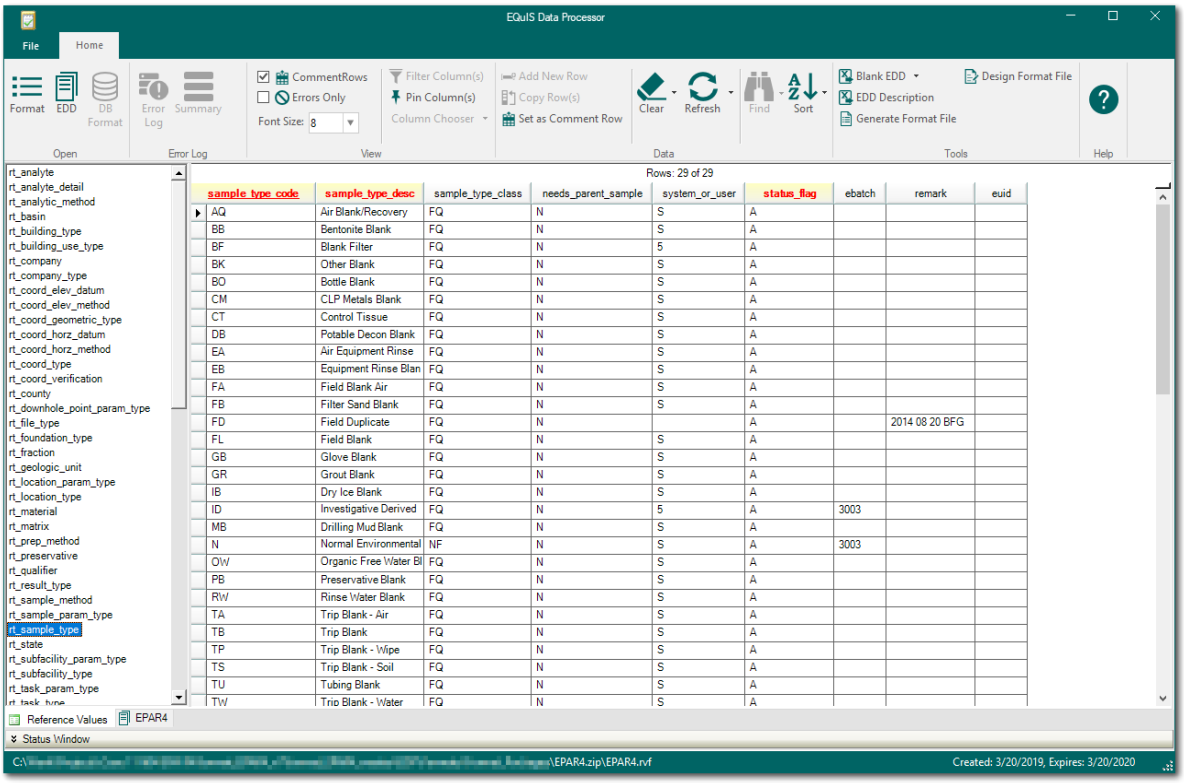


The EPA Region 4 EDD sections are displayed along the left side of the window. Each of the EDD sections listed in EDP corresponds to the EDD files described in the *EPA Region 4 Format File Guide*.

An empty table is displayed in the right pane showing the field names associated with the highlighted section type along the top. Information about each field is provided when the cursor is placed over the column header or field name (as indicated in the example below).



Select the **Reference Values** tab to view the current valid values that are acceptable in the EPA Region 4 database. The valid reference values for the RT_SAMPLE_TYPE table are displayed in the example image below.



3.2 Blank EDDs and EDD Descriptions

The Blank EDD and EDD Description tools within EDP can facilitate data processing. The Blank EDD tool creates an empty Excel template that follows the format file's specific layout for data entry. The EDD Description tool defines the specifications of the format. These tools can be used to help generate EDDs to submit to the EPA Region 4.

To create a blank EDD in the EPAR4 format, select **Blank EDD** from the Tools group of the Home ribbon. Selecting the arrow next to the Blank EDD option will provide the ability to export a blank EDD with either Valid Values or Parent Values (or both). In the Save dialog, name the blank EDD file, choose the desired location, and select **Save**.

A Microsoft Excel file is created with the format columns and field types, along with the color-coding as displayed in EDP to indicate required fields, primary keys, and look-up values. Separate worksheets are named for each section of the format file. The blank EDD file may be loaded into EDP for data checking after data has been entered.

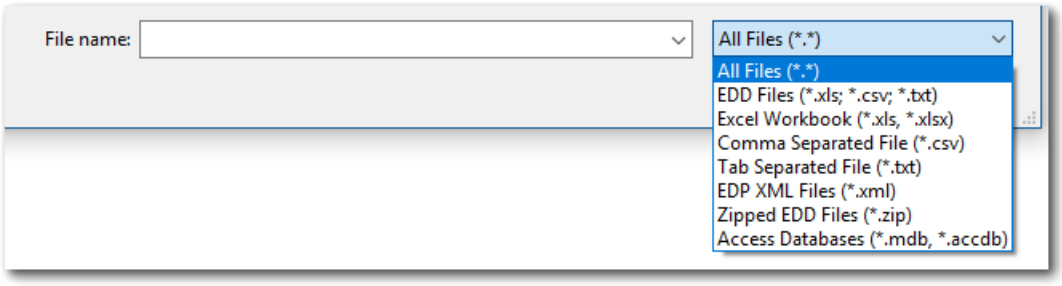
To create an EDD description file of the EPAR4 format, select **EDD Description** from the Tools group of the Home ribbon. In the Save dialog, name the EDD description file, choose the desired location, and select **Save**.

Information about the format file is exported into a Microsoft Excel spreadsheet, including whether a field is a primary key (i.e., required), data type, related look-up tables, and EQuIS Schema mapping. The EDD description file includes a separate worksheet detailing each section of the EDD format file.

3.3 Loading EDD Files

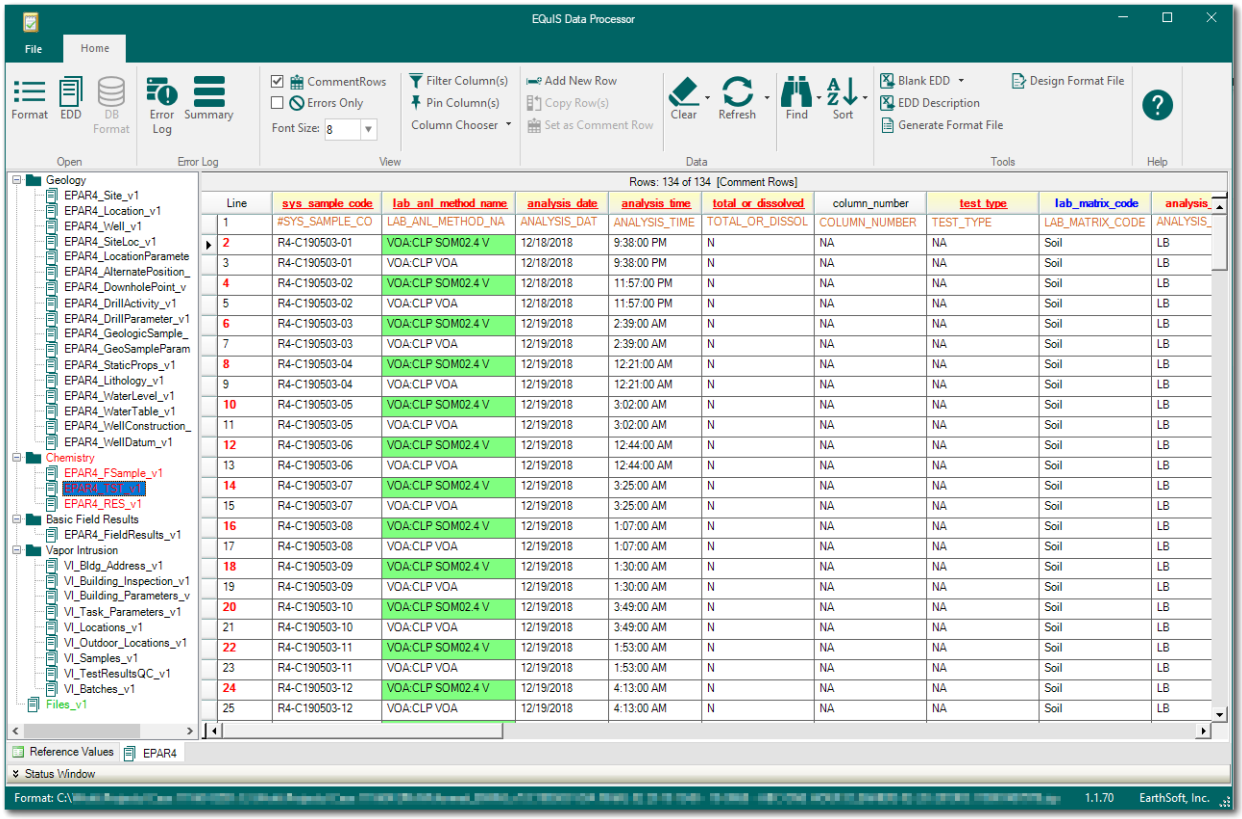
A variety of EDD file types can be loaded into EDP and checked, including:

- Individually created EDD files.
- A single Access database created with individual tables named according to the naming conventions.
- An Excel spreadsheet with tabs named according to the naming conventions.
- A zip package of individual EDD files.



To load EDD files, select **EDD** from the Open group of the EDP Home ribbon. Use the Browse window to locate the EDD file and select **Open** or double-click on the file to open it. The data file will load to EDP and be checked for errors during loading. This make take a few minutes depending on the size of the dataset. A loading dialog box will display progress of the EDD.

Data will load into one or more sections of the format, depending on which sections were populated in the EDD file. Populated section names will be Red or Green, depending on if they contain errors or are error-free, respectively. Empty sections of the format will remain black. Data will be displayed in the tables and any detected errors will be shaded.



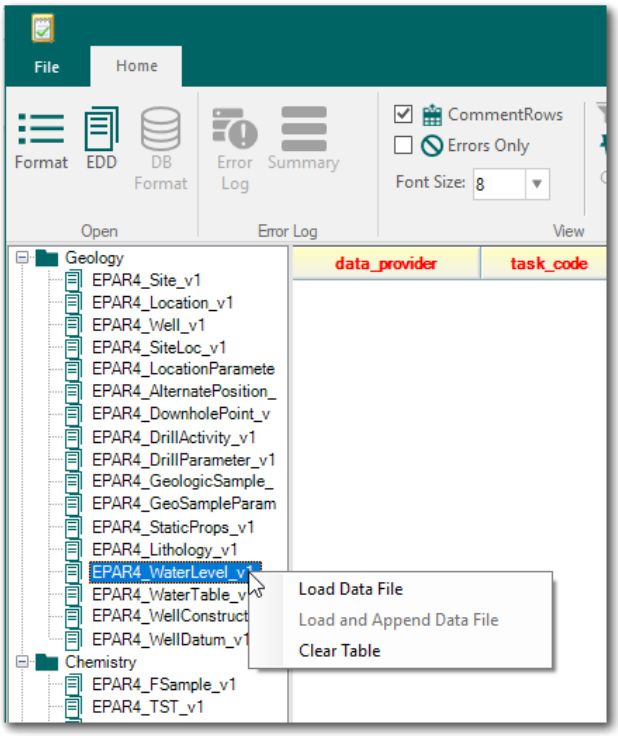
Warning: If the data file contains header rows, EDP will identify fields in the header rows as errors, unless each header row is preceded by the default hash character (#) in the first column.

Note: To be 508 Compliant, add bitmap images for the appearance of errors (see Section 3.3.2 to edit appearance options and view examples).

3.3.1 Load EDD File Directly to Section

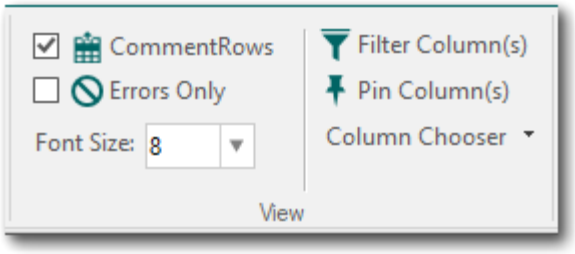
A single EDD can be loaded directly to a format section by right-clicking on the desired format table in format list. Select **Load Data File** to open the file browser. Use the Browse window to locate the EDD file and select **Open** or double-click on the file to open it. The selected EDD file will load directly to the format section and be checked for errors during loading. Data will be displayed in the table and any detected errors will be shaded.

If data already exists in the table, the **Load and Append Data File** option will be visible.



3.3.2 Managing EDD Display

A variety of display options are available in the View group of the EDP Home ribbon. These features can assist you with seeing and resolving errors in the EDD.



Display of comment rows is controlled by checking and unchecking the **Comment Rows** box.

To view only the data rows with errors, check the **Errors Only** box. To restore all the rows, uncheck the **Errors Only** box. EDDs may contain thousands of records and large EDDs with an exorbitant number of errors may cause EDP to appear as not responding when switching from “Errors Only” to viewing all data.

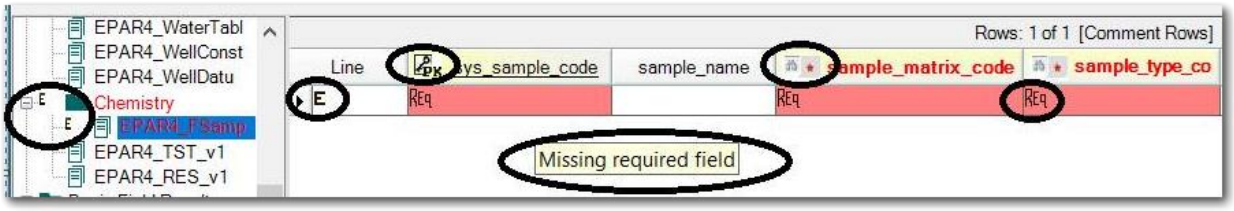
Some tables have many columns of data. Use the **Pin Column(s)** option to be able to pin any column to the left-hand side of the table's grid view to make viewing data easier.

Filter a column by any value within that column by selecting the **Filter Column(s)** option. A filter search box appears below each column header in the grid view. If the field is numeric, then the filter operator appears as "=". If the field is text, then the filter operator appears as "A".

Tip: The default appearance of the Column Headers, Errors, and Informational messages can be modified to enhance the user experience and/or be 508 compliant. This functionality can be accessed by clicking on the EDP **File** menu, then selecting the **Options** button and expanding the **Appearance** menu.

Changing the appearance to add an image provides the user with multiple means to denote an error or understand the significance of a field, such as being a primary key. Within the Appearance menu, select the column header or error type to change and then select **Image**.

In the examples below, a symbol has been added in the header to denote a required field and symbols have been added to indicate the presence of errors. Hovering the mouse will pop-up the error. The 'Req' text within the sample_matrix_code of the image will not affect the data and will remove itself once the cell error has been corrected.



3.4 EDD Data File Checks

EDP checks for errors both within a single EDD and between related EDD files. EDP highlights errors with different colors to signify the error type, facilitating error correction. Errors may be corrected easily and directly within EDP.

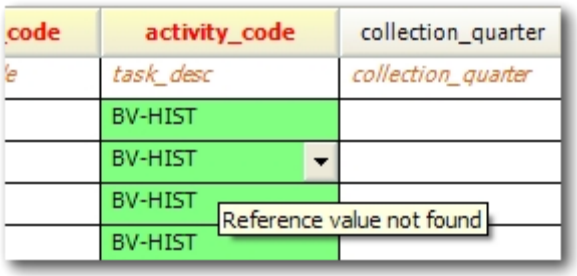
EDD errors often occur because of issues with formatting, logic, or terminology. Along with the EPA Region 4 business rule verifications, EDP checks data for the following potential issues:

- Reference Values
- Field Lengths
- Required Fields
- Data Types
- Valid Dates
- Duplicate Rows
- Parent-Child Relationships

This section describes these errors in more detail. Section 3.6 provides the process to resolve the errors.

3.4.1 Reference Value Not Found

The value in the field does not match the values listed in the reference file downloaded from the EPA. If the field value is correct, research to ensure that the value does not already exist as a valid reference valid written in a different way (e.g., analytes have many synonyms, and analytical methods may be written in similar

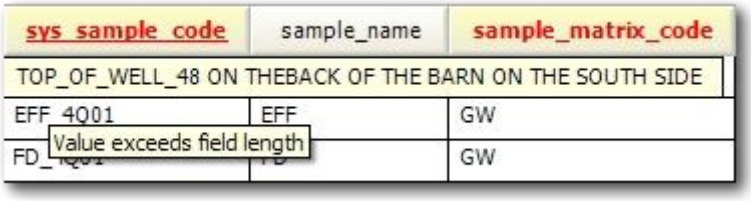


ways but reference the same method). After careful consideration and research, if the value is correct and not listed as a reference value, follow the guidelines in Section 2.6 of the *EPA Region 4 Format File Guide* to request that the value be added to the EPA Region 4 valid value tables. Send that request along with the necessary accompanying information to R4DART@epa.gov. The DART Coordinator will review the request and forward it, if appropriate, to the correct administrator for review and inclusion in the system.

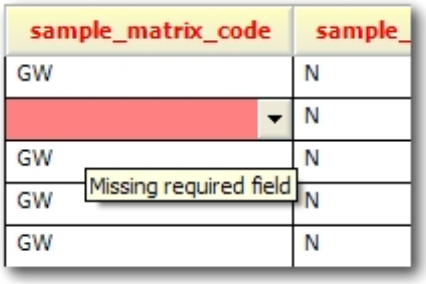
Do NOT submit your data until the request has been approved and you are notified that the values were added. Doing so may cause your data to be rejected for failing to pass EDP. If a new *EPAR4.rvf* file is not provided, check the website for and download the new one when available. Replace the old *EPAR4.rvf* file with the updated file and recheck your EDD before submitting to R4DART.

3.4.2 Value Exceeds Field Length

The number of characters of the value entered in the field exceeds the maximum allowed number of characters. Place your cursor over the column header to view the description of the field that will include the field length. For further detail of each field, see Section 3 of the *EPA Region 4 Format File Guide*.



3.4.3 Missing Required Field



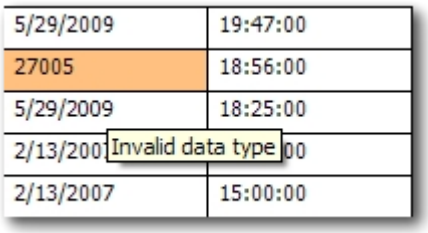
The field *must* be populated with a value. The field cannot be left null (i.e., blank). See Section 3 of the *EPA Region 4 EDD Format File Guide* for information on required fields.

Note that the field name at the top of the column is written in red. This indicates that the field is required and that the EDD will not pass EDP unless all values in this column are populated correctly. Placing your cursor over the column header will also bring up the description that includes the line: "NOTE: This is a required field."

A common problem arises with the CAS RN field during the conversion process from one type of file to another, such as comma separated to Excel. In this process, any CAS RN that may appear as a date may be converted to a date. For example, the CAS RN for Potassium is 7440-09-7 and will be converted to 9/7/7440.

3.4.4 Invalid Data Type

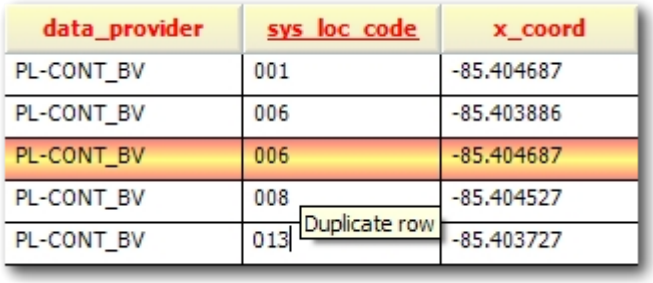
The value is not the appropriate data type. Each field has a specific data type that must be used, such as text, date/time, or numeric. If the appropriate data type for a field is Date/Time, then the value must be a valid date format such as the MM/DD/YYYY HH:MM:SS format. See the Data Type description in Section 3 of the *EPA Region 4 EDD Format File Guide* for the appropriate data type.



Another common problem arises with Date/Time fields during the conversion process from one type of file to another. In this process, Date/Time fields may be incorrectly converted to an integer Julian date. Be sure to check your Date fields to make sure they are appropriately classified to avoid errors.

3.4.5 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 EPAR4_Location_v1 file has the sys_loc_code field as the primary keys. Two records that both have 006 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 006.

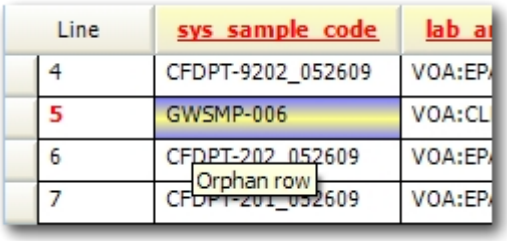


Laboratories frequently report data from the same event in multiple packages, sometimes creating duplications of sample records. In these cases, if all data is processed through EDP at the same time, duplicate records will appear in the EPAR4_FSample_v1 EDD. These duplicate records will need to be deleted prior to submitting the data to EPA Region 4. Data will not pass the EDP checker with duplicate rows.

Refer to Section 2.4 of the *EPA Region 4 EDD Format File Guide* for further discussion of data integrity and duplicate records.

3.4.6 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 accurately and the parent record must exist in the corresponding file.

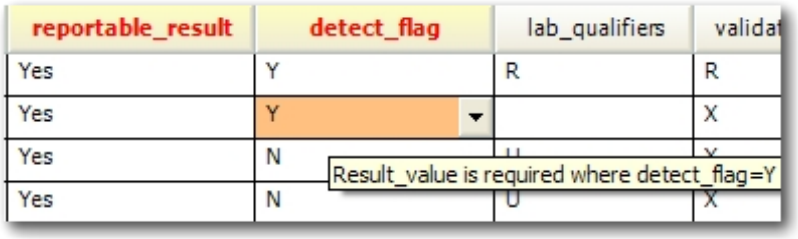


For example, each row in the EPAR4_TST_v1 table must include a sys_sample_code that corresponds to a sys_sample_code reported in the EPAR4_FSample_v1 table. If a record in the EPAR4_TST_v1 table has a sys_sample_code of GWSMP-006, then a record must also be included in the EPAR4_FSample_v1 table with a sys_sample_code of GWSMP-006. If a record in the EPAR4_TST_v1 table has a sys_sample_code that is *not* included in the EPAR4_FSample_v1 table, an "Orphan Row" error will be identified.

Likewise, each row in the EPAR4_RES_v1 table must have a matching "parent" row in the EPAR4_TST_v1 table. There are six (6) fields that establish the relationship between the results and test records, and they are: sys_sample_code, lab_anl_method_name, analysis_date, analysis_time, total_or_dissolved, and test_type. See Table 2-5 and Section 2.3 of the *EPA Region 4 EDD Format File Guide* for further discussion of child/parent records.

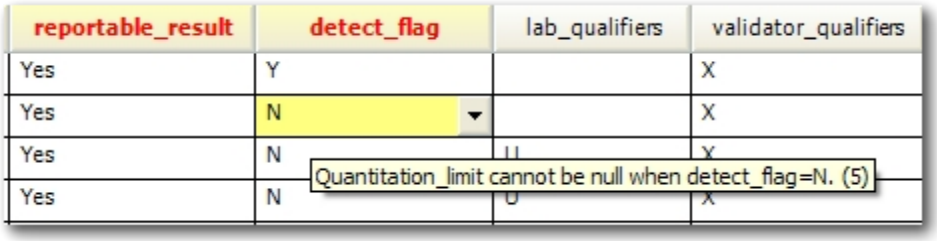
3.4.7 Result_value Is Required When detect_flag = Y

Identifies records that have the detect_flag (EPAR4_RES_v1) value of "Y" yet there is no value reported in the result_value field. 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 (EPAR4_RES_v1) and the detect_flag has a value of "Y", the result_value field must be populated with the numeric result value (i.e., field cannot be left blank).



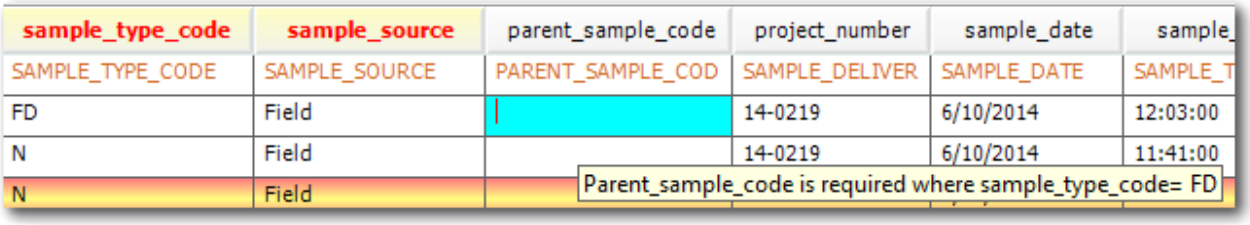
3.4.8 Quantitation_limit Cannot Be Null When detect_flag = N

Identifies records with a detect_flag (EPAR4_RES_v1) value of "N" and the quantitation_limit field is null. All records that have a value of "N" in the detect_flag field must have the quantitation_limit field populated with the appropriate detection limit value (i.e., field cannot be left null).



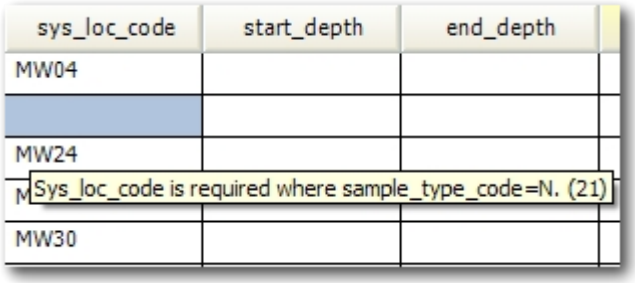
3.4.9 Parent_sample_code Is Required Where sample_type_code = FD

Identifies records that have a sample_type_code (EPAR4_FSample_v1) of "FD" but are missing the appropriate parent_sample_code. In the image below, the 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 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 EPAR4_FSample_v1 file (i.e., there must be a record for the original sample and a separate record for the duplicate sample).



3.4.10 Sys_loc_code Is Required When sample_type_code = N

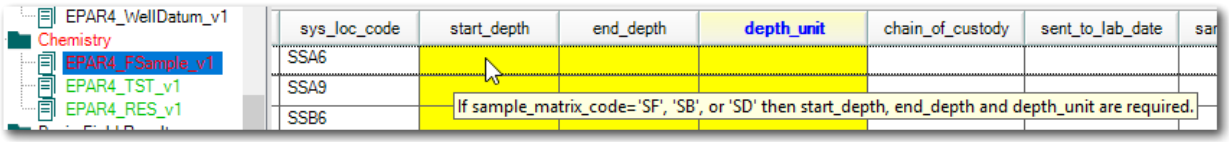
A location identifier (i.e., sys_loc_code) must be provided for all samples that are normal environmental samples. Therefore, all records in the EPAR4_FSample_v1 file that have a sample_type_code of "N" must also have the sys_loc_code field populated (i.e., this field cannot be left blank).



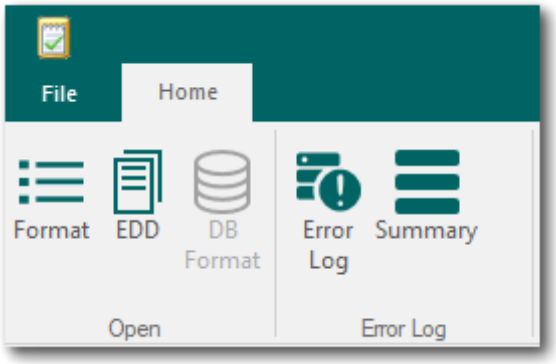
A matching sys_loc_code with coordinates in the WGS84 (Longitude/Latitude) system must exist in the database or be submitted within the EDD package with the corresponding samples. Data Providers using the Scribe software are required to populate #R4DART# for the station name of various field QC type samples. Before submitting the EDD to R4DART, you must remove this placeholder from your data.

3.4.11 Depths are Required When sample_matrix_code = SF, SB, or SD

Identifies records that have a sample_matrix_code (EPAR4_FSample_v1) of "SF", "SB", or "SD" and has null values for start_depth, end_depth and/or depth_unit. If sample_matrix_code = SF, SB, or SD, then start_depth, end_depth and depth_unit are required.



3.5 Error Logs



It is easy to generate a summary list of data errors and reference value errors to assist in troubleshooting. EDP produces an error log that can be saved as an HTML formatted file.

On the EDP Home ribbon in the Error Log group, 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 example below). Use the Browse window to locate the desired location and select **Save**. The error log file will then be saved in the selected folder.

Table	# of Rows	Column	Value	Message	Type
EPAR4_FSample_v1	1	sample_matrix_code	[NULL]	Missing required field	ERROR
EPAR4_FSample_v1	1	sys_loc_code	[NULL]	Sys_loc_code is required where sample_type_code='N'.	ERROR
EPAR4_FSample_v1	1	sample_matrix_code	BadMatrix	Reference value not found	ERROR
EPAR4_FSample_v1	1	parent_sample_code	[NULL]	Parent_sample_code is required where sample_type_code= FD	ERROR
EPAR4_FSample_v1	1	~	~	Duplicate row	ERROR
EPAR4_TST_v1	1	sys_sample_code	R4-E142404-01ORPAN	Orphan row	ERROR
EPAR4_RES_v1	1	detect_flag	Y	Result_value is required where detect_flag='Y' and result_type_code='TRG', 'TIC' and CAS_RN <> 'R4-6500'.	ERROR
EPAR4_RES_v1	20	sys_sample_code	R4-E142404-01	Orphan row	ERROR
EPAR4_RES_v1	20	lab_anl_method_name	OCP:EPA 8081	Orphan row	ERROR
EPAR4_RES_v1	20	analysis_date	7/3/2014	Orphan row	ERROR
EPAR4_RES_v1	20	analysis_time	04:10:00	Orphan row	ERROR
EPAR4_RES_v1	20	total_or_dissolved	N	Orphan row	ERROR
EPAR4_RES_v1	20	test_type	NA	Orphan row	ERROR

3.6 Resolve Errors

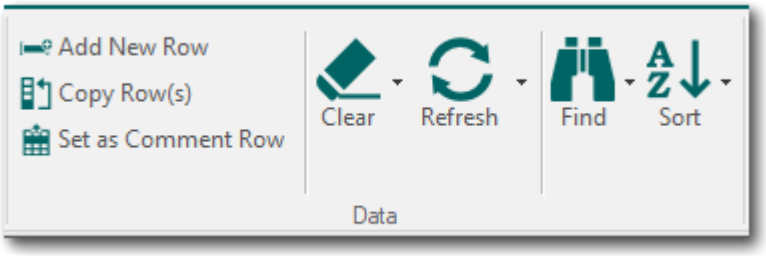
Data are checked for errors by EDP as the EDD files are loading. Populated section names will be Red (or display the image added with the Appearances Option) or Green (or display no image), depending on if they contain errors or are error-free, respectively. For any sections with errors, troubleshoot and resolve data errors.

Single-click on the top-most Red section to view the data in that section. Since data in the various tables are related, it is important to resolve data errors in a top-down manner. The fields with errors will be shaded different colors depending on the type of error. Place the cursor over an error to show a description for the type of error.

Note: Two different cells with the same error type may or may not be resolved the same way.

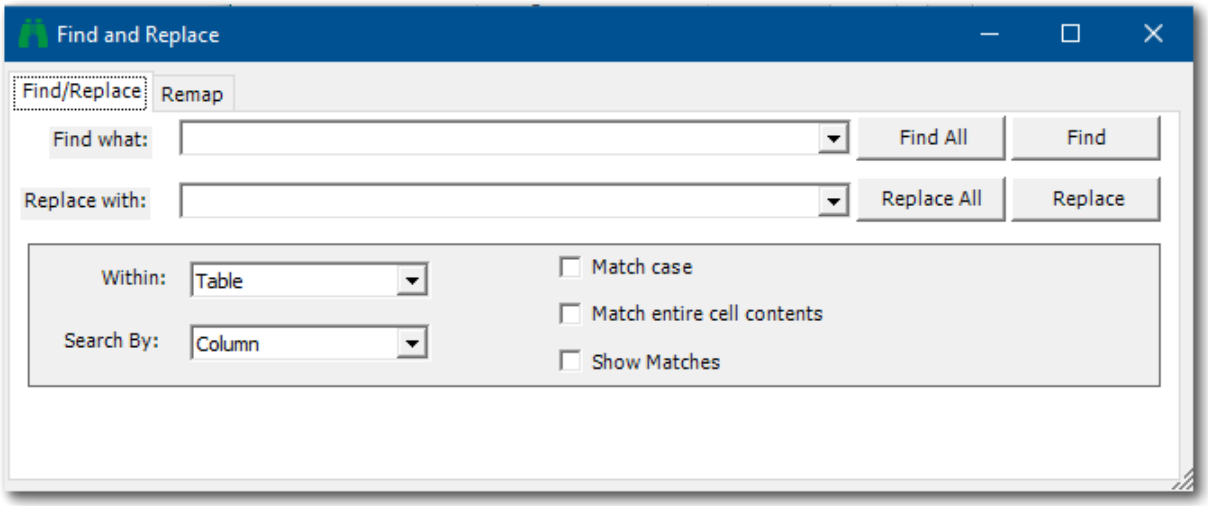
3.6.1 EDP Tools

EDP provides tools in the Data group on the Home ribbon that can assist with resolving data errors.

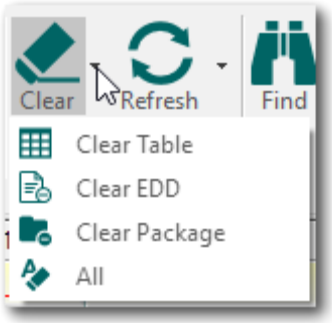


If an intended header row appears as an error, highlight the header row by clicking to the left of the row and then select the **Set as Comment Row** button. This will add hash character (#) in the first column of the row to set as a comment row.

The "Find and Replace" function allows users to search the file for a specified value and then replace it with another value. This function is useful when there are a number of similar values that need to be changed. Select the **Find** button to open the "Find and Replace" dialog box. Type the value to be replaced in the "Find what" field and the new value in the "Replace with" field. Select the desired search options and then click either **Replace** or **Replace All**.



To clear the data from EDP, select the **Clear** drop-down button and then select **Clear EDD**. The EDD file will be cleared from the EDP viewer. Data can also be removed from individual tables.



Note: Clearing data from EDP will not delete the EDP file; the action only removes the data from the viewer.

3.6.2 Methods to Resolve Errors

Two methods can be used to correct errors within EDD files: Edit files outside of EDP or correct errors within EDP.

1. Exit EDP and then open the EDD file using a text editor or spreadsheet application, correct the errors, resave the file, and then re-load the EDD into EDP to ensure no further errors. The detailed error log file is extremely useful for this method.

2. The preferred method is to correct errors directly within EDP. Correct errors by clicking in the field with the error and entering 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 right-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. It may be necessary to highlight the row by clicking on the tab at the far left of the row and then selecting **Refresh> Selected Rows** from the Data group on the Home ribbon. This will "reload" the checker for just that row with the new data entered. You may also refresh the entire table. For large data sets, refreshing the entire table may take a long time and may cause EDP to appear as not responding while it processes.

While EDP is a robust data checker, EDDs should also be reviewed for potential errors that might not be flagged as errors by EDP. Examples of these other potential issues are incorrect sampling dates, matrix types being inconsistent with sample type, etc.

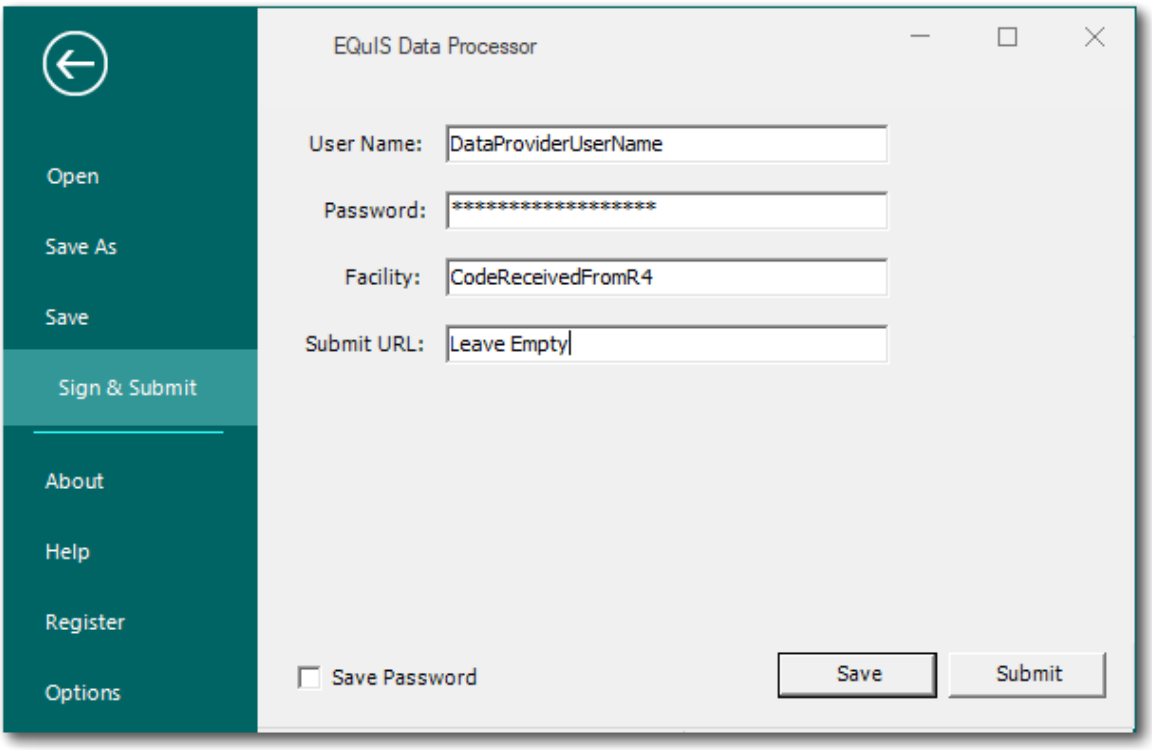
3.7 Saving Changes to the EDD File

To save the changes made to the EDD data file, click the EDP **File** Menu 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: When saving as an individual file type such as .csv or .txt, only the individual EDD table selected in the left-hand format list will be saved. To save all of the EDD tables to one package, change the "Save as Type" to one of the selections that will allow for multiple tables, such as an Excel workbook, Zipped EDD file, or Access database.

3.8 Sign and Submit

The "Sign and Submit" function of EDP allows all files within an EDD to be compiled into the final Data Package, which is subsequently submitted to the EPA Region 4. Click on the EDP **File** Menu and select **Sign & Submit**.



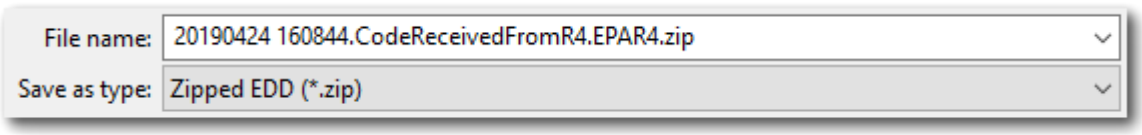
During the "Sign and Submit" process, EDP will export and name all loaded tables according to EDP's required file naming conventions before compressing them into a single archive zip file. Also included in the .zip file is the Data Provider's .usr file, referred to as the user certificate.

The user certificate is created automatically using the User Name and Password as entered into the Sign and Submit form. Your user name and password *must* match *exactly* the user name and password assigned in the R4DART EQuIS Enterprise for each Data Provider. If you do not have a valid user name and password for DART, email R4DART@epa.gov and provide the Site, RPM, data provider company, and data provider email address.

You should also have been notified of the facility code for the site you are preparing the EDD submittal. Enter the facility code into the Facility box.

If you have not been assigned this information or received the proper facility code, contact the R4DART coordinator before completing this process to submit your data.

After entering the correct authentication information, click the **Save** button. [Note that the Submit option with the Submit URL is not used by EPA Region 4]. Select the folder location where you would like the zip archive file to be saved.



Data Providers need to update the default filename of the Data Package being created to use the EPA Region 4 required EDD file naming convention. Update the prefix of the filename from the date/time stamp to include the following pieces of information connected with underscores (see example filename below):

- Task Code (PRP use PYYYY-nnnn as provided by R4DART; Fund-lead use YY-nnnn as provided by EPA work order)
- Site Name
- Company Code
- Date Submitting (YY MM DD)
- Type (Loc-Location/Ch-Chemistry/FR-FieldResults/WL-WaterLevel/GEO-Geology/ WC-Well and Well Construction /VI-Vapor Intrusion)

Click the **Save** button to save the archive zip file to the selected folder.

Note: The EPA electronic mail system will reject files with a .zip format due to security concerns. To match how the EPA Region 4 laboratory submits EDDs, you must rename the EDD file extension from .zip to .dat before emailing the EDD to the EPA Region 4. **Files must be submitted with the ".dat" extension**.

Note: The EPA electronic mail system will reject files with more than three (3) periods. Thus, it is important to connect the file naming elements using underscores rather than periods.

Example EDD Filename

P2015-0001_ChemicalCompanyABC_PL-DataProvider_19 01 08_Ch.110001224773.EPAR4.dat

where,

- Task Code is "P2015-001"
- Site Name is "Chemical Company ABC"
- Company Code is "PL-DataProvider"
- Date Submitting is January 8, 2019 as "190108"
- Type of data submission is "Ch" for Chemistry
- Facility Code is "110001224773"
- Format used is "EPAR4"
- File extension has been changed from ".zip" to ".dat"

4 Submitting and Resubmitting the EDD Data Package

After the EDD files have been checked by EDP and the "Sign and Submit" process has been completed, the Data Package is ready for submittal to EPA Region 4. Data packages are submitted by attaching the .dat file to the original email requesting the data and forwarding the email to EPAR4@EQuISonline.com. This may either be the Project Log Summary report email or the Request for Data Submittal Summary template email. If the file is named improperly or not attached to the requesting email, it will be returned to the data provider.

Some causes of common data rejections are:

1. When Location EDDs are not submitted prior to submitting environmental chemistry,
2. Geology EDDs are submitted when the station locations do not already exist in the EPA Region 4 database, and
3. If depths are missing for soil and sediment sample types and the Scribe COC XML file was not sent to R4DART.

Another type of request for additional data from your RPM may be due to Water Level EDDs submitted when there is no corresponding Well Datum information to calculate the water level elevations. The Well Datum EDD is required when submitting water levels for the first time and any time a well is modified.

Fund lead projects begin with an automated email being sent to the project leads for projects scheduled through R4LIMS. For these types of projects, there will be a list of existing station IDs and their related coordinates sent to you as they exist for the site in the EPA Region 4 EQuIS database.

If you receive an error log with information that your EDD submission was rejected and need to send a corrected EDD, you are to use the initial file name with simply a revision character between the Type and Facility Code file name sections. Acceptable examples of second and third submissions necessary due to errors in the EDD for the same data are provided below:

Initial submitted file:
P2015-0001_ChemicalCompanyABC_PL-DataProvider_190108_Ch.110001224773.EPAR4.dat

Second submitted file of the corrected initial submission:
P2015-0001_ChemicalCompanyABC_PL-DataProvider_190108_Ch_**A**.110001224773.EPAR4.dat

Third submitted file if the second is also rejected:
P2015-0001_ChemicalCompanyABC_PL-DataProvider_190108_Ch_**B**.110001224773.EPAR4.dat

5 Updating the Reference Value File

Periodically, the EPA Region 4 will post an updated reference value file (.rvf) on the EarthSoft website. Follow the steps below to update the reference values in the EDP application:

1. Download the most recent reference value file from the EarthSoft website: https://earthsoft.com/products/edp/edp-format-for-epar4/.

2. Create a backup of your existing EPAR4 format files.

3. Replace the existing reference value file, *EPAR4.rvf*, with the recent downloaded file.
    a. If you maintain the format within a .zip, you may update the .zip file by coping the *EPAR4.rvf* into the zip replacing the existing file.
    b. If you extracted the files to a folder, replace the *EPAR4.rvf* within the folder.

4. The next time EDP is started, the new reference values will be loaded.

6 Updating the Format File

When the EPA Region 4 makes changes to their EDD format, the EDP application will need to be updated with a new format file. Follow the steps below to update the EPAR4 format file in the EDP application:

1. Download the most recent format file from the EPA Region 4 EarthSoft website: https://earthsoft.com/products/edp/edp-format-for-epar4/.

2. Create a backup of your existing EPAR4 format files.

3. Replace the existing EPAR4 format file(s)
    a. If you maintain the format within a .zip, replace the .zip file.
    b. If you extracted the files to a folder, extract and replace all the files into the folder.

4. The next time EDP is started, the new format files will be loaded.

Note: The EPA Region 4 does not expect to make changes very often; however, if changes are made, R4DART will provide notification of the format changes via email to all data providers at the email address you submitted when beginning EQuIS database work on the site.