

EARTHSOFT - DNREC

EQuIS Data Processor Tutorial



Contents

1. THE STANDALONE EQUIS DATA PROCESSOR	3
1.1 GETTING STARTED	3
1.2 FORMAT FILE	4
1.3 UNDERSTANDING THE DNREC IMPORT FORMAT	5
1.4 REFERENCE VALUE FILE	6
1.5 DATA FILES	8
1.6 LOADING INDIVIDUAL CSV OR TXT DATA FILES.....	9
1.7 LOADING EXCEL, ACCESS, OR ZIP DATA FILES	9
1.7.1 Excel Files	9
1.7.2 Access Database Files	11
1.7.3 ZIP Data Files	11
1.7.4 Load and Append Multiple Data Files	12
1.8 CLEAR TABLES.....	13
1.9 FINDING ERRORS	13
1.10 SAVING MODIFICATIONS	15
1.11 ERROR RESOLUTION	15
1.11.1 Column Headers Errors.....	15
1.11.2 Reporting Reference Value Errors and Error Logs	16
1.12 TOOLS	18
1.13 SIGN AND SUBMIT	20
1.14 EDP OPTIONS	21
1.14.1 Remember Last Format File	21
1.14.3 EDP Application Appearance.....	23

[This manual is a guide for DNREC - Delaware Department of Natural Resources - on usage of the Standalone EQuIS Data Processor with DNREC's EQuIS database]

1. The Standalone EQUIS Data Processor

The Standalone EQUIS Data Processor (EDP) is used to check data prior to loading into the Delaware Department of Natural Resources' (DNREC's) EQUIS database. EDP was designed to assist data providers with checking their data prior to submitting it to their clients to ensure high data quality as well as to streamline the data submittal process.

EDP checks all of the following data quality issues:

- Required Fields
- Field Lengths
- Data Types
- Valid Dates
- Reference Values
- Duplicate Rows
- Range Checking

1.1 Getting Started

When using EDP, data providers need to have the DNREC format files, DNREC's reference values, and the data they are submitting in the DNREC Electronic Data Deliverable (EDD) format. This section describes how to use EDP Standalone for checking and submitting data to DNREC. You can obtain the DNREC format files and reference values files online or by contacting DNREC directly.

DNREC EDP Download page:

http://www.earthsoft.com/en/downloads/DNREC/download_EDP.asp

DNREC EDP Registration page:

http://www.earthsoft.com/en/support/edp_DNREC.asp

After you obtain these format files, EarthSoft recommends that you store the files in a folder designated for DNREC in the C:/Program Files/EarthSoft/EQUIS/Formats directory.

Start EDP by selecting **Programs > EarthSoft > EQUIS Data Processor** from the **Start** menu. If you are running the EQUIS Data Processor in evaluation mode, click **OK** on the **EQUIS Data Processor Evaluation** screen. The EQUIS Data Processor opens on the main EDP screen.

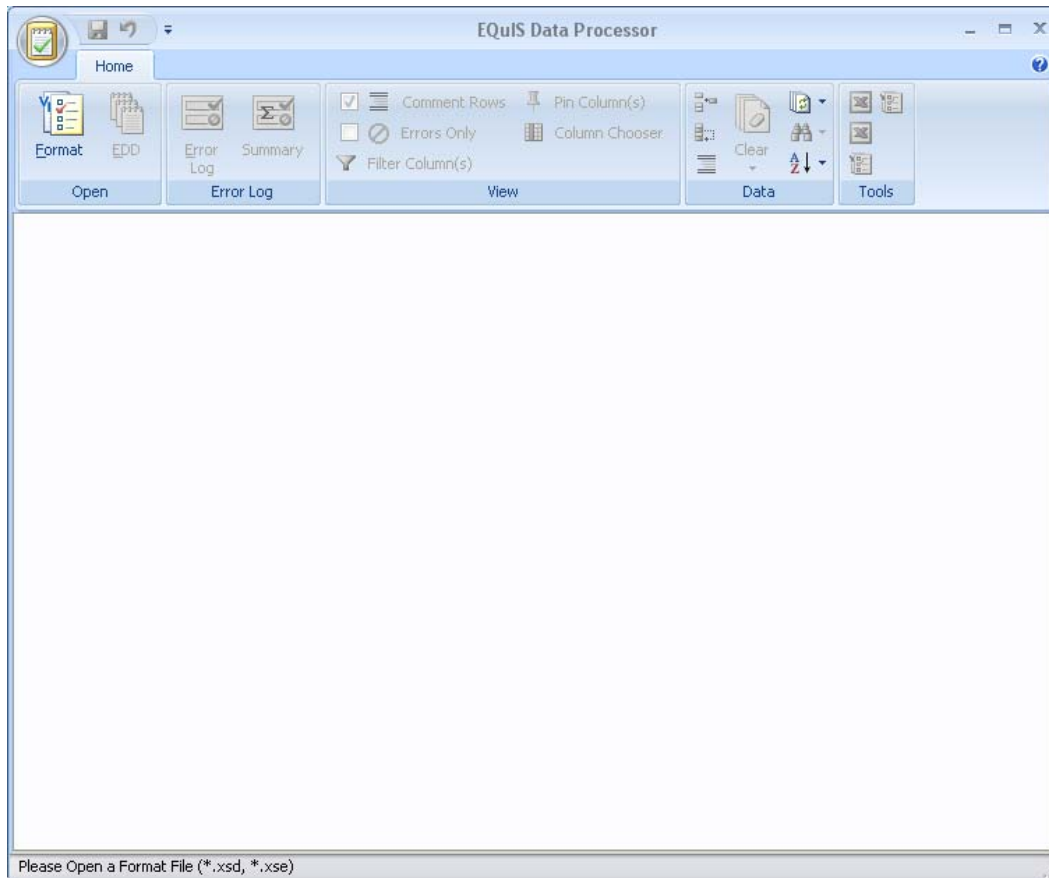



Figure 1 – EQuIS Data Processor

1.2 Format File

The **Format File** contains the definitions that belong to the DNREC EDD format for each individual section. The DNREC format ensures that the data provided to DNREC meets DNREC’s data quality standards. The DNREC EDD format file consists of the following files that reside in the same folder (typically the \EarthSoft\EQuIS\Formats\DNREC\ folder):

- DNREC.xse
- DNREC-enum.xsd
- DNREC.vbe

To open a Format File:

1. Click **Format File**  located in the upper left hand corner of EDP, and browse to the **DNREC.xse** file (\EarthSoft\EQuIS\ Formats\DNREC\ folder).
2. Select the DNREC.xse format file and click **Open**.

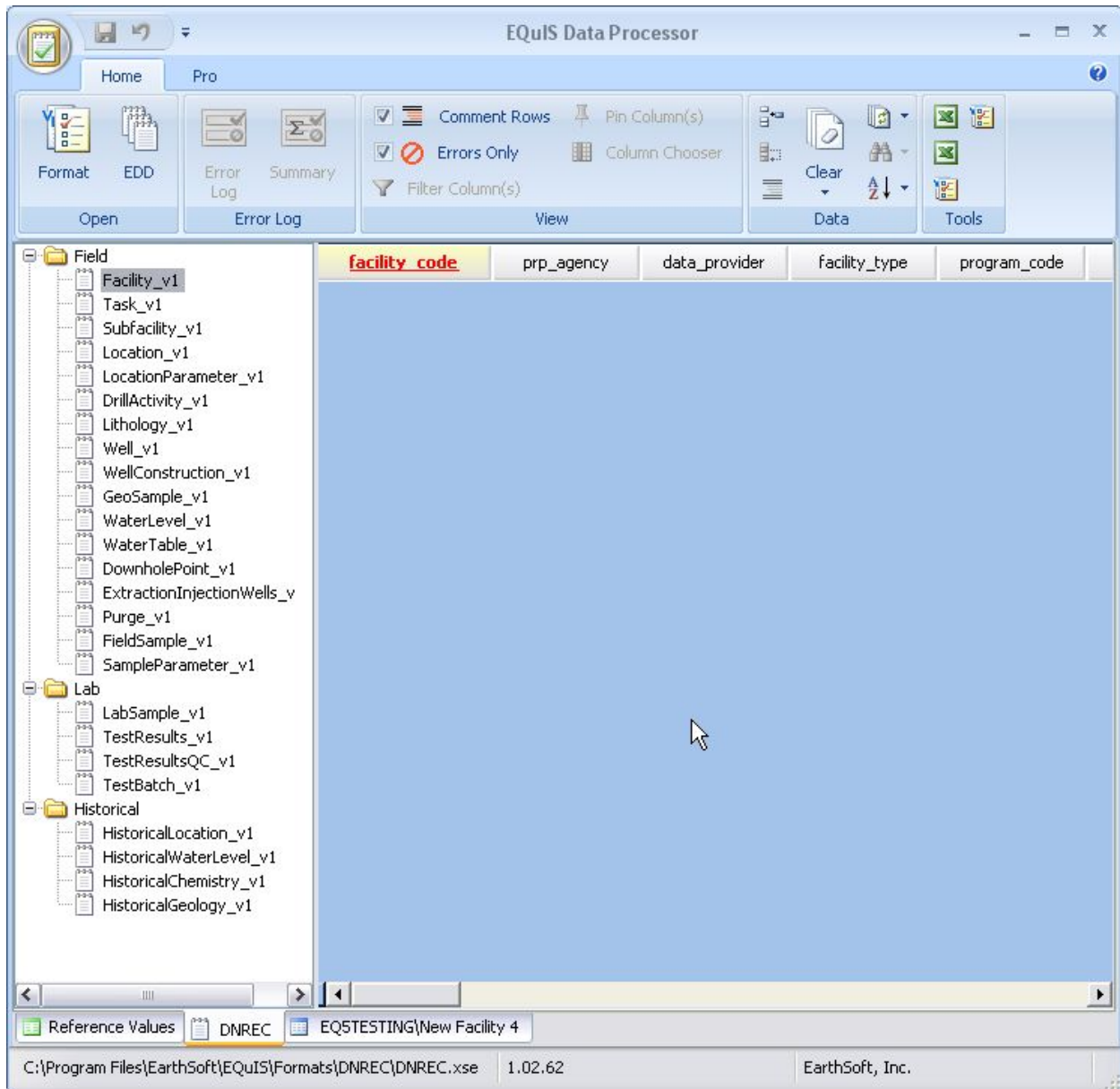


Figure 2 – DNREC.xse Format File

Note: If you use the same **Format File** repeatedly, there is a preference you can set in the **Options** menu so that you always open that particular **Format File** when EDP starts. For more information, see 1.14 EDP Options in this manual.

1.3 Understanding the DNREC Import Format

The DNREC format includes tools to assist users in understanding the format's requirements. These tools include: tool tips, color-coded column headers, and drop-down boxes that enforce reference values.

To view the DNREC Import Format, click the middle tab at the bottom left of the main page.

In the DNREC Format, tool tips are displayed in the column headers. To access these tool tips, hold your mouse over the column headers' name. Use these tool tips to help you determine what values to use when populating each field, as well as to determine which reference value table is used to populate the field.

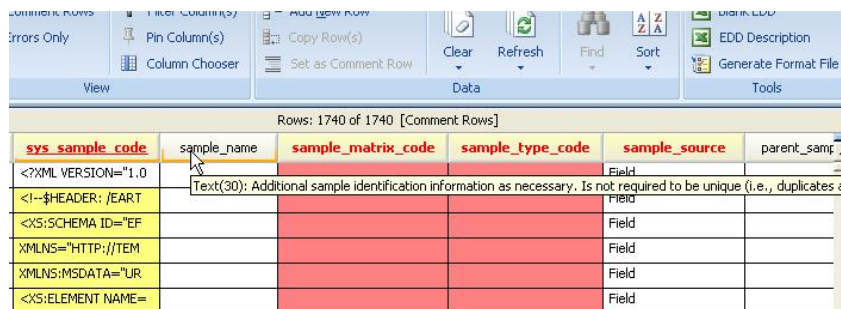


Figure 3 - EDP Tooltip Example

Another aid to help you understand a format, is the color coding that is used to indicate what each column represents. The defaults are:

- **(Underlined) Red** – Column headers that are underlined and use red font (default color) are **required primary key** fields. The primary key fields are required and must be unique.
- **Red** – Column headers with red font (default color) are **required** fields.
- **Blue** – Column headers with a blue font (default color), are **reference value** fields.

Note: Required fields may also be reference value fields; when this occurs, the column header is displayed in a red font.

You can modify the default font colors in the **Options Tool** in EDP (see 1.14 EDP Options in this manual).

Finally, the DNREC format provides you with drop-down lists that appear when you place your cursor over any **Reference Value** or **Enumeration Value** field. These drop-down lists display the approved reference values from the reference table indicated in the tool tip (as shown in Figure 3 - EDP Tooltip Example) or values from the **enum file**. If an error indicating a **'Missing Reference Value'** is displayed, you can select one of the values from the drop-down lists to resolve the error.

1.4 Reference Value File

The **Reference Value** file specifies what reference values, or valid values, are allowed in the data file. This file is a lookup list that displays the values that you can choose from for a particular field.

Note: This file is extremely important for data management and the communication of data across many organizations.

To review the **Reference Values**, click on the first tab at the bottom left corner of the main page.

You do not have to manually open the reference value file to ensure that clients load the appropriate set of reference values for each format. Instead, each time you open a format file, EDP goes to the same folder as the format file and looks for a file with the extension .RVF (the reference value file). If a file is present, that file is opened and used. If there is no *.RVF file, then you are prompted to browse to the appropriate *.RVF file.

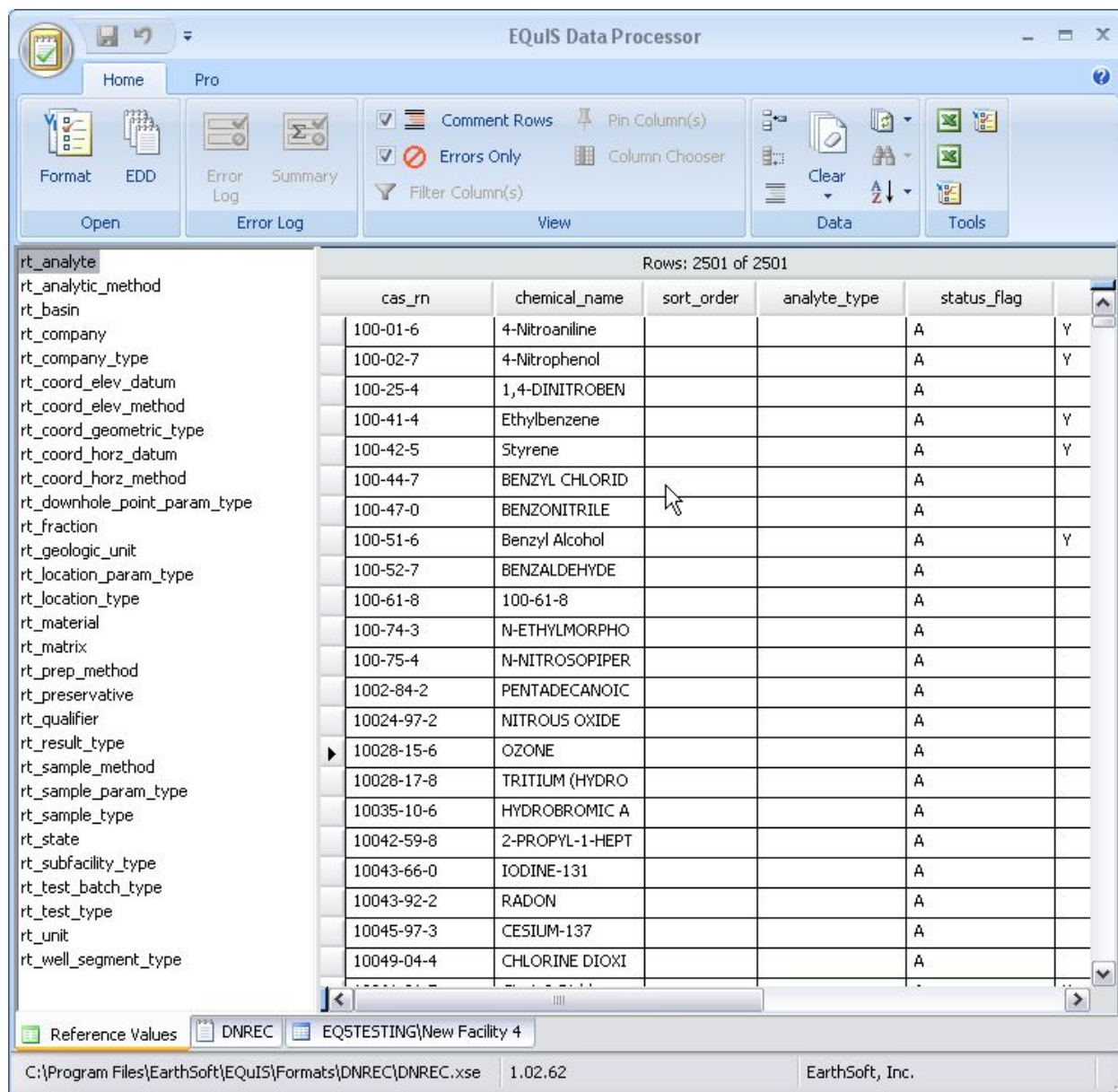
When EDP loads the format file, the **Reference Values** tab, located in the lower left corner of EDP, shows the various tables and valid values defined within each table. Note that values displayed on this tab are *not editable*, and if values need to be added, data submitters need to notify DNREC and request that a new .rvf file be created.

To facilitate finding the appropriate reference value, click the column headers to sort the **Reference Values**.

After you open the format file, verify that the most up-to-date DNREC reference values file is loaded into EDP.

To verify that the correct reference value file was loaded, select the **Reference Values** tab. The path to the reference value file as well as the date it was created is displayed in the lower left corner.

This manual uses the default folder for the DNREC EDD files located in the \\EarthSoft\EQuIS\Formats\DNREC\ folder.



The screenshot shows the EQiS Data Processor software interface. The main window displays a list of 2501 rows of data. The columns are: cas_rn, chemical_name, sort_order, analyte_type, and status_flag. The data includes various chemical names such as 4-Nitroaniline, 4-Nitrophenol, 1,4-DINITROBEN, Ethylbenzene, Styrene, BENZYL CHLORID, BENZONITRILE, Benzyl Alcohol, BENZALDEHYDE, 100-61-8, N-ETHYLMORPHO, N-NITROSOPIPER, PENTADECANOIC, NITROUS OXIDE, OZONE, TRITIUM (HYDRO, HYDROBROMIC A, 2-PROPYL-1-HEPT, IODINE-131, RADON, CESIUM-137, and CHLORINE DIOXI.

cas_rn	chemical_name	sort_order	analyte_type	status_flag
100-01-6	4-Nitroaniline			A
100-02-7	4-Nitrophenol			A
100-25-4	1,4-DINITROBEN			A
100-41-4	Ethylbenzene			A
100-42-5	Styrene			A
100-44-7	BENZYL CHLORID			A
100-47-0	BENZONITRILE			A
100-51-6	Benzyl Alcohol			A
100-52-7	BENZALDEHYDE			A
100-61-8	100-61-8			A
100-74-3	N-ETHYLMORPHO			A
100-75-4	N-NITROSOPIPER			A
1002-84-2	PENTADECANOIC			A
10024-97-2	NITROUS OXIDE			A
10028-15-6	OZONE			A
10028-17-8	TRITIUM (HYDRO			A
10035-10-6	HYDROBROMIC A			A
10042-59-8	2-PROPYL-1-HEPT			A
10043-66-0	IODINE-131			A
10043-92-2	RADON			A
10045-97-3	CESIUM-137			A
10049-04-4	CHLORINE DIOXI			A

Figure 4 – Reference Value File

1.5 Data Files


The data file contains the data that needs to be checked. The data file can be an individual Tab delimited (.txt) or Comma Separated (.csv) file. Alternatively, you can also load an Excel workbook (.xls), Access database (.mdb), or Zip (.zip) file containing several .txt or .csv files. To use these alternate data file formats (.xls, .mdb, .zip, etc.) with EDP, you must include the EDD format name in the data file name. Including the format name allows the data file to be checked against the correct format. Additional naming convention details are provided in the sections that follow.

1.6 Loading Individual CSV or TXT Data Files

You can load data files into the EDP by selecting the **EDD** button from the main EDP toolbar and browsing to the EDD file you want to load into EDP. When multiple sections exist within a given **Format File**, and you are loading individual data files in the .csv or .txt formats, the data file selected is loaded into the currently selected section.

For example, in the DNREC **Import Format** tab, the section **Facility_v1** is the first section listed. This section is highlighted by default upon opening the format. Any data file opened without changing the format section selection will be loaded into this section, even if this data should be loaded into a different section.

To load a data file:

1. Select the **LabSample_v1** section of the **Lab** format.
2. Click **EDD**  from the EDP Toolbar.
3. Browse to the folder where data resides and select the appropriate file.
4. Click **Open**.

The data check occurs in the background while the data file is being loaded. When using EDP, EarthSoft recommends loading data into the DNREC format's sections in the order listed to ensure that all data checks are completed properly. The sections are listed in parent-child order, such that if the last section is loaded before the first, errors may appear which require you to refresh the data to enable EDP to relate the sections.

1.7 Loading Excel, Access, or ZIP Data Files

You can also load Excel, Access, and Zipped data files into EDP by selecting the **EDD** button from the main EDP toolbar and browsing to the EDD file you would like to load into EDP. Unlike .csv or .txt files, when loading Excel, Access, and Zipped data files, the section of the format currently selected is not relevant. Instead, EDP uses the *name* of the individual Excel worksheet, Access table, or file included in the zip file, to open the file into the appropriate section(s) of the format. This functionality allows users to open data into more than one section of an EDD at one time, thus streamlining the data checking process.

The following sections outline the specifications that you need to follow when loading data from an Excel, Access, or ZIP file.

1.7.1 Excel Files

When using Excel files with EDP, the Excel file (.xls) can use any nomenclature, however you must name the data sheets according to the EDD section names. For example, the Excel spreadsheet shown in Figure 5 - Excel Spreadsheet, contains data in the DNREC format. The data sheet below shows how the worksheet names must correspond to the section names of the DNREC format. In this example, the first tab is named Facility_v1 to

match the first section of the DNREC format. Note the additional sheets are also named for their respective EDD sections (Task_v1, Subfacility_v1, Location_v1, etc.).

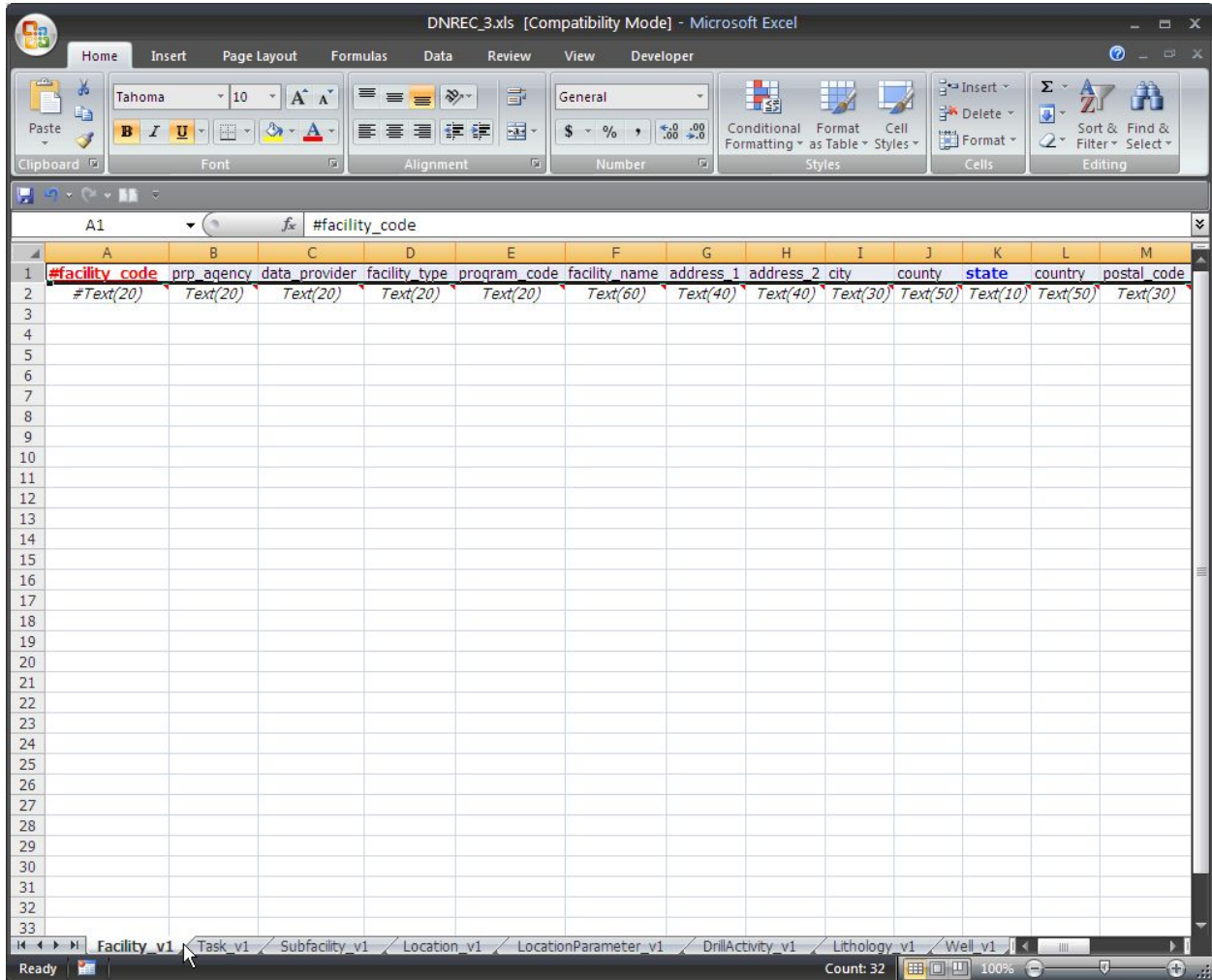


Figure 5 - Excel Spreadsheet

1.7.2 Access Database Files

Similarly to Excel files, you can use Access database file (.mdb) to load data into EDP. The individual Access database file can use any nomenclature, but you must name the data tables according to the EDD format name. For example, the Access database shown in Figure 6 – Access Database contains tables which use the DNREC Lab Data section's names as the table names.

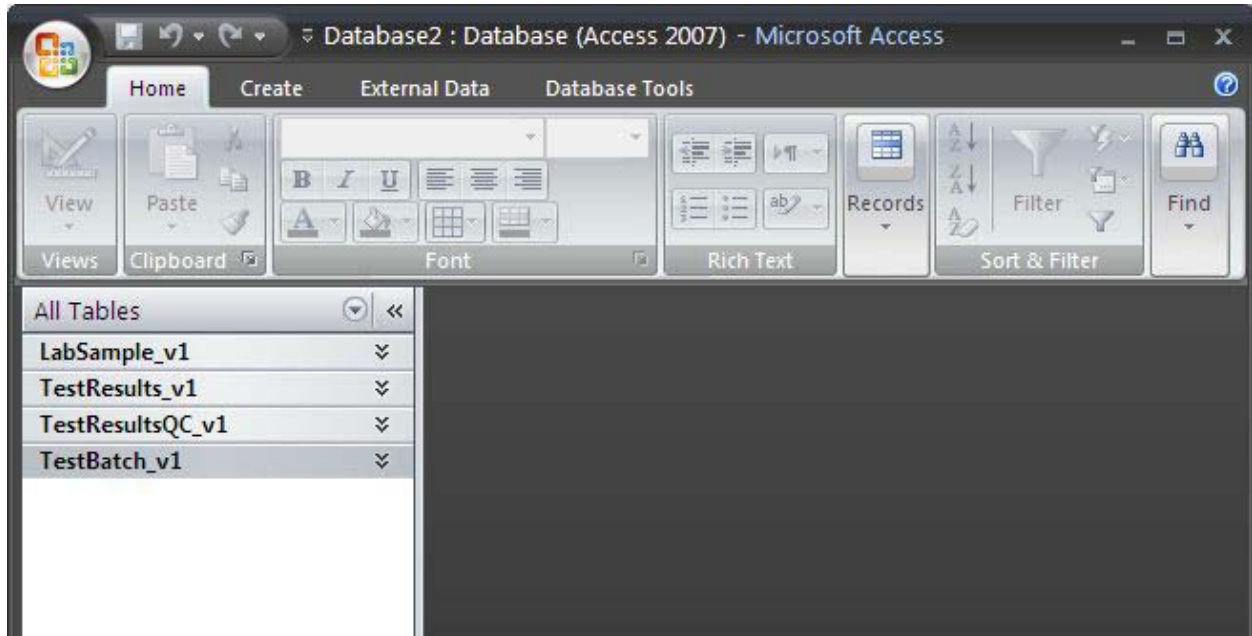


Figure 6 – Access Database

1.7.3 ZIP Data Files

When using Zipped data files with EDP, you must name the data files (in the *.zip file) with the EDD format name. For example, the zipped file shown in Figure 7 – Loading Zip Files contains data files which use the DNREC Lab Data section's names as the text file names.

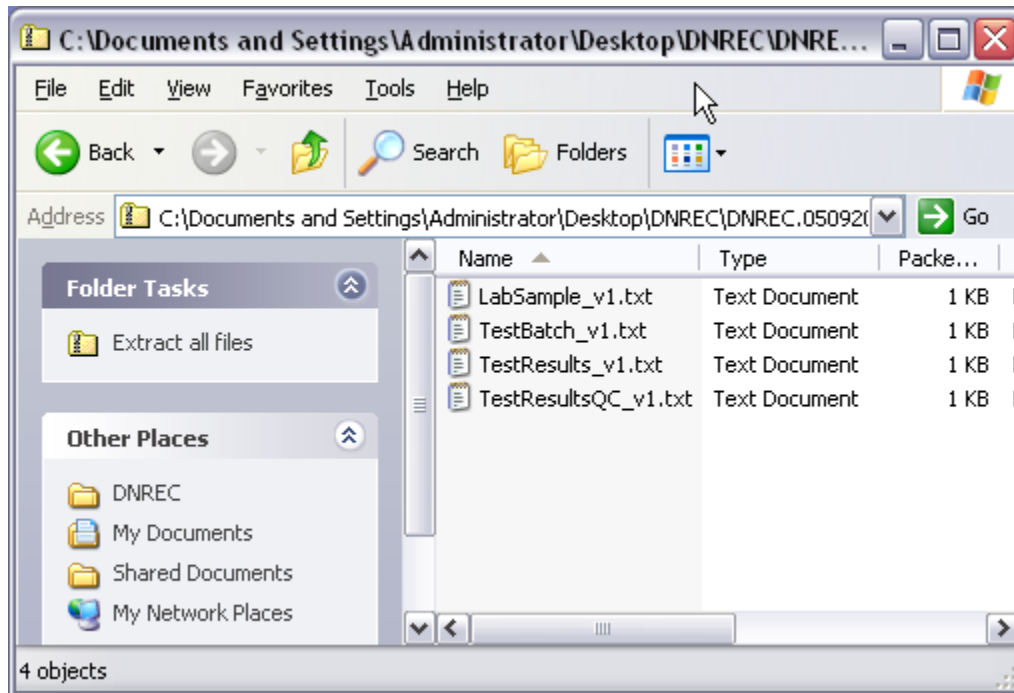


Figure 7 – Loading Zip Files

1.7.4 Load and Append Multiple Data Files

When you load data files, it is also possible to load multiple data files of the same format during the same “EDP session” (as shown in Figure 8 – Load and Append Data File). For example, if after loading the **Field Sample** data you receive another Sample_v1 data file which should be included with the current data set, you can use the **Load and Append Data File** utility to simultaneously check both files in EDP.

Note: This functionality is only enabled once a data file is loaded.

To load and append data:

1. Right-click on a section name to bring up a list that includes the **Load and Append Data File** utility.
2. Browse to the data folder and select the next data file to load.

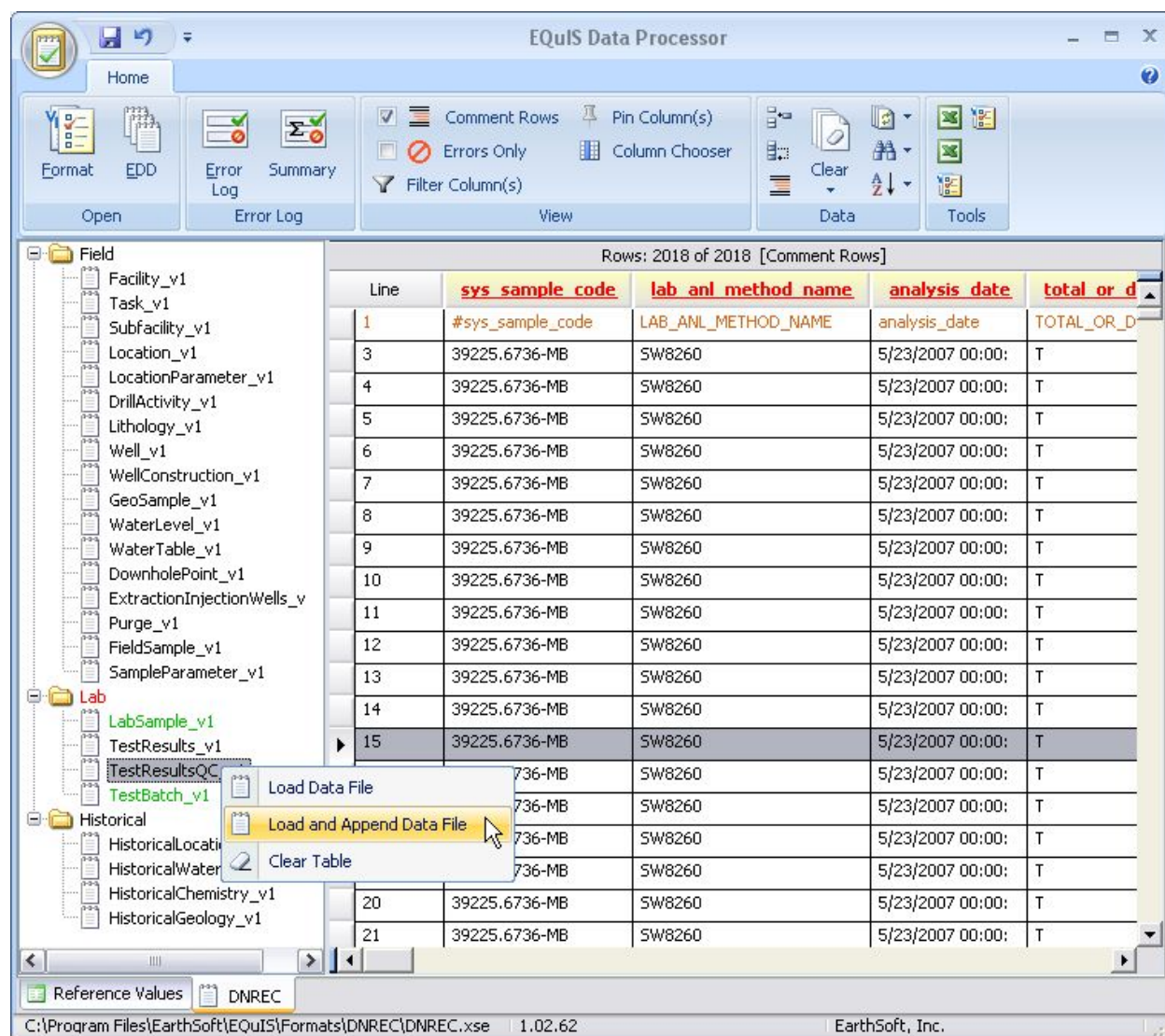


Figure 8 – Load and Append Data File

Using the **Load and Append** utility appends the new data file to the end of the original data file.

1.8 Clear Tables

If you accidentally load a data file using the wrong section, you can clear all data rows by right-clicking on the format name and selecting **Clear Table** or clicking **Clear** from the **Data** section of the EDP toolbar and selecting **Clear EDD**.

1.9 Finding Errors

Once you load an EDD into EDP, you can identify which section(s) contain at least one error because the sections' name listed on the left side of the screen will be highlighted in red.

Note that in Figure 9 – Displaying Errors, the **Lab** section is in red. Also, any data that do not satisfy the checking requirements identified in the **Format File** are highlighted in the data grid. Finally, the line numbers along the left side of the data grid are red in lines where an error exists.

To obtain additional details about what each error means, hold your cursor over the highlighted cell in the data grid.

Note: You can change the colors assigned to each error. For additional information on changing these color codes, see 1.14 EDP Options section in this manual.

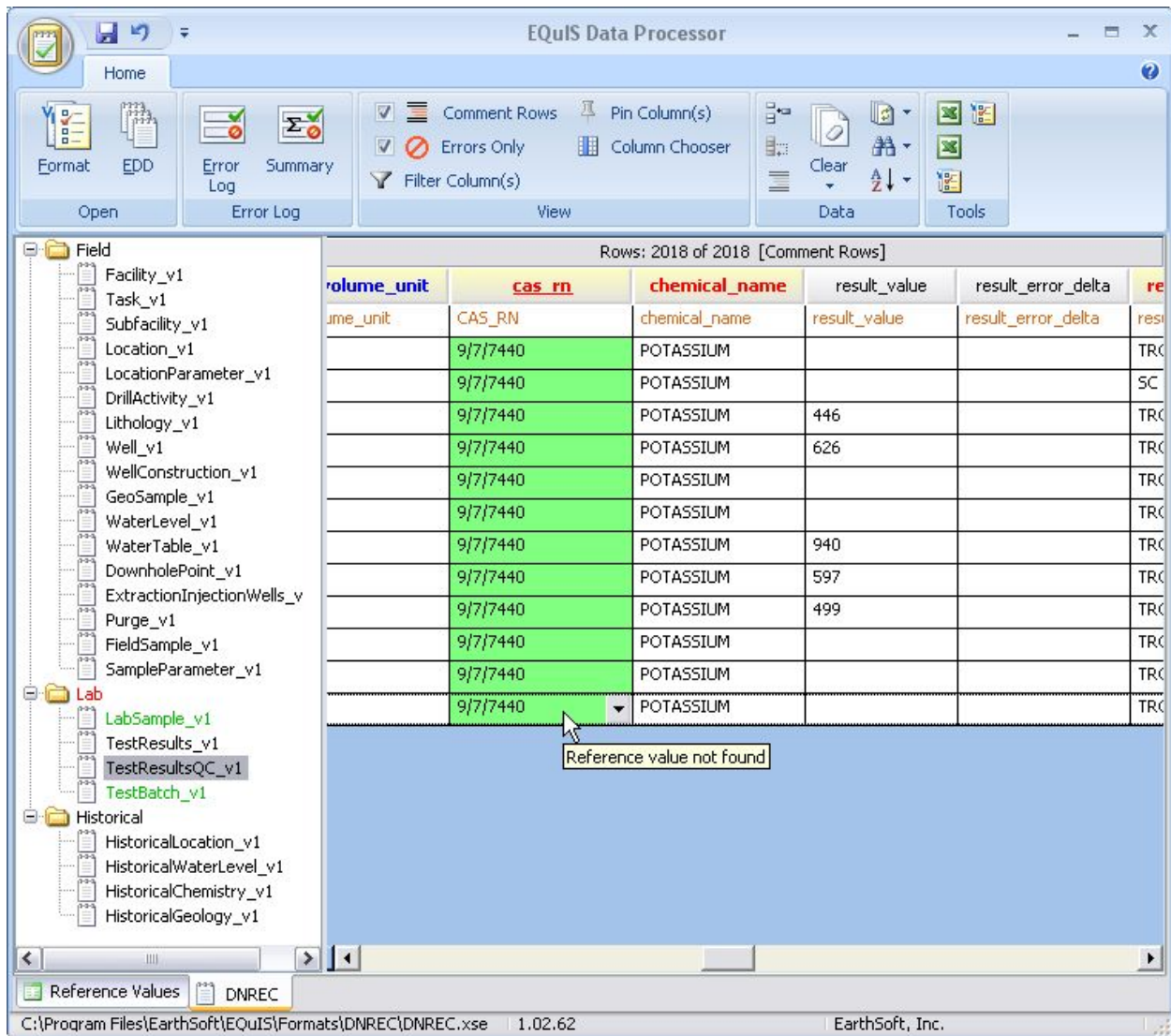


Figure 9 – Displaying Errors

1.10 Saving Modifications

Before you begin, it is important to understand that edits made to data files within EDP are **not** automatically saved. If you are modifying data files using the edit feature available in EDP, then you should save your data files frequently during an EDP Session.

To save modifications made to data in EDP:



1. Select the **Application** menu located in the upper left corner of EDP.
2. Select **Save > EDD**.
3. Select the File.
4. Select **Save**.
5. Select **Yes** to replace existing file.

1.11 Error Resolution

As discussed in section 1.9 Finding Errors, errors that are found in EDP are highlighted with a specified color to signify the type of error. You can correct these data errors within EDP if the **Edit** feature was enabled during installation of EDP.

- To make modifications to any cell within the data file after it is loaded, highlight the cell and type the modification.
- To correct **Reference Value** errors, use the drop-down lists associated with the field where the error occurred.

See 1.4 Reference Value File in this manual for additional information.

1.11.1 Column Headers Errors

You may have noticed that many of the errors in the data files loaded in these exercises are attributed to the two header rows that are contained in each of the data files. These two rows contain column header information, useful to ensure that the correct data populates each column, and to check that the appropriate data file was loaded into the correct section of the format.

You can instruct EDP to ignore any header rows in the data grid.

To ignore header rows:

1. Select the EFW2FSample format.
2. Highlight the header rows
3. Right-click and select **Set as Comment Row**.

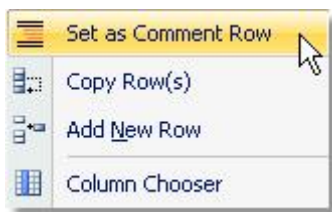


Figure 10 - Set as Comment Row

The text in the first two rows becomes italicized and includes a pound sign (#) directly before the first letter of the column header. The colors of the cells in these rows also changes indicating that they are no longer problematic.

Repeat this step for the remaining loaded files.

You can skip the step of setting a row as a column header if the data file's comment rows include this pound sign (#) prior to being loaded into EDP.

Line	sys_sample_code	sample_name	sample_matrix_co	sample_type
1	#sys_sample_code	sample_name	sample_matrix_code	sample_type
2	#Text[40]	Text[30]	Text[10]	Text[20]

Figure 11 - Example of Pound Sign (#) Comment Indicator

You can choose to not view the **Comment Rows** using the option on the **View** group. Note the check next to the word "Comment Rows". If you click the **Comment Row** option, the check is removed and the **Comment Rows** will not be displayed. To view the comment rows again, toggle the switch.

1.11.2 Reporting Reference Value Errors and Error Logs


As outlined in the Reference values section of this manual (1.4 Reference Value File), using consistent nomenclature is extremely important for data management. Therefore, reference value and enumeration files are part of DNREC's format file to make resolving reference value errors as easy as possible. New reference values may need to be added to the database. However, these new reference values will need to be reported to the DNREC data management team.

Using Standalone EDP, you can easily provide a summary list of missing reference values for the database manager to resolve. This can be useful for data providers to do when requesting that new reference values be added to the RVF file, enabling them to successfully submit data. Once all data issues are resolved, other than the new reference values, you can create an error summary outlining the new reference values.

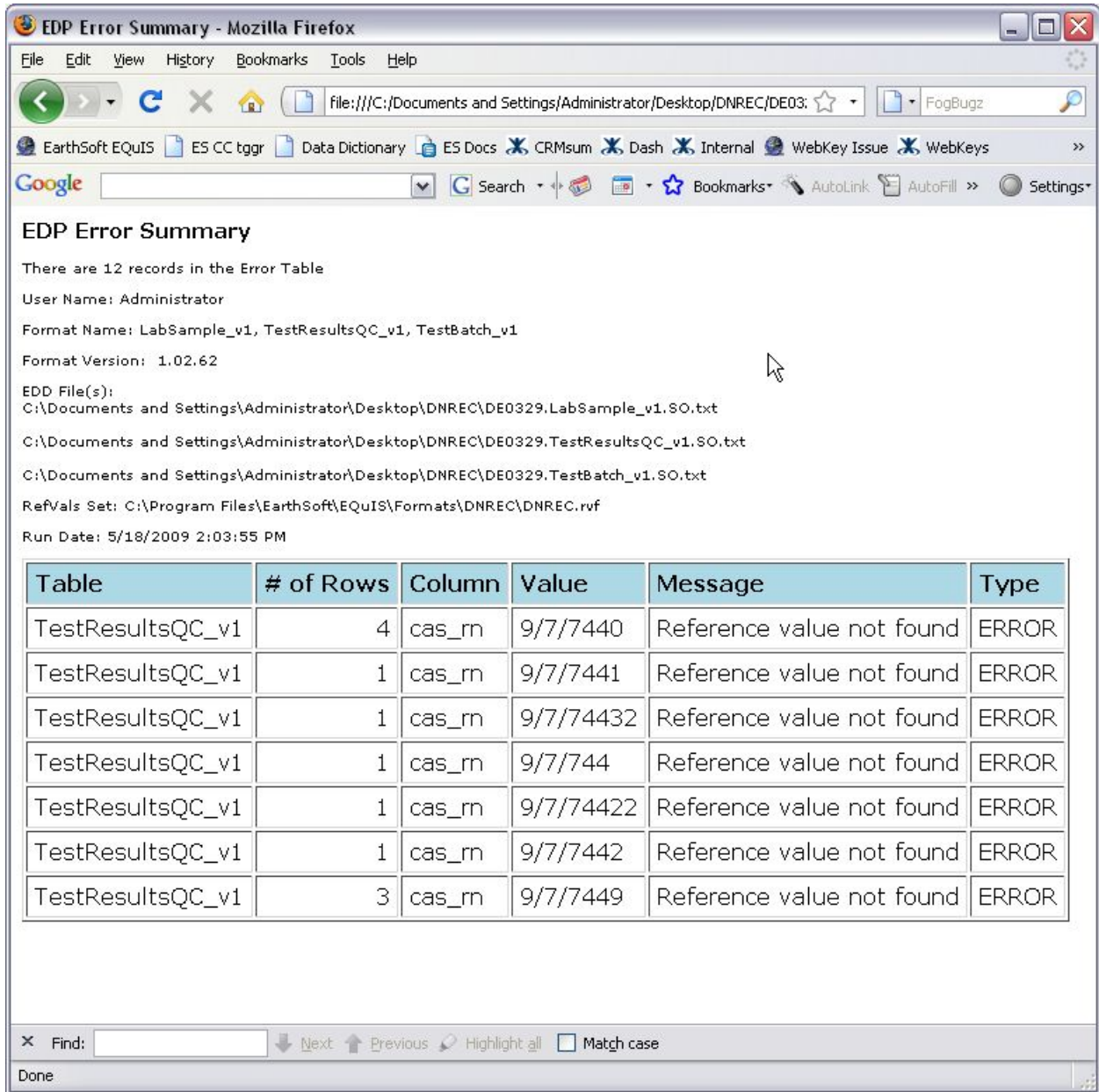
To create an Error Summary:

1. Resolve all data issues other than the reference value errors.

Note: You cannot modify the list of reference values from Standalone EDP.

2. Click **Error Log (Summary)**  in the **Error Log** group.
3. If desired, in the **Save Error Log As** dialog box, change the default name or location of the error log file.

After you save the .htm file, it opens the .html file in your default internet browser.



EDP Error Summary

There are 12 records in the Error Table

User Name: Administrator

Format Name: LabSample_v1, TestResultsQC_v1, TestBatch_v1

Format Version: 1.02.62

EDD File(s):
 C:\Documents and Settings\Administrator\Desktop\DNREC\DE0329.LabSample_v1.SO.txt
 C:\Documents and Settings\Administrator\Desktop\DNREC\DE0329.TestResultsQC_v1.SO.txt
 C:\Documents and Settings\Administrator\Desktop\DNREC\DE0329.TestBatch_v1.SO.txt

RefVals Set: C:\Program Files\EarthSoft\EQUIS\Formats\DNREC\DNREC.rvf

Run Date: 5/18/2009 2:03:55 PM

Table	# of Rows	Column	Value	Message	Type
TestResultsQC_v1	4	cas_rm	9/7/7440	Reference value not found	ERROR
TestResultsQC_v1	1	cas_rm	9/7/7441	Reference value not found	ERROR
TestResultsQC_v1	1	cas_rm	9/7/74432	Reference value not found	ERROR
TestResultsQC_v1	1	cas_rm	9/7/744	Reference value not found	ERROR
TestResultsQC_v1	1	cas_rm	9/7/74422	Reference value not found	ERROR
TestResultsQC_v1	1	cas_rm	9/7/7442	Reference value not found	ERROR
TestResultsQC_v1	3	cas_rm	9/7/7449	Reference value not found	ERROR

Find: Next Previous Highlight all Match case

Done

Figure 12 – EDP Error Summary Report

Providing the error log to the database manager displays a summary of any missing reference values that may need to be added to the database.

1.12 Tools

There are several options available in EDP toolbar to assist with data review.

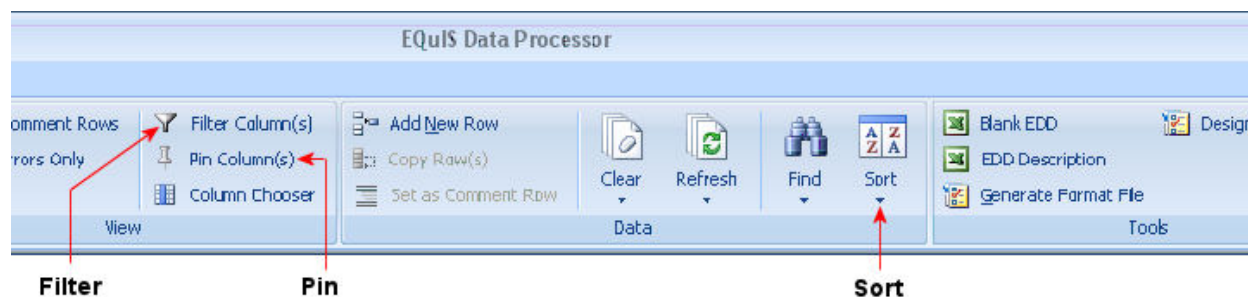




Figure 13 – Filter, Pin, and Sort Tools

Use **Sort**  to sort the data in ascending or descending order. For example, if you want to sort the result data based on chemical name, you can highlight the chemical name column and click **Sort**.

Use **Filter**  to narrow the data set for viewing subsets of the data. Click **Filter**, and each column is enabled with a pull-down menu which allows you to filter on the available values for that field. Do this by clicking on the filter icon, then chose the selection you wish to use. Click the Filter again and select All to remove the filter. You can also select **Custom** to build your own selection criteria.

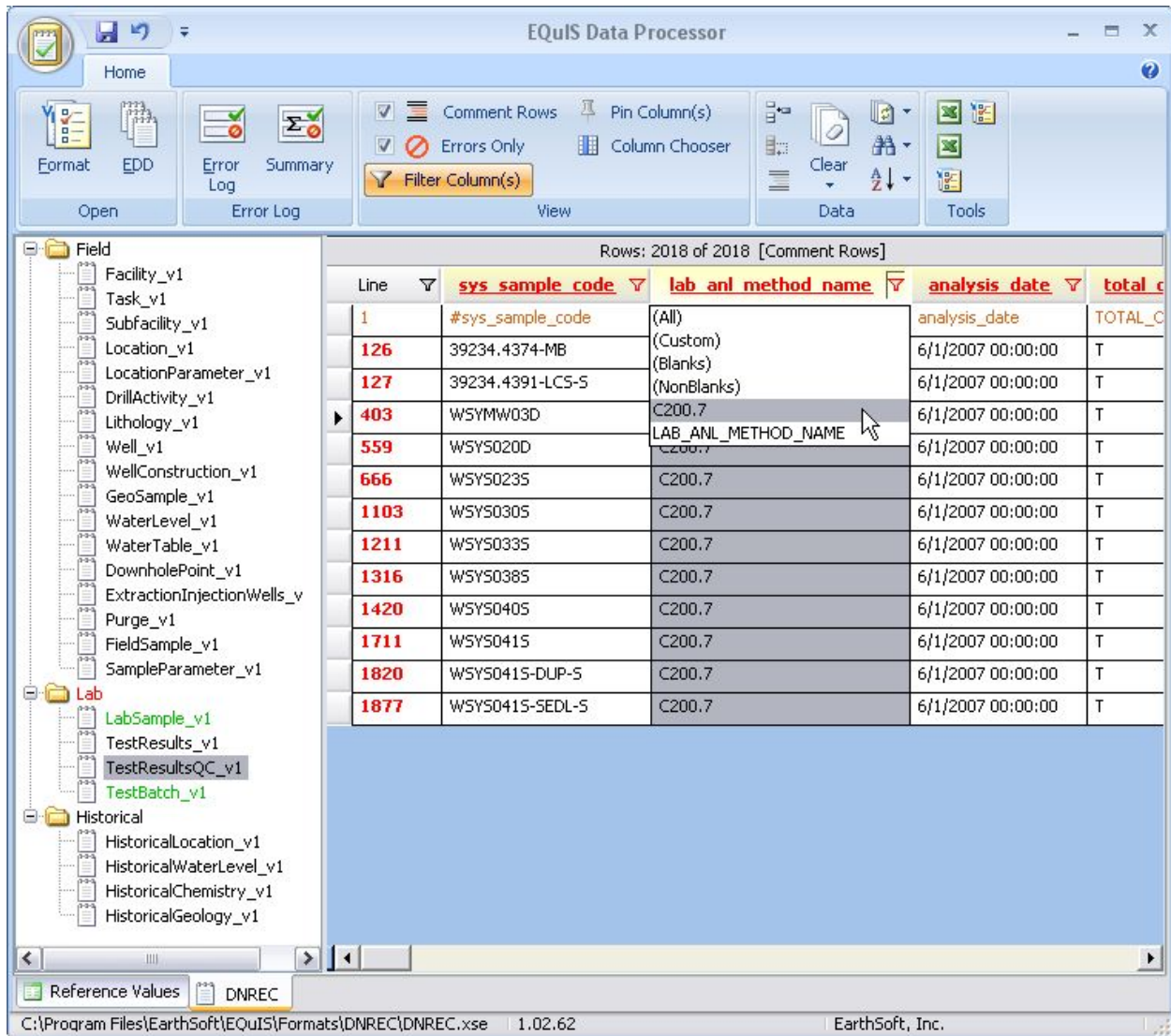



Figure 14 – Using the Filter Tool

Use **Pin**  to assist with viewing large data sets. Click **Pin** and a "pin" is enabled adjacent to each field. When selected, the pin holds that column in place while viewing the related fields for each record. As you scroll to the right, the "pinned" column remains stationary so you can review the data associated with each sample.

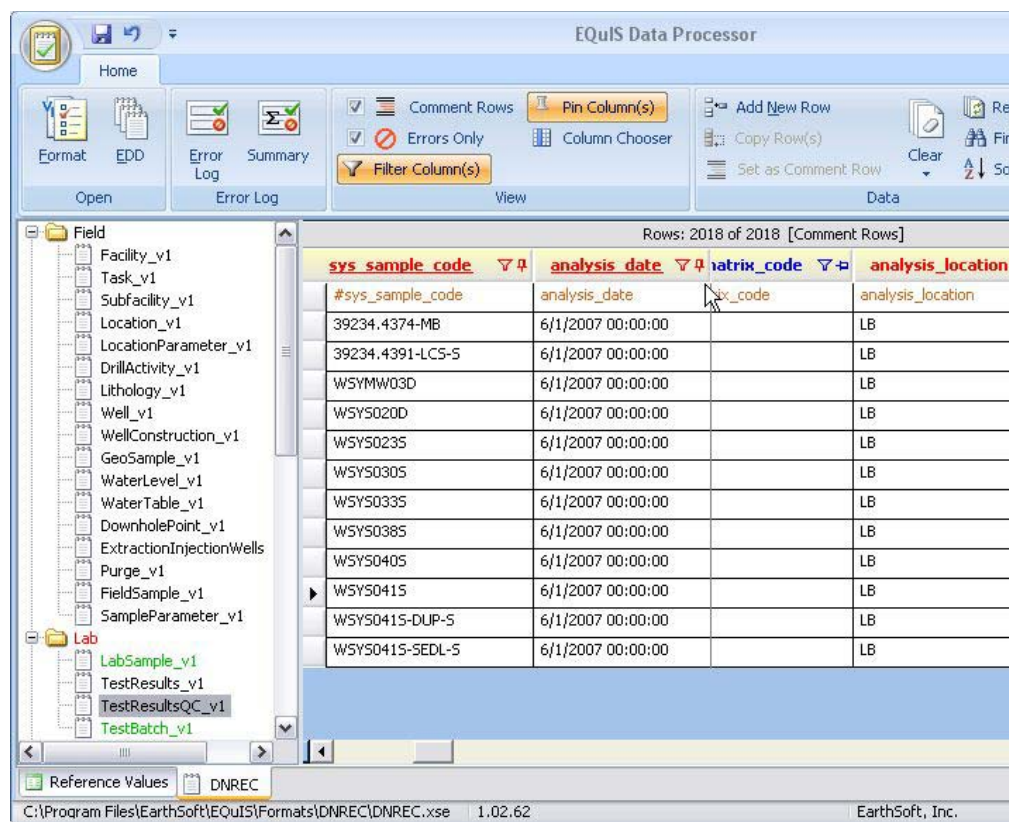


Figure 15 – Using the Pin Tool

1.13 Sign and Submit

After using the tools outlined above to resolve all of the issues in a set of Data Files, the data is ready to be submitted for loading into DNREC’s EQUS 5 database. The **Sign and Submit** tool was designed to facilitate submittal of data to EQUS Enterprise EDP. To use the **Sign and Submit** feature requires a user name and password that you can request on DNREC’s Enterprise site at <https://apps.dnrec.state.de.us/equis/>.

Please click the ‘Register New User’ hyperlink and enter the information. After submitting a new user request, users must wait for Enterprise Administrator approval before they can sign into the application.

After the user account has been activated, follow the instructions below to to use the **Sign and Submit** feature:

1. Select **Sign and Submit** from the **Application**  menu in the upper left corner.



Figure 16 - Sign and Submit Window

2. Enter your **User Name** and **Password**.
This information is used to create a user certificate file that EQuIS Enterprise uses to ensure that a valid user is submitting data to the appropriate facility.
3. Select **Save or Submit**.
 - Selecting **Save** prompts you to save the file to a directory on your network.
 - Selecting **Submit** submits the data to DNREC's EQuIS Database.

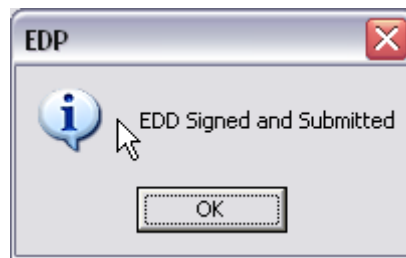


Figure 17 - Sign and Submit Verification Window

4. Select **OK**.

1.14 EDP Options


You can modify:

- The default functionality of EDP so that EDP will not recall the last Format File used
- Use a different comment indicator code
- Increase the number of errors that EDP checks for when Data Files are initially loaded into EDP.
- Change the EDP Application Appearance

1.14.1 Remember Last Format File

By default, EDP was designed to always recall the last format file the application used. If you do not plan to use the same Format File repeatedly, there is a preference you can set in the **Options** menu so the last Format File used is not always opened when EDP starts.

To set the Remember Last Format File option:

1. Select the **Application**  menu, then click **Options** on the bottom right, beside the **Exit**.
2. On the **Options** dialog, click **General**.

The option **RecallLastUsedFormatFile** has a default setting of **Yes**. This indicates that the last format file used in EDP will be opened during the following EDP session.

You can change the setting to **No** to have EDP open to a blank screen each time EDP is launched.

Note: If **RecallLastUsedFormatFile** is set to **Yes** and the specified Format File is not available or cannot be found, a warning message is displayed and you have the opportunity to select a different Format File

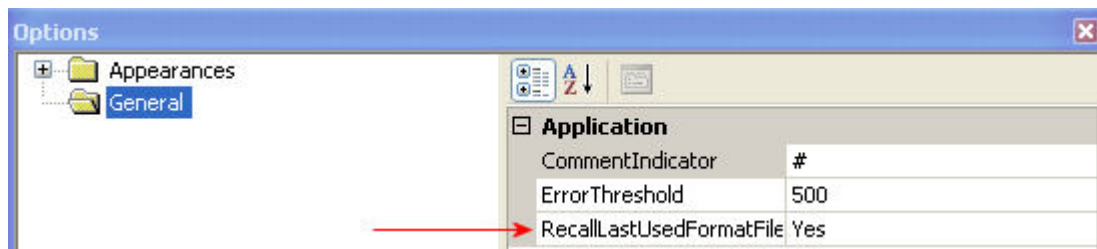


Figure 18 - Option to recall last format opened in EDP.

1.14.2 Modify Comment Indicator Code & Error Threshold

You can also modify the comment row indicator. The default comment row indicator is a pound sign (#). Users who export data from a Microsoft Access database, find that the pound sign is not a good symbol to use. You can also modify the number of errors that the EDP checks when loading the data.

To change the comment row indicator:

1. Follow steps 1 and 2 from "To set the Remember Last Format File option:"
2. Type a new value in the **CommentIndicator** field.

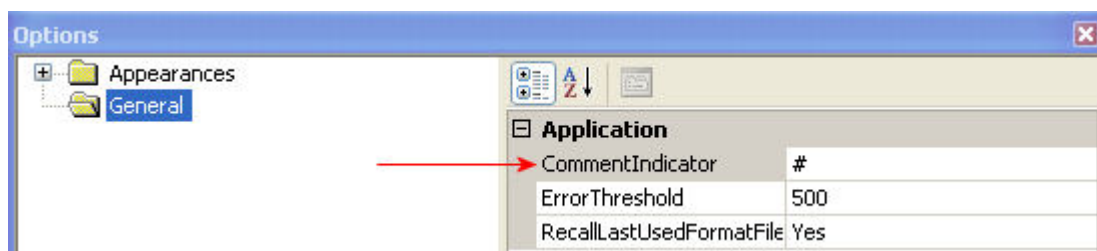



Figure 19 – CommentIndicator Field

1.14.3 EDP Application Appearance

You can modify the default appearance of Column Headers, Errors, and Informational Messages to facilitate usability.

To modify any of these items' appearance:

1. Select the **Application**  menu and then click **Options**.
2. Expand the node (+) to the left of **Appearances**.
3. Select an attribute you would like to modify. For example, '**Reference Value Cell Errors**' under **Errors** > **Cell Errors** in the **Appearance** tree located in the left of the **Options** screen.
4. Click in the cell that appears in the right hand portion of the **Options** window next to '**BackColor**'.
5. Click the Down Arrow to display the colors selection options for cell background colors.
6. Select a color.
7. Click **OK**.

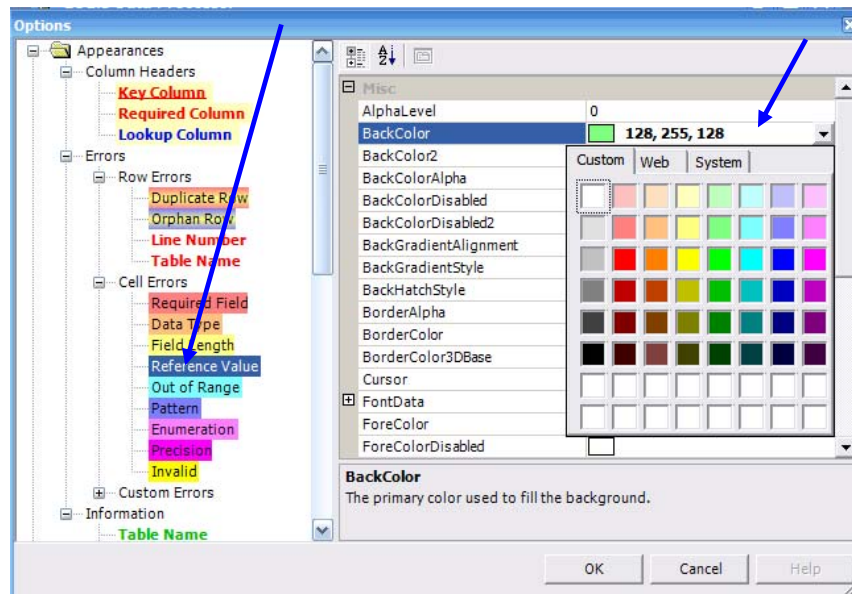


Figure 20 - Example of Modifying Appearance in the Options Window.

You can repeat this process for any **Appearance** settings that are applied to cells in EDP.