

Guide to the NYSDEC Automated Data Loading System

August 7, 2024

The New York State Department of Environmental Conservation (NYSDEC) has implemented a new Automated Data Loading System (ADL) to quickly review and upload EQUIS Electronic Data Deliverables (EDDs) that data providers submit to NYSDEC. This Guide will provide you with basic instructions for creating and submitting an EDD to the ADL system; describe the process used by ADL to receive, review, and upload or reject EDDs; and discuss the two kinds of error logs you might receive if an EDD that you submit is rejected by ADL.

Creating and Submitting EDDs with ADL

The process of creating EDDs for submittal to ADL is the same as the older process of creating EDDs to submit to NYSDEC for manual review and upload by NYSDEC staff, with two important exceptions. First, the version of the EQUIS Data Processor (EDP) that you use must be 7.23.1 or later. Second, the EDD must be created with the NYSDEC EDD Format version 5 (“Format v5”). You can download Standalone EDP v.7.23.3 and Format v5 from EarthSoft at the “New York State DEC EDP Format” page: <https://earthsoft.com/products/edp/edp-format-for-nysdec/>.

If you use EDP in EQUIS Professional as part of your company’s EQUIS system, please download and follow the [NYSDEC EDD Format Files README Instructions](#) on the EarthSoft NYSDEC download page to properly configure EDP Professional to use Format v5.

Note that ADL will automatically reject an EDD that does not have the NYSDEC_v5_MERGE commit type at the end of the file name. It will also reject any EDD with a facility code that is not present in the NYSDEC EQUIS database.

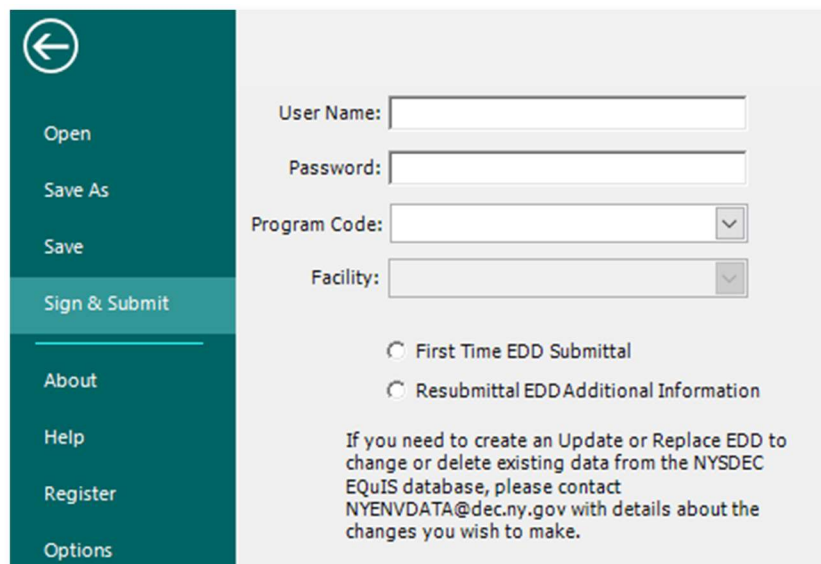
ADL will only accept Merge EDDs. Update and Replace EDDs must still be submitted to NYENVDATA@dec.ny.gov for manual processing.

To prepare and submit a Merge EDD to NYSDEC through our ADL system, follow these steps:

1. Mount Format v5 to EDP version 7.23.1 or later.
2. Load your data into EDP. If the data does not appear in the correct Format sections in EDP, check the .txt file names in the EDD zip file or the section tabs in the Excel file to make sure all section names end in “_v5”.
 - a. If the section names in an EDD zip file end in “_v4”, convert the sections to Format v5:
 - i. Extract the zipped EDD file. A folder will appear with the name of the EDD.
 - ii. Open the folder and change the end of the name of each .txt file from “_v4” to “_v5”.
 - iii. Highlight all of the .txt files in the folder, right click one of them, and select “Send To” > “Compressed (zipped) folder.”

- iv. Copy the original EDD file name, paste it into the file name of the new zipped EDD, and change the end of the new EDD file name from “NYSDEC_MERGE” to “NYSDEC_v5_MERGE”.
- b. If your data is in an Excel workbook, rename the worksheet tabs so each section name ends in “_v5”.
3. Check the EDD for errors in the Standalone EDP software and correct any errors. ADL will not correct errors in EDDs, it will only detect errors and consequently reject the data package.
4. Make certain that any sample location codes or subfacility codes referenced in EDD sections are either (a) listed in the Location_v5 or Subfacility_v5 section, along with any required information; or (b) are already present in the NYSDEC EQUIS database. If you need to know what location or subfacility codes are already in the NYSDEC EQUIS database for your facility, please contact the NYSDEC EQUIS Team at NYENVDATA@dec.ny.gov.
5. Separate Vapor Intrusion (VI) data sections into their own EDD. Attempts to upload VI sections with Field or Chemistry sections can sometimes cause errors.
6. After ensuring that your EDD is free of errors, use the Sign & Submit Tool in EDP with the user ID and password that you received from NYSDEC to create a new Merge EDD with your EQUIS User Certificate. Your User Certificate is a .usr file in the EDD that contains an encrypted version of your EQUIS password. When your EDD is received by the ADL system, your User Certificate will provide the credentials to show that you are authorized to upload data to the NYSDEC EQUIS database.

The Sign & Submit Tool in EDP v.7.23.3 with NYSDEC Format v5 should look like this:



Note that there are only two radio buttons to select the type of EDD instead of the four buttons in previous versions of EDP. The Sign & Submit Tool has been redesigned to only create Merge EDDs with the commit type “NYSDEC_v5_MERGE” at the end of the EDD file name.

If you need to create an Update or Replace EDD to change or delete data that is already in the NYSDEC EQUIS database, contact NYENVDATA@dec.ny.gov with the details.

If the Sign & Submit Tool does not look like the screenshot above, if the tool does not have drop-down menus for the Program Code and Facility, or you get an error message when you try to save your EDD, you have old EDP settings in the Settings.xml file of your EarthSoft Roaming folder that are interfering with the operation of EDP 7.23.3. To fix the problem:

- 1) Close EDP Standalone.
 - 2) Open Windows Explorer and find your EQUIS Roaming Profile with the settings file [C:\Users\\[your Windows username\]\AppData\Roaming\EarthSoft\Settings.xml](file://C:/Users/[your Windows username]/AppData/Roaming/EarthSoft/Settings.xml).
 - 3) Rename Settings.xml to something like SettingsBackup.xml.
 - 4) Restart EDP Standalone, load in your EDD data, and confirm that the Sign & Submit Tool looks and functions properly.
 - 5) Repeat these steps if problems with the Sign & Submit Tool are still there.
7. After saving your new EDD, confirm that it has the correct structure in the file name:
[\[Date Stamp\] \[Time Stamp\].\[Facility Code\].NYSDEC_v5_MERGE](#)

Note that any additional information, such as “Q3 Sampling”, must go at the **beginning** of the file name before the date stamp. The **Facility Code** must exactly match a facility in our database. If your facility is not in the drop-down menu of the Sign and Submit Tool because the facility was recently added to the NYSDEC database, you can choose a different facility code and change it to the correct facility in the EDD file name after you save the EDD. **Note that an incorrect facility code in an EDD submitted to ADL will result in the EDD being rejected, or worse, your data being uploaded for the wrong facility in the database.**

8. Attach your Merge EDD to an email message, add brief descriptions of the EDD to the subject line and text of the email, and send your EDD to NYENVEDD@dec.ny.gov. Please copy the NYSDEC Project Manager for your facility and any other persons who should know that you submitted the EDD. All emails sent to NYENVEDD will be archived for future reference.

After you submit your EDD to ADL, you will receive automated emails from NYENVDATA.dec.ny.gov telling you the status of your data package.

- If the EDD file name has the correct structure, a facility code recognized by ADL, and [NYSDEC_v5_MERGE](#) as the commit type at the end, you will get a message saying that your EDD has been received and is being checked for errors. You should receive another email a few minutes later stating whether your EDD has been uploaded or rejected due to errors. These messages are discussed in detail below.
- If your EDD has a facility code that is not in the NYSDEC database, you will receive a message saying, “the corresponding facility was not found.” Check and correct the facility code in your EDD file name and resubmit the EDD.

- If your EDD file name ends in something other than “.NYSDEC_v5_MERGE”, you will get a message that your EDD “does not match the required package naming conventions.” Correct the commit type in the file name and resubmit the EDD.

Once ADL accepts your EDD, there are three possible messages that you will receive:

- The EDD was found to be free of errors and has been uploaded to the EQuIS database. Congratulations, you have successfully used the Automated Data Loading System!
- ADL determines that there are errors in the EDD that violate the business rules of Format v5. You will receive a message stating that the EDD could not be uploaded because of errors, and the email will have an EDP Error Log attached. This type of error log is described below.
- Your EDD was found to be free of errors when checked against Format v5, but errors occurred when ADL tried to upload your EDD to the EQuIS database. This happens most often because the EDD referred to information in the database that wasn't already there. The message will include a very complex Commit Error Log that can be difficult to decipher, but guidance for interpreting this kind of error log is provided below.

Format v5 Errors Reported in the EDP Error Log

When the ADL system receives an EDD, the system performs two error checks. These checks are done by Enterprise EDP, the version of EDP that uploads EDDs to the EQuIS database. The first is a check for errors that violate the requirements of Format v5. Any errors found at this stage should have been spotted and corrected in EDP standalone or Professional before the EDD was submitted to NYSDEC, but occasionally data providers have submitted EDDs with known errors. If Enterprise EDP finds any Format v5 errors in your EDD, the data package will be rejected and a message will be sent to you saying that your EDD failed to upload due to errors. The message will have an EDP Error Log attached that will list the following information:

1. The **Tables** (sections) of the EDD where the errors occur.
2. The specific **Line** and **Column** locations of the errors.
3. The **Values** that triggered the errors, color coded as they would be in EDP Standalone.
4. The **Error Messages**. These messages describe each error in clear language.
5. The **Type** of message; this will almost always be 'ERROR'.

The EDP Error Log will look something like this:

| Table | Line | Column | Value | Message | Type |
|-----------|------|-----------------------|---------------------|--|-------|
| Sample_v4 | 20 | depth_unit | [NULL] | Missing required field | ERROR |
| Sample_v4 | 20 | depth_unit | [NULL] | If sample_matrix_code='SO', 'SS' or 'WG' and sample_type_code='N' or 'FD' then start_depth, end_depth and depth_unit are required. | ERROR |
| Sample_v4 | 20 | end_depth | ft | Invalid data type | ERROR |
| Sample_v4 | 20 | chain_of_custody | 2021/08/10 00:00:00 | Value exceeds field length | ERROR |
| Sample_v4 | 20 | sample_receipt_date | Pirnak | Invalid data type | ERROR |
| Sample_v4 | 20 | sampling_company_code | [NULL] | Missing required field | ERROR |
| Sample_v4 | 20 | sample_method | SM | Reference value not found | ERROR |
| Sample_v4 | 20 | task_code | [NULL] | Missing required field | ERROR |
| Sample_v4 | 89 | depth_unit | [NULL] | Missing required field | ERROR |
| Sample_v4 | 89 | depth_unit | [NULL] | If sample_matrix_code='SO', 'SS' or 'WG' and sample_type_code='N' or 'FD' then start_depth, end_depth and depth_unit are required. | ERROR |
| Sample_v4 | 89 | end_depth | ft | Invalid data type | ERROR |

If you receive this kind of error log, load the EDD back into EDP and correct the errors, then use the Sign & Submit Tool to save the data set as a new EDD. ADL will not accept EDDs with duplicate names. If you do not understand an error message, contact the NYSDEC EQUIS Team at NYENVDATA@dec.ny.gov.

Interpretation of a Commit Error Log

If there are no errors found when Enterprise EDP checks the EDD with Format v5, the system will attempt to upload, or commit, the EDD to the EQUIS database. At that point, a different kind of error check is carried out by Enterprise EDP in which the codes that are referenced in every line of the EDD are compared to the same kinds of codes that already exist in the database for that facility. You may have previously received messages from the EQUIS Team about EDDs that

could not be uploaded because sample location information was missing from the database; this is the stage where those errors appear.

A Commit Error Log can appear intimidating and can be a challenge to interpret, because unlike an EDP Error Log, a Commit Error Log does not just show the EDD rows with errors and provide clear explanations of those errors, nor does it provide the exact locations of those errors in the EDD. Instead, the Commit Error Log shows the database tables into which Enterprise EDP tried to load data and the results of those attempts. Enterprise EDP processes every row of data for uploading into the corresponding database tables, starting with the first row to appear at the top of the list of Format sections and working down. For example, if an EDD with data for samples and analytical results has data in the Sample_v5, TestResultQC_v5, and Batch_v5 sections, the data from those sections will be processed for uploading in that order. This process continues even if data in preceding rows could not be uploaded. Consequently, if a few rows of data in the Sample section cannot be upload because those rows referenced Location codes (sys_loc_codes) that aren't in the database, those failures will cause a cascade of errors as the analytical test data for those samples also cannot be uploaded, followed by the analytical result data for the failed tests, and so on.

The resulting Commit Error Log can look like this, with cascading error messages continuing for dozens or hundreds of rows:

| Status | Table | Row | Message |
|--------|----------------------|--------------------------------|--|
| INFO | dt_sdg | | 1 row(s) processed |
| INFO | dt_test_batch | | 17 row(s) processed |
| INFO | dt_task | | 1 row(s) processed |
| INFO | dt_sample | | 57 row(s) processed |
| ERROR | dt_sample | 161 20354928 IW-7-20220630 ... | ('161', 'IW-7') cannot be inserted into child table dt_sample (facility_id, sys_loc_code) because a corresponding parent row does not exist in table dt_location (facility_id, sys_loc_code) [FK_dt_sample__65370702]. |
| ERROR | dt_sample | 161 20354929 IW-9-20220629 ... | ('161', 'IW-9') cannot be inserted into child table dt_sample (facility_id, sys_loc_code) because a corresponding parent row does not exist in table dt_location (facility_id, sys_loc_code) [FK_dt_sample__65370702]. |
| INFO | sample_id | | 123 sample_id(s) updated |
| INFO | dt_test | | 66 row(s) processed |
| ERROR | dt_test | 161 20355006 20354928 ... | ('161', '20354928') cannot be inserted into child table dt_test (facility_id, sample_id) because a corresponding parent row does not exist in table dt_sample (facility_id, sample_id) [FK_dt_test__7C1A6C5A]. |
| ERROR | dt_test | 161 20355007 20354929 ... | ('161', '20354929') cannot be inserted into child table dt_test (facility_id, sample_id) because a corresponding parent row does not exist in table dt_sample (facility_id, sample_id) [FK_dt_test__7C1A6C5A]. |
| INFO | test_id | | 5067 test_id(s) updated |
| INFO | at_test_batch_assign | | 125 row(s) processed |
| ERROR | at_test_batch_assign | 161 20355006 ANALYSIS ... | ('161', '20355006') cannot be inserted into child table at_test_batch_assign (facility_id, test_id) because a corresponding parent row does not exist in table dt_test (facility_id, test_id) [FK_at_test_batch_as__4C0144E4]. |
| ERROR | at_test_batch_assign | 161 20355006 PREP ... | ('161', '20355006') cannot be inserted into child table at_test_batch_assign (facility_id, test_id) because a corresponding parent row does not exist in table dt_test (facility_id, test_id) [FK_at_test_batch_as__4C0144E4]. |
| ERROR | at_test_batch_assign | 161 20355007 ANALYSIS ... | ('161', '20355007') cannot be inserted into child table at_test_batch_assign (facility_id, test_id) because a corresponding parent row does not exist in table dt_test (facility_id, test_id) [FK_at_test_batch_as__4C0144E4]. |
| ERROR | at_test_batch_assign | 161 20355007 PREP ... | ('161', '20355007') cannot be inserted into child table at_test_batch_assign (facility_id, test_id) because a corresponding parent row does not exist in table dt_test (facility_id, test_id) [FK_at_test_batch_as__4C0144E4]. |
| INFO | dt_field_sample | | 57 row(s) processed |
| ERROR | dt_field_sample | 161 20354928 ... | ('161', '20354928') cannot be inserted into child table dt_field_sample (facility_id, sample_id) because a corresponding parent row does not exist in table dt_sample (facility_id, sample_id) [FK_dt_field_sample__035179CE]. |
| ERROR | dt_field_sample | 161 20354929 ... | ('161', '20354929') cannot be inserted into child table dt_field_sample (facility_id, sample_id) because a corresponding parent row does not exist in table dt_sample (facility_id, sample_id) [FK_dt_field_sample__035179CE]. |
| INFO | dt_result | | 2471 row(s) processed |
| ERROR | dt_result | 161 20355006 100-41-4 ... | ('161', '20355006') cannot be inserted into child table dt_result (facility_id, test_id) because a corresponding parent row does not exist in table dt_test (facility_id, test_id) [FK_dt_result__0880433F]. |
| ERROR | dt_result | 161 20355006 100-42-5 ... | ('161', '20355006') cannot be inserted into child table dt_result (facility_id, test_id) because a corresponding parent row does not exist in table dt_test (facility_id, test_id) [FK_dt_result__0880433F]. |

The columns of the Commit Error Log are the following:

- **Status** shows the points in the upload process where either data was loaded into a database table (INFO) or data failed to upload due to errors (ERROR). **Look for the first group of ERROR rows.**
- **Table** shows the table in the database that Enterprise EDP was attempting to load with new data from the EDD. **Note** that names of these tables do not always correspond exactly with the table (section) names in the EDD, but they do provide indications of where to find the problems in the EDD.

- **Row** is the row in the database table where the error occurred. **This is not a row in the EDD.** The codes in the Row entries do, however, provide information that will help in locating the sources of the errors in the EDD.
- **Message** is the error message. Although the error messages contain several codes that can be challenging to interpret, these messages provide important information about the nature of the errors and how to correct those errors.

The way to interpret and use the information in a Commit Error Log is to remember that **the first group of errors at the top of the log shows the actual errors in the EDD:**

| Status | Table | Row | Message |
|--------|-----------|--------------------------------|--|
| INFO | dt_sample | | 57 row(s) processed |
| ERROR | dt_sample | 161 20354928 IW-7-20220630 ... | ('161', 'IW-7') cannot be inserted into child table dt_sample (facility_id, sys_loc_code) because a corresponding parent row does not exist in table dt_location (facility_id, sys_loc_code) [FK_dt_sample__65370702]. |
| ERROR | dt_sample | 161 20354929 IW-9-20220629 ... | ('161', 'IW-9') cannot be inserted into child table dt_sample (facility_id, sys_loc_code) because a corresponding parent row does not exist in table dt_location (facility_id, sys_loc_code) [FK_dt_sample__65370702]. |

In this example, the first errors appear at the point where data was being processed for loading into the dt_sample data table. These data came from the Sample_v5 section of the EDD. **Note** that the EDD was rejected even though 57 rows of data were successfully processed for the dt_sample table before Enterprise EDP encountered the errors. **No part of an EDD will be uploaded to the EQulS database if the EDD contains any errors.**

Let's examine the first of these error entries in more detail.

The **Row** entry contain three codes. The first two (161, 20354928) are internal codes used by EQulS to designate facilities and samples – you can ignore these. **The third code in this entry, however, is the sys_sample_code 'IW-7-20220630' from the EDD.** When you load the EDD into EDP Standalone or Professional, you can locate the sample causing this error by searching for this sys_sample_code in the Sample_v5 section.

The **Message** entry contains three items of important information about the nature of the error and how to fix it. Closely examine the text of the message:

('161', 'IW-7') cannot be inserted into child table dt_sample (facility_id, sys_loc_code) because a corresponding row does not exist in table dt_location (facility_id, sys_loc_code) [FK_dt_sample__65370702].

The first part of the message to look at is in the first set of parentheses, “('161', 'IW-7')”. The first number is the *internal* identification number for the facility that is used by EQulS to connect all data for that facility in the database. This number bears no resemblance to the facility code used in the file name of the EDD and by the NYSDEC program that manages the facility, and you can ignore it. **The second code, 'IW-7', should look familiar as a location code.**

The next section of text, “cannot be inserted into child table dt_sample,” shows that a piece of data in the Sample_v5 section was dependent on corresponding data in another table in the database. This parent-child relationship between data tables is common throughout the EQulS database.

The next set of information is in the second set of parentheses, "(facility_id, sys_loc_code)". This identifies the codes in the first set of parentheses, showing that **'IW-7' is in fact a sys_loc_code**.

The next set of text and parentheses, "because a corresponding row does not exist in table dt_location (facility_id, sys_loc_code)," shows the cause of the error – *the sys_loc_code at the beginning of the message, 'IW-7', is not in the dt_location data table*. **A sample with a location reference (sys_loc_code), such as a normal sample or field duplicate, cannot be loaded into the EQulS database unless there is already information about that location in the dt_location table of the database.**

You can ignore the internal EQulS codes that appear in brackets at the end of the error message.

Recall that Enterprise EDP tries to upload data into the EQulS database in the order that the data appears in the EDD. If the location 'IW-7' (and all of its mandatory data) had been recorded in the Location_v5 section, which appears above the Sample_v5 section, that location would have already been ready for the database by the time EDP reached the Sample_v5 section.

The second error message contains the same information for missing location 'IW-9', which is referenced by sample 'IW-9-20220629' as shown in the **Row** column of the error log. The way to fix these errors and be able to upload this EDD is now clear – **submit all required data for the locations 'IW-7' and 'IW-9' in a separate EDD, then rename and resubmit your original EDD**. Adding the missing locations to your original EDD could cause other errors. Enterprise EDP will not accept two EDDs with the same name, but you can rename your original EDD simply by adding "Resubmittal" before the date stamp in the EDD file name.

Learning to Interpret and use the information in a Commit Error Log takes practice, but focus on the first group of errors every time and you will get the hang of it. If you need help interpreting one of these error logs, please forward the error log to NYENVDATA@dec.ny.gov and the NYSDEC EQulS Team will be happy to assist you.

The new EQulS Automated Data Loading System at NYSDEC will process EDDs much faster than the EQulS Team has been able to accomplish manually, but you will need to get used to interpreting the error logs you receive in automated email messages if any of your EDDs do not make it into the EQulS database. Please do not be intimidated by this – you can always reach out for help to the NYSDEC EQulS Team at NYENVDATA@dec.ny.gov.

Thank you for all the hard work you do to provide essential remediation data to NYSDEC. Your efforts contribute greatly to improving the health and natural environment of the people and wildlife of New York State.