



What Checks are Run Against My EIS CERS XML File?

EPA-454/B-19-014
May 2014

What Checks are Run Against My EIS CERS XML File?

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Air Quality Assessment Division
Research Triangle Park, NC

The Emissions Inventory System

What Checks are Run Against My EIS CERS XML File?

Introduction

The Emissions Inventory System process employs a number of different types of checks and validations on the EIS CERS XML file and the data contained within the file. The purpose of this section of the EIS Users Manual is to outline the various types of checking, identify when the checking occurs, point out what will cause the entire file or just some of the data to be rejected, and list the types of feedback you will receive at each step of the process.

Types of Checking and Validation

1. Schema Validation

Schema validation is the process by which the XML file is determined to be “well formed.” That is, does the schema meet the requirements of the EIS CERS XML schema and how it should be constructed. This includes cardinality rules, which are outlined in the NEIP’s *Section 5: Submitting XML Data to the EIS*. Schema validation does not check the data within the file.

While schema validation occurs as the second type of checks run during the submission process, it can happen in a few different ways. For users that rely on the EIS Bridge and the EIS Staging Tables, schema validation will be run when the staging tables are loaded into the EIS Bridge for conversion to an EIS CERS XML file. If a schema validation error is generated, it will appear in the EIS Bridge and the EIS CERS XML file will be saved to the file you designated at the start of the process with an indication that the file is “Invalid”.

A second place where the schema can be validated is EPA’s CDX Node site. You can access the site at <http://tools.epacdxnode.net/>. The schema validation performed in the EIS Bridge is the same validation that occurs at this site.

The final place where schema validation occurs is at the CDX Node when you submit your file to EIS. If a schema validation error occurs, the entire file is rejected and will not be submitted to EIS. You will receive a notification from the CDX Node containing the errors.

A few data validation checks occur at the CDX Node as well. These checks, six in total, verify user identifiers, date data types, percent data types, and program system codes. You will receive feedback if any of these items generate an error.



2. Identify Submission Information

The first thing EIS does when it receives an EIS CERS XML file is the following:

- Identify the submitter of the file.
- Identify the category of data in the file.
- Identify if the file is going to the QA or Production environment.
- Assign a transaction identifier to the file.

All of this information, along with the date and time the file was received, is stored in EIS. If EIS cannot determine the submitter of the file, the category of data within the file, or the intended environment, the entire file is rejected and feedback is sent to the user.

3. Data Transformation

After the submission information is verified and the schema is validated, the file is transformed to data. This is the process by which all the tags and supporting information in the EIS CERS XML file are removed, leaving only the emissions inventory data. A failure by EIS to transform the EIS CERS XML file to data will cause the processing of the data to end, and will cause the entire file to be rejected. Feedback is generated and a notification is sent to the user that the feedback can be viewed and downloaded from the EIS Gateway.

4. Quality Assurance Checks

Known as Business Logic Checks, quality assurance checks are applied against the data in the EIS CERS XML file. The same quality assurance checks are applied in both the QA and Production Environments. There are three levels of criticality for these checks.

- **Critical.** Any data element that fails a critical check is rejected. If a rejected data element is required for a component, the entire component is rejected. The notification details errors and suggests corrective actions. EIS is designed so that as little data as possible is rejected.
- **Warning.** Submitted data that fail a warning check are accepted but warrant further investigation and review by an EPA analyst. A warning notification details potential errors. Checks in this level may be upgraded to critical in a later inventory cycle.
- **Informational.** Submitted data that receive an informational message are accepted as provided. Additional information about the checked data will be made available.

The QA checks are applied based on the category of data in the EIS CERS XML file. Feedback is generated with the results of the QA checks and a notification is sent to the user that the submitted data and the resulting feedback can be downloaded from the EIS Gateway.

For a complete list of the quality assurance checks, see the NEIPS's *Appendix 5: QA Checks*.

United States
Environmental Protection
Agency

Office of Air Quality Planning and Standards
Air Quality Assessment Division
Research Triangle Park, NC

Publication No. EPA-454/B-19-014
May 2014
