Windsor Solutions, Inc.

SLEIS

FILE IMPORT SPECIFICATION FOR BATCH IMPORTS

SLEIS Version 3.3

File Import Specification for Batch Imports

Updated: December 2024

WINDSOR SOLUTIONS | Environmental + Health Information Systems

This page intentionally left blank

Version Control

Date	Author	Changes
10/24/2019	Windsor Solutions	Initial version for batch import process for OKDEQ on SLEIS 2.5
10/30/2020	Windsor Solutions	Added SCC column to Processes.csv
09/02/2022	Windsor Solutions	Confirm 2.8 compatibility.
12/29/2023	Windsor Solutions	Rev to 3.0 – added columns for unit and process descriptions, review comments
12/03/2024	Windsor Solutions	Rev to 3.3 – added report year, company, and facility info columns to ProcessSupplementalParameters section

Contents

Introduction

The purpose of this document is to describe the State and Local Emissions Inventory System (SLEIS) file import process and flat file specification for batch import processing at the company level. It includes standards and conventions to use when constructing an import file, describes the physical structure of each import file type, and provides basic instruction on how to create a SLEIS import file.

The contents of this document cover the following file import related topics:

File Import Process	Defines the file import process, including an overview and specifics for each step in the process.
CSV File Format	Describes the general design conventions of the Comma Separated Values (CSV) file format.
Data File Specification	Details the flat file specification for each data file type used in the file import process.
Appendix	Specifies data file reference values, and offers brief instructions on editing CSV file formats.

File Import Process

Overview

File imports will be initiated and completed using the SLEIS user interface. This interface provides functionality for:

1) Selecting the company's facilities' emissions reports that will receive the contents of the imported data files
2) Invoking the file import process for the selected emissions reports
3) Selecting the specific data files to load for the selected emissions reports
4) Executing the file import process, which is composed of the following three steps:
    a. Data file format validation
    b. Business rule validation
    c. Database load processing
5) Viewing the results of the file import process, such as data file format errors and business rule violations

Data File Format Validation

Each SLEIS data file must adhere to a pre-defined comma separated values (CSV) file format. Upon initiation of the file import process, the following data file format validation processing will be performed:

1) Validate that header columns in each CSV file are properly named, and that all appropriate columns exist in the file (refer to the *Data File Specification* section).
2) Validate that each CSV file line and field is properly formatted (refer to the *CSV File Format* section).
3) Validate that each CSV file line can be parsed into individual field values (refer to the *Data File Specification* section).

Business Rule Validation

After an attempt to import files, the user will receive an e-mail indicating how many of the facility reports were successfully imported, an attachment that shows the status of each of the facility reports, and a link to a page where the user can download the full list of validation errors. If there are no validation errors, the business rules engine will process the individual records extracted from the data file.

Database Load Processing

Following successful business rule validation, the data extracted from the imported files will be saved to the SLEIS database. Reference the appropriate *Data File Specification* section of this document for additional details.

Data insert and overwrite/replacement for process and emissions data will occur at the unit process level, and the file import process can be run multiple times for a report, inserting and replacing data as needed.

For example:

- File import 1 contains emissions data for emission unit A, B, and C
- File import 2 contains emissions data for emission unit C, D, and E

In this case, data for emission unit C would be overwritten by the second file import. Data for emission units A, B, D, and E would be inserted once, and not overwritten/replaced.

CSV File Format

A CSV data file format will be used to import data into SLEIS. This file format provides for the storage of structured data in a relatively simple format that can be easily created through automated (e.g., exported from a facility's information system) or manual (e.g., MS Excel, or even a basic text editor) means.

The CSV file format design conventions used by each SLEIS data file include:

1) Each record will be contained on one line of the file, terminated by a line feed, or a carriage return and line feed pair, with each field separated by a comma.
2) The first record in the CSV file will contain a column name header in each of the fields.
3) There is no comma following the final field in the record.
4) In case field values have embedded commas, embedded double-quote characters, intentional leading/trailing space characters, or other reserved characters, fields will always be enclosed within double-quote characters, whether necessary or not.
5) Leading and trailing spaces or tabs adjacent to commas (not within double-quotes) will be trimmed.
6) Any embedded double-quote characters must be represented by a pair of double-quote characters.

   For example, a CSV file containing:

```
“header 1”  , “header 2”  , “header 3” , “header 4” ,  “header 5”
“value 1”, “”, “value 3, value 3, value 3” , “   value 4   ” ,   “value “”value”” 5”
```

   Would be processed into individual rows and columns as:

header 1	header 2	header 3	header 4	header 5
value 1		value 3, value 3, value 3	value 4	value “value” 5

Data File Specification

Since the batch file import process can include more than one facility, the data files do require meta-data for report identification (e.g., reporting year, company identifier, facility identifier, etc.).

The emission report import payload will be composed of the following CSV data files, and will be processed in the order given:

1) Processes.CSV
2) ProcessEmissions.CSV
3) ProcessSupplementalParameters.CSV

Processes.CSV

General business rules:

- Will not allow import of process emissions with parent emissions units that have been "shutdown" in current reporting year or earlier
- Will not allow process emissions with unit process last/final inventory year less than current reporting year

Note, all columns must be completed except ones highlighted in grey. The ones highlighted in green are populated from SLEIS during the template download to help distinguish between records; any data shown in the green fields will be ignored during the import.

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
ReportYear	Integer	2023	Required.	(Multiple screens)
CompanyId	Text value	9999	Required.	(Multiple screens)
CompanyName	Text value	Test Company	Ignored during import. This field will be populated after downloading template but is for information purposes only.	(Multiple screens)
FacilityID	Text value	88888	Required.	(Multiple screens)
FacilityName	Text value	Test Facility	Ignored during import. This field will be populated after downloading template but is for information purposes only.	(Multiple screens)
EmissionUnitId	Text value	EU1	Required.	Process: Emission Unit Identifier

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
			The unique emission unit identifier for the facility.	
EmissionUnitDesc	Text value	Test EU	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Process: Emission Unit Description
ProcessId	Text value	PR1	Required. The unique process identifier for the emission unit for the facility.	Process: Process Identifier
ProcessDesc	Text value	Distillate Oil (Diesel)	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Process: Description
ProcessSCC	Text value	10100101	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Process: SCC
IsReported	Text value of: • TRUE • FALSE	TRUE	Defaults to TRUE if not specified. NOTE: If IsReported value is FALSE, no process data or process emissions data will be imported for this process (if any has been specified).	Process: Process is Reported?
IsConfidential	Text value of: • TRUE • FALSE	FALSE	Defaults to FALSE if not specified. NOTE: If IsConfidential value is TRUE, throughput quantity and process emission factors will be treated as confidential business information (CBI), and not reported to the EPA.	Process: Process is CBI?
ThroughputQuantity	Floating point number	7183.175	Annual Throughput/Process Rate **OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (25,10)	Process: Annual Throughput
ThroughputUnit	Text value*		*Refer to the valid reference values provided by the agency for this attribute in either the ReferenceThroughputValues.CSV file, or by	Process: Throughput Unit of Measure

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
			referring to the equivalent Process Emissions page field within the application. **OKDEQ Business Rules:** • Required	
ThroughputType	Text value*		*Refer to the valid reference values provided by the agency for this attribute in either the ReferenceThroughputValues.CSV file, or by referring to the equivalent Process Emissions page field within the application. **OKDEQ Business Rules:** • Required	Process: Throughput Type
ThroughputMaterial	Text value*		*Refer to the valid reference values provided by the agency for this attribute in either the ReferenceThroughputValues.CSV file, or by referring to the equivalent Process Emissions page field within the application. **OKDEQ Business Rules:** • Required	Process: Throughput Material
Comments	Text value up to 4000 characters in length		**OKDEQ Business Rules:** • Conditional. Optional if IsReported = TRUE, otherwise required	Process: Comments
ReviewComments	Text value up to 4000 characters in length		Ignored during import. This field will be populated after downloading template but is for information purposes only. If the report had been returned for amendments, any review comments made by agency staff would appear here.	Process: Review Comments
OpStartTime	Time, including AM or PM	8:00 AM	**OKDEQ Business Rules:** • Not used. Leave blank.	Operations: Standard Start Time
OpStopTime	Time, including AM or PM	5:00 PM	**OKDEQ Business Rules:** • Not used. Leave blank.	Operations: Standard End Time

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
AvgHrsPerDay	Floating point number, >=0 and <= 24	22.5	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (3,1)	Operations: Average Hours/Day
AvgDaysPerWeek	Floating point number, >=0 and <=7	6.5	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (2,1)	Operations: Average Days/Week
AvgWeeksPerYear	Floating point number, >=0 and <=52	50.53	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (3,1)	Operations: Average Weeks/Year
ActualDaysPerPeriod	Floating point number, >=0 and <=366	250	**OKDEQ Business Rules:** • Not used. Leave blank. • Precision and scale (P,S) is limited to (4,1)	Operations: Actual Days/Year
ActualHrsOperation	Floating point number >=1 and <= 8784	7233	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (5,1)	Operations: Actual Hours/Year
DecToFebPercent	Floating point number, >= 0 and <= 100	24.5	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (4,1)	Operations: Season Operations December-February (%)
MarToMayPercent	Floating point number, >= 0 and <= 100	25.5	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (4,1)	Operations: Season Operations March-May (%)
JunToAugPercent	Floating point number, >= 0 and <= 100	25	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (4,1)	Operations: Season Operations June-August (%)
SepToNovPercent	Floating point number, >= 0 and <= 100	25	**OKDEQ Business Rules:** • Required • Precision and scale (P,S) is limited to (4,1)	Operations: Season Operations September-November (%)
TotalOzoneSeasonDays	Integer, 0-153	153	Days operated May-September. **OKDEQ Business Rules:** • Not used. Leave blank.	Operations: Total Ozone Season Days (May-September)

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
			• Precision and scale (P,S) is limited to (3,0)	
TotalSummerSeasonDays	Integer, 0-92	92	Days operated June-August. **OKDEQ Business Rules:** • Not used. Leave blank. • Precision and scale (P,S) is limited to (2,0)	Operations: Total Summer Season Days (June-August)
TotalCOSeasonDays	Integer, 0-91	91	Days operated December-February. **OKDEQ Business Rules:** • Not used. Leave blank. • Precision and scale (P,S) is limited to (2,0)	Operations: Total CO Season Days (December-February)

ProcessEmissions.CSV

General business rules:

- At least one pollutant must be specified for each process reported in the Processes.csv file
- Pollutants cannot be duplicated (reported twice) for the same emission unit and process

Note, all columns must be completed except ones highlighted in grey. The ones highlighted in green are populated from SLEIS during the template download to help distinguish between records; any data shown in the green fields will be ignored during the import.

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
ReportYear	Integer	2018	Required.	(Multiple screens)
CompanyId	Text value	9999	Required.	(Multiple screens)
CompanyName	Text value	Test Company	Ignored during import. This field will be populated after downloading template but is for information purposes only.	(Multiple screens)
FacilityID	Text value	88888	Required.	(Multiple screens)
FacilityName	Text value	Test Facility	Ignored during import. This field will be populated after downloading template but is for information purposes only.	(Multiple screens)

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
EmissionUnitId	Text value	EU1	Required. The unique emission unit identifier for the facility.	Process: Emission Unit Identifier
EmissionUnitDesc	Text value	Test EU	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Process: Emission Unit Description
ProcessId	Text value	PR1	Required. The unique process identifier for the emission unit for the facility.	Process: Process Identifier
ProcessDesc	Text value	Test PR	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Process: Process Description
PollutantCode	Text value*	NOX	Required. *Refer to the valid reference values provided by the agency for this attribute in either the ReferenceDataValues.CSV file (search for *PollutantCode* in the AttributeName column), or by referring to the equivalent Process Emissions page field within the application.	Emissions: Pollutant Code
CalculationMethod	Text value*	4_0	Required. *Refer to the valid reference values provided by the agency for this attribute in either the ReferenceDataValues.CSV file (search for *CalculationMethod* in the AttributeName column), or by referring to the equivalent Process Emissions page field within the application.	Emissions: Calculation Method
EmissionFactor	Floating point number	3.3579	In Lbs per EmissionFactorUnit **OKDEQ Business Rules:** • Required if CalculationMethod code is '3_1', '3_2', '4_0', '4_1', '4_2', '7_0', '7_1', '7_2',	Emissions: Emission Factor (Lbs/Unit)

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
			'8_1', '8_2', '8_3', '10_1', '10_2', '10_3', '12_1', '12_2', '12_3', '44_0', '44_1', '44_2' • Not allowed (leave blank) if CalculationMethod code anything other than ones listed above • Precision and scale (P,S) is limited to (28,15)	
EmissionFactorUnit	Text value*		*Refer to the valid reference values provided by the agency for this attribute in either the ReferenceDataValues.CSV file, or by referring to the equivalent Process Emissions page field within the application. NOTE: The EmissionFactorUnit value supplies the denominator of the implied "Lbs" per EmissionFactorUnit of the EmissionFactor value. **OKDEQ Business Rules:** • Required if CalculationMethod code is '3_1', '3_2', '4_0', '4_1', '4_2', '7_0', '7_1', '7_2', '8_1', '8_2', '8_3', '10_1', '10_2', '10_3', '12_1', '12_2', '12_3', '44_0', '44_1', '44_2' • Not allowed (leave blank) if CalculationMethod code anything other than ones listed above	Emissions: Emission Factor Unit
EmissionQty	Floating point number	2978.45678	Required. Estimated emissions in Tons. NOTE: If a CalculationMethod is specified that requires an EmissionFactor and EmissionFactorUnit, AND the EmissionFactorUnit matches the Processes.csv Throughput Unit, the EmissionQty will be calculated upon import, and the EmissionQty from the import file will be ignored. If CaclulationMethod is specified that DOES NOT require EmissionFactor, EmissionFactor and	Emissions: Estimated Emissions (Tons)

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
			EmissionFactorUnit cannot be specified, and EmissionQty must be specified. **OKDEQ Business Rules:** • Must be >= 0 • Precision and scale (P,S) is limited to (28,15)	
StackTestDate	Date format		**OKDEQ Business Rules:** • Required if CalculationMethod code is '4_0', '4_1', '4_2', '44_0', '44_1', '44_2'	Emissions: Stack Test Date
Comments	Text value up to 4000 characters in length		Optional comments/notes regarding the process emission record.	Emissions: Comments

ProcessSupplementalParameters.CSV

General business rules:

- This is an optional import file
- Values entered will be validated based on configurations stored in the database for each parameter (variable). For example, fields that are percentages must be a valid number greater than 0 and less than or equal to 100.

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
ReportYear	Integer	2018	Required.	(Multiple screens)
CompanyId	Text value	9999	Required.	(Multiple screens)
CompanyName	Text value	Test Company	Ignored during import. This field will be populated after downloading template but is for information purposes only.	(Multiple screens)
FacilityID	Text value	88888	Required.	(Multiple screens)
FacilityName	Text value	Test Facility	Ignored during import. This field will be populated after downloading template but is for information purposes only.	(Multiple screens)
EmissionUnitId	Text value	EU1	Required	Process: Emission Unit Identifier
EmissionUnitDesc	Text value	Caterpillar Diesel	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Emission Units: Description
ProcessId	Text value	PR1	Required	Process: Process Identifier
ProcessDesc	Text value	Distillate Oil (Diesel)	Ignored during import. This field will be populated after downloading template but is for information purposes only.	Process: Description
ParameterCode	Text value	A	Required. Leave as is. This field will be populated after downloading template but is for information purposes only.	Process: Supplemental Calculation Parameters

Column Header	Allowed Values	Example	Business Rules/Notes	Process Emissions Page Field (Tab: Field Label)
ParameterName	Text value	% Ash	Leave as is. This field will be populated after performing a Download Template operation, but is for information purposes only.	Process: Supplemental Calculation Parameters
Value	Floating point number	2.6	Required	Process: Supplemental Calculation Parameters

Appendix

Reference Data Values

Reference data used to validate values within the imported data files is specific to each agency, and is not documented within this specification. The *Download Template* function will generate the following CSV files:

1. Processes.CSV

   Data from the currently selected facilities' reports, formatted per specification in the *Data File Specification* section for this CSV file type.

2. ProcessEmissions.CSV

   Data from the currently selected facilities' reports, formatted per specification in the *Data File Specification* section for this CSV file type.

3. ProcessSupplementalParameters.CSV

   Data from the currently selected report, formatted per specification in the *Data File Specification* section for this CSV file type.

4. ReferenceDataValues.CSV (see below)

   A reference file that contains the list of allowed values that are used in order to pass validation rules per the specification in the *Data File Specification* for any CSV file type.

Column Header	Example	Comments/Notes
AttributeName	AnnThroughputUnit	Corresponding column header/attribute name as documented in the data file specification.
Value	GAL	Code/value to input into the import data file.
Description	GALLONS	Descriptive text of the code/value specified in the value column.
Notes		Additional descriptive text or instructions on the attribute value.

5. ReferenceThroughputValues.CSV (see below)

   A reference file used to select a valid throughput combination for the Processes.CSV file. It contains the list of valid throughput combinations based on Source Classification Code that are used in order to pass validation rules. To use, find the SCC number of the Process, select the correct ThroughputUnit, ThroughputType, and ThroughputMaterial (columns B, C, and D), then copy and paste them into Processes.CSV file (columns F, G, and H).

Column Header	Example	Comments/Notes
SCC	10100101	The Source Classification Code value. Search this column for the SCC for which allowed throughput combinations are needed.
ThroughputUnit	TON	Throughput unit of measure code to input into the import data file.
ThroughputType	I	Throughput type code to input into the import data file. I = Input, O = Output, E = Existing

Column Header	Example	Comments/Notes
ThroughputMaterial	640	Throughput material code to input into the import data file.
MaterialType	Anthracite	Descriptive text of the throughput material name. It will not be used in the import data file, but will be used as a reference when selecting material code (above).

Editing CSV Files

Ideally, generation of SLEIS-compatible CSV files will be handled programmatically through a facility's existing information system(s) or possibly through Microsoft Excel macros/scripting.

For manual editing of the CSV files, the default editor for most users will likely be Microsoft Excel, which recognizes the CSV file format and automatically loads and formats the data. Any text editor may also be used (e.g., Notepad, etc.) to edit CSV files, as well as any number of free CSV file editors found online.

Note that one significant limitation of editing CSV files with Microsoft Excel is that an undesired data conversion will occur when opening a CSV file with leading 0's (zeros) in alpha-numeric fields. For example, a process name, or code value with all numeric values will be converted to a numeric value with leading zeros removed upon opening the file. This can cause values such as '001' to be converted to the number 1 (no leading zeroes).

To edit a CSV file in MS Excel without having leading zeros removed from fields, perform the following steps:

1. Extract the individual .CSV files from the .ZIP file generated using the Export or Download Template function to your local system (e.g., C:\Users\[username]\Desktop)
2. Open the MS Excel application (Note: NOT by double-clicking the file, just open the application itself)
3. Click on the *Data* tab
4. On the *Data* tab, in the *Get External Data* section, click *From Text* on the ribbon bar. A file selection dialog will be opened.
5. Navigate to where you saved the files and select one (e.g., Processes.csv), then click *Import*. The Text Import Wizard dialog is opened.
6. Click the box next to the text *My data has headers*.
7. Ensure the *Delimited* file type option is selected, then click *Next*.
8. Ensure *Comma* is the only item selected in the *Delimiters* section and click the *Next* button.
9. In the *Data Preview* section, select the field that could contain leading zeros (e.g., the EmissionUnitId and ProcessId columns) and select the *Text* option in the *Column data format* section. Repeat for all columns that could contain leading zeros you want to retain in the value.
10. Click *Finish*.

Note: The file can then be edited and saved as an MS Excel document, but before importing it into SLEIS, you will need to save as .CSV file format (through *File -> Save As* dialog in MS Excel, choose .csv file format).