



# 2017 NEI Final Plan, Appendix 6: EIS Requirements Document - Enhancement to Option Group / Option Set Functionality



EPA-454/R-19-011C  
April 2019

2017 NEI Final Plan, Appendix 6: EIS Requirements Document - Enhancement to Option  
Group / Option Set Functionality

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

# EIS Requirements Document

Issue Description: Enhancement to Option Group / Option Set Functionality

EIAG Point(s) of Contact: Rich Mason

## Summary

In the Source Classification Code table, EIAG can define SCCs that have a hierarchical nature. That is there may be a “General” group, but then more specific SCCs within the same group. These relationships are defined by the “Option Group / Option Set” (OGOS) fields in the SOURCE\_CLASSIFICATION\_CODE table. During the selection process, if all data were compiled across multiple data sets, there is the potential for the double-counting of data sources. For example: The EPA may report data to a “General” SCC, but the State / Local / or Tribal (SLT) organization submits to the detail SCCs. Without the OGOS evaluation, both sets of data would be included in the selection.

The current OGOS rules employed in the Selection assumes that if a SLT submits data, they are submitting data for the entire group and no additional data sets are to be used to “back-fill” any SCCs within the same option set. The desired function is for the selection to back-fill any SCCs within the same option set.

## Definitions

### Option Group:

The Option Group defines the general categories for which an SCC hierarchy exists. Examples: “Cattle”, “Commercial”, “Poultry”, “Goat”, etc.

### Option Set:

An Option Set is a group (1 or more) SCCs that exist within an Option Group. There may be multiple levels of an Option Set. The highest-ranking option set is determined alphabetically. Using the “Goat” Option Group example from above, the option sets (in priority order) may look like this:

| Option Group | Option Set | SCC Level 3 Description | SCC Level 4 Description  |
|--------------|------------|-------------------------|--------------------------|
| Goat         | A          | Goats Waste Emissions   | Not Elsewhere Classified |
| Goat         | B          | Goats Waste Emissions   | Milk Goats               |
| Goat         | B          | Goats Waste Emissions   | Angora Goats             |

### Ranking Dataset

The Ranking Data Set is the defined by the Selection Definition in EIS. When creating a selection, the list of data sets to be used within the selection must be specified. These are assigned in priority order. The data set that has the highest priority that is found for the given Option Group will define what level of Option Set to use.

## Business Rules to be Used

1. The OGOS rules cannot be applied to the Point Data Category
2. If no OGOS exist for an SCC, the following hierarchy will be used:
  - a. Highest Ranking Data Set
  - b. Location
  - c. SCC
  - d. Emissions Type Code
  - e. Pollutant
3. If a OGOS does exists, the selection process will determine what Option Sets within an Option Group to use based on the following:
  - a. Location
  - b. Pollutant
  - c. Option Group
  - d. Minimum Option Set Level (Length to Evaluate and the Group Letter. "3-B" for Example)
  - e. Data Set Ranking
  - f. Highest Ranking Option Set within the Group
4. The structure of the Option Sets will be assumed to be as follows:
  - a. The first position must be a capital letter ("A", "B", "C", etc.) and must be populated if there is an Option Group Defined for the SCC.
  - b. The second position (if provided) must be a number between 1 and 9.
  - c. The third position (if provided) must be a capital letter ("A", "B", "C", etc.)
  - d. If there is a second position provided, there must also be a third position provided.
  - e. No more than 3 levels may be specified
  - f. The first-position letter assignment should be thought of as "This is the set that I want to have a higher priority over another" rather than "'A' is assigned to the total grouping".
5. The ranking of the Option Sets will be as Follows:
  - a. The lowest alphabetical value for the first position of the Option Set value
    - i. An "A1A" SCC will rank higher than a "B" SCC within the Option Group
  - b. Option Sets with 1 character rank higher than those with 3 characters
    - i. If SLTs supply SCCs with both an "A" and an "A1A", then only the "A" SCCs will be used.
  - c. Option Sets with 3-character Option Sets are Grouped by the first 2 characters in the Option Set description. Within each of these groups, the third position will determine the ranking.
    - i. Example:
      1. In the selection definition, the SLT data set ranks higher than the EPA data set.
      2. An SLT reports SCCs with the following Option Sets within an Option Group: A1A, A2A, A2B, A3B, and A3C.
      3. The EPA Submits data for this Option Group with the following Option Sets: A, B, A1A, A2A, A3A, A4A

4. Since the SLT has the priority, it will set what level of Option Sets to be used. In this case, it is the second level of the Option Sets (since they did not supply any “A” or “B” Option Set SCCs). The Option Sets to Evaluate within the Option Group are:
  - a. A1 (From SLT)
  - b. A2 (From SLT)
  - c. A3 (From SLT)
  - d. A4 (From EPA)
5. The highest-ranking Option Sets within each of these groups are:
  - a. A1A (From SLT)
  - b. A2A (From SLT)
  - c. A3B (From SLT)
  - d. A4A (From EPA)

## Examples

The following examples are to illustrate how different data scenarios would be implemented with these rules. They are strictly for informational purposes only and do not reflect any specific known data submission scenario. A key to the table in all the examples are:

- ✓ = The agency Submitted Data
- = The data point will be selected

### Example 1:

Selection Definition (Listed in Priority Order):

1. SLT Data
2. EPA Dataset 1
3. EPA Dataset 2

Submitted Data by Each Organization

*Option Set = RWC\_Hydronic*

| SCCs in Option Set | SCC Description                                   | Option Set Code | SLT Data | EPA Dataset 1 | EPA Dataset 2 |
|--------------------|---|-----------------|----------|---------------|---------------|
| 2104008600         | Hydronic heater: general, all types               | A               |          | ✓             |               |
| 2104008640         | Hydronic heater: meets NESCAUM phase II standards | B               | ✓        |               | ✓             |
| 2104008630         | Hydronic heater: pellet-fired                     | B               |          |               | ✓             |
| 2104008610         | Hydronic heater: outdoor                          | B               |          |               | ✓             |
| 2104008620         | Hydronic heater: indoor                           | B               | ✓        |               |               |

What Data Gets Included and Why (or Why Not)?

- None of the “A” Option Set SCCs will be chosen because the SLT Data is the highest-ranking data set. The highest-ranking Option Set Code they submitted was “B”, thus making the “Minimal Option Set Level” for this example “1 – B” (one character long, set code of “B”)
- 2104008640 – Selected because it was submitted by the SLT and is in the “B” Option Set

- 2104008630 – Selected because it is the highest-ranking data set for the SCC and is in the “B” Option Set
- 2104008610 – Selected because it is the highest-ranking data set for the SCC and is in the “B” Option Set
- 2104008620 – Selected because it was submitted by the SLT and is in the “B” Option Set

### Example 2:

Selection Definition (Listed in Priority Order):

1. SLT Data
2. EPA Dataset 1
3. EPA Dataset 2

Submitted Data by Each Organization

*Option Set = WW Treatment*

| SCCs in Option Set | SCC Description                                   | Option Set Code | SLT Data | EPA Dataset 1 | EPA Dataset 2 |
|--------------------|---|-----------------|----------|---------------|---------------|
| 2630010000         | Industrial Total Processed                        | B1A             |          | ✓             |               |
| 2630020000         | Public Owned Total Processed                      | B2A             |          |               | ✓             |
| 2630030000         | Residential/Subdivision Owned Total Processed     | B3A             | ✓        |               | ✓             |
| 2630020001         | Public Owned Flaring of Gases                     | B2B             |          |               | ✓             |
| 2630020010         | Public Owned Wastewater Treatment Processes Total | B2B             | ✓        |               | ✓             |
| 2630050000         | Public Owned Land Application - Digested Sludge   | B2B             | ✓        |               |               |

What Data Gets Included and Why (or Why Not)?

- The first thing to consider in this example is that since all the option set codes are three digits long, the Minimum Option Set Level is “3-B”. So, we need to evaluate the first two digits to see where they fall in the hierarchy.
- 2630010000 is selected because it is in the “3-B” level and is the only value reported in the “B1” group
- 2630020000 is not selected because the SLT data set is higher than it is within the “3-B” level.
- 2630030000 is selected because it is in the “3-B’ level and the only one reported at the “B3” group
- 2630020001 is selected because it is in the “3-B’ level and the only one of this SCC reported at the “B2B” group
- 2630020010, and 2630050000 are selected because they are from the highest-ranking data set within the “3-B” level. It is these records that set “B2B” as the highest-ranking Option Set Code within the Minimum Option Set Level.

### Example 3:

Selection Definition (Listed in Priority Order):

1. SLT Data
2. EPA Dataset 1

### 3. EPA Dataset 2

Submitted Data by Each Organization

*Option Set = P and P Product Tran*

| SCCs in Option Set | SCC Description                         | Option Set Code | SLT Data | EPA Dataset 1 | EPA Dataset 2 |
|--------------------|---|-----------------|----------|---------------|---------------|
| 2505000000         | All Transport Types Total: All Products | A               |          | ✓             |               |
| 2505000030         | All Transport Types Crude Oil           | B               |          |               | ✓             |
| 2505000060         | All Transport Types Residual Oil        | B               |          |               | ✓             |
| 2505000090         | All Transport Types Distillate Oil      | B               |          |               | ✓             |
| 2505000120         | All Transport Types Gasoline            | B               |          |               | ✓             |
| 2505000150         | All Transport Types Jet Naphtha         | B               |          |               | ✓             |
| 2505000180         | All Transport Types Kerosene            | B               |          |               | ✓             |
| 2505000900         | All Transport Types Tank Cleaning       | B               |          |               | ✓             |
| 2505010000         | Rail Tank Car Total: All Products       | B1A             | ✓        |               |               |
| 2505010030         | Rail Tank Car Crude Oil                 | B1B             |          | ✓             |               |
| 2505010060         | Rail Tank Car Residual Oil              | B1B             |          | ✓             |               |

What Data Gets Included and Why (or Why Not)?

- The first thing to consider in this example is determining the Minimum Option Set Level. Out of this Option Group, the SLT is the highest-ranking data set. The level where they submitted data is “3-B”.
- None of the Option Set “A” or “B” values will be selected because of the Minimum Option Set Level Being “3-B”.
- The 2 “B1B” values are not selected because the SLT reported values at the “B1A” level.
- The only value that is selected here is for SCC 2505010000.

Example 4:

Selection Definition (Listed in Priority Order):

1. SLT Data
2. EPA Dataset 1
3. EPA Dataset 2

Submitted Data by Each Organization

*Option Set = P and P Product Tran*

| SCCs in Option Set | SCC Description                         | Option Set Code | SLT Data | EPA Dataset 1 | EPA Dataset 2 |
|--------------------|---|-----------------|----------|---------------|---------------|
| 2505000000         | All Transport Types Total: All Products | A               |          | ✓             |               |
| 2505000030         | All Transport Types Crude Oil           | B               |          |               | ✓             |
| 2505000060         | All Transport Types Residual Oil        | B               |          |               | ✓             |
| 2505000090         | All Transport Types Distillate Oil      | B               |          |               | ✓             |
| 2505000120         | All Transport Types Gasoline            | B               |          |               | ✓             |
| 2505000150         | All Transport Types Jet Naphtha         | B               |          |               | ✓             |
| 2505000180         | All Transport Types Kerosene            | B               |          |               | ✓             |



| SCCs in Option Set | SCC Description                   | Option Set Code | SLT Data | EPA Dataset 1 | EPA Dataset 2 |
|--------------------|-----------------------------------|-----------------|----------|---------------|---------------|
| 2505000900         | All Transport Types Tank Cleaning | B               |          |               | ✓             |
| 2505010000         | Rail Tank Car Total: All Products | B1A             | ✓        |               |               |
| 2505020000         | Marine Vessel Total: All Products | B2A             |          | ✓             | ✓             |
| 2505030000         | Truck Total: All Products         | B3A             |          | ✓             | ✓             |
| 2505040000         | Pipeline Total: All Products      | B4A             |          | ✓             | ✓             |
| 2505010030         | Rail Tank Car Crude Oil           | B1B             |          | ✓             |               |
| 2505010060         | Rail Tank Car Residual Oil        | B1B             |          | ✓             |               |
| 2505020030         | Marine Vessel Crude Oil           | B2B             | ✓        |               |               |
| 2505020041         | Marine Vessel Lube Oil-Barge      | B2B             |          | ✓             |               |
| 2505030030         | Truck Crude Oil                   | B3B             |          | ✓             | ✓             |
| 2505030060         | Truck Residual Oil                | B3B             | ✓        | ✓             | ✓             |
| 2505040030         | Pipeline Crude Oil                | B4B             | ✓        |               |               |
| 2505040060         | Pipeline Residual Oil             | B4B             |          |               | ✓             |

#### What Data Gets Included and Why (or Why Not)?

- The first thing to consider in this example is determining the Minimum Option Set Level. Out of this Option Group, the SLT is the highest-ranking data set. The level where they submitted data is “3-B”.
- None of the Option Set “A” or “B” values will be selected because of the Minimum Option Set Level being “3-B”.
- The 2 “B1B” values are not selected because the SLT reported values at the “B1A” level.
- SCC 2505010000 is selected because it is the highest-ranking dataset within the “B1A” group in the “3-B” Minimum Option Set Level.
- None of the B2A, B3A, or B4A SCC are selected because there is a higher-ranking data set that reported data at a different level.
- 2505020030 and 2505020041 are selected because they are the highest-ranking data sets in the B2B grouping.
- 2505030030 and 2505030060 are selected because they are the highest-ranking data sets in the B3B grouping of SCCs.
- 2505040030 and 2505040060 are selected because they are the highest-ranking data sets in the B4B grouping of SCCs.

---

United States  
Environmental Protection  
Agency

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

Publication No. EPA-454/R-19-011C  
April 2019

---