

# Emissions Data Model Product Development Team

## Project Team Report – Phase II

### Table of Contents

Team Participants .....	1
Project Overview.....	1
Phase I - Recap .....	2
Phase II.....	2
Phase II – Task.....	4
Task 1 – Data solutions and Documentation .....	4

### Team Participants

Agency	Name	Email
Alabama Dept of Environmental Management	Anna Watkins Wood	alwatkins@adem.alabama.gov
Idaho Department of Natural Resource	Marnie Stein	marnie.stein@dnr.iowa.gov
Massachusetts Dept of Environmental Protection	Mark Wert, Co-Chair	mark.wert.state.ma.us
Mississippi Dept of Environmental Quality	Elliott Bickerstaff	ebickerstaff@mdeq.ms.gov
Mississippi Dept of Environmental Quality	Matt Carpenter	mcarpenter@mdeq.ms.gov
Mississippi Dept of Environmental Quality	Deborah Boleware	dboleware@mdeq.ms.gov
North Carolina Dept of Environment and Natural Resource	Tammy Manning	tammy.manning@ncdenr.gov
South Carolina Dept of Health and Environmental Control	Dave McClard	mcclarad@dhec.sc.gov
Wyoming Dept of Environmental Quality	Ben Way	ben.way@wyo.gov
US EPA	Jonathan Miller	miller.jonathan@epa.gov
US EPA	Brandon Little	little.brandon@epa.gov
US EPA	Sally Dombrowski, Co-Chair	dombrowski.sally@epa.gov

### Project Overview

As part of the Combined Air Emission Reporting (CAER) project, the EPA and State, Local and Tribal (SLT) air programs are working together to identify opportunities to reduce redundancy, improve quality, and increase efficiency in the reporting of air emissions from facilities. As part of CAER, The Product Design

Team (PDT) comprises members from SLT and EPA programs that are carrying out specific R&D projects that further the implementation objectives of the overall CAER project. The PDT has the responsibility for

- Segmenting the CAER implementation work
- Prioritizing supporting research activities
- Setting up the smaller Research and Development (R&D) teams to do the work
- Setting the project scope and expectations for these teams
- Enabling these teams to be successful
- Integrating the outcomes of the R&D teams into future activities and the proposed vision for the CAER future state.

The R&D projects involve a range of policy and program research activities related to identifying program needs, analyzing business rules and quality assurance/quality control procedures across programs, review of program regulatory requirements, and other program characteristics and functions that are important considerations for creating a shared emissions system under the CAER proposed future state.

#### Phase I - Recap

The Emissions Data Model project team researched, and documented emissions-related data elements and functionality needed to support a possible common emission form (CEF) reporting design structure in a shared emissions platform. The identified state-specific data elements will be sufficient to allow for broad usage by SLTs and applicable EPA reporting programs.

The SLT team members distributed an on-line survey to emissions inventory contacts nationwide, including localities and tribes. The team received 50 responses. The full results of Phase I can be found on the [CAER website](#).

#### Phase II

The Emissions Data Model team identified four tasks to be accomplished in Phase II of this project.

1. **Data solutions and documentation** – The Phase I data survey highlighted many individual issues that need to be investigated, resolved, and documented to specify a CEF. The effort to resolve any one of these issues is modest but together represent a significant amount of work. The tasks range from specifying field size/type to describing calculation functionality for annual and ozone season emissions. This task included both documenting SLT-specific data identified in the Phase 1 survey and identifying differences in SLT and NEI for fields in common (e.g., field size). This task also included identifying fields/features that will need to have some mechanism for SLT control for filers in their individual states and the nature of that control. For example, a mechanism is needed for a SLT to opt-in/out of a particular CEF feature and to control the flow of data after it is collected by the CEF. Part of this task included identifying issues that may need to be postponed for later consideration (i.e., might be too time-consuming or undefined to be addressed in time for a pilot). The results from this task are the raw materials for building a CEF.

2. **Business Rules for CEF** – Document discussions on basic business rules needed for the CEF. Business rules are specific to each SLT, and so specify those rules determines needs for the controls to configure the CEF, set available emission factors and for preventing filers from updating certain data fields when completing the form.. The business rules will take into consideration the PDT R&D team findings on confidential business information (CBI).
3. **Workplan for development of a CEF** – Collaborate with the EPA IT lead to develop a checklist of the documentation needed to create the CEF. The items in this list would be those necessary for a pilot and would aid in the management of the pilot development process. The goal of this task is to (1) assess the gap between what specifications have been developed or are underway and what is needed to specify a system that is fully functional for the pilot SLT, and (2) specify what work is needed to fill that gap. This task must inherently address SLT/NEI needs for collecting facility data. The results would be used to organize a procurement process. This task was intended to document the specifications needed for a more broadly usable CEF.
4. **Targeted pre-pilot assessment** – An original Phase 2 concept was to assess the specific needs of a potential pilot SLT in preparation for the pilot project, provided that a pilot SLT could be identified. Such an approach would apply the results of task 1 above to a real test case. The task was to investigate the specific needs of the pilot state and compare them against the CEF as documented thus far. Those needs would be a subset of the needs for a final CEF. The results of the task would identify any additional specifications and documentation needed to ensure the CEF will successfully meet the needs of the pilot SLT (including the handling of facility data). The objective is to specify a minimally viable product targeted at the pilot state.

During our review on Task 1, the Data Model Team quickly realized the amount of work which Task 1 alone would encompass, and the decision was made to only complete Task 1 in Phase II. Task 2, Business Rules, was deferred to Phase III and the CAER pilot. Task 3 and 4, the Workplan for the CEF and Targeted Pre-Pilot Assessment, were passed back to the full PDT for completion under the CEF Pilot.

## Phase II – Task

### Task 1 – Data solutions and Documentation

#### *Phase 1 Outstanding Survey Questions*

As a result of the on-line survey to emission inventory contacts during Phase 1, several questions arose that required additional contact/discussion for clarification. These questions and resulting team recommendations from the discussions are as follows:

1. Addition of Pollutant Codes to the Emission Inventory System (EIS) – The EIS has a limited number of pollutants that are accepted and included in the National Emission Inventory. Because of this limitation in the current state, SLTs need to create multiple inventories to meet EPA requirements. Survey results showed as many as 800+ pollutants beyond the EIS pollutants were used by SLTs. EIS developers stated that a process could be developed to add the additional pollutants to the EIS and still limit what is included in the NEI. This would allow SLTs to submit all and any pollutants collected. All pollutants would be stored and available in EIS but would not necessarily be included in the NEI.
  - i. **Recommendation** – Allow all pollutants to be accepted by EIS
2. Addition of Percent Sulfur and Percent Ash Data Fields to EIS – Percent Sulfur and Percent Ash already exist in EIS. The three States which requested this addition were informed and are satisfied. One state did request that the lower range check that currently exists in EIS, be changed to 0.0001 to account for low sulfur requirements and standards.
  - i. **Recommendation** - Change to the QA range check will be considered for the 2020 NEI.
3. Billable/Nonbillable – Clarification was needed after several respondents to the survey requested the addition of a billable/nonbillable data field or flag. Findings:
  - i. SLTs should have the option to make billable data fields available to facility filers for filers to denote which emissions are billable
  - ii. A billable data field can be collected at the facility type, unit type, or pollutant, as specified by the SLT.
  - iii. A consideration is that some SLTs have rules which limit the tonnage billed. Actual ton should not be displayed which adds a level of complexity to this addition.
  - iv. **Recommendation** – Allow SLTs to display billable flags to the facility filer as part of the facility attributes on the CEF. The Facility IPT Team will be advised of this additional field request. Since this field can also be on the pollutant level as well,, it will be considered for the emissions level of the CEF
4. Insignificant Source/Activity – An insignificant source is a whole facility that may be collected by an SLT, but the emissions data are not submitted to EIS. An insignificant activity is a unit or process which is below reporting thresholds (e.g., an SLT HAP reporting threshold) and is not reported to EIS but may be reported by the facility to the EPA’s Toxic Release Inventory (TRI) system. An insignificant source/activity can become significant and would be identified as such on the facility permit. Some insignificant activities are aggregated into one unit/process and reported to EIS using a Source Classification Code which describes the group.

- i. **Recommendation** – The CEF must have the ability for individual SLTs to designate which data (facilities, units, processes, and pollutants) are sent to EPA and which are to be held in the SLT system only.

#### *Facility and Point Data Field Matching Exercise*

Iowa, Massachusetts, North Carolina, and South Carolina volunteered to do data field matching between their state systems, the new FRS data model and the EIS. This matching included definition, data type, data size and whether a data field was required to be reported by a filer. These volunteers were also asked to include data fields which exist in their systems, but do not exist in EIS or FRS. While this sampling was small, the states represented on Data Model Team had varying degrees of inventory system sophistication. Those systems spanned between paper submissions to highly integrated. All team members added input and comments to recommended additions.

- ii. **Recommendation** – Appendix 1, CEF Data Fields.xlsx, shows the result of the comparison and recommendations for additional data fields. Facility data field recommendations have been passed onto the FRS Data Model Team with additions noted. The emissions portion of Appendix 1 shows SLT recommended additions that the team is recommending be added to EIS as noted in the comment fields. Also noted are those data fields that are recommended to be included in the CAER Pilot. Please refer to the READ ME worksheet for definitions of column headings and color coding.

#### Permit Data Fields

A large portion of the data fields that were recommended by SLTs fell under the purview of permitting. These data fields are listed under the tab “Permitted” in Appendix 1.

Since E-Enterprise has kicked-off the E-Permitting project, it was felt that these particular data fields would be forwarded to this team for review and inclusion under the E-Permitting project. As the CAER project moves forward, these fields may be determined to be a necessary part of CAER, but at this time the team did not feel that such a recommendation was advisable.

#### *Calculator Requirements*

In creating the list of required data fields for the CEF, the team started discussions about how these data fields could be used to estimate of emissions. The overall CAER working design included emissions estimation as one of the functions of the emissions “Calculator” needed for the project. These discussions were lengthy, and while defining specifications for the Calculator was not one of our tasks for this Phase, we felt that the results of the discussions should be captured. Appendix 2. Common Emissions Form (CEF) Calculator Requirements documents those discussions and our recommendations.