NEI Augmentation Subgroup Recommendations Report

Introduction

In 2014, E-Enterprise for the Environment (E-Enterprise) identified five initial scoping projects that highlighted important state-federal data partnerships. One of these projects was the Combined Air Emissions Reporting (CAER) project which focused on how air reporting could be streamlined across programs. During the scoping phase, the CAER project identified a subset of incremental steps required to streamline data reporting, including data augmentation within the National Emission Inventory (NEI).

The NEI is used by the Environmental Protection Agency (EPA) to support evaluating the National Ambient Air Quality Standards (NAAQS), assessing interstate transport of air pollutants, air toxics programs, and for international reporting. It is also used by state, local, and tribal (SLT) air agencies as a starting point for their State Implementation Plan (SIP) development and by other federal agencies, researchers, and environmental groups to understand sources of air pollution. Because the NEI is the basis of so much work, it is important that the inventory be as complete and accurate as possible. If EPA determines that the data they receive from SLTs are not sufficient to create a complete NEI, EPA undertakes a process to augment the submitted data.

One inefficiency of the NEI program, identified in the CAER scoping project, is the time it takes for EPA staff to perform downstream augmentation of SLT submitted data. A select group of SLT air agencies worked with EPA as part of the NEI Augmentation team to define all data criteria which SLT submittals must meet. This group set out to minimize, and hopefully eliminate, the need for EPA augmentation of submitted data.

During this process the group focused on determining (1) why EPA augments SLT data, (2) why SLTs may not be able to provide EPA with the data needed, (3) how EPA can improve, reduce, or streamline data augmentation until it can be reduced or eliminated, and (4) what is needed to assist SLTs in providing a more complete NEI submission.

The key objectives of this project were:

- To have a better understanding of why EPA augments SLT data and why SLTs are unable to provide full NEI datasets, eliminating those obstacles where possible.
- To create a tool or reference which will minimize or eliminate the need for EPA augmentation in the targeted agencies for the current and future NEIs.
- To identify changes to EPA's augmentation where it can be improved or eliminated. EPA would implement those changes for the current and future NEI (in house) where possible.

Section 1: Background and Initial Fact-finding

In 2015, the State of Arizona surveyed SLTs to get a sense of how emissions are collected and whether SLTs augmented facility submitted data. The following questions were asked:

1. How are data submitted by SLTs to the Emission Inventory System (EIS)?
2. What software do you use to manage your point source data?
3. Do you augment data received by industry?
4. Is there a reason why SLTs cannot augment industry data?
5. Is your reporting data mandated by regulation?
6. What is your ability to change regulations regarding reporting data?
7. When are facility data required to be reported to SLTs?
8. Who makes Source Classification Code (SCC) assignments used in the inventory?
9. Who calculates emissions used in the inventory?

52% of SLTs use either Excel-based or SLT developed software to manage point source emission data, and only 13% have an integrated e-permitting system (i.e., integration of e-file permit application, e-tracking, and e-pay with emissions inventory system). 11% of SLTs use the State and Local Emissions Inventory System (SLEIS) to submit data; nearly 70% use the EIS' Bridge Tool. SLTs indicated that they augment reported data received from industry only 30% of the time, and, when SLTs do, two-thirds of the data augmented was PM. 54% of SLTs indicated that there were no obstacles to them augmenting data received from industry, and reasons for obstacles cited by the remainder of respondents included legal challenges, lack of resources, State policy, technological limitations and conflicting fee structures. Appendix A outlines the survey in more detail including the questions asked, responses received, and graphical summaries of those responses.

Section 2: A Search for Solutions

In order to search for solutions, the team needed to first have a common understanding of why and how EPA augments SLT submitted data. EPA presented a webinar to the team to explain why augmentation is conducted, the methods used to calculate the values added, how SLTs can identify the augmented data through process level reports, and how SLTs can remove or change augmented data.

EPA requires a complete inventory in order to satisfy various data customers. The NEI is used as the foundation for many projects and is dependent on the data being as complete as possible. Augmentation is filling gaps and not replacing SLT submitted data. SLT data such as PM and VOC are used as bounding inputs to drive PM and HAP augmentation.

The Air Emissions Reporting Rule (AERR) is the regulatory requirement that drives data collection for the NEI. The AERR states;

> *(1) Required pollutants for triennial reports of annual (12-month) emissions for all sources and every-year reports of annual emissions from Type A sources:*
>
> *(i) Sulfur dioxide (SO_2).*
>
> *(ii) Volatile organic compounds (VOC).*

(iii) Nitrogen oxides (NO_X).
(iv) Carbon monoxide (CO).
(v) Lead and lead compounds.
(vi) Primary $PM_{2.5}$. As applicable, also report filterable and condensable components.
(vii) Primary PM_{10}. As applicable, also report filterable and condensable components.
(viii) Ammonia (NH_3).

EPA interprets "as applicable" to mean where it is emitted, and not if it is available.

During this presentation, EPA discussed how the draft NEI is created through a selection hierarchy of SLT and EPA datasets. EPA prefers to use SLT submitted data first whenever possible in the selection process. The only reason why EPA would not select SLT data first is if the SLT value is an outlier and the SLT has not corrected the data, or if EPA has test data which is deemed to be of higher quality.

The PM Tool (available on CHIEF) and HAP Augmentation Tool (available in EIS) were reviewed. Very few of the SLTs had used or even looked at these tools previously. It was noted that all of the SLTs felt that knowing how to use these tools would be beneficial. EPA also demonstrated how SLTs could review what data had been augmented in the selection during their comment period by using the EIS process level report and data set name.

The demonstration had immediate benefits:

1. SLT team members had a better understanding of the PM requirements in the AERR;
2. SLT team members understood that augmentation was not a replacement of SLT data but only to fill gaps not reported by SLTs;
3. That PM and HAP augmentation were based on SLT submitted PM and VOC. Chrome speciation was based on SLT submitted total chrome.
4. That an understanding of the PM Tool and the HAP Augmentation Tool could be helpful for SLTs in either augmenting their own data or at least understanding how EPA calculated augmented data;
5. That training was important prior to the SLT comment period of the selected draft NEI to understand how to identify augmented data and also how to comment on the removal or change of augmented data.

Also, it is important to note that Toxic Release Inventory (TRI) augmentation was not significantly reviewed as part of this effort, however several known outstanding issues were discussed and have been identified for the full CAER team to address in the future.

Section 3: Group Recommendations

The NEI Augmentation team met several times and provided time for SLT team members to discuss various challenges with data collection, local regulations, and other potential barriers to providing full NEI datasets to EPA. The following recommendations were developed from those discussions and the survey data collected from other SLTs.

The purpose of data augmentation is not to alter reported data, but rather to fill in perceived information gaps to complete the emissions inventory as required by the AERR and/or as desired by data users. Eliminating augmentation of data submitted to EPA is unlikely in the near-term, but we could make headway on reduction by focusing efforts on the following areas:

1. Encourage submittal of complete datasets by facilities
2. Provide SLTs training on NEI process and augmentation
3. Provide SLTs tools to fill data gaps themselves
4. Create a more open and collaborative EPA augmentation process

Encourage Initial Submittal of Complete Datasets by Facilities

The NEI originates with datasets provided by regulated facilities to SLT agencies. As such, the best way to avoid augmenting these datasets is to ensure that complete datasets are initially collected from the facilities. To this end, the team recommends the following:

- EPA will provide and promote an **Expected Pollutant** list in EIS, as well as on EPA's public CHIEF website. This list will define what pollutants are expected to be reported by Source Classification Code (SCC), and provide clarity to SLTs on what data are expected to be reported to EPA. It will also identify what pollutants will be augmented if data are not submitted.
- SLTs are encouraged to share the Expected Pollutant list with their facilities as a reference tool.
- Incorporate into the EIS the ability to inform SLTs which pollutants are missing from their submitted data (i.e. as a report of missing pollutants). This draft report has been developed but has not been loaded to production in EIS. A sample of this reported can be found in Appendix B.

Provide SLTs Training on NEI Process and Augmentation

Team discussions exposed inconsistent understanding by SLTs regarding the NEI process, particularly the role of augmentation and how they could participate. Neither the EPA nor SLTs were aware of this knowledge disconnect. To address this, the team recommends the following:

- EPA should provide clarity on their augmentation process through recorded webinar training (e.g., training on how to use PM calculator, etc.).
- EPA should provide documentation and instructions on the augmentation tools available for SLT use on the CHIEF website.
- EPA should provide an EIS report which will compare the expected pollutant list to the SLT submitted data.
- EPA should provide clear guidance as to how the NEI draft should be reviewed. The NEI draft is the selection created for SLT comments prior to publication to the public.
    - Identify where data have been augmented;
    - Comment or request adjustment on an incorrect augmentation.

Provide SLTs Tools to Fill Data Gaps Themselves before Submitting Data to EPA

Through team discussions it became apparent that there are many potential obstacles to SLTs collecting complete NEI datasets from facilities. These include local regulations that dictate how/what data are

collected, lack of facility expertise in data submittal, and limited SLT resources. To work towards resolving these issues, the team recommends:

- If complete NEI datasets cannot, for whatever reason, be provided directly by regulated facilities SLT agencies may augment their data, using the EPA augmentation tools, so as to include everything from the **Expected Pollutant** list.
    - If SLT regulation or policies prevent the collection of everything from the **Expected Pollutant** list, these should be identified. Pollutants unable to be submitted will be augmented.
- EPA should provide the PM Augmentation Tool and HAP Augmentation Tool to SLT agencies, including:
    - Tool names, descriptions, technical documentation, and download location
    - Use cases
        - SLTs are interested in augmenting their own data in some situations – these tools will help with that process
    - Training plans: webinars, recorded materials, written material
    - Resources may be allocated toward updating tools so that they can be used by states, and address any needed training

Create a More Open and Collaborative EPA Augmentation Process

Because existing SLT regulations and policies are many and varied, it's likely that some SLTs will not be able to provide EPA with complete NEI datasets (i.e., everything from the **Expected Pollutants** list). To address these scenarios, the team recommends:

- If EPA does not receive datasets that contain all of the **Expected Pollutants**, they will perform augmentation to fill in the missing data.
- EPA will provide draft versions of the NEI to include augmented datasets for SLT review.
- SLTs will be encouraged to participate in the review of these augmented datasets
    - SLTs should use this as an opportunity to amend their data collection processes to reduce or eliminate areas where EPA is augmenting data.
- EPA and SLTs should work together to identify those local regulations or policies that prevent SLTs from either collecting everything from the **Expected Pollutants** list or augmenting data once it's collected. If common gaps are identified across states, EPA and SLTs should work nationally to identify a path forward.

Additional Considerations

Team members highlighted a number of other pertinent considerations for the CAER team in the context of reporting procedures. Additional considerations could be candidates for additional training on EPA tools and technical resources. In Appendix C, these additional ideas have been broken up into two themes: (1) areas where the team expressed specific points of confusion to be addressed between SLTS and EPA - perhaps by way of training; and (2) specific points to be addressed in the context of state-to-state coordination.

Section 4: Next steps

The recommendations in this report have been reviewed by the CAER team and circulated to other SLT air agency staff via ECOS Wire and in coordination with the National Association of Clean Air Agencies and other groups. Comments received during this review process will be considered by the NEI Augmentation team and incorporated as appropriate.

The recommendations in this report are already being implemented in the scheduled webinars on how to review the 2014 v1 NEI draft during the first week in June, 2016. These webinars will include how to identify and comment on augmented data.

This project received $20,000 to develop a tool or resource to aid in reducing augmentation of SLT data. These funds are being used towards (1) the development of a report to identify missing expected pollutants; and, (2) to scope the integration of the PM Calculator into EIS for use by SLT in calculating PM-FIL and PM-CON portions of PM-PRI when developing their inventories.

The NEI Augmentation team determined that the overall quality assurance and quality control (QA/QC) process requires more discussion between SLT and EPA staff. The CAER team also came to this conclusion during a two-day workshop. This topic has been forwarded to the CAER Implementation Team, thus specific recommendations on changes to the QA/QC process are not included in this report.

Conclusion

The NEI Augmentation team's work showcased the enhanced clarity that collaboration and shared learning between peers at SLTs and EPA can bring to the nuanced and at times opaque challenges of shared information. There was broad agreement that most SLTs were comfortable with EPA augmenting their data where necessary. Their concerns lie with a lack of mutual understanding about what was being augmented, when it was taking place, and how they could find out what areas had been augmented after augmentation had taken place. The recommendations developed by this team will continue to inform the CAER effort as they move forward with the final implementation report.

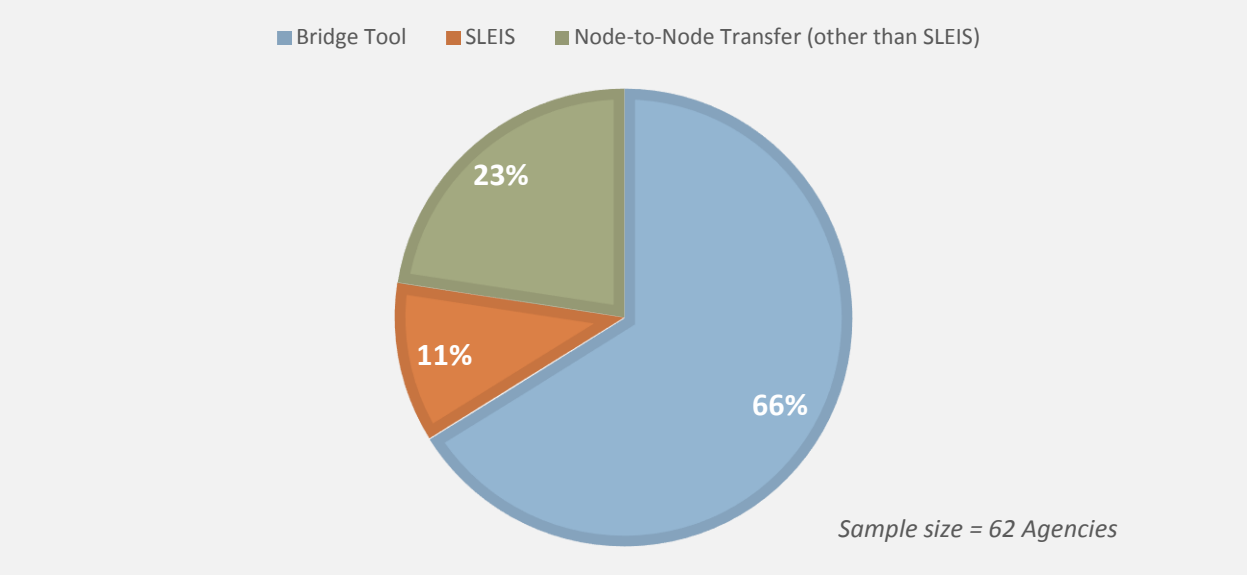
Appendix A: *SLT Survey Response Summary*

In spring 2015, Arizona circulated a survey to 62 organizations from states, small and regional municipalities and tribes on how they collect and report data on air emissions to EIS. The following questions and accompanying graphics were developed by the team using survey results. They provided helpful context and are presented here as a more detailed review of the issues at hand. Appendix D contains a list of organizations that participated in this survey.

List of Graphics

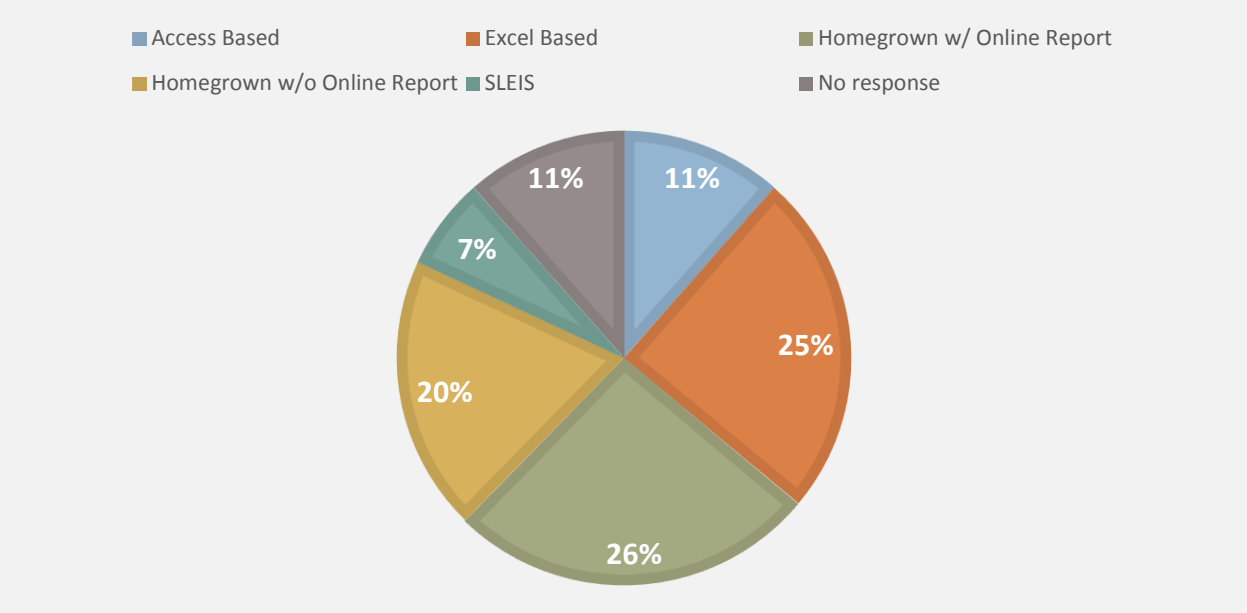
1. How is data submitted by State/Local Agency to EIS?
2. What software do you use to manage point source emission data?
3. Is there any reason (regulation, office, policy, politics, etc.) why you cannot augment data received from industry?
4. Do you augment data received from industry?
5. Is your reporting date mandated by regulation?
6. What is your ability to change regulations regarding reporting data?
7. When is reporting required?
8. Who makes source classification code assignments used in inventory?
9. Who calculates emissions used in inventory?

Chart 1: How is data submitted by State/Local Agency to EIS?



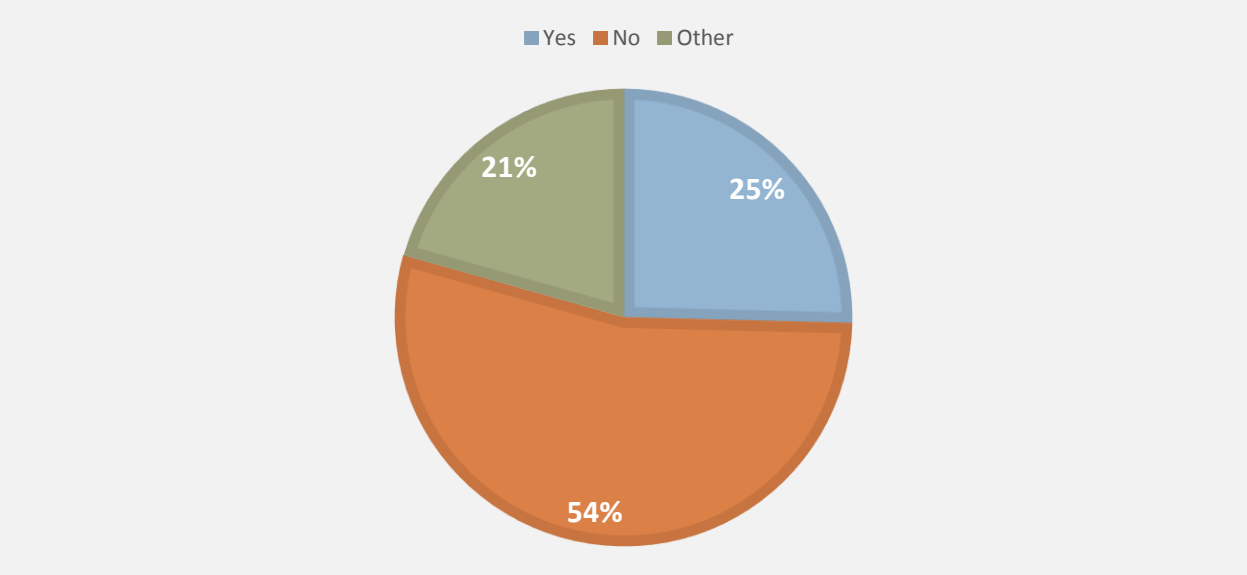
Nearly 70% of respondents submit data using a bridge tool, 11% use SLEIS (State and Local Emissions Inventory System) and a further 23% use a non-SLEIS node-to-node transfer.

Chart 2: What software do you use to manage point source emission data?



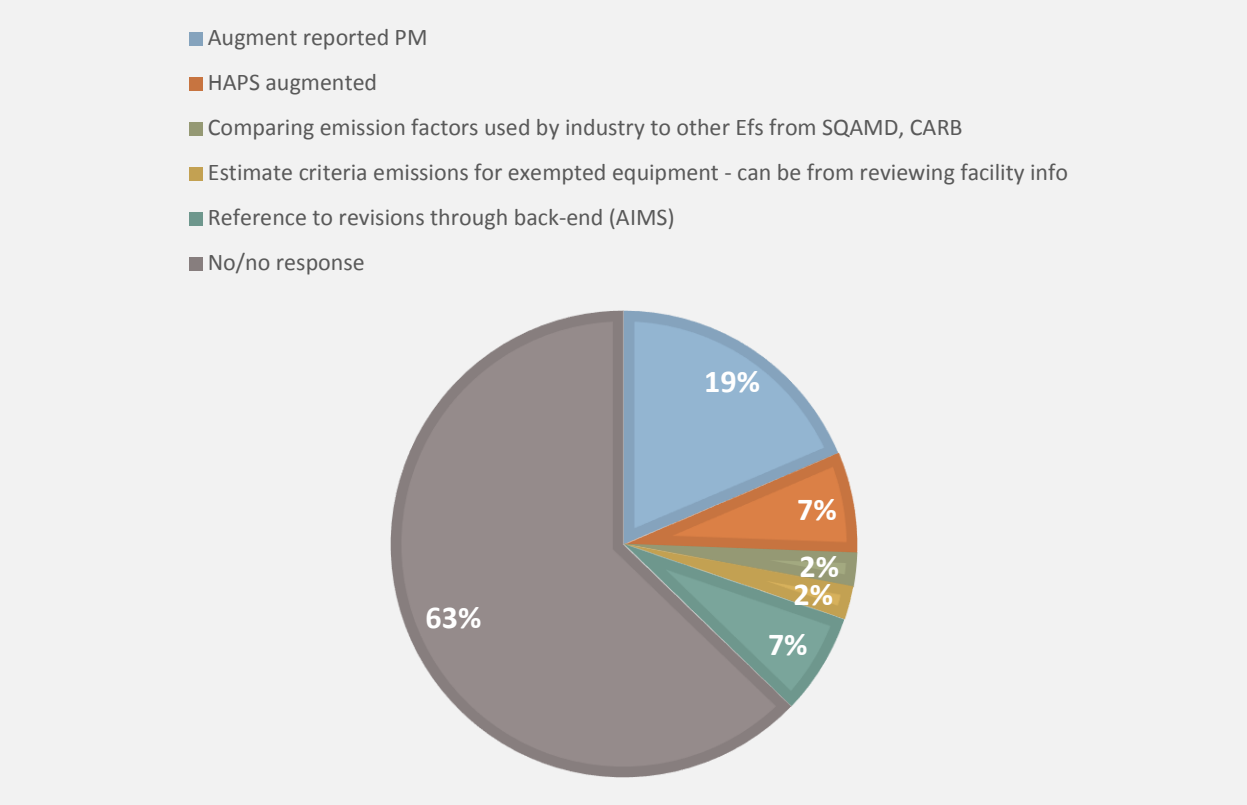
46% of survey respondents use homegrown software to manage point source emission data. The next largest group (20%) use an Excel-based system, while only 7% percent used SLEIS and 11% used an access-based software.

Chart 3: Is there any reason (regulation, office, policy, politics, etc.) you cannot augment data?



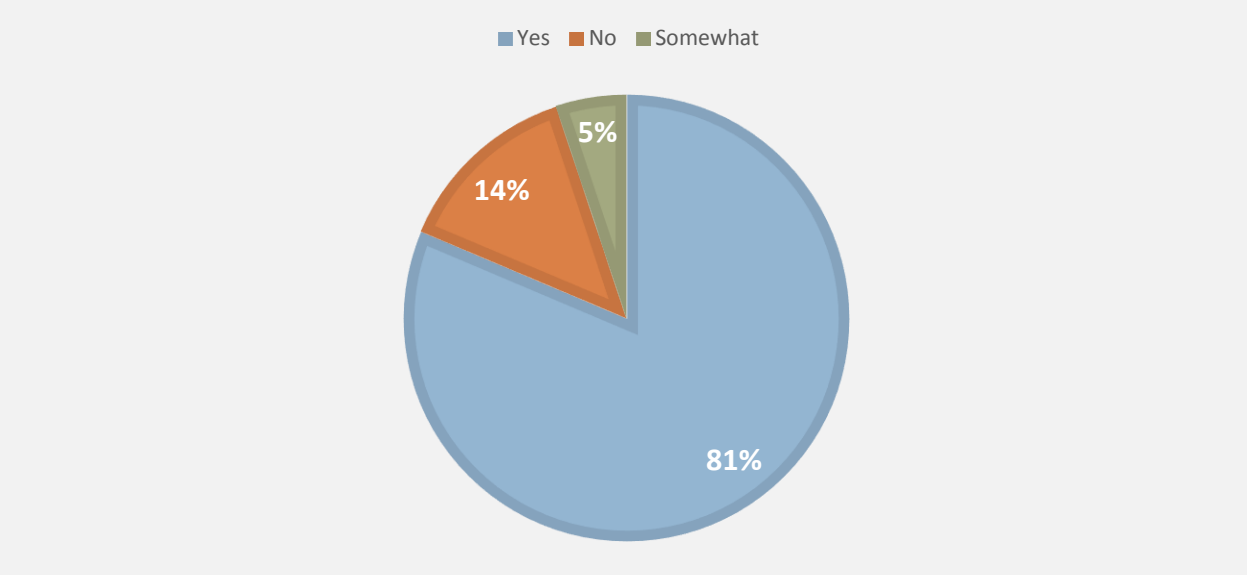
54% of survey respondents stated there is no reason they can't augment data they receive from industry. The remaining 46% indicated that there was either a complication (21%) or a discrete reason they could not augment industry data (25%).

Chart 4: Do you augment data received from industry?



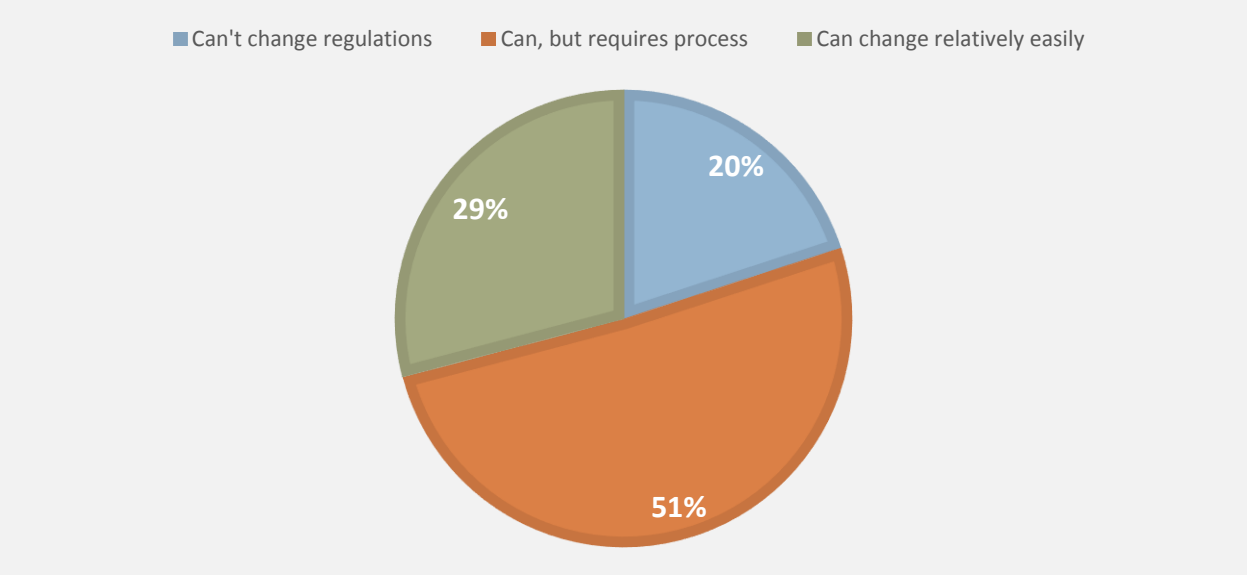
63% of those surveyed either indicated that they do not augment data received from industry or did not provide an answer to the question. 19% of all respondents augment reported PM, 7% augment HAPS and 7% reference “back-end” revisions through AIMS.

Chart 5: Is your reporting date mandated by Regulation?



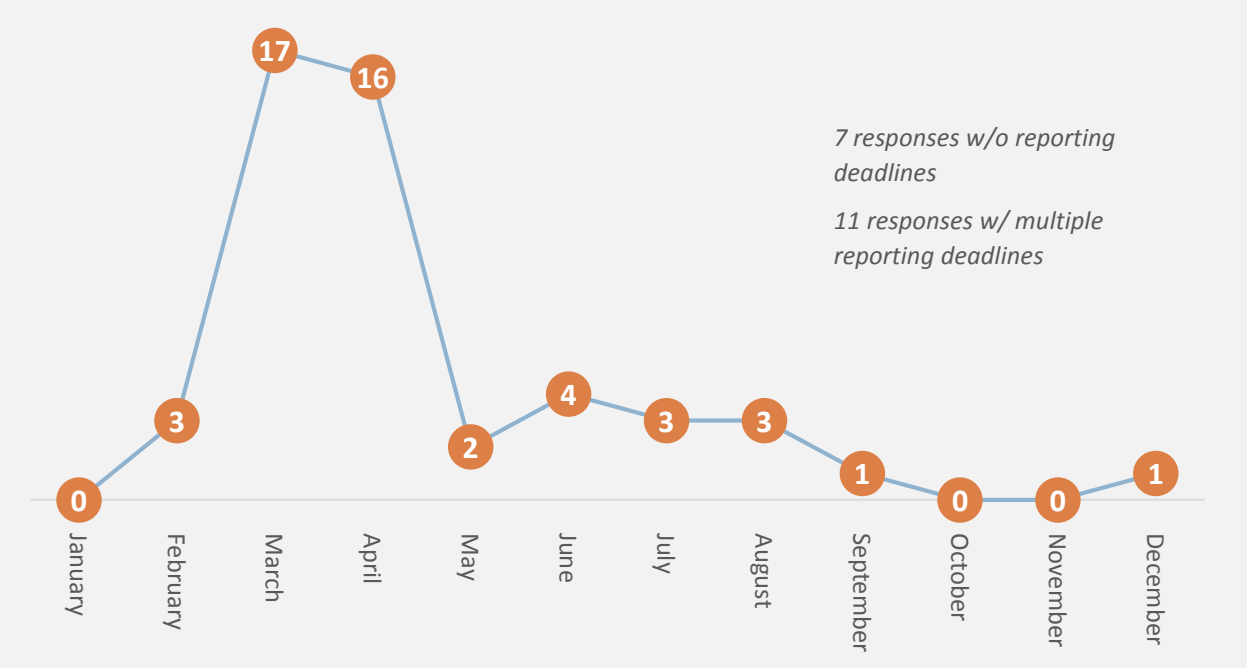
In order to gauge flexibility of reporting date, survey participants were asked whether their reporting date was mandated by regulation. 81% indicated it was, 14% noted it wasn't and 5% described specific complications involved in changing the reporting date.

Chart 6: What is your ability to change regulations regarding reporting data?



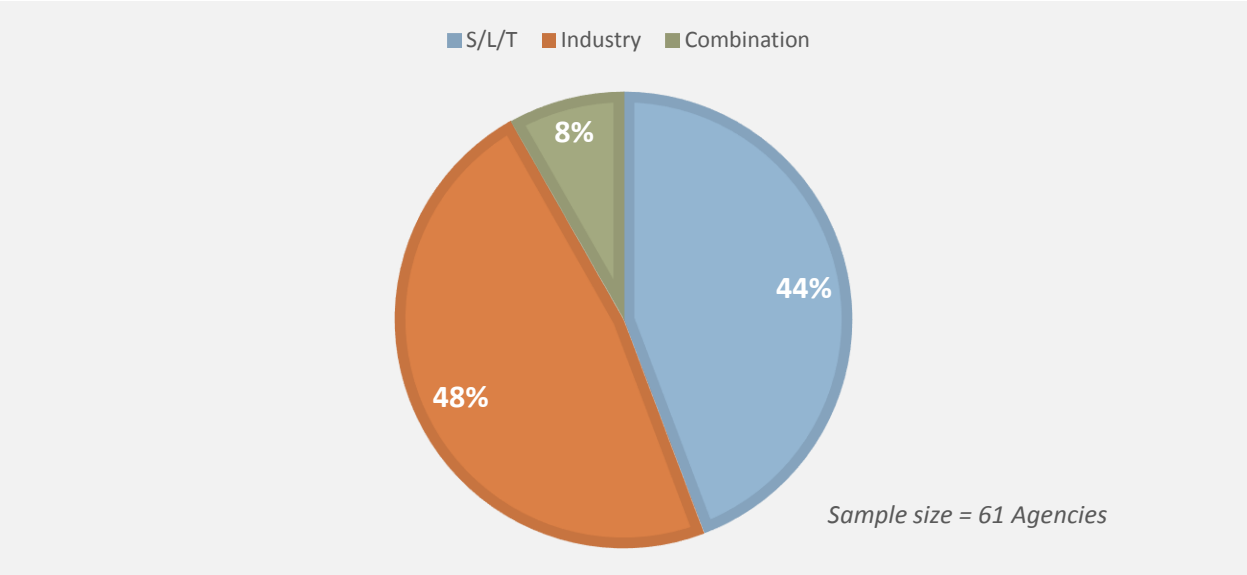
51% of respondents indicated that they could change their reporting date, but it would require a process, legislative or otherwise, to do so. 29% could change their regulations regarding reporting date relatively easily, but 20% would be unable to make a change.

Chart 7: When is reporting required?



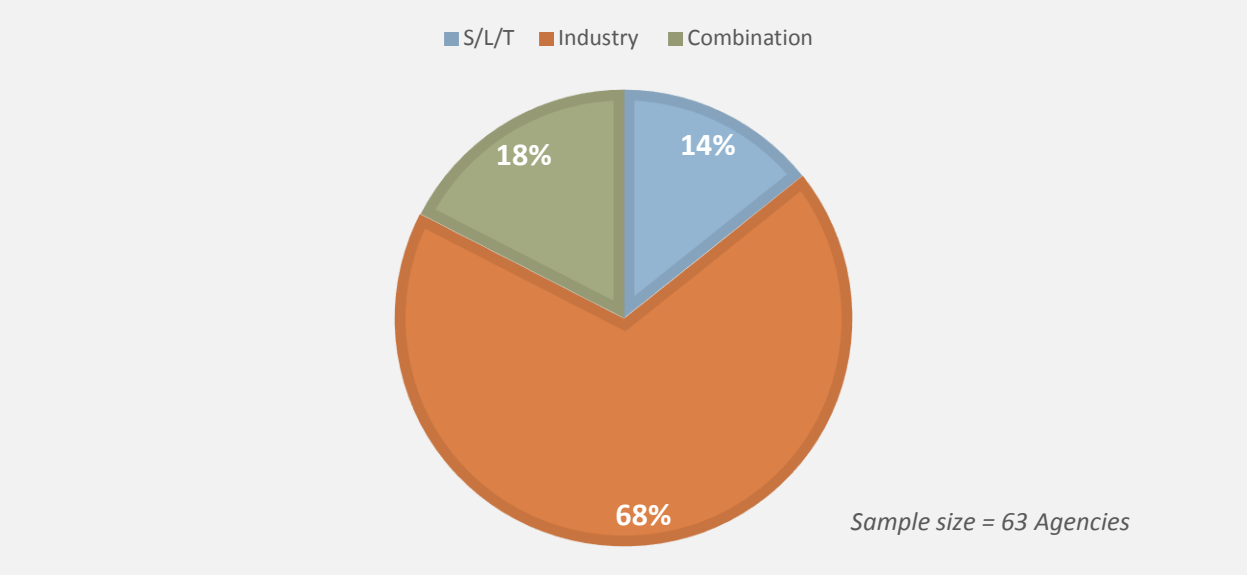
This chart indicates when facility data is due to SLTs in order to meet the reporting requirement to EIS. March and April were by far the most popular months. It is also important to note that 11 respondents have multiple reporting deadlines and are not represented in this chart.

Chart 8: Who makes SCC assignments used in inventory?



48% of source classification code (SCC) assignments used in the inventory were made by industry, 44% were made by SLTs, and 8% of assignments were made by a combination of industry and SLTs.

Chart 9: Who calculates emissions used in inventory?



In 68% of cases industry calculates emissions. SLTs calculate emissions 14% of the time, and SLTs and industry share emissions calculations 18% of the time.

Appendix B: *Sample Expected Pollutant/SLT Submitted Data Comparison Report*

EIS Facility	Facility Name	PSC	Agency Facility ID	EIS Unit ID	Agency Unit ID	EIS Process ID	Agency Process ID	SCC	Missing Pollutant Code	Pollutant Description
536011	Emory University	GADNR	08900233	4842713	924	4606914	1	10300601	110543	Hexane
536011	Emory University	GADNR	08900233	4842713	924	4606914	1	10300601	CO	Carbon Monoxide
536011	Emory University	GADNR	08900233	4842713	924	4606914	1	10300601	VOC	Volatile Organic Compounds
536011	Emory University	GADNR	08900233	4842713	924	46062014	2	10300501	CO	Carbon Monoxide

A draft report has been developed to aid SLTs in identifying which if any pollutants may be missing at the process level. This comparison report will compare the published expected pollutant list to the SLT submitted data. A sample of the resulting comparison report is above.

In this sample, the report has identified that for this given facility/unit, process 1 and 2 are both missing CO. While process 1 is also missing Hexane and VOC. This report could not only be used as a tool to identify what is missing, but also to identify what may be augmented. This report would also include the SCC level descriptions but have been excluded in this example.

Appendix C: Future Considerations

EPA and States Coordination Clarification

- Not all state facilities are required to report species of PM according to their permits.
- Some states lack the regulatory authority to collect HAP and PM from their facilities. (The AERR does specify the requirement to report all species of PM).
- States are concerned about the potential for changes to facility fee structures (based on data requirements). This may need additional discussions and/or communication about current business rules.
- Possible web services, similar to short-term win on SCC web services, which would allow search, download, and report comments on expected pollutants.
- Address facilities comments that augmentation creates inconsistencies with other EPA reported data which may or may not be reported to SLTs.
- Improving augmentation through:
    - Directing reporting by facilities;
    - Test data collection via CEDRI to improve emission factors and in turn augmentation ratios;
    - Use of facility or process specific augmentation factors, if available.

State to State Coordination

- Facilities may not have the expertise to handle more refined PM and/or HAP data – data collection can benefit from state or local agency assistance and examples of best practices from other states.
- Facilities that lack site-specific data for one or all of their processes may benefit from coordinating with similar facilities in other states that do have site-specific data.
- The augmentation tools used by SLTs and/or EPA may introduce inaccuracy into the data.
    - Approximate methods (via augmentation) are not a substitute for test data, especially when data are being used for human health measurements/data collection/decision making.

Appendix D

The following organizations participated in the SLT survey.

- Alabama Department of Environmental Management
- Alaska Division of Air Quality
- Allegheny County Health Department
- Arkansas Department of Environmental Quality
- California Air Resources Board
- City of Omaha
- Clark County, Nevada
- Colorado Department of Public Health and Environment
- Connecticut Department of Energy and Environmental Protection
- District of Columbia Department of Energy and Environment
- Delaware Department of Natural Resources and Environmental Control - Division of Air Quality
- Forsyth County Office of Environmental Assistance and Protection
- Idaho Department of Environmental Quality
- Iowa Department of Natural Resources - Air Quality Bureau
- Jefferson County Department of Health
- Knox County Air Quality Management
- Lane Regional Air Protection Agency
- Louisiana Department of Environmental Quality
- Louisville Metro Air Pollution Control District
- Maine Department of Environmental Protection
- Maricopa County Air Quality Department
- Maryland Department of the Environment
- Massachusetts Department of Environmental Protection
- Mecklenburg County Air Quality
- Nashville Metro Public Health Department
- Michigan Department of Environmental Quality – Air Quality Division
- Minnesota Pollution Control Agency
- Mississippi Department of Environmental Quality
- Montana Department of Environmental Quality
- Missouri Department of Environmental Quality
- North Carolina Department of Environment and Natural Resources – Division of Air Quality
- Nebraska Department of Environmental Quality
- New Jersey Department of Environmental Protection
- North Dakota Department of Health
- Ohio Environmental Protection Agency
- Oklahoma Department of Environmental Quality
- Olympic Region Clean Air Agency
- Oregon Department of Environmental Quality

- Pennsylvania Department of Environmental Protection
- Pima County Department of Environmental Quality
- Pinal County Air Quality Control District
- Puerto Rico Environmental Quality Board
- Puget Sound Clean Air Agency
- Rhode Island Department of Environmental Management
- South Carolina Department of Health and Environmental Compliance
- State of Georgia Environmental Protection Division - Air Protection Branch
- State of Utah, Department of Environmental Quality - Division of Air Quality
- Texas Commission on Environmental Quality
- Tennessee Department of Environment and Conservation
- Ventura County California, Air Pollution Control District
- Virginia Department of Environmental Quality
- Vermont Department of Environmental Conservation
- Washoe County Health District, Air Quality Management Division
- Western North Carolina Regional Air Quality Agency
- Wisconsin Department of Natural Resources
- West Virginia Department of Environmental Protection
- Wyoming Department of Environmental Quality - Air Quality Division