

1999  
National Emission Inventory  
Preparation Plan  
- Revised (February 2001)-

(page left intentionally blank)

## **1.0 INTRODUCTION**

### **1.1 Who Is the Intended Audience for This Document?**

This document was prepared primarily as an internal planning tool for EPA's Emission Factor and Inventory Group (EFIG) in preparation of the 1999 National Emission Inventory (NEI) for criteria and hazardous air pollutants (HAPs). However, it also serves to inform our State and local agency partners (S/Ls) of our plans for the 1999 NEI preparation and the opportunities for providing input to, and for reviewing, the NEI. EFIG also hopes to work with Tribal representatives in the future to improve emission inventory data for Indian lands.

### **1.2 How Will This Document Help Me?**

This document describes the procedures and the schedule that EFIG will follow in order to incorporate the emissions data received from a variety of sources into a comprehensive national inventory for both criteria and HAPs for the year 1999. The primary suppliers of emissions data to EFIG will continue to be the S/Ls. EFIG will also use data from other sources, such as EPA's Emission Standards Division (ESD), Clean Air Markets Division (CAMD), Office of Transportation and Air Quality (OTAQ) and Toxic Release Inventory (TRI) program. This document describes when and how the various data sources will be used to produce a complete inventory and the quality checks that EFIG will perform on that inventory. This document also describes when and how the data should be initially submitted, reviewed by submitters and other stakeholders, and subsequently corrected by S/Ls in order to produce an accurate inventory.

### **1.3 Does This Document Establish Reporting Requirements?**

No. This document is provided to our regulatory partners as information only, so that interested parties who may be supplying or using the inventory data can better understand EFIG's goals and objectives, schedule, plans for filling in missing data, and quality checking procedures. By documenting the procedures and the schedule to be followed, EFIG hopes to provide the S/Ls, EPA's Regional Offices and Emission Standards Division, and other stakeholders with the opportunity to participate in the process in the most timely and effective manner. We envision that a complete and accurate national inventory suitable for regional scale modeling can be a very cost-effective tool for all of the data suppliers and stakeholders, as well as EPA and the regional planning organizations (RPO's).

The Clean Air Act and implementing regulations at 40 CFR contain legally binding requirements, and this document does not substitute for those provisions or regulations, nor is it a regulation itself. Thus, it does not impose binding, enforceable inventory reporting requirements on any party and may not apply to a particular situation based upon the circumstances. EPA decisionmakers retain the discretion to adopt approaches on a case-by-case basis that differ from this document where appropriate. Therefore, interested parties are free to raise questions and objections about the appropriateness of the application of this document to a particular situation. Moreover, although planned submittal dates and review periods in this document were developed to be in accord with the proposed Consolidated Emissions Reporting Rule (CERR) that was

published on May 23, 2000 (65 FR 33268), it is not intended to describe the reporting requirements of the proposed CERR.<sup>1, 2, 3</sup>

#### **1.4 How Is This Document Organized?**

Section 2.0 of this document describes the intended uses and data quality objectives of the inventory. Section 3.0 describes the data sources that EPA would prefer to use for each sector of the inventory and what we will use in lieu of these preferred data. This section therefore highlights the areas where S/L inputs would make the largest improvements in the quality of the inventory. Section 4.0 describes the steps that EFIG will follow to collect the data, fill-in data gaps, check the data quality, supply the data for review, and finally publish the data. This section also provides the dates for each step, and therefore highlights the times when S/Ls can review and correct data for their jurisdiction.

#### **1.5 What Are the Highlights of This Document?**

Readers should note the following highlights, described further in this document:

- Timing for Inventory Preparation - Time allotted for each step of the process is ambitiously short, in order to produce an inventory in a useful time frame.
- Iterative Process - The 1999 NEI will be updated periodically to provide more accurate versions of the inventory over time. Scheduled completion dates are:
  - < Preliminary 1999 National Emission Inventory (Version 1) - Dec 1, 2000
  - < 1999 National Emission Inventory (Version 2) - June 1, 2002
  - < 1999 National Emission Inventory (Version 3) - June 1, 2003
- Integrating Inventories - EFIG will be handling the HAPs and criteria emission inventories in a more integrated fashion this year. There are valid reasons for continuing to handle some aspects of these two pollutant classes individually, but there are also many advantages to bringing these two previously separate inventories together. We would like to encourage the data submitters to help us in realizing these advantages. One benefit of integrating the HAPs and criteria inventories would be a common set of facility and emission point identifiers.

---

<sup>1</sup>For information on the CERR, go to [www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/) and under Emission Inventories, click on Consolidated Emissions Reporting Rule.

<sup>2</sup>For information on criteria pollutant emission inventory reporting requirements, go to [www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/) and under Emission Inventories, click on Emission Inventory Publications, Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter NAAQS and Regional Haze Regulations.

<sup>3</sup>For information on the NEI Input Format, go to [www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/) and under Emission Inventories, click on Emission Inventory Data, Submitting Data to EPA.

## 2.0 USES OF THE INVENTORY

### 2.1 How Will the National Emission Inventory Be Used?

The 1999 National Emission Inventory will include both criteria pollutants and their precursors (SO<sub>x</sub>, VOC, NO<sub>x</sub>, CO, PM<sub>10</sub>, PM<sub>2.5</sub>, and NH<sub>3</sub>), and HAPs (188). The 1996 NET and NTI continue to be the basis for numerous modeling efforts including regional and local scale air quality modeling, as well as human exposure modeling. Examples of these uses include: the National Air Toxics Assessment (NATA), EPA's regulatory analysis and implementation of TIER 2 standards for automobile engines, and the regulatory analysis for the Heavy Duty Engine and Diesel Fuel Sulfur rule.

Upon completion, the 1999 National Emission Inventory is expected to be used for the following:

- Initial inventory for EPA regional and local scale modeling efforts to predict ambient concentrations, exposures, and the resultant risks to human health and the environment such as planned future National Air Toxics Assessments;
- Initial inventory for S/L and regional planning organization modeling efforts;
- Emission estimates for the EPA Air Quality and Emission Trends report; and
- Use for other EPA, S/L, and public analyses requiring emission inventories. Some examples include: Implementation of the following areas of the 1990 Clean Air Act Amendments - Section 112(f), Residual Risk, and Section 112(c)(3), Section 112(k), Urban Area Source Program, the Industrial SO<sub>2</sub> Report to Congress, published every 5 years; and tracking progress toward program goals of the federal Government Performance and Results Act (GPRA).

To support these uses, the National Emission Inventory must be comprehensive, covering all criteria pollutants and HAPs for all areas of the United States. It must also cover all significant emission sources, including all stationary and mobile sources. We are attempting to harmonize the nomenclature for HAP and criteria pollutant source categories for inventory purposes. Regardless of the terminology, the inventories must include the following types of sources:

<b>Category</b>	<b>Cutoff</b>
<i>HAPs</i>	
Major	Stationary sources with potential to emit $\geq 10$ tons/yr of any HAP or 25 tons/yr of more than one HAP <sup>4</sup>
Area/Other	Nonmajor stationary sources
Onroad	Mobile sources licensed for use on highways or roadways
Nonroad	Other mobile sources (construction, lawn/garden, boats, trains, airplanes, etc.)

---

<sup>4</sup>As defined in Section 112(a) of the Clean Air Act.

Category	Cutoff
<i>Criteria</i>	
Point	Stationary sources with actual emissions >100 tons/yr of any criteria pollutant, and smaller sources wherever reported on an individual facility basis
Area	All other stationary sources
Onroad	Mobile sources licensed for use on highways or roadways
Nonroad	Other mobile sources (construction, lawn/garden, boats, trains, airplanes, etc.)

## 2.2 What Versions of the National Emission Inventory Will Be Used to Produce the National Air Quality and Emissions Trends Reports?

The following list describes the versions of the national inventory that will be used as the basis for the annual Trends Brochures and Reports for the years 2000 through 2004:

Brochure/ Report Date	Data Due Date*	Yr of Data**	Inventory Version
Jul/Dec 2000	Jun/Nov 2000	1999	99 Preliminary National Emission Inventory (V1) - with 1999 CEM Utility Data (SO <sub>2</sub> and NO <sub>x</sub> ) - with 1999 Onroad & Nonroad emissions from OTAQ
Jul/Nov 2001	Jun/Aug 2001	2000	99 Preliminary National Emission Inventory (V1) - emissions projected to 2000 - with 2000 CEM Utility Data (SO <sub>2</sub> and NO <sub>x</sub> ) - with 2000 Onroad/Nonroad emissions from OTAQ
Jul/Nov 2002	Jun/Aug 2002	2001	99 National Emission Inventory (V2) - emissions projected to 2001 - with 2001 CEM Utility Data (SO <sub>2</sub> and NO <sub>x</sub> ) - with 2001 Onroad/Nonroad emissions from OTAQ
Jul/Nov 2003	Jun/Aug 2003	2002	02 Preliminary National Emission Inventory (V1) - emissions projected to 2002 (from 99NEI V3) - with 2002 CEM Utility Data (SO <sub>2</sub> and NO <sub>x</sub> ) - with 2002 Onroad/Nonroad emissions from OTAQ
Jul/Nov 2004	Jun/Aug 2004	2003	02 Preliminary National Emission Inventory (V1) - emissions projected to 2003 - with 2003 CEM Utility Data (SO <sub>2</sub> and NO <sub>x</sub> ) - with 2003 Onroad/Nonroad emissions from OTAQ

\* Date Inventory data is due for incorporation into Draft Brochure or Report.

\*\* Latest year of data that is presented in the Trends Report.

### 3.0 NATIONAL DATA PRIORITIES

The purpose of this section is to identify areas in which EPA and S/Ls should focus the greatest attention in preparing and providing data for the 1999 National Emission Inventory. This section will: 1) discuss which data are most important in preparing the 1999 National Emission Inventory; and 2) describe the sources of data that we intend to draw upon, and the priority order in cases where there are multiple data sources.

#### 3.1 What Are the Overall Priorities (By Data Source)?

EPA's sources of data for the 1999 National Emission Inventory in priority order are:

- 1) S/L Emission Inventories
  - a) point sources (except SO<sub>2</sub>, Hg, and NO<sub>x</sub> emissions from electric power generation, and all HAPs from Municipal Waste Combustors)
  - b) area sources
  - c) mobile sources
  - d) biogenics (only for criteria pollutants)
- 2) EPA's Clean Air Markets Division CEM Data Base (Electric Power Generation SO<sub>2</sub> and NO<sub>x</sub> emissions)
- 3) EPA's Emission Standards Division MACT Data (mercury emissions from electric power generators, HAP emissions from municipal waste combustors, and HAP emissions from other MACT source categories)
- 4) EPA's Office of Transportation and Air Quality (onroad mobile source VMT and nonroad mobile source emissions for all pollutants)
- 5) EPA's Toxics Release Inventory (HAPs from certain point and area sources)
- 6) Data from other federal agencies
- 7) Adjustments to previous year emission estimates

#### 3.2 What Are the Priorities by Source Category?

##### 3.2.1 Point Sources

Electric Power Generating Utilities - Data will come from 3 sources in the following order:

- a) CEM Data and Hg from 1999 Utility Study - Because of the high degree of accuracy, emissions of SO<sub>2</sub> and NO<sub>x</sub> from the CEM Data Base (facilities covered by Part 75; may include facilities traditionally thought of as non-utilities), and emissions of Hg from the 1999 Mercury Utility Study (covers all electric generating units) will be used in the data base. To gain support from the S/Ls for these values, we will do the following:
  - Provide S/Ls with documentation on our procedures for calculating emissions from these data sources, including what we do to address data gaps and how many significant and non-significant data gaps occurred in the CEM Data Base for '97, '98.
  - Provide S/Ls with our data including emissions, heat input, ORIS ID, operating hours, data with significant gaps, which are CEM/nonCEM

- Offer to meet with S/Ls (with participation of CAMD and ESD) to discuss differences, explaining up front that EPA is compelled to use these numbers because of regulatory significance.
- b) S/L Emission Inventories - Electric utility emission estimates for pollutants other than SO<sub>2</sub>, NO<sub>x</sub>, and Hg, and emission estimates for all pollutants for facilities not covered by Part 75 (generally electric generating units less than 25 megawatts), will be taken from S/L agencies wherever available. S/Ls should use ORIS ID in reporting their information, and, where emissions are not directly reported to S/L agencies, emission estimates should be based on heat input operating hours from Part 75 or fuel use from EIA 767.
- c) EIA767 Data - Emissions for all other pollutants and facilities will be estimated based on data from the Department of Energy's EIA767 collection. During draft inventory review, EPA will describe the process used to estimate these emissions and will provide emissions values, heat input operating hours, SCC's, and ef's used to estimate emissions.

Non-Utilities - Data for non-utility point sources will come from 4 data sources in the following priority order:

- a) S/L Emission Inventories (except for municipal waste combustors as noted below)
- b) ESD MACT Data (HAPs) - *HAP emissions from municipal waste combustors will be used directly and will not be replaced with S/L's data because these data sets are highly accurate and the result of extensive source testing.*
- c) TRI (HAPs)
- k) Emission estimates generated using emission factors and activity data
- e) Adjusted emission estimates from previous year inventories (based on indicators such as economic growth)

### **3.2.2 Stationary Area Sources**

Same data sources/priorities as for Non-Utility Point Sources

### **3.2.3 Mobile Sources**

Data will come from data sources in the following priority order:

- a) S/Ls Emission Inventories (prepared from local activity data and EPA-approved emission factor models)
- b) FHWA VMT/OTAQ activity estimates and EPA emission factors
- c) Adjust emission estimates from previous year inventories (based on indicators such as economic growth or OTAQ-provided national emission estimates)

### **3.2.4 Biogenic Sources**

Data will come from S/Ls where available. Otherwise, estimates will be generated by EPA using the Biogenic Emission Inventory System (BEIS) Model.



## 4.0 DATA INCORPORATION PROCEDURES

### 4.1 What Will EPA Do to Help S/Ls Prepare for Data Incorporation?

In addition to providing information to S/Ls on how to prepare criteria and HAP emission estimates (located on the CHIEF web site), on the process for submitting data to EPA, and on EPA's plans for data incorporation, EFIG will distribute a preliminary set of 1999 HAP and criteria emission estimates for S/Ls to consider as they prepare their 1999 emission inventories. [Note: This preliminary inventory will be made publicly available on the CHIEF Web site and will serve as the basis for 1999 criteria emission estimates in the National Air Quality and Emission Trends Reports, prior to completion of the NEI Version 2.] These preliminary data will be provided to S/Ls in the NEI Input format. Included with this distribution will be point source HAP data provided to EFIG in the NEI Input Format by October 2000 that is the result of the MACT (Maximum Achievable Control Technology) standard-setting process managed by EPA's Emission Standards Division. Also included will be HAP data provided by for the chemical industry provided by the American Chemical Council. Data for the 1996 NET and NTI are currently available on EPA's ftp site: [ftp.epa.gov/EmisInventory](ftp://ftp.epa.gov/EmisInventory)

In preparation for data incorporation, EPA Emission Inventory Contacts in the Regional Offices will be contacting their S/L counterparts in early 2001 to gather the following information:

- The S/L agencies that will be submitting 1999 EI data to EPA for incorporation into the National Emission Inventory, and the appropriate S/L contacts
- The data that they plan to submit (pollutants, geographic coverage, categories)
- The format/mechanism they will be using to submit data (e.g., NIF V2.0)
- The date they plan to transfer the data to EPA (e.g., all point and area source categories by April 15, 2001; all mobile source categories by April 15, 2002)

*Note: Data must be transferred prior to June 1, 2001 to be incorporated into the National Emission Inventory Version 2; or prior to June 1, 2002 for Version 3.*

### 4.2 What Are the Steps for Incorporating Data Into the National Emission Inventory?

Incorporating data into the NEI is done as a 9-step process, as illustrated in Figure 4.1. The figure provides an overview of the process of incorporating data into the inventory and the expected timing of the key data processing steps. More detailed descriptions of each step are provided later in the section. The steps, in order of implementation, are:

1. Data Submittal
2. Quality Checks for Format
3. Complete the Database
4. Quality Checks for Content
5. Publish Draft 1999 National Emission Inventory
6. External Public Review
7. S/Ls Submit Data Revisions
8. Final Quality Checks
9. Publish the Update of 1999 National Emission Inventory

The illustration also highlights some related parallel activities that EFIG will perform that may be of interest to S/L agencies. One of these is the National Air Quality and Emission Trends Report that EPA prepares annually, which will eventually include data from the 1999 NEI. In addition, we are making preparations to identify and implement one unique Plant ID for each facility. Previously, we have stored one Plant ID for a facility in the NET (criteria pollutant) database, and a separate Plant ID for the same facility in the NTI (HAPs) database. Our objective is to eliminate instances of multiple ID's for the same physical plant. We are mapping the facilities in the 1996 NET, to the corresponding facilities in the 1996 NTI. This crosswalk will be provided in early 2001 to help S/Ls integrate their criteria and HAPs inventories during preparation of their 1999 and subsequent inventories.

In the future, this will influence our storage of S/L plant emissions data in the National Emission Inventory. S/Ls are encouraged to prepare a 1999 integrated inventory that eliminates the need for crosswalks between HAP and criteria inventory facilities. In lieu of an integrated inventory, S/Ls are encouraged to share with EPA any crosswalk information they have that relates facilities in their 1999 HAP and criteria inventories. If inventories are not integrated, and no crosswalk information is available from a S/L agency, EPA will create a crosswalk for the 1999 inventories and will use this as the basis for establishing a unique ID. Facility ID's from EPA's Facility Registry System will be evaluated for use as a unique facility identifier in the 1999 NEI.

After the summary illustration in Figure 4.1, Table 4.1 lists the major milestone activities, the respective roles and responsibilities intended for the participants in this data incorporation process, and the 'time window' of opportunity for S/L participation. The timetable to accomplish the sequence of events described in this document is ambitious for all participants. However, it is the best compromise at this time, given EPA's responsibility to continuously improve the national criteria and HAPs inventory - to update the data as often as possible, within the resources available each year, and in a manner that makes it ready to process for air quality modeling and risk assessment purposes. As we finalize and publish one update of the NEI, we immediately begin the next update cycle, which starts again with S/Ls transferring 1999 data updates that they want EPA to incorporate in the NEI.

While the data submittal schedule is aligned with the submittal dates prescribed in the Consolidated Emissions Reporting Rule (CERR), if data are not available according to the dates in this document, the contingency is to move forward with the latest data we have as the basis for updating our National Emission Inventory. It is expected that data not submitted by the due date may be submitted and incorporated into the National Emission Inventory in the subsequent (annual) update cycle.

### **4.3 What Are the Details for Each of the 9 Steps?**

#### **4.3.1 *Data Submittal***

##### Who Submits the Data?

The agency, either State or Local, that compiles the criteria and HAP emission inventory data should submit the data directly to EPA. *The exact EPA Office and location (e.g., Regional*

*Office, Office of Environmental Information, or Emission Factor and Inventory Group; ftp site, etc.) is to be determined.* Even in cases where the submitted data were developed by an outside source ( i.e., contractor or regional planning organization), we prefer to work directly with an S/L contact when addressing issues identified during the quality checks and any revisions that may be necessary to make the data set acceptable for incorporation into the National Emission Inventory. This will help ensure that the agency has direct knowledge of any modifications to its data set that will be represented and published in the National Emission Inventory.

### What Data Do S/Ls Submit and When?

The Consolidated Emissions Reporting Rule (CERR) describes the data elements that should be reported to EPA once the rule is final and in effect. Prior to the rule being in effect, S/Ls should report inventory data per existing requirements (SIP and annual criteria point source inventories) and any other data that they want reflected in the National Emission Inventory. The June 1 due date noted in this NEI Preparation Plan is the same as that prescribed in the proposed CERR. While the data submittal schedule here is aligned with the submittal dates prescribed in the proposed CERR, if the data due dates are not met, our contingency is to move forward with the latest data we have as the basis for updating our National Emission Inventory. It is expected that data which is not submitted by the due date, may be submitted and incorporated into the National Emission Inventory in the subsequent (annual) update cycle.

The data elements needed to complete the National Emission Inventory are listed in the Appendix B. The data elements listed are the same as those referenced in the proposed CERR.

### How Do S/Ls Submit Their Data?

The CERR lists the acceptable data transfer formats. The most widely used format will be the NEI Input Format (NIF). The most current version of the NIF, and its user documentation is posted under the EPA CHIEF Web site ([www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/)). Through June 2001, the NIF is available in two versions - V 1.2 and V 2.0 . The more recent V 2.0 has some minor structural changes that improve its implementation, and adds some data elements relevant to HAPs data transfer. The NIF V 1.2 will be retired after the June 2001 update cycle.

The following guide will determine which NIF version S/Ls should use for the June 1, 2001 update cycle: If submitting . . .

- HAPs data only - NIF V 2.0 ;
- Criteria and HAPs data - NIF 2.0 ;
- Criteria data only - NIF V 1.2 or NIF V2.0 .

The procedure for moving S/L formatted data files to EFIG electronically is outlined at the same Web site location as information about the NIF. In addition to the NIF data file(s), an electronic copy of the S/L's Inventory Submittal Form must be included to provide EFIG the information needed to log-in the data set for processing. The Submittal Form is also posted at the Web site.

Figure 4.1 Data Processing Steps

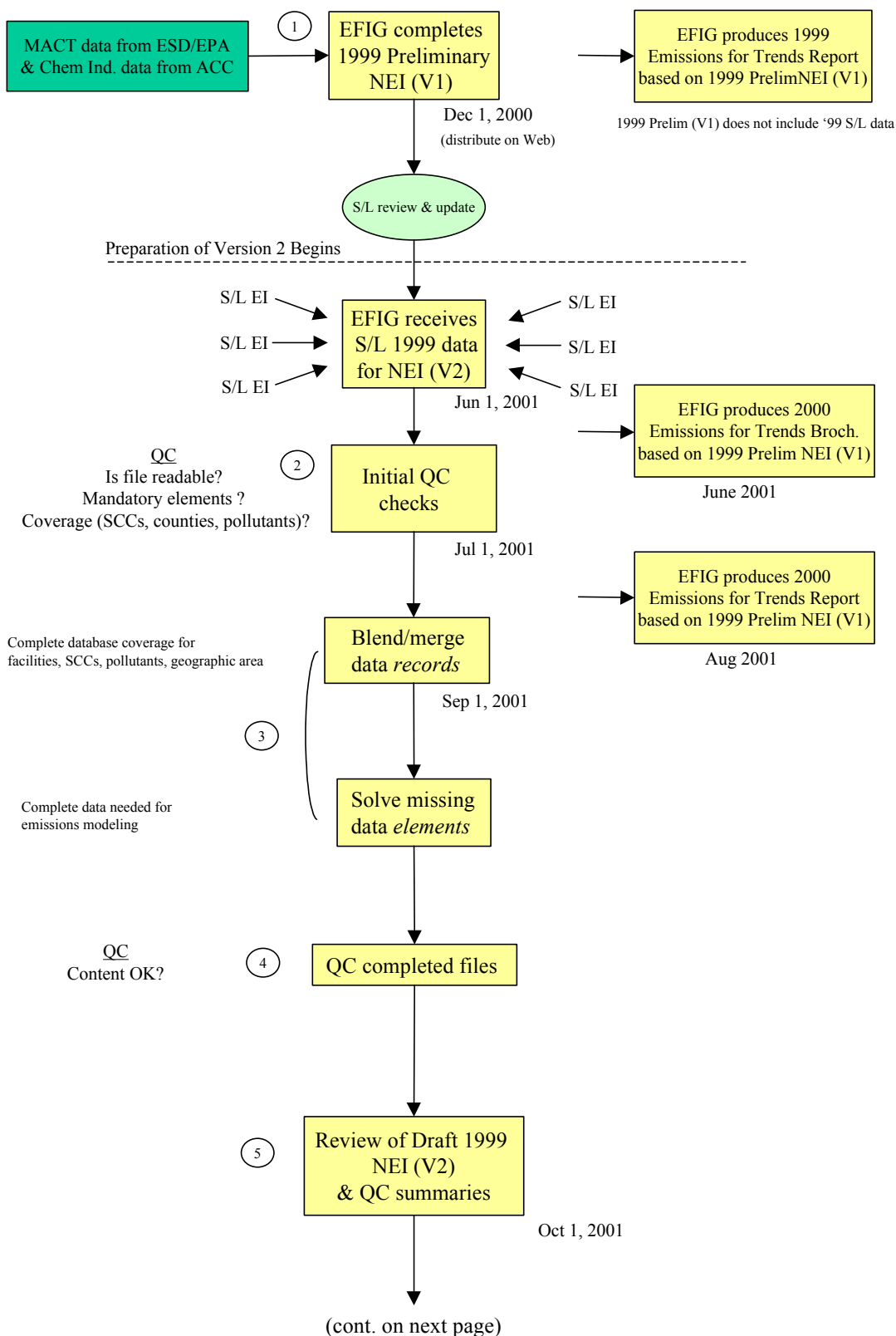
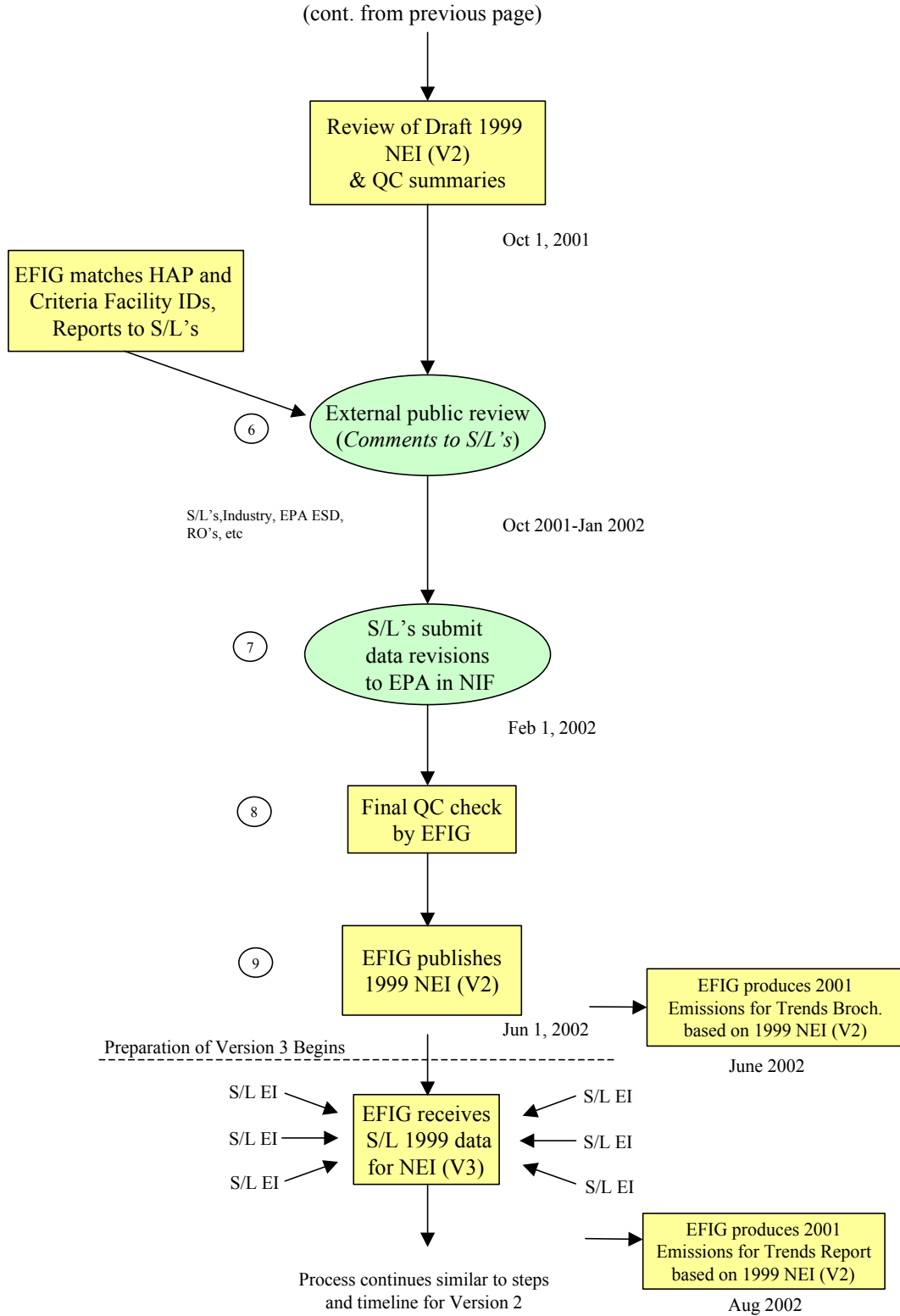


Figure 4.1 Data Processing Steps



**Table 4.1 Summary of Major Milestone Activities and Dates, and Roles & Responsibilities**

<b>Activity</b>	<b>By Whom</b>	<b>When</b>	<b>Expected Result</b>
Distribute preliminary 1999 National Emission Inventory (Version 1)	EFIG	Winter 2000	Will emphasize what is in the 1999 natl' ei <i>without</i> S/Ls data; Will prompt S/L's preparation of 1999 criteria and HAP ei data
1. Transfer 1999 EI data to EPA for incorporation into NEI Version 2	S/Ls	By 06/01/01	S/L's use prescribed data transfer format; 1999 S/L data incorp. into natl ei can begin
2. Initial QC of S/L data for format	EFIG, S/L data submitter	07/01	Confirmation that data file is readable and minimal data is present and adequate, or iteration to make it so
3. Complete the National Emission Inventory	EFIG	Fall 2001	Add any data necessary to support emissions modeling
4. QC the completed database for content	EFIG, Regional Offices, S/L data submitter	Fall 2001	Confirm if data content is reasonable overall, reports to S/Ls, ROs
5. Publish Draft National Emission Inventory V2	EFIG	Oct 1, 2001	Draft database accessible for public review
6. External public review	S/Ls, other EPA programs, industry, general public	Oct 2001 - Jan 2002	Significant comments are forwarded to the S/L's; S/L's prepare corrections for resubmittal as necessary
7. Transfer 1999 data revisions to EPA	S/Ls	By Feb 1, 2002	S/L's use prescribed data transfer format and resubmittal procedure; 1999 S/L data corrections to natl ei can begin
8. Final QC of 1999 S/L data corrections	EFIG, S/Ls	Spring 2002	Corrected data file is readable; data content is reasonable overall
9. Publish final Version 2 of 1999 National Emission Inventory	EFIG	June 2002	The 'preliminary' 1999 national ei distributed by EFIG in Fall 2000 is now updated with S/L data
1. Transfer 1999 EI data to EPA for incorporation into Version 3	S/Ls	June 2002	S/L's use prescribed data transfer format; 1999 S/L data updates to natl EI

The procedure for sending the files electronically to EPA will change somewhat to implement the Central Receiving Facility (CRF) for air emissions inventory data. The CRF is being designed and tested now. When the CRF is ready to use, it will be announced on the CHIEF Web site, and the new steps in the electronic transfer process will be defined. The acceptable data transfer formats will stay the same. The primary thing expected to change with the CRF, is *where* the S/Ls send their data files. This will likely include the implementation of an individual and private electronic mailbox for each agency which will also enable automatic receipt notification and efficient transfer of QC reports regarding each S/L's data files.

#### How Will S/Ls Know Their Data Was Received?

Upon receipt of the S/L inventory data, EPA will email an acknowledgment to the S/L contact with the following information:

- Date the data was received
- Summary of what was received (e.g., file names)
- EPA's next steps (e.g., applying quality checks for format)
- When the S/L should expect further contact from EPA

#### How Will EPA Track What Was Received?

Upon receipt from State/ Local Agencies, the files will be downloaded and stored in a "read-only" archive. A copy of the original file will be placed into a secure working directory, accessible only to EFIG. The working file will be submitted to EFIG's data processing steps (e.g., initial QC checks, blend/ merge of data records, and solving for missing data elements). For each data file processed, the EFIG will maintain an internal data processing log to describe the status and results of each processing step.

This data handling process will be repeated in a separate secure manner for data revisions received as a result of the public review process. Upon completion of that process, the files will be moved to a secure EFIG directory, and from this staging area, will be loaded into the EPA NEI database, and subsequently posted onto a secure EPA Internet site for public access. All EPA archive and working drive directories will be backed-up periodically in a secure manner.

#### **4.3.2 Quality Checks For Format**

The initial set of checks done on the data will ensure that the file can be processed. The quality checks seek answers to the questions below regarding the file format. The logic for these checks is based primarily on the NEI Input Format specifications, the use conventions described in the User Instructions and its Appendix A, and the NIF Codes Tables. [All of these are presently located on the Emission Factor and Inventory Group Web site at [www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/).]

### What Types of Checks Will be Applied to the Files?

- Does the file conform to the format specification? - The initial checks performed on each submitted data set will verify that the file format is correct, and therefore readable for further processing (e.g., field widths, begin / end position, data types).
- Are mandatory data elements reported? - The presence or absence of mandatory data elements will be confirmed. Some of the mandatory data fields are the *primary keys* in each record that help relate and maintain the individual records together in a file for subsequent processing. The mandatory data also provide a minimal set of information needed to represent data for an agency in the national data base.
- Does the data set contain what the S/L agency said they are submitting? - The data in the file will be compared to the Inventory Submittal Form that was provided with the file to verify the noted and intended coverage for geographic area, pollutants, source categories, and temporal information.

### How will EPA track the data problems initially encountered with data files?

EFIG will keep a log of errors and problems encountered with each of the 1999 data submissions, and will provide those to the S/L agency when communicating with the agency.

### Will the quality checks be available for S/Ls to apply to their data sets?

In addition to documenting the quality checks for manual application, we will develop automated software in Microsoft Access to perform many of these format checks and will post the software program on the CHIEF Web site as it is complete and operating successfully. Appendix A lists the quality checks that we will use for each data file.

### **4.3.3 Complete the Data Base**

#### Why? and What is Complete?

In order for the 1999 National Emission Inventory to be a dependable and reliable starting point for use in regional scale modeling and human exposure modeling, the data must be comprehensive for:

- all criteria pollutants, precursors, and the 188 HAPs (for information on reporting for compound groups, refer to [www.epa.gov/ttn/chief/nti/ntiq&a.pdf](http://www.epa.gov/ttn/chief/nti/ntiq&a.pdf));
- all 50 states / counties, D.C., and the territories of Puerto Rico and the Virgin Islands (HAPs only); and
- all significant sources of emissions including point (major and non-major), area, mobile, and biogenic source categories.



Ultimately, we are trying to complete the National Emission Inventory for the following data - (these completion objectives may vary between criteria and HAPs data and those differences are described in Appendix B):

- specific geographic location of point sources;
- facility emissions at point/ segment/ stack level;
- operating schedule data for facility;
- stack parameters for facility;
- annual and daily emissions for major/point facilities (at the point/segment/stack level);
- annual and daily emissions for area and mobile sources.

Annual emissions are for a calendar year period. Daily emissions refer to a typical work week's daily emissions for a peak nonattainment season such as summer for ozone or winter for CO.

#### How? and What Techniques Do We Use?

As only a portion of this information may be provided by the S/L agencies, EFIG must gather and prepare additional information to complete the data base. Completing the data base consists of a 2-step process:

- 1) Add records to fill in missing facilities or source categories, or to fill in for missing geographic areas (e.g., where data were not reported for entire counties), and
- 2) Add, or solve for, data elements missing in existing records.

Each step is described briefly below. Appendix B documents more specifically the methods we will use to solve for missing data elements.

#### *Blend / Merge Data Records*

EFIG will prepare the 1999 National Emission Inventory using the sources of data described in Section 3.1. Our techniques and guidelines for using these information sources to compile and complete the point, area, mobile, and biogenics data respectively, are described in detail in Appendix B - *EFIG's Methodology for Completing the Database*.

#### *Solve For Missing Data Elements*

Appendix B outlines the methods that will be used to complete the data elements missing in existing records. The data elements listed are recognizable from the NEI Input Format. Many of the data elements are identified as "Mandatory", where they must be provided by the S/L agency, or as "Solvable", where a method exists for completing the missing data. These methods are indicated in the table of Appendix B.

#### **4.3.4 *Quality Checks for Data Content***

While the first set of quality checks looked at format issues, this second set of checks will focus on the integrity of the data content.

### What types of content checks will be performed on the data?

These quality checks will examine the following:

- Conditional Fields - Fields required by other fields in the same table. For example, if there is a PCT Capture Efficiency in the CE Table, then there should also be a Primary Device Type in the CE Table. See Appendix A of DIP for other Conditional Fields.
- Acceptable Codes - The Pollutant Code (and all acceptable codes) should come from the Code Table in Appendix A of NIF User Instructions. The User Convention Notes for each field in the NIF state which fields are Coded Fields and refer to the Code Table in Appendix A of NIF User Instructions.
- Numeric Values In Acceptable Range - For example, the Annual Average Days Per Week in the EP Table should be  $< 7$ . See Appendix A of this document for other Acceptable Ranges.
- Inter-File Format - Fields required by other fields in different tables. If there is an Emission Record in the EM Table, then there should be an associated activity record in the AC Table. See Appendix A of this document for other Inter-File Format requirements.
- Inter-Source Relationships - See Appendix A.
- Inter-Pollutant Relationships - See Appendix A.

#### **4.3.5 Publish Draft National Emission Inventory**

Once the draft inventory has been completed and quality check results have been addressed, the inventory will be made available for public review. The draft National Emission Inventory will be provided, along with extensive documentation, in the following formats and forums:

- Detailed data set (NEI Input Format) - at: [ftp.epa.gov/EmisInventory](ftp://ftp.epa.gov/EmisInventory)
- Summary data sets (Excel or Access Format) - at: [ftp.epa.gov/EmisInventory](ftp://ftp.epa.gov/EmisInventory)  
- E.g., emissions by facility, by source category, etc., organized by State/county

#### **4.3.6 External Public Review**

The purpose of the 4-month external review is to give interested parties an ample opportunity to confirm or correct the draft version of the data prior to processing it for use in air quality modeling. Interested parties may include S/Ls, regional planning organizations, EPA program offices (i.e., Regional Offices, ESD, OTAQ, ), industry, and other public entities. While questions and comments concerning emissions data may be communicated directly to EPA, we will in most cases not act on those questions/comments until the commenter has coordinated them through the S/L agency that has air quality jurisdiction for that area so that the S/L agency can confirm the appropriate response. In order to ensure respectful and responsible depiction of

the S/L data in the National Emission Inventory, we will only accept data corrections directly from, and therefore validated by, the S/Ls.

The external review process will result in each S/L agency submitting data corrections, where necessary, for EPA to place in the National Emission Inventory. The resubmittal procedure is described below.

#### ***4.3.7 S/Ls Submit Data Revisions***

Revisions must be submitted to EPA by the S/L agency that submitted the original data set. Other agencies, organizations, companies or individuals may wish to request revisions to the draft National Emission Inventory, or to comment on the draft inventory to indicate where revisions are necessary. Such requests or comments must be coordinated through the appropriate S/L agency. This is necessary to avoid EPA receiving conflicting revisions. Revisions will be accepted only in the NEI Input Format, with flags added to indicate whether the revision is a change (i.e., changes to specific data elements), add (addition of specific data elements to an existing plant or source category record), or a delete (i.e., removing an entire record). It should be noted that this step in the process is for the purpose of revising data that was included in the draft inventory. While missing data elements may be added during this step, plants or source categories (i.e., entire records) should not be added, but should be provided to EFIG prior to June 1, 2002, for incorporation into Version 3 of the National Emission Inventory.

Revisions to the 1999 National Emission Inventory will be subjected to a rigorous review process in order to ensure internal consistency. Specifically, EPA/EFIG will perform the following steps upon receipt of S/L data revisions:

- Review the documentation provided by the commenter to ensure that it is consistent with the actual changes to the inventory submitted;
- Verify that the change/add/delete designations of the revisions are accurate (e.g., that a record designated for addition is not in fact a revision to an existing record);
- Verify that there is no source category overlap between the existing draft and revised data (area source categories);
- Perform reality checks on emission estimates, by source category and HAP, to identify outliers and determine the validity of such estimates;
- Verify that added pollutants are HAPs; and
- Verify that added pollutants have correct or valid Chemical Abstracts Service numbers.

For the most part, revisions provided by S/Ls will be incorporated to produce the inventory. When questions arise over specific revisions, the reviewers will be contacted by EFIG. Memos describing how specific revisions are addressed into the National Emission Inventory will be sent to the S/L reviewers.

In general for point sources, revisions will be incorporated into the National Emission Inventory using the following methodology:

1. All changes and additions will be made for:

- Process throughput information
- Zip codes
- Control/capture efficiencies
- Control device information
- SCC/process descriptions
- Stack parameters
- Unit descriptions
- lat/long or UTM coordinates
- FIPS codes
- SIC codes

2. Emission record additions and changes to existing emission records will be automatically processed if the original data were provided by the S/Ls.

3. Facilities marked for deletion will be evaluated to determine whether the deletions should be processed.

C If the draft data are provided by S/Ls, we will process the deletion.

C If the draft data are provided by TRI or MACT engineers, we will further evaluate the proposed deletions. We will use trade association journals to determine if facilities were in operation in 1999. If a facility was closed in 1999, we will delete the facility.

C If a TRI or MACT facility is recommended for deletion, but is open in 1999, and no duplicate facility is present in the draft inventory, we will not process the deletion.

C If a TRI or MACT facility is recommended for deletion, is open in 1999, and a duplicate facility is present in the draft inventory, we will then evaluate the HAPs emitted from both facilities. If the HAP emission records are duplicated, we will process the deletions. HAP emission records not duplicated will be retained. We will then merge the HAPs for the duplicate sites using a unique site ID.

Note: Additions of new facilities not in the draft National Emission Inventory should be provided just prior to June 1, 2002 for incorporation into Version 3.

#### **4.3.8 Final QC**

A final set of quality checks will be applied to the revised National Emission Inventory. These checks will be a repeat of the same checks applied previously concerning format (Section 4.3.2) and content (Section 4.3.4).

#### **4.3.9 Publish the 1999 National Emission Inventory**

Version 2 of the 1999 NEI will be published June 2002 in the following forums:

- AIRS Data Web site (national/state/county data summaries and detailed data queries)
- FTP site (raw data files in NEI Input Format)
- Trends Report (national level emissions by source category)

Each of these sites can be accessed from [www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/), by clicking on Emission Inventories, Emission Inventory Data, Download Data.

#### **4.4 What Were the Most Common Problems Encountered With Incorporating S/L Data into the 1996 NTI and NET?**

The problems most often encountered with data received from State/local agencies for 1996 are listed below. Agencies should take appropriate steps to ensure that such problems are corrected in the 1999 data prior to transferring data to EPA.

- Data set does not correctly follow NEI Input Format
- Missing Primary Key fields - e.g., stack ID for a Stack record (refer to NEI Input Format for Primary Key fields)
- Missing Mandatory fields - e.g., City, State, Zip Code (refer to NEI Input Format for Mandatory fields)
- Missing Emission Release Point records - even if there are no stacks for a facility, there needs to be at least one release point identified per facility (e.g., vent or fugitive release).
- Missing related records - e.g., Emissions record with no Emissions Process record (technical term for this is “orphan child”).
- Duplicate key data - e.g., two Emission records with same key/identifier and same amount of emissions. For this example, it would be unclear whether this was actually a single emissions point that was inadvertently duplicated, or two distinct emissions points that were inadvertently given the same identifier.
- Transposed latitude and longitude coordinates - e.g., latitude coordinates in the longitude field, and vice versa.
- Duplicate records - e.g., plants listed twice, sometimes under different names
- Insufficient quality checking of the data prior to transferring the data to EPA



# Appendix A

## Quality Check Routines for Processing S/L Data

(page left intentionally blank)



## I. Quality Checks for File Format

The logic for these checks is based primarily on the NEI Input Format (NIF) specifications (NIF V1.2), the use conventions described in the NIF User Instructions and its Appendix A, and the NIF Codes Tables. All of these are presently located on the Emission Factor and Inventory Group (EFIG) Web site at [www.epa.gov/ttn/chief/](http://www.epa.gov/ttn/chief/) under Emission Inventories, Emission Inventory Data, Submitting Data to EPA. [Note: Additional checks are being developed to correspond to changes made in NIF V2.0. These additional data checks will be described in an addendum to this plan and will be posted by EFIG on the same web site by Spring 2001.]

### Steps 1 - 3

1. Query routines are applied to each data file to check for proper file format (NIF V1.2). The data files MUST pass these checks in order for the files to be further processed for incorporation into the NEI.

### POINT SOURCES

#### A Data format and field widths

File	Field	Check For:	Consequence	How Found
All	All	> max length	blanks in mandatory fields	import error if NIF file type - .txt or .dbf; length query for numerics and table design for text if NIF file type - .mdb
All	Any numeric value	scientific notation	blank in numeric field	import error if NIF file type - .tat or .dbf,

#### B Mandatory fields must be filled

File	Field (V1.2)	Consequence	How Found
All	Record Type	Can't Analyze, <b>referential integrity compromised</b>	null queries
All	State FIPS, County FIPS	Can't QC, analyze, or map without locational data <b>ref. int. comp.</b>	null queries
TR	Organization name	Can't track back emissions to source	null queries
TR	Transaction Type	data not replaced	null queries

File	Field (V1.2)	Consequence	How Found
TR	Inventory Year	Can't QC, analyze, or trend w/o temporal data	null querys
TR	Contact name, phone, creation date, submission number	NEED if questions	null querys
TR, AC, EM	Start Date, End Date, Day Type	Can't QC, analyze, or trend w/o temporal data	null querys
TR	source type	Need to QC and categorize	null querys
TR	format version	format for import	null querys
SI, ER, EP, CE, AC, EM	AFS/NEDS Plant ID, SIC	Can't QC, analyze, or map without locational data or track back emissions to source, <b>ref. int. comp.</b>	null querys
SI	Facility Name, street, city, state, zip	Can't QC, analyze, or map without locational data	null querys
EU, ER, EP, CE, AC, EM	AIRS Point ID	Can't QC, analyze, or map without locational data or track back emissions to source, <b>ref. int. comp.</b>	null querys
ER, EP	Stack ID	Can't QC, analyze, or map without locational data or track back emissions to source, <b>ref. int. comp.</b>	null querys
ER	x & y coordinate, xy coordinate type	Can't QC, analyze, or map without locational data	null querys
EP, CE, AC, EM	Segment ID, SCC	Can't QC, analyze, or map without locational data or track back emissions to source, <b>ref. int. comp.</b>	null querys
CE, EM	Pollutant code	Can't QC if missing, can't analyze by pollutant, <b>ref. int. comp.</b>	null querys
EM	Inventory Type Code		null querys
EM	Emission numeric value	can't do anything w/o	null querys
EM	emission unit numerator	need to analyze	null querys
EM	Emission Type	need to QC, gives dates	null querys

C Compare with State provided Submittal Form

I. geographic coverage

Submittal Form	TR	EM	Comment
coverage on SF	query unique counties	query # records/cty	compare cty info to SF

ii. pollutant coverage

Submittal Form	EM	Comment
coverage on SF	query # records/pollutant	compare pollutant info to SF

iii. source category coverage

EM	Historical	Comment
query # records/1st 3 digits of SCC	query # records/1st 3 digits of SCC	compare current to historical

iv. temporal coverage

Submittal Form	TR	EM	Comment
coverage on SF	query unique start and end dates	query unique start and end dates, inventory type, and emission type	compare temporal coverage to SF

D Referential Integrity

The NEI is a relational database. The following one to many relationships between files/tables must be created and enforced. For example, the relationship between the TR file/table and the SI file/table is a one to many relationship, i.e. there is one TR record for many SI records or one unique county record for possibly many sites. A relationship is created unless one of the following two errors occurs:

1- orphan records (For example, a record in the site table exists without a record in the TR table for that county).

2- records not unique for key fields (A county was listed twice in the TR table).

TR-SI	SI-EU	EU-ER	EU-EP	EP-AC	AC-EM
-------	-------	-------	-------	-------	-------

All of the above are strict 1 to many relationships. The CE table has a conditional 1 to many relationship with the EM table because not every emission is controlled, yet one control record can connect with several EM records which occurred during different periods. Therefore, if there is a CE record there should be a 1 to many relationship between that record and many corresponding EM records.

2. Notify data submitter by e-mail of completion / results of Format checks.  
e-mail to include:

- A the Microsoft Access Database created or .mdb file;
- B table summary identifying problems and the queries used to find them;
- C written summary identifying what we can fix here and how, what they need to fix and how, and if necessary to resubmit.

3. Repeat 1 and 2 until Format checks are error free.

Note: File/Table Abbreviations:

- TR = Transmittal
- SI = Site
- EU = Emission Unit
- ER = Emission Release Point
- EP = Emission Process
- CE = Control Equipment
- AC = Activity
- EM = Emission

## II. Quality Checks for Data Content

### Steps 1 - 3

1. Query routines applied to each data file (NIF V1.2):

#### A Acceptable Codes

File	Field (V1.2)	Correct	Found	Consequence
All	Record Type	TR, SI, EU, ER, EP, CE, AC, EM	unique code query	Can't Analyze, referential integrity
All	State, County FIPS	See AIRS GeoCommon Code Table	unique code query	locational QC and analysis
TR	Transaction Type	00 - original, 05 - replace	unique code query	data not replaced
TR, AC, EM	Day Type	0-8 See Appendix A	unique code query	
TR	Source Type	See source Type Code Table	unique code query	can't categorize analysis by source
TR	Contact Type	0-5 See contact Type Code Table	unique code query	can't back track for QC
SI	Facility Category	0-3 See facility category Code Table	unique code query	can't categorize analysis by fac type
SI	ORIS facility code	DOE unique identifier for electric generating units see <a href="http://www.eia.doe.gov">www.eia.doe.gov</a>	unique code query	can't categorize analysis by fac type
SI, EU	SIC	Standard Industrial Classification Code. See <a href="http://www.census.gov/epcd/www/sic.html">www.census.gov/epcd/www/sic.html</a>	unique code query	can't categorize analysis by SIC
SI, EU	NAICS	N. Amer. Industry Classification System. See <a href="http://www.census.gov/epcd/www/sic.html">www.census.gov/epcd/www/sic.html</a>	unique code query	can't categorize analysis by NAICS
SI	Address Type Code	0-6 See address type code table	unique code query	can't back track for QC

File	Field (V1.2)	Correct	Found	Consequence
EU	ORIS boiler Code	DOE unique identifier for electric generating units see <a href="http://www.eia.doe.gov">www.eia.doe.gov</a>	unique code query	can't categorize analysis by boiler
EU	Emission Unit Type	0-18 See Emission Unit Code Table (v1.2 only)	unique code query	can't categorize analysis by unit type
EU	Design Capacity Unit Numerator and Denominator	See Units Code Table	unique code query	can't analyze unit capacity
ER	Emission Release Point Type	0-6 See Emission Release Point Type Code Table	unique code query	can't categorize analysis by ERP type
ER	XY Coordinate Type	LATLON or UTM See XY coordinate Type Code Table	unique code query	locational QC and analysis
EP, CE, AC, EM	SCC - Source Category Classification Code	See <a href="http://www.epa.gov/ttn/chief/codes/index.html">http://www.epa.gov/ttn/chief/codes/index.html</a>	unique code query	can't categorize analysis by SCC
CE, EM	Pollutant Code	See Pollutant Code Table	unique code query	can't stratify by pollutant
CE	Primary and Secondary Device Code	See Control Device Type Code	unique code query	Can't analyze CE
AC	Throughput Unit Numerator	See Units Code Table	unique code query	can't analyze activity data
AC, EM	Material	See Material Code Table	unique code query	can't analyze activity data
AC, EM	Material I/O	3 possible- See Material I/O Code Table	unique code query	can't analyze activity data
EM	Inventory Type	7 possible - See Inventory Type Code Table		
EM	Emission Unit Numerator	See Units Code Table	unique code query	can't analyze emission data

File	Field (V1.2)	Correct	Found	Consequence
EM	Emission Type	See Emission Type Code Table	unique code query	Can't QC, gives dates
EM	Factor Unit Numerator and denominator	See Units Code Table	unique code query	can't analyze factor data
EM	Factor Type code	See Factor Type Code Table	unique code query	can't analyze by factor type
EM	Calculation Method code	See Factor Calculation Method Code Table	unique code query	can't analyze by factor calc method
EM	EF Reliability Indicator	A - F: See Reliability Indicator Code Table	unique code query	can't analyze/utilize by quality rating
EM	Rule Effectiveness Method	11 possible - See Rule Effectiveness Method Code Table	unique code query	can't analyze by rule effect meth

#### B Numeric Values In Acceptable Range

File	Field (V1.2)	Acceptable	Found	Consequence
TR, AC, EM	End Date - last month and day of season being inventoried (YYYYMMDD)	>= Start Date beginning month and day of season being inventoried	range query	Can't QC or Analyze
TR, AC, EM	End Time (HHMM)	>= Start Time	range query	Can't QC or Analyze
TR, AC, EM	Start Date and End Date	within Inventory Year	range query	Can't QC or Analyze
SI	city, state, and zip	within St, County if address type code is Physical	range query	

File	Field (V1.2)	Acceptable	Found	Consequence
ER	Stack Height (in feet)	[0-100] if ER Point Type = fugitive or 01, (0-700] if ER Point Type = non fugitive or 02-06	range query	
ER	Stack Diameter (in feet)	(0-50]	range query	
ER	Exit Gas Temperature (degrees Fahrenheit)	(50-1500]	range query	
ER	Exit Gas Velocity (in feet per second)	(0-100]	range query	
ER	Exit gas Flow Rate (in cubic feet per second)	(0-200,000)	range query	
ER	computational check	Exit gas Flow Rate = $(J * \text{Stack Diameter}^2 * \text{Exit Gas Velocity}) / 4$		
ER	x coordinate (in decimal degrees or UTM Easting Km)	Should fall w/I range for state	range query	
ER	y coordinate (in decimal degrees or UTM Northing Km)	Should fall w/I range for state	range query	
EP	winter, spring, summer, fall throughput PCT	Sum should = 100 PCT	range query	Print out list if < 98%
EP (annual average), AC (period)	annual average or period days/week	(0-7]	range query	
EP (annual average), AC (period)	annual average or period weeks/year	(0-52]	range query	



File	Field (V1.2)	Acceptable	Found	Consequence
EP (annual average), AC (period)	annual average or period hours/day	(0-24]	range query	
EP (annual average), AC (period)	annual average or period hours/year	(0-8760] or 8784 if leap year	range query	
CE	PCT Capture Efficiency	<= 100	range query	
CE	Primary PCT Control Efficiency and Total Capture Control Efficiency	<100	range query	
EM	Emission Numeric Value	>=0	range query	
EM	Emission Numeric Value - depends on pollutant code	see emission range table at <a href="http://www.epa.gov/ttn/chief/eidocs/index.html#pack">www.epa.gov/ttn/chief/eidocs/index.html#pack</a> with NIF v 2.0 user package (to be provided)	range query	
EM	Emission Numeric Value	For PM, PM2.5 < or = PM 10	range query	
EM	Emission Numeric Value	Depending on Inventory type, Seasonal < or = Annual	range query	

2. Notify data submitter by e-mail of completion / results of Content checks.  
e-mail to include:
  - A the Microsoft Access Database created or .mdb file;
  - B table summary identifying problems and the queries used to find them;
  - C Summary identifying what we can fix here and how, what they need to fix and how, and if necessary to resubmit.
  
3. Repeat 1 and 2 until Content checks are error free.

### **III. Final QC**

Repeat steps in I-II above.

## Appendix B

### Methodology for Completing the Data Base

(page left intentionally blank)

# I. Methodology For Blending / Merging Records

## Point Sources

The methodology for compiling the point source data in the 1999 National Emission Inventory uses the following hierarchy:

- 1) Obtain HAP and criteria emissions data from the S/Ls.\*
- 2) Supplement S/L data for electric generating utilities with MACT and CEM data.
- 3) Supplement S/L HAP data with MACT inventory data.
- 4) Supplement S/L and MACT HAP data with TRI data.
- 5) Supplement S/L, MACT and TRI HAP and criteria data with EPA preliminary emission estimates (based on projections from previous year emissions) where necessary.

\*Note: Facility-specific MACT inventory information for two MACT source categories - electric utility mercury emissions and HAP emissions from municipal waste combustors will be used directly and will not be replaced with S/L data because the MACT data sets are highly accurate and the result of extensive source testing. Likewise, data from the Acid Rain CEM data base for electric utility NO<sub>x</sub> and SO<sub>2</sub> emissions will be used directly and will not be replaced with S/L data because of the high level of accuracy.

As described in the list above, HAP data for the same source category could come from multiple information sources (e.g., S/Ls, MACT, etc.). With the exception of mercury from utilities and municipal waste combustors noted above, EFIG will use the following criteria to determine which source of HAP data to use for a facility in multiple data sets:

- Base year 1999.
- Actual emissions reported rather than potential, allowable, maximum.
- Contains location data (Latitude/Longitude coordinates, UTM coordinates).
- Has SCCs associated with emissions.
- Has stack data.

Potential facility data gaps in S/L emission inventories will likely include:

- Entire counties missing, and
- Missing facilities.

To assess the completeness of S/L-submitted emission inventories for facility coverage within a geographic area, EFIG will use the 1996 NTI and NET inventories, MACT inventory data and TRI. If S/Ls are aware of facilities in the 1996 NTI and NET that have closed, are duplicate facilities, or changed their names, the S/Ls should notify EFIG of this information in order to reflect these changes in the 1999 National Emission Inventory. This information can be summarized in a memo or spreadsheet format but would not be part of the NEI Input Format submittal. For facilities in geographic areas that are not submitted to EFIG and for missing individual facilities, EFIG will fill point source facility gaps using the steps 3 - 5 of the hierarchy presented above.

One of the goals of the 1999 National Emission Inventory is to compile a comprehensive inventory of emissions of the 188 HAPs. In the process of merging point source data from the 3

HAP data sources (S/L, MACT, and TRI data), EFIG will include all HAPs emitted from a facility. If a facility is duplicated in one or more data sources, EFIG will evaluate the HAPs emitted from the processes and emission units at the duplicate facilities. EFIG will retain process, emission unit and individual stack emissions of HAPs. If HAPs emitted from stacks, emission units, processes or at the facility level are duplicated, S/L data will be used and duplicate data from MACT or TRI data will be deleted. If only a subset of HAPs are included for duplicate facilities, EFIG will retain the subset of HAPs from S/L data and then merge the remaining HAPs from the duplicate sites to create a unique site with a unique set of HAPs. Similarly EFIG will evaluate data at the process, emission unit, and individual stack level. Data will be retained at the most refined level. For example, MACT data are available for two petroleum refinery MACT standards - catalytic cracking units and vents, and TRI data are available at the facility level. The more refined MACT data for the two categories would be compiled in the National Emission Inventory rather than total facility data from the TRI.

For criteria pollutants, because the data generally come from a single source (S/L), we do not fill in missing pollutant emissions by merging data from other information sources. Instead, we estimate the missing pollutant emissions based on projections from previous year emissions. This methodology is described in Section II, Methodology For Data Completion.

## **Area Sources**

To begin compiling the area source data, EFIG will first identify area source categories that emit HAPs and criteria pollutants. The 1996 NTI and NET will be the starting point for area source category identification. The methodology for compiling the area source data in the 1999 National Emission Inventory uses the following hierarchy.

1. Obtain emissions data from the S/Ls (HAPs and criteria)
2. Supplement S/L data with MACT inventory data (HAPs)
3. Supplement S/L and MACT inventory data with EFIG generated data using emission factors and activity data (HAPs and criteria)

Emissions from a limited number of area source categories in the 1999 National Emission Inventory will be estimated using national data sets and AP-42 factors for those areas for which S/Ls do not provide data. These source categories include:

- Fuel Combustion
- Solvent Usage
- Storage and Transport
- Waste Disposal
- Fugitive Dust
- Other Combustion (wildfires, prescribed burning, agricultural burning)
- Animal Husbandry
- Fertilizer Application

Each of these categories may be made up of multiple processes. EFIG may estimate emissions for certain processes within these categories on a case by case basis where emission estimates are

not provided by S/Ls. EFIG will evaluate emission factors for these categories for their completeness, representativeness, and overall quality. In some cases, EFIG will use emission factors from published documents such as AP-42. For other source categories, more recent emission factors based on recent test data gathered from the MACT program, S/Ls, or industry will be used. EFIG will obtain most source activity data from published sources such as government statistical documents and databases (e.g., Energy Information Administration fuel consumption reports, U.S. Forest Service reports on fires and burned acreage, and waste disposal reports published by EPA), industry trade publications, and commercially published business directories and journals. Census data from the U.S. Department of Commerce will be used to estimate emissions using per capita and per employee emission factors. Similar to the development of the 1996 NTI and NET, EFIG will use information about raw material usage and production levels supplied by industry and trade groups. These data will be extrapolated as needed to represent emissions on a county, state and national scale. EFIG will also use available S/L activity data. Wherever possible, EFIG will estimate area source emissions for HAPs and criteria pollutants in an integrated fashion (i.e., will use the same activity data and methodology).

### **Mobile Sources**

The methodology for compiling the on-road and non-road mobile source data in the 1999 NTI uses the following hierarchy:

1. Obtain HAP and criteria emissions data from the S/Ls.
2. Supplement S/L's inventory data with data from OTAQ and the Federal Highway Administration (Highway Performance Monitoring System - vehicle miles traveled)

For on-road mobile sources, EFIG will require the following from S/Ls:

- average daily or annual VMT,
- mobile model input files, and
- average daily and/or annual emissions

All three of these data sets are useful in preparing the National Emission Inventory and preparing specific data sets for air dispersion modeling. S/L should compare their onroad VMT estimates at the State or Local level with VMT from HPMS to confirm that they are within 10%. When submitting VMT and emissions data to EPA, S/L's should inform EPA about the findings of their comparison. If outside 10%, EFIG will use our discretion as to whether or not to use the S/L-provided VMT. In such a case, EFIG will hold discussions where possible with the S/L, OTAQ and FHWA to determine the best approach.

For non-road mobile sources, EFIG will accept from S/Ls, emission estimates created using the latest version of the NONROAD Model. EFIG is planning on running the NONROAD Model at the county level for the entire U.S. for 1999. These emissions, as well as the NONROAD Model inputs used to create them (esp. temperature, RVP, and activity) will be supplied to S/Ls for review. S/Ls may use these emissions as their nonroad submission to the NEI if they do not wish to change any of the inputs. These emissions will be used in the NEI for all areas for which S/Ls do not submit alternative nonroad data.

EPA's Office of Transportation and Air Quality will provide direction on how emission estimates should be developed and advice on which emission factors and speciation profiles (for HAP emission estimates) should be used.



## II. Solve for Missing Data Elements

### Point Sources

After an inventory of facilities and associated emissions is compiled by EFIG, missing data elements will be defaulted by EFIG. Missing data elements may include:

1. Stack parameters;
2. Facility location data [address, county FIPS and names, latitude/longitude or Universal Transverse Mercator (UTM) coordinates]; and
3. MACT codes (The NTI associates MACT codes with every MACT source category to the major and area source data. The tagging of data with MACT codes allows EPA to determine reductions attributable to the MACT program.).

Because the NET and NTI will be used for modeling, the association of stack parameters and location data to each facility emission estimate of major sources is required. EFIG will obtain default stack parameters from the following sources in the hierarchy presented.

1. Facility specific stack data in 1996 NTI, and NET inventories
2. Derived stack data (derive velocity, diameter, and flow rate if 2 of 3 variables are provided)
3. Default stack data associated to SCCs (file provided by ORD)
4. Default stack data associated to SIC codes (file provided by OP)
5. Model plant defaults provided by MACT engineers

EFIG will assume that the release point is a stack for facilities with no information on the types of release (stack vs. fugitive). The NEI Input Format contains flags to identify data fields that use defaults.

The NET and NTI requires that point sources include facility location data. If latitude/longitude coordinate problems exist (missing or incorrect), EFIG will fill location data gaps using the following hierarchy of methods.

1. Using existing inventories - 1996 NTI, TRI, NET inventory
2. Using zip codes to identify latitude/longitude coordinates
3. Prioritizing remaining facilities based on the size of emissions. Small facilities (e.g., emissions of total HAPs less than 10 tons/year of HAPs) will be defaulted to the county centroid

The NET and NTI also requires that county FIPS codes be associated with point source facilities. EFIG will not accept any data that does not have an associated county FIPS code.

MACT codes are 4-digit codes that are assigned to all MACT categories and sub-categories. The MACT code is assigned to all facilities and processes within a MACT category. The MACT implementation code is then used to indicate control due to 112/129 standards. EFIG will assign MACT code and MACT implementation codes to facilities if codes are not assigned by S/Ls. The MACT codes will be assigned using the following methodology.

1. Facility lists provided by MACT engineers

2. SCCs that are applicable to MACT processes within a facility provided by MACT engineers
3. SIC codes that are applicable to specific MACT categories provided by MACT engineers

Table B-1 lists the data elements in the NEI Input Format. Each mandatory data element is coded “mandatory submission” (MS) or “data solvable” (DS). Data elements coded MS must be supplied by the S/L agency for EFIG to process the data. Data elements coded DS are elements that EFIG will add to the S/L data if they are not supplied by the S/Ls. Table B-1 also includes a brief description of the method EFIG will use to complete the necessary DS data elements. A more detailed description of how EFIG will solve for missing criteria emissions data and temporally allocate emissions follows Table B-1.

Table B-1: Point Source Data Completion Methods

NIF Table	Data Element	Element Type	Data Completion Method
(all)	Record Type	MS	
Activity	Process rate throughput		
Activity	Start date time	MS	Use 1/1/99 for annual data
Activity	End date time	MS	Use 12/31/99 for annual data
Activity	Throughput method		
Activity	Unit of measure for throughput		
Activity	Period average days per week		
Activity	Period average weeks per period		
Activity	Period average hours per day		
Activity	Period hours per period		
Control Equipment	Measurement type for total capture control efficiency		
Control Equipment	Control device type	MS if controls reported	
Control Equipment	Total capture control efficiency		
Emission Process	S/L Process ID	MS (for HAPS, MS if reporting data at process level)	
Emission Process	Process MACT implementation code	DS	ESD default

NIF Table	Data Element	Element Type	Data Completion Method
Emission Process	SCC	MS (criteria)	
Emission Process	Process MACT code	DS	ESD default
Emission Process	Process description		
Emission Process	Winter Throughput Pct	DS(criteria)	Default Temporal Allocation File
Emission Process	Spring Throughput Pct	DS(criteria)	Default Temporal Allocation File
Emission Process	Summer Throughput Pct	DS(criteria)	Default Temporal Allocation File
Emission Process	Fall Throughput Pct	DS(criteria)	Default Temporal Allocation File
Emission Process	Annual average days per week	DS(criteria)	Default to 7
Emission Process	Annual average weeks per year		
Emission Process	Annual average hours per day	DS(criteria)	Default to 24
Emission Process	Annual average hours per year		
Emission Release Point	S/L Stack ID	MS	
Emission Release Point	Exit gas velocity	DS	Default using existing inventories, SCC defaults, SIC defaults, model plant parameters, calculate if diameter and flow rate known
Emission Release Point	Exit gas temperature	DS	Default using existing inventories, SCC defaults, SIC defaults, model plant parameters
Emission Release Point	Exit gas flow rate	DS	Default using existing inventories, SCC defaults, SIC defaults, model plant parameters, calculate if velocity and diameter known
Emission Release Point	Nonstack vertical dimension		

NIF Table	Data Element	Element Type	Data Completion Method
Emission Release Point	Stack diameter	DS	Default using existing inventories, SCC defaults, SIC defaults, model plant parameters, calculate if velocity and flow rate known
Emission Release Point	Nonstack horizontal dimension		
Emission Release Point	Stack height	DS	Default using existing inventories, SCC defaults, SIC defaults, model plant parameters
Emission Release Point	Nonstack dimensional units		
Emission Release Point	Emission release point description		
Emission Release Point	Emission release point type	DS	Default to fugitive
Emission Release Point	Latitude or UTM Northing	DS	Match to current NTI, TRI or NET
Emission Release Point	Longitude or UTM Easting	DS	Match to current NTI, TRI, or NET
Emission Release Point	UTM Zone	MS if UTM coordinates supplied	
Emission Release Point	XY coordinate type (lat/long or UTM)	MS if coordinates supplied	
Emission Release Point	Stack fence line distance		
Emission Unit	S/L Emission Unit ID	MS (for HAPS, MS if reporting data at unit level)	
Emission Unit	Emission unit description		
Emission Unit	Emission unit type		
Emission Unit	ORIS ID for boiler	DS (for electric utilities)	From NET and Hg Utility Study
Emissions	Numerical value of emission estimate	MS	
Emissions	Control status	MS (HAPS)	If control device reported, assign controlled Assign unknown if blank
Emissions	Units of emission estimate	MS	

NIF Table	Data Element	Element Type	Data Completion Method
Emissions	Pollutant Code (CAS# or NTI pollutant code)	MS	
Emissions	Emission estimate type (actual, potential, etc.)	DS	For HAPS, assign unknown if blank. For criteria assign actual if blank
Emissions	Emission data level (facility, unit, process, stack)	MS	
Emissions	Emission estimate method code		
Emissions	Emission time period	MS	
Emissions	Rule effectiveness method		
Emissions	Rule effectiveness	DS	Default to 100
Site	S/L site ID	MS	
Site	Site Address Zip Code		
Site	NTI Site ID	MS (HAPS, Mandatory for facilities in 1996 NTI)	
Site	Source type (major, area, mobile)	MS	Assign using CAA definition
Site	Country FIPS	DS	Assume US
Site	SIC code	MS	
Site	Site MACT code	DS	ESD
Site	Site name	MS	
Site	Site description		
Site	Site MACT Implementation code	DS	ESD
Site	Contact name	MS	
Site	NAICS code	DS	Use SIC/NAICS lookup
Site	County FIPS	MS	
Site	State FIPS	MS	
Site	Site Address Street Line 1	MS	
Site	Site Address Street Line 2		
Site	Site Address Street Line 3		
Site	Site Address City	MS	

NIF Table	Data Element	Element Type	Data Completion Method
Site	NTI Unique Facility ID	DS	Use 1996 NTI
Site	Address type (corporate, site)	MS	
Site	Dun & Bradstreet #		none
Site	ORIS ID	DS	From NET and Hg Utility Study
Site	TRI ID	DS (HAPS)	
Site	Zip code	MS	
Transmittal	Contact alternate phone number		
Transmittal	Organization submitting data (S/L, ESD, TRI)	MS	
Transmittal	Contact name	MS	
Transmittal	Contact phone number	MS	
Transmittal	Contact fax number		
Transmittal	Contact email address	MS	
Transmittal	Transaction date	DS	Assume date of receipt if blank
Transmittal	Incremental submission #	MS	
Transmittal	Transaction comments		
Transmittal	Transaction type	MS	
Transmittal	Inventory year	DS	Assume 1999 if blank

## Emissions (Criteria)

### 1. Solve for Pollutants - Repeat Process for Each Pollutant Not Present

A. If annual emissions are known, solve annual emissions for other pollutants

Or

If daily emissions are known, solve daily emissions for other pollutants.

B. Perform plant level match to current NET data

If there is a match, distribute plant level emissions to segments in New (S/L submitted) NEI data based on segment level distribution of known pollutant (NO<sub>x</sub>) in New NEI Data

C. If there is no plant level match to current NET data

Develop emissions using uncontrolled emission factor ratios to calculate uncontrolled emissions. This method applies SO<sub>2</sub> or PM-10 ratios to NO<sub>x</sub>. NO<sub>x</sub> was the pollutant

utilized to calculate the ratio because (1) the types of sources likely to be important SO<sub>2</sub> and PM-10 emitters are likely to be similar to important NO<sub>x</sub> sources and (2) the generally high quality of the NO<sub>x</sub> emissions data. Ratios of SO<sub>2</sub>/NO<sub>x</sub> and PM-10/NO<sub>x</sub> based on uncontrolled emission factors were developed. These ratios were multiplied by uncontrolled NO<sub>x</sub> emissions to determine either uncontrolled SO<sub>2</sub> or PM-10 emissions.

PM-2.5 emission estimates were developed based on the PM-10 estimates using source-specific uncontrolled particle size distributions and particle size specific control efficiencies for sources with PM-10 controls. To estimate PM-2.5, uncontrolled PM-10 was first estimated by removing the impact of any PM-10 controls on sources in the inventory. Next, the uncontrolled PM-2.5 was calculated by multiplying the uncontrolled PM-10 emission estimates by the ratio of the PM-2.5 particle size multiplier to the PM-10 particle size multiplier. (These particle size multipliers represent the percentage to total particulate below the specified size.) Finally, controls were reapplied to sources with PM-10 controls by multiplying the uncontrolled PM-2.5 by source/control device particle size specific control efficiencies.

All NH<sub>3</sub> emission estimates remained unchanged from the estimates currently in the NET.

## 2. Solve for Temporal Emissions

Annual or Daily Emissions (whichever) is not known will be calculated using the following using operating schedule information (see equations 1a and 1b).

To calculate daily from annual -

$$EMIS_{ASD} = EMIS_{ANN} * SUMTHRU * (1/13 * DPW) \quad [1a]$$

To calculate annual from daily -

$$EMIS_{ANN} = EMIS_{ASD} / (SUMTHRU * (1/(13 * DPW))) \quad [1b]$$

Where:

- EMIS<sub>ANN</sub> = Annual Emissions
- EMIS<sub>ASD</sub> = Typical Summer Day Emissions
- SUMTHRU = Summer Throughput Percentage
- DPW = Days Per Week in Operation
- 13 = Number of Weeks in Summer

## Area Sources

Table B-2 lists the in the NEI Input Format. Each mandatory data element is coded “mandatory submission” (MS) or “data solvable” (DS). Data elements coded MS must be supplied by the S/L agency for EFIG to process the data. Data elements coded DS are elements that EFIG will add to the S/Ls data if they are not supplied by the S/L. Table B-2 also includes a brief

description of the method EFIG will use to solve for the necessary DS data elements. A more detailed description of how EFIG will solve for missing criteria emissions data and temporally allocate emissions follows Table B-2.

Table B-2: Area and Nonroad Mobile Source Data Completion Methods

NIF Table	Data Element	Element Type	Data Completion Method
all	Record type	MS	
Activity	Inventory Start date	MS	Use 1/1/99 for annual data
Activity	Inventory End date	MS	Use 12/31/99 for annual data
Control equipment	Control device type	MS if controls reported	
Control equipment	Total capture control efficiency		
Emissions	CAS # or NTI pollutant code for HAPs without CAS #s	MS	
Emissions	Pollutant name	DS	Assign based on CAS# or pollutant code
Emissions	Control status	MS	
Emission	Emission type (actual, potential, etc.)	DS	For HAPS, assume actual annual if blank. For criteria, assume actual if blank
Emissions	Numerical value of emission estimate	MS	
Emissions	Units of emission estimate (emission unit numerator in NIF)	MS	
Emission process	Category MACT code	DS	ESD



NIF Table	Data Element	Element Type	Data Completion Method
Emission process	Category MACT Implementation code	DS	ESD
Emission process	Emission process description		
Emission process	NAICS code	DS (if SIC supplied)	Use SIC/NAICS lookup
Emission process	SIC code		
Emission Process	SCC/AMS code	MS	
Transmittal	Contact name	MS	
Transmittal	Contact phone number	MS	
Transmittal	Contact fax number		
Transmittal	Contact email address	MS	
Transmittal	Contact alternate phone number		
Transmittal	County FIPS	MS	
Transmittal	County name	DS	Base on county FIPS
Transmittal	Incremental submission number	MS	
Transmittal	Inventory year	DS	Assume 1999 if blank
Transmittal	Organization submitting data	MS	
Transmittal	Source type (major, area, mobile)	MS	Assign using CAA definitions
Transmittal	State FIPS	MS	
Transmittal	Transaction comments		none
Transmittal	Transaction date	DS	Assume date of receipt if blank

NIF Table	Data Element	Element Type	Data Completion Method
Transmittal	Transaction type	MS	
Transmittal	Transmittal source (S/L, EPA)	DS	Assign based on source of data

## Solving for Missing Data - Emissions

### 1. Solve for Pollutants - Repeat Process for Each Pollutant Not Present

A. If annual emissions are known, solve annual emissions for other pollutants.

Or

If daily emissions are known, solve daily emissions for other pollutants.

B. Perform State/county/SCC level match to current NET data

If there is a match use emissions from current NET.

C. If there is State/county/SCC match to current NET data

Develop emissions using uncontrolled emission factor ratios to calculate uncontrolled emissions. This method used SO<sub>2</sub> or PM-10 ratios to NO<sub>x</sub>. NO<sub>x</sub> was the pollutant utilized to calculate the ratio because (1) the types of sources likely to be important SO<sub>2</sub> and PM-10 emitters are likely to be similar to important NO<sub>x</sub> sources and (2) the generally high quality of the NO<sub>x</sub> emissions data. Ratios of SO<sub>2</sub>/NO<sub>x</sub> and PM-10/NO<sub>x</sub> based on uncontrolled emission factors were developed. These ratios were multiplied by uncontrolled NO<sub>x</sub> emissions to determine either uncontrolled SO<sub>2</sub> or PM-10 emissions.

PM-2.5 emission estimates were developed based on the PM-10 estimates using source-specific uncontrolled particle size distributions and particle size specific control efficiencies for sources with PM-10 controls. To estimate PM-2.5, uncontrolled PM-10 was first estimated by removing the impact of any PM-10 controls on sources in the inventory. Next, the uncontrolled PM-2.5 was calculated by multiplying the uncontrolled PM-10 emission estimates by the ratio of the PM-2.5 particle size multiplier to the PM-10 particle size multiplier. (These particle size multipliers represent the percentage to total particulate below the specified size.) Finally, controls were reapplied to sources with PM-10 controls by multiplying the uncontrolled PM-2.5 by source/control device particle size specific control efficiencies.

All NH<sub>3</sub> emission estimates remained unchanged from the estimates currently in the NET.

### 2. Solve for Temporal Emissions

Annual or Daily Emissions (whichever) is not known will be calculated using the following using EPA's default temporal allocation factor file (TAFF) (see equations 2a and 2b). The TAFF contains national default temporal factors by SCC.

To calculate daily from annual -

$$EMIS_{ASD} = EMIS_{ANN} * (SUMFAC * WKDYFAC) \quad [2a]$$

To calculate annual from daily -

$$EMIS_{ANN} = EMIS_{ASD} / (SUMFAC * WKDTFAC) \quad [2b]$$

Where:

- EMIS<sub>ANN</sub> = Annual Emissions
- EMIS<sub>ASD</sub> = Typical Summer Day Emissions
- SUMFAC = Summer Season Factor from TAFF (by SCC)
- WKDYFAC = Summer Weekday Factor from TAFF (by SCC)

Table B-3: Mobile Source Onroad Data Completion Methods

NIF Table	Data Element	Element Type	Data Completion Method
all	Record type	MS	
Activity	Throughput (vmt)	MS	
Activity	Throughput Units Numerator	MS	
Emissions	CAS # or NTI pollutant code for HAPs without CAS #s	MS	
Emissions	Emission type	Nice to Have, DS	Assume actual annual emissions if blank
Emissions	Numerical value of emission estimate	MS	
Emissions	SCC/AMS code	MS	

NIF Table	Data Element	Element Type	Data Completion Method
Emissions	Units of emission estimate (emission unit numerator in NIF)	MS	
Transmittal	Contact name	MS	
Transmittal	Contact phone number	MS	
Transmittal	Contact fax number		
Transmittal	Contact email address	MS	
Transmittal	Contact alternate phone number		
Transmittal	County FIPS	MS	
Transmittal	Country FIPS	DS	Assume US
Transmittal	Incremental submission number	MS	
Transmittal	Inventory year	DS	Assume 1999 if blank
Transmittal	Organization submitting data	MS	
Transmittal	Source type (major, area, mobile)	MS	
Transmittal	State FIPS	MS	
Transmittal	Transaction comments		none
Transmittal	Transaction date	DS	Assume date of receipt if blank

NIF Table	Data Element	Element Type	Data Completion Method
Transmittal	Transaction type	MS	
Transmittal	Transmittal source (S/L, OTAQ)	DS	Assign based on source of data