NEI Quality Assurance and Data Augmentation for Point Sources

US EPA
Emissions Monitoring and Analysis Division
Emission Factor and Inventory Group
RTP, NC 27711

May 26, 2004

TABLE OF CONTENTS

List of Tables

SECTION 1. INTRODUCTION

The Emission Factor and Inventory Group (EFIG) in the Environmental Protection Agency (EPA) compiles the National Emission Inventory (NEI) for hazardous air pollutants (HAPs) and criteria air pollutants (CAPs). The NEI plays an important role in air quality management activities such as emission trends, rule and policy development and risk assessment. To support many of the these functions, the NEI must contain the data necessary for air quality modeling. The NEI for HAPs is compiled in order to determine if Clean Air Act (CAA) programs are successful in reducing emissions and human health and environmental risk due to HAPs emissions. Title I, Section 110 of the CAA requires states to submit emission inventories for CAPs as part of their State Implementation Plans.

The NEI contains estimates of facility-specific HAP and CAP emissions and their source-specific parameters necessary for modeling such as location and facility characteristics (stack height, exit velocity, temperature, etc.). Complete source category coverage is needed, and the NEI contains estimates of emissions from stationary point and nonpoint (i.e., stationary sources such as residential heating that are inventoried at the county level) and mobile source categories. The NEI contains individual stack and fugitive estimates at individual geocoordinates for point sources. County level estimates are provided in the NEI for nonpoint and mobile sources. Point source categories of HAPs include major and area sources as defined in Section 112 of the CAA. Non-point source categories of HAPs include area sources and other stationary sources that may be more appropriately addressed by other programs rather than through regulations developed under certain air toxics provisions (Sections 112 or 129) in the CAA.

The major steps involved in compiling the 2002 NEI include:

- Submittal of 2002 inventory data by state and local agencies, tribes, industry, and EPA offices;
- Blending/Merging of data from multiple data sources;
- Augmentation of data for missing data elements;
- QC/QA of data;
- Preparation of draft 2002 NEI for external and internal review;
- Incorporation of external and internal review comments on the draft 2002 NEI and incorporation of new inventory data submitted during review period; and
- Preparation of final 2002 NEI.

Important steps in preparing point source files are Quality Assurance (QA) and augmentation of data. QA and data augmentation are needed to prepare point source files for use in air quality and risk and exposure modeling. In addition, the QA of data and data augmentation are needed in order for the NEI to meet recent EPA data standards and Office of Management and Budget's (OMB) Information Quality Guidelines. This report focuses on the QC/QA and augmentation of data in the NEI point source inventory.

We conduct a variety of QA activities to identify pont source records with referential integrity problems, duplicate records, and records with missing or out-of-range parameters which are needed for air quality and exposure modeling. After conducting QA, we augment data using a methodology in this report. We summarize the errors found and reports back to the data providers on the QA findings. We tracks errors, their resolution and all communications on a QA/QC form, through emails, and in a phone log as part of documentation. These tracking mechanisms help to ensure the transparency and reproducibility of the NEI. This is a requirement of OMB Information Quality Guidelines, but it also helps us establish an electronic trail for each record in the NEI. We have created a QA/QC process and tracking database to provide feedback reports to data providers at regular intervals during the QA of the data. We archive all files submitted by data providers, records removed during QA, records augmented, and each iteration of the draft and final NEI.

We first resolve records with referential integrity problems and duplicates. Then after identifying parameters and data fields with missing or out-of-range values during the QA of files, we augment the data using the methodology in this document. The NEI data files identify all fields of data that have defaulted parameters. Data augmentation occurs at different times in the compilation of the NEI. For example, augmentation of location coordinates occurs prior to blending/merging of data, while augmentation of stack parameters occurs after blending/merging of data.

This document revises the previous memorandum, "NEI Quality Assurance and Data Augmentation Steps for Point Sources", June 2003.

Data providers submit their data to EPA using the NEI Input Format (NIF) or an extensible markup language (XML) format. Both formats are described and located at http://www.epa.gov/ttn/chief/nif/index.html. The 2002 point source data will be submitted to EPA in NIF V3.0 or NEI XMLV3.0 for incorporation into the 2002 NEI. The point source NIF V3.0 and NEI XML V3.0 contain many data fields. Table 1 describes the eight tables in the NIF for point sources.

Table 2 presents a list of all fields in the NIF V3.0 Point Source Tables. Table 2 indicates if the field has a methodology for data augmentation in this document. For all fields of data in Table 2 that do not have a data augmentation procedure, data submitters will need to populate the fields of data. Table 2 also indicates those fields of data in which we will only conduct QA and will be unable to default invalid values.

Table 3 shows additional data fields in the NEI that are not included in the NIF V3.0 Tables. Data for all of these fields will be generated by EFIG. These fields may be in auxiliary files or may be included in NIF tables as output data fields.

The QA and Augmentation Procedure for fields of data in Tables 2 and 3 are presented in the Sections 2 - 14 of the report. Fields of data have been grouped by function into the following sections of the report.

Section 2	Inventory information
Section 3	Geographic information
Section 4	Facility information
Section 5	Geocoordinates
Section 6	Unit and Process IDS
Section 7	Source Description
Section 8	Emission Release Description and Parameters
Section 9	Unit Codes
Section 10	Activity
Section 11	Control Equipment
Section 12	Emissions
Section 13	Other Fields that undergo QA but are not augmented
Section 14	Default Flags

Any comments or questions concerning this report should be sent to Anne Pope, pope.anne@epa.gov.

Table 1. NIF V3.0 Point Source Table Description

NIF Table	Abbreviation	Description
Transmittal	TR	organization submitting to EPA, data year, source types
Site	SI	plant location and IDs
Emission Unit	EU	unit or device generating point emissions
Emission Process	EP	operational process generating emissions
Emission Period	PE	activity data and time period for releases
Emission Release Point	ER	mechanism releasing emissions to the air including stack parameters and geographic coordinates
Control Equipment	CE	control equipment and efficiency
Emission	EM	amount of emissions per pollutant

Table 2. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
PE	Actual Throughput	Throughput of process activity		
TR	Affiliation Type	Report certifier NIF Code Table	Augmentation Procedure	2.7
EP	Annual Average Days per Week	Average days per week process is operating	Augmentation Procedure	10.4
EP	Annual Average Hours per Day	Average hours per day process is operating	Augmentation Procedure	10.5
EP	Annual Average Hours per Year	Average hours per year process is operating	Augmentation Procedure	10.6
EP	Annual Average Weeks per Year	Average weeks per yr process is operating	Augmentation Procedure	10.7
EP	Ash Content	Ash content of fuel		
SI	City	City of physical location	Augmentation Procedure	4.8
TR	Contact Person Name	EPA's single pont of contact regarding questions on inventory submittal		
TR	Contact Phone Number	Phone number for contact person		
EM	Control Status	Controlled or Uncontrolled	Augmentation Procedure	11.1

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
CE	Control System Description	Description of control devices		
ER	Coordinate Data Source Code	Data source providing latitude and longitude NIF Code Table	Augmentation Procedure	5.7
SI	Country	Country Name	Augmentation Procedure	3.5
EU	Design Capacity	Operational Capacity for emission unit	Augmentation Procedure	9.2
EU	Design Capacity Denominator	Unit of measure for design capacity NIF Code Table - Unit codes	Augmentation Procedure	9.1 9.2
EU	Design Capacity Numerator	Unit of measure for design capacity NIF Code Table - Unit codes	Augmentation Procedure	9.1 9.2
SI	Dun and Bradstreet Number	Dun and Bradstreet number for facility	Augmentation Procedure	4.14
EM	EF Reliability Indicator	AP-42 Rating NIF Code Table	QA	13
TR	Electronic Address Text	Email address		
TR	Electronic Address Type Name	Email, Internet, Intranet, HTTP, ftp, etc NIF Code Table		
EM	EM Reliability indicator	Quality of emissions estimate		

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
EM	Emission Calculation Method Code	How were emissions derived? NIF Code Table	QA	13
EM	Emission Data level	Level HAP emissions reported at - site, unit, stack, process	Augmentation Procedure	6.1
EM	Emission Numeric Value		Augmentation Procedure	12
EP	Emission Process Description	Emission process description		
ER	Emission Release Point Description			
EP, ER, EM	Emission Release Point ID	ID for point/location where emissions are released to the air	Augmentation Procedure	6.4
ER	Emission Release Point Type	Code for physical configuration of the release point NIF Code Tables	Augmentation Procedure	8.1
EM	Emission Type	Temporal description of emissions reported (annual, weekday, etc.) NIF Code Table	Augmentation Procedure	10.1
EU	Emission Unit Description	Emission unit description		
EU, EP, PE, CE, EM	Emission Unit ID	ID assigned to identify emission units within a facility	Augmentation Procedure	6.2

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
EM	Emission Unit Numerator	Unit of measure for emissions NIF Code Table - Unit codes	Augmentation Procedure	9.1
PE, EM	End Date	End date of period of emissions	Augmentation Procedure	10.2
PE, EM	End Time	End time of period	Augmentation Procedure	10.3
ER	Exit Gas Flow Rate	Flow rate of an exit gas (cu ft/sec)	Augmentation Procedure	8.2
ER	Exit Gas Temperature	Temperature of an exit gas stream (degrees F)	Augmentation Procedure	8.2
ER	Exit Gas Velocity	Velocity of an exit gas (ft/sec)	Augmentation Procedure	8.2
SI	Facility Category	Major or Area NIF Code Table	Augmentation Procedure	7.7
SI	Facility Name	Name of Facility		
SI	Facility Registry ID	EPA Facility Registry ID assigned by EPA's OEI	Augmentation Procedure	4.12
EM	Factor Numeric Value			
EM	Factor Unit Denominator	Unit of measure for emission factor denominator NIF Code Table - Unit codes	Augmentation Procedure	9.1
EM	Factor Unit Numerator	Unit of measure for emission factor numerator NIF Code Table - Unit codes	Augmentation Procedure	9.1

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
EP	Fall Throughput Percent	Percentage of process that operates during the fall months (Sept,Oct, Nov)	Augmentation Procedure	10.12
TR	Format Version	NIF version	Augmentation Procedure	2.8
CE	Fourth Device Type Code	Fourth control device NIF Code Tables	Augmentation Procedure	11.2
EM	HAP Performance Level	Actual, Allowable, Maximum, Potential NIF Code Table	Augmentation Procedure	12.5
EP	Heat Content	Heat content of fuel in MMBTU per tons of coal, per 1000 barrels of oil, or million SCG of natural gas		
ER	Horizontal Accuracy Measure	Measure in meters of accuracy of the latitude and longitude coordinates	Augmentation Procedure	5.7
ER	Horizontal Area Fugitive	Release height above terrain of fugitive releases		
ER	Horizontal Collection Method Code	Method used to determine latitude and longitude NIF Code Table	Augmentation Procedure	5.7
ER	Horizontal Reference Datum Code	Reference date used to determine latitude and longitude coordinates NIF Code Table	Augmentation Procedure	5.7
TR	Incremental Submission Number	Number of submission		

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
TR	Inventory year	Year of inventory	Augmentation Procedure	2.4
TR	Inventory Type Code	Does data contain HAPs, CAPs or both? NIF Code Table	Augmentation Procedure	2.5
SI	Location Address	Physical location of the main entrance of facility	Augmentation Procedure	4.7
PE, EM	Material	Type of material processed NIF Code Tables	QA	13
PE, EM	Material I/O	Is material input or output? NIF Code Tables	QA	13
EU	Maximum Nameplate Capacity	Rated design capacity at 100% operation		
SI, EU	NAICS Code	North American industry Classification Code NIF Code Table	Augmentation Procedure	7.1
SI	NTI Site ID	1996 NTI ID		
TR	Organization Name	Name of organization submitting data		
EU	ORIS Boiler ID	Unique ID for electric generating facilities		
SI	ORIS Facility Code	Unique ID for electric generating facilities	Augmentation Procedure	4.11
CE	Pct Capture Efficiency	Percent capture efficiency of control devices		

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
PE	Period Days per Week	Average number days per week process is active within period; Parameter is applied when emission period is less than annual	Augmentation Procedure	10.8
PE	Period Hours per Day	Average number hours per day process is active within period; Parameter is applied when emission period is less than annual	Augmentation Procedure	10.9
PE	Period Hours per Period	Average number hours per period process is active within period; Parameter is applied when emission period is less than annual	Augmentation Procedure	10.10
PE	Period Weeks per Period	Average number weeks per period process is active within period; Parameter is applied when emission period is less than annual	Augmentation Procedure	10.11
CE, EM	Pollutant Code	Pollutant Code NIF Code Table	Augmentation Procedure	12.1
CE	Primary Device Type Code	Primary control device NIF Code Tables	Augmentation Procedure	11.2
CE	Primary PCT Control Efficiency	Percent effectiveness of primary control system		
EP, PE, CE, EM	Process ID	ID assigned to identify emission processes within a facility	Augmentation Procedure	6.3

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
EP	Process MACT Code	Maximum Achievable Control Technology source category code NIF Code Tables	Augmentation Procedure	7.4
EP	Process MACT Compliance Code	Major/Area classification under Sections 112 and 129 NIF Code Table	Augmentation Procedure	7.6
TR, SI, EU, EP, PE, ER, CE, EM	Record Type	Type of NIF record	Augmentation Procedure	2.1
ER	Reference Point Code	Place where latitude and longitude are established (individual stack, plant entrance, plant centroid, etc.) NIF Code Table	Augmentation Procedure	5.7
TR	Reliability Indicator	Score for quality of inventory		
EM	Rule Effectiveness	Measure of percent effectiveness of control device		
EM	Rule Effectiveness Method	Rule effectiveness method NIF Code Table	QA	13
EP	SCC	Source Classification Code NIF Code Tables	Augmentation Procedure	7.3
CE	Secondary Device Type Code	Secondary control device NIF Code Tables	Augmentation Procedure	11.2

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
SI, EU	SIC Code	Standard Industrial Classification Code - being replaced by NAICS NIF Code Tale	Augmentation Procedure	7.2
SI	Site Description	Comments on facility		
ER	Source Map Scale Number	Map scale	Augmentation Procedure	5.7
TR	Source type	Point NIF Code Table	Augmentation Procedure	2.6
EP	Spring Throughput Percent	Percentage of process that operates during the spring months (Mar, Apr, May)	Augmentation Procedure	10.12
ER	Stack Diameter	Stack Diameter (ft)	Augmentation Procedure	8.2
ER	Stack Fence line Distance	Distance from stack to fence line in ft		
ER	Stack Height	Stack Height (ft)	Augmentation Procedure	8.2
PE, EM	Start Date	Start date of period of emissions	Augmentation Procedure	10.2
PE, EM	Start Time	Start time of period	Augmentation Procedure	10.3
SI	State	State of physical location	Augmentation Procedure	4.9

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
TR, SI, EU, EP, PE, ER, CE, EM	State and County FIPS	FIPS code for the state and county NIF Code Table	Augmentation procedure	3.1
SI, EU, EP, PE, ER, CE, EM	State Facility ID	ID assigned by data submitter to identify a facility		
SI, EU, EP, PE, ER, CE, EM	Submittal Flag	Submittal Status of record - O, A, D, R	Augmentation Procedure	2.2
EP	Sulfur Content	Sulfur content of fuel		
EP	Summer Throughput Percent	Percentage of process that operates during the summer months (June, July, Aug)	Augmentation Procedure	10.12
TR	Telephone Number Type Name	Fax, Mobile, Pager, Office NIF Code Table	QA	13
CE	Third Device Type Code	Third control device NIF Code Tables	Augmentation Procedure	11.2
PE	Throughput Unit Numerator	Unit of measure for throughput NIF Code Table - Unit codes	Augmentation Procedure	9.1
CE	Total Capture Control Efficiency	Collective efficiency for all controls	Augmentation Procedure	11.3
TR	Transaction Comments	General comments regarding transmittal		

Table 2 continued. Fields in NIF V3.0 Point Source Tables

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
TR	Transaction Creation Date	Creation date of transmittal data	QA	13
TR	Transaction Type	Original or Revision NIF Code Table	Augmentation Procedure	2.3
SI	TRI ID	TRI ID for facility	Augmentation Procedure	4.13
TR, SI, EU, EP, PE, ER, CE, EM	Tribal Code	Codes for American Indian tribes and Alaskan Native entities NIF Code Table	Augmentation Procedure	3.2
ER	UTM Zone	Zone number	Augmentation Procedure	5.4
EP	Winter Throughput Percent	Percentage of process that operates during the winter months (Dec, Jan, Feb)	Augmentation Procedure	10.12
ER	X Coordinate	Longitude (- decimal degrees)	Augmentation Procedure	5.2
ER	XY Coordinate Type	Type of coordinate system NIF Code Table	Augmentation Procedure	5.5
ER	Y Coordinate	Latitude (+ decimal degrees)	Augmentation Procedure	5.3
SI	Zip Code	Zip Code of physical location	Augmentation Procedure	4.10

Table 3. Additional Point Source Data Fields in the NEI that are not in the NIF V3.0

Field Name	Field Description	NIF Table Relationship	Report Section
Annual Emissions (tons)	Create annual emissions in tons	EM	12.9
Coordinate Default Methodology	How latitude/longitude coordinates are defaulted	ER	5.6
Data Source	Source of Data	EM	12.12
Default Flags	Default flags added for data fields augmented	all	14
Emission Data Quality Rating	Qualitative rating for emission estimate	EM	12.11
MACT Code Assignment	Method used to assign MACT Codes	ER	7.5
MSA	Metropolitan Statistical Area	SI	3.4
NAICS/SIC Code Default Methodology	How NAICS and SIC Codes are defaulted	SI, EU	7.1
NEI Facility ID	NEI Facility ID assigned to all facilities in NEI - needed for merged HAP and CAP	SI, EU, EP, PE, ER, CE, EM	4.1
NEI Site Latitude	Latitude of Facility	SI	5.8
NEI Site Longitude	Longitude of Facility	SI	5.9
NEI Standardized Facility Name	Facility Name associated with NEI Site ID	SI	4.2
NEI Standardized Address	Address associated with NEI Site ID	SI	4.3
NEI Standardized City	City associated with NEI Site ID	SI	4.4
NEI Standardized State	State associated with NEI Site ID	SI	4.5
NEI Standardized Zip Code	Zip Code associated with NEI Site ID	SI	4.6

Table 3, continued. Additional Point Source Data Fields in the NEI that are not in the NIF V3.0

Field Name	Field Description	NIF Table Relationship	Report Section
Period Type	Are emissions annual or non-annual?	EM	12.8
Stack Latitude	Convert Coordinates to Latitude if in UTM	ER	5.7
Stack Longitude	Convert Coordinates to Longitude if in UTM	ER	5.7
Stack Parameter Default Flag	Indicates which stack parameters have been defaulted	ER	8.3
Submittal date	Represents the date of submittal to show the date of data in each record	SI, EU, EP, PE, ER, CE, EM	2.9
Total Capture Control Efficiency Default Flag	Indicates how total capture control efficiencies have been defaulted	CE	11.4
Urban/Rural Flag	Is county Urban or Rural using Section 112k/112c3 definition - U1, U2, R	SI	3.3

SECTION 2. INVENTORY INFORMATION

This section includes the following data fields.

- Record Type
- Submittal Flag
- Transaction Type
- Inventory Year
- Inventory Type Code
- Source Type
- Affiliation Type
- Format Version
- Submittal Date

2.1 Record Type

The Record Type indicates the type of record in each table. The Record Type for each record within a table is the table abbreviation (TR, SI, EU, EP, PE, ER, CE, EM). If the Record Type is missing or invalid, we will default the field based on the type of table.

2.2 Submittal Flag

The Submittal Flag indicates the status of each record in the NIF tables. Acceptable values include the following codes.

O - Original
A - Add
D - Delete
R - Revise

If the Submittal Flag is missing in the original submission of data in June 2004, we will populate the field with "O". If the Submittal flag is missing in the revision of data in May 2005, we will populate the field for new records with "A" or "RA" for revised records. The deleted records would not appear in the final NEI tables as these records would have been removed.

2.3 Transaction Type

The Transaction Type indicates if the inventory is an original submittal or a revision to the draft 2002 NEI. Acceptable values include the following codes.

00 - Original
05 - Replace

If this field is missing in the original submission of data in June 2004, we will populate the field with "00". If the Submittal flag is missing in the revision of data in May 2005, we will populate the field for new records with "05" for revised inventory.

2.4 Inventory Year

The Inventory Year indicates the year of the inventory submitted to EPA. If the Inventory Year is not included in the Transaction Table, then we will assume that the year of the data is 2002 and augment this data field.

2.5 Inventory Type Code

The Inventory Type indicates whether the submittal contains CAPs, HAPs or both HAPs and CAPs. Acceptable values include the following codes.

NIF Code	Pollutants
CRIT:	CAPs
HAP:	HAP
CRITHAP:	CAPs and HAPs

We will verify that the inventory type matches the pollutants submitted in the Emissions Table. If the code is in error, we will correct the Inventory Type Code.

2.6 Source Type

The Source Type indicates whether the submittal contains point, nonpoint, onroad or nonroad records. The Source Type should be coded as "POINT" for point source submittals. If Source Type is missing or in error, we will populate this field with "POINT".

2.7 Affiliation Type

The only valid code for Affiliation Type is "Report Certifier" in the NIF. This indicates that the Contact Person identified in the Transmittal Table is EPA's point of contact to answer questions about the data submitted. If Affiliation Type is blank or in error, we will populate this field with "Report Certifier".

2.8 Format Version

The Format Version is the Version of NIF that is used to format and submit NIF files. For the 2002 NEI, the format version should be reported as "3.0". If the Format Version is blank or in error, we will augment this field as "3.0".

2.9 Submittal Date

We will add this field to all tables to indicate the date of the data for each record. For example for the emissions table, the draft emissions estimate would contain a date of 20041001. If the estimate is not revised, the submittal date would remain 20041001. If the estimate was revised during the external review period, the date would be revised to 20050601 in the final inventory. As revisions are made in time, then the submittal date would be revised to reflect date of revision.

SECTION 3. GEOGRAPHIC INFORMATION

This section includes the following data fields.

- State/County FIPS Code
- Tribal Code
- Urban/Rural flag
- MSA
- Country

EFIG will prepare a State/County table with the following data fields.

- EPA Region
- State/County FIPS Code
- State Abbreviation
- State Name
- County Name
- Urban/Rural Classification
- MSA

This table will be used in various summary and look-up tables associated with the 2002 NEI.

3.1 State and County FIPS Codes

State and County FIPS codes are used by a variety of organizations to identify the state and county within databases. The State and County FIPS Code is a mandatory data field in the NEI. The State and County FIPS code represents the county or county equivalent and the state or state equivalent of the United States.

For State and County FIPS codes, we will first compare the State and County FIPS Codes reported by data submitters to 2002 State/County FIPs NIF Code Table and identify records with errors in FIPS code. 2002 revisions to FIPS codes in the NEI include the following.

CO: Adding FIPS code for Broomfield County, CO (08014). Broomfield County (08014) was created from parts of Adams (08001), Boulder (08013), Jefferson (08059), and Weld (08123) counties.

VA: Removing independent city (county-equivalent) of Clifton Forge, VA (51560). Clifton Forge reverted to town status and is now part of Alleghany County, VA (51005) rather than a separate county-equivalent surrounded by Alleghany County.

For invalid FIPS codes in CO, we will contact the data providers and correct invalid FIPS codes after consultation with the data provider. For invalid VA FIPS Code in Clifton Forge, we will augment data by deleting the old FIPS code and replacing it with the FIPS code for Alleghany County.

If a facility has records with valid FIPS codes , we will default records having missing or invalid FIPS

code with the FIPS code from the records within the facility having valid FIPS code. For facilities that do not have any records with valid FIPS codes, we will first use the NEI Historical Facility Table to see if the facility is present in the table. If it is present, we will default the State and County FIPS code to the FIPS Code in the NEI Historical Facility Table.

For facilities missing State and County FIPS Codes that are not in the NEI Historical Facility Table, we will contact the data submitter to correct the FIPS code.

3.2 Tribal Code

The Tribal Code, a mandatory data field in NIF, represents American Indian Tribes and Alaskan Native entities. The tribal code is used to identify facilities that are located on tribal lands.

For Tribal Codes, we will compare the Tribal Codes reported by data submitters to 2002 NIF Tribal Code Table and identify records with invalid or missing Tribal Codes. For invalid or missing Tribal Codes, we will first use the NEI Facility File and see if the facility is present in the file. If it is present, we will default the Tribal Code using the NEI Historical Facility File. For facilities that are not in the NEI Historical Facility File, we will augment Tribal Codes for point sources in the NEI by mapping site latitude and longitude of all facilities and identifying facilities on tribal lands using Tribal GIS, a GIS software which contains boundaries for federally recognized tribes. We will then augment Tribal Codes for facilities on tribal lands in the NEI that do not have a Tribal Code after tribes review the Tribal Code assignment during the review of the draft NEI. We will specifically ask tribes to review the assignment of Tribal Codes to facilities during the external review process.

Because the Tribal Code is a primary key in the NEI, we will default nontribal data to code “000".

3.3 Urban/Rural Flag

The Integrated Urban Air Toxics Strategy is an important part of EPA’s national air toxics program. The NEI is used in analyses relating to Section 112(k) and Section 112(c)(3) requirements in the CAA. Please note that the definition of "urban"does not necessarily apply for regulatory or implementation purposes. For more information on EPA’s Urban Air Toxics Program, go to the following site, www.epa.gov/ttn/atw/urban/urbanpg.html .

We will classify all counties in the 2002 NEI as Urban-1, Urban-2, or Rural, based on 2002 Census data. The classifications are defined as:

- Urban-1 (U1): Counties included in an MSA with MSA population greater than 250,000 (for counties associated with more than one MSA in the six New England states, the largest MSA population was used);

- Urban-2 (U2): Counties in which the U.S. Census Bureau designates more than 50 percent of the total county population as urban; and
- Rural (R): Counties in which the U.S. Census Bureau designates less than 50 percent of the total county population as urban. Counties in which the urban population exactly equaled the rural population were classified as R.

The following population and geographic information available from the U.S. Census Bureau (via their Internet web page) will be used in the U1/U2/R classifications for each of the counties defined in the 2002 NEI

1. Total county, Minor Civil Division (MCD), and incorporated place population estimates (U.S. Census Bureau).

2. Total Metropolitan Statistical Area (MSA) population, and county and place populations associated with each MSA (U.S. Census Bureau).

Urban-1 (U1) counties are included in an MSA with MSA population greater than 250,000.

The U.S. Census Bureau does not estimate county urban population between censuses (1990 and 2000). Therefore a surrogate for county urban population will be calculated by summing the population from all incorporated places within a county with population 2,500 or more (the only part of the urban area definition for which 2002 estimates are available). If over 50% of the county population live in incorporated places with population 2,500 or more, the county will be classified as Urban-2 (U2).

Rural (R) counties have a surrogate urban population that is less than 50 percent of the total county population.

Counties will be classified for 2002 as U1, U2, or R using the same criteria as used for previous classifications. For counties in the six New England states that are associated with more than one MSA, the MSA with the largest population will be assigned for use in the U1/U2/R classification process. Table 4 shows the MSAs that were assigned in the 1999 NEI for these areas.

3.4 MSA

We will identify the MSA of counties in the 2002 NEI.

3.5 Country

Country is not a mandatory data field in NIF. If NEI data are provided to other countries, then it is useful to populate this data field. The Country name is the name that represents a primary geopolitical unit of the world. The default for the 2002 NEI is "USA".

Table 4. MSAs Assigned to Counties in New England States Associated with More Than One MSA for Use In The County U1/U2/R Classification Process

FIPS State Code (State Abbreviation)	FIPS County Code	County Name	FIPS MSA Code (1999)	Metropolitan Statistical Area Name
09 (CT)	005	Litchfield County	5602	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMSA
	007	Middlesex County	5602	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMSA
	011	New London County	3280	Hartford, CT MSA
	015	Windham County	1122	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA
23 (ME)	031	York County	1122	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA
25 (MA)	005	Bristol County	1122	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA
	013	Hampden County	1122	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA
44 (RI)	009	Washington County	6480	Providence-Fall River-Warwick, RI-MA MSA

SECTION 4. FACILITY INFORMATION

This section includes the following data fields.

- NEI Facility ID
- NEI Standardized Facility Name
- NEI Standardized Address
- NEI Standardized City
- NEI Standardized State
- NEI Standardized Zip Code
- ORIS Facility Code
- Location Address
- City
- State
- Zip Code
- Facility Registry ID
- TRI ID
- Dun and Bradstreet Number

The 2002 NEI will combine HAPs and CAPs into a single inventory. In order to compile a merged file, we will prepare a NEI Historical Facility Table. For each facility in the NEI, the table currently contains the following data fields.

- State/County FIPS Code
- Tribal Code
- NEI Facility ID
- NEI Standardized Facility Name
- NEI Standardized Address
- NEI Standardized City
- NEI Standardized State
- NEI Standardized Zip Code
- NEI Site Latitude
- NEI Site Longitude
- NEI Primary SIC Code/NAICS code
- 2002 Preliminary NEI for CAPs State Facility ID
- 1999 NEI for HAPS Unique Facility ID
- 1999 NEI for HAPs State Facility ID
- 2002 Preliminary NEI for CAPs site name
- 1999 NEI for HAPs site name
- 2002 Preliminary NEI for CAPs Location Address
- 1999 NEI for HAPs Location Address
- 2002 Preliminary NEI for CAPs City
- 1999 NEI for HAPs Location City
- 2002 Preliminary NEI for CAPs State

- 1999 NEI for HAPs State
- 2002 Preliminary NEI for CAPs Zip Code
- 1999 NEI for HAPs Zip Code
- 2002 Preliminary NEI for CAPs site latitude and longitude
- 1999 NEI for HAPs site latitude and longitude
- 2002 Preliminary NEI for CAPs primary SIC Code/NAICS code
- 1999 NEI for HAPs primary SIC Code/NAICS code
- 2002 ORIS Facility Code
- FRS ID
- FRS Facility Name
- FRS Address
- FRS City
- FRS State
- FRS Zip Code
- FRS Latitude/Longitude
- FRS State/County FIPS Code
- TRI ID
- Dun and Bradstreet Number
- Latitude and Longitude EPA data Standards fields
- Historical Names for Facility and Ownership

When the 2002 data submittals are received in June 2004, the table will be updated to include 2002 information for facilities in the current table and to add new data for facilities not in the table. The augmentation of the facility data fields in the NEI Historical Facility Table and in the NIF tables is discussed below.

4.1 NEI Facility ID

The NEI Facility ID provides a single unique ID for each facility in the NEI. Data providers sometimes use different site IDs for CAPs and HAPs emitted from the same facility. The NEI also merges data provided by different sources for pollutants emitted from individual processes within a facility . The NEI Facility ID is assigned to facility records submitted from different data sources in order to detect duplicate estimates among the records. The NEI Facility ID is critical to blending HAP and CAP emissions from different data sources and for augmenting VOC and PM emissions from HAP emission estimates.

We will first determine if a facility is in the NEI Historical Facility Table. If it is in the existing NEI Historical Facility Table, then we will use the existing NEI Facility ID. If the facility is not in the table, then we will assign a NEI Facility ID and add all of the information for the facility to the NEI Historical Facility Table. This information will be used to augment other data fields associated with facilities.

We have matched facilities in the 2002 Preliminary NEI for CAPs and in the 1999 NEI for HAPs and prepared a crosswalk with between the two data sets. We used a combined automated and manual

approach for matching facilities in the 2002 Preliminary NEI for CAPs and the 1999 NEI for HAPs. We developed a NEI Historical Facility Table, which will be used as a starting point for the 2002 NEI. This table will evolve into a historical facility database where changes in names/ownerships and subsidiaries will be tracked over time. We would like to add available historical data prior to 2002 to track facilities in earlier inventories but do not know if this is possible at this time.

We will use the same approach as used to match 99 HAP and 2002 Preliminary CAP facilities for facilities that are submitted in June 2004 that are not in the NEI Historical Facility Table. The facility matching process is described below.

The process of matching facilities in the point source NEI consists of two activities:

- matching facilities in the NEI and assigning a common ID, and
- matching NEI facilities to the Facility Registry System.

Matching Facilities in the NEI

An automated facility-matching program is first run to identify the best matches of common facilities using state/county FIPS code, facility name, facility address and site latitude/longitude coordinate pairs (a site average of individual emission release point coordinates). The automated matching assigns a common ID to facilities if they matched on all of these parameters. The program then looks for additional high probability matches by matching exactly on all parameters except the site latitude/longitude coordinate pairs. In this case, the program searches first for coordinate pairs located within 0.001 degrees of one another. If no match is found, the program interactively relaxes this condition for additional matches until it reaches an upper bound of coordinates that are 0.1 degrees apart. 0.1 degrees was found empirically to be the limit of 100% confidence in automated matching for the 1999 NEI. The procedure starts with exact matches (0 degrees difference) and works up to 0.1 degrees in increments of 0.0001 degrees to insure that the closest facility is always matched in the case where multiple plants may exist. However, some facilities that might be colocated within 0.1 degrees are not automatically matched if there is more than one potential match between facilities. The automated facility matches are compiled into a facility file and removed from the facility universe.

For the remaining unmatched facilities, an algorithm is applied that strips out punctuation and leading/trailing spaces, drops insignificant punctuation, and standardizes corporate names, from the facility name. Additional searches using these standardized names are then conducted by: comparing facility names on a case sensitive basis, identifying similar sounding facility names in each county, and matching on similar name and address and exact site latitude/longitude coordinate pairs. After using exact coordinate pairs, the program identifies additional potential matches using similar names, similar address, and similar coordinate pairs with latitude/longitude varying from 0.001 to 0.1 degrees. The matching stops when any ambiguity is found. The program generates candidate lists of potential matches for manual review. A number of these candidate lists are generated by using an artificial intelligence technique called heuristic fuzzy pattern matching to match facility names and addresses with small typographical differences.

Candidate lists of potential facility matches are then prepared and reviewed manually to determine if a valid match exists. Examples of candidate lists include the following.

- Same state/county FIPS code, facility name and address, but coordinates differ by > 0.1 degrees
- Same state/county FIPS code, facility name and coordinates, but different or missing address
- Same state/county FIPS code, address and coordinates, but different facility name
- Same state FIPS code, facility name, address and coordinates, but different county FIPS code
- First 5 letters of facility name are same with same state/county FIPS code, address and coordinates

The same type of candidate lists are also generated for potential "fuzzy" facility name and address matches. Candidate lists also include SIC codes.

These candidate lists are evaluated one at a time in order of highest to least certainty. To verify matching of facilities in the candidate lists, we researched individual companies, their locations, acquisition history and name changes on the internet. We consulted several external databases including company websites, Mapquest, internet yellow pages, TRI explorer, FRS, and business directories. These web searches help to identify closures and matches. The person reviewing the candidate lists verifies that the list has been manually validated and then compiles matches from each candidate facility list into the file with automated facility matches.

Assigning FRS IDS

EPA is required to include Facility Registry System (FRS) IDS in the NEI in order to meet the new EPA data standards for Facility Identification. (For more information on EPA data standards, go to: http://oaspub.epa.gov/edr/epastd$.startup) The standards relevant to the NEI are the: SIC/NAICS, Latitude/Longitude, Chemical Identification, Facility Identification and Contact Data Standards. To implement the Facility Identification standard, EPA must map NEI facilities to the FRS facilities maintained by EPA's Office of Environmental Information (OEI). The assignment of FRS IDS to NEI facilities involves the following steps.

- EFIG submits NEI facility list to OEI
- OEI assigns draft FRS IDS to NEI facilities
- EFIG reviews OEI's assignment of draft FRS IDS to NEI facilities
- EFIG and OEI discuss differences in facility matching and resolve data problems
- OEI finalizes FRS ID assignment to NEI facilities
- EFIG incorporates FRS ID into NEI facility file

The process of assigning FRS ID to the 2002 NEI will not occur until after the external revisions are received and the final list of facilities in the 2002 NEI is compiled.

4.2 NEI Standardized Facility Name

The NEI Standardized Name provides a single public or commercial name for a facility. The Facility Name provided by different data submitters for the same facility may differ in HAP and CAP NIF Site

Tables. This often occurs because of the use of subsidiary names, abbreviations, or old names that do not reflect current name and ownership of facilities in inventories. We will use the NEI Historical Facility Table to augment this data field.

4.3 NEI Standardized Address

The NEI Standardized Address provides a single physical address for a NEI facility. The Location Address provided by different data submitters for the same facility may differ in HAP and CAP NIF Site Tables. This often occurs because a road may be identified using a street name and as a highway number in different inventories. It also occurs when addresses change for a facility. We will use the NEI Historical Facility Table to augment this data field.

4.4 NEI Standardized City

The NEI Standardized City provides a single name of the city, town, or village or other locality in the physical address for a facility. The City provided by different data submitters for the same facility may differ in HAP and CAP NIF Site Tables. This sometime occurs if a new city/town is incorporated or city limits are extended and a facility has new address and inventory files are not updated. We will use the NEI Historical Facility Table to augment this data field.

4.5 NEI Standardized State

The NEI Standardized State data field indicates the state associated with the NEI Facility ID. We will use the NEI Historical Facility Table to augment this data field.

4.6 NEI Standardized Zip Code

The NEI Standardized Zip Code data field indicates the zip code associated with the NEI Facility ID. The Zip Code provided by different data submitters for the same facility may differ in HAP and CAP NIF Site Tables. This sometime occurs when new zip codes are assigned by the US Postal Service and the inventory files are not updated. We will use the NEI Historical Facility Table to augment this data field.

4.7 Location Address

The Location Address describes the physical (geographic) location of the front door or main entrance of a facility site, including urban-style street address or rural address. It is needed to default location coordinates for a facility if they are missing or invalid. If the Location Address is missing and the facility is in the NEI Historical Facility Table, we will use the NEI Standardized Address to default the Location Address.

4.8 City

The City is the name of the city, town, or village or other locality of the physical location of the facility. The City is needed to verify the reported zip code. The zip code is needed to default location coordinates for a facility if they are missing or invalid. If the City is missing and the facility is in the NEI Historical Facility Table, we will use the NEI Standardized City to default this field.

4.9 State

The State is the state abbreviation of the physical location of the facility. The state abbreviation can be used to QA the Zip Code, State and County FIPS Code, and Latitude and Longitude.

We will QA the State by comparing it to the State and County FIPS Code. If the State is missing or invalid and the facility is in the NEI Historical Facility Table, we will use the NEI Standardized State to default this field. If the facility is not in the Historical Facility Table, we will use the State and County FIPS Code to default the state abbreviation. If the State is incorrect, it is likely that the Address, City and Zip Code are also incorrect and that a mailing address was used rather than the physical location of the facility.

4.10 Zip Code

The Zip Code is needed to default geocoordinates for a facility if they are missing or invalid. The Zip Code is the combination of the 5-digit Zone Improvement Plan (ZIP) code and the four-digit extension code (if available) that represents the geographic segment that is a subunit of the ZIP Code, assigned by the U.S. Postal Service.

We will QA the quality of the Zip Code using a Zip Code QA file available at the following address, http://www.epa.gov/ttn/chief/net\2002inventory.html. The file contains the zip code, state abbreviation, post office, and latitude and longitude centroids for the zip codes. The source of the zip code file is ESRI and the file uses zip code information from the US Postal Service and FIPS code information from the US Department of Commerce, National Institute of Information and Technology. Zip Codes and State and County FIPS codes reported in the Site records are compared with zip codes and FIPS codes in the Zip Code QA file. If the reported State and County FIPS Code does not match the FIPS Code in Zip Code QA file, then the reported Zip Code is assumed to be invalid.

If Zip Codes are incorrect or missing, we will use the NEI Standardized Zip Code in the NEI Historical Facility Table to default the Zip Code. If global errors such as dropping leading zeroes for a state or transposition errors exist, then we will correct the Zip Codes. We will also use the Zip Code QA file to replace Zip Codes if the facility is not in the NEI Historical Facility Table. If we are unsure of how to correct errors (e.g., zip code does not match town or latitude/longitude), then we will contact the data providers.

4.11 ORIS Facility Code

The ORIS Facility Code is a unique ID for electric generating units. The ORIS Facility Code is assigned by the Department of Energy (DOE)'s Energy Information Administration(EIA). We will compare the reported ORIS Facility Code for a facility to the ORIS Facility Code in the NEI Historical Facility Table. If the ORIS Facility Code is missing for electric generating units, then we will use the 2002 EGU ORIS Facility Code in the NEI Historical Facility Table to default the missing ORIS Facility Codes. If the comparison indicates that the ORIS Facility Codes are different in the NEI Historical Facility Table and the data submitted in NIF tables to EPA, then we will use DOE information to correctly assign ORIS Facility Codes in the NIF Site Table and to revise the ORIS Facility Codes in the NEI Historical Facility Table.

4.12 Facility Registry ID

EPA's OEI assigns Facility Registry IDs to provide a single ID to identify facilities across EPA databases. We will use the NEI Historical Facility Table to augment this data field if data submitters do not provide it.

4.13 TRI ID

The Toxic Release Inventory (TRI) ID is assigned by EPA's OEI to facilities that report under SARA 313 Right To Know Act. TRI is one of the sources of NEI data for point sources emitting HAPs and ammonia. We will use the NEI Historical Facility Table to augment this data field for TRI facilities if data submitters do not provide it.

4.14 Dun and Bradstreet Number

The Dun and Bradstreet Number is a number assigned to facilities in the Dun and Bradstreet database. The Data Universal Numbering System (DUNS) number is assigned by Dun and Bradstreet to uniquely identify business establishments. The Dun and Bradstreet Number is reported in the TRI. We will use the NEI Historical Facility Table to augment this data field if the data are available.

SECTION 5. GEOCOORDINATES

This section includes the following data fields.

- X coordinate (longitude)
- Y coordinate (latitude)
- UTM zone
- XY Coordinate Type
- Coordinate Default Methodology
- Stack Latitude
- Stack Longitude
- Horizontal Collection Method Code
- Horizontal Reference Datum Code
- Coordinate Data Source Code
- Reference Point Code
- Horizontal Accuracy Measure
- Source Map Scale Number
- NEI Site Latitude
- NEI Site Longitude

5.1 Geocoordinates QA and Default Procedure

The first step in blending and merging of point sources is to QA geocoordinates and correct erroneous coordinates. Latitude and longitude are needed to correctly place facility emission release points and associated emissions into specific geographic domains (grid cell, census tract, etc.) for proper emissions modeling.

All UTM coordinates will first be converted to latitudes and longitudes based on 1984 World Geodetic System datum (WGS84).

The longitude should be reported in units of decimal degrees with a negative sign and the latitude should be reported in units of decimal degrees with a positive sign. We will QA the reported latitude and longitude to determine if units are reported in degrees-minutes-seconds. If the latitude and longitude are reported in degrees-minutes-seconds, we will convert the latitude and longitude to decimal degrees using the following equations.

Latitude (decimal degrees) = Latitude degrees + Latitude minute/60.0 + Latitude second/3600.0

Longitude (decimal degrees) = Longitude degrees + Longitude minute/60.0 + Longitude second/3600.0

If latitude and longitude are reversed, a common error in the 1999 NEI submittals, we will revise the latitude and longitude. If the reported value does not have a decimal, we will use the 1999 NEI for HAPs County FIPS Lookup Table to determine if placing decimal in the reported values will result in a valid latitude and longitude. The 1999 NEI for HAPs County FIPS Lookup Table contains the

following fields:

- State and County FIPS,
- Latitude and Longitude, and
- UTM East, UTM North, and UTM Zone.

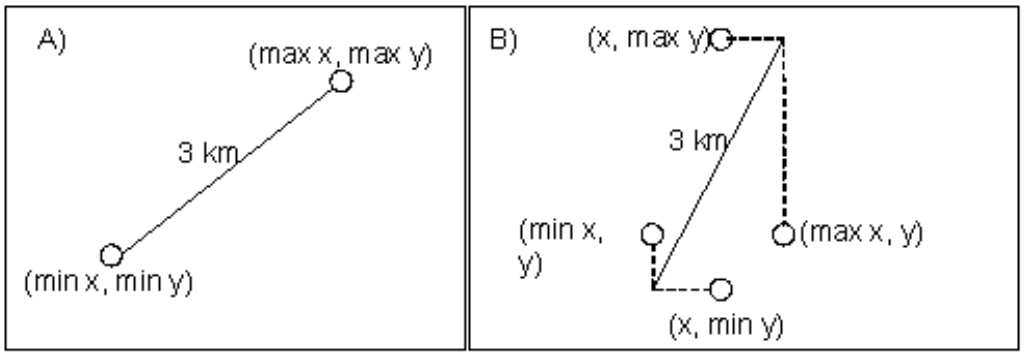
If we cannot correct the reported geocoordinates, we will contact the data provider.

After converting UTMs to latitude and longitude and correcting latitudes and longitudes reported in wrong units, we will use a routine to assess the validity of the latitude and longitude values, to replace values if necessary, and to fill-in missing data points. The QA of location coordinates is a multi-step process. The first step is to make sure that all of the emission release points within a facility are within a reasonable distance of one another. The second step uses geographic information system (GIS) overlays to evaluate each coordinate pair with respect to its county boundary. The stages of the routine are described below.

Step 1. Find and replace the latitude/longitude of emission release points within a facility that are located at distances greater than 3.0 km of other release points in the facility.
This step includes determining if an individual release point within a facility is within 3.0 km of other emission release points. (i.e., to ensure no stacks within a single facility are located miles apart).

If there is only one emission release point, the process is complete and we will proceed to Step 2.

If there is more than one emission release point, we identify the emission release points with the highest and the lowest latitudes and emission release points with the highest and lowest longitudes at the site. We calculate the distances between the highest and lowest latitudes and between the highest and lowest longitudes. Please refer to the diagram below for more details on how we conduct this analysis.



This method may but does not necessarily measure the true distance between emission release points. For example, if a facility has 2 emission release points, the true distance is measured by examining minimum and maximum latitudes/longitudes at all emission release points described by Example A in the diagram. If, for example, a facility has 4 emission release points as in Example B of the diagram, the distance calculated does not represent the distance between any two specific emission release points,

but rather defines the maximum distance that could exist between any two emission release points given all of the points in the set (Example B). This analysis identifies sites whose emission release points are potentially far apart and whose coordinates need correction, but it is not designed to identify the specific outlying emission release points.

A. If the greatest distance between latitudes and between longitudes is less than 3.0 km, the process is complete and we will proceed to Step 2.

B. If the distances between the highest and lowest latitudes or the highest and lowest longitudes are greater than 3.0 km, the latitudes and longitudes of all emission release points within a facility are evaluated. If there are records whose latitudes and longitudes are found to be at a distance of more than 3.0 km, the SIC or NAICS codes are examined to determine if the distance between emission release points is technically correct.

- For a source category, if it is acceptable that the distance between emission release points is greater than 3.0 km, the process is complete and we will proceed to Step 2.
- For a source category, if it is determined that the distance between emission release points should not be more than 3.0 km, then the following steps are conducted.
    1. The distances between all emission release points within the facility is calculated and outlier(s) are identified. For example if four emission release points are present (A, B, C and D), then the distances between the latitudes and longitudes of A & B, A & C, A & D, B & C, B & D, and C & D are calculated.
    2. An average site latitude/longitude is calculated using only the acceptable coordinates. This average site latitude/longitude is then used to replace the inaccurate latitude/longitude values. After this step, the process is complete and we will proceed to Step 2.

Step 2. Find and replace latitude/longitude of emission release points that are out-of-county boundary or missing.

This step includes the use of a GIS overlay to plot each latitude/longitude value and compare it to the physical boundaries of the FIPS county to which the value is associated (i.e., to ensure no stacks are located in the oceans or in far away states). Detailed county boundaries using a scale of 1 to 100,000 or better are used in the GIS overlay.

If the plotted release point is within 10 km of the county, the point is assumed to be valid and neither latitude/longitude nor county FIPS code are corrected. The process is complete.

If the plotted release point is found to exist more than 10 km outside of the county or if the latitude/longitude is missing, then the latitude/longitude of each emission release point is replaced using the following hierarchy. Only one method is used for missing or out-of-boundary latitudes/longitudes within a facility, i.e., method A is not used for one emission release point at a facility and method D for another emission release point at the same facility.

A. *Use Facility Specific Data* - A check is first completed to see if there are any other emission release points at the facility that exist within the10 km zone. The county FIPS code for the emission release points that are outside the 10 km boundary is compared with county FIPs codes of other release points within the facility that exist within the 10 km boundary. If the other valid emission release points within a facility have a different county FIPS code from the emission release points that are in question, then the county FIPS code is changed and the latitude/longitude of the emission release points rechecked to see if the latitude/longitude of emissions release points in question are now valid.

For invalid coordinates, if any emission release points in the facility exist within the 10 km boundary, an average site location is estimated using the latitudes/longitudes of the group of emission release points located within the 10 km zone. The average site latitude/longitude and county FIPS code are used to replace invalid coordinates. The process is complete.

B. *Use Geocoding software* – If none of the reported emission release point latitudes/ longitudes exist within the ten kilometer boundary of the county according to the facility's FIPS code, geocoding software is used. More information on Tele Atlas North America "EZ- Locate" geocoding software is located at www.geocode.com/.

The first step in using Geocoding software is to check the quality of the zip code provided in the inventory with a Zip Code QA file. The Zip Code QA file can be found at the following address, www.epa.gov/ttn/chief/net/2002inventory.html.The source of the Zip Code QA file is ESRI and the file uses zip code information from the US Postal Service and FIPS code information from the US Department of Commerce, National Institute of Information and Technology. Zip Codes and State and County FIPS codes reported in the Site records are compared with zip codes and FIPS codes in the Zip Code QA file. If the reported State and County FIPS Code does not match the FIPS Code in Zip Code QA file, then we would not use the reported Zip Code in the Geocoding process. If Zip Codes are incorrect or missing, we will use the NEI Standardized Zip Code in the NEI Historical Facility Table to default the Zip Code. If global errors such as dropping leading zeroes for a state or transposition errors exist, then we will correct the Zip Codes. We will also use the Zip Code QA file to replace Zip Codes if the facility is not in the NEI Historical Facility Table. If we are unsure of how to correct errors (e.g., zip code does not match town or latitude/longitude), then we will contact the data providers.

After correcting Zip Codes, a list of sites that have emission releases points with missing or erroneous latitude/ longitude is compiled. The file contains site name, the physical addresses of the sites, state and county FIPS code. The file is submitted to the geocoding software. Using the inventory record's street address, the geocoding software matches using the following hierarchy.

1. First the software standardizes the street address and looks for an exact address match.
2. If an exact match cannot be found, the software then tries to match at the single street block, as defined by Geocoder's documentation. The latitude/longitude are located at the centroid of the single street block.
3. If a single street block match cannot be found, then the software tries to match on the 5-digit zip code plus 2 digits. The latitude/longitude are located at the centroid of the 5 digit zip code plus 2 digits.
4. If a match to 5-digit zip code plus 2 digits cannot be found, then the software tries to match to the 5-digit zip code alone. The latitude/longitude are located at the centroid of the 5-digit zip code.
5. If a match to 5-digit zip code cannot be found, then the software tries to match to 3-digit zip code. The latitude/longitude are located at the centroid of the 3- digit zip code.
6. If a match to a 3-digit zip code cannot be found, then the software provides an "ambiguous" match which is a match to multiple non-standardized street segments.

If the geocoding software finds valid latitude/longitude data, the process is considered to be complete and all latitude/longitude pairs are assumed to be valid. The State and County FIPS code will be changed if necessary to match the geocoded state and county FIPS code only after we contact the data provider and verify the county FIPS code. If the data provider indicates the geocoded FIPS code is incorrect, we will not use the geocoded latitude/longitude. If geocoded latitude/longitude are used, the process is considered complete.

C. *Use NEI Historical Facility Table* - If Geocoder cannot find a valid latitude/longitude for the facility, then the NEI Historical Facility Table will be used. If the identified state and county in NEI Historical Facility Table for a facility does not match the reported state and county in the inventory, the NEI Historical Facility Table data will not be used. If the NEI Historical Facility Table contains valid latitude/longitude data with corresponding valid county, the process is considered complete and all the latitude/longitude data are assumed to be valid.

D. *Assign Site Release Point at County Centroid* - If, after each of these stages, an emissions release point latitude/longitude data set is still found to be missing or invalid, the latitude and longitude of the county centroid will be used to default the latitude and longitude.

5.2 X Coordinate (longitude)

The longitude should be reported in decimal degrees with a negative sign. We will use the Geocoordinate QA and Default Procedure to QA and default longitudes.

5.3 Y coordinate (latitude)

The latitude should be reported in decimal degrees with a positive sign. We will use the Geocoordinate QA and Default Procedure to QA and default latitudes.

5.4 UTM Zone

The UTM Zone number will be verified and corrected based on the State and County FIPS code. We will use the 1999 NEI for HAPs County FIPS Lookup Table to QA the UTM zone and to replace any invalid UTM Zones.

5.5 XY Coordinate Type

We will verify the XY Coordinate Type and correct erroneous types. If latitude and longitude are reported, the XY Coordinate Type is LATLON. If UTM coordinates are reported, the XY Coordinate Type is UTM. After we convert all coordinates to latitude and longitude, we will populate this field with coordinate type of LATLON in the final 2002 NEI.

5.6 Coordinate Default Methodology

All defaulted latitudes and longitudes will be identified in the in the Emissions Release Point records in the NEI. We will use the coordinate default flags shown in Table 5.

Table 5. Coordinate Default Flags

Code	Code Description
Exact	Match is to within a unique intersection or within a single side of a single street block.
Near	Match is to a single street block but the correct placement within block is unknown.
Zip code+2	Match to a 5-digit zip code, plus the first two digits of the 4-digit extension.
Zipcode5	Match to a 5-digit zip code
Zipcode3	Match to multiple 3-digit zip codes based on postal service Sectional Center Facility (SCF).
Ambig	Match is to multiple street segments.
Cntycent	County centroid.
NEI Fac Table	Coordinate found in the NEI Historical Facility Table
Site-Avg.	Average of accurate coordinates of other emission release points at the same site

5.7 Latitude/Longitude EPA Data Standard Elements

The EPA's Latitude/Longitude Data Standard consists of the group of data elements used for recording horizontal and vertical coordinates and associated metadata that define a point on earth. Table 6 summarizes these elements. This standard will help users gauge the accuracy and reliability of a given set of coordinates. The primary responsibility for populating these fields lies with the data submitters, as it is difficult if not impossible to discern the origin of a latitude/longitude without being the primary author

of the data.

Stack Latitude
We will report stack latitude in decimal degrees with a positive sign after QA and defaulting of geocoordinates have occurred.

Stack Longitude
We will report stack longitude in decimal degrees with a negative sign after QA and defaulting of geocoordinates have occurred.

Horizontal Collection Method Code
If the Horizontal Collection Method code is missing, we will populate this field code as "027" (The information is not known) when latitude and longitude coordinates that are not defaulted are retained in the 2002 NEI. When we default latitude and longitude using Geocoding software, we will only be able to populate these fields whenever latitude/longitudes were obtained from the TeleAtlas Geocoding EZ Locator Service (http://geocode.com). Table 7 presents default flags from Geocoder software and default values for the latitude and longitude data standards.

Horizontal Reference Datum Code
We will only populate the Horizontal Reference Datum Code with codes in Table 7 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (*http://geocode.com*). If the Horizontal Reference Datum Code is not reported by data submitters and we do not default latitude and longitude using Geocoder software, we will not be able to populate this field unless EPA's OEI provides us a methodology to populate this data field.

Coordinate Data Source Code
If the Coordinate Data Source Code is missing, we will populate this data field based on the data source using the codes in the NIF code tables.

Reference Point Code
If the Reference Point Code is missing, we will populate this field code as "108" (Points not represented by 101-107) when latitude and longitude coordinates that are not defaulted are retained in the 2002 NEI. We will populate the Reference Point Code with codes in Table 7 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (*http://geocode.com*).

Table 6. EPA Latitude/Longitude Data Standards.

Latitude/Longitude Standard	Description	Comments
Latitude Measure	Y Coordinate - The measure of the angular distance on a meridian north or south of the equator.	+78.123456 The number of decimal positions recorded is determined by the precision of the measurement.
Longitude Measure	X Coordinate - The measure of the angular distance on a meridian east or west of the prime meridian.	-123.234561 The number of decimal positions recorded is determined by the precision of the measurement
Source Map Scale Number	The number that represents the proportional distance on the ground for one unit of measure on the map or photo.	Only used when a map has been used to determine latitude/longitude. e.g. 125,000
Horizontal Collection Method Code	Method used to determine the latitude and longitude coordinates for a point on the earth.	e.g., 001 = address-matching house number, 018 on interpolation map, 028 = Global Positioning Method, with unspecified parameters.
Horizontal Accuracy Measure	The measure of the accuracy (in meters) of the latitude and longitude coordinates.	
Horizontal Reference Datum Code	The code that represents the reference datum used in determining latitude and longitude coordinates.	001 = North American Datum of 1927 002 = North American Datum of 1983 003 = World Geodetic System of 1984
Reference Point Code	The code that represents the place for which geographic coordinates were established.	e.g. 101 = Entrance point of a facility or station.; 105 = Point where substance is processed, treated, settled, or stored.; 106 = Point where a substance is released.
Coordinate Data Source Code	The code that represents the party responsible for providing the latitude and longitude coordinates	e.g. EPA Headquarters, a state agency, tribal organization, EPA regional office etc.

Table 7. Geocoder Default Flags and Default Values for Latitude/Longitude Data Standard

Code	Description	Source Map Scale	Horizontal Collection Method Code & Description	Horizontal Reference Datum	Horizontal Accuracy (meters)	Coordinate Data Source Code
Exact	Match is to within a unique intersection or within a single side of a single street block.	24000	002 - Determination method based on address matching-block face.	001 - North American Datum of 1927	12	080 or 084*
Near	Match is to a single street block but the correct placement within block is unknown.	24000	003 - Determination method based on address matching-street centerline.	001 - North American Datum of 1927	50	080 or 084*
Zipcode+2	Match to a 5-digit zip code, plus the first two digits of the 4-digit extension.	24000	038 - Determination method based the center of an area defined by the 5-digit ZIP code and its 2-digit geographic segment extension.	001 - North American Datum of 1927	100	080 or 084*
Zipcode5	Match to a 5-digit zip code.	24000	026 - Determination method based on zipcode-centroid.	001 - North American Datum of 1927	10000	080 or 084*
Zipcode3	Match is to a 3-digit zip code.	24000	021 - Determination method based on interpolation-other.	001 - North American Datum of 1927	1000	080 or 084*
SCF3	Match to multiple 3-digit zip codes based on postal service Sectional Center Facility (SCF).	24000	021 - Determination method based on interpolation-other.	001 - North American Datum of 1927	1000	080 or 084*
Ambig	Match is to multiple street segments.	24000	007 - Determination method based on address matching-other.	001 - North American Datum of 1927 001	20000	080 or 084*
Countycent	County centroid.	24000	021 - Determination method based on interpolation-other; 030 - based on a digital map source (TIGER).	001 - North American Datum of 1927	N/A	080 or 084*

* Coordinates are derived from USPS, Census Bureau Tiger server, or Eagle's TeleAtlas. These correspond to codes 080 (government agency) and 084 (contracting organization).

Horizontal Accuracy Measure

We will only populate the Horizontal Accuracy Measure with values in Table 7 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (*http://geocode.com*). If the Horizontal Accuracy Measure is not reported by data submitters and we do not default latitude and longitude, we will not be able to populate this field unless EPA's OEI provides us a methodology to populate this data field.

Source Map Scale Number

This data field is only applicable when a map has been used to determine latitude and longitude. We will only populate the Source Map Scale Number with codes in Table 7 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (*http://geocode.com*). If the Source Map Scale Number is not reported by data submitters and we do not default latitude and longitude, we will not be able to populate this field unless EPA's OEI provides us a methodology to populate this data field.

5.8 NEI Site Latitude

An average NEI Site Latitude will be estimated using the latitudes of the group of valid nondefaulted latitudes of emission release points within a facility. If all coordinates are defaulted, then the defaulted site latitude will be used.

5.9 NEI Site Longitude

An average NEI Site Longitude will be estimated using the longitudes of the group of valid nondefaulted longitudes of emission release points within a facility. If all coordinates are defaulted, then the defaulted site longitude will be used.

SECTION 6. UNIT AND PROCESS IDS

This section includes the following data fields.

- HAP Emission Data Level
- Emission Unit ID
- Process ID
- Emission Release Point ID

Due to the CERR not including HAPs, some of the key fields that are mandatory for CAPs are not mandatory for state and local agencies and tribes reporting HAPs. For the 2002 NEI, process-level data are required for CAPs, but not for HAPs. While data are also preferred at the process-level for HAPs, if process-level data are unavailable, then the agency or tribe may report HAP data at the site, unit, or emission release point level, and we will process the data for incorporation in the 2002 NEI. The possible reporting options for HAPs and defaults are shown in Table 8.

Table 8. HAP Emission Data Levels

Emission Data Level	Mandatory Reported Fields	EFIG Defaulted Fields
Site	Site ID, Emission Release Point ID	Emission Unit ID, Process ID
Unit	Site ID, Emission Unit ID, Emission Release Point ID	Process ID
Stack	Site ID, Emission Unit ID, Emission Release Point ID	Process ID
Process	Site ID, Emission Unit ID, Process ID, Emission Release Point ID	

6.1 Emission Data Level

If the Emission Data Level is missing, we will default the field based on what level the emissions are reported.

6.2 Emission Unit ID

If the HAP Emission Data Level is reported as Site and Emission Unit IDs are not reported, we will assign Emissions Unit IDs to all of the records in the EU, EP, PE, CE and EM NIF tables.

6.3 Process ID

If HAP Emission Data Level is reported as Site, Unit, or Stack and Process IDs are not reported, we will assign Process IDs to all records in the EP, PE, CE, and EM NIF tables.

6.4 Emission Release Point ID

If the Emission Release Point IDs are not reported, we will assign Emission Release Point IDs to records after we contact data providers.

SECTION 7. SOURCE DESCRIPTION

This section includes the following data fields.

- NAICS Code
- SIC Code
- NAICS/SIC Code Default Flag
- SCC
- Process MACT Code
- Process MACT Compliance Code
- MACT Code Assignment
- Facility Category

7.1 NAICS Code

Standard Industrial Classification (SIC) codes are being replaced by the North American Industry Classification (NAICS) codes that were adopted by Canada, Mexico, and the United States in 1997. The NAICS is a classification of business establishments by economic activity. It supercedes the SIC. The NAICS code consists of 6 digits which are arranged hierarchically.

- ***Two digits*** - Economic sector (NAICS Sector Code)
- ***Three digits*** - Economic subsector (NAICS Subsector Code)
- ***Four digits*** - A group of related industries within the economy (NAICS Industry Code)
- ***Five digits*** - An industry within the economy (NAICS Industry Code)
- ***Six digits*** - A subdivision of an industry (NAICS Code)

NAICS codes may be reported in the NEI as 2, 3, 4, 5, or 6 digits. NAICS Codes are reported at the site level and at the emission unit level. The Primary NAICS Code is reported in the SI table for site level and the NAICS Code Unit Level is reported in the EU table for emission unit level. The Primary NAICS Code is mandatory in the NIF, and the NAICS Code Unit Level is not mandatory.

To meet the EPA's NAICS Data Standard, we adapted the Census Bureau's 1987 SIC to 2002 NAICS crosswalk (see http://www.census.gov/epcd/naics02/) and applied it to the 1999 NEI for HAPs. NAICS codes supplied by the data submitters were not overwritten in the 1999 NEI. We prepared a NAICS/SIC crosswalk for the NEI by modifying the Census Bureau crosswalk to accommodate those situations in which one SIC code maps to multiple NAICS codes. Where all the NAICS codes associated with one SIC code shared the first 5 digits, the SIC code is mapped to this 5 digit NAICS code. If no common 5 digit NAICS code existed, we applied the common 4 digit NAICS code, and so on. In those cases where all of the NAICS codes associated with an SIC code did not share the same 5, 4, 3 or 2 digit NAICS code, then the most common 5, 3, 4, 2 digit NAICS code was selected. The SIC/NAICS crosswalk is presented in a table, NAICS02toSIC87_NEI99, in the State Lookup Tables file, 99nei_lkup_states.zip, found at: http://www.epa.gov/ttn/chief/net/1999inventory.html.

We will first QA the reported NAICS codes by using the Census Bureau's list of NAICS codes (see http://www.census.gov/epcd/naics02/). We will default all Primary NAICS Code.

We will use the following hierarchy to default the Primary NAICS Code.

1. If valid SIC codes are reported, we will use the SIC/NAICS crosswalk used in the 1999 NEI for HAPs to augment the NAICS Codes.
2. For facilities that have invalid or missing SIC codes or SIC codes that are not in the SIC/NAICS crosswalk , we will use facility-specific NAICS codes from third-party services such as PowerBusiness by InfoUsa. The PowerBusiness database is available on a subscription basis and contains census-derived data for over 14 million U.S. businesses and is searchable by name, address, business type and SIC and NAICS codes.
3. For facilities that have invalid or missing SIC codes, have SIC codes are not in the SIC/NAICS crosswalk, or are not in Power Business database and are in the NEI Historical Facility Table, we will use the NEI Historical Facility Table to augment the Primary NAICS Code.

We will only default invalid NAICS Code Unit Level if valid SIC Code Unit Level is reported and a NAICS Code match exists for the SIC Code in the 1999 NEI for HAPs SIC/NAICS crosswalk. We will not default missing NAICS Code Unit Level.

We will assign NAICS/SIC default flags shown in the Table 9 to indicate how the NAICS codes are defaulted. Default flags are included for NAICS codes in the Site and Emission Unit Tables.

7.2 SIC Code

SIC codes may be reported in the NEI as 2, 3, or 4 digits.

- ***Two digits*** - Major Group Number
- ***Three digits*** - Industry Group Number
- ***Four digits*** - Industry Number

SIC Codes are reported at the site level or at the emission unit level. The Primary SIC Code is reported in the SI table for site level and the SIC Code Unit Level is reported in the EU table for emission unit level. SIC codes are not mandatory in the NEI.

We will first QA the reported SIC codes by using the Census Bureau's list of SIC codes (see http://www.census.gov/epcd/naics02/). A common error found in the 1999 NEI was the use of outdated SIC codes. If SIC codes are invalid, we will first use the 1987 SIC manual to revise old SIC codes. We will default Primary SIC Codes if possible.

We will use the following hierarchy to default the Primary SIC Code.

1. For facilities included in the NEI Historical Facility Table, we will use the NEI Historical Facility Table to augment the Primary SIC Code.
2. For new facilities not included in the NEI Historical Facility Table, we will use facility-specific SIC codes from third-party services such as PowerBusiness by InfoUsa. The PowerBusiness

database is available on a subscription basis and contains census-derived data for over 14 million U.S. businesses and is searchable by name, address, business type and SIC and NAICS codes.

3. For new facilities not in the NEI Historical Facility Table and not in the Power Business database, if valid NAICS codes are reported and there is a one-to-one match between NAICS and SIC codes, we will use the SIC/NAICS crosswalk used in the 1999 NEI for HAPs to augment the SIC Codes.

For the SIC Code Unit Level, we will only default missing SIC Codes if valid NAICS codes are reported and there is a one-to-one match between NAICS and SIC codes by using the SIC code/NAICS crosswalk used in the 1999 NEI for HAPs.

We will assign the SIC default flags shown in the Table 9 to indicate how the SIC codes are defaulted. Only codes 01 and 10 are valid as SIC default flags. Default flags are included for SIC codes in the Site and Emission Unit Tables

TABLE 9. NAICS/SIC Code Default Flags

Code	Match Type	Code Description
01	one to one	One SIC maps to only one NAICS code.
02	one to many	One SIC maps to many NAICS code all of which share the first 5-digits
03	one to many	One SIC maps to many NAICS code. Have chosen the most common 5 digit NAICS among these.
04	one to many	One SIC maps to many NAICS code all of which share the first 4-digit.
05	one to many	One SIC maps to many NAICS code. Have chosen the most common 4 digit NAICS among these.
06	one to many	One SIC maps to many NAICS code all of which share the first 3-digits.
07	one to many	One SIC maps to many NAICS code. Have chosen the most common 3 digit NAICS among these.
08	one to many	One SIC maps to many NAICS code all of which share the first 2-digits.
09	one to many	One SIC maps to many NAICS code. Have chosen the most common 2 digit NAICS among these.
10	exact	1999 NEI facility file

7.3 SCC

SCCs are reported at the process level. SCCs are mandatory for CAPs and must be reported at the 8 digit level for CAPs. SCCs are not mandatory for HAPs and may be reported in the NEI as 3, 6, or 8 digits for HAPs.

We will QA the reported SCCs by using the NIF Code Lookup Table. For missing or invalid SCCs, we will use the Emission Process Description to default the SCCs.

7.4 Process MACT Code

The MACT Code identifies processes that are or will be subject to MACT standards and new area source standards. For a list of MACT standards, please see www.epa.gov/ttn/atw/eparules.html. For a list of area source categories, please see www.epa.gov/ttn/atw/urban/arearules.html.

MACT codes will be used in the blending/merging step to identify potential duplicate processes for a facility where data are provided by multiple data sources.

MACT codes will be assigned to processes based on the following hierarchy.

1. MACT Code identified by State or local agency or tribe
2. Emissions data provided by the MACT engineers
3. Facility list provided by the MACT engineer
4. SCC/MACT Code dictionary
5. SIC/MACT Code dictionary

In all cases, any one process and thus any emission record will be associated with one and only one MACT category. If data are supplied by the EPA for a particular MACT category or a facility list is provided, then the appropriate MACT code will be assigned and SIC codes and SCCs will not be used to default the MACT code. If SCC or SIC code defaults are used, we either use SCC or SIC code defaults for a process within a facility and do not mix SCC and SIC code defaults at a single facility. Thus once we default a process within a facility based on SCC, we do not default MACT codes for other records missing SCCs that have a SIC code. For example, 3 following processes are present at a facility.

Process 1	SCC 10100101	SIC 4911
Process 2		SIC 4911
Process 3		SIC 4911

We would default Process 1 to MACT Code 1808-1, utility coal boiler, but we would not default Process 2 and 3 to a MACT Code.

The SCC/SIC Code/MACT dictionary used in the 1999 NEI was developed by comparing all of the SCCs and SIC Codes with information on types of sources that may be subject to each MACT standard. EPA MACT engineers then reviewed the NEI to revise the dictionary and to revise the facilities associated with individual MACT categories. The comments of the MACT engineers were incorporated into the 1999 NEI.

For 2002 NEI, we revised the 1999 SCC/SIC Code/MACT dictionary to include the new area sources categories.

If state and local agencies and tribes do not provide MACT codes in their 2002 submittals, we will default MACT codes. The 1999 NEI MACT code assignment will be a starting point for the 2002 NEI.. If new facilities are reported in 2002 that do not contain MACT Codes, we will use facility lists or SCC/SIC Code/MACT Code dictionary to default MACT codes.

We will assign MACT Codes to both HAP and CAP records in the 2002 NEI.

7.5 MACT Code Assignment

The basis of MACT Code assignment in the 2002 NEI will be indicated in the data field, MACT Code Assignment using the following values.

1. ESD
2. State and local agency or tribe
3. SCC default
4. SIC default

7.6 MACT Compliance Code

The MACT Compliance Code will be used to identify whether a source is subject as a major or an area source to CAA Section 112 and Section 129 standards. The MACT Compliance Code and MACT Code can be used by modelers to identify for a MACT source category major and area sources that are subject to standards. For projections, major and area sources within a MACT category can have different percent reductions. Acceptable codes are:

01 Major Source (>10/25 tpy), Compliance date has not yet occurred;
02 Major Source (>10/25 tpy), Compliance date has occurred;
03 Area Source (<10/25tpy) category listed in, and subject to, Sections 112 &129 standards;
04 Area Source (<10/25tpy) category listed in, but not subject to, Sections 112 &129 standards as a synthetic minor;
05 Area Source (<10/25tpy) category listed in, but not subject to, Sections 112 &129 standards as true area or natural minor source;
06 Rule only affects major sources; area may be flagged; and
07 Rule only covers certain HAPs, all HAPs flagged.

We will default this field to 03 for the area source categories subject to new area source standards. We will default the code to 01 and 02 for major sources subject to MACT standards.

7.7 Facility Category

The Facility Category indicates whether a facility that emits HAPs is a major or an area source. HAP

emissions are defined as either major or area in the Clean Air Act. Major sources are defined as stationary sources that:

- have the potential to emit 10 tons per year (tpy) or more of one HAP; or
- have the potential to emit 25 tpy or more of any combination of HAPs.

Area sources are defined as sources with annual emissions below these thresholds.
For more information about the definition of major and area sources, please see www.epa.gov/ttn/atw/pollsour.html.

If Facility Category is missing for a facility that emits HAPs, we will calculate the total emissions of HAPs for the facility by summing the reported emissions of individual HAP to determine if a facility is a major or an area source. If more than one HAP Performance Level is reported, we will select an emissions value for only one HAP Performance Level to use in calculating the total emissions of the facility. The value of the reported emissions using the following hierarchy will determine if a facility is major or area.

1. Potential
2. Actual
3. Allowable
4. Maximum

After calculating the facility emissions, we will populate missing fields with valid NIF Codes. The valid NIF codes for the Facility Category are:

01 - Major sources, and
02 - Area sources.

SECTION 8. EMISSION RELEASE DESCRIPTION AND PARAMETERS

This section includes the following data fields.

- Emission Release Point Type
- Stack Parameters (Stack Height, Stack Diameter, Exit Gas Temperature, Exit Gas Velocity, Exit Gas Flow Rate)
- Stack Parameter Default Flag

In preparing emissions for grid modeling, valid parameters for the physical characteristics of each release point (stack height, diameter, temperature, velocity, and flow) are necessary to correctly place facility release points and associated emissions into vertical layers for proper air quality modeling. Gaussian dispersion models need stack parameters to characterize the plume, which is needed to estimate proper concentrations from these models. The first step is to QA the Emission Release Point Type. After we augment for any invalid or missing Emission Release Point Types, we use a routine to assess the validity of the stack parameters, to replace values if necessary, and to fill-in missing data points.

8.1 Emission Release Point Type

The Emission Release Point Type identifies whether emissions are released as a stack or fugitive. The valid NIF codes for the Emission Release Point Type are:

01 - Fugitive,
02 - Vertical,
03 - Horizontal,
04 - Goose Neck,
05 - Vertical with Rain Cap, and
06 - Downward Facing Vent

The Emission Release Point Type is needed to determine how to augment missing or out-of-range stack parameters.

We will evaluate the SCCs of Emission Release Point Types to determine if the Emission Release Point Type is reported correctly as a fugitive or a stack by using the SCC/Emission Release Point Type Crosswalk. Two tables are in the crosswalk - SCCs that are always fugitive and SCCs that are always non-fugitive (stack). The Crosswalk is available at the following address, www.epa.gov/ttn/chief/net/2002inventory.html. If the Emission Release Point Type is found to be inconsistent with the reported SCC, we will revise the Emissions Release Point Type. When we change a non-fugitive(stack) value to fugitive, we will use NIF Code 01 for Emissions Release Point Type. When we change a fugitive value to non-fugitive, we will use NIF Code 02 for a vertical stack.

8.2 Stack Parameters QA and Default Procedure

Stack parameters include stack height, stack diameter, exit gas temperature, exit gas velocity, and exit gas flow rate.

We will employ a routine that compares each emission release point parameter to a minimum and maximum range of values and when that parameter is missing or is found to exist outside of that range, we will augment the parameter. We will also check non-fugitive stack parameters for internal consistency between:

- stack height and diameter, and
- stack diameter, exit gas velocity, and exit gas flow rate.

When internal consistency is not met, we will replace the parameters.

The following steps summarize the process of finding and replacing missing, out-of-range, or internally inconsistent stack parameters.

Step 1: <u>For fugitive emission release points, replace stack parameters</u>

For fugitive emission release points, we will first compare the existing height against the following range thought to be representative of the minimum and maximum values allowable for most fugitive emission release points. Although the values must be reported in English units and will be defaulted to English units, we have provided metric conversions for convenience.

Fugitive Release Height:	0.1 to 100 ft	0.03048 to 30.48 m

If the height is valid, we will keep the height and replace all other stack parameters with the defaulted values listed below. If the height is invalid, we will replace all stack parameters with the defaulted values.

Stack Height:	10 ft	3.048 m
Stack Temperature:	72 ^{o}F	295.37 ^{o}K
Stack Diameter:	0.003ft	0.00091 m
Stack Velocity:	0.0003 ft/sec	0.000091 m/sec
Stack Flow:	0 cu ft/sec	0 cu m/sec

Step 2: <u>For non-fugitive emission release points, find out-of-range or missing stack parameters</u>

For non-fugitive emission release points, we will first compare existing stack parameters against a set of the following ranges thought to be representative of the minimum and maximum values allowable for most emission release points. Although the values must be reported in English units and will be defaulted to English units, we have provided metric conversions for convenience.

Stack Height:	0.1 to 1200 ft	0.03048 to 365.76 m
Stack Temperature*	50 to 1,800 ^{o}F	283.15 to 982.37 ^{o}K
Stack Diameter	0.1 to 50 ft	0.03048 to 15.24 m
Stack Velocity	0.1 to 100 ft/sec	0.03048 to 30.48 m/sec

*Stack Temperature must be greater than 250 ^{o}F for the following SCCs and MACT Codes.

- SCCs: 10*, 20*, 501001*, 501005*, 502001*, 502005*, 503001*, 503005*
- MACT codes: 0105, 0107, 0107-1, 0107-2, 0107-3, 0107-4, 0108, 0801, 0801-1, 0801-2, 0801-3, 0801-4, 1801, 1802, 1808-1, 1808-2, 1808-3

First we will identify missing or out-of range parameters. Then we will evaluate the source category to determine if out-of-range parameters may be plausible. If any parameter is missing or out-of range, the parameter will be replaced using the procedures described in Step 4.

If all parameters are found to exist within the bounds of the emission release point ranges, we will proceed to Step 3.

Step 3: For non-fugitive emission release points, find inconsistencies in stack parameters

We will determine any inconsistencies in stack parameters by conducting the following two steps.

A. For stack diameter, we will compare the stack diameter to the stack height. For non-fugitive emission release points, the stack height may not be less than stack diameter.

B. We will determine the internal consistency between diameter, velocity and flow rate using the following equation.

Stack Flow [cu ft/sec] = (Π [Pi] * (Stack Diameter [ft] / 2) ^ 2) * Stack Velocity [ft/sec]

If the calculated flow and the reported flow are within 10 % of one another, then internal consistency is assumed to be valid.

If all parameters are found to exist within the bounds of the emission release point ranges in Step 2, and the consistency checks (A) and (B) in Step 3 are satisfied, no additional steps will be taken.

If any parameter is missing or out-of range, or if the parameters fail the internal consistency tests, the parameter will be replaced using the procedures described in Step 4.

Step 4: Replace stack parameters for non-fugitive emission release points

The first step in replacing stack parameters is to determine if there are problems with stack height or diameter. Because stack height and diameter are the physical parameters that are most easily measured or estimated, when there are problems with these parameters, then the entire set of stack parameters are deemed questionable. If either height or diameter is missing or out-of range, or if the stack diameter is greater than stack height, then all 5 parameters will be defaulted using national default sets of physical parameter data. No additional steps are taken once all 5 parameters are defaulted.

If stack height and diameter do not need replacement, then velocity and flow rate are evaluated next. If velocity and flow rate are not internally consistent, we will QA the flow rate to determine if it was reported in cubic feet per minute rather than cubic feet per second as required in the NIF. In the 1999 NEI, we found that several data submitters reported flow rate in cubic feet per minute rather than cubic feet per second. We will correct flow rates reported in cubic feet per minute to cubic feet per second and than evaluate the flow rate and velocity for internal consistency.

If the internal consistency is not met for velocity, flow rate, and diameter, Table 10 provides instructions on how we will replace missing, out-of-range values, or internally inconsistent values for velocity and

flow rate based on different reported scenarios. Velocity and flow rate are augmented either by calculation or the use of national defaults.

Table 10. Stack Parameter Data Replacement Matrix (X = Data value present)

Diameter	Velocity	Flow Rate	Action
X	X	X	1. Check that velocity is within range. A. If velocity is within range and flow rate does not meet internal consistency for diameter, velocity and flow rate, then: ▸ Calculate flow rate using internal consistency formula. B If velocity is not within range, then: ▸ Calculate velocity using internal consistency formula. ▸ Check that calculated velocity is within range. If so, then default to calculated velocity. ▸ If calculated velocity is not within range, then default all 5 parameters using national default set.
X	-	X	1. Calculate velocity using internal consistency formula. 2. Check that calculated velocity is within range. A. If calculated velocity is not within range, then: ▸ Default all 5 parameters using national default sets.
X	X	-	1. Check that velocity is within range. A. If velocity is within range, then: ▸ Calculate flow rate using internal consistency formula. B. If velocity is not within range, then: ▸ Default all 5 parameters using national default sets.
X	-	-	1. Default velocity using national default sets. 2. Calculate flow rate using internal consistency formula.
-	X	X	1. Default all 5 parameters using national default sets.

Finally, in cases where all 5 parameters have not been defaulted, and velocity and flow rate have been evaluated and replaced if necessary, temperature is evaluated. If temperature is missing or out-of-range, then the temperature is defaulted using national default sets of physical parameter data.

The following four National default sets of physical parameter data are available for non-fugitive emissions.

1. SCC match
2. MACT Codes
3. facility level SIC code match
4. national default for release points, if no SCC, MACT Code or SIC Code match is possible

From Version 3 of the 1999 NEI for HAPs and 1999 NEI for CAPs, default look-up tables were generated by SCC, MACT Code and SIC code to report the average value calculated for stack height, diameter, temperature and velocity for all emission release point types that are coded as stacks. Only valid values were used in the averaging. The following records were removed from the 1999 NEI files prior to preparing SCC, MACT Code, and SIC Code stack parameter default look-up tables.

- Records with SCCs incorrectly reported as stack
- Records with national default
- Records with velocity greater than 100 ft/sec
- Records with temperatures < 250 ^{o}F with SCC 10*, 20*, 501001*, 501005*, 502001*, 502005*, 503001*, and 503005*
- Records with temperatures < 250 ^{o}F with MACT codes 0105, 0107, 0107-1, 0107-2, 0107-3, 0107-4, 0108, 0801, 0801-1, 0801-2, 0801-3, 0801-4, 1801, 1802, 1808-1, 1808-2, and 1808-3

The stack flow rate was calculated by using the average diameter and average velocity and the equation in Step 3.

Default stack parameters are available for more than 3,600 SCCs, 125 MACT codes and more than 800 SIC codes. Separate look-up tables are prepared for SCCs, MACT Codes, and SIC Codes. These files can be found at the following address. www.epa.gov/ttn/chief/net/2002inventory.html

The following hierarchy is used to replace missing or invalid stack parameters for non-fugitive emissions..

1. SCC match unless the SCC is 39999999
2. MACT Code match
3. facility level SIC code match
4. national default for release points, if no SCC, MACT Code or SIC Code match is possible

If multiple SCCs (or MACT Codes or SIC Codes) are available for a single Emission Release Point, we will use the default record having the lowest stack height to modify and replace that out-of-range release parameters.

If no SCC, MACT Code, or SIC code match is possible, we will use the following national default values for the stack parameters. Although the values must be reported in English units and will be defaulted to English units, we provide metric conversions for convenience.

Stack Height:	10 ft	3.048 m
Stack Temperature:	72 ^{o}F	295.37 ^{o}K
Stack Diameter:	1 ft	0.3048 m
Stack Velocity:	15 ft/sec	5.572 m/sec
Stack Flow:	12 cu ft/sec	339.802 cu m/sec

8.3 Stack Parameter Default Flag

All defaulted stack parameters will be identified in the Emission Release Point record. We will use the following coding system to identify the source of default stack parameters.

0 = Original value (not a default)
1 = SCC default
2 = SIC code default
3 = National default
4 = Calculated value
5= MACT Code default

A single default field will be used to represent the source of all five stack parameters. The codes will be presented in this field in the following order: stack height, exit gas temperature, stack diameter, exit gas velocity, exit gas flow rate. Thus, the code “00014" indicates that stack height, exit gas temperature, and stack diameter are original values, exit gas velocity is SCC defaults, and exit gas flow rate is calculated based on the stack diameter and exit gas velocity values.

SECTION 9. UNIT CODES

This section includes the following data fields.

- Design Capacity Numerator
- Design Capacity Denominator
- Emission Unit Numerator
- Factor Unit Denominator
- Factor Unit Numerator
- Throughput Unit Numerator

Unit codes are used in a variety of data fields in the NIF. Table 11 associates NIF Data Fields with the Unit Fields. For example, if Design Capacity is reported, then the Design Capacity Numerator and Denominator must be reported.

Table 11. Fields in NIF Requiring Unit Codes

Fields Needing Unit Codes	Unit Code Fields
Design Capacity	Design Capacity Numerator Design Capacity Denominator
Emission Numeric Value	Emission Unit Numerator
Factor Numeric Value	Factor Unit Numerator Factor Unit Denominator
Actual Throughput	Throughput Unit Numerator

9.1 Unit Codes

We will QA the reported Unit Codes. If the reported Unit Code values are incorrect, then we will make corrections using valid codes. This type of error occurred frequently in the development of the 1999 NEI. If a new code is reported that is not in the NIF Code Table, we will add the new code to the NIF Unit Code Table. If we cannot determine what the Unit Code is or if the Unit Code is missing for a reported value, we will contact the data provider to obtain correct unit and then we will augment the data field with the correct Unit Code.

9.2 Design Capacity, Design Capacity Numerator and Denominator

We will conduct additional QA of Design Capacity and Design Capacity Numerator and Denominator for industrial boilers, turbines and IC engines. Design Capacity information is needed for modeling that considers regional control strategies.

The steps of the process include the following.

1. Identify industrial boilers, turbines and IC engines by using SCCs of 102* and 202*.
2. QA reported Design Capacity Numerator and Denominator. The Design Capacity Numerator/Design Capacity Denominator should be MMBTU/HR. If Design Capacity Numerator/Design Capacity Denominator is reported as BTU/HR, we will QA the Design Capacity Value and replace with correct value as reported in units of MMBTU/HR.
3. If Design Capacity and Design Capacity Numerator and Denominator are missing and capacity information is reported in the Emission Process Description or in the Emission Unit Description, we will use this information to populate the Design Capacity, Design Capacity Numerator, and Design Capacity Denominator fields.
4. For records that are missing Design Capacity and Design Capacity Numerator and Denominator with SCCs of 102* and 202* except 10200402, 10200403, 10200502, 10200503, 10200602, and 10200603, and emissions greater than 100 tpy NOX, we will send a list of these facilities to the data provider and encourage them to provide capacity information.

SECTION 10. ACTIVITY DATA FIELDS

This section includes the following data fields.

- Emission Type
- Start Date
- End Date
- Start Time
- End Time
- Annual Average Days per Week
- Annual Average Hours per Day
- Annual Average Hours per Year
- Annual Average Weeks per Year
- Period Days per Week
- Period Hours per Day
- Period Hours per Period
- Period Weeks per Period
- Winter Throughput Percent
- Spring Throughput Percent
- Summer Throughput Percent
- Fall Throughput Percent

In addition to the correct placement of emissions, air quality modeling attempts to represent the actual physical and chemical processes as they occur over a specific duration of time. As such, it is important that the temporal allocation of emissions be as accurate as possible. Temporal allocation can be thought of as an accounting of emissions variation over time. The simplest temporal allocation is for a steady-state emissions source that continually releases emissions at the same rate all the time. Under actual conditions, however, steady-state emission sources are quite rare. Instead, emissions sources may operate only in the winter, not operate on Sundays, or their activity may peak during certain hours of the day.

Temporal allocations based on seasonal throughput percentages, weeks per year, days per week, and hours per day allow emissions variability to be correctly modeled during the desired modeling periods. For example, an emission source reporting activity for fifty-two weeks per year, five days a week, and eight hours a day would be allocated a temporal emission release between the hours of 9 AM and 5 PM, Monday through Friday, each week of the year.

If operating parameters are not reported by data providers to assign emissions to month, week, day, or hour, default SCC-based temporal profiles are used to represent when emissions from source types typically occur. Documentation of the procedures and allocation profiles currently used in this allocation can be found at the following address. (www.epa.gov/ttn/chief/emch/temporal) It is the intent of EPA to use actual reported operating parameters to assign when individual sources release their emissions. Currently, however, many instances exist where provided data are outside of acceptable ranges or

missing completely. This results in the assignment of a flat profile of twenty four hour a day, seven days a week, fifty two weeks a year operation (i.e., the same emission release each hour of each day of the year). Having operating parameters provided with the emissions inventory will result in a better, more accurate, source specific distribution of emissions on a diurnal, daily, weekly, and seasonal level.

10.1 Emission Type

The Emission Type is a temporal description of the emission reported. Acceptable values include the following codes.

20	Average Sunday
21	Average Monday
22	Average Tuesday
23	Average Wednesday
24	Average Thursday
25	Average Friday
26	Average Saturday
27	Average Weekday
28	Average Weekend Day
29	Average Day
30	Entire Period

We will QA the data field and default missing or invalid codes to “30".

10.2 Start and End Dates

If the Start Date and End Date are not included in data submittals, we will default the Start and End Date using the Inventory Year. For example if the Inventory Year is 2002, we will default the Start Date to 20020101 and the End Date to 20021231.

10.3 Start and End Times

If the Start Time and End Time are invalid in data submittals, we will default the Start Time to 0000 and the End Time to 2359.

10.4 Annual Average Days Per Week

If Annual Average Days Per Week is invalid (greater than 7), we will default invalid values to 7.

10.5 Annual Average Hours Per Day

If Annual Average Days Per Week is invalid (greater than 24), we will default invalid values to 24.

10.6 Annual Average Hours Per Year

We will QA this data field by using the following equation.

Annual Ave Hrs Per Yr = Annual Ave Hrs Per Day * Annual Ave Days Per Week * Annual Ave Weeks Per Yr

The maximum value is 8760. If Annual Average Hours Per Year is invalid, we will default invalid values by using the equation if valid Annual Average Hours Per Day, Annual Average Days Per Week, and Annual Average Weeks Per Year are reported. If these data fields are not reported, then we will default invalid values to 8760.

10.7 Annual Average Weeks Per Year

If Annual Average Weeks Per Year is invalid (greater than 52), we will default invalid values to 52.

10.8 Period Days per Week

We will QA this data field using Start and End Dates for records. If the period is one week or less based on the Start and End Dates, we will default missing or invalid values to the number of days in a week based on the Start and End Dates. If the period is greater than one week based on Start and End Dates, we will default missing and invalid values to 7.

10.9 Period Hours per Day

We will QA this field using Start and End Times. If Start and End Times are present, we will default missing and invalid values based on Start and End Times. If Start and End Times are missing, we will default missing and invalid values to 24.

10.10 Period Hours per Period

We will QA this field using Start and End Dates and Start and End Times. We will first calculate the number of days of the period for the record. If Start and End Times are reported, we will use the Start and End Times to calculate the number of hours in a day the process operates. If the Start and End Times are not reported, we will assume the facility operates 24 hrs a day. We will default missing or invalid values by multiplying the number of days per period by the number of hours per day.

10.11 Period Weeks per Period

We will QA this field using the Start and End Dates. We will default invalid values based on Start and End Date.

10.12 Winter, Spring, Summer, and Fall Throughput Percents

These data fields are needed by EPA and others to calculate emissions for an average day in a specific season from annual emissions. If seasonal throughput percentages are not reported, then we will use SCC-based temporal profiles found at the following address. (www.epa.gov/ttn/chief/emch/temporal) NOTE that EPA will not retain seasonal emissions from data submitters unless seasonal throughput percentages are included in NIF submission.

Throughput Percentages should be reported as an integer between 0 and 100. We will first QA the reported values and change any percentages reported as decimal to integer. We found some values being reported as a decimal less than 100.

Then we will use the following methodology to QA Throughput Percentages.

1. Determine if the reported seasons of the Throughput Percentage match the Start and End Date period.
2. Sum the reported Seasonal Throughput Percentages. The sum of Winter, Spring, Summer and Fall Throughput Percents should equal 100.

When the Throughput Percents are invalid, we will use the following methodology to default these data fields.

- If the sum is greater than100, we will default Throughput Percentages by allocating percentages based on the Start and End Dates. For example, if the Start and End Dates are April 2002 - August 2002, then only Spring and Summer Throughput Percentages should be reported and we would default the invalid fields to Spring Throughput Percent of 40 and Summer Throughput Percent to 60.
- If the sum is less than 100%, we will default missing percentages by summing reported percentages and equally allocating remaining percentages based on Start and End Dates. For example, if the reporting period is April 2002 - November 2002 and Summer Throughput Percentage is reported as 90, we would default Spring Throughput Percent to 5 and Fall Throughput Percent to 5.

SECTION 11. CONTROL EQUIPMENT

This section includes the following data fields.

- Control Status
- Primary Device Type Code
- Secondary Device Type Code
- Third Device Type Code
- Fourth Device Type Code
- Total Capture Control Efficiency
- Total Capture Control Efficiency Methodology Flag

Information about control equipment and efficiencies is needed for modeling. This information can be used in regulatory assessments as well as projections. Control information allows one to look at the potential for future reductions within a facility or source category.

11.1 Control Status

We will default the Control Status to indicate if emissions are controlled or uncontrolled. The valid values are "CONTROLLED" and "UNCONTROLLED". For missing values, we will use the following methodology.

- If a primary control device code is not reported in the CE table, we will leave the field blank.
- If any control device code is reported as any code other than 000 in the CE table, we will default this field as "CONTROLLED".
- If the control device code of 000 is reported in the CE table, we will populate this field as "UNCONTROLLED".

11.2 Primary Device Type Code, Secondary Device Type Code, Third Device Type Code, and Fourth Device Type Code

We will QA the Device Type Codes. For missing or invalid Device Type Codes, we will use the Control System Description to default the Device Type Codes. For control equipment not included in the Control Device Code Table, we will add new codes.

11.3 Total Capture Control Efficiency

To facilitate the use of data in dispersion and exposure modeling, we will default missing Total Capture Control Efficiencies. The Total Capture Control Efficiency represents the collective (aggregate) value for all control devices and is reported as a percent decimal.

We will QA the numerical value and revise if decimal instead of percent is used. We found that some states reported the numerical value as decimal only, for example they reported 0.985 rather than 98.5. We will also QA the numeric value for a range of 1 - 100.

We will augment the Total Capture Control Efficiency by reviewing the Primary Percent Control Efficiency, Percent Capture Efficiency, and Total Capture Control Efficiency fields. Table 12 presents the methodology that will be used to populate this data field.

Table 12. Total Capture Control Efficiency Methodology

Total Capture Control Flag Code	Reported Data Fields	Total Capture Control Efficiency Methodology
01	Primary Percent Control Efficiency = 0 Percent Capture Efficiency = 0 Total Capture Control Efficiency = 0	Total Capture Control Efficiency = 0
02	Total Capture Control Efficiency	Total Capture Control Efficiency = Reported Total Capture Control Efficiency
03	Percent Capture Efficiency	Total Capture Control Efficiency = Percent Capture Efficiency
04	Percent Capture Efficiency Total Capture Control Efficiency	Total Capture Control Efficiency = Total Capture Control Efficiency
05	Percent Capture Efficiency Primary Percent Control Efficiency	Total Capture Control Efficiency = Percent Capture Efficiency * Primary Percent Control Efficiency
06	Primary Percent Control Efficiency	Total Capture Control Efficiency = Primary Percent Control Efficiency
07	Primary Percent Control Efficiency Total Capture Control Efficiency	Total Capture Control Efficiency = Total Capture Control Efficiency
08	Total Capture Control Efficiency Percent Capture Efficiency Primary Percent Control Efficiency	1. Calculate Total Capture Control Efficiency by multiplying Percent Capture Efficiency and Primary Percent Control Efficiency 2. Compare reported and calculated Total Capture Control Efficiencies. Use the greater value.

11.4 Total Capture Control Efficiency Methodology Flag

All Total Capture Control Efficiencies will be identified in the Control Equipment record using the Total Capture Control Flag codes listed in Table 12.

SECTION 12. EMISSIONS

This section includes the following topics.

- Pollutant Code
- PM Augmentation
- VOC from HAPs
- PM from HAPs
- Emissions Numeric Value
- CAP Emissions
- HAP Emissions
- QC of Emissions
- HAP Performance level
- Period Type
- Annual Emissions (tons)
- Non-Annual Emissions
- Emission Data Quality Rating
- Data Source
- Historical Emissions Table

A variety of QA and data augmentation occurs related to the pollutant emissions. We will conduct QA of reported pollutant codes and their emissions values. We will augment emissions of some pollutants, and we will generate annual emission estimates from reported data for missing CAPs. In addition, we will assign an emission data quality rating to every record and identify the source of data. A Historical Facility Emissions Table will be compiled to show original emissions submitted by all data sources, draft 2002 estimates, revisions received in Spring 2005, final estimate in 2005 version of the 2002 NEI, and any revisions received over time. The Data Source indicates who provided the emission estimate (state or local agency, EPA, Industry, Tribes, etc.)

12.1 Pollutants

The 2002 NEI includes Hazardous Air Pollutants (HAPs) and Criteria Air Pollutants (CAPs). Data providers should report pollutants using NIF pollutant codes. We will add any additional pollutants that are HAPs submitted by data providers to the NIF pollutant code table.

HAPs Reporting

Section 112(b) of the CAA contains a list of 188 HAPs. HAPs are generally defined as those pollutants that are known or suspected to cause serious health problems, including cancer. Section 112(b) of the Clean Air Act currently identifies a list of 188 pollutants as HAPs, www.epa.gov/ttn/atw/orig189.html. EPA's ATW web site presents more information on HAPs, their effects, and EPA's programs to reduce HAPs. (www.epa.gov/ttn/atw/basicfac.html)

The NEI includes emissions data for all 188 HAPs. In addition to numerous specific chemical species and compounds, the list of 188 HAPs includes several compound groups (e.g., individual metals and their compounds, polycyclic organic matter (POM), and glycol ethers). Many of the uses of the NEI depend upon data for individual compounds within these groups rather than aggregated data on each group as a whole.

Table 13 lists all of the specific pollutants and compound groups included in the 1999 NTI along with their Chemical Abstract Services (CAS) numbers (for individual compounds). For risk assessments, individual speciated HAPs are needed because the toxicity associated with an individual compound within a compound group varies widely. For trends purposes, we are also asked to track the emissions over time for the Section 112b list. For modeling files generated from the NEI, we report speciated pollutants using the NIF Code Table. For trends analyses and summary NEI data, we prepare files using the Section 112b HAPs and aggregated speciated NEI pollutants using the HAP Category Name. The HAP Category Name is the same name for HAPs that are not listed as compound groups in Section 112b, for example, formaldehyde. For a pollutant such as methylmercury, the HAP Category Name is mercury and compounds.

Table 13. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
79345	1,1,2,2-Tetrachloroethane	1,1,2,2-Tetrachloroethane
79005	1,1,2-Trichloroethane	1,1,2-Trichloroethane
57147	1,1-Dimethyl Hydrazine	1,1-Dimethylhydrazine
58899	1,2,3,4,5,6-Hexachlorocyclyhexane	1,2,3,4,5,6-Hexachlorocyclyhexane (All Stereo Isomers, Including Lindane)
120821	1,2,4-Trichlorobenzene	1,2,4-Trichlorobenzene
96128	1,2-Dibromo-3-Chloropropane	1,2-Dibromo-3-Chloropropane
122667	1,2-Diphenylhydrazine	1,2-Diphenylhydrazine
106887	1,2-Epoxybutane	1,2-Epoxybutane
75558	1,2-Propylenimine	1,2-Propylenimine (2-Methylaziridine)
106990	1,3-Butadiene	1,3-Butadiene
542756	1,3-Dichloropropene	1,3-Dichloropropene
1120714	1,3-Propanesultone	1,3-Propane Sultone
106467	1,4-Dichlorobenzene	1,4-Dichlorobenzene
540841	2,2,4-Trimethylpentane	2,2,4-Trimethylpentane
95954	2,4,5-Trichlorophenol	2,4,5-Trichlorophenol
88062	2,4,6-Trichlorophenol	2,4,6-Trichlorophenol
94757	2,4-Dichlorophenoxy Acetic Acid	2,4-D (2,4-Dichlorophenoxyacetic Acid)(Including Salts And Esters)
51285	2,4-Dinitrophenol	2,4-Dinitrophenol
121142	2,4-Dinitrotoluene	2,4-Dinitrotoluene
584849	2,4-Toluene Diisocyanate	2,4-Toluene Diisocyanate
53963	2-Acetylaminofluorene	2-Acetylaminofluorene
532274	2-Chloroacetophenone	2-Chloroacetophenone
79469	2-Nitropropane	2-Nitropropane
91941	3,3'-Dichlorobenzidene	3,3'-Dichlorobenzidene
119904	3,3'-Dimethoxybenzidine	3,3'-Dimethoxybenzidine

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
119937	3,3'-Dimethylbenzidine	3,3'-Dimethylbenzidine
101144	4,4'-Methylenebis(2-Chloraniline)	4,4'-Methylenebis(2-Chloroaniline)
101779	4,4'-Methylenedianiline	4,4'-Methylenedianiline
101688	4,4'-Methylenediphenyl Diisocyanate	4,4'-Methylenediphenyl Diisocyanate (MDI)
534521	4,6-Dinitro-o-Cresol	4,6-Dinitro-o-Cresol (Including Salts)
92671	4-Aminobiphenyl	4-Aminobiphenyl
60117	4-Dimethylaminoazobenzene	4-Dimethylaminoazobenzene
92933	4-Nitrobiphenyl	4-Nitrobiphenyl
100027	4-Nitrophenol	4-Nitrophenol
75070	Acetaldehyde	Acetaldehyde
60355	Acetamide	Acetamide
75058	Acetonitrile	Acetonitrile
98862	Acetophenone	Acetophenone
107028	Acrolein	Acrolein
79061	Acrylamide	Acrylamide
79107	Acrylic Acid	Acrylic Acid
107131	Acrylonitrile	Acrylonitrile
107051	Allyl Chloride	Allyl Chloride
62533	Aniline	Aniline
10025919	Antimony Trichloride	Antimony Compounds
1309644	Antimony Trioxide	Antimony Compounds
1327339	Antimony Oxide	Antimony Compounds
1345046	Antimony Trisulfide	Antimony Compounds
16925250	Sodium hexafluoroantimenate	Antimony Compounds
7440360	Antimony	Antimony Compounds
7783702	Antimony Pentafluoride	Antimony Compounds
92	Antimony & Compounds	Antimony Compounds
1303282	Arsenic Pentoxide	Arsenic Compounds(Inorganic Including Arsine)
1327522	Arsenic Acid	Arsenic Compounds(Inorganic Including Arsine)
1327533	Arsenic Trioxide	Arsenic Compounds(Inorganic Including Arsine)
3141126	Arsenous Acid	Arsenic Compounds(Inorganic Including Arsine)
7440382	Arsenic	Arsenic Compounds(Inorganic Including Arsine)
7784421	Arsine	Arsenic Compounds(Inorganic Including Arsine)
93	Arsenic & Compounds (Inorganic Including Arsine)	Arsenic Compounds(Inorganic Including Arsine)
1332214	Asbestos	Asbestos
71432	Benzene	Benzene (Including Benzene From Gasoline)
92875	Benzidine	Benzidine
98077	Benzotrichloride	Benzotrichloride
100447	Benzyl Chloride	Benzyl Chloride
109	Beryllium & Compounds	Beryllium Compounds
1304569	Beryllium Oxide	Beryllium Compounds
13510491	Beryllium Sulfate	Beryllium Compounds
7440417	Beryllium	Beryllium Compounds
7787497	Beryllium Fluoride	Beryllium Compounds
57578	Beta-Propiolactone	Beta-Propiolactone

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
92524	Biphenyl	Biphenyl
117817	Bis(2-Ethylhexyl)Phthalate	Bis(2-Ethylhexyl)Phthalate (Dehp)
542881	Bis(Chloromethyl)Ether	Bis(Chloromethyl) Ether
75252	Bromoform	Bromoform
10108642	Cadmium Chloride	Cadmium Compounds
10124364	Cadmium Sulfate	Cadmium Compounds
10325947	Cadmium Nitrate	Cadmium Compounds
125	Cadmium & Compounds	Cadmium Compounds
1306190	Cadmium Oxide	Cadmium Compounds
1306236	Cadmium Sulfide	Cadmium Compounds
7440439	Cadmium	Cadmium Compounds
7790809	Cadmium Iodide	Cadmium Compounds
156627	Calcium Cyanamide	Calcium Cyanamide
133062	Captan	Captan
63252	Carbaryl	Carbaryl
75150	Carbon Disulfide	Carbon Disulfide
56235	Carbon Tetrachloride	Carbon Tetrachloride
463581	Carbonyl Sulfide	Carbonyl Sulfide
120809	Catechol	Catechol
133904	Chloramben	Chloramben
57749	Chlordane	Chlordane
7782505	Chlorine	Chlorine
79118	Chloroacetic Acid	Chloroacetic Acid
108907	Chlorobenzene	Chlorobenzene
510156	Chlorobenzilate	Chlorobenzilate
67663	Chloroform	Chloroform
107302	Chloromethyl Methyl Ether	Chloromethyl Methyl Ether
126998	Chloroprene	Chloroprene
10034829	Sodium Chromate(VI)	Chromium Compounds
10060125	Chromium Chloride	Chromium Compounds
10101538	Chromic Sulfate	Chromium Compounds
10294403	Barium Chromate	Chromium Compounds
10588019	Sodium Dichromate	Chromium Compounds
11103869	Zinc Potassium Chromate	Chromium Compounds
11115745	Chromic Acid	Chromium Compounds
12018018	Chromium Dioxide	Chromium Compounds
12018198	Chromium Zinc Oxide	Chromium Compounds
1308130	Zinc Chromate	Chromium Compounds
1308141	Chromium Hydroxide	Chromium Compounds
1308389	Chromic Oxide	Chromium Compounds
1333820	Chromium Trioxide	Chromium Compounds
13530659	Zinc Chromate	Chromium Compounds
13530682	Chromic Sulfuric Acid	Chromium Compounds
136	Chromium & Compounds	Chromium Compounds
13765190	Calcium Chromate	Chromium Compounds
14307358	Lithium Chromate	Chromium Compounds

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
14977618	Chromyl Chloride	Chromium Compounds
16065831	Chromium III	Chromium Compounds
18540299	Chromium (VI)	Chromium Compounds
21679312	Chromium (III)-AA	Chromium Compounds
50922297	Zinc Chromite	Chromium Compounds
7440473	Chromium	Chromium Compounds
7738945	Chromic Acid (VI)	Chromium Compounds
7775113	Sodium Chromate	Chromium Compounds
7778509	Potassium Dichromate	Chromium Compounds
7788967	Chromyl Fluoride	Chromium Compounds
7789006	Potassium Chromate	Chromium Compounds
7789062	Strontium Chromate	Chromium Compounds
7789095	Ammonium Dichromate	Chromium Compounds
10124433	Cobalt Sulfate	Cobalt Compounds
1307966	Cobalt Oxide	Cobalt Compounds
1308061	Cobalt Oxide (II,III)	Cobalt Compounds
1317426	Cobalt Sulfide	Cobalt Compounds
1345160	Cobalt Aluminate	Cobalt Compounds
136527	Cobalt 2-ethylhexanoate	Cobalt Compounds
139	Cobalt & Compounds	Cobalt Compounds
16842038	Cobalt Carbonate	Cobalt Compounds
61789513	Cobalt Naphtha	Cobalt Compounds
618	Cobalt Hydrocarbonyl	Cobalt Compounds
7440484	Cobalt	Cobalt Compounds
140	Coke Oven Emissions	Coke Oven Emissions
141	Benzene Soluble Organics (BSO)	Coke Oven Emissions
142	Methylene Chloride Soluble Organics (MCSO)	Coke Oven Emissions
8007452	Coal Tar	Coke Oven Emissions
106445	p-Cresol	Cresol/Cresylic Acid (Mixed Isomers)
108394	m-Cresol	Cresol/Cresylic Acid (Mixed Isomers)
1319773	Cresol	Cresol/Cresylic Acid (Mixed Isomers)
331	Cresols (Includes o, m, & p)/Cresylic Acids	Cresol/Cresylic Acid (Mixed Isomers)
95487	o-Cresol	Cresol/Cresylic Acid (Mixed Isomers)
98828	Cumene	Cumene
13943583	Potassium Ferrocyani	Cyanide Compounds
13967505	Gold (I) Potassium Cyanide	Cyanide Compounds
140294	Benzyl Cyanide	Cyanide Compounds
14220178	Potass Nickel Cyanid	Cyanide Compounds
143339	Sodium Cyanide	Cyanide Compounds
144	Cyanide & Compounds	Cyanide Compounds
151508	Potassium Cyanide	Cyanide Compounds
37187647	Gold Cyanide	Cyanide Compounds
506649	Silver Cyanide	Cyanide Compounds
544923	Copper Cyanide	Cyanide Compounds

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
554074	Gold Potassium Cyanide	Cyanide Compounds
557211	Zinc Cyanide	Cyanide Compounds
57125	Cyanide	Cyanide Compounds
74908	Hydrogen Cyanide	Cyanide Compounds
78820	2-Methyl-Propanenitrile	Cyanide Compounds
72559	Dde (1,1-Dichloro-2,2-Bis(p-Chlorophenyl) Ethylene)	Dde (1,1-Dichloro-2,2-Bis(p- Chlorophenyl) Ethylene)
334883	Diazomethane	Diazomethane
132649	Dibenzofuran	Dibenzofuran
84742	Dibutyl Phthalate	Dibutyl Phthalate
111444	Dichloroethyl Ether	Dichloroethyl Ether (Bis[2-Chloroethyl]Ether)
62737	Dichlorvos	Dichlorvos
111422	Diethanolamine	Diethanolamine
64675	Diethyl Sulfate	Diethyl Sulfate
131113	Dimethyl Phthalate	Dimethyl Phthalate
77781	Dimethyl Sulfate	Dimethyl Sulfate
79447	Dimethylcarbamoyl Chloride	Dimethylcarbamoyl Chloride
155	Dioxins	Dioxins/Furans (total, non TEQ)
262124	Dibenzo-p-Dioxin	Dioxins/Furans (total, non TEQ)
30402143	Total Tetrachlorodibenzofuran	Dioxins/Furans (total, non TEQ)
30402154	Total Pentachlorodibenzofuran	Dioxins/Furans (total, non TEQ)
34465468	Hexachlorodibenzo-p-Dioxin	Dioxins/Furans (total, non TEQ)
36088229	Total Pentachlorodibenzo-p-Dioxin	Dioxins/Furans (total, non TEQ)
37871004	Total Heptachlorodibenzo-p-Dioxin	Dioxins/Furans (total, non TEQ)
38998753	Total Heptachlorodibenzofuran	Dioxins/Furans (total, non TEQ)
41903575	Total Tetrachlorodibenzo-p-Dioxin	Dioxins/Furans (total, non TEQ)
55684941	Total Hexachlorodibenzofuran	Dioxins/Furans (total, non TEQ)
609	Dibenzofurans (Chlorinated) {PCDFs}	Dioxins/Furans (total, non TEQ)
610	Dioxins, Total, w/o Individ. Isomers Reported {PCDDs}	Dioxins/Furans (total, non TEQ)
622	Hexachlorodibenzo-p-Dioxins, Total	Dioxins/Furans (total, non TEQ)
623	Polychlorinated Dibenzo-p-Dioxins, Total	Dioxins/Furans (total, non TEQ)
624	Polychlorinated Dibenzofurans, Total	Dioxins/Furans (total, non TEQ)
1746016	2,3,7,8-Tetrachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
19408743	1,2,3,7,8,9-Hexachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
3268879	Octachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
35822469	1,2,3,4,6,7,8-Heptachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
39001020	Octachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
39227286	1,2,3,4,7,8-Hexachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
40321764	1,2,3,7,8-Pentachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
51207319	2,3,7,8-Tetrachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
55673897	1,2,3,4,7,8,9-Heptachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
57117314	2,3,4,7,8-Pentachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
57117416	1,2,3,7,8-Pentachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
57117449	1,2,3,6,7,8-Hexachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
57653857	1,2,3,6,7,8-Hexachlorodibenzo-p-Dioxin	Dioxins/Furans as 2,3,7,8-TCDD TEQs
600	2,3,7,8-TCDD TEQ	Dioxins/Furans as 2,3,7,8-TCDD TEQs
60851345	2,3,4,6,7,8-Hexachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
67562394	1,2,3,4,6,7,8-Heptachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
70648269	1,2,3,4,7,8-Hexachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
72918219	1,2,3,7,8,9-Hexachlorodibenzofuran	Dioxins/Furans as 2,3,7,8-TCDD TEQs
106898	1-Chloro-2,3-Epoxypropane	Epichlorohydrin (1-Chloro-2,3-Epoxypropane)
140885	Ethyl Acrylate	Ethyl Acrylate
51796	Ethyl Carbamate Chloride	Ethyl Carbamate (Urethane) Chloride (Chloroethane)
75003	Ethyl Chloride	Ethyl Chloride
100414	Ethyl Benzene	Ethylbenzene
106934	Ethylene Dibromide	Ethylene Dibromide (Dibromoethane)
107062	Ethylene Dichloride	Ethylene Dichloride (1,2-Dichloroethane)
107211	Ethylene Glycol	Ethylene Glycol
75218	Ethylene Oxide	Ethylene Oxide
96457	Ethylene Thiourea	Ethylene Thiourea
151564	Ethyleneimine	Ethyleneimine (Aziridine)
75343	Ethylidene Dichloride (1,1-Dichloroethane)	Ethylidene Dichloride (1,1-Dichloroethane)
383	Fine Mineral Fibers	Fine Mineral Fibers
608	Ceramic Fibers (Man-Made)	Fine Mineral Fibers
613	Glasswool (Man-Made Fibers)	Fine Mineral Fibers
616	Slagwool (Man-Made Fibers)	Fine Mineral Fibers
617	Rockwool (Man-Made Fibers)	Fine Mineral Fibers
50000	Formaldehyde	Formaldehyde
1002671	Diethylene Glycol Ethyl Methyl Ether	Glycol Ethers
10137969	Ethyleneglycol Mono-2-Methylpentyl Ether	Glycol Ethers
10137981	Ethyleneglycolmono-2,6,8-Trimethyl-4-Nonyl Ether	Glycol Ethers
10143530	Diethylene Glycol Ethylvinyl Ether	Glycol Ethers
10143541	Diethylene Glycol Mono-2-Cyanoethyl Ether	Glycol Ethers
10143563	Diethyleneglycol-Mono-2-Methyl-Pentyl Ether	Glycol Ethers
10215335	3-Butoxy-1-Propanol	Glycol Ethers
109864	Ethylene Glycol Methyl Ether	Glycol Ethers
110496	Ethylene Glycol Monomethyl Ether Acetate	Glycol Ethers
110714	1,2-Dimethoxyethane	Glycol Ethers
110805	Cellosolve Solvent	Glycol Ethers
111104	Methoxyethyl Oleate	Glycol Ethers
111159	Cellosolve Acetate	Glycol Ethers
111762	Butyl Cellosolve	Glycol Ethers
111773	Diethylene Glycol Monomethyl Ether	Glycol Ethers

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
111900	Diethylene Glycol Monoethyl Ether	Glycol Ethers
111966	Diethylene Glycol Dimethyl Ether	Glycol Ethers
112072	2-Butoxyethyl Acetate	Glycol Ethers
112152	Carbitol Acetate	Glycol Ethers
112254	2-(Hexyloxy)Ethanol	Glycol Ethers
112276	Triethylene glycol	Glycol Ethers
112345	Diethylene Glycol Monobutyl Ether	Glycol Ethers
112356	Methoxytriglycol	Glycol Ethers
112367	Diethylene glycol diethyl ether	Glycol Ethers
112492	Triethylene Glycol Dimethyl Ether	Glycol Ethers
112505	Ethoxytriglycol	Glycol Ethers
112594	N-Hexyl Carbitol	Glycol Ethers
120558	Diethylene Glycol Dibenzoate	Glycol Ethers
122996	Phenyl Cellosolve	Glycol Ethers
124174	Butyl Carbitol Acetate	Glycol Ethers
140056	Methyl Cellosolve Acetylricinoleate	Glycol Ethers
143226	Triglycol Monobutyl Ether	Glycol Ethers
1589497	3-Methoxy-1-Propanol	Glycol Ethers
16672392	Di(Ethylene Glycol Monobutyl Ether) Phthalate	Glycol Ethers
171	Glycol Ethers	Glycol Ethers
18912806	Diethylene Glycol Monoisobutyl Ether	Glycol Ethers
20706256	2-Propoxyethyl Acetate	Glycol Ethers
23436193	1-Isobutoxy-2-Propanol	Glycol Ethers
23495127	Ethyleneglycol Monophenyl Ether Propionate	Glycol Ethers
2807309	Propyl Cellosolve	Glycol Ethers
3121617	Methyl Cellosolve Acrylate	Glycol Ethers
3775857	Ethylene Glycol Bis(2,3-Epoxy-2-Methylpropyl) Ether	Glycol Ethers
4206615	Diethylene Glycol Diglycidyl Ether	Glycol Ethers
4439241	Isobutyl Cellosolve	Glycol Ethers
629141	Ethylene Glycol Diethyl Ether	Glycol Ethers
662082	Ethylene Glycol Monobenzyl Ether	Glycol Ethers
67425	(Ethylenebis(Oxyethylenenitrilo)) Tetraacetic Acid	Glycol Ethers
693210	Diethylene Glycol Dinitrate	Glycol Ethers
7529273	Ethylene Glycol Diallyl Ether	Glycol Ethers
764487	Ethylene Glycol Monovinyl Ether	Glycol Ethers
764998	Diethylene Glycol Divinyl Ether	Glycol Ethers
7795917	Ethylene Glycol Mono-Sec-Butyl Ether	Glycol Ethers
929373	Diethylene Glycol Monovinyl Ether	Glycol Ethers
76448	Heptachlor	Heptachlor
118741	Hexachlorobenzene	Hexachlorobenzene
87683	Hexachlorobutadiene	Hexachlorobutadiene
77474	Hexachlorocyclopentadiene	Hexachlorocyclopentadiene

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
67721	Hexachloroethane	Hexachloroethane
822060	Hexamethylene Diisocyanate	Hexamethylene Diisocyanate
680319	Hexamethylphosphoramide	Hexamethylphosphoramide
110543	Hexane	Hexane
302012	Hydrazine	Hydrazine
7647010	Hydrochloric Acid	Hydrochloric Acid (Hydrogen Chloride [Gas Only])
7664393	Hydrogen Fluoride	Hydrogen Fluoride (Hydrofluoric Acid)
123319	Hydroquinone	Hydroquinone
78591	Isophorone	Isophorone
10031137	Lead Arsenite	Lead Compounds
10099748	Lead Nitrate	Lead Compounds
12060003	Lead Titanate	Lead Compounds
12626812	Lead Titanate Zircon	Lead Compounds
1309600	Lead Dioxide	Lead Compounds
1314416	Lead (II, IV) Oxide	Lead Compounds
1317368	Lead (II) Oxide	Lead Compounds
1335257	Lead Oxide	Lead Compounds
1335326	Lead Subacetate	Lead Compounds
13814965	Lead Fluoroborate	Lead Compounds
18454121	Lead Chromate Oxide	Lead Compounds
195	Lead & Compounds	Lead Compounds
27253287	Lead Neodecanoate	Lead Compounds
301042	Lead Acetate	Lead Compounds
598630	Lead Carbonate	Lead Compounds
602	Lead Compounds (Inorganic)	Lead Compounds
603	Lead Compounds (Other Than Inorganic)	Lead Compounds
61790145	Lead Naphthenate	Lead Compounds
620	Lead Dioxide, Unknown CAS #	Lead Compounds
7428480	Lead Stearate	Lead Compounds
7439921	Lead	Lead Compounds
7446142	Lead Sulfate	Lead Compounds
7446277	Lead Phosphate	Lead Compounds
7758976	Lead Chromate	Lead Compounds
7784409	Lead Arsenate	Lead Compounds
78002	Tetraethyl Lead	Lead Compounds
88	Alkylated Lead	Lead Compounds
108316	Maleic Anhydride	Maleic Anhydride
10101505	Permanganic acid	Manganese Compounds
10377669	Manganese Nitrate	Manganese Compounds
1313139	Manganese Dioxide	Manganese Compounds
1317346	Manganese Trioxide	Manganese Compounds
1317357	Manganese Tetroxide	Manganese Compounds
1336932	Manganese Napthenate	Manganese Compounds
198	Manganese & Compounds	Manganese Compounds
7439965	Manganese	Manganese Compounds

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
7722647	Potassium permanganate	Manganese Compounds
7783166	Manganesehypophosphi	Manganese Compounds
7785877	Manganese Sulfate	Manganese Compounds
8030704	Manganese Tallate	Manganese Compounds
199	Mercury & Compounds	Mercury Compounds
200	Elemental Gaseous Mercury	Mercury Compounds
201	Gaseous Divalent Mercury	Mercury Compounds
202	Particulate Divalent Mercury	Mercury Compounds
22967926	Mercury (Organic)	Mercury Compounds
593748	Methyl Mercury	Mercury Compounds
62384	Mercury Acetato Phen	Mercury Compounds
7439976	Mercury	Mercury Compounds
7487947	Mercuric Chloride	Mercury Compounds
67561	Methanol	Methanol
72435	Methoxychlor	Methoxychlor
74839	Methyl Bromide	Methyl Bromide (Bromomethane)
74873	Methyl Chloride	Methyl Chloride (Chloromethane)
71556	Methyl Chloroform	Methyl Chloroform (1,1,1-Trichloroethane)
78933	Methyl Ethyl Ketone	Methyl Ethyl Ketone (2-Butanone)
74884	Methyl Iodide	Methyl Iodide (Iodomethane)
108101	Methyl Isobutyl Ketone	Methyl Isobutyl Ketone (Hexone)
624839	Methyl Isocyanate	Methyl Isocyanate
80626	Methyl Methacrylate	Methyl Methacrylate
1634044	Methyl Tert-Butyl Ether	Methyl Tert-Butyl Ether
75092	Methylene Chloride	Methylene Chloride (Dichloromethane)
60344	Methylhydrazine	Methylhydrazine
121697	N,N-Dimethylaniline	N,N-Dimethylaniline
68122	N,N-Dimethylformamide	N,N-Dimethylformamide
91203	Naphthalene	Naphthalene
10101970	Nickel (II) Sulfate Hexahydrate	Nickel Compounds
12035722	Nickel Subsulfide	Nickel Compounds
12054487	Nickel Hydroxide	Nickel Compounds
12710360	Nickel Carbide	Nickel Compounds
1271289	Nickelocene	Nickel Compounds
13138459	Nickel Nitrate	Nickel Compounds
1313991	Nickel Oxide	Nickel Compounds
1314063	Nickel Peroxide	Nickel Compounds
13462889	Nickel Bromide	Nickel Compounds
13463393	Nickel Carbonyl	Nickel Compounds
13770893	Nickel Sulfamate	Nickel Compounds
226	Nickel & Compounds	Nickel Compounds
3333393	Nickel Carbonate	Nickel Compounds
373024	Nickel Acetate	Nickel Compounds
6018899	Nickel Diacetate TET	Nickel Compounds
604	Nickel Refinery Dust	Nickel Compounds
7440020	Nickel	Nickel Compounds

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
7718549	Nickel Chloride	Nickel Compounds
7786814	Nickel Sulfate	Nickel Compounds
NY059280	Nickel (NI 059)	Nickel Compounds
98953	Nitrobenzene	Nitrobenzene
62759	N-Nitrosodimethylamine	N-Nitrosodimethylamine
59892	N-Nitrosomorpholine	N-Nitrosomorpholine
684935	N-Nitroso-N-Methylurea	N-Nitroso-N-Methylurea
90040	o-Anisidine	o-Anisidine
95534	o-Toluidine	o-Toluidine
56382	Parathion	Parathion
123911	p-Dioxane	p-Dioxane
82688	Pentachloronitrobenzene	Pentachloronitrobenzene (Quintobenzene)
87865	Pentachlorophenol	Pentachlorophenol
108952	Phenol	Phenol
75445	Phosgene	Phosgene
7803512	Phosphine	Phosphine
10025873	Phosphorus Oxychloride	Phosphorus Compounds
101020	Triphenyl Phosphite	Phosphorus Compounds
10294561	Phosphorous Acid	Phosphorus Compounds
115866	Triphenyl Phosphate	Phosphorus Compounds
12136913	Phosphorous Nitride	Phosphorus Compounds
13011546	Phosphorous Salt	Phosphorus Compounds
1314245	Phosphorus Trioxide	Phosphorus Compounds
1314563	Phosphorus Pentoxide	Phosphorus Compounds
1314803	Phosphorus Pentasulfide	Phosphorus Compounds
2921882	Phosphorothioic Acid	Phosphorus Compounds
398	Phosphorus & Compounds	Phosphorus Compounds
7664382	Phosphoric Acid	Phosphorus Compounds
7719122	Phosphorus Trichloride	Phosphorus Compounds
7723140	Phosphorus	Phosphorus Compounds
7779900	Zinc Phosphate	Phosphorus Compounds
78308	Triorthocresyl Phosphate	Phosphorus Compounds
92203026	Phosphoric Acid,Rx P	Phosphorus Compounds
85449	Phthalic Anhydride	Phthalic Anhydride
1336363	Polychlorinated Biphenyls	Polychlorinated Biphenyls (Aroclors)
189559	Dibenzo[a,i]Pyrene	Polycyclic Organic Matter as non-15-PAH
189640	Dibenzo[a,h]Pyrene	Polycyclic Organic Matter as non-15-PAH
191300	Dibenzo[a,l]Pyrene	Polycyclic Organic Matter as non-15-PAH
192654	Dibenzo[a,e]Pyrene	Polycyclic Organic Matter as non-15-PAH
192972	Benzo[e]Pyrene	Polycyclic Organic Matter as non-15-PAH
195197	Benzo(c)phenanthrene	Polycyclic Organic Matter as non-15-PAH
198550	Perylene	Polycyclic Organic Matter as non-15-PAH
203123	Benzo(g,h,i)Fluoranthene	Polycyclic Organic Matter as non-15-PAH
203338	Benzo(a)fluoranthene	Polycyclic Organic Matter as non-15-PAH
205823	B[j]Fluoranthen	Polycyclic Organic Matter as non-15-PAH
224420	Dibenzo[a,j]Acridine	Polycyclic Organic Matter as non-15-PAH

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
2381217	1-Methylpyrene	Polycyclic Organic Matter as non-15-PAH
2422799	12-Methylbenz(a)Anthracene	Polycyclic Organic Matter as non-15-PAH
247	Methylbenzopyrenes	Polycyclic Organic Matter as non-15-PAH
248	Methylchrysene	Polycyclic Organic Matter as non-15-PAH
26914181	Methylanthracene	Polycyclic Organic Matter as non-15-PAH
284	Extractable Organic Matter (EOM)	Polycyclic Organic Matter as non-15-PAH
3697243	5-Methylchrysene	Polycyclic Organic Matter as non-15-PAH
40	16-PAH	Polycyclic Organic Matter as non-15-PAH
42397648	1,6-Dinitropyrene	Polycyclic Organic Matter as non-15-PAH
42397659	1,8-Dinitropyrene	Polycyclic Organic Matter as non-15-PAH
5522430	1-Nitropyrene	Polycyclic Organic Matter as non-15-PAH
56495	3-Methylcholanthrene	Polycyclic Organic Matter as non-15-PAH
57976	7,12-Dimethylbenz[a]Anthracene	Polycyclic Organic Matter as non-15-PAH
607578	2-Nitrofluorene	Polycyclic Organic Matter as non-15-PAH
625	Naphthenes (Cyclo)	Polycyclic Organic Matter as non-15-PAH
7496028	6-Nitrochrysene	Polycyclic Organic Matter as non-15-PAH
779022	9-Methylbenz(a)Anthracene	Polycyclic Organic Matter as non-15-PAH
832699	1-Methylphenanthrene	Polycyclic Organic Matter as non-15-PAH
90120	1-Methylnaphthalene	Polycyclic Organic Matter as non-15-PAH
91576	2-Methylnaphthalene	Polycyclic Organic Matter as non-15-PAH
91587	2-Chloronaphthalene	Polycyclic Organic Matter as non-15-PAH
120127	Anthracene	Polycyclic Organic Matter as 15-PAH
129000	Pyrene	Polycyclic Organic Matter as 15-PAH
191242	Benzo[g,h,i,]Perylene	Polycyclic Organic Matter as 15-PAH
206440	Fluoranthene	Polycyclic Organic Matter as 15-PAH
208968	Acenaphthylene	Polycyclic Organic Matter as 15-PAH
83329	Acenaphthene	Polycyclic Organic Matter as 15-PAH
85018	Phenanthrene	Polycyclic Organic Matter as 15-PAH
86737	Fluorene	Polycyclic Organic Matter as 15-PAH
102	Benzo[b+k]Fluoranthene	Polycyclic Organic Matter as 7-PAH
103	Benz(a)Anthracene/Chrysene	Polycyclic Organic Matter as 7-PAH
193395	Indeno[1,2,3-c,d]Pyrene	Polycyclic Organic Matter as 7-PAH
205992	Benzo[b]Fluoranthene	Polycyclic Organic Matter as 7-PAH
207089	Benzo[k]Fluoranthene	Polycyclic Organic Matter as 7-PAH
218019	Chrysene	Polycyclic Organic Matter as 7-PAH
234	PAH, Total	Polycyclic Organic Matter as 7-PAH
246	Polycyclic Organic Matter	Polycyclic Organic Matter as 7-PAH
50328	Benzo[a]Pyrene	Polycyclic Organic Matter as 7-PAH
53703	Dibenzo[a,h]Anthracene	Polycyclic Organic Matter as 7-PAH
56553	Benz[a]Anthracene	Polycyclic Organic Matter as 7-PAH
56832736	Benzofluoranthenes	Polycyclic Organic Matter as 7-PAH
75	7-PAH	Polycyclic Organic Matter as 7-PAH
106503	p-Phenylenediamine	p-Phenylenediamine
123386	Propionaldehyde	Propionaldehyde
114261	Propoxur	Propoxur (Baygon)
78875	Propylene Dichloride	Propylene Dichloride (1,2-Dichloropropane)

Table 13, continued. 1999 NEI HAPs

NEI Pollutant Code	NEI Pollutant Name	HAP Category
75569	Propylene Oxide	Propylene Oxide
91225	Quinoline	Quinoline
106514	Quinone	Quinone (p-Benzoquinone)
24267569	Iodine-131	Radionuclides (Including Radon)
400	Radionuclides (Including Radon)	Radionuclides (Including Radon)
605	Radionuclides	Radionuclides (Including Radon)
606	Radon And Its Decay Products	Radionuclides (Including Radon)
7440611	Uranium	Radionuclides (Including Radon)
12640890	Selenium Oxide	Selenium Compounds
253	Selenium & Compounds	Selenium Compounds
7446084	Selenium Dioxide	Selenium Compounds
7446346	Selenium Monosulfide	Selenium Compounds
7488564	Selenium Disulfide	Selenium Compounds
7782492	Selenium	Selenium Compounds
7783008	Selenous Acid	Selenium Compounds
7783791	Selenium Hexafluoride	Selenium Compounds
100425	Styrene	Styrene
96093	Styrene Oxide	Styrene Oxide
127184	Tetrachloroethylene	Tetrachloroethylene (Perchloroethylene)
7550450	Titanium Tetrachloride	Titanium Tetrachloride
108883	Toluene	Toluene
95807	Toluene-2,4-Diamine	Toluene-2,4-Diamine
8001352	Toxaphene	Toxaphene (Chlorinated Camphene)
79016	Trichloroethylene	Trichloroethylene
121448	Triethylamine	Triethylamine
1582098	Trifluralin	Trifluralin
108054	Vinyl Acetate	Vinyl Acetate
593602	Vinyl Bromide	Vinyl Bromide
75014	Vinyl Chloride	Vinyl Chloride
75354	Vinylidene Chloride	Vinylidene Chloride (1,1-Dichloroethylene)
106423	p-Xylene	Xylenes (Mixed Isomers)
108383	m-Xylene	Xylenes (Mixed Isomers)
1330207	Xylenes (Mixture of o, m, and p Isomers)	Xylenes (Mixed Isomers)
95476	o-Xylene	Xylenes (Mixed Isomers)

Preferred Reporting of HAPs

We encourage data providers to report emissions for specific compounds, both for individual HAP species and for pollutants within compound groups. CAS numbers are preferred to identify pollutants reported to the EPA. If data providers cannot report pollutants using CAS numbers, aggregated compound group emissions will be accepted. If emissions are reported for groups of compounds, then we will use simplifying assumptions regarding speciation within the group in order to use these data as inputs to models.

Recommendations for reporting data for specific groups of compounds are summarized below in a hierarchy of most preferred method to least preferred. For pollutant groups, only one reporting strategy per HAP group per source should be used. Simultaneous use of more than one reporting strategy (e.g., reporting both individual chromium compounds and total chromium for the same source) may result in the same emission being counted twice resulting in a potential overestimation of emission levels and risk. CAS numbers referenced to in the text are also shown in Table 13 with their associated pollutant names.

- **Metal and cyanide groups**:

1. Report emissions and associated CAS numbers of all individual metal and cyanide compounds; e.g., report emissions and associated CAS numbers of arsenic oxide, lead arsenate, etc., rather than emissions of arsenic compounds as a whole. All individual compounds should be reported as the mass of the total compounds, not just the metal within the compound.
2. If individual metal compounds cannot be reported, a less preferred method for chromium, lead, mercury and nickel is to separately report two forms of widely-varying toxicity. If you use this approach, report only the mass of emissions of the metal, not of the entire metal compound.
   Chromium – Separate chromium compounds into hexavalent chromium (CAS #18540299) and trivalent (CAS #1606583).
   Lead – Separate lead compounds into organic and inorganic.
   Mercury – Separate mercury compounds into particulate, gaseous elemental and gaseous divalent.
   Nickel - Separate nickel compounds into nickel subsulfide (CAS #12035722) and other nickel (CAS #7440020).
   Other metal and cyanide groups - Report total emissions of the group in terms of the mass of the metal or cyanide alone, and report under the CAS number of the metal or cyanide.
3. Alternatively, but far less preferred, report total emissions of the group in terms of mass of total emissions, and report under the pollutant group number for "metal and compounds" or "cyanide and compounds". Do not include metals or cyanide already reported using the more preferred methods above, in order to avoid the possibility of double counting emissions.

- **POM** : POM compounds are formed primarily from combustion and are present in the atmosphere in particulate and gaseous forms. Table 14 identifies individual POM Compounds of interest to EPA. POM compounds are generally grouped into POM reported as 7-PAH, POM reported as 15-PAH, and POM reported as non-15-PAH. In order to generate a total POM estimate, you must sum 15-PAH and non-15-PAH together. You should not sum 7-PAH and 15-PAH together as 15-PAH includes 7-PAH. Table 14 shows the specific compounds constituting 7-PAH and 15-PAH groups.

1. Report emissions and associated CAS numbers of as many individual POM compounds as possible, rather than as total PAH or total POM. Most important to report individually are the 7-PAH compounds listed in Table 14. We also encourage the reporting of other individual POM compounds for which cancer assessments are available (also listed in Table 14).
2. If you cannot report all individual PAHs, then report 7-PAH as a subgroup.

3. If you cannot report emissions of 7-PAH, then report total POM (total POM includes total PAH). Since naphthalene is listed individually as a HAP, do not include any individually-reported naphthalene as total POM.
4. If you follow any other scheme than one listed in this hierarchy, clearly identify what it is.

Table 14. POM Compounds

7-PAH	15-PAH (includes 7-PAH compounds)	Other POM Compounds for which we have cancer assessments
Benz(a)anthracene	Acenaphthene	Carbazole
Benzo(a)pyrene	Acenaphthylene	Dibenz(a,h)acridine
Benzo(b)fluoranthene	Anthracene	Dibenz(a,j)acridine
Benzo(k)fluoranthene	Benzo(ghi)perylene	7H-Dibenzo(c,g)carbazole
Chrysene	Fluoranthene	Dibenzo(a,e)pyrene
Dibenz(a,h)anthracene	Fluorene	Dibenzo(a,i)pyrene
Indeno(1,2,3-cd)pyrene	Phenanthrene	Dibenzo(a,l)pyrene
	Pyrene	7,12-Dimethylbenz(a)anthracene
		1,6-Dinitropyrene
		1,8-Dinitropyrene
		3-Methylcholanthrene
		5-Methylchrysene
		5-Nitroacenaphthalene
		6-Nitrochrysene
		2-Nitrofluorene
		1-Nitropyrene
		4-Nitropyrene

- **Dioxins/Furans**: Dioxins and Furans can be grouped by 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) equivalents (TEQ). TEQ factors are multipliers for some dioxin and furan congeners to estimate a common basis of toxicity for chlorinated dioxins and furans. In the NEI, we group

dioxins and fuans into TEQ and non-TEQ groups (See Table 13). There are 2 sets of TEQ factors used to estimate TEQ - NATO TEQ factors and WHO TEQ factors. TEQ factors change as better toxicity data become available.

1. Report mass emissions and associated CAS numbers of all individual congeners of both chlorinated dibenzodioxins (CDDs) and chlorinated dibenzofurans (CDFs).
2. If you cannot report individual CDD and CDF congeners, report dioxins and furans as TEQ under the HAP name "dioxins/furans as TEQ". (Note: Although the CAA specifically lists only 2,3,7,8-TCDD as a HAP, other CDDs and CDFs qualify as HAPs within the POM group. Because some of these other congeners are also potent carcinogens, EPA will use the TEQ approach to evaluate CDDs and CDFs as a group.). Do not include dioxins/furans in the reporting of POM emissions. Identify the TEQ approach that you are using - I (NATO) TEQ or WHO TEQ.
3. If you cannot report emissions using the TEQ approach, report individual congener emissions where possible and report any remaining emissions as total "dioxins" or total "dibenzofurans".

- **Glycol Ethers***:* Glycol ethers include mono- and di- ethers of ethylene glycol, diethylene glycol, and triethylene glycol. Polymers are excluded for the glycol ether category.

1. Report emissions for individual glycol ethers with their associated CAS numbers. Ethylene Glycol Butyl Ether (EGBE) has been delisted as a HAP and should not be included in the glycol ethers group. A good reference for determining if a compound is a glycol ether is: http://daq.state.nc.us/toxics/glycol . Another reference is the Toxic Release Inventory (TRI) guidance on glycol ethers, http://www.epa.gov/opptintr/tri/glycol.pdf . Note that, historically, many compounds have been mistakenly included in the glycol ethers compound group. Table 15 identifies compounds that were reported to EPA as glycol ethers in the development of the 1999 NEI that are not actually glycol ethers and should not be included in 2002 NEI reporting.
2. If you cannot report individual glycol ether emissions, report total emissions of glycol ethers as a group under "glycol ethers".

- **Xylenes and Cresols***:* Xylenes include o-, m-, and p- isomers. Cresols include o-, m-, and p-isomers and cresylic acid.

1. Report emissions for individual xylene and cresol isomers with their associated CAS numbers. Do not report any emissions for total xylenes or cresols to avoid double counting
2. If you cannot report individual emissions of xylenes or cresols, report total emissions of xylenes or cresols as a group under "xylenes (mixture of o, m, and p isomers)" (CAS #1330207) or "cresols/cresylic acids".

Table 15. List of Compounds Often Mistaken as Glycol Ethers

Compound	CAS Number
1,1-Dimethoxyethane	534156
1-Ethoxy-2-propanol	1569024
3-Ethoxy-1-propanol	111353
Diethylene glycol	111466
Diethylene glycol di(3-aminopropyl) ether	4246519
Dipropylene glycol monomethyl ether	34590948
Glycols, polyethylene, polypropylene monobutyl ether (nonionic)	9038953
Isopropyl glycol	109591
Nonyl phenyl polyethylene glycol ether	9016459
Propylene glycol methyl ether acetate	108656
Propylene glycol monomethyl ether	107982
Propylene glycol t-butyl ether	57018527
Triethylene glycol	112276

QA of HAP Pollutant Codes

We will QA the reported Pollutant Codes. If the reported codes are incorrect, we will contact the data provider for a crosswalk of HAPs reported to CAS Numbers and pollutant names. If a reported HAP is not in the current NIF Pollutant Code Table, we will add the pollutant to the NIF Code Table.

We will only retain reported pollutants that are HAPs in the NEI. We will retain all reported records in our archive files.

NEI HAP Lookup Table

We will prepare a HAP Lookup Table with the following data fields.

- NEI Pollutant Code
- NEI Pollutant Name
- HAP Category Name
- Is Pollutant listed under Section 112k and 112c3 Urban Area Source Category Strategy that was published in July 1999 in Federal Register, 64FR38705. For more information, please see http://www.epa.gov/ttn/atw/urban/urbanbackground.html#A1 .

- I TEQ Factor
- WHO TEQ Factor
- VOC HAP
- PM HAP

Chemical Data Standards Lookup Table in the NEI

The EPA's Chemical Data Standard provides for the use of common identifiers throughout the Agency for all chemical substances regulated or monitored by EPA environmental programs. This standard provides unique, unambiguous, chemically correct common names for all chemicals substances and groupings in EPA's system and are provided here as they enable users to search for chemical substances across EPA programs and their databases. The Chemical Identification Standards Database is an excellent source of CAS numbers for HAPs. The Chemical Identification Standard consists of the following data elements:

- ***Chemical Abstracts Service Registry Number*** - the unique number assigned by Chemical Abstracts Service (CAS) to a chemical substance;
- ***Chemical Substance Systematic Name (9th Collective Index Name)*** - the name assigned to a chemical substance that describes it in terms of its molecular composition;
- ***EPA Chemical Identifier*** - The unique number assigned (in CRS) when CAS number not available; and
- ***EPA Chemical Registry Name*** - the name EPA has selected as the name to be commonly used by EPA in referring to a chemical substance.

We will compile a NEI Chemical Data Standards Table (Table 16) to meet EPA's Chemical Data Standards. The NEI Chemical Data Standards Table contains each of these elements for every NEI pollutant code and for the corresponding NEI HAP Category.

Table 16. NEI Chemical Data Standards Table

Field Name	Description
NEI Pollutant Code	Unique code assigned to NEI pollutant
NEI Pollutant Name	HAP name for NEI pollutant
NEI HAP Category	Grouping of related NEI pollutants
CASRN	Chemical Abstracts Service Registry Number
CASRN_compact	Chemical Abstracts Service Registry Number without dashes
ChemSystematicName	Chemical Substance Systematic Name (9th Collective Index Name)
EPAChemicalID	EPA Chemical Identifier

Table 16, Continued. NEI Chemical Data Standards Table

Field Name	Description
EPAChemRegistryName	EPA Chemical Registry Name
HAPCAT_CASRN	Chemical Abstracts Service Registry Number (assigned to HAP category)
HAPCAT_CASRN_compact	Chemical Abstracts Service Registry Number without dashes (assigned to HAP category)
HAPCAT_ChemSystematicName	Chemical Substance Systematic Name (9th Collective Index Name)(assigned to HAP category)
HAPCAT_EPAChemcialID	EPA Chemical Identifier (assigned to HAP category)
HAPCAT_EPAChemRegistryName	EPA Chemical Registry Name (assigned to HAP category)

CAPs Reporting

CAPs include: Volatile Organic Compounds (VOC), Nitrogen Oxides (NOX), Carbon Monoxide (CO), Sulfur Dioxide(SO2), Particulate Matter (PM), and Lead. Ammonia (NH3) will also be included in the 2002 CAP NEI. Table 17 describes PM compounds that may be reported to EPA by data providers for the 2002 NEI. We will generate emission estimates for filterable and primary PM-2.5, filterable and primary PM-10 and condensible PM in the 2002 NEI.

We will QA the reported Pollutant Codes. If the reported codes are incorrect, we will correct invalid codes.

We will QA reported CAP emissions to identify missing CAPs and invalid emissions. For each individual emission record, the sum of PM-condensibles and PM2.5 filterables cannot be greater than the sum of PM2.5-primary, and the sum of PM-condensibles and PM10-filterables cannot be greater than the sum of PM10-primary.

We will augment CAPs and their emissions for three cases in the 2002 NEI.

1. PM augmentation for filterable, condensible and primary PM-10 and PM-2.5
2. VOC and PM augmentation when HAPs are reported that are VOC or PM compounds and VOC and PM emissions are not reported
3. CAPs augmentation if a partial submission of CAPs is provided in a data submittal, e.g., a state agency only submits NOX.

Table 17. PM Compounds Reported in the 2002 NEI

NEI Code	Pollutant	Pollutant Description
PM-CON	Primary PM Condensible portion only (all < 1 micron)	Material that is vapor phase at stack conditions, but which condenses and/or reacts upon cooling and dilution in the ambient air to form solid or liquid PM immediately after discharge from the stack.
PM-FIL	Primary PM, Filterable portion only	Particles that are directly emitted by a source as a solid or liquid at stack or release conditions and captured on the filter of a stack test train.
PM-PRI	Primary PM, Includes filterables and condensibles PM-PRI= PM-FIL + PM-CON	Particles that enter the atmosphere as a direct emission from a stack or an open source. It is comprised of two components: Filterable PM and Condensible PM.
PM10-FIL	Primary PM10, Filterable portion only	Particles with an aerodynamic diameter equal to or less than 10 micrometers that are directly emitted by a source as a solid or liquid at stack or release conditions and captured on the filter of a stack test train.
PM10-PRI	Primary PM10, Includes filterables and condensibles, PM10- PRI = PM0-FIL + PM-CON	Particles with an aerodynamic diameter equal to or less than 10 micrometers that enter the atmosphere as a direct emission from a stack or an open source. It is comprised of two components: Filterable PM and Condensible PM. (As specified in § 51.15 (a)(2), these two PM components are the components measured by a stack sampling train such as EPA Method 5.)
PM25-FIL	Primary PM2.5, Filterable portion only	Particles with an aerodynamic diameter equal to or less than 2.5 micrometers that are directly emitted by a source as a solid or liquid at stack or release conditions and captured on the filter of a stack test train.
PM25-PRI	Primary PM2.5, Includes filterables and condensibles PM25-PRI= PM25-FIL + PM-CON	Particles with an aerodynamic diameter equal to or less than 2.5 micrometers that enter the atmosphere as a direct emission from a stack or an open source. It is comprised of two components: Filterable PM and Condensible PM. (As specified in § 51.15 (a)(2), these two PM components are the components measured by a stack sampling train such as EPA Method 5.)

12.2 PM Augmentation - This section will be developed in the fall of 2004

We will develop a methodology to default missing PM compounds and incorporate the methodology into this report later this year. Table 18 describes how we will default missing PM. We will develop a PM condensible default table by SCCs that includes a ratio of filterable PM to condensible PM. These ratios will be based on emission factors in AP-42/FIRE. All PM emissions estimates that are defaulted in the 2002 NEI will include default flags to identify that the emissions are defaulted.

Table 18. PM Augmentation in the 2002 NEI

PM Reported	PM to Default	Augmentation Methodology
PM25-PRI PM25-FIL PM10-PRI PM10-FIL PM-CON		
PM-PRI	PM-CON PM25-PRI PM25-FIL PM10-PRI PM10-FIL	PM-CON = PM-PRI x % Con PM-FIL = PM-PRI - PM-CON Generate PM10-FIL using PM-FIL with PM Calculator Generate PM25-FIL using PM-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL
PM10-PRI	PM-CON PM25-PRI PM25-FIL PM10-FIL	PM-CON = PM10-PRI x %Con PM10-FIL = PM10-PRI - PM-Con Generate PM25-FIL using PM10-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL
PM10-FIL	PM-CON PM25-PRI PM25-FIL PM10-PRI	PM-CON = PM10-FIL x ratio of PM-CON/PM10-FIL Generate PM25-FIL using PM10-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL
PM10-FIL PM10-PRI	PM-CON PM25-PRI PM25-FIL	PM-CON= PM10-PRI - PM10-FIL Generate PM25-FIL using PM10-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL
PM10-FIL PM-CON	PM10-PRI PM25-PRI PM25-FIL	Generate PM25-FIL using PM10-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL
PM-CON PM10-PRI	PM10-FIL PM25-PRI PM25-FIL	PM10-FIL= PM10-PRI - PM-CON Generate PM25-FIL using PM10-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL
PM25-PRI	PM-CON PM25-FIL PM10-PRI PM10-FIL	PM-CON = PM25-PRI x %Con PM25-FIL = PM25-PRI - PM-Con Generate PM10-FIL using PM25-FIL with PM Calculator PM10-PRI = PMCON + PM10-FIL

Table 18, continued. PM Augmentation in the 2002 NEI

PM Reported	PM to Default	Augmentation Methodology
PM25-FIL	PM-CON PM25-PRI PM10-FIL PM10-PRI	PM-CON = PM25-FIL x ratio of PM-CON/PM25-FIL Generate PM10-FIL using PM25-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL
PM25-FIL PM25-PRI	PM-CON PM10-PRI PM10-FIL	PM-CON= PM25-PRI - PM25-FIL Generate PM10-FIL using PM25-FIL with PM Calculator PM10-PRI = PMCON + PM10-FIL
PM25-FIL PM-CON	PM10-PRI PM25-PRI PM10-FIL	Generate PM10-FIL using PM25-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL
PM-CON PM25-PRI	PM25-FIL PM25-PRI PM10-FIL	PM25-FIL= PM25-PRI - PM-CON Generate PM10-FIL using PM25-FIL with PM Calculator PM10-PRI = PMCON + PM1025-FIL
PM10-FIL PM25-FIL	PM-CON PM25-PRI PM10-PRI	PM-CON = PM10-FIL x ratio of PM-CON/PM10-FIL PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL
PM-CON	PM25-PRI PM25-FIL PM10-PRI PM10-FIL	PM10-FIL = PM-CON x ratio of PM10-FIL/PM-CON Generate PM25-FIL using PM10-FIL with PM Calculator PM25-PRI = PMCON + PM25-FIL PM10-PRI = PMCON + PM10-FIL

12.3 VOC and PM Augmentation from HAPs

We will compare reported PM and VOC emissions with sum of VOC HAPs and sum of PM HAPs at the facility level. Facilities will be identified where the sum of VOC HAPs emissions is greater than reported VOC emissions and the sum of PM HAPs emissions is greater than reported PM emissions. We will use the HAP Lookup Table to identify which HAPs are VOC and/or PM constituents.

For facilities who have HAP VOC and HAP PM emissions but do not have VOC or PM emissions, we will augment CAP inventories to include the HAP VOC emissions and HAP PM emissions. We will include default flags to identify emissions where PM or VOC have been augmented using HAP data.

If the sum of HAP VOC or the sum of HAP PM emissions are more than 20% greater than reported VOC or PM emissions, we will compile a list of these facilities with their emissions data for further QA. It is not known at this time if we will be able to resolve the discrepancy. At a minimum, the table will need to contain the following information.

- Facility information
- Source category information

- VOC emissions from NEI for CAPs
- PM emissions from NEI from CAPs
- Sum of VOC HAPs emissions from NEI for HAPs
- Sum of PM HAPs emissions from NEI for HAPs
- % difference between VOC emissions from NEI for CAPs and sum of VOC emissions from NEI for HAPs
- % difference between PM emissions from NEI for CAPs and sum of PM emissions from NEI for HAPs
- Resolution if any

If VOC emissions are greater than HAP VOC emissions or PM emissions are greater than HAP PM emissions, we assume that VOC and PM emissions include all VOC HAP and PM HAP emissions plus emissions of additional non-HAPs.

12.4 CAP Augmentation for Missing CAPs

When a data provider submits a partial submission of CAPs for a facility, we will determine if emissions of other CAPs should be reported. If the data provider reports emissions of other CAPs as "0" and indicates in the documentation that "0" means no emissions as opposed to being a placeholder, we will not add missing CAPs. For missing CAPs, we will use the reported Throughput and SCC with available emission factors to generate estimates for other CAPs. If Throughput and SCCs are not reported, or emission factors are not available for a process, then we will use the 2002 Preliminary NEI for CAP Point Source estimates released in January 2004.

12.5 HAP Performance Level

Data providers may report HAP emissions as actual, allowable, potential or maximum. In Trends analyses, actual emissions are preferred whereas in risk assessments, maximum emissions are preferred.

More than one performance level may be reported for each pollutant and process. For example, an agency may reported actual and potential emissions of a single HAP from a process. When we prepare files for various uses, we use a hierarchy to select only one type of emissions to avoid double counting. For 1999 NATA modeling, the following hierarchy was used.

1. Actual
2. Allowable
3. Potential
4. Maximum

If this field is invalid or missing for HAP emission records, we will default the reported emissions as actual.

12.6 HAP Emissions QC

After we have blended and merged point source data, we will evaluate emissions records associated with facilities that have individual HAP compounds reported and HAP category groups reported (e.g., where chromium and compounds and hexavalent chromium are reported for the same facility and process). If duplicate emissions are identified, then we will delete records with the compound group emissions and retain speciated compounds. For cases where a single data source reports emissions of individual compounds and compound group, the following actions will be taken.

- If compound group emissions are less than the sum of emissions of individual compounds within a compound group, we will delete compound group emissions.

- If compound group emissions are greater than or equal to the sum of individual compounds within the group, we will take no action.

12.7 Emission Content QC

After we have compiled a draft point source inventory, we will conduct a variety of content QC on an internal draft inventory prior to preparing the draft inventory files for external review. QA/QC activities include identifying and correcting erroneous emissions data. For the most part, we will identify outliers with very high emissions estimates. We have developed a series of internal QA/QC activities to target outliers and duplicate emissions. We will conduct the following activities to QC the emissions data.

- Evaluation of significant changes between the 1999 NEI and 2002 NEI data, and/or extreme variation within the 2002 data. This will include comparing 1999 emission estimates to 2002 estimates by pollutant for each facility, source category, county, and state. These big picture summaries will highlight source categories, states, and facilities with potential problems.

- Identification of top emitters for each pollutant nationwide, ranking each facility based on its emissions of each pollutant on a national basis and listing the top emitters for pollutant/source category combination nationwide. Statistical tools and GIS will be used to perform these activities.

Outliers are usually difficult to spot - what appears to be a high emissions value may in fact be acceptable for a particular facility or source category. To aid in detecting these errors, the emissions data will be compared to the range of values in the NEI and the percent contribution to total emissions. A summary table with the list of facilities that appear multiple times as top emitters for different pollutants will help identify point sources with outliers. These high values may be due to a series of outliers or duplicated emission records. The high emissions may also be correct for that facility and category. Thus, these summary data will need to be closely reviewed before any records are deleted. In some cases, the state/local agency submitting the data will be contacted to discuss the quality of the estimates. Comparative QC will include analyses such as:

- "Top 10" lists by pollutant and source category - include highest emitting facilities and source categories nationally and within a state and county or region;
- Emissions ranked lists and maps by pollutant - include counties within a state ranked by pollutant and source category emissions; and
- Previous Year Comparison - compare pollutant emissions nationally and for state, county, source category and facility from previous inventories with the draft 2002 NEI.

Statistical QC will include analyses such as:

- Statistical tables (min, max, average, standard deviation, and percentile values) by source category - compare pollutant specific statistical values nationally and for state to individual counties for source categories and individual facilities;
- Emissions distribution by pollutant - prepare source category, national, county level, and state level emissions values frequency histogram; and
- Contribution of facilities and source categories to multiple pollutants - compare list of pollutants emitted from a source category to individual facilities.

We will use the following specific files to help QC the data.

- 99-2002 Site List - provides a list of sites found in the 1999 base year NEI, the 2002 draft, and sites common to both the 1999 and 2002 inventories. By sorting this list by state, county, and facility name, we will evaluate the sites listed in each county and detect potential duplicates. If a site is found in the 1999 version but is not in the 2002 draft, reviewers can verify that the facility closed.
- 2002 County Emissions Summary - provides a snapshot of the emissions for each pollutant in each county by sector. These summary tables will enable us to target states and counties for detailed evaluation and to prioritize our QC.
- 2002 Facility Emissions Summary- summarizes emissions of each pollutant emitted from each facility. These summary tables will enable us to target facilities for detailed evaluation.
- 2002 Facility Data Source Summary - provides summary emissions data for each facility/pollutant/source category combination where more than one data source (state or local agency, TRI, ESD, etc.) was available. The report indicates which source was selected for inclusion in the draft. It will allow us to quickly compare data from the different sources.

We will prepare a pollutant/source category crosswalk table to QC pollutants emitted from source categories. For CAPs, we will identify source categories by SCC and SIC code. For HAPs, we will identify source categories by MACT Code, SIC Code, and SCC. This table is not available at this time. We will expand the crosswalk table to include emissions ranges.

After we complete emissions QC, we will identify potential errors and contact the data provider. If the data provider indicates that emissions should be revised, then we will revise estimates prior to the distribution of the draft. If the data provider does not respond, then we will note potential errors in the draft data distribution files.

12.8 Period Type

The 2002 NEI will contain annual emission estimates for HAPs and CAPs and non-annual emissions if reported for CAPs. If non-annual emissions are reported and annual emissions are not reported, we will generate an annual emission estimate.

We will add a data field, Period Type, to the Emissions Table to distinguish between annual and non-annual emission estimates in the 2002 NEI. The flag will indicate cases where non-annual emissions will be used to generate annual emission estimates.

We will use the Emission Type, Start Date, and End Date to determine the Period Type. In determining if an emission is annual or non-annual, we will use the following procedure.

- If there is only **one emissions record** for a given pollutant for a unique emission release point (tribal code, state and county FIPS code, site id, unit id, process id, emissions release point id), we will base the Period Type on the Emission Type and period of emissions.
  - If the Emissions Type is 30 and the Start Date is 01012002 and End Date is 12312002, the Period Type is annual.
  - If the Emission Type is 30 and emissions are reported for a period of a season or less, then the Period Type is assumed to be non-annual.
  - If the Emission Type is 30 and emissions are reported for a period greater than a season, then the Period Type is assumed to be annual.
  - If the Emission Type is 20 - 29, the Period Type is non-annual.
- When **multiple records with Emission Type 30 and other Emission Types** are reported for a unique emission release point, we will evaluate the Emission Types and Start and End Dates. If an Emission Type of 30 is reported and the Start and End Dates are 20020101 and 20021231, we will assume the period type is annual for this record and the period type of other reported records is non-annual. In most cases in which the data provider supplied a partial year (ozone season) in the 1999 NEI, they also supplied an annual record.
- When **multiple records with Emission Type 30 and multiple Start and End Dates** are reported for a unique emission release point, we will evaluate the Start and End Dates. If the Start and End Dates for a record are 20020101 and 20021231, we will identify the Period Type as annual for this record and non-annual for all other records. When emissions are provided for each season and no annual estimate is provided, the Period Type is non-annual for each of the records.

12.9 Annual Emissions

The 2002 NEI will contain annual emission estimates for HAPs and CAPs. We will retain the reported emissions and units in the NIF tables. We will add a field, Annual Emissions (tpy), to the Emissions Table.

Annual emissions will be converted to tons if reported in other units. Table 19 shows the conversions units used to generate emissions in tons per year.

Table 19. Unit Conversions for Tons

Reported Unit	Conversion Factors for Tons
E3TON	1000
MILLIGRAM	1.1025E-09
G	1.1025E-06
KG	0.0011025
LB	0.005
TON	1
UG	1.1025E-12
E2LB	0.05
NG	1.1025 E-15
TONNE	1.1025

We will use the methodology in Table 20 to calculate the Annual Emissions in tons/year. We will only use one methodology for each emission release point. We will give preference to complete years, type 30. In most cases in which a state supplied a partial year estimate (e.g., ozone season) in 1999, they also supplied an annual emission estimate.

Table 20. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Annual	30	Use reported emissions
Non-Annual Emissions only	30	All Individual Periods are present for a year; or If all individual periods are not present for a year and the data provider indicates Non-Annual Emissions for reported periods are the same as Annual Emissions. Annual Emissions = Sum of Numeric Emissions Value of each non-overlapping period Ex. Emissions are reported for each of the 12 months, Annual Emissions = Sum of monthly emissions

Table 20, continued. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Non-Annual Emissions only	30	Start Date and End Dates in Period are Not 20020101 and 20021231 Data provider indicates Non-Annual Emissions are not the Annual Emissions ▸ Start Time = 0000 and End Time = 2359 or Start and End Times are missing Annual emissions = Emissions Numeric Value * (365/ Number of Days in Period) ▸ Start and End Times are for less than a 24 hr period Annual Emissions = Numeric Emissions * (365/Number of Days in Period) * (Hrs/Day)
Non-Annual Emissions only	29	Data provider indicates Non-Annual Emissions are the same as Annual Emissions. ▸ Start Time = 0000 and End Time = 2359 or Start and End Times are missing Annual Emissions = Numeric Emissions * (Number of Days/Period) ▸ Start and End Times are for less than a 24 hr period Annual Emissions = Numeric Emissions * Number of Days/Period) * (Hrs/Day)

Table 20, continued. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Non-Annual Emissions only	29	Start Date and End Dates in Period are Not 20020101 and 20021231 Data provider indicates Non-Annual Emissions are not the Annual Emissions. ▸ Start Time = 0000 and End Time = 2359 or Start and End Times are missing Annual emissions = Emissions Numeric Value * (365/ Number of Days in Period) ▸ Start and End Times are for less than a 24 hr period Annual Emissions = Numeric Emissions * (365/Number of Days in Period) * (Hrs/Day) Ex. Emissions are for 20020501 - 20020831 for 24 hr/day Annual Emissions == Emissions (365/123)

12.10 Non-Annual Emissions

We will not retain non-annual emissions in the final NEI from data submitters unless seasonal throughput percentages are included in NIF submission.

We will QA emissions for emission release point records that contain both annual and non-annual emission estimates to identify cases where non-annual emissions exceed annual emission values. Because emissions may be reported and/or calculated for different time periods, e.g., annual, daily, seasonal, etc., and for the same or different pollutants, verification is needed to ensure that non-annual emissions are not higher than annual emissions. We will delete non-annual emission records where non-annual emissions exceed annual emissions.

EPA will not augment for invalid or missing non-annual emissions in the NEI.

12.11 Emissions Data Quality Rating

We will add an emissions data rating to every emission record in the point source NIF Emissions Table. The data rating will help to provide users of the NEI some sense of the reliability of an emission estimate. We are developing a simplified rating scheme so that a score can be assigned to each estimate (e.g., on a scale of 1-5). This enhances the transparency of the data and also satisfies the requirement that EPA perform an assessment of the data. The scoring will also help EFIG determine which data point to retain in the 2002 NEI when there are multiple estimates for the same facility. It may also help EFIG understand which source categories need improvement.

This rating scheme will not be in-depth as other systems such as DARS (Data Attribute Rating System), but will consider the following factors in assigning a score.

1. *Completeness of data* - Has the submitter provided enough information to enable the reviewer to repeat the calculation, assess emission factors and/or calculation methods? The fields that provide this information are:
   - Actual Throughput,
   - Throughput Unit Numerator,
   - Emissions Reliability Indicator,
   - Emission Calculation Method,
   - Factor Numeric Value,
   - Factor Unit Numerator,
   - Factor Unit Denominator, and
   - Emission Factor Reliability Indicator.

2. *Emission Calculation Method* - Estimates based on continuous monitoring should receive higher scores than data based on less accurate methods, e.g., "engineering judgment."

3. *Age of data* - In some cases, the NEI will have an emissions estimate from an earlier year, not the current inventory year. Having "old" data is preferred to having a data gap. However, a 1999 estimate should have a lower rank than a 2002 estimate.

4. *Qualitative Information* - Additional information is often provided by data sources that is not reflected in the database (e.g., municipal waste combustor estimates from EPA are based on source testing). We are also considering the breadth of data, i.e, did the source of this estimate provide a large number of HAPs relative to other sources for this category? For example, does EPA refinery data have more HAPs per facility than data provided by other sources (state and local agencies, TRI and industry)?

5. *Specificity of Data* - An estimate which provides process level emissions is better than aggregated facility level HAP emissions. For PM, state provided data for PM2.5 PRI and PM10 PRI would be assigned a hgiher rating than EPA-augmented data.

12.12 Data Source

We will add a data field to the Emissions Table to indicate the source of the reported emission value. The following codes will be used.

S	State agency
L	Local agency
T	Tribe
CAMD	EGU estimates based on 2002 ETS/CEM SO2, NOX and heat input values (utility related data) provided by EPA's Clean Air Markets Division data
767	EGU estimates based on 2002 Form EIA-767 information (utility related data)
CAMD/767	EGU estimates based on 2002 Form EIA-767 information, updated with 2002 ETS/CEM SO2, NOX and heat input values (utility related data) provided from EGU data based on EPA's Clean Air Markets Division data
CIBO	Council of Industrial Boiler Owners data compiled by EPA's Clean Air Markets Division
PM	EFIG generated data for PM augmentation PM2.5, PM10, Filterables and Condensibles
VOC HAP	EFIG generated data for VOC from HAPs
PMHAP	EFIG generated data for PM from HAPs
TRI	TRI
I	Industry
ESD	MACT or Residual Risk Data provided by ESD
99NEI	99 NEI
02PRE	2002 Preliminary CAP
CAP	Other CAPs generated by EPA from reported SCC and Throughput and emission factors from AP-42 and FIRE

12.13 Maintaining an Emissions History Table

We will create and maintain an Emissions History Table which tracks the changes in total emissions for each facility/site/pollutant combination in iterative versions of the 2002 NEI. This table will enable users to see past and current total facility emissions for each pollutant. The Emissions History Table will contain the following elements.

- Status - Indicates if a facility has closed or was added after the initial version
- State and County FIPS Code
- Tribal Code
- NEI Facility ID
- State Facility ID
- NEI Standardized Facility Name
- Pollutant Code
- Pollutant Name
- Emissions Value 1 - Emissions in TPY associated with version 1
- Data Source 1- Source of emissions value in version 1

- Ver_Date 1 - Version number concatenated with date of release
- Emissions Value 2 - Emissions in TPY associated with version 2
- Data Source 2 - Source of emissions value in version 2
- Ver_Date 2 - Version number concatenated with date of release
- .....

This table will be aggregated to the site/pollutant code level and will provide an emission estimate for all processes associated with each site. As new versions of the NEI are created, a new column will be added to this table providing the revised estimates and the data source for the newest version.

If a site closed (and is deleted from the inventory), the historical record will be retained, and the status column will be updated to "D." If a site is added that did not exist in previous versions, the status will be updated to "A" and the columns related to the most recent version (and every version, thereafter) will be filled in. Maintaining a derived table like this can quickly become unwieldy if too many data elements are added or the level of disaggregation is too detailed. This table is not meant to provide all the information that the NEI contains, but to provide a quick overview of the changes in emissions values over the course of the revision process. It is meant to help the user target facilities for further investigation, not provide all of the supporting information.

The Emissions History Table will be supplemented by the "holding" tables which contain the specific and individual changes submitted during each review period. These holding tables mirror the NIF tables and contain complete NIF records. These records have a "submittal flag" which indicates the change type ("A", "RA" or "D") and contain a source field which describes the submitter of the change. We will improve this table by adding a comment field. This field would be populated in those cases in which a user has provided a specific reason for a change.

In addition to the "holding" tables, when we use the "merge" algorithm to choose one value from among multiple data sources for the same pollutant/facility, we save the "rejected values" to subsidiary tables along with the list of data source choices. We will experiment with merging this information with the Emissions History Table. If supplementing the Emissions History Table with the "rejected values" is not efficient, we will retain them as separate stand-alone tables.

SECTION 13. OTHER FIELDS THAT UNDERGO QA BUT ARE NOT AUGMENTED

There are data fields that we will QA reported values and codes. If the codes are invalid, we will delete the values and not augment the values.

These following data fields will undergo QA but will not be augmented.

- EF Reliability Indicator
- Emission Calculation Method Code
- Material
- Material I/O
- Rule Effectiveness Method
- Telephone Number Type Name
- Transaction Creation Date

SECTION 14. DEFAULT FLAGS

Default Flags will be added to all records in each table to indicate values of fields that have been deleted, revised or augmented. Specific default flags that were discussed in previous sections will be added for the following fields.

- Coordinate Default Methodology (Latitude/Longitude)
- NAICS/SIC Code Default Flag
- MACT Code Assignment
- Stack Parameter Default Flag
- Total Capture Control Efficiency Methodology Flag
- Data Source

For all fields of data that are defaulted, a Default Flag will be added to each NIF Table to indicate the fields of data that are defaulted. Each data field in a table will be assigned a number from 01 to xx, and we will indicate if a field is defaulted by reporting its number. For example, the following fields are in the Control Equipment Table:

01 Record Type
02 State and County FIPS Code
03 State Facility Identifier
04 Emission Unit ID
05 Process ID
06 Pollutant Code
07 Primary PCT Control Efficiency
08 PCT Capture Efficiency
09 Total Capture Control Efficiency
10 Primary Control Device Type Code
11 Secondary Control Device Type Code
12 Control System Description
13 Third Control Device Type Code
14 Fourth Control Device Type Code
15 Submittal
16 Tribal Code

If we defaulted the Emission Unit ID, Pollutant Code, and Primary Control Device Code, then we would populate the Default Flag as “04,06,10".