



NEI Quality Assurance and Data Augmentation for Point Sources

EPA-454/R-20-001
February 2006

NEI Quality Assurance and Data Augmentation for Point Sources

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

TABLE OF CONTENTS

<u>Sections</u>		<u>Page</u>
1.	Introduction	1
2.	Inventory Information	16
2.1	Record Type	16
2.2	Submittal Flag	16
2.3	Transaction Type	16
2.4	Inventory Year	17
2.5	Inventory Type Code	17
2.6	Source Type	17
2.7	Affiliation Type	17
2.8	Format Version	17
2.9	Revision Date	17
3.	Geographic Information	18
3.1	State and County FIPS Code	18
3.2	Tribal Code	19
3.3	Urban/Rural Classification	20
3.4	MSA Name and Code	21
3.5	Country	21
4.	Facility Information	22
4.1	NEI Unique Facility ID	23
4.2	NEI Standardized Facility Name	26
4.3	NEI Standardized Address	26
4.4	NEI Standardized City	26
4.5	NEI Standardized State	26
4.6	NEI Standardized Zip Code	26
4.7	Location Address	26
4.8	City	27
4.9	State	27
4.10	Zip Code	27
4.11	ORIS Facility Code	27
4.12	Facility Registry ID	28
4.13	TRI ID	28
4.14	Dun and Bradstreet Number	28
4.15	BART Source and BART Unit	28
4.16	IPM Facility and IPM Unit	29

5.	Geocoordinates	30	
5.1	Geocoordinates QA and Default Procedures	30	
5.2	X Coordinate (longitude)	34	
5.3	Y Coordinate (latitude)	34	
5.4	UTM Zone	34	
5.5	XY Coordinate Type	34	
5.6	Location Default Methodology	35	
5.7	Latitude/Longitude EPA Data Standards	35	
	Stack Latitude	35	
	Stack Longitude	36	
	Horizontal Collection Method Code	37	
	Horizontal Reference Datum Code	37	
	Coordinate Data Source Code	37	
	Reference Point Code	37	
	Horizontal Accuracy Measure	37	
	Source Map Scale Number	37	
5.8	NEI Site Latitude	39	
5.9	NEI Site Longitude	39	
5.10	Latitude/Longitude QC Report	39	
6	Unit and Process IDs	42	
6.1	Emission Data Level	42	
6.2	Emission Unit ID	42	
6.3	Process ID	43	
6.4	Emission Release Point ID	43	
7	Source Description	44	
7.1	NAICS Code	44	
7.2	SIC Code	45	
7.3	NAICS/SIC Code Default Methodology	46	
7.4	SCC	47	
7.5	Process MACT and Section 112(k) Area Source Standards Code	47	
7.6	MACT and Section 112(k) Area Source Standards Code Assignment	50	
7.7	MACT and Section 112(k) Area Source Standards Compliance Code	51	
7.8	Facility Category	52	
8	Emission Release Point Description and Parameters	53	
8.1	Emission Release Point Type	53	53
8.2	Stack Parameters QA and Default Procedure	53	
8.3	Stack Parameter Default Flag	58	
8.4	Stack Parameters QC Report	58	

9	Unit Codes	61
9.1	Unit Codes	61
9.2	Design Capacity, Design Capacity Numerator and Denominator	61
10.	Activity Data Fields	63
10.1	Emission Type	64
10.2	Start and End Dates	64
10.3	Start and End Times	64
10.4	Year of Emission Estimate	64
10.5	Annual Average Days Per Week	65
10.6	Annual Average Hours Per Day	65
10.7	Annual Average Hours Per Year	65
10.8	Annual Average Weeks Per Year	65
10.9	Period Days Per Week	65
10.10	Period Hours Per Day	65
10.11	Period Hours Per Period	65
10.12	Period Weeks Per Period	66
10.13	Winter, Spring, Summer, and Fall Throughput Percentages	66
11.	Control Equipment	67
11.1	Control Status	67
11.2	Primary Device Type Code, Secondary Device Type Code, Third Device Type Code, and Fourth Device Type Code	67
11.3	Total Capture Control Efficiency	68
11.4	Total Capture Control Efficiency Methodology Flag	69
12.	Emissions	70
12.1	Pollutants	70
	NEI Pollutant Code Table	70
	CAPHAP Data Field	72
	Chemical Data Standards Lookup Table in the NEI	72
	HAPs Reporting	73
	Preferred Reporting of HAPs	74
	QA of HAP Pollutant Codes	78
	CAPs Reporting	79
12.2	PM Augmentation	80
	Steps in PM Augmentation	81
	Creation of PM Ratios	87
	References	93
12.3	VOC and PM Augmentation from HAPs	93
12.4	CAP Augmentation for Missing CAPs	94
12.5	HAP Augmentation from VOC and PM	94
12.6	HAP Performance Level	94
12.7	QC of HAP Category Group Emissions	95

12.8	Emissions Type Period	100
12.9	Annual Emissions	101
12.10	Non-Annual Emissions	104
12.11	Emission Ton Value	105
12.12	Emission QC	105
	Emission Content QC Report	107
	Other Summary Files to Assist in Emission Content QC	108
	Toxicity Weighting of HAPs	111
12.13	Emission Data Quality Rating for HAPs	113
12.14	Data Source	116
12.15	Maintaining an Emissions History Table	118
13.	Other Fields that undergo QA but are not augmented	120
14.	Default Flags	121

List of Tables

Tables	Page
1. NIF V3.0 Point Source Table Description	3
2. Fields in NIF V3.0 Point Source Tables	4
3. Additional Point Source Data Fields in the NOF that are not in the NIF V3.0	14
4. Additional Point Source Data Fields in NEI Auxiliary Files that are not in NIF V3.0 or NOF	15
5. MSAs Assigned to Counties in New England States Associated with More Than One MSA for Use In The County U1/U2/R Classification Process	21
6. Location Default Flags	35
7. EPA Latitude/Longitude Data Standards.	36
8. Geocoder Default Flags and Default Values for Latitude/Longitude Data Standard Elements	38
9. HAP Emission Data Levels	42
10. NAICS/SIC Code Default Flags	46
11. MACT Code Facility List Flags	49
12. Stack Parameter Data Replacement Matrix	56
13. Fields in NIF Requiring Unit Codes	61
14. Total Capture Control Efficiency Methodology	68
15. NEI Pollutant Code Table	71
16. NEI Chemical Data Standards Table	73
17. POM Compounds	76
18. List of Compounds Often Mistaken as Glycol Ethers	78
19. PM Compounds Reported in the 2002 NEI	79
20. PM Cases and Required Steps to Augment PM	84
21. PM Augmentation: SCCs Starting with “402” and the Surrogate SCCs	90
22. PM Augmentation: SCCs Starting With “5” That Have Combustion Sources (Incinerators and Flares) That Can Use Average Emission Factors from External Combustion Sources and Stationary Internal Combustion Sources	91
23. Methodology Used to Generate Annual Emissions	101
24. Unit Conversions for Tons	105
25. Table Structure for ALL NEI	109

SECTION 1. INTRODUCTION

The Emission Inventory and Analysis Group (EIAG) 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. The Clean Air Act (CAA) includes many mandates for the EPA related to CAPs and HAPs. The NEI is a tool that EPA can use to meet the CAA mandates. To support many of these functions, the NEI must contain the data necessary for air quality modeling.

The NEI is a critical component of the EPA's National Air Toxics Program. The CAA presents a list of 188 HAPs for which EPA is to identify their sources, quantify their emissions by source category, develop regulations for each source category, and assess public health and environmental impacts after the regulations are put into effect. The NEI for HAPs is compiled in order to determine if CAA programs are successful in reducing emissions and human health and environmental risk due to HAPs emissions. The initial objective is to make the NEI HAP data available to EPA modelers for use in the National Air Toxics Assessment (NATA) and for use in residual risk assessments conducted by EPA.

The NEI for CAPs are used in State Implementation Plans (SIPs), compliance demonstrations, emissions trading, and in modeling activities designed to evaluate ambient air concentrations. Title I, Section 110 of the CAA requires states to submit emission inventories for CAPs as part of their SIPs. EPA collects NEI CAP data under the Consolidated Emissions Reporting Rule (CERR) (40 CFR Part 51). The CERR specifies two sets of reporting thresholds for CAPs. Type A (large sources) must report annually, while Type B sources must report every three years. For the 2002 NEI, EPA collected information on both Type A and Type B sources. (For more information on CERR, see: [/www.epa.gov/ttnchie1/cerr/index.html](http://www.epa.gov/ttnchie1/cerr/index.html))

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 (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. Nonpoint 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)/Quality Control (QC) 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. In response to guidelines issued by OMB under Section 515(a) of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (P.L. 106-554; H.R. 5658), EPA issued the *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency*. This report focuses on the QA and augmentation of data in the NEI point source inventory.

We conduct a variety of QA activities to identify point 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 the methodology in this report. We summarize the errors found and provide reports back to the data providers on the QA findings. We track 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. The feedback reports include the following 4 QC reports.

Data QC Report - summarizes structural problems with the data, e.g, data errors, incorrect codes, data integrity issues, etc.

Latitude/Longitude QC Report - lists latitude/longitudes that failed QC review and provides EPA defaulted coordinates.

Stack Parameters QC Report - lists stack parameters that failed QC review and provides EPA defaulted parameters.

Emission Content QC Report - compares 2002 emissions data to 1999 emissions data, and highlights potential outliers, missing facilities, and missing and erroneous pollutants.

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.

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 were submitted to EPA in NIF V3.0 or NEI XMLV3.0 for incorporation into the 2002 NEI. Table 1 describes the eight tables in the NIF for point sources.

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 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 for 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 included in the NEI Output Format (NOF) Tables. All NIF data fields are included in the NOF. Table 4 shows additional data fields in the NEI that are included in auxiliary files. Data for all fields in Tables 3 and 4 will be generated by EIAG.

The QA and Augmentation Procedure for fields of data in Tables 2 - 4 are presented in 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

This document revises the previous report, “NEI Quality Assurance and Data Augmentation for Point Sources”, February 1, 2005. Any comments or questions concerning this report should be sent to Anne Pope, pope.anne@epa.gov.

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	10.5
EP	Annual Average Hours per Day	Average hours per day process is operating	Augmentation Procedure	10.6
EP	Annual Average Hours per Year	Average hours per year process is operating	Augmentation Procedure	10.7

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

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
EP	Annual Average Weeks per Year	Average weeks per yr process is operating	Augmentation Procedure	10.8
EP	Ash Content	Ash content of fuel		
SI	City	City of physical location	Augmentation Procedure	4.8
TR	Contact Person Name	EPA's single point 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
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

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

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
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		
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	7.8
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
EP	Fall Throughput Percent	Percentage of process that operates during the fall months (Sept, Oct, Nov)	Augmentation Procedure	10.13
TR	Format Version	NIF version	Augmentation Procedure	2.8
CE	Fourth Device Type Code	Fourth 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
ER	Fugitive Dimensions Unit	Unit of measure NIF Code Table - Unit codes	Augmentation Procedure	9.1
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	Horizontal area of fugitive emissions		
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		
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

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

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
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		
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	10.9
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	10.10
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	10.11
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	10.12
CE, EM	Pollutant Code	Pollutant Code NIF Code Table	Augmentation Procedure	12.1

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

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
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
EP	Process MACT Code	Maximum Achievable Control Technology source category code NIF Code Tables	Augmentation Procedure	7.5
EP	Process MACT Compliance Code	Major/Area classification under Sections 112 and 129 NIF Code Table	Augmentation Procedure	7.7
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
ER	Release Height Fugitive	Release height above terrain of fugitive emissions		
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	7.4
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	10.13
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
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

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

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
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	10.13
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		
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	10.13
ER	X Coordinate	Longitude (- decimal degrees)	Augmentation Procedure	5.2

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

NIF Table Abbreviation	Field Name	Field Description	Augmentation Procedure	Report Section
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 NOF that are not in the NIF V3.0

Field Name	Field Description	NIF Table Relationship	Report Section
Augmented Data	Default flags added to indicate defaulted fields of data	all	14
BART Source	Does facility contain BART units?	SI	4.15
BART Unit	Is unit a BART unit?	EU	4,15
CAP HAP	Is pollutant CAP or HAP?	EM	12.1
Data Rating	Qualitative rating for emission estimate	EM	12.13
Data Source	Source of Data	EM	12.14
Emission Ton Value	Create emissions in tons	EM	12.11
Emissions Type Period	Are emissions annual or non-annual?	EM	12.8
IPM Facility	Does facility contain IPM units?	SI	4.16
IPM Unit	Is unit an IPM unit?	EU	4.16
Location Default Flag	How latitude/longitude coordinates are defaulted	ER	5.6
MACT Code Flag	Method used to assign MACT Codes	ER	7.6
NAICS Code Default Flag	How NAICS and SIC Codes are defaulted	SI, EU	7.3
Revision Date	Latest date of record - will be creation or revision date	all	2.9
SIC Code Default Flag	How NAICS and SIC Codes are defaulted	SI, EU	7.3
Stack Parameter Default Flag	Indicates which stack parameters have been defaulted	ER	8.3
Total Capture Control Efficiency Default Flag	Indicates how total capture control efficiencies have been defaulted	CE	11.4
Year	Year of emission estimate	EM	10.4

Table 4. Additional Point Source Data Fields in NEI Auxiliary Files that are not in NIF V3.0 or NOF

Field Name	Field Description	NIF Table Relationship	Report Section
MSA Code	Metropolitan Statistical Area Code	SI	3.4
MSA Name	Metropolitan Statistical Area Name	SI	3.4
NEI Unique Facility ID	NEI Unique 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
Stack Latitude	Convert Coordinates to Latitude if in UTM	ER	5.7
Stack Longitude	Convert Coordinates to Longitude if in UTM	ER	5.7
Urban/Rural Classification	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
- Revision 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 NIF code 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.

<u>NIF Code</u>	<u>Pollutants</u>
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 Revision 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 20050201. If the estimate is not revised, the date would remain 20050201. If the estimate was revised during the external review period, the date would be revised to 20050501 in the final inventory. As revisions are made in time, then the revision date will 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 Classification
- MSA Name and Code
- Country

We will prepare a State/County FIPS Code Table with the following data fields.

- State/County FIPS Code
- State FIPS
- County FIPS
- State/County FIPS Code Description
- Start Date - date FIPS code became valid
- End Date - date FIPS code became invalid for use
- Valid for NIF 2.0
- Valid for NIF 3.0
- Comments - Notes about changes
- EPA Region
- County Name
- State Abbreviation
- Urban/Rural Classification
- 2002 Population (number of people) - estimates from US Census Bureau
- MSA Name - from U.S. Census Bureau
- MSA Code - from U.S. Census Bureau
- County Centroid Latitude
- County Centroid Longitude
- County Maximum Latitude
- County Minimum Latitude
- County Maximum Longitude
- County Minimum Longitude
- State Maximum Latitude
- State Minimum Latitude
- State Maximum Longitude
- State Minimum Longitude

This table is found in the 2002 State Lookup File, 02nei_lkup_states.zip, found at:
<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

3.1 State and County FIPS Codes

State and County Federal Information Processing Standards (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 2002 NEI 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 2002 NEI Facility Table.

For facilities missing State and County FIPS Codes that are not in the 2002 NEI 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 2002 NEI Facility Table and see if the facility is present in the file. If it is present, we will default the Tribal Code using the 2002 NEI Facility Table. For facilities that are not in the 2002 NEI Facility Table, 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 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 non-tribal data to "000".

3.3 Urban/Rural Classification

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 of 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 of 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 is assigned for use in the U1/U2/R classification process. Table 5 shows the MSAs that were assigned in the 1999 NEI for these areas.

Table 5. 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

3.4 MSA Name and Code

We will identify the MSA of counties in the 2002 NEI. We will include the MSA Name and MSA Code in the State/County FIPS Code table.

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".

SECTION 4. FACILITY INFORMATION

This section includes the following data fields.

- NEI Unique 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
- BART Source and Unit
- IPM Facility and Unit

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

- State/County FIPS Code
- Tribal Code
- State FIPs Abbreviation - State abbreviation for State/County FIPS Code
- County FIPS Name - County Name for State/County FIPS Code
- Tribal Name - Name of Tribe for Tribal Code
- State Facility ID - Local identifier provided in original state, local agency, tribal, industry, TRI, or MACT data submittal
- Facility Name (associated with State Facility ID)
- Location Address (associated with State Facility ID) - Physical location of the main entrance of the site
- City (associated with State Facility ID) - City of the physical location of the site
- State (associated with State Facility ID) - State abbreviation of the physical location of the site
- Zip Code (associated with State Facility ID) - U.S. Postal Service zip code of the physical location of the site
- NEI Unique Facility ID - Common identifier assigned to all sites determined to be the same facility
- NEI Standardized Facility Name - Preferred name when one or more names are available for the same facility
- NEI Standardized Address - Preferred zip code when one or more location addresses are available for the same facility

- NEI Standardized City - Preferred city when one or more cities are available for the same facility
- NEI Standardized State - Preferred state abbreviation when one or more states are available for the same facility
- NEI Standardized Zip Code - Preferred zip code when one or more zip codes are available for the same facility
- NEI Site Latitude - Average of all latitudes for the site
- NEI Site Longitude - Average of all longitudes for the site
- NEI Primary SIC Code
- NEI Primary NAICS Code
- ORIS Facility ID - assigned by EPA to electric generating units
- FRS State/County FIPS Code - Facility Registry Identifier state and county FIPS assigned to facility by OEI
- FRS ID - Facility Registry Identifier assigned to facility by OEI
- FRS Facility Name - Facility Registry Identifier facility name assigned to facility by OEI
- FRS Address - Facility Registry Identifier location address assigned to facility by OEI
- FRS City - Facility Registry Identifier city assigned to facility by OEI
- FRS State - Facility Registry Identifier state abbreviation assigned to facility by OEI
- FRS Zip Code - Facility Registry Identifier zip code assigned to facility by OEI
- FRS Latitude - Facility Registry Identifier latitude assigned to facility by OEI
- FRS Longitude - Facility Registry Identifier longitude assigned to facility by OEI
- TRI ID - assigned by EPA's Office of Environmental Information (OEI)
- Dun and Bradstreet Number - Dun and Bradstreet number for the facility
- Historical Names for Facility and Ownership
- Facility Status - indicates if facility is closed or inactive
- 1999 CAP - denotes if the site is in the 1999 NEI for CAPs
- 1999 HAP -denotes if the site is in the 1999 NEI for HAPs
- 2002 Preliminary NEI - denotes if the site is in the 2002 Preliminary NEI for CAPs
- 2002 Draft NEI - denotes if the site was in the 2002 NEI 02012005 Draft
- 2002 NEI CAP 02012006 - denotes if the site is in the current version (Version 1.0, 02012006) of the 2002 NEI for CAPs
- 2002 NEI HAP 02012006 - denotes if the site is in the current version (Version 1.0, 02012006) of the 2002 NEI for HAPs
- Comments - provides additional information about some facilities

The 2002 NEI Facility Table will evolve into a historical facility database where changes in names/ownerships and subsidiaries will be tracked over time. The 2002 Facility Table is located in a zipped file, nei2002_point_facility_ids_history_date, at:

ftp://ftp.epa.gov/EmisInventory/2002finalnei/point_sector_data/ .

4.1 NEI Unique Facility ID

The NEI Unique 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 Unique Facility ID is assigned to facility records submitted from different data sources in order to detect duplicate estimates among the records. The NEI Unique 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 and HAP emissions from VOC and PM emissions.

We will first determine if a facility is in the 1999 NEI Facility Table. If it is in the 1999 NEI Facility Table, then we will use the existing NEI Unique Facility ID. If the facility is not in the table, then we will assign a NEI Unique Facility ID and add all of the information for the facility to the 2002 NEI Facility Table. This information will be used to augment other data fields associated with facilities.

We 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 to generate a preliminary 2002 NEI Facility Table that was used as a starting point for the 2002 NEI. 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 will use this same approach for facilities that are submitted for inclusion in the 2002 NEI that are not in the 2002 Preliminary NEI Facility Table.

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

- matching facilities in the NEI and assigning a common NEI Unique Facility ID, and
- matching NEI facilities to EPA's Facility Registry System (FRS).

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 co-located 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 include SIC and NAICS 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#1](http://oaspub.epa.gov/edr/epastd$.startup#1)) The standards relevant to the NEI are the: SIC/NAICS, Latitude/Longitude, Chemical Identification, Facility Identification and Contact Information. To implement the Facility Identification standard, EIAG 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.

- EIAG submits NEI facility list to OEI
- OEI assigns draft FRS IDs to NEI facilities
- EIAG incorporates FRS IDs into NEI facility file

The process of assigning FRS IDs to the 2002 NEI will not occur until after 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 2002 NEI 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 an address changes for a facility. We will use the 2002 NEI 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 2002 NEI 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 Unique Facility ID. We will use the 2002 NEI 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 Unique 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 2002 NEI 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 2002 NEI 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 2002 NEI 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 2002 NEI Facility Table, we will use the NEI Standardized State to default this field. If the facility is not in the 2002 NEI 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:

ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/.

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. 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 2002 NEI 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 2002 NEI 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 2002 NEI Facility Table. If the ORIS Facility Code is missing for electric generating units, then we will use the 2002 Electric Generating Unit (EGU) ORIS Facility Code in the 2002 NEI Facility Table to default the missing ORIS Facility Codes. If the comparison indicates that the ORIS Facility Codes are different in the 2002 NEI 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 2002 NEI 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 2002 NEI 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 2002 NEI 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 2002 NEI Facility Table to augment this data field if the data are not available.

4.15 BART Source and BART Unit

The identification of Best Available Retrofit Technology (BART) emission units in the NEI is needed to support modeling for EPA's regional haze rule. The BART requirements apply to units within facilities built between 1962 and 1977 in 26 source categories that have the potential to emit 250 tons or more of visibility- impairing pollutants. (For more information on BART, go to: <http://www.epa.gov/air/visibility/factsheet.html> .)

State and local agencies, tribes, and Regional Planning Organizations (RPOs) have begun to identify BART units in their inventories. We will flag BART facilities and emission units within facilities in the 2002 NEI from available information provided by state and local agencies, tribes and RPOs.

4.16 IPM Facility and IPM Unit

EPA uses the Integrated Planning Model (IPM) to analyze the projected impact of environmental policies on the electric power sector. EPA uses the IPM to evaluate the cost and emissions impacts of proposed policies to limit emissions of sulfur dioxide (SO₂), nitrogen oxides (NO_X), carbon dioxide (CO₂), and mercury (Hg) from the electric power sector. (For more information on IPM, go to: <http://www.epa.gov/airmarkets/epa-ipm/>.)

The identification of IPM emission units in the NEI is needed to support IPM analyses. We will flag IPM facilities and emissions units within facilities in the 2002 NEI using information provided by EPA staff.

SECTION 5. GEOCOORDINATES

This section includes the following data fields:

- X coordinate (longitude)
- Y coordinate (latitude)
- UTM zone
- XY Coordinate Type
- Location 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).

Longitude should be reported in units of decimal degrees with a negative sign and 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.

$$\text{Latitude (decimal degrees)} = \text{Latitude degrees} + \text{Latitude minute}/60.0 + \text{Latitude second}/3600.0$$

$$\text{Longitude (decimal degrees)} = \text{Longitude degrees} + \text{Longitude minute}/60.0 + \text{Longitude second}/3600.0$$

If latitude and longitude are reversed, a common error in NEI submittals, we will transpose the latitude and longitude. If the reported value does not have a decimal, we will use the minimum and maximum latitude and longitude fields in the State/County FIPS Code Table to determine if

placing a decimal in the reported values will result in a valid latitude and longitude. If we cannot correct the reported geocoordinates, we will contact the data provider.

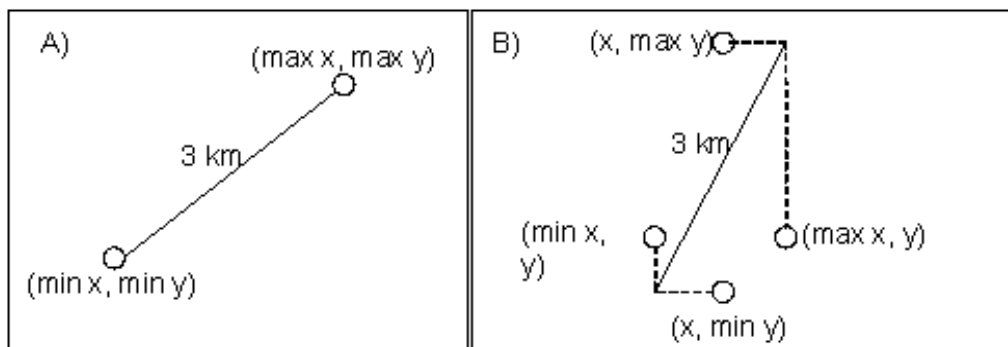
After converting UTM's to latitude and longitude and correcting latitudes and longitudes reported in wrong units, we will use a routine to assess the validity of the latitude and longitude values, to replace values if necessary, and to fill-in missing data points. The QA of location coordinates is a 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 will 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 will 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 will 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. 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 point(s).

- 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 feasible.
 - ▶ 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 list of SIC Codes and NAICS Codes that the distance between emission release points may be greater than 3.0 km, download the table “SIC NAICS 3km ERP_Acceptable” from:
ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/ .
 - ▶ 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 records are returned to the data provider for assistance in identifying the source of error. We will then 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 State/County FIPS Code 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 State/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 erroneous 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 Geocoding software* – If the reported emission release point latitudes/ longitudes failed any of the above steps and revised information from the data source is not provided,

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:

ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/

. 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 will be compared to 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 will not use the reported Zip Code in the Geocoding process. If global errors such as dropping leading zeroes for a state or transposition errors exist, then we will correct the Zip Codes to the extent possible. 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 where possible, we will compile a list of sites that have emission release points with missing or erroneous latitude/longitude. The file contains site name, the physical address of the sites, and state and county FIPS code. The file is submitted to the geocoding software. Using the inventory record’s street address, the geocoding software provides matches using the following hierarchy:

1. First the software standardizes the street address and looks for an exact address match. The latitude/longitude are located as an exact 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 that match the recorded FIPS, the process is considered to be complete and the geocoded latitude/longitude pairs are assumed to be valid. If the latitude/longitude data do not match the recorded FIPS, the geocoded coordinates are not used. The State and County FIPS code will be changed only after we contact the data provider and verify the coordinates and FIPS code. If geocoded latitude/longitude are used, the process is considered complete.

- B. *Use 2002 NEI Facility Table* - If Geocoder cannot find a valid latitude/longitude for the facility, then the 2002 NEI Facility Table will be used. If the identified state and county in 2002 NEI Facility Table for a facility does not match the reported state and county in the inventory, the 2002 NEI Facility Table data will not be used. If the 2002 NEI 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.
- C. *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 coordinates. County centroids are found in the State/County FIPS Code Table.

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 where needed based on the State and County FIPS code.

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 draft and final 2002 NEI.

5.6 Location Default Methodology

All defaulted latitudes and longitudes will be identified in the Emissions Release Point records in the NEI. We will use the coordinate default flags shown in Table 6.

Table 6. Location 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.
Zipcode+2	Match is to a 5-digit zip code, plus the first two digits of the 4-digit extension.
Zipcode3	Match is to multiple 3-digit zip codes based on postal service Sectional Center Facility (SCF).
Zipcode5	Match is to a 5-digit zip code.
CNTYCENT	County centroid is used.
NEIHFT	Coordinate is found in the 2002 NEI Facility Table.
SITEAVG	A Site Average of accurate coordinates of other emission release points at the same site is used.
State_Rev	The coordinates were revised by state or local agency or tribe.

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 7 summarizes these elements. This standard will help users gauge the accuracy and reliability of a given set of coordinates. (For more information on EPA data standards, go to: [http://oaspub.epa.gov/edr/epastd\\$.startup#1](http://oaspub.epa.gov/edr/epastd$.startup#1)) 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.

Table 7. 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 = 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.	Indicates the maximum distance on the ground from coordinates to true physical location, e.g., 3.
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.

Horizontal Collection Method Code

If the Horizontal Collection Method code is missing, we will populate this field code as “027” (i.e., 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 be able to populate these fields based on documentation from TeleAtlas Geocoding EZ Locator Service (www.geocode.com/). Table 8 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 8 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (www.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 8 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (www.geocode.com/).

Horizontal Accuracy Measure

We will only populate the Horizontal Accuracy Measure with values in Table 8 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (www.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 8 when latitude/longitudes are obtained from the TeleAtlas Geocoding EZ Locator Service (www.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.

Table 8. Geocoder Default Flags and Default Values for Latitude/Longitude Data Standard Elements

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*
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*
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*
CNTYCENT	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*
SITEAVG	Site average.		021 - Determination method based on interpolation-other.	001 - North American Datum of 1927	N/A	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). The Reference Point Code for all defaults is 108 (points not represented by codes 101-107).

5.8 NEI Site Latitude

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

5.10 Latitude/Longitude QC Report

We will prepare a QC report that contains a list of emission release point records that failed our QA of latitude and longitude and provides our defaulted coordinates. We will send these reports to the data providers and will ask data providers to review the defaults and provide corrected addresses or latitudes/longitudes. Replacing coordinates or correcting bad address information could improve the quality of the data needed for modeling..

The QC report will contain the following two files.

- *LATLON Report* will contain 3 sections. Section I will provide a count of the affected records by error type. Section II will summarize the corrections made by default method. Section III will provide definitions of some of the errors and default types.
- *LATLON QA database* will contain 3 tables. The table, "LatLonRevisions", will contain all coordinates that failed one or more of our tests and the default applied to correct the error. Field explanations for "LatLonRevisions" as well as full default definitions will be found in tables "tblFields" and "DefaultCodes," respectively.

Data providers will be able review the defaulted latitude/longitude and provide corrected addresses and/or coordinate pairs if desired. Revisions will be provided by data providers by filling in the fields starting with "REVISED" at the end of the LatLonRevisions table. We will ask data providers to focus their review on defaulted coordinates labeled as "CNTYCENT" as these are the least accurate coordinate type. We will also ask data submitters to look more closely at those sites with emissions release points that were greater than 3km apart (i.e., those with "Fail" in the "ERPComment" field) as we will not be able to identify which of the points are outliers.

If data providers do not provide changes, we will use the defaulted values in the 2002 NEI.

The LatLon Revisions Table will contain the following fields of data and revision columns for data providers to complete.

- State Name
- State FIPS

- State County FIPS
- State Facility ID
- Emission Release Point ID
- Original X Coordinate
- Original Y coordinate
- Facility Name
- Street Address
- City
- State
- Zip Code
- NullZero - identifies records submitted with null coordinates, 0 coordinates or UTM coordinates without a UTM zone
- Converted - indicates if the X and Y Coordinates were originally provided in UTM or as DDMSS and were converted to decimal degrees
- ERP Comment - result of emission release point test to indicate whether ERPs are within 3 km of each other. Results are pass/fail; although some SIC/NAICS codes indicate facilities that could legitimate, and these are marked "Pass due to SIC/NAICS".
- GISCOMatch - QA Test: Indicates if coordinates when plotted in a GIS fall into the county associated with the emissions release point. Results are pass/fail.
- Within 10 km - QA Test: Indicates where coordinates are when plotted in a GIS with respect to the county associated with the emissions release point. Value of 0 indicates that the point falls into its assigned county.
- GISstestFIN - QA Test Results: Indicates whether coordinates ultimately pass/fail GISstest given "GISCoMatch" and "Within10km."
- ZipCode Check - QA Test: Indicates if the zip code associated with the site matches the county associated with the site. Also indicates any corrections made to zipcode based on NEI Historical Facility Table.
- Bad Address - Identifies records with addresses that are not physical addresses or addresses that can be used for geocoding.
- GISFIPS - The county FIPS associated with X and Y coordinates when plotted.
- Dist2County - Distance from plotted XY coordinates to the county within which it is reported to lie.
- Default - Code indicating how the X and Y coordinates were defaulted.
- New X Coordinate - Defaulted X coordinate. We will use defaulted coordinate unless a revision is provided.
- New Y Coordinate - Defaulted Y coordinate. We will use defaulted coordinate unless a revision is provided.
- XY Type - Defaulted coordinate type. We will use defaulted coordinate type unless a revision is provided.
- Latitude/Longitude EPA Data Standard Elements - Includes Horizontal Collection Method, Horizontal Accuracy Measure, Horizontal Reference Datum, Reference Point Code, Source Map Scale, and Coordinate Data Source Code
- Original Data Provider - person who provided original coordinates

- Revised Data Fields: Reviewers may provide corrections for EPA to evaluate in these data fields.
 - ▶ Data provider of revised coordinates - required data field for all revisions
 - ▶ Revised Street Address
 - ▶ Revised City
 - ▶ Revised State
 - ▶ Revised Zip Code
 - ▶ Revised X Coordinate
 - ▶ Revised Y Coordinate
 - ▶ Revised UTM Zone
 - ▶ Revised XY Type
 - ▶ Revised Latitude/Longitude EPA Data Standard Elements - Horizontal Collection Method, Horizontal Accuracy Measure, Horizontal Reference Datum, Reference Point Code, Source Map Scale, and Coordinate Data Source Code

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 an 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 9.

Table 9. HAP Emission Data Levels

Emission Data Level	Mandatory Reported Fields	EIAG 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 the level that 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 code. 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 in the NEI 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 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 was 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 contained in the table, NAICS02toSIC87_NEI99, is found in the 2002 State Lookup File, 02nei_1kup_states.zip, at: <ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

We will first QA the reported NAICS codes by using the Census Bureau's list of NAICS codes (see www.census.gov/epcd/naics02/) A common error that we found in the 99 NEI was the reporting of the SIC Code as the NAICS Code. We will compare the SIC and NAICS codes and identify records where the SIC and NAICS codes are the same and determine if the SIC Code is valid.

We will default all Primary NAICS codes in the Site Table. 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 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, SIC code, and NAICS code.
3. For facilities that have invalid or missing SIC codes and are not in the Power Business database but are in the 2002 NEI Facility Table, we will use the 2002 NEI 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 SIC/NAICS crosswalk. We will not default missing NAICS Code Unit Level.

We will assign NAICS/SIC default flags shown in Table 10 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 and 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 2002 NEI.

We will first QA the reported SIC codes by using the Census Bureau's list of SIC codes (see 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 2002 NEI Facility Table, we will use the 2002 NEI Facility Table to augment the Primary SIC Code.

2. For new facilities not included in the 2002 NEI 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 2002 NEI 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, NAICS02toSIC87_NEI99, in the 2002 State File, 02nei_lkup_states.zip, is found in the 2002 State Lookup File, 02nei_lkup_states.zip, at: <ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .
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/NAICS crosswalk.

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

7.3 NAICS/SIC Code Default Methodology

Table 10 presents the NAICS and SIC Code Default Flags that will be used in the Site and Emission Unit Tables.

TABLE 10. 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.

Table 10 continued. NAICS/SIC Code Default Flags

Code	Match Type	Code Description
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

7.4 SCC

Source Classification Codes (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.

We will QA the reported SCCs by using the SCC Code Lookup Table posted at: <http://www.epa.gov/ttn/chief/code/index.html> . For missing or invalid SCCs, we will return records to the data provider for assistance in revising SCCs during the review of the draft 2002 NEI. If data providers do not revise invalid SCCs, we will use the Emission Process Description to default the SCCs. If we revise the SCCs, we will indicate that SCCs have been revised.

7.5 Process MACT and Section 112(k) Area Source Standards Code

The MACT and Section 112(k) Area Source Standards Codes identify processes that are or will be subject to MACT standards and new Section 112 (k) area source standards. MACT and Section 112(k) Area Source Standards codes (hereafter, referred to as MACT codes), flags and compliance codes are applied to processes in the NEI. 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/urbanpg.html .

We will prepare a MACT Code table with the following fields:

- MACT Source Category
- MACT Code
- proposed new code (Yes or No)
- 112k area source category (Yes or No)
- Group (EPA group responsible for category)
- 10 year MACT (Yes or No)
- Residual Risk (Yes or No)
- Contact (EPA staff lead)

- Contact Email
- Promulgation Date
- Implementation Date
- Major Source MACT Compliance Code
- Area Source MACT Compliance Code
- Comment

The MACT Code Table will be included in the 2002 State Lookup File, 02nei_lkup_states.zip, found at: <ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

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 are assigned at the process level in the NEI.

The first preference in assigning MACT codes will be given to the code assigned by the original data provider. Thus, when the state/local agency or tribe supplies a code, we will retain the MACT code and identify the data source as “STATE- BASED”. When we receive data from an EPA Emissions Standard Division (ESD) lead, we will assign the appropriate MACT Code and identify the data source “ESD-BASED”. We will identify EGU data generated by EPA as “EIG-BASED”. Please note that EPA’s Emissions Standards Division is now known as the Sector Policies and Programs Division.

If state and local agencies and tribes do not provide MACT codes in their 2002 submittals, we will default MACT codes by using Facility lists, or applying SCC, SIC code, and NAICS code defaults.

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

- Emissions data provided by EPA
- Emissions data provided by State or local agency or tribe
- EGU data provided by EPA
- Facility list provided by EPA’s ESD
- SCC/MACT Code dictionary
- SIC/MACT Code dictionary
- NAICS/MACT Code dictionary

The complete steps for assigning MACT Codes are described below.

1. If ESD-BASED file, apply MACT Code supplied by ESD and flag as ESD-BASED.
2. If State-BASED file, retain MACT Code supplied by the state or local agency or tribe and flag as STATE-BASED.
3. If EGU-based file, retain MACT or Area Source Code supplied by the EIAG and flag as EIG-BASED.

4. Apply Facility list codes to all processes at the facility. Do not overwrite ESD-, STATE-, or EIG-BASED codes. The facility lists were supplied by ESD and ESD's contractor, RTI. In some cases, the list was derived from ESD-BASED MACT data in the 1999 NEI. Codes assigned by these lists are flagged as "FACILLIST*." Table 11 describes the specific source of each facility list and its corresponding flag.

Flags beginning with "FACILLIST-MLT" indicate that the facility appeared on one or more (i.e., multiple) category lists. The look-up table, "Multiple MACT Codes", includes a list of all MACT Codes assigned to a Multi-MACT Category. The table, "Multipl-MACT Source of Facility Lists", provides a key for the MACT flags beginning with "FACILLIST-MLT" The two tables will be included in the 2002 State Lookup File, 02nei_lkup_states.zip, found at:
<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

TABLE 11. MACT Code Facility List Flags

Facility List Flag	Description
FACILLIST-99NEI	Facility list derived from 1999 NEI
FACILLIST-ESD02	Facility supplied by ESD in 2002
FACILLIST-RTI	Facility list supplied by RTI
FACILLIST-MLTx	Facility appears on multiple lists

5. Apply SCC-Default list to assign MACT and Area Source Flags to non-coded processes. If SCC-default list assigns a different MACT or Area source category than a Facility list, then let that code overwrite the Facility list supplied code. Use "SCC-DEFAULT" as the flag.
6. Apply SIC-Default list to assign MACT Flags to all processes at a site. However, if any one process already has a MACT code, do not assign ANY SIC-based codes to that site. If SIC-default list assigns a different MACT category than a Facility list, then let that code overwrite the Facility list supplied code. Use "SIC-DEFAULT" as the flag.
7. Apply NAICS-Default list to assign MACT and Area Source Flags to all processes at a site. However, if any one process already has a MACT code, do not assign ANY NAICS-based codes to that site. If NAICS-default list assigns a different MACT category than a Facility list or SIC-default, then let that code overwrite the previous codes. Use "NAICS-DEFAULT" as the flag.

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, NAICS codes, and SCCs will not be used to default the MACT code. If SCC, SIC code, or NAICS code defaults are used, we will either use SCC, SIC code, or NAICS code defaults for a process within a facility and will not mix SCC, SIC code, and NAICS code defaults at a single facility. Thus once we default a process within a facility based on SCC, we will not default MACT codes for other records missing SCCs that have a SIC or NAICS code. For example, the three 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/NAICS Code/MACT Code default dictionary used in the 2002 NEI was developed by comparing all of the SCCs, SIC Codes, and NAICS Codes with information on types of sources that may be subject to each MACT and Section 112(k) Area Source standard. EPA staff will review the draft 2002 NEI to revise the dictionary and to revise the facilities associated with individual categories. For 2002 NEI, we revised the 1999 SCC/SIC Code/MACT Code dictionary to include the new area sources categories. SCC/MACT Code, SIC Code/MACT Code and NAICS Code/MACT Code defaults used in the 2002 NEI are found in the tables, “SCC to MACT Defaults”, “SIC to MACT Defaults”, and “NAICS to MACT Defaults”, in the 2002 State Lookup File, 02nei_lkup_states.zip, found at: ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/ . Separate SCC MACT default tables are available for the Hazardous Organic NESHAP and petroleum Refinery MACT Categories in the Lookup file.

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

7.6 MACT and Section 112(k) Area Source Standards Code Assignment

The basis of MACT and Section 112(k) Area Source Standards Code (hereafter, referred to as MACT Code) assignment in the 2002 NEI will be indicated in the data field, MACT Flag Assignment, using the following values.

- EFIG-Based
- ESD-Based
- FACILLIST-99NEI
- FACILLIST-ESD002
- FACILLIST-MLT50
- FACILLIST-MLT51
- FACILLIST-MLT52
- FACILLIST-MLT53
- FACILLIST-RTI
- NAICS Default

SCC Default
SIC Default
State-Based

7.7 MACT and Section 112(k) Area Source Standards Compliance Code

The MACT and Section 112(k) Area Source Standards Compliance Code (hereafter, referred to as MACT Compliance Code) will be used to identify whether a source is subject as a major or an area source of HAPs 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 or Section 112(k) Area Source Standard source category major and area sources of HAPs that are subject to standards. For projections, major and area sources of HAPs within a MACT or Section 112(k) Area Source Standards category can have different percent reductions.

Acceptable NIF MACT Compliance 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.

The MACT Compliance Code will be assigned as follows.

1. If the process is associated with an Section 112(k) Area Source category code and this category is only a Section 112(k) Area Source category, assign compliance code "03".
2. If a process is associated with a category code that could either be a MACT or Section 112(k) Area Source category code and the facility is a major source, assign the compliance code listed under "Major Source compliance code" in the MACT lookup table.
3. If a process is associated with a category code that could either be a MACT or Section 112(k) Area Source category and the facility is an area source, assign compliance code "03".
4. If a process is associated with a category code that is only a MACT Code and the facility is a major source, assign the compliance code listed under "Major Source MACT Compliance Code" in the MACT Code lookup table.
5. If a process is associated with a category code that is only a MACT Code and the

facility is a area source, assign the compliance code listed under “Area Source MACT Compliance Code” in the MACT Code lookup table.

7.8 Facility Category

The Facility Category indicates whether a facility is a major or an area source of HAP emissions. 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 the 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 HAPs 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 HAP source, and
- 02 - Area HAP source

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 will 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: ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/. 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: For fugitive emission release points, replace stack parameters

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 °F	295.37 °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: For non-fugitive emission release points, find out-of-range or missing stack parameters

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 1250 ft	0.03048 to 381.00 m
Stack Temperature*	50 to 1,800 °F	283.15 to 982.37 °K
Stack Diameter	0.1 to 50 ft	0.03048 to 15.24 m
Stack Velocity	0.1 to 150 ft/sec	0.03048 to 45.72 m/sec

*Stack Temperature should be greater than 100 °F for the following SCCs and MACT Codes if controls are not present.

- 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.

$$\text{Stack Flow [cu ft/sec]} = (\Pi [\text{Pi}] * (\text{Stack Diameter [ft]} / 2) ^ 2) * \text{Stack Velocity [ft/sec]}$$

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 then evaluate the flow rate and velocity for internal consistency.

If the internal consistency is not met for velocity, flow rate, and diameter, Table 12 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 12. Stack Parameter Data Replacement Matrix (X = Data value present)

Diameter	Velocity	Flow Rate	Action
X	X	X	<ol style="list-style-type: none"> 1. Check that velocity is within range. <ol style="list-style-type: none"> A. If velocity is within range and flow rate does not meet internal consistency for diameter, velocity and flow rate, then: <ul style="list-style-type: none"> ▶ Calculate flow rate using internal consistency formula. B. If velocity is not within range, then: <ul style="list-style-type: none"> ▶ 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	<ol style="list-style-type: none"> 1. Calculate velocity using internal consistency formula. 2. Check that calculated velocity is within range. <ol style="list-style-type: none"> A. If calculated velocity is not within range, then: <ul style="list-style-type: none"> ▶ Default all 5 parameters using national default sets.
X	X	-	<ol style="list-style-type: none"> 1. Check that velocity is within range. <ol style="list-style-type: none"> A. If velocity is within range, then: <ul style="list-style-type: none"> ▶ Calculate flow rate using internal consistency formula. B. If velocity is not within range, then: <ul style="list-style-type: none"> ▶ Default all 5 parameters using national default sets.
X	-	-	<ol style="list-style-type: none"> 1. Default velocity using national default sets. 2. Calculate flow rate using internal consistency formula.
-	X	X	<ol style="list-style-type: none"> 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 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

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 °F with SCC 10*, 20*, 501001*, 501005*, 502001*, 502005*, 503001*, and 503005*
- Records with temperatures < 250 °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 tables are in the file, 2002 NEI Stack Parameter Default, and can be found at the following address:

ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/ .

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, metric conversions are provided for convenience.

Stack Height:	10 ft	3.048 m
Stack Temperature:	72 °F	295.37 °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

We will send a Stack Parameter QA report to data providers. The report contains all of the emissions release point records submitted and identifies which parameters were defaulted as a result of our QA. We will request that data providers review the defaulted records and revise the records if they do not agree with the defaulted values. If we do not receive any comments from data providers, we will retain the defaulted values.

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
- 6 = State/Local Agency or Tribe revision to the defaulted value
- 8 = CAMD Value

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 based on SCC defaults, and exit gas flow rate is calculated based on the stack diameter and exit gas velocity values.

8.4 Stack Parameters QC Report

We will prepare a QC report that contains a list of emission release point records that failed our QA of stack parameters and provides our defaulted parameters. We will send these reports to the data providers and will ask data providers to review the defaults and provide corrected stack

parameters. Replacing stack parameters could improve the quality of the data needed for modeling.. Data providers can review the defaults and provide corrected stack parameters if desired. If we do not receive corrections, we will use the defaulted values in the 2002.

We will prepare a file, "StackParameterQA", that contains tables and forms for submitting revisions.

The following forms will be included in the file.

- Stack Parameter QC Report - Summarizes QC of all stack parameters and provides information on the number of records defaulted for each stack parameter
- View/Revised Defaulted Emission Release Type Records - contains a list of all Emission Release Point records that have the Emissions Release Type defaulted based on SCCs; the report presents the original and defaulted Emission Release Type and reviewers may revise these records by revising the SCCs of the records
- View/Revised Converted Exit Gas Flow Rate Records - contains a list of all Emission Release Point records where the Exit Gas Flow Rate have been defaulted using the stack flow equation in Section 8.2; the report presents the original and defaulted parameters and reviewers may revise these parameters
- View/Revised Stack Parameters Records - contains a list of all Emission Release Point records where the stack parameters have been defaulted; the report presents the original and defaulted parameters and reviewers may revise these parameters

The following tables will be included in the StackParameterQA file.

- Default tables
 - ▶ National defaults for fugitive releases
 - ▶ National defaults for stacks
 - ▶ MACT Code defaults
 - ▶ SCC defaults
 - ▶ SIC Code defaults
- Stack Default Codes for all 5 stack parameters
- Stack Parameter Revision Table - contains all stack records (both defaulted and non-defaulted records)

The Stack Parameter Revision contains the following data fields. Revisions will be provided by data providers by filling in the fields ending with "REVISED".

- StateCountyFIPS
- StateFacilityIdentifier
- EmissionReleasePointID
- TribalCode
- FacilityName
- LocationAddress
- City
- State
- ZipCode

- StackDefault Codes
- EmissionReleasePointType_ORIGINAL
- EmissionReleasePointType_DEFAULTED
- EmissionReleasePointType_REVISIED
- StackHeight_ORIGINAL
- StackHeight_DEFAULTED
- StackHeight_REVISIED
- StackDiameter_ORIGINAL
- StackDiameter_DEFAULTED
- StackDiameter_REVISIED
- ExitGasTemperature_ORIGINAL
- ExitGasTemperature_DEFAULTED
- ExitGasTemperature_REVISIED
- ExitGasVelocity_ORIGINAL
- ExitGasVelocity_DEFAULTED
- ExitGasVelocity_REVISIED
- ExitGasFlowRate_ORIGINAL
- ExitGasFlowRate_DEFAULTED
- ExitGasFlowRate_REVISIED
- MACT Code
- SCC
- SIC Code
- ConvertFlowConsistency - Indicates if calculated flow rate is inconsistent with other velocity
- Data Source
- Data Provider - person who provided data
- State File Name

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
- Fugitive Dimensions Unit
- Throughput Unit Numerator

Unit codes are used in a variety of data fields in the NIF. Table 13 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 13. 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
Horizontal Area Fugitive Release Height Fugitive	Fugitive Dimensions Unit
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 the 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
- Year of Emission Estimate
- 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:
<http://www.epa.gov/ttn/chief/emch/temporal/index.html> . 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 data submitted to EPA 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 emissions 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 Year of Emission Estimate

The Year of the emission estimate indicates the actual year of the estimate whereas the Inventory Year for the 2002 NEI is 2002. For example, some agencies collect HAP data every two years (1999, 2001, 2003, 2005, etc.), and they have reported 2001 data for the 2002 NEI. In this example, the Inventory Year in the 2002 NEI is 2002 and the Year of Emission Estimate is 2001.

We will populate the Year of the emission estimate using the Start and End Dates.

10.5 Annual Average Days Per Week

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

10.6 Annual Average Hours Per Day

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

10.7 Annual Average Hours Per Year

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

$$\text{Annual Ave Hrs Per Yr} = \text{Annual Ave Hrs Per Day} * \text{Annual Ave Days Per Week} * \text{Annual Ave Weeks Per Yr}$$

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.

10.8 Annual Average Weeks Per Year

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

10.9 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.

10.10 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.

10.11 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.

10.12 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 Dates.

10.13 Winter, Spring, Summer, and Fall Throughput Percentages

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:

<http://www.epa.gov/ttn/chief/emch/temporal/index.html> . NOTE: 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 Percentages should equal 100.

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

- If the sum is greater than 100, 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 Percentage of 40 and Summer Throughput Percentage 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 Percentage to 5 and Fall Throughput Percentage 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 QA and default the Control Status to indicate if emissions are controlled or uncontrolled. The valid values are “CONTROLLED”, “UNCONTROLLED”, and “UNKNOWN”.

We will QA the Control Status field by examining the control device codes in the CE table. Invalid values occur for the following two cases.

- All control device codes are “000” and the control status is “CONTROLLED”
- All control device codes are not “000” and the control status is “UNCONTROLLED”.

For invalid Control Status values, we will contact the data provider.

For missing values, we will use the following methodology.

- If a primary control device code is not reported in the CE table, we will default the status as “UNKNOWN”.
- 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 a decimal value instead of a 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 - 99.99.

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 14 presents the methodology that will be used to populate this data field.

Table 14. 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	<ol style="list-style-type: none"> 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 14.

SECTION 12. EMISSIONS

This section includes the following topics.

- Pollutant Code
- HAPs Emissions Reporting
- CAPs Emissions Reporting
- PM Filterable and Condensable Augmentation for PM2.5 and PM10
- VOC and PM Augmentation from HAPs
- CAP Augmentation for Missing CAPs
- HAPs Augmentation from VOC and PM
- HAP Performance level
- QC of HAP Category Group Emissions
- Emissions Type Period
- Annual and Non-Annual Emissions
- Emissions Ton Value
- Emissions QC
- Emission Data Quality Rating
- Emissions Data Source
- Historical Emissions Table
- Toxicity Weighting of HAPs

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. In addition, we will assign an emission data quality rating to every HAP record and identify the source of data for HAPs and CAPs. 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.

NEI Pollutant Code Table

We will prepare a NEI Pollutant Code Table (Table 15). The NEI Pollutant Code Table is contained in the 2002 State Lookup File, 02nei_lkup_states.zip, found at:

<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

More information about specific fields in Table 15 will be discussed later in this section.

Table 15. NEI Pollutant Code Table

Field Name	Description
Pollutant Code	NEI pollutant code Chemical Abstracts Services (CAS) number for individual compounds; 2 and 3 digit numerical codes for HAP Category groups and 3 letter Abbreviation for CAPs
Pollutant Code Description	NEI pollutant description
HAP Category Name	For HAP pollutant groups, HAP group name
Chemical Formula	Chemical formula for pollutants that are not compound groups
Molecular Weight	Molecular Weight for pollutants that are not compound groups
Active	Yes or No: Is Pollutant Code Active?
Last Modification Date	Date code was added or deleted from NIF Code Tables
Valid for NIF 2.0	Yes or No
Valid for NIF 3.0	Yes or No
Comments	
33 Urban HAP	Yes or No: 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 .
33 Urban HAP 1990 baseline NTI	Yes or No
NATO TEQ	NATO TEQ Factor
WHO TEQ	WHO TEQ Factor
I TEQ	I TEQ Factor
Ozone Precursor	Yes or No: Is HAP a VOC?
PM Precursor	Yes or No: Is HAP a PM?
HAP or CAP	Identify whether pollutant is a HAP or CAP
Pollutant Rank	Hierarchy used to select pollutants and emissions within a HAP pollutant group where emissions of both individual compounds within a group and a compound group are reported for an emissions record

CAPHAP Data Field

In the NOF EM Table, we will identify whether the pollutant in each record is a HAP or a CAP. We will populate this field as CAP or HAP and will use the NEI Pollutant Code Table to populate this field. Because the NEI contains both CAPs and HAPs, EIAG made the decision that emissions data for lead, which is both a CAP and a HAP, will be treated in the NEI as a 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 enables users to search for chemical substances across EPA programs and databases. (For more information on EPA data standards, go to: [http://oaspub.epa.gov/edr/epastd\\$.startup#1](http://oaspub.epa.gov/edr/epastd$.startup#1)) 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 Chemical Registry System (CRS) by EPA when CAS number is 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 each NEI HAP Category. EPA's Office of Environmental Information (OEI) assigns the following fields of data in Table 16.

- CRS ID
- ChemSystematicName
- EPA Chemical ID
- EPAChemRegistryName

OEI also verifies the CASRN.

We will use the 99 NEI Chemical Data Standards Table for pollutants as a starting point and will add any new pollutants to the Table. For new pollutants in 2002, we will submit a list of pollutants to OEI for assignment of Chemical Data Standards information. The process of assigning Chemical Registry System data fields in the 2002 NEI will not occur until after external revisions are received and the final list of pollutants in the 2002 NEI is compiled.

The NEI Chemical Data Standards Table is contained in the 2002 State Lookup File, 02nei_lkup_states.zip, found at:

<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

Table 16. NEI Chemical Data Standards Table

Field Name	Description
CRS ID	Chemical Registry System ID
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 (9 th Collective Index Name)
EPAChemicalID	EPA Chemical Identifier
EPAChemRegistryName	EPA Chemical Registry Name
HAPCAT_CRS ID	Chemical Registry System ID (assigned to HAP category)
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 (9 th Collective Index Name)(assigned to HAP category)
HAPCAT_EPAChemicalID	EPA Chemical Identifier (assigned to HAP category)
HAPCAT_EPAChemRegistryName	EPA Chemical Registry Name (assigned to HAP category)

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.

(<http://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/pollsour.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.

For risk assessments, individual speciated HAPs are needed because the toxicity associated with an individual compound within a compound group varies widely. For risk, exposure and dispersion modeling, we prepare files containing speciated pollutants using the NEI Pollutant Code and Description. For trends purposes, we are also asked to track the emissions over time for the Section 112b list. For trends analyses and summary NEI data, we prepare files using the Section 112b HAPs and aggregate 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 and the NEI Pollutant Description is methylmercury.

The 2002 NEI NEI Pollutant Code Table lists all of the specific pollutants and compound groups included in the 2002 NEI along with their Chemical Abstract Services (CAS) numbers for individual compounds; information about if the pollutant is a constituent of VOC, PM or both; TEQ factors; and other information. The NEI Pollutant Code Table is contained in the 2002 State Lookup File, 02nei_lkup_states.zip, found at:

<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

Please refer to this table for the following discussion of HAPs in the NEI.

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 and potential overestimation of emission levels and risk. CAS numbers referenced to in the text are also shown in 2002 NEI Pollutant Code Table 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 (NEI pollutant codes 7440473 and 136) – Separate chromium compounds into hexavalent chromium (NEI Pollutant Code 18540299) and trivalent (NEI Pollutant Code 16065831).

Lead (NEI pollutant code 195) – Separate lead compounds into organic lead (NEI pollutant code 603) and inorganic lead (NEI pollutant code 602).

Mercury (NEI pollutant codes 7439976, 199, and 22967926)– Separate mercury compounds into particulate divalent mercury (NEI Pollutant Code 202), elemental gaseous mercury (NEI Pollutant Code 200) and gaseous divalent mercury (NEI Pollutant Code 201). Default mercury speciation factors are found at: <ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

Nickel Compounds (NEI pollutant codes 226)- Separate nickel compounds into nickel subsulfide (NEI Pollutant Code 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 17 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 17 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 17. We also encourage the reporting of other

individual POM compounds for which cancer assessments are available (also listed in Table 17).

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 17. POM Compounds

7-PAH	15-PAH (includes 7-PAH compounds)	Other POM Compounds for which we have cancer assessments
Benz(a)anthracene	Benz(a)anthracene	Carbazole
Benzo(a)pyrene	Benzo(a)pyrene	Dibenz(a,h)acridine
Benzo(b)fluoranthene	Benzo(b)fluoranthene	Dibenz(a,j)acridine
Benzo(k)fluoranthene	Benzo(k)fluoranthene	7H-Dibenzo(c,g)carbazole
Chrysene	Chrysene	Dibenzo(a,e)pyrene
Dibenz(a,h)anthracene	Dibenz(a,h)anthracene	Dibenzo(a,i)pyrene
Indeno(1,2,3-cd)pyrene	Indeno(1,2,3-cd)pyrene	Dibenzo(a,l)pyrene
	Acenaphthene	7,12-Dimethylbenz(a)anthracene
	Acenaphthylene	1,6-Dinitropyrene
	Anthracene	1,8-Dinitropyrene
	Benzo(ghi)perylene	3-Methylcholanthrene
	Fluoranthene	5-Methylchrysene
	Fluorene	5-Nitroacenaphthalene
	Phenanthrene	6-Nitrochrysene
	Pyrene	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 furans into TEQ and non-TEQ groups (See 2002 NEI Pollutant Code Table). There are 3 sets of TEQ factors used to estimate TEQ - I TEQ factors, 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 Clean Air Act 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 TEQ, NATO TEQ or WHO TEQ. Separate pollutant codes are in the NEI Pollutant Code Table for I TEQ (NEI pollutant code 626), NATO TEQ or WHO TEQ (NEI pollutant code 627).
 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 from 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/tri/guide_docs/2000/glycol2000.pdf . Note that, historically, many compounds have been mistakenly included in the glycol ethers compound group. Table 18 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”.

Table 18. 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

- **Xylenes and Cresols:** Xylenes include ortho-, meta-, and para- isomers. Cresols include ortho-, meta-, and para- 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)” (NEI Pollutant Code 1330207) or “cresols/cresylic acids” (NEI Pollutant Code 331).

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, NEI Pollutant Code Table, and NEI Chemical Data Standards Table.

We will augment HAPs and their emissions in the 2002 NEI if VOC and PM are reported for a facility, but no HAPs are reported, and a reliable estimation methodology exists.

CAPs Reporting

CAPs include: Volatile Organic Compounds (VOC), Nitrogen Oxides (NOX), Carbon Monoxide (CO), Sulfur Dioxide (SO₂), and Particulate Matter (PM). Ammonia (NH₃) will also be included in the 2002 CAP NEI. Lead is both a HAP and a CAP and has been included in previous discussions on HAPs Reporting. Table 19 describes PM compounds that may be reported to EPA by data providers for the 2002 NEI. We will generate missing emission estimates for filterable and primary PM-2.5, filterable and primary PM-10 and condensible PM in the 2002 NEI.

Table 19. 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.)

Table 19, continued. PM Compounds reported in the 2002 NEI

NEI Code	Pollutant	Pollutant Description
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.)

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 two cases in the 2002 NEI.

1. PM augmentation for filterable and primary PM-10, filterable and primary PM-2.5, and condensible PM
2. CAPs augmentation if a partial submission of CAPs is provided in a data submittal, e.g., a state agency only submits NOX.

12.2 PM Augmentation

All PM emissions estimates that are defaulted in the 2002 NEI will include data source codes to identify that the emissions are defaulted.

PM terms are related to each other via the following three equations.

$$\text{PM-PRI} = \text{PM-FIL} + \text{PM-CON}$$

$$\text{PM10-PRI} = \text{PM10-FIL} + \text{PM-CON}$$

$$\text{PM25-PRI} = \text{PM25-FIL} + \text{PM-CON}$$

Steps in PM Augmentation

From those three equations, we developed seven math steps and 14 ratios that will be applied to augment the reported PM terms in the 2002 point source NEI.

Before we run PM augmentation math steps, we will first review the data for inconsistencies. If values are found to be inconsistent, they will be replaced. These replacement values will have a data source code of "A-R." The consistency checks and replacement actions are as follows:

1. If $PM10-PRI > 0$ and $PM25-PRI > PM10-PRI$ (and $PM10-FIL$, $PM25-FIL$ and $PM-CON$ are null or 0), then set $PM25-PRI = PM10-PRI$.
2. If $PM10-PRI = 0$ and $PM25-PRI > PM10-PRI$ (and $PM10-FIL$, $PM25-FIL$ and $PM-CON$ are null or 0), then set $PM10-PRI = PM25-PRI$.
3. If $PM10-FIL > 0$ and $PM25-FIL > PM10-FIL$ (and $PM10-PRI$, $PM25-PRI$ and $PM-CON$ are null or 0), then set $PM25-FIL = PM10-FIL$.
4. If $PM10-FIL = 0$ and $PM25-FIL > PM10-FIL$ (and $PM10-PRI$, $PM25-PRI$ and $PM-CON$ are null or 0), then set $PM10-FIL = PM25-FIL$.
5. If $PM10-PRI > 0$ and $PM10-FIL > PM10-PRI$ (and $PM25-PRI$, $PM25-FIL$ and $PM-CON$ are null or 0), then set $PM10-FIL = PM10-PRI$.
6. If $PM10-PRI = 0$ and $PM10-FIL > PM10-PRI$ (and $PM25-PRI$, $PM25-FIL$ and $PM-CON$ are null or 0), then set $PM10-PRI = PM10-FIL$.
7. If $PM25-PRI > 0$ and $PM25-FIL > PM25-PRI$ (and $PM10-PRI$, $PM10-FIL$ and $PM-CON$ are null or 0), then set $PM25-FIL = PM25-PRI$.
8. If $PM25-PRI = 0$ and $PM25-FIL > PM25-PRI$ (and $PM10-PRI$, $PM10-FIL$ and $PM-CON$ are null or 0), then set $PM25-PRI = PM25-FIL$.
9. If $PM10-PRI > 0$ and $PM25-FIL > PM10-PRI$ (and $PM10-FIL$, $PM25-PRI$ and $PM-CON$ are null or 0), then set $PM25-FIL = PM10-PRI$.

Math Steps 1-7 will be used to augment PM.

Math Step 1

PM-CON term is needed. $PM-PRI$ and $PM-FIL$ are available
 $PM-CON = PM-PRI - PM-FIL$

Math Step 2

PM-CON term is needed. $PM10-PRI$ and $PM10-FIL$ are available
 $PM-CON = PM10-PRI - PM10-FIL$

Math Step 3

PM-CON term is needed. PM25-PRI and PM25-FIL are available
 $PM-CON = PM25-PRI - PM25-FIL$

Math Step 4

PM10-FIL term is needed. PM10-PRI and PM-CON are available
 $PM10-FIL = PM10-PRI - PM-CON$

Math Step 5

PM10-PRI term is needed. PM10-FIL and PM-CON are available
 $PM10-PRI = PM10-FIL + PM-CON$

Math Step 6

PM25-FIL term is needed. PM25-PRI and PM-CON are available
 $PM25-FIL = PM25-PRI - PM-CON$

Math Step 7

PM25-PRI term is needed. PM25-FIL and PM-CON are available
 $PM25-PRI = PM25-FIL + PM-CON$

If Math Steps 1 through 7 are not adequate to augment all of the missing PM term(s), additional augmentation steps will be needed that require the use of ratios developed from available emissions and particle size distribution data. These ratios will be needed when only one PM term is available, and two or more terms need to be augmented. Examples of how we will use the PM ratios are shown below.

$$FIL \times Ratio_{CON/FIL} = CON$$

$$PRI \times Ratio_{CON/PRI} = CON$$

$$CON \times Ratio_{FIL/CON} = FIL$$

$$CON \times Ratio_{PRI/CON} = PRI$$

Ratios 1 through 14 will be used to complete PM augmentation.

Ratio 1

PM25-FIL term is needed. PM10-FIL is available
 $PM25-FIL = PM10-FIL \times Ratio_{PM25-FIL/PM10-FIL}$

Ratio 2

PM10-FIL term is needed. PM25-FIL is available
 $PM10-FIL = PM25-FIL \times Ratio_{PM10-FIL/PM25-FIL}$

Ratio 3

PM-CON term is needed. PM10-PRI is available
 $PM-CON = PM10-PRI \times Ratio_{PM-CON/PM10-PRI}$

Ratio 4

PM25-PRI term is needed. PM10-PRI is available

$$\text{PM25-PRI} = \text{PM10-PRI} \times \text{Ratio}_{\text{PM25-PRI/PM10-PRI}}$$

Ratio 5

PM-CON term is needed. PM25-PRI is available

$$\text{PM-CON} = \text{PM25-PRI} \times \text{Ratio}_{\text{PM-CON/PM25-PRI}}$$

Ratio 6

PM10-PRI term is needed. PM25-PRI is available

$$\text{PM10-PRI} = \text{PM25-PRI} \times \text{Ratio}_{\text{PM10-PRI/PM25-PRI}}$$

Ratio 7

PM-CON term is needed. PM10-FIL is available

$$\text{PM-CON} = \text{PM10-FIL} \times \text{Ratio}_{\text{PM-CON/PM10-FIL}}$$

Ratio 8

PM25-FIL term is needed. PM10-FIL is available

$$\text{PM25-FIL} = \text{PM10-FIL} \times \text{Ratio}_{\text{PM25-FIL/PM10-FIL}}$$

Ratio 9

PM-CON term is needed. PM25-FIL is available

$$\text{PM-CON} = \text{PM25-FIL} \times \text{Ratio}_{\text{PM-CON/PM25-FIL}}$$

Ratio10

PM10-PRI is needed. PM-CON is available

$$\text{PM10-PRI} = \text{PM-CON} \times \text{Ratio}_{\text{PM10-PRI/PM-CON}}$$

Ratio 11

PM10-FIL is needed. PM-CON is available

$$\text{PM10-FIL} = \text{PM-CON} \times \text{Ratio}_{\text{PM10-FIL/PM-CON}}$$

Ratio 12

PM-CON is needed. PM-PRI is available

$$\text{PM-CON} = \text{PM-PRI} \times \text{Ratio}_{\text{PM-CON/PM-PRI}}$$

Ratio 13

PM10-FIL is needed. PM-FIL is available

$$\text{PM10-FIL} = \text{PM-FIL} \times \text{Ratio}_{\text{PM10-FIL/PM-FIL}}$$

Ratio 14

PM25-FIL is needed. PM-FIL is available

$$\text{PM25-FIL} = \text{PM-FIL} \times \text{Ratio}_{\text{PM25-FIL/PM-FIL}}$$

Prior to the application of these math steps and ratios, we will analyze the data and separate the data into cases. The cases will determine which combination of ratios and math steps to apply. Table 20 presents an overview of these cases and their augmentation methodology.

Table 20. PM Cases and Required Steps to Augment PM

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 Ratio 12 PM-FIL = PM-PRI - PM-CON, Math Step 1 PM10-FIL = PM-FIL x Ratio 13 PM25-FIL = PM-FIL x Ratio 14 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-PRI	PM-CON PM25-PRI PM25-FIL PM10-FIL	PM-CON = PM10-PRI x Ratio 3 PM10-FIL = PM10-PRI - PM-Con, Math Step 4 PM25-FIL = PM10-FIL x Ratio 8 PM25-PRI = PM-CON + PM25-FIL, Math Step 7
PM10-FIL	PM-CON PM25-PRI PM25-FIL PM10-PRI	PM-CON = PM10-FIL x Ratio 7 PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM25-PRI	PM-CON PM25-FIL PM10-PRI PM10-FIL	PM-CON = PM25-PRI x Ratio 5 PM25-FIL = PM25-PRI - PM-Con, Math Step 6 PM10-FIL = PM25-FIL x Ratio 2 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM25-FIL	PM-CON PM25-PRI PM10-FIL PM10-PRI	PM-CON = PM25-FIL x Ratio 9 PM10-FIL = PM25-FIL x Ratio 2 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM-CON	PM25-PRI PM25-FIL PM10-PRI PM10-FIL	PM10-FIL = PM-CON x Ratio 11 PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM-FIL	PM-CON PM25-PRI PM25-FIL PM10-PRI PM10-FIL	PM10-FIL = PM-FIL x Ratio 13 PM-CON = PM10-FIL x Ratio 7 PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-FIL PM10-PRI	PM-CON PM25-PRI PM25-FIL	PMCON = PM10-PRI - PM10-FIL, Math Step 2 PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PMCON + PM25-FIL, Math Step 7

Table 20 continued. PM Cases and Required Steps to Augment PM

PM Reported	PM to Default	Augmentation Methodology
PM10-FIL PM-CON	PM10-PRI PM25-PRI PM25-FIL	PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-FIL PM25-PRI	PM-CON PM25-FIL PM10-PRI	PM-CON= PM10-FIL x Ratio7 PM25-FIL = PM25-PRI - PM-Con, Math Step 6 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-PRI PM25-FIL	PM-CON PM25-PRI PM10-FIL	PM-CON= PM10-PRI x Ratio 3 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-FIL = PM10-PRI - PM-CON, Math Step 4
PM10-PRI PM-CON	PM10-FIL PM25-PRI PM25-FIL	PM10-FIL= PM10-PRI - PM-CON, Math Step 4 PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PM-CON + PM25-FIL, Math Step 7
PM25-FIL PM25-PRI	PM-CON PM10-PRI PM10-FIL	PM-CON= PM25-PRI - PM25-FIL, Math Step 3 PM10-FIL = PM25-FIL x Ratio2 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM25-FIL PM-CON	PM10-PRI PM25-PRI PM10-FIL	PM10-FIL = PM25-FIL x Ratio 2 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM-CON PM25-PRI	PM25-FIL PM25-PRI PM10-FIL	PM25-FIL= PM25-PRI - PM-CON, Math Step 6 PM10-FIL = PM25-FIL x Ratio 2 PM10-PRI = PM-CON + PM1025-FIL, Math Step 5
PM10-FIL PM25-FIL	PM-CON PM25-PRI PM10-PRI	PM-CON = PM10-FIL x Ratio 7 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-PRI PM25-PRI	PM-CON PM25-FIL PM10-FIL	PM-CON = PM10-PRI x Ratio 3 PM25-FIL= PM25-PRI - PM-CON, Math Step 6 PM10-FIL= PM10-PRI - PM-CON, Math Step 4
PM-PRI PM-FIL	PM-CON PM25-PRI PM25-FIL PM10-PRI PM10-FIL	PM-CON = PM-PRI - PM-FIL, Math Step 1 PM10-FIL = PM-FILx Ratio13 PM25-FIL = PM-FIL x Ratio14 PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-FIL PM10-PRI PM25-FIL	PM-CON PM25-PRI	PM-CON = PM10-PRI - PM10-FIL, Math Step 2 PM25-PRI = PM-CON + PM25-FIL, Math Step 7
PM10-FIL PM10-PRI PM25-PRI	PM-CON PM25-FIL	PM-CON = PM10-PRI - PM10-FIL, Math Step 2 PM25-FIL= PM25-PRI - PM-CON, Math Step 6

Table 20 continued. PM Cases and Required Steps to Augment PM

PM Reported	PM to Default	Augmentation Methodology
PM10-PRI PM25-PRI PM25-FIL	PM-CON PM10-FIL	PM-CON= PM25-PRI - PM25-FIL, Math Step 3 PM10-FIL= PM10-PRI - PM-CON, Math Step 4
PM10-FIL PM25-FIL PM-CON	PM25-PRI PM10-PRI	PM25-PRI = PM-CON + PM25-FIL, Math Step 7 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-FIL PM25-FIL PM25-PRI	PM-CON PM10-PRI	PM-CON= PM25-PRI - PM25-FIL, Math Step 3 PM10-PRI = PM-CON + PM10-FIL, Math Step 5
PM10-FIL PM10-PRI PM-CON	PM25-PRI PM25-FIL	PM25-FIL = PM10-FIL x Ratio 1 PM25-PRI = PM-CON + PM25-FIL, Math Step 7
PM10-PRI PM25-PRI PM-CON	PM25-FIL PM10-FIL	PM25-FIL= PM25-PRI - PM-CON, Math Step 6 PM10-FIL= PM10-PRI - PM-CON, Math Step 4
PM10-FIL PM10-PRI PM25-FIL PM25-PRI	PM-CON	PM-CON= PM25-PRI - PM25-FIL, Math Step 3
PM10-FIL PM10-PRI PM25-PRI PM-CON	PM25-FIL	PM25-FIL= PM25-PRI - PM-CON, Math Step 6
PM10-FIL PM10-PRI PM25-FIL PM-CON	PM25-PRI	PM25-PRI = PM-CON + PM25-FIL, Math Step 7
PM10-PRI PM25-FIL PM25-PRI PM-CON	PM10-FIL	PM10-FIL= PM10-PRI - PM-CON, Math Step 4
PM10-FIL PM25-FIL PM25-PRI PM-CON	PM10-PRI	PM10-PRI = PM-CON + PM10-FIL, Math Step 5

Creation of PM Ratios

We developed the ratios 1 through 14 from the following sources.

- PM calculator

- AP-42
- 2002 Electric Generating Unit (EGU) NEI data developed by EIAG and used to supplement the state and local agency and tribal inventories.

Some of these ratios are SCC and control device specific. We will use a hierarchy when we use the ratios. Unfortunately, in some cases, particularly for PM-CON, only generic ratios can be used. Information about the ratios and the hierarchy of their use in PM Augmentation are shown below.

1. EGU Data

- a. PM10 and PM25 primary and filterable
- b. SCC and control device specific
- c. Developed by EIAG using PM Calculator and AP-42 data
- d. Applicable only to SCCs that start with 101 and 201

2. PM Calculator

- a. Applicable only to filterable PM; Does not include PM-CON ratios
- b. SCC and control device specific

3. AP-42

- a. Only source of PM-CON ratios; Not limited to filterable PM
- b. SCC specific and sometimes control device specific
- c. Limited number of SCCs;
- d. General/generic PM ratios based on SCC 1-digit least preferred
- e. Developed using emission factors (preferred) or particle distribution

The file, PM Augmentation Ratios, contains the ratios and is available at at:
ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/ .

A detailed description of how we created PM ratios is presented below.

- **EGU Data**
Ratios developed using EGU data are SCC and control device specific. We developed ratios by using the available PM emissions in the EGU data disaggregated by SCC. We then averaged ratios according to SCCs. However EGU ratios only applied to SCCs starting with 101 and 201. Details on the development of the PM emission estimates for EGUs can be found in the “Documentation for the 2002 Electric Generating Unit (EGU) National Emissions Inventory (NEI)” (Eastern Research Group and E.H. Pechan, 2004). The 2002 NEI EGU Point Source Documentation can be found at:
<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point> .
- **PM-Calculator**
Ratios developed using the PM Calculator are both SCC and control device specific. The PM Calculator is EPA’s software for estimating various PM terms from other PM terms (U.S. EPA, 2004a). The PM Calculator is applicable to point sources only, and requires

the user to input uncontrolled emissions (either total PM-FIL or PM10-FIL) for each source, the SCC and the control device, if any. The program then calculates controlled emissions for PM25-FIL and PM10-FIL. This method is preferred because it is SCC and control device specific. Because the PM Calculator is limited in the PM terms it can calculate, PM ratios will have to be found using other sources of data.

The PM Calculator can only calculate controlled or uncontrolled PM10-FIL and PM25-FIL using uncontrolled PM10-FIL or PM-FIL. The PM Calculator does not directly import NIF files, so we developed generic ratios from the output file by inputting a simple dummy file. We uploaded the dummy file, with all of the SCCs and control code device codes in the 2002 point source inventory, into the PM Calculator. Emissions of 100 tons were entered as the emissions in the dummy file for simplicity. The PM Calculator then created an output file that listed for SCC and control device codes the following information.

- PM uncontrolled
- PM10 uncontrolled
- PM25 uncontrolled
- PM10 controlled
- PM25 controlled
- PM10 control efficiency (CE)
- PM25 CE.

Then we developed ratios by simply dividing the term needed by the term available.

To the extent possible, we developed PM-FIL ratios using the PM Calculator for all of the SCCs and control device codes in the 2002 point source state, local, and tribal agency submittals. When a ratio based on SCC and control device could not be developed using PM Calculator, we used AP-42 emission factors or particle size distributions.

- **Emission Factors from AP-42**

We developed PM ratios using the emission factors from the EPA's AP-42 (U.S. EPA, 2004b). There were, however, a few drawbacks associated with the use of AP-42 emission factors in this manner. In some cases, the ash content or the sulfur content in the fuel used in the process were needed to calculate the emission factors. Unfortunately, most of the state, local, and tribal records did not include sulfur content or ash content.

To overcome this shortcoming, we used assumptions from the 1999 NEI for CAPs. (Strait et al., 2003). For coal combustion, we assumed that the sulfur content was 1% and ash content was 8% for SCCs that start with 1 and 2. We also assumed that PM-CON was only controlled with flume gas desulfurization. To develop the PM ratios, we grouped PM emission factors for various PM terms based on their SCCs. This required PM emission factors for both the available PM and the needed PM term.

In some cases, ash content and sulfur content were not needed to develop PM ratios. We developed PM ratios for these SCCs by looking for emission factors for the various PM

terms available. However, there still were situations when there were not enough data, and we had to use other surrogate data to develop PM ratios.

- **Cumulative Particle Distribution from AP-42**

AP-42 also has particle distribution data for some SCCs based on the 6 and 8 digit level. These are less specific than the ratios developed from the PM Calculator and AP-42 emission factors, and will only be used when there are no corresponding ratios from EGU Data, PM Calculator, or AP-42 emission factors.

To develop these ratios, we made the following assumptions:

- Percent distribution less than or equal to 1 micron are PM-CON emissions
- Percent distribution less than or equal to 2.5 microns are PM25-PRI emissions (since they would also include emissions less than or equal to 1 micron)
- Percent distribution less than or equal to 10 microns are PM10-PRI emissions (since they would also include emissions less than or equal to 1 micron)
- The percent for PM-PRI will always be 100%

The data available from the AP-42 includes the percent distribution less than or equal to 1 micron, the percent distribution less than or equal to 2.5 microns, and the percent distribution less than or equal to 10 microns. Therefore, the data were easily converted into ratios. First, we labeled the percent distributions as PM-CON, PM25-PRI, and PM10-PRI. Then we calculated PM-FIL, PM25-FIL, and PM10-FIL using the available terms. Finally, we calculated the ratios by dividing the needed term by the available term. For example, if PM-PRI is available and PM10-FIL is needed, you can simply multiply PM-PRI by the ratio of PM10-FIL/PM-PRI, to calculate PM10-FIL.

- **Other SCC Emission Factors**

For SCCs that start with 3 and 4, PM emissions can occur from fuel fired equipment, process heaters, incinerators, and coating ovens. For SCCs that start with 3 or 4 and have combustion sources, we developed surrogate ratios based on emission factors for industrial external combustion boilers. These surrogates were matched on the fuel type used. Table 21 shows some of the SCCs that were used as surrogates for the SCCs that start with 402 and have combustion sources. For coating ovens, we assumed that natural gas is the fuel used.

SCCs that start with 5, solid waste disposal, did not have the necessary emission factors (or cumulative particle distribution data discussed below) in AP-42. Therefore, we used average emission factors for external combustion sources and stationary internal combustion sources. A list of these SCCs is provided in Table 22.

Table 21. PM Augmentation: SCCs Starting with “402” and the Surrogate SCCs

SCC	SCC_L3	SCC_L4	SCC to Use	Description of SCC to Use
402008	Coating Oven - General		102006	Industrial External Fuel Combustion, Natural Gas
40200801	Coating Oven - General	General	102006	Industrial External Fuel Combustion, Natural Gas
40200803	Coating Oven - General	Baked > 175F **	102006	Industrial External Fuel Combustion, Natural Gas
40200810	Coating Oven - General	General	102006	Industrial External Fuel Combustion, Natural Gas
40200820	Coating Oven - General	Prime/Base Coat Oven	102006	Industrial External Fuel Combustion, Natural Gas
40200830	Coating Oven - General	Topcoat Oven	102006	Industrial External Fuel Combustion, Natural Gas
40200840	Coating Oven - General	Two Piece Can Curing Ovens: General	102006	Industrial External Fuel Combustion, Natural Gas
40200845	Coating Oven - General	Three Piece Can Curing Ovens	102006	Industrial External Fuel Combustion, Natural Gas
40200870	Coating Oven - General	Color Coat Oven	102006	Industrial External Fuel Combustion, Natural Gas
40200898	Coating Oven - General	General	102006	Industrial External Fuel Combustion, Natural Gas
40200899	Coating Oven - General	See Comment **	102006	Industrial External Fuel Combustion, Natural Gas
40201001	Coating Oven Heater	Natural Gas	102006	Industrial External Fuel Combustion, Natural Gas
40201002	Coating Oven Heater	Distillate Oil	102005	Industrial External Fuel Combustion, Distillate Oil
40201122	Fabric Coating/Printing	Fabric Print:Dryer: Fuel-fired	102006	Industrial External Fuel Combustion, Natural Gas
40204340	Fabric Coating, Dip Coating	Drying/Curing	102006	Industrial External Fuel Combustion, Natural Gas
40290013	Fuel Fired Equipment	Natural Gas: Incinerator/Afterburner	102006	Industrial External Fuel Combustion, Natural Gas
40290023	Fuel Fired Equipment	Natural Gas: Flares	102006	Industrial External Fuel Combustion, Natural Gas

Table 22. PM Augmentation: SCCs Starting With “5” That Have Combustion Sources (Incinerators and Flares) That Can Use Average Emission Factors from External Combustion Sources and Stationary Internal Combustion Sources

SCC	SCC_L2	SCC_L3	SCC_L4
501001	Government	Municipal Incineration	
50100101	Government	Municipal Incineration	Starved Air: Multiple Chamber
50100102	Government	Municipal Incineration	Mass Burn: Single Chamber
50100103	Government	Municipal Incineration	Refuse Derived Fuel
50100104	Government	Municipal Incineration	Mass Burn Refractory Wall Combustor
50100105	Government	Municipal Incineration	Mass Burn Waterwall Combustor
50100106	Government	Municipal Incineration	Mass Burn Rotary Waterwall Combustor
50100107	Government	Municipal Incineration	Modular Excess Air Combustor
50100108	Government	Municipal Incineration	Fluidized Bed: Refuse Derived Fuel
50100410	Government	Landfill Dump	Waste Gas Destruction: Waste Gas Flares
50100411	Government	Landfill Dump	Waste Gas Destruction: Incinerators
50100505	Government	Other Incineration	Medical Waste Incinerator, unspecified type, Infectious wastes only
50100506	Government	Other Incineration	Sludge
50100515	Government	Other Incineration	Sludge: Multiple Hearth
50100516	Government	Other Incineration	Sludge: Fluidized Bed
502001	Commercial/ Institutional	Incineration	
50200101	Commercial/ Institutional	Incineration	Multiple Chamber
50200102	Commercial/ Institutional	Incineration	Single Chamber
50200501	Commercial/ Institutional	Incineration: Special Purpose	Med Waste Controlled Air Incin-aka Starved air, 2-stg, or Modular comb
50200502	Commercial/ Institutional	Incineration: Special Purpose	Med Waste Excess Air Incin - aka Batch, Multiple Chamber, or Retort
50200503	Commercial /Institutional	Incineration: Special Purpose	Medical Waste Rotary Kiln Incinerator
50200504	Commercial/ Institutional	Incineration: Special Purpose	Medical Waste Incinerator, unspecified type (use 502005-01, -02, -03)
50200505	Commercial/ Institutional	Incineration: Special Purpose	Medical Waste Incinerator, unspecified type, Infectious wastes only
50200507	Commercial/ Institutional	Incineration: Special Purpose	VOC Contaminated Soil
50200515	Commercial/ Institutional	Incineration: Special Purpose	Sewage Sludge Incinerator: Multiple Hearth
50200518	Commercial/ Institutional	Incineration: Special Purpose	Sewage Sludge Incinerator: Single Hearth Cyclone
50200601	Commercial/ Institutional	Landfill Dump	Waste Gas Flares ** (Use 5-01-004-10)

Table 22. PM Augmentation: SCCs Starting With “5” That Have Combustion Sources (Incinerators and Flares) That Can Use Average Emission Factors from External Combustion Sources and Stationary Internal Combustion Sources

SCC	SCC_L2	SCC_L3	SCC_L4
50300101	Industrial	Incineration	Multiple Chamber
50300102	Industrial	Incineration	Single Chamber
50300107	Industrial	Incineration	Trench Burner: Tires
50300112	Industrial	Incineration	Mass Burn Waterwall Combustor
50300501	Industrial	Incineration	Hazardous Waste
50300502	Industrial	Incineration	Hazardous Waste Incinerators: Fluidized Bed
50300503	Industrial	Incineration	Hazardous Waste Incinerators: Liquid Injection
50300506	Industrial	Incineration	Sludge
50300520	Industrial	Incineration	Sewage Sludge Incinerator: High Pressure, Wet Oxidation
50300599	Industrial	Incineration	Fuel Not Classified
50300601	Industrial	Landfill Dump	Waste Gas Flares
50300789	Industrial	Liquid Waste	Sludge Digester Gas Flare

- **1st Digit SCC AP-42 Data**

At the most generic level, we calculated ratios based on the 1st digit of the SCC if specific SCC ratios could not be found. We developed these from AP-42 emission factors and cumulative particle distribution. Again, ratios based on emission factors are preferred over those based on cumulative particle distribution. These 1st digit SCC based ratios also have very generic control efficiencies that can be applied. For this method, we used the same assumptions and methods for the 6 and 8 digit SCCs. In addition, we also developed multipliers that can be used to convert the uncontrolled emissions into controlled emissions. We will only use these very generic ratios as a last resort to augment PM.

The EPA has revised the PM emissions from natural gas combustion in the final version of the 2002 NEI. The revision decreased the emissions of PM10-PRI, PM2.5-PRI, and PM-CON. The reason for this adjustment is that EPA believes that the current AP-42 factors for condensible emissions are too high. The EPA is basing this adjustment on some limited data from the draft EPA dilution method that is similar to conditional test method (CTM) 39

(<http://www.epa.gov/ttn/emc/ctm.html>) that measures PM10-PRI and PM2.5-PRI directly.

The data that this adjustment is based on can be found at:

[http://www.nyserda.org/programs/Environment/emepreports.asp#Fine Particulates](http://www.nyserda.org/programs/Environment/emepreports.asp#Fine%20Particulates).

The file, “Ratios to adjust PM values in NEI for natural gas combustion, August 2005”, contains the PM natural gas ratios by SCC and is available at:

ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/ .

The Data Source codes for emissions that have PM augmentation are “A-R” and “A”. “A” denotes augmented PM data that used original reported PM in augmentation procedures. “A-R” denotes augmented PM data that required the replacement of existing PM data submitted to NEI due to inconsistency in the reporting of PM by data submitters.

References

Eastern Research Group and E.H. Pechan, 2004. “Documentation for the 2002 Electric Generating Unit (EGU) National Emissions Inventory (NEI).” Prepared for the Emission Inventory Group, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Strait et al. 2003. PM Augmentation Procedures for the 1999 Point and Area Source NEI. Presented at the 12th International Emission Inventory Conference "Emission Inventories - Applying New Technologies," San Diego, CA. April 29 - May 1, 2003.

U.S. EPA, 2004a. PM Calculator. Emission Inventory Group, U.S. Environmental Protection Agency, Research Triangle Park, NC. Accessed February 23, 2004.

U.S. EPA, 2004b. Compilation of Air Pollutant Emission Factors, AP-42, Fifth Edition, Volume I: *Stationary Point and Area Sources*. <http://www.epa.gov/ttn/chief/ap42/index.html>

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 NEI Pollutant Code Table to identify which HAPs are VOC and/or PM constituents.

For facilities that have HAP VOC and HAP PM emissions but do not have VOC or PM emissions, we will not augment CAP inventories to include the HAP VOC emissions and HAP PM emissions in the 2002 NEI. This is an area of ongoing methodological development that will be addressed in the 2008 NEI.

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 will 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 to default emissions of missing CAPs.

12.5 HAP Augmentation from VOC and PM

We will augment HAPs emitted from boilers for facilities that have reported VOC or PM boiler emissions, but no HAP boiler estimates. If throughput is reported, we will use emission factors and reported throughput to generate HAP emissions. If throughput is not available, we will develop ratios of HAP emission factors to CAP emissions factors and apply the ratios to reported CAP emissions to develop HAP emission estimates.

Any HAP boiler emission estimates defaulted from VOC and PM emissions will be assigned a data source code of “A-B”.

12.6 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. Acceptable NIF code values are:

01 Actual

- 02 Allowable
- 03 Maximum
- 04 Potential

When we prepare files for various uses, we use a hierarchy to select only one type of emissions to avoid double counting. For 2002 NATA modeling, the following hierarchy will be used.

- Actual
- Allowable
- Potential
- Maximum

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

12.7 QC of HAP Category Group Emissions

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, we will delete records with the compound group emissions and retain speciated compounds. In cases where a single data source reports emissions of individual compounds and compound group, the following actions will be taken.

- **Antimony:**

1. If speciated Antimony emissions are greater than or equal to Antimony and Compounds (NEI pollutant code 92) emissions, delete Antimony and Compounds record.
2. If Antimony and Compounds emissions are greater than speciated Antimony emissions, delete speciated Antimony records.
3. If emissions are speciated and no Antimony and Compounds emissions are reported (Antimony and Antimony Trioxide for example), keep all records.

- **Arsenic:**

1. If speciated Arsenic emissions are greater than or equal to Arsenic and Compounds (NEI pollutant code 93) emissions, delete Arsenic and Compounds record.
2. If Arsenic and Compounds emissions are greater than speciated Arsenic emissions, delete individual pollutant records.
3. If emissions are speciated and no Arsenic and Compounds emissions are reported (Arsenic and Arsine for example), keep all records.

- **Beryllium:**
 1. If speciated Beryllium emissions are greater than or equal to Beryllium and Compounds (NEI pollutant code 109) emissions, delete Beryllium and Compounds record.
 2. If Beryllium and Compounds emissions are greater than speciated Beryllium emissions, delete speciated Beryllium records.
 3. If emissions are speciated and no Beryllium and compounds emissions are reported (Beryllium and Beryllium oxide for example), keep all records.

- **Cadmium:**
 1. If speciated Cadmium emissions are greater than or equal to Cadmium and Compounds (NEI pollutant code 125) emissions, delete Cadmium and Compounds record.
 2. If Cadmium and Compounds emissions are greater than speciated Cadmium emissions, delete Cadmium record.
 3. If emissions are speciated and no Cadmium and compounds emissions are reported (Cadmium and Cadmium oxide for example), keep all records.

- **Chromium:**
 - A. If Chromium VI (NEI pollutant code 18540299) and Chromium III (NEI pollutant code 16065831) emissions are greater than or equal to Chromium and Compounds (NEI pollutant code 136) emissions, delete Chromium and Compounds record.

$$\text{Cr VI} + \text{Cr III} \geq \text{Cr and Compounds}; \text{ delete Cr and Compounds}$$
 - B. If Chromium VI and Chromium III emissions are greater than or equal to Chromium (NEI pollutant code 7440473) emissions, delete Chromium record.

$$\text{Cr VI} + \text{Cr III} \geq \text{Cr}; \text{ Delete Cr}$$
 - C. If emissions are speciated and no Chromium and compounds emissions are reported (Chromium VI and Chromium III for example), keep all records.
 No revision
 - D. If Chromium and Chromium VI emissions are reported and Chromium VI emissions are less than Chromium emissions, generate Chromium III emission record by subtracting Chromium VI emissions from Chromium emissions. Delete Chromium record.

$$\text{Cr} - \text{CrVI} = \text{CrIII}; \text{ delete Cr and add CrIII}$$
 - E. If Chromium and Compounds and Chromium VI emissions are reported and Chromium VI emissions are less than Chromium and Compounds emissions, generate Chromium III emission record by subtracting Chromium VI emissions from Chromium and Compounds emissions. Delete Chromium and Compounds record.

$$\text{Cr and Compounds} - \text{CrVI} = \text{CrIII}; \text{ delete Cr and Compounds and add CrIII}$$

- F. If Chromium and Chromium III emissions are reported and Chromium III emissions are less than Chromium emissions, generate Chromium VI emission record by subtracting Chromium III emissions from Chromium emissions. Delete Chromium record.
 $Cr - CrIII = CrVI$; delete Cr and add CrVI
- G. If Chromium and Compounds and Chromium III emissions are reported and Chromium III emissions are less than Chromium and Compounds emissions, generate Chromium VI emission record by subtracting Chromium III emissions from Chromium and Compounds emissions. Delete Chromium and Compounds record.
 $Cr \text{ and Compounds} - CrIII = CrVI$; delete Cr and Compounds and add CrVI
- **Cobalt:**
 1. If speciated Cobalt emissions are greater than or equal to Cobalt and Compounds (NEI pollutant code 139) emissions, delete Cobalt and Compounds record.
 2. If Cobalt and Compounds emissions are greater than speciated Cobalt emissions, delete speciated Cobalt records.
 3. If emissions are speciated and no Cobalt and compounds emissions are reported (Cobalt and Cobalt oxide for example), keep all records.
 - **Cresols:**

In the 2002 NEI, data providers either reported the 3 isomers of Cresols (ortho-cresol, NEI pollutant code 95487; meta- cresol, NEI pollutant code 108394; and para-cresol, NEI pollutant code 106445) or they reported Cresols as a group (NEI pollutant code 1319773). The toxicity is the same for all isomers or cresol and cresylic acid.

 1. If speciated Cresol emissions are greater than or equal to Cresols emissions, delete Cresols records.
 2. If Cresols emissions are greater than individual isomers of Cresol emissions, delete individual isomer Cresol records. This may happen if Cresols are reported and only one or two isomers of cresols are reported.
 3. If emissions are speciated and no Cresols emissions are reported, keep all records.
 - **Cyanide:**
 1. If speciated Cyanide emissions are greater than or equal to Cyanide and Compounds (NEI pollutant code 144) emissions, delete Cyanide and Compounds record.
 2. If Cyanide and Compounds emissions are greater than speciated Cyanide emissions, delete speciated Cyanide records.
 3. If emissions are speciated (Cyanide and Hydrogen Cyanide for example), keep all records.

- **Glycol Ethers:**

In our analysis of point sources, glycol ethers emissions were always much greater than individual ethers reported for same emission release point records. We assumed data providers did not double count individual ethers and glycol ether group. Glycol ethers all have same toxicity which is not very high.

The procedure for reconciling glycol ethers emissions includes the following steps.

- A. If speciated glycol ether emissions are greater than or equal to Glycol Ethers (NEI pollutant code 171) emissions, delete Glycol Ethers records.
- B. If Glycol Ethers emissions are greater than speciated Glycol Ethers emissions, keep all records.
- C. If emissions are speciated and no Glycol Ethers emission are reported, keep all records.

- **Lead:**

1. If speciated Lead emissions are greater than or equal to Lead and Compounds (NEI pollutant code 195) emissions, delete Lead and Compounds record.
2. If Lead and Compounds emissions are greater than speciated Lead emissions, delete speciated Lead records.
3. If emissions are speciated and no Lead and compounds emissions are reported (Lead and Lead Oxide for example), keep all records

- **Manganese:**

1. If speciated Manganese emissions are greater than or equal to Manganese and Compounds (NEI pollutant code 198) emissions, delete Manganese and Compounds.
2. If Manganese and Compounds emission are greater than speciated Manganese emissions, delete speciated Manganese records.
3. If emissions are speciated and no Manganese and compounds emissions are reported(Manganese and Manganese Trioxide for example), keep all records.

- **Mercury:**

1. If speciated Mercury emissions are greater than or equal to Mercury and Compounds (NEI pollutant code 199) emissions, delete Mercury and Compounds record.
2. If Mercury and Compounds emissions are greater than speciated Mercury emissions, delete speciated Mercury records.
3. If emissions are speciated and no Mercury and Compounds emissions are reported (Mercury and Mercuric Oxide for example), keep all records.

4. If Mercury (NEI pollutant code 7439976) emissions, Mercury (organic) (NEI pollutant code 22967926) emissions and Mercury and Compounds emissions are all equal, delete Mercury (organic) emissions and Mercury and Compounds records.

- **Nickel:**

1. If speciated Nickel emissions are greater than or equal to Nickel and Compounds (NEI pollutant code 226) emissions, delete Nickel and Compounds record.
2. If Nickel and Compounds emissions are greater than speciated Nickel emissions, delete speciated Nickel records.
3. If emissions are speciated and no Nickel and compounds emissions are reported (Nickel and Nickel Sulfide for example), keep all records.

- **POM:**

In the 2002 NEI, data providers report POM for a variety of PAHs. Please refer to the NEI Pollutant code table for a complete list of POM compounds. The NEI Pollutant Code Table is contained in the 2002 State Lookup File, 02nei_lkup_states.zip, found at:

<ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> .

1. If speciated PAH emissions are greater than or equal to POM (NEI pollutant codes 130498292, 249, 246, 75, 40, or 284) emissions, delete POM records.
2. If POM emissions are greater than speciated PAH emissions, delete speciated PAH records.
3. If emissions are speciated and no POM emissions are reported (benzoapyrene and anthracene for example), keep all records.

- **Selenium:**

1. If speciated Selenium emissions are greater than or equal to Selenium and Compounds (NEI pollutant code 253) emissions, delete Selenium and Compounds record.
2. If Selenium and Compounds emissions are greater than speciated Selenium emissions, delete speciated Selenium records.
3. If emissions are speciated and no Selenium and compounds emissions are reported (Selenium and Selenium oxide for example), keep all records.

- **Xylenes:**

In 2002 NEI, data providers either reported emissions of the 3 isomers of Xylenes (ortho- xylene, NEI pollutant code 95476; meta-xylene, NEI pollutant code 108383; and para-xylene, NEI

pollutant code 106423) or they reported Xylenes as a group (NEI pollutant code 1330207). The toxicity is the same for all isomers or xylenes.

1. If individual Xylene isomer emissions are greater than or equal to Xylenes emissions, delete Xylenes records.
2. If Xylenes emissions are greater than the sum of individual isomers of Xylene emissions, delete isomers of Xylene records.
3. If emissions are speciated and no xylenes emissions are reported, keep all records.

When we prepare files for various uses, we will use the Pollutant Rank data field in the NEI Pollutant Code Table to select pollutants and emissions within a HAP Category where emissions of both individual compounds within a group and a compound group are reported for an emission record.

Any Chromium VI or Chromium III emission estimates that we have created will be assigned a data source code of "A-C".

12.8 Emissions Type Period

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.

The Emissions Type Period in the NOF EM Table distinguishes between annual and non-annual emission estimates in the 2002 NEI.

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 Emissions Type Period 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 Emissions Type Period is annual.
 - If the Emission Type is 30 and emissions are reported for a period of a season or less, then the Emissions Type Period 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 Emissions Type Period is assumed to be annual.
 - If the Emission Type is 20 - 29, the Emissions Type Period 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 Emissions Type Period is annual for this record and the Emissions Type Period 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 Emissions Type Period 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 Emissions Type Period 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 use the methodology in Table 23 to calculate the Annual Emissions. 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), they also supplied an annual emission estimate. However, if only a partial year estimate is supplied, it will be retained in the annual file and expressed as annual emissions. Retaining this partial year estimate, prevents the loss of the emissions data from the annual file.

Table 23. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Annual	30	Use reported emissions

Table 23, continued. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Non-Annual	30	<p>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.</p> <p>Annual Emissions = Sum of Numeric Emissions Value of each non-overlapping period</p> <p>Ex. Emissions are reported for each of the 12 months, Annual Emissions = Sum of monthly emissions</p>
Non-Annual	30	<p>Start Date and End Dates in Period are Not 20020101 and 20021231 Data provider indicates Non-Annual Emissions are not the Annual Emissions, but can be scaled to the entire year:</p> <ul style="list-style-type: none"> ▶ Start Time = 0000 and End Time = 2359 or Start and End Times are missing <p>Annual emissions = Emissions Numeric Value * (365/ Number of Days in Period)</p> <ul style="list-style-type: none"> ▶ Start and End Times are for less than a 24 hr period <p>Annual Emissions = Numeric Emissions * (365/Number of Days in Period) * (Hrs/Day)</p> <p>If the data provider cannot be reached, set Annual emissions = Emissions Numeric Value.</p>
Non-Annual	29	<p>Data provider indicates Non-Annual Emissions are the same as Annual Emissions.</p> <ul style="list-style-type: none"> ▶ Start Time = 0000 and End Time = 2359 or Start and End Times are missing <p>Annual Emissions = Numeric Emissions * (Number of Days/Period)</p> <ul style="list-style-type: none"> ▶ Start and End Times are for less than a 24 hr period <p>Annual Emissions = Numeric Emissions * Number of Days/Period) * (Hrs/Day)</p>

Table 23 continued. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Non-Annual	29	<p>Start Date and End Dates in Period are Not 20020101 and 20021231 Data provider indicates Non-Annual Emissions are not the Annual Emissions, but can be scaled to the entire year:</p> <ul style="list-style-type: none"> ▶ 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) <p>Ex. Emissions are for 20020501 - 20020831 for 24 hr/day Annual Emissions == Emissions (365/123)</p> <p>If the data provider cannot be reached, set Annual Emissions = Numeric Emissions * (Number of Days/Period).</p>
Non-Annual	27 (average weekday)	<p>Data provider indicates Non-Annual Emissions are the same as Annual Emissions.</p> <ul style="list-style-type: none"> ▶ Start Time = 0000 and End Time = 2359 or Start and End Times are missing Annual Emissions = Numeric Emissions * (Number of Days/Period) * 5/7 ▶ Start and End Times are for less than a 24 hr period Annual Emissions = Numeric Emissions * Number of Days/Period) * (Hrs/Day) * 5/7

Table 23 continued. Methodology Used to Generate Annual Emissions

Period Type	Emission Type	Methodology
Non-Annual	28 (average weekend)	<p>Data provider indicates Non-Annual Emissions are the same as Annual Emissions.</p> <ul style="list-style-type: none"> ▶ Start Time = 0000 and End Time = 2359 or Start and End Times are missing <p>Annual Emissions = Numeric Emissions * (Number of Days/Period) * 2/7</p> <ul style="list-style-type: none"> ▶ Start and End Times are for less than a 24 hr period <p>Annual Emissions = Numeric Emissions * Number of Days/Period) * (Hrs/Day) * 2/7</p>

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

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.

12.11 Emission Ton Value

We will convert emissions to tons if reported in other units and report the ton value in the Emission Ton Value field in the NOF EM Table. Table 24 shows the conversions factors that we will use to generate emissions in tons.

Table 24. Unit Conversions for Tons

Reported Unit	Conversion Factors for Tons
E3TON	1000
MILLIGRAM	1.1025E-09

Reported Unit	Conversion Factors for Tons
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

12.12 Emission Content QC

After we have compiled a draft point source inventory, we will conduct a variety of emissions content QC on an internal draft inventory prior to preparing the draft inventory files for external review. The purpose of Emissions Content QA/QC is to identify and correct erroneous emissions and add missing emissions. We will identify potential problems including:

- Outliers
- Missing facilities
- Missing pollutants from facilities
- Pollutants that should not be reported for facilities

Our Emissions Content QA/QC will consist of both Comparative and Statistical QC. The Comparative QC includes the evaluation of significant changes between the 1999 NEI and 2002 NEI data, and/or extreme variation within the 2002 data. We will compare 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. Statistical QC includes using statistical tables to identify 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. Comparative and Statistical QC will be used to identify outliers, missing pollutants from facilities and erroneous pollutants reported for facilities. Comparative QC will be used to identify missing facilities.

Specific Comparative QC activities will include:

- “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;
- Previous Year Emissions Comparison - compare pollutant emissions nationally and for state, county, source category and facility from 1999 NEI with the draft 2002 NEI; and

- 99-2002 Site List - provides a list of sites found in the 1999 base year NEI, sites in the 2002 draft NEI, and sites common to both the 1999 and 2002 inventories.

Specific Statistical QC activities will include:

- 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;
- Contribution of facilities and source categories to multiple pollutants - compare list of pollutants emitted from a source category to individual facilities; and
- Pollutant/Source Category Crosswalk Comparison - compare emissions of pollutants to Pollutant/Source Category crosswalks generated from the 1999 NEI..

Outliers are often difficult to spot - what may appear to be a high emissions value may in fact be acceptable for a particular facility or source category. To aid in detecting these errors, we will use both comparative and statistical QC activities. We will compare emissions data to the range of values in the draft 2002 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.

To identify missing facilities, we will use the 99-2002 Site List. By sorting this list by state, county, and facility name, we will evaluate the sites listed in each county and detect missing facilities. If a site is found in the 1999 version but is not in the 2002 draft, reviewers can verify that the facility closed or supply data for the missing facility during the draft review period.

We will use identify missing and erroneous pollutants for individual facilities using the pollutant/source category crosswalk tables and statistical tables. We prepared pollutant/source category crosswalk tables to QA pollutants emitted from source categories. For CAPs, the crosswalk identifies source categories by SCC. For HAPs, the crosswalk identifies source categories by MACT Code, SIC Code, and SCC. We will explore expanding the crosswalk table to include emissions ranges. The two pollutant source category crosswalk files, HAP Source Category Crosswalk and CAP Source Category Crosswalk, are available at:

ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/augmentation_point_sector/ . The HAP Source Category Crosswalk contains 3 tables: Facility MACT, Facility SCC, and Facility SIC. The CAP Source Category Crosswalk contains 1 table - Facility SCC. The following data fields are included in the pollutant/source category crosswalk tables.

- Source Category - MACT Code, SCC, or SIC Code
- Pollutant - HAP Category Name for HAPS or Pollutant Name for CAPs

- Number of Facilities - number of facilities reporting emissions of pollutant for a source category
- Number of Total Facilities - number of total facilities for a source category
- % Facilities - % of Number Facilities reporting emissions of pollutant/Total Number of Facilities for a source category
- Max of Emissions(TPY) - maximum emissions for facility emitting a pollutant within the category
- Min of Emissions(TPY) - minimum emissions for facility emitting a pollutant within the category
- 99 Total Emissions(TPY) - total emissions for pollutant from source category

Emission Content QC Report

We will prepare a Emission Content QC report for each state and tribe. The report will compare 2002 emissions data to 1999 emissions data and highlight potential outliers, missing facilities, missing pollutants and erroneous pollutants. We will send these reports to the data providers and also post the reports during the review period of the draft 2002 NEI. We will ask data providers to review the reports and provide revisions to us. The following tables will be included in the state emissions QC files.

- Missing Facility Report - facilities present in 1999 NEI but missing from 2002 NEI
- Missing County Report - counties with facilities present in 1999 NEI but missing from 2002 NEI
- Counties with No Data - No facilities were present in these counties in 1999 NEI and 2002 NEI
- Missing CAPs at State Level - CAPs missing from 2002 state inventory
- Missing Facility CAPs- CAPs present for these facilities in 1999 NEI but missing from 2002 NEI
- 1999 and 2002 Facility CAP Comparison - comparison of emissions at facility level in 1999 NEI and 2002 NEI
- 1999 and 2002 Facility HAP Comparison - comparison of emissions at facility level in 1999 NEI and 2002 NEI
- 1999 and 2002 State CAP Comparison - comparison of emissions at state level in 1999 NEI and 2002 NEI
- 1999 and 2002 State HAP Comparison - comparison of emissions at state level in 1999 NEI and 2002 NEI
- State Facility CAP Counts - Comparison of number of facilities emitting CAPs in 1999 NEI and 2002 NEI
- State Facility HAP Counts - Comparison of number of facilities emitting HAPs in 1999 NEI and 2002 NEI
- Emissions Percentage CAPs - Facilities where % of total state emissions of a pollutant (facility emissions/state point source emissions) are greater than 25%
- Emissions Percentage HAPs - Facilities where % of total state emissions of a pollutant (facility emissions/state point source emissions) are greater than 25%
- Top 10 Facilities of CAPs
- Top 10 Facilities of HAPs

- Data from the 2002 NEI Facility File - to assist in identifying locations and facility IDs

Other Summary Files to Assist in Emission Content QC

The following emission summary files will also be compiled from the draft 2002 NEI to help facilitate QC of the NEI, and to help put the emission estimates into perspective by state, county, source category, and facility.

- All NEI
 - HAPs, Annual
 - CAPs Annual
 - CAPs NonAnnual
- NEON HAP Summary Tables
- NEI Emissions History Table (See Section 12.15)

For NEI point source data users who are interested in more than one state data set, three ALL NEI summary files will be available that combine key data elements from all of the state and tribal data sets in a single database.

- ALL_NEI_CAP_ANNUAL - This file will contain annual emissions data for the 2002 NEI for CAPs. Emissions data will be presented in tons/year. However, records will retain their original start/end dates and emission types for reference.
- ALL_NEI_CAP_NONANNUAL - This file will contain nonannual emissions data for the 2002 NEI for CAPs. This file will contain the non-annual emissions records supplied by state and local agencies and tribes. The emissions will be presented in their original units and will not be converted to tons.
- ALL_NEI_HAP_ANNUAL - This file will contain annual emissions data for the 2002 NEI for HAPs. Emissions data will be presented in tons/year. However, records will retain their original start/end dates and emission types for reference.

Table 25 contains the table structure for the ALL NEI files.

Table 25. Table Structure for ALL NEI

Name	CAP	HAP
POP_FLAG	X	X
EPARegion	X	X
strTribalCode	X	X
Tribal_Name	X	X
ST_ABBR	X	X
Cnty_Name	X	X
strStateCountyFIPS	X	X

Name	CAP	HAP
strFacilityRegistryIdentifier	X	X
strStateFacilityIdentifier	X	X
strNEI Unique Facility ID	X	X
strFacilityName	X	X
strSICPrimary	X	X
strNAICSPrimary	X	X
ergNAICSflag	X	X
strEmissionUnitID	X	X
strProcessID	X	X
strSCC	X	X
strEmissionReleasePointID	X	X
strEmissionReleasePointType	X	X
lngStartDate	X	X
lngEndDate	X	X
strEmissionType	X	X
strPollutantCode	X	X
Pollutant_Code_Desc	X	X
CAP_Category_Name	X	
HAP_Category_Name		X
Emissions(TPY)	X	X
Emissions(TPY)-TEXT	X	X
strControlStatus	X	X
strLocationAddress	X	X
strCity	X	X
strState	X	X
strZipCode	X	X
sngStackHeight	X	X
sngExitGasTemperature	X	X
sngStackDiameter	X	X
sngExitGasVelocity	X	X

Table 25 continued. Table Structure for ALL NEI

Name	CAP	HAP
sngExitGasFlowRate	X	X
ergStackDefaultFlag	X	X
ergDataSource	X	X
strHAP EmissionsPerformanceLevel		X
strCAPEmissionsPerformanceLevel	X	
dblXCoordinate	X	X
dblYCoordinate	X	X
strHorizontalCollectionMethodCode	X	X
strHorizontalAccuracyMeasure	X	X
strHorizontalReferenceDatumCode	X	X
strReferencePointCode	X	X
strSourceMapScaleNumber	X	X
strCoordinateDataSourceCode	X	X
ergLocationDefaultFlag	X	X

The NEON HAP files contain emission summaries for each 188 HAP category, the sum of the 188 HAPs, and the sum of the 33 urban HAPs. Emissions are also presented for each individual HAP species in all files except for the county emission summary, the source category summary and the point source facility summary files. The data structure for the NEON summary files is in the file, neon_table_specs092003.xls, and is available at:

<ftp://ftp.epa.gov/pub/EmisInventory/finalnei99ver3/haps/summaries/>. The 7 NEON HAP files include the following summaries.

- **County Emission Summary** - The county emission summary presents the NEI HAP emissions by state, and county for major, area, onroad, and nonroad sources. Major and area sources are also summarized as MACT vs. non-MACT source categories.
- **Source Category Summary** - The source category summary presents the NEI HAP emissions by state, and county for major, area, onroad, and nonroad sources. The area sources are delineated as point or nonpoint. Each stationary source category is presented by MACT code, SIC Code, or just source category name if there is no applicable MACT or SIC code.
- **Point Source Facility Summary** - The point source facility summary presents the NEI HAP point source emissions by NEI Unique Facility ID and individual site ID for major and area

point sources. Included with each facility record is the address, site latitude/ longitude, emission type (actual, allowable, potential, etc.), MACT and/or SIC code, the source of the emission estimate.

- **Point Source Stack Summary** - The point source stack summary presents the NEI HAP emissions by NEI Unique Facility and individual site for major and area point sources. Included with each record is the emission type (actual, allowable, potential, etc.), emission unit ID, process ID, emission release point ID, SCC, MACT and/or SIC code, emission release point type (stack/vent or fugitive), and latitude/longitude of the emission release point and the source of the emission estimate.
- **Nonpoint Stationary Source Summary** - The nonpoint stationary source summary presents the NEI HAP emissions by state, county, and area source category. Included with each record is the emission type (actual, allowable, potential, etc.), SCC, MACT, and/or SIC code.
- **Nonroad Summary** - The nonroad summary presents the NEI HAP emissions by state, county, and aggregated nonroad source categories.
- **Onroad Summary** - The onroad summary presents the NEI HAP emissions by state, county, and aggregated nonroad source categories.

Toxicity Weighting

Toxicity weighting of HAP emissions is a useful tool to identify potential outliers and problems with HAP emissions. Emissions of HAPs can be toxicity weighted for cancer and noncancer effects. The National Air Toxics Assessment contains more information on risk characterization at the following address. <http://www.epa.gov/ttn/atw/nata1999/riskbg.html> .

Exposure modeling should be conducted to provide an accurate risk characterization. Toxicity weighting of emissions provides a quick and less accurate relative ranking of emissions by risks. Toxicity weighting does not account for dispersion and transport of HAPs, secondary formation of HAPs, or exposure within micro-environments to HAPs. Because toxicity weighting does not have units, the results are dimensionless and only a relative ranking of pollutants is possible using this approach. We toxicity weighted draft NEI HAP emissions to identify facilities with the largest risk in order to conduct additional QC.

The methodology to toxicity weight emissions data includes the following steps.

1. **For all individual NEI pollutants except Chromium and Compounds (pollutant code136) and Chromium (pollutant code 7440473):** Multiply emissions by Metal_CN Speciation Factor to extract metal and cyanide mass for toxicity weighting. The Metal_CN Speciation Factors are found in the file, "Tox wt factors for 2002 NEI 06282006", available at: <ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> . *Note:* Some NEI HAPs are speciated into more than one metal or into a metal and cyanide (CN). The

Metal_CN Speciation factors were developed by using the molecular weight of the pollutants to extract only the mass of metal element or HCN from compounds. The toxicity weighting factors for metals are assigned to elemental metal and the toxicity weighting factors for cyanides are assigned to hydrogen cyanide (HCN). NEI estimates for metals must first be multiplied by the Metal_CN Speciation Factor to provide emission estimates for toxicity weighting. For all non-metal and non-cyanide compounds, the metal speciation factor is 1. In addition, the table groups NEI POM compounds into the 8 POM groups that have toxicity factors.

2. **For NEI pollutants Chromium and Compounds (pollutant code136) and Chromium (pollutant code 7440473):** Use chromium speciation file, "Cr Speciation 06282006.mdb" available at: <ftp://ftp.epa.gov/EmisInventory/2002finalnei/documentation/point/> to speciate source category emissions into hexavalent chromium (Cr(VI)) and trivalent chromium (Cr(III)) emissions. Cr(VI) emissions are generated by multiplying Chromium and Compounds (pollutant code136) and Chromium (pollutant code 7440473) by the Hexavalent Chromium Speciation Multiplier; and Cr(III) emissions are generated by multiplying Chromium and Compounds (pollutant code136) and Chromium (pollutant code 7440473) by the Trivalent Chromium Speciation Multiplier. Emissions by source category are needed in order to use chromium speciation file. In order to toxicity weight chromium emissions, all emissions must be allocated to Cr(VI) and Cr(III). The following hierarchy is used to assign default speciation profiles to chromium source categories.
 - A. Use MACT Code default speciation profiles first
 - B. Use SCC speciation profiles if MACT code default speciation profiles are unavailable. NOTE nonroad SCCs are truncated in table.
 - C. Use SIC Code speciation profiles if MACT code and SCC default speciation profiles are unavailable.
 - D. Use default speciation multiplier of 0.34 for Cr(VI) and 0.66 for Cr(III) if MACT Code, SCC and SIC Code default speciation profiles are unavailable.
3. **For Cancer Tox weighting:** Multiply emissions generated from steps 1 and 2 by the Unit Risk Estimate (URE) found in the file, "Tox wt factors for 2002 NEI 06282006".
NOTE: Cancer risks between pollutants are additive. Cancer and NonCancer risks are not additive.
4. **For NonCancer Tox weighting:** Divide emissions generated from steps 1 and 2 by the Reference Concentration (RfC) found in the file, "Tox wt factors for 2002 NEI 06282006".
NOTE: NonCancer risks between pollutants are additive. Cancer and NonCancer risks are not additive.

An example toxicity weighted emissions calculation is shown below..

NEI: 1 ton of lead arsenate is emitted by a facility.

Metal_CN Speciation Factors: 0.2146 Arsenic and 0.5935 lead
Lead emissions for toxicity weighting: 1 ton x 0.5935 = 0.5935 tons lead
Arsenic emissions for toxicity weighting: 1 ton x 0.2146 = 0.2146 tons arsenic

UREs: Arsenic 0.0043
RfCs: Arsenic 0.00003; Lead 0.0015

Cancer Risk: Arsenic: $0.2146 \times 0.0043 = 0.0092278$
Noncancer Risk: Arsenic: $0.2146/0.00003 = 7153.3$
Lead: $0.5935/0.0015 = 395.6$

From this example, while the arsenic emission are smaller than lead, the toxicity weighted emissions results in greater risks for arsenic than lead.

12.13 Emissions Data Quality Rating for HAPs

The Data Rating will help to provide users of the NEI some sense of the reliability of an emission estimate. We developed a simplified rating scheme so that a qualitative grade from A to E can be assigned to each HAP estimate. This enhances the transparency of the data and also satisfies the requirement that EPA perform an assessment of the data. The scoring may help us determine which data point to retain in the final 2002 NEI when there are multiple estimates for the same facility. It may also help us to 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.

The Emissions Data Quality Rating is based on the following factors..

- Completeness of Data
- Emission Calculation Methods
- Age of Data
- Emission Data Level
- Specificity of Data
- SCC

To evaluate these six inventory factors, we will first extract the following data fields from NOF to make a single table:

State County FIPS
Tribal Code
State Facility Identifier
Emission Unit ID
Process ID
Pollutant Code
HAP Category Name
Emission Release Point ID

Start Date
End Date
Emission Type
Data Source
Emissions (TPY)
Control Status
EM Reliability Indicator
Factor Numeric Value
Factor Unit Numerator
Factor Unit Denominator
Emission Calculation Method Code
EF Reliability Indicator
Emission Data Level
SCC
Actual Throughput
Throughput Unit Numerator

The methodology for assigning a rating is described below.

- ***Completeness of Data***

This factor addresses whether 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.

A maximum of 10 points could be earned for completeness of data. If the following fields are not null then 10 points will be awarded. The sum of the points earned is divided by 5 to obtain a completeness of data number.

- Actual Throughput & Throughput Unit Numerator (both must be non-null)
- EM Reliability Indicator
- Factor Numeric Value, Factor Unit Numerator, Factor Unit Denominator (all 3 must be non-null)
- EF Reliability Indicator
- Control Status (only uncontrolled and controlled are given points)

- ***Emission Calculation Method Code***

This factor addresses the quality of the emission estimate based on estimation method. For example, estimates based on continuous monitoring receive higher scores than data based on less accurate methods, such as “engineering judgment.”

The emissions calculation method code determines the amount of points used in the rating. The maximum number of points given for emission calculation method code is 10. If the field is null or invalid, no points are given.

Emission Calculation Method Code	Emission Calculation Method Description	Points Given
01	CEMS - Continuing Monitoring System	10
02	Engineering Judgment	2
03	Material Balance	5
04	Stack Test	8
05	EPA Speciation Profile	4
06	State/Local/Tribal Speciation Profile	4
07	Manufacturer Specification	4
08	EPA Emission Factor	5
09	State/Local/Tribal Emission Factor	5
10	Site-Specific Emission Factor	5
11	Vendor Emission Factor	5
12	Trade Group Emission Factor	5

- ***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. This factor accounts for the age of the estimate.

A maximum of 5 points will be given for the age of data. If the data has a start date or end date of 2002, then 5 points will be given. Any other year will be given zero.

- ***Emission Data Level***

Emission Data Level is very important for modeling. An estimate which provides process level emissions is better than aggregated facility level HAP emissions. A maximum of 5 points will be given for the emission data level.

Emission Data Level	Points Given
Process	5
Stack	3
Unit	3
Site	1

- ***Specificity of Data***

This factor addresses pollutant groups. Reporting of individual HAPs is preferred to reporting HAP compound groups such as arsenic and compounds.

A maximum of 5 points will be given for the specificity of the data. If the pollutant is ranked greater than 1 in the pollutant code table, no points will be given. All others will receive 5 points.

- ***SCC***

SCCs are very important for modeling. This factor addresses whether SCCs are present.

A maximum of 5 points will be given for the SCC. If the SCC is not null and not 39999999, then 5 points will be given. All others will receive zero points.

- ***Composite Score and Grading***

All of the different scores will be added together and each record will receive the following grades based on the score:

Score	Grade
33-40	A
25-32	B
17-24	C
9-16	D
0-8	F

We will assign the Data Rating for all HAP point source records in the final 2002 NEI.

12.14 Data Source

The Data Source in the NOF Emissions Table indicates the source of the reported emission value. The following codes will be used.

767	EGU estimates based on 2002 Department of Energy (DOE) Form EIA-767 information (utility related data)
767/CAMD	EGU record in both 2002 DOE Form EIA-767 and 2002 CAMD ETS/CEM; with SO ₂ , NO _x , and heat input values from CAMD ETS/CEM; condensable PM, and primary PM ₁₀ and PM _{2.5} are recalculated using EPA's CAMD heat input values
767/CAMD1	EGU Combined cycle record (heat recovery steam generators (HRSG) + combustion turbine) in 2002 CAMD ETS/CEM used; the HRSG record in 2002 DOE Form EIA-767 eliminated. SO ₂ , NO _x , and heat input values from CAMD ETS/CEM; other emissions estimated
767/CAMD2	EGU Combined cycle record (steam turbine + combustion turbine) in 2002 ETS/CEM used; the steam record in 2002 Form EIA-767 eliminated. SO ₂ , NO _x , and heat input values from ETS/CEM; other emissions estimated.
A	Augmented PM data

A-B	Augmented boiler data
A-C	Augmented chromium Data
A-R	Augmented PM data replaced existing PM data submitted to NEI
B	Tribal data
BCAMD1	EGU data were received from the tribal agency. The tribal agency's NOX and SO2 emission values were replaced with the ETS values.
C-01-F	Wyoming did not have any PM and NOX data in their submittal. Data was brought forward from the 2001 inventory.
CAMD	EGU record only in CAMD 2002 ETS/CEM for SO2, NOx, and heat input values; other emissions estimated.
CAMD1	EGU data was not received from the state, local agency, or tribal agency. The ETS records were used.
I	Industry data
L	Local agency data
LCAMD1	EGU data were received from the local agency. The local agency's NOX and SO2 emission values were replaced with the ETS values.
LCAMD3	EGU data were received from the local agency. The local agency's NOX and SO2 value was used where the ETS value was zero.
M	MACT, Residual Risk, or Area Source data provided by ESD
N	99 NEI
P	Preferred MACT, Residual Risk, or Area Source data provided by ESD
PCAMDHG	EGU Mercury data . No data were received from state and local agencies and tribes. The ESD mercury values were given to the ETS record.
PI	Preferred industry data.
PLHG	EGU Mercury data were received from the local agency. The local agency provided a total mercury record. The record was replaced with three individual mercury species records from ESD.
PSHG	EGU Mercury data were received from the state. The state provided a total mercury record. The record was replaced with three individual mercury species records from ESD.
R	Regional Planning Organization data
S	State agency data
SCAMD1	EGU data were received from the state. The state's NOX and SO2 emission values were replaced with the ETS values.
SCAMD2	EGU Data were received from the state. The state did not provide NOX, SO2, or either. NOX and/or SO2 records were created with the same IDs and the ETS values were used.
SCAMD3	EGU data were received from the state. The state's NOX and SO2 value was used where the ETS value was zero.
SCAMD4	EGU Data were received from the state. It was not clear which of the state's units matched to the ETS units. The state values were maintained, because the emissions were close to the ETS values.
T	2002 Toxics Release Inventory data

12.15 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
- Pollutant Code
- Pollutant Name
- HAP Category Name
- Original P (tpy) - Emissions supplied for Preferred category by ESD for MACT, Residual Risk or Area Source
- Original S (tpy) - Emissions supplied by State
- Original L (tpy) - Emissions supplied by Local Agency
- Original B (tpy) - Emissions supplied by Tribe
- Original R (tpy) - Emissions supplied by Regional Planning Organization
- Original T (tpy) - Emissions supplied by TRI
- Original M (tpy) - Emissions supplied by ESD for MACT, Residual Risk or Area Source
- OrigCAMD_NS (tpy) - Emissions supplied by CAMD for NOX and SO2 for EGUs
- OrigE (tpy) - Emissions for EGUS for all other pollutants except NOX and SO2 provided by EPA
- OrigA (tpy) - Emissions augmented by EPA
- 02012005 Draft Emissions (tpy)
- 02012005 Draft Data Source
- 02012006 Final Emissions (tpy)
- 02012005 Final Data Source
- Revised Version Date - date of future revisions
- Revised Emissions Value (tpy) - emissions value of future revisions
- Revised Data Source - data source of future revisions
- Emissions Change (tpy) - difference between revised and current emissions
- Comments

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 will 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.

The 2002 NEI Emissions History Table will be located in a zipped file, nei2002_point_facility_emissions_history, at:
ftp://ftp.epa.gov/EmisInventory/2002finalnei/point_sector_data/ .

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

For all fields of data that are defaulted, a Default Flag will be added to each NOF 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".

United States
Environmental Protection
Agency

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

Publication No. EPA-454/R-20-001
February 2006
