APPENDIX F: NEI QA/QC PROCESS

I. Quality Checks for File Format

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

<u>Steps 1 - 3</u>

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

POINT SOURCES

A Data format and field widths

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

B Mandatory fields must be filled

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

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

C Compare with State provided Submittal Form

I. geographic coverage

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

ii. pollutant coverage

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

iii. source category coverage

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

iv. temporal coverage

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

D Referential Integrity

The NEI is a relational database. The following one to many relationships between files/tables must be created and enforced. For example, the relationship between the TR file/table and the SI file/table is a one to many relationship, i.e. there is one TR record for many SI records or one unique county record for possibly many sites. A relationship is created unless one of the following two errors occurs:
1- orphan records (For example, a record in the site table exists without a record in the TR table for that county).
2- records not unique for key fields (A county was listed twice in the TR table).

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

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

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

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

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

Note: File/Table Abbreviations:

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

II. Quality Checks for Data Content

Steps 1 - 3

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

A Acceptable Codes

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

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

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

B Numeric Values In Acceptable Range

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

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

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

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

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

III. Final QC

Repeat steps in I-II above.

I. Methodology For Blending / Merging Records

Point Sources

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

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

*Note: Facility-specific MACT inventory information for two MACT source categories - electric utility mercury emissions and HAP emissions from municipal waste combustors will be used directly and will not be replaced with S/L data because the MACT data sets are highly accurate and the result of extensive source testing. Likewise, data from the Acid Rain CEM data base for electric utility NOx and SO2 emissions will be used directly and will not be replaced with S/L data because of the high level of accuracy.

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

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

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

- Entire counties missing, and
- Missing facilities.

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

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

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

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

Area Sources

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

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

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

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

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

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

Mobile Sources

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

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

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

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

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

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

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

II. Solve for Missing Data Elements

Point Sources

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

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

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

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

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

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

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

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

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

1. Facility lists provided by MACT engineers

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

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

Table B-1: Point Source Data Completion Methods

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

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

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

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

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

Emissions (Criteria)

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

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

B. Perform plant level match to current NET data
If there is a match, distribute plant level emissions to segments in New (S/L submitted) NEI data based on segment level distribution of known pollutant (NOx) in New NEI Data

C. If there is no plant level match to current NET data
Develop emissions using uncontrolled emission factor ratios to calculate uncontrolled emissions. This method applies SO2 or PM-10 ratios to NOx. NOx was the pollutant

utilized to calculate the ratio because (1) the types of sources likely to be important SO2 and PM-10 emitters are likely to be similar to important NOx sources and (2) the generally high quality of the NOx emissions data. Ratios of SO2/NOx and PM-10/NOx based on uncontrolled emission factors were developed. These ratios were multiplied by uncontrolled NOx emissions to determine either uncontrolled SO2 or PM-10 emissions.

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

All NH3 emission estimates remained unchanged from the estimates currently in the NET.

2. Solve for Temporal Emissions

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

To calculate daily from annual -

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

To calculate annual from daily -

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

Where:

$EMIS_{ANN}$ = Annual Emissions
$EMIS_{ASD}$ = Typical Summer Day Emissions
SUMTHRU = Summer Throughput Percentage
DPW = Days Per Week in Operation
13 = Number of Weeks in Summer

Area Sources

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

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

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

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

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

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

Solving for Missing Data - Emissions

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

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

B. Perform State/county/SCC level match to current NET data
If there is a match use emissions from current NET.

C. If there is State/county/SCC match to current NET data
Develop emissions using uncontrolled emission factor ratios to calculate uncontrolled emissions. This method used SO2 or PM-10 ratios to NOx. NOx was the pollutant utilized to calculate the ratio because (1) the types of sources likely to be important SO2 and PM-10 emitters are likely to be similar to important NOx sources and (2) the generally high quality of the NOx emissions data. Ratios of SO2/NOx and PM-10/NOx based on uncontrolled emission factors were developed. These ratios were multiplied by uncontrolled NOx emissions to determine either uncontrolled SO2 or PM-10 emissions.

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

All NH3 emission estimates remained unchanged from the estimates currently in the NET.

2. Solve for Temporal Emissions

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

To calculate daily from annual -

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

To calculate annual from daily -

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

Where:

$EMIS_{ANN}$ = Annual Emissions
$EMIS_{ASD}$ = Typical Summer Day Emissions
SUMFAC = Summer Season Factor from TAFF (by SCC)
WKDYFAC = Summer Weekday Factor from TAFF (by SCC)

Table B-3: Mobile Source Onroad Data Completion Methods

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

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

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