

The Development of the 1999 National Emissions Inventory for HAPs

Anne Pope

Emission Factor and Inventory Group, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711

Richard Billings, Stephanie Finn and Darcy Wilson

Eastern Research Group, Inc., P.O. Box 2010, 1600 Perimeter Park, Morrisville, NC 27560

ABSTRACT

Requirements of the Clean Air Act (CAA) and Government Performance Results Act (GPRA) have established the need for a more comprehensive hazardous air pollutant (HAP) emissions inventory effort that can be used to track progress by the Environmental Protection Agency (EPA) over time in reducing HAPs in ambient air. To estimate risk and HAP emission reductions, the EPA compiles the National Emission Inventory (NEI) for HAPs to provide a model-ready emissions inventory. The EPA has previously compiled a baseline 1990 and 1996 National Toxics Inventory (predecessor to NEI for HAPs) and is now completing the development of the 1999 NEI for HAPs.

The NEI contains estimates of facility-specific HAP 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 and mobile source categories. Point source categories include major and area sources as defined in section 112 of the CAA. Nonpoint source categories include area sources and other stationary sources that may be more appropriately addressed by other programs rather than through regulations developed under sections 112 or 129 in the CAA.

The sources of data in the NEI for HAPs are:

- State and local agency data and tribal data,
- Industry data,
- Data gathered by the EPA's Emission Standards Division while developing Maximum Achievable Control Technology (MACT) standards,
- Toxic Release Inventory (TRI) data,
- Data developed by the EPA's Office of Transportation and Air Quality,
- Estimated nonpoint emissions for sources not included in the state and local agency or tribal data, and
- 1996 National Toxics Inventory data for facilities and categories not included in the sources of data above.

The compilation of the 1999 NEI for HAPs requires many steps including the following key processing activities.

- Submittal of 1999 HAP inventory data by state and local agencies, and tribes;
- Blending/Merging of data from multiple data sources;
- Augmentation of blended data for missing data elements;
- Quality Control/Quality Assurance (QC/QA) of the data;
- Preparation of draft NEI for HAPs for external and internal review;
- Incorporation of external and internal review comments; and
- Preparation of final NEI for HAPs.

This paper discusses the steps in compiling of the 1999 NEI for HAPs.

INTRODUCTION

Requirements of the Clean Air Act (CAA) and Government Performance Results Act (GPRA) have established the need for a more comprehensive hazardous air pollutant (HAP) emissions inventory effort that can be used to track progress by the Environmental Protection Agency (EPA) over time in reducing HAPs in ambient air. To estimate risk and HAP emission reductions, the EPA compiles the National Emission Inventory (NEI) for HAPs to provide a model-ready emissions inventory. The EPA has previously compiled a baseline 1990 and 1996 National Toxics Inventory (NTI), predecessor to NEI for HAPs, and is now completing the development of the 1999 NEI for HAPs. Recently, the EPA used the 1996 NTI in the 1996 National Scale Air Toxic Assessment (NSATA) and in residual risk analyses required under section 112(f) of the CAA. The 1996 NTI was also used as the starting point for some state and local agencies that wished to maintain their own HAP emission inventories.

The EPA is compiling a 1999 HAP emission inventory as part of its 1999 NEI development process. The 1999 NEI contains both HAP and criteria emissions data. EPA and state and local agencies will use the 1999 NEI for HAPs to assess progress in reducing HAP emissions and exposure to HAPs. Specific uses of the 1999 NEI for HAPs include residual risk and 1999 NSATA. The NEI will be updated every three years (1999, 2002, 2005, etc.)

The various CAA and GPRA needs for air toxics emission data include estimates of emissions of the 188 HAPs at the national, regional, county, facility and process levels. The NEI contains estimates of facility-specific HAP 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 and mobile source categories. Point source categories include major and area sources as defined in section 112 of the CAA. Nonpoint source categories include area sources and other stationary sources that may be more appropriately addressed by other programs rather than through regulations developed under sections 112 or 129 in the CAA. The NEI is thus designed to provide a model-ready emission inventory of anthropogenic sources of the 188 HAPs listed in the CAA in order to facilitate comprehensive dispersion and exposure modeling.

To meet the various needs of CAA and GPRA programs, the EPA established the following goals for developing the 1999 NEI for HAPs:

- Produce NEI for HAPs in timely manner;

- Use standardized format for data inputs and outputs;
- Consistently use the best available data;
- Improve Quality Assurance/Quality Control (QA/QC) of the inventory;
- Reduce the number of stationary nonpoint source categories by increasing the number of categories inventoried as point sources; and
- Integrate HAP and criteria emissions for source categories where feasible.

To compile a HAP inventory that meets these goals, the EPA encourages the participation of state and local agencies, tribes, industry and the public in the submittal of inventory data and in the review of the draft inventory. This paper discusses the steps in compiling of the 1999 NEI for HAPs and provides information on how the EPA's inventory partners can participate in the development of the NEI for HAPs to improve the quality of the inventory.

NEI SOURCE CATEGORY AND HAP DEFINITION

In order to provide model ready inventory to support CAA and GPRA analyses, NEI emissions need to be clearly associated with individual HAPs and source categories.

Reporting of HAP Groups

HAPs are generally defined as those pollutants that are known or suspected to cause serious health problems. Section 112(b) of the CAA currently identifies a list of 188 pollutants as HAPs. (www.epa.gov/ttn/uatw/pollsour.html) EPA's Unified Air Toxics Web Site (UATW) presents more information on HAPs, their effects, and EPA's programs to reduce HAPs. (www.epa.gov/ttn/uatw/basicfac.html)

In addition to numerous specific chemical species and compounds, the list of 188 HAPs includes seventeen 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 for the groups. The draft 1999 NEI for HAPs contains more than 500 individual compounds.

One of the major issues encountered in compiling the 1996 NTI was the reporting of information on HAP groups. If aggregated emissions were reported for the HAP groups, various assumptions had to be made about the HAP compound group, which introduced uncertainties in the use of the 1996 NTI data for air quality modeling and risk characterization. Uncertainties in air quality modeling arise because individual compounds in some of the HAP groups (e.g., mercury compounds) have substantially different fate and transport characteristics which strongly affect the modeling results. In order to model the atmospheric deposition of mercury, the NEI needs to include three different species of mercury because they do not all transport or react the same once they are in the atmosphere. Uncertainties in risk characterization also arise because HAP exposures and associated human health effects can vary enormously among the specific compounds within many of the HAP groups. For example, some compounds within the polycyclic organic matter (POM) group are relatively non-toxic, while others are highly potent carcinogens.

To reduce uncertainties and potential overestimation of risk in future NEI-based assessments,

the EPA requested that organizations providing data for the 1999 NEI report emissions for specific compounds, both for individual HAP species and for HAPs within compound groups. If emissions of individual pollutants within HAP groups were reported, the EPA accepted aggregated compound group emissions. However, the EPA will have to use simplifying assumptions regarding speciation in order to use these data as inputs to models.

In the "Questions and Answers for the 1999 NTI", recommendations for reporting data for specific groups of compounds are summarized in a hierarchy of most preferred method to least preferred.¹ For pollutant groups, only one reporting strategy per HAP group per source should be used to avoid potential overestimation of emission levels and risk. The preferred methodology for all compound groups is to report emissions and associated CAS numbers of all individual species (e.g., report emissions and associated CAS numbers of arsenic oxide, lead arsenate, etc., rather than emissions of arsenic compounds as a whole.).

For metals, the EPA encouraged data submitters to report all individual compounds as the mass of the total compound, not just the metal within the compound. If individual metal compounds could not be reported, a less preferred method for chromium, lead, mercury and nickel is to report the mass of emissions of the metal, not of the entire metal compound.

- Chromium - Separate chromium compounds into hexavalent chromium (CAS #18540299), trivalent (CAS #1606583) and other chromium.
- Lead - Separate lead compounds into organic and inorganic.
- Mercury - Separate mercury compounds into particulate, gaseous elemental, and gaseous divalent.
- Nickel - Separate nickel compounds into nickel subsulfide (CAS #12035722) and other nickel (CAS #7440020).

For POM, the EPA requested data submitters to report emissions of the individual compounds rather than emissions of total polycyclic aromatic hydrocarbons (PAH) or total POM. The most important PAH compounds to report individually are the 7-PAH compounds listed in Table 1. The EPA also encouraged the reporting of other individual POM compounds for which cancer assessments are available. If emissions of all individual PAH could not be reported, then data submitters were encouraged to report 7-PAH as a subgroup.

For dioxins/furans, the EPA encouraged data submitters to report mass emissions and associated CAS numbers of all individual congeners of both chlorinated dibenzodioxins (CDDs) and chlorinated dibenzofurans (CDFs). If emissions of individual CDD and CDF congeners could not be reported, data submitters were encouraged to report dioxins and furans as 2,3,7,8-tetrachlorodibenzodioxin (TCDD) toxic equivalents (TEQ).

For xylenes, cresols and glycol ethers, the EPA encouraged data submitters to report mass emissions and associated CAS numbers for individual xylene and cresol isomers and glycol ethers. If individual emissions of xylenes or cresols could not be reported, data submitters were encouraged to report total emissions of xylenes or cresols as a group under "xylenes (mixture of

Table 1. POM Compounds

7-PAH	Other POM Compounds for which we have cancer assessments
Benz(a)anthracene	Carbazole
Benzo(a)pyrene	Dibenz[a,h]acridine
Benzo(b)fluoranthene	Dibenz[a,j]acridine
Benzo(k)fluoranthene	7H-Dibenzo[c,g]carbazole
Chrysene	Dibenzo[a,e]pyrene
Dibenz(a, h)anthracene	Dibenzo[a,i]pyrene
Indeno(1,2,3-cd)pyrene	Dibenzo[a,l]pyrene
	7,12-Dimethylbenz[a]anthracene
	1,6-Dinitropyrene
	1,8-Dinitropyrene
	3-Methylcholanthrene
	5-Methylchrysene
	5-Nitroacenaphthene
	6-Nitrochrysene
	2-Nitrofluorene
	2-Nitrofluorene
	1-Nitropyrene
	4-Nitropyrene

o, m, and p isomers)” (CAS #1330207) or “cresols/cresylic acids”. For glycol ethers, data submitters should use EPA guidance on glycol ethers to identify compounds that are glycol ethers.² In addition, ethylene glycol butyl ether (EGBE) has been delisted as a HAP and should not be included in the glycol ethers compound group.

Reporting of Source Categories

The NEI contains HAP emission estimates for point, nonpoint, and mobile source categories. Point sources in the NEI are sources for which the specific location is known. Point source categories include major and area sources as defined in section 112 of the CAA. Point sources have stack and fugitive emissions. Nonpoint source categories include area sources and other stationary sources that may be more appropriately addressed by other programs rather than

through regulations developed under sections 112 or 129 of the CAA. Nonpoint sources in the NEI also include area sources that are not identified as point sources. Nonpoint and mobile source emissions are reported as aggregates at the county level in the NEI. The NEI for HAPs includes a data field that identifies emissions as major, area/other, or mobile.

The following section summarizes the NEI source category definitions.

- Major sources, as defined by section 112 of the CAA, are stationary sources that emit or have the potential to emit 10 tons per year or more of any listed HAP or a 25 tons per year or more of a combination of listed HAPs. (www.epa.gov/ttn/uatw/pollsour.html) When estimates of potential emissions are not available, the NEI identifies point sources as major based on reported emissions being at or above 10/25 tons per year. The NEI includes facility data for all major sources. Examples of major sources include electric utility plants, chemical plants, steel mills, oil refineries, and hazardous waste incinerators. These sources may release air toxics from equipment leaks, when materials are transferred from one location to another, or during discharge through emissions stacks or vents.
- Area sources, as defined by section 112 of the CAA, are stationary sources that emit or have the potential to emit less than 10 tons per year of a single HAP or less than 25 tons per year of a combination of HAPs. (www.epa.gov/ttn/uatw/pollsour.html) When estimates of potential emissions are not available, the NEI identifies point sources as area based only on reported emissions being below 10/25 tons per year. The NEI includes facility data for some area sources and aggregated emission estimates at the county level for the remaining area sources. Area sources are regulated under toxics provisions in the CAA. Examples of area sources include neighborhood dry cleaners and gas stations. Though emissions from individual area sources are often relatively small, collectively their emissions can be of concern particularly where large numbers of sources are located in heavily populated areas.
- Other stationary sources are sources that may be more appropriately addressed by other programs rather than through regulations developed under air toxics provisions such as sections 112 or 129 in the CAA. Examples of other sources include wildfires and prescribed burning whose emissions are being addressed through the burning policy agreed to by EPA and the U.S. Department of Agriculture (USDA). The NEI includes aggregated emission estimates at the county level for these other sources.
- Mobile source categories include on-road vehicles, non-road 2- and 4- stroke and diesel engines, off road vehicles, aircraft, locomotives, and commercial marine vessels. The NEI includes aggregated emission estimates at the county level for mobile sources.

To evaluate EPA's progress in reducing air toxic emissions through the Maximum Achievable Control Technology (MACT) standards and to identify sources that may be modeled as part of residual risk assessments, operations within facilities that are subject to MACT standards are identified in the NEI by MACT codes. The tagging of data with MACT codes allows EPA to determine reductions attributable to the MACT program. The NEI associates MACT codes with major and area source data. MACT codes are assigned at the process level or at the site level in the point source data, e.g., the MACT code for municipal waste combustors is assigned at the site

level whereas the MACT code for petroleum refinery catalytic cracking is assigned at the process level. MACT codes are also assigned to source categories in the nonpoint source file. The EPA requested state and local agencies and tribes to include MACT codes as part of their submittal of 1999 HAP emission inventory data. If they did not include MACT codes in their inventories, then the EPA assigned MACT codes.

For dispersion and exposure modeling, it is preferable to model all stationary sources as point sources. To the extent possible, the EPA encouraged organizations to provide facility-specific emissions data for all point sources, regardless of whether they are classified as major or area. At a minimum, the EPA requested that all major sources, including both MACT and non-MACT sources be reported as point sources in the 1999 NEI. The EPA also encouraged organizations to report facility-specific emissions data for all MACT source categories regardless of whether a facility is classified as major or area. The NEI retains all facility-specific data as point sources, regardless of the magnitude of the emissions, to facilitate more comprehensive assessments. Therefore, no reporting thresholds exist for point sources in the NEI.

COMPILATION OF 1999 NEI

Figure 1 illustrates the major steps involved in compiling the 1999 NEI. Key processing activities include:

- Submittal of 1999 HAP inventory data by state and local agencies, tribes, industry;
- Blending/Merging of data from multiple data sources;
- Augmentation of blended data for missing data elements;
- QC/QA of the data;
- Preparation of draft NEI for HAPs for external and internal review;
- Incorporation of external and internal review comments; and
- Preparation of final NEI for HAPs.

Figure 1 presents an overall flowchart of the major steps, but individual activities within QC/QA, blending/merging of data, and data augmentation do not occur sequentially. Certain activities are also performed simultaneously during the development of the 1999 NEI. This paper provides a description of the key steps, but does not describe all activities within each step chronologically.

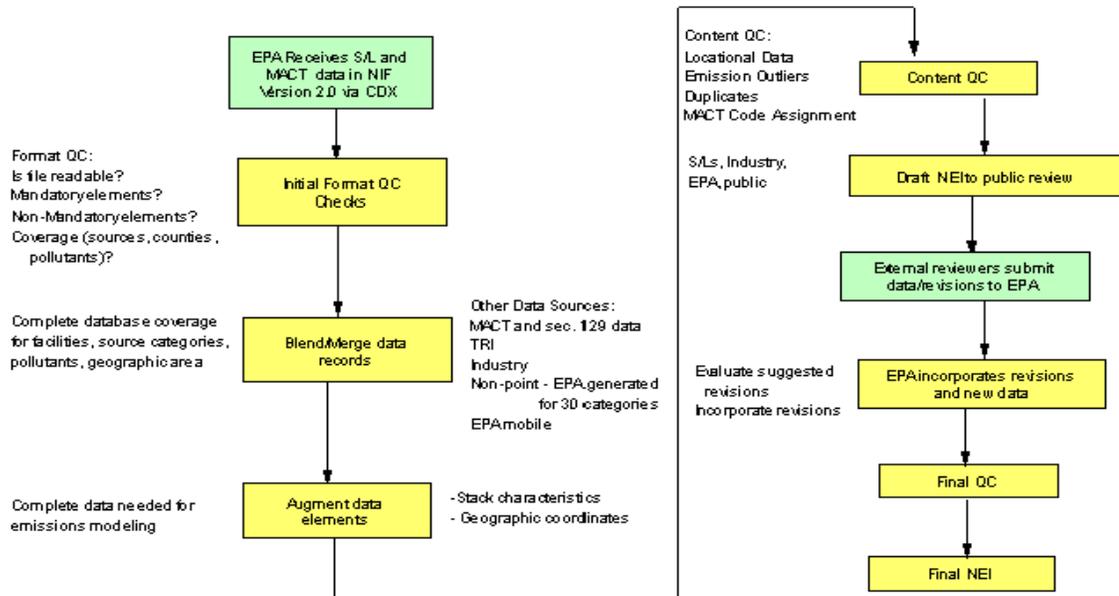
Data Submittal

The sources of data in the NEI for HAPs are:

- State and local agency data and tribal data,
- Industry data,
- Data gathered by the EPA's Emission Standards Division while developing MACT standards,
- Toxic Release Inventory (TRI) data,
- Data developed by the EPA's Office of Transportation and Air Quality,
- Estimated nonpoint emissions for sources not included in the state/local/tribal data, and
- 1996 NTI data for facilities in states and local areas that did not provide data to EPA and were not in 1999 TRI or MACT data.

The EPA prefers to use state and local agency and tribal HAP inventory data to compile the NEI. The NEI includes data for all 50 states, the District of Columbia, Puerto Rico, and Virgin Islands.

Figure 1. 1999 NEI Compilation Steps



American Samoa and Guam did not provide data for the draft 1999 NEI for HAPs, but their data will be accepted if provided by June 1, 2002. Because complete data are not available from some states, tribes, and territories, the EPA prepares the NEI for HAPs using additional data sources. The EPA requested organizations to submit their 1999 HAP emission inventory data in the NEI Input Format (NIF) Version 2.0.³

The EPA is very pleased with the participation of state and local agencies that provided HAP inventory data. Figure 2 shows states that provided 1999 HAP emission inventory data to the EPA. Table 2 indicates that 48 agencies located in 39 states provided stationary source data to the EPA by June 1, 2001. The EPA incorporated these data in the first draft of the 1999 NEI for HAPs. State and local agencies, tribes, and industry may submit additional or revised inventory to EPA by June 1, 2002 to be incorporated into the final version of the 1999 NEI for HAPs.

Blending/Merging of Data from Multiple Data Sources

In order for the 1999 NEI to contain reliable data for air dispersion and exposure modeling, it must be comprehensive for the 188 HAPs emitted from point, nonpoint and mobile sources in all states, DC, and territories. As noted in the point source draft 1999 NEI for HAPs documentation,

Figure 2. Agencies Who Submitted 1999 HAP Emission Inventory Data/Revisions to EPA by February 2002

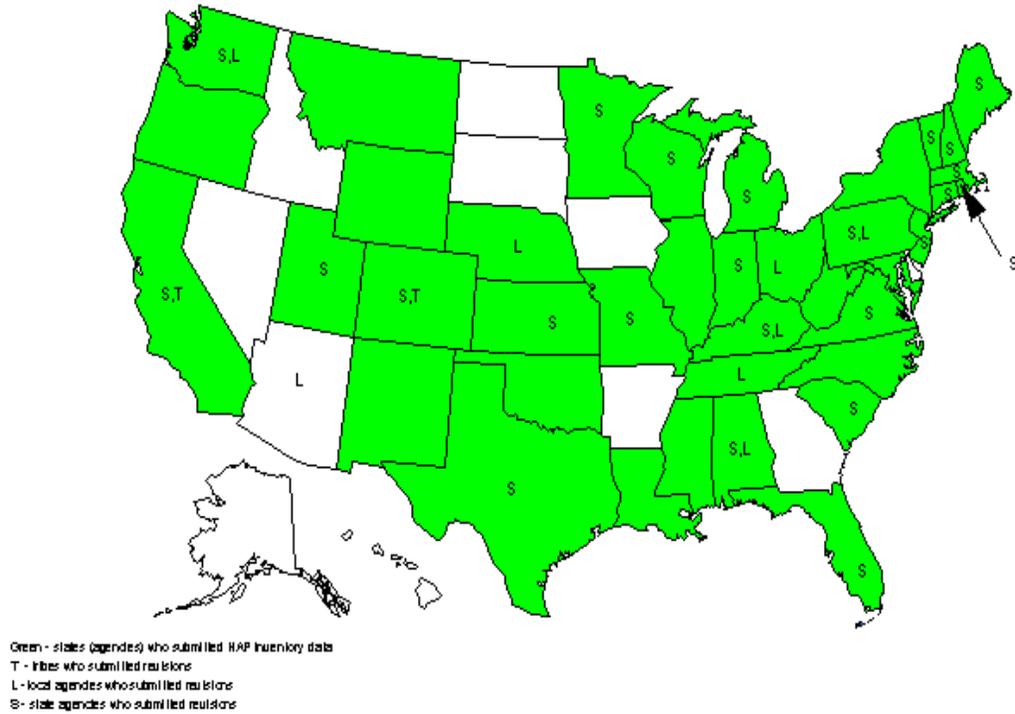


Table 2. Summary of Stationary HAP Inventory Data Received from Agencies and Tribes by June 2001

Agency Type	# Point Inventories	#NonPoint Inventories	# Stationary Inventories
State	37	11	38
Local	10	2	10
Tribe			
Total Agencies	47	13	48
Total States (State/Local/Tribe)	39	13	39

state and local agencies provided a large amount of data for the first draft of the 1999 NEI.⁴ Because facilities and source categories were missing from some state and local agency data, the EPA merged additional data from other sources to compile the draft 1999 NEI for HAPs.

Point Sources

An automated QC tool was used to check each state and local agency, MACT source category, and TRI file for format and data field errors.⁵ Duplicate records were then removed, along with records that had null and zero emission values. Referential integrity violations, invalid codes and erroneous locational data were then corrected or added if possible. Then the files were used in the data blending/merging process. State and local agency files were compared with MACT source category files and the TRI database. Duplicate facilities were identified based on the state, county, facility name, address, location (latitude/longitude), and common IDs. An automated facility-matching program was run to identify common facilities. For unmatched facilities, the EPA applied an algorithm that stripped out punctuation and leading/trailing spaces, dropped insignificant punctuation, standardized corporate names, compared facility names on a case sensitive basis and identified similar sounding facility names in each county with exact locational data, then similar coordinates. Candidate pairs were then reviewed manually.

To better understand the matching and blend/merge process, it is important to distinguish between the data fields, “site id” and “facility id” as defined in the NEI for HAPs. In the NEI for HAPs, there can be multiple sites associated with the same NTI Unique Facility ID. The NTI Unique Facility ID is currently stored in the Federal Facility ID field in the site table of the NIF Version 2.0. There are two reasons for this one-to-many relationship between facilities and sites:

- Multiple data sources have supplied data for the NEI for the same facility; or
- One source supplied multiple site records for co-located facilities.

For example, a state may have submitted a set of records for a facility with Site ID AL001. The EPA may have provided MACT data for the same facility under Site ID EM234. Although these data are for the same facility, its emissions are for different processes within the facility and do not duplicate the emissions data submitted by the state. A common NTI Unique Facility ID was then assigned to the two different Site IDs. Not only is it easier to make this assignment than change the Site IDs in the remaining tables, it preserves the original site IDs. This assists reviewers in tracking their data during the review process, aids users in tracing the origin of the data and helps EPA in comparing data for the same sites from year to year.

MACT codes were assigned at either the site or process level (but not both) based on:

- emissions data provided by the MACT engineer,
- a facility list provided by the MACT engineer, or
- the Standard Industrial Classification (SIC) code or the Source Classification Code (SCC)

If a MACT code was assigned at the process level, some processes at a site might be associated with a MACT category while others might not be. In all cases, any one process and thus any emission record can be associated with one and only one MACT category. If data were supplied by the EPA for a particular MACT category or a facility list was provided, then the appropriate MACT code was assigned and SIC codes and SCCs were not used to default the MACT code.

In the merging of the different data sets, where data were supplied for the same facility, MACT category, and pollutant from two or more data sources, only one data source was selected. In choosing which records to keep, the hierarchy for the most part was as follows:

- State and local agency data were preferred over MACT data,
- MACT data were preferred over TRI data, and
- TRI data were preferred over 1996 base year NTI data.

An exception to this approach was given to municipal waste combustors data and mercury data for coal-fired utilities supplied by MACT engineers because of extensive source testing conducted by the EPA and industry. During the Blend/Merge process, the EPA is conservative on retaining facilities. Where facilities could not be conclusively matched, multiple facilities are retained to enable the external reviewers to make the final judgment.

Figure 3 illustrates the Blend/Merge Process for Point Sources.

Nonpoint Sources

To begin compiling the nonpoint stationary source data, the EPA first identified nonpoint stationary source categories that emit HAPs present in the 1996 NTI. The methodology for compiling the nonpoint stationary source inventory used the following hierarchy:

1. Obtain HAP emission data from state and local agencies and tribes.
2. Supplement state/ local/tribal data with MACT nonpoint source category data.
3. Supplement state/local/tribal and MACT inventory data with EPA generated data using emission factors and activity data (The nonpoint source draft 1999 NEI for HAPs documentation presents the methodology for these categories)⁶.
4. Reconcile point and nonpoint emission estimates for the same categories by reducing or removing nonpoint emissions.

Mobile Sources

The EPA prepared the mobile source estimates for the draft 1999 NEI for HAPs.

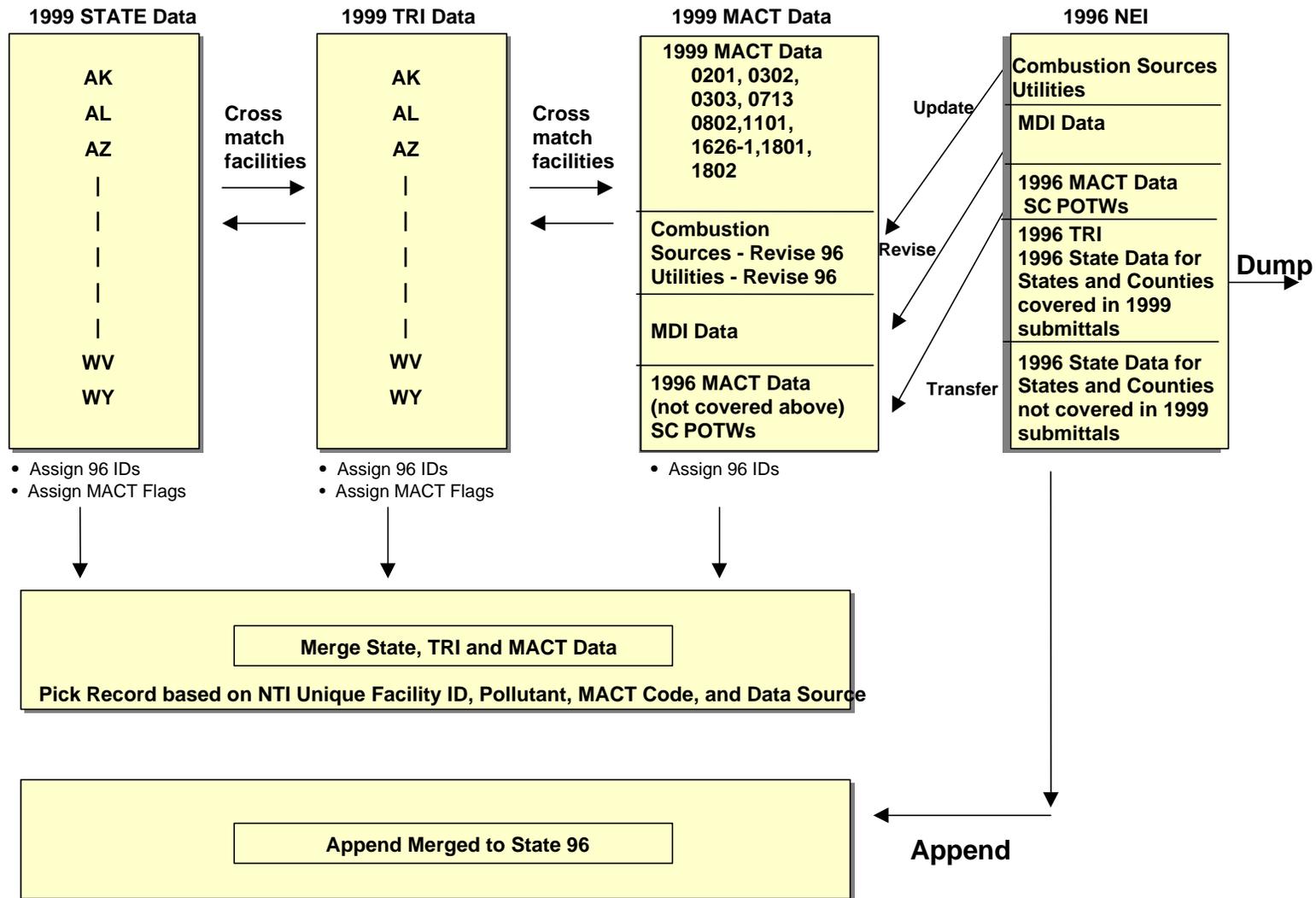
Data Augmentation

To support air dispersion and exposure modeling, the point source NEI for HAPs should at a minimum contain the following data:

- source identification (name and ID),
- specific geographic location (latitude, longitude, address, state, county),
- source description (release type – stack/fugitive, source type – major/area),
- facility emissions at the facility/unit/process/emission release point (stack) level,
- emissions type (actual, allowable, potential, maximum),
- stack parameters (height, diameter, velocity, temperature, flow rate),
- operating schedule,
- source category information (SIC code, MACT code, SCC), and
- control information (device and efficiency).

Data augmentation occurs at different times in the process. After investigating errors identified

Figure 3. 1999 Draft Point Source NEI (Toxics): Blend/Merge Process



by EPA's automated QC tool and Geographic Information System (GIS) tools and prior to the blending and merging of data, the EPA augmented invalid code values and locational coordinates. Stack parameters were augmented after the blending and merging of data. The EPA used specific guidance on augmenting missing fields of data.⁷ Flags were added in the database to identify the augmented fields of data.

Locational Defaults

Latitude and longitude are needed to correctly place facility release points and associated emissions into specific geographic domains for emissions modeling. The draft 1999 NEI for HAPs contains 185,540 records in the emission release point table. QC/QA of these records indicates that 79% of the records were located in the correct county. The EPA first defaulted missing or invalid latitude/longitude within a facility with multiple release points from one emission release point at the site to all emission release points at the site. (These records are included in the 79%) For the remaining 40,083 records, the latitude/longitude was augmented by attempting to find a valid latitude/longitude using geocoding software⁸ or the Facility Registry System's (FRS) database of EPA plant information.⁹ The EPA used Geocoder to find latitudes/longitudes for 26,378 records (14% of total records). Of these records, Geocoder assigned the following matches.

- Exact- Match to within a unique intersection or within a single side of a single block, 12,593 records (47.7% of Geocoder matches)
- Near - Match to a single street block, but the correct placement within the block is unknown, 476 records (1.8% of Geocoder matches)
- Zip code + 2 - Match to a 5 digit zip code + the first two digits of the zip code, 629 matches (2.4% of Geocoder matches)
- Zip code5 - Match to a 5 digit zip code, 12,402 records (47% of Geocoder matches)
- Zip code3 - Match to a 3 digit zip code, 135 records (0.51% of Geocoder matches)
- Ambig - Match to one of possible multiple street segments, 143 records (0.54% of Geocoder matches)

For 13,014 records (7%) with missing or invalid latitude/longitude, the EPA defaulted coordinates to the county centroids. There are 691 records that do not contain any coordinate information because they are portable facilities that move from county to county within a state.

Stack Parameters Defaults

Stack parameters (stack height, stack diameter, temperature, velocity, and flow rate) are needed for air dispersion modeling and reside in the emission release point table. QC/QA of these records in the draft 1999 NEI for HAPS shows the following.

- 23% of records have all 5 parameters
- 11 % of records have 4 parameters
- 1% of records have 3 parameters
- <1% of records have 2 parameters
- 3% of records have 1 parameter
- 28% of records have no parameters for stack emissions
- 34% have no parameters fugitive emissions

For missing or out-of-range parameters of fugitive releases, the lower end of the parameter range is used to default data (0.01 ft for stack height and 50 degrees F for temperature, 0.01 ft stack diameter, 0.01 ft/sec stack velocity, and 0.01 cu ft/sec flow rate). For non-fugitive releases, parameters are defaulted using:

- calculations for missing flow rate, velocity and diameter, if only one of the three parameters is missing, or
- NEI default look-up tables generated by SCC and SIC codes.

When there are multiple replacement values available from SCCs and SIC codes for a specific stack, the default record having the lowest stack height is used. If SCCs or SIC codes are not available, national default values are used (10 ft stack height, 72 degrees F temperature, 1 ft stack diameter, 15 ft/sec stack velocity, and 12 cu ft/sec flow rate).

QC/QA

The EPA conducted a variety of internal activities to QC the data provided by state and local agencies, MACT engineers and TRI.¹⁰ These included using:

- Automated QC format tool to identify potential errors with data integrity, code values, and range
- Geographic Information System (GIS) tools to verify facility locations
- QC Content Analysis to identify potential errors with emissions estimates - pollutant-, source category-, and facility-level emission estimates were reviewed to identify outliers and duplicate emissions and sites

QC/QA of the data by the EPA occurred at different times during the compilation of the 1999 NEI for HAPs. The EPA ran Format QC and GIS QC at multiple times (prior to blending and merging of the draft data, prior to data augmentation, prior to incorporating revisions, etc.) during the development of the inventory. The EPA conducted Content QC to identify potential errors with emission estimates during the external review of the first draft of the 1999 NEI for HAPs. After performing Content QC, revised records were compiled into several holding tables for use during the revision process. This section discusses the three main types of QC/QA conducted by the EPA.

QC Format Tool

The EPA developed and used the automated QC format tool to check all data submitted to EPA for format and data field errors. The QC format tool verifies that minimum data requirements for file acceptance are present. The tool also performs data field checks related to the codes, numeric data ranges and locational data at the state level. The EPA accepted data with data field errors as these could be corrected with minimal effort. After completing QC of format, the EPA then corrected referential integrity violations, invalid codes, and erroneous locational data for county assignments.

Evaluation of Locational Data

The EPA conducted QC of locational data for each facility. To facilitate this evaluation, coordinates provided in Universal TransMercator units (UTMs) were first converted to latitude/longitude in degrees. After all coordinates were converted, obvious errors in latitude

and longitude were corrected.

Next, the EPA examined all of the coordinates associated with one facility to see if these coordinates might be too far apart and therefore indicate possible incorrect assignment of the NTI Unique Facility ID. The standard deviation for all latitudes and longitudes associated with one facility were calculated and a list of all facilities in which the standard deviation for either coordinate was $>.02$ degrees was compiled. These records were set aside for manual review. The EPA also created a set of facilities that had identical latitude/longitudes, but different facility IDs to find possible duplicate facilities.

The last stage in the locational data review process was to use a GIS program to overlay the corrected latitude/longitude pairs with the reported county's boundaries. This step included the locating of each latitude/longitude value and comparing it to the physical boundaries of the county to which the value is associated. If the release point was within 5 kilometers of an outside boundary of the county, the point was assumed to be valid.

Content QC

The EPA performed content QC to identify emission outliers, duplicate facilities and duplicate emissions. Significant changes between 1996 and 1999 emissions and/or extreme variation within the 1999 draft data were evaluated. This included comparing 1996 HAP estimates to 1999 draft estimates for each facility, comparing total emissions for each state between 1996 and 1999, and comparing total emissions for each MACT category between 1996 and 1999. These big picture summaries highlighted source categories, states, and facilities with potential problems. Then the EPA performed QA/QC to specifically highlight individual facilities. The facility QC included identifying the top emitters for each HAP nationwide, ranking each facility based on its emissions of each HAP on a national basis, and listing the top emitters for each HAP/MACT combination nationwide.

Preparation of Draft NEI for HAPs for External and Internal Review

To improve the quality of the first draft of the NEI for HAPs, the EPA made the draft available for review from October 1, 2001 – February 1, 2002. Reviewers were asked to identify duplicate facilities, sites, and emission records; identify facilities not operating in 1999; and to provide emissions data for missing facilities.

The draft 1999 NEI for HAPs was posted on the following ftp site,

<ftp://ftp.epa.gov/EmisInventory/draftnei99ver2/haps>.

The ftp site contained the following subdirectories:

- README file.
- Documentation.
- State data files in NIF Version 2.0. and
- Data Summary Reports.

README

The README file contains an overview of the files posted on the ftp site, organization of files,

description of files, helpful hints on how to review data, instructions on submitting revisions, schedule, and software needed to review the data. Reviewers are strongly encouraged to first read the README before downloading any files or reviewing the draft inventory.

Documentation

Separate reports are available that provide documentation for the draft 1999 NEI for HAPs point, nonpoint and mobile source inventories. Reviewers are encouraged to read the documentation to better understand how the draft inventory is compiled and the sources of data.

State Data Files in NIF Version 2.0

The ftp site has separate point and nonpoint zipped files for each state including Washington DC, Puerto Rico, and Virgin Islands. ACCESS 97 data files are provided in NIF Version 2.0.

Data Summary Reports

A series of reports summarize the data in the draft 1999 to facilitate its review. These include the following 4 reports.

- 1) 96-99 Site List – provides a list of sites found in the 1996 base year NEI, the 1999 draft, and sites common to both the 1996 and 1999 inventories. By sorting this list by state, county, and facility name, reviewers could evaluate the sites listed in each county and detect potential duplicates. If a site is found in the 1996 version but is not in the 1999 draft, reviewers could verify that the facility closed.
- 2) 99 NTI County Emissions Summary - provides a snapshot of the emissions for each HAP in each county (with emissions divided into stationary area and major subtotals). This report helps reviewers identify states and counties to investigate HAP emissions by source sector.
- 3) 99 NTI Facility Emissions Summary - provides a detailed list of HAP emissions per facility. This report helps to target facilities to investigate further by using the State Data Files.
- 4) 99 NTI Data Source Summary - provides summary emissions data for each facility/HAP/MACT combination where more than one data source (state or local agency, TRI, ESD) is available. The report indicates which source was selected for inclusion in the draft and allows a quick comparison of data from the different sources.

Incorporation of External and Internal Review Comments

The EFIG received revisions for stationary source inventories from 34 agencies located in 28 states by February 2002. As of March 1, 2002, agencies and tribes in 40 states have either provided 1999 HAP inventory data or revisions to the EFIG. Revisions and/or inventories were also received from MACT engineers for 92 of the 134 MACT source categories and from 9 facilities or industrial trade associations by February 2002. Figure 2 and Table 3 presents a summary of revisions received by February 2002 from state and local agencies.

Revisions to the draft 1999 NEI for HAPs are being subjected to a rigorous review process. Specifically the following steps are being performed:

Table 3. Summary of Stationary HAP Inventory Revisions Received from Agencies and Tribes by February 2002

Agency Type	# Point Inventories	#NonPoint Inventories	# Stationary Inventories
State	23	4	24
Local	9	1	9
Tribe		1	1
Total Agencies	32	6	34
Total States (State/Local/Tribe)	27	5	28

- QC format of revised and new records;
- Review of documentation provided by the commenter to ensure that it is consistent with the actual changes to the inventory submitted;
- Verification that the add/revise/delete designations of the revisions are accurate (e.g., that a record designated for addition is not in fact a revision to an existing record);
- Verification that there is no source category overlap between the draft and revised data (nonpoint stationary categories); and
- Review of the revised emission estimates by source category and HAP to identify outliers and determine validity of such estimates.

Revisions are grouped into the following three categories:

- Changes to draft emission records;
- Additions of new facilities or additions of HAPs to a facility; and
- Deletion of draft NEI facilities or deletions of HAPs within a facility.

For the most part, revisions provided by state and local agencies and tribes will be incorporated to produce the final inventory. When questions arise over specific revisions, the EPA will contact the reviewers. Memos describing how specific revisions are addressed will be sent to the reviewers.

Point source revisions received by February 2002 will be incorporated in the final 1999 NEI for HAPS using the following methodology.

1. Revised files in NIF Version 2.0 will be processed using the QC Format tool.
2. All revisions and additions in correct format, with valid codes, and within valid data ranges will be made for most non-emissions fields.
3. Additions of new facilities not in the draft 1999 NEI for HAPS will be made.
4. Emission record additions and revisions to existing emission records will be automatically processed if the original data were provided by the reviewers.
5. Facilities marked for deletion will be evaluated to determine whether the deletions should be processed.
 - If draft NEI data were provided by state/local agencies, the EPA will process the deletion.
 - If draft NEI data were provided by TRI or MACT databases, the EPA will further

evaluate the proposed deletions. If a facility was not in operation, the EPA will delete it.

- If a TRI or MACT facility is recommended for deletion, but was open in 1999, and no duplicate facility is present in the draft NEI, the EPA will not process the deletion.
- If a TRI or MACT facility is recommended for deletion and was open in 1999, and a duplicate facility is present in the draft NEI, the EPA will then evaluate the HAPS emitted from processes within the duplicated facilities. If the HAP emission records are duplicated, the EPA will process the proposed deletions. The HAP emission records that are not duplicated will be retained. The EPA will then merge the HAPs for the duplicate sites using a NTI Unique Facility ID.

Preparation of Version 3.0 Draft 1999 NEI for HAPs

The next version of the draft 1999 NEI for HAPs will be available for review with draft 1999 ambient concentration estimates from the ASPEN model used in the NSATA during October 2002 – February 1, 2003. This version will contain revisions received by February 2002 and new inventory data received by June 2002. The steps to preparing the version 3.0 draft will be similar to those described above.

Preparation of Final 1999 NEI for HAPs

After the revisions to the draft version 3.0 1999 NEI for HAPs are incorporated, the EPA will conduct final QC and produce the 1999 NEI for HAPs. The final NEI will be publicly available in the following two formats.

- Air DATA Web Site – data summaries, www.epa.gov/air/data/, and
 - NEI ftp site – state NIF Version 2.0 files and documentation, <ftp://ftp.epa.gov/EmisInventory/>
- The final 1999 NEI for HAPs will be available in June 2003.

CONCLUSION

The 1999 NEI for HAPs will be used to support a variety of CAA programs and analyses. The EPA has compiled the NEI for HAPs using the best available data. However since there is no requirement in the CAA for state and local agencies to submit HAP inventory data to the EPA, the EPA must use several sources of data. The blending and merging of HAP inventory data from several sources results in a greater need for enhanced QC/QA of the inventory. During the development of the 1999 NEI for HAPs, the EPA greatly improved internal QC/QA. In order to improve the quality of the NEI for HAPs, the EPA works with its partners in state and local agencies, tribes, industry and the public during the acquisition of inventory data and review of draft inventories. The EPA looks forward to working with its partners in the process of compiling the 1999 NEI for HAPs.

REFERENCES

1. Questions and Answers for the 1999 NTI, www.epa.gov/ttn/chief/nti/ntiq&a.pdf.
2. Toxic Release Inventory List of Toxics Chemicals within the Glycol Ethers Category, EPA-745-R-00-004, <http://www.epa.gov/opptintr/tri/glycol.pdf>.

3. NEI Input Format (NIF) Version 2.0, www.epa.gov/ttn/chief/nif.index.html.
4. Documentation for the Draft 1999 Base Year Point Source National Emission Inventory for HAPs. <ftp://ftp.epa.gov/EmisInventory/draftnei99ver2/haps/documentation/>.
5. More information on EPA's QA/QC National Emission Inventory Input Format tool is available at <http://www.epa.gov/ttn/chief/nif/index.html#qa>.
6. Documentation for the Draft 1999 Base Year NonPoint Source National Emission Inventory for HAPs. <ftp://ftp.epa.gov/EmisInventory/draftnei99ver2/haps/documentation/>.
7. National Emission Inventory QA and Augmentation Memo, www.epa.gov/ttn/chief/emch/invent/index.html
8. More information on the Tele Atlas North America "EZ-Locate" geocoding software is located at <http://www.geocode.com/>.
9. More information on the EPA Facility Registry System (FRS) is located at <http://www.epa.gov/enviro/html/facility.html>.
10. Pope, A., Finn, S., Wilson, D., and Oh, J. "QA/QC – An Integral Step in the Development of the 1999 National Emission Inventory for HAPs", 2002 Air and Waste Management Association Annual Meeting, June 2002.

Key Words

Air Toxics, Hazardous Air Pollutants, Emissions inventory, NEI, NTI