QC/QA of 1999 NEI for HAPs

Anne Pope
pope.anne@epa.gov

April 17, 2002

What is the NEI?

EPA's National Emission Inventory

- National Emission Inventory of Criteria Pollutants (and Precursors) and HAPs
- Replaces NTI (HAPs) and NET (criteria air pollutants)
- Published by EPA's Office of Air Quality Standards and Planning

Geographic coverage	50 states, DC, Puerto Rico, Virgin Islands
Source category coverage	major - facility specific (point) area & other- facility specific (point) and county (non-point) mobile on-road & non-road - county
Number of HAPs	188 HAPs (Approximately 500 speciated HAPs)
Base year	1990, 1996, 1999 - Highest quality data 1996 and forward
Resolution of data; (level of detail)	facility/stack - major and some area sources county - mobile, other and some area sources
Temporal Coverage	annual emissions
Data System/Format	ORACLE
Uses	National, Regional and Local Scale Air Quality & Human Exposure Modeling, Residual Risk Assessments, CAA Programs (PBT, Urban, GW, etc.), Regulatory Impact Analysis, Trends, etc.

HAP Inventory Source Categories

- **Point Sources (facility)**
  - Major Sources (>10/25 tpy)
    - MACT
    - Non-MACT
  - Area Sources (<10/25 tpy)
    - MACT
    - Non-MACT
- **Non-Point Stationary Sources (county)**
  - Area/Other Sources (<10/25 tpy)
    - MACT
    - Non-MACT
- **Mobile Sources (county)**
  - On-road
  - Non-road

Stationary Point Sources

Point Sources - Sources are inventoried individually at stack/process level

- Source Identification
  - Facility Name and Id
- Source Location
  - Latitude, Longitude
  - Address, State/County
- Source Description
  - Release Type- Stack or Fugitive
  - Stack Parameters, Capacity
  - Source Type - Major or Area
- Process Description
  - Standard Industrial Classification (SIC) Code
  - Source Classification Code (SCC)
  - MACT Code
- Activity
  - Operating schedule
  - Throughput
  - Temporal Data
- Control Device
  - Equipment Type
  - Efficiency
- Emissions By Pollutant Species
  - Amount
  - Methods including emission factor
  - Emission Type - Actual, Allowable, Potential, Maximum

Stationary NonPoint Sources

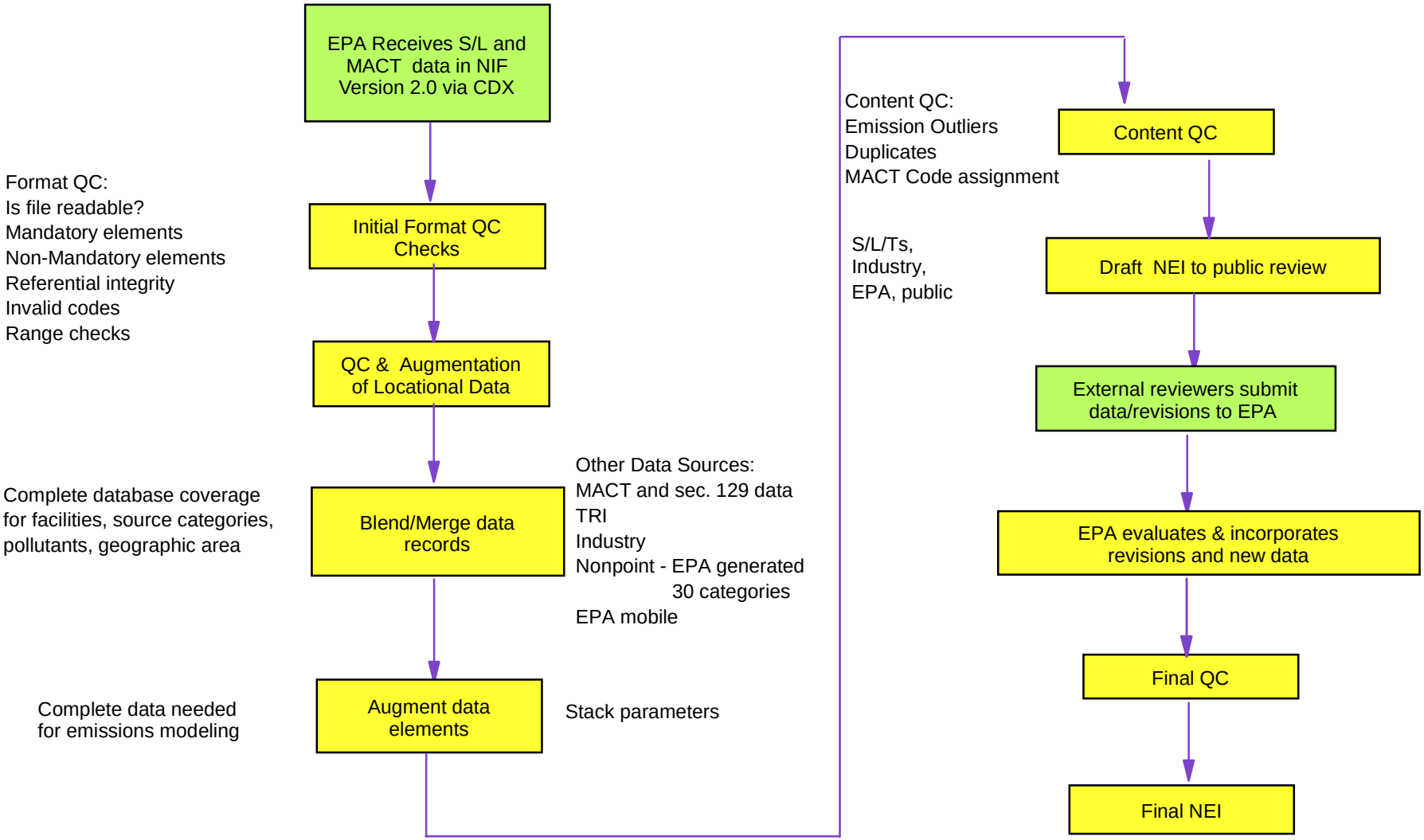
NonPoint Sources - Typically too small to be inventoried as point sources

- Source Identification
  - Source Category
- Source Location
  - County/State
- Source Category Description
  - Source Type - Major or Area
  - Source Classification Code (SCC)
  - MACT Code
- Activity
  - Throughput
  - Temporal Data
- Control Device
  - Equipment Type
  - Efficiency
- Emissions By Pollutant Species
  - Amount
  - Methods including emission factor
  - Emission Type - Actual, Allowable, Potential, Maximum

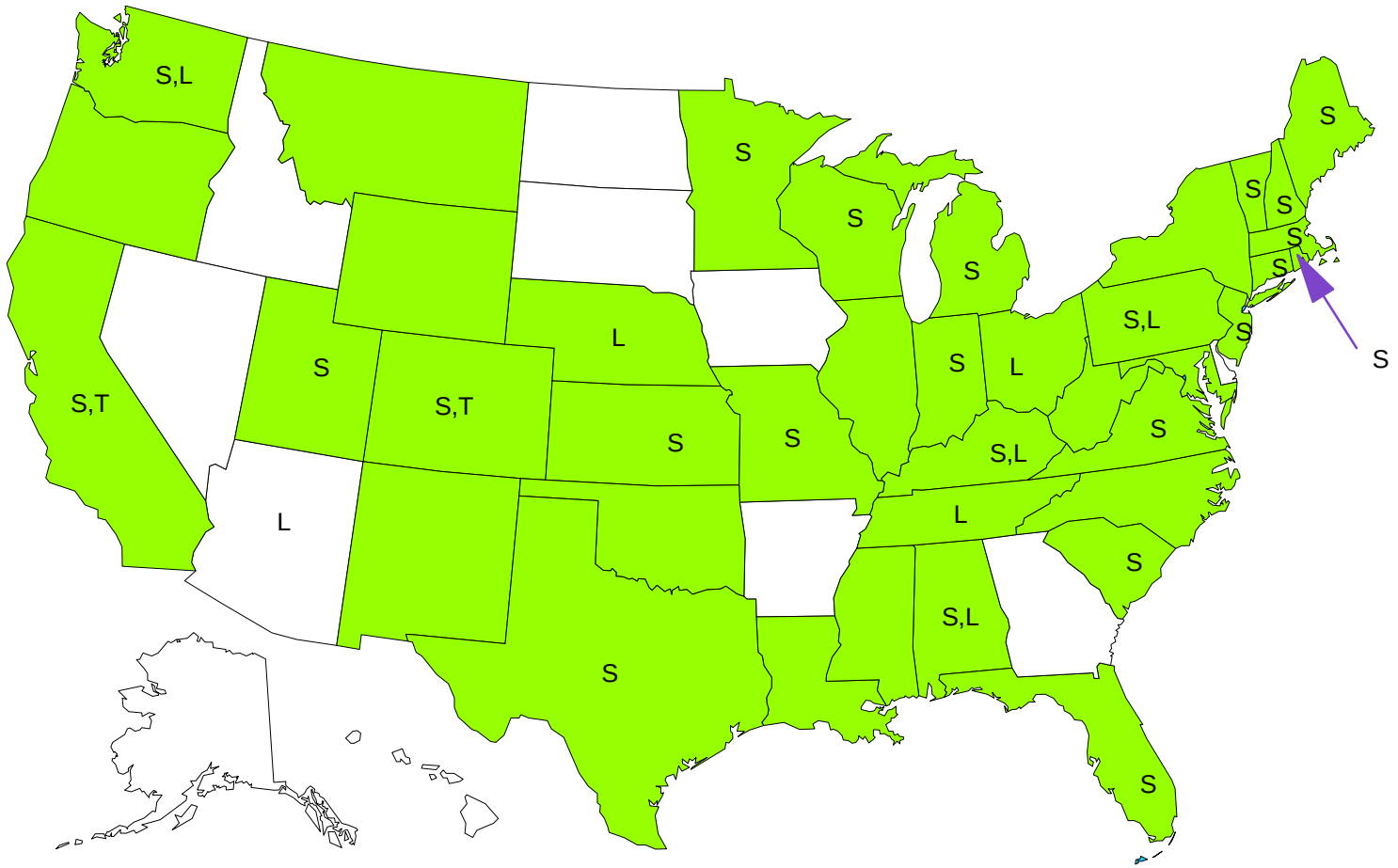
Goals for the 1999 NTI

- Stick to schedule
- Use standardized format for data inputs and outputs
  - NIF Version 2.0 to send in data
  - Outputs - NIF Version 2.0, summary data tables
- Consistently use best data
  - Hg from Utilities and MWC data - we will only use MACT data
  - Other sources - use State and Local agency and Tribal data if available
- Improve QA/QC
  - Reduce number of missing sources
  - Reduce errors in geographic coordinates
  - Reduce number of missing stack parameters
  - Improve consistency in nonpoint sources across nation
- Reduce number of nonpoint source categories by increasing number of categories inventoried as point sources
- Integrate criteria and HAP emissions where feasible

Compilation of NEI



Agencies Who Submitted 1999 HAP Inventory Data/Revisions to EPA by February 2002



Green - states (agencies) who submitted HAP inventory data
T - tribes who submitted revisions
L - local agencies who submitted revisions
S - state agencies who submitted revisions

Summary of 99 HAP Data Received from S/L/T

Agency Type	# Point Inventories Submitted by June 2001	# NonPoint Inventories Submitted by June 2001	# Stationary Inventories Submitted by June 2001	# Point Inventories Submitted by February 2002	# NonPoint Inventories Submitted by February 2002	# Stationary Inventories Submitted by February 2002
State	37	11	38	23	4	24
Local	10	2	10	9	1	9
Tribe					1	1
Total Agencies	47	13	48	32	6	34
Total States (State/Local/ Tribal)	39	13	39	27	5	28

March 1, 2002 - Agencies and tribes located in 40 states have either provided 1999 HAP inventory data or revisions to EPA.

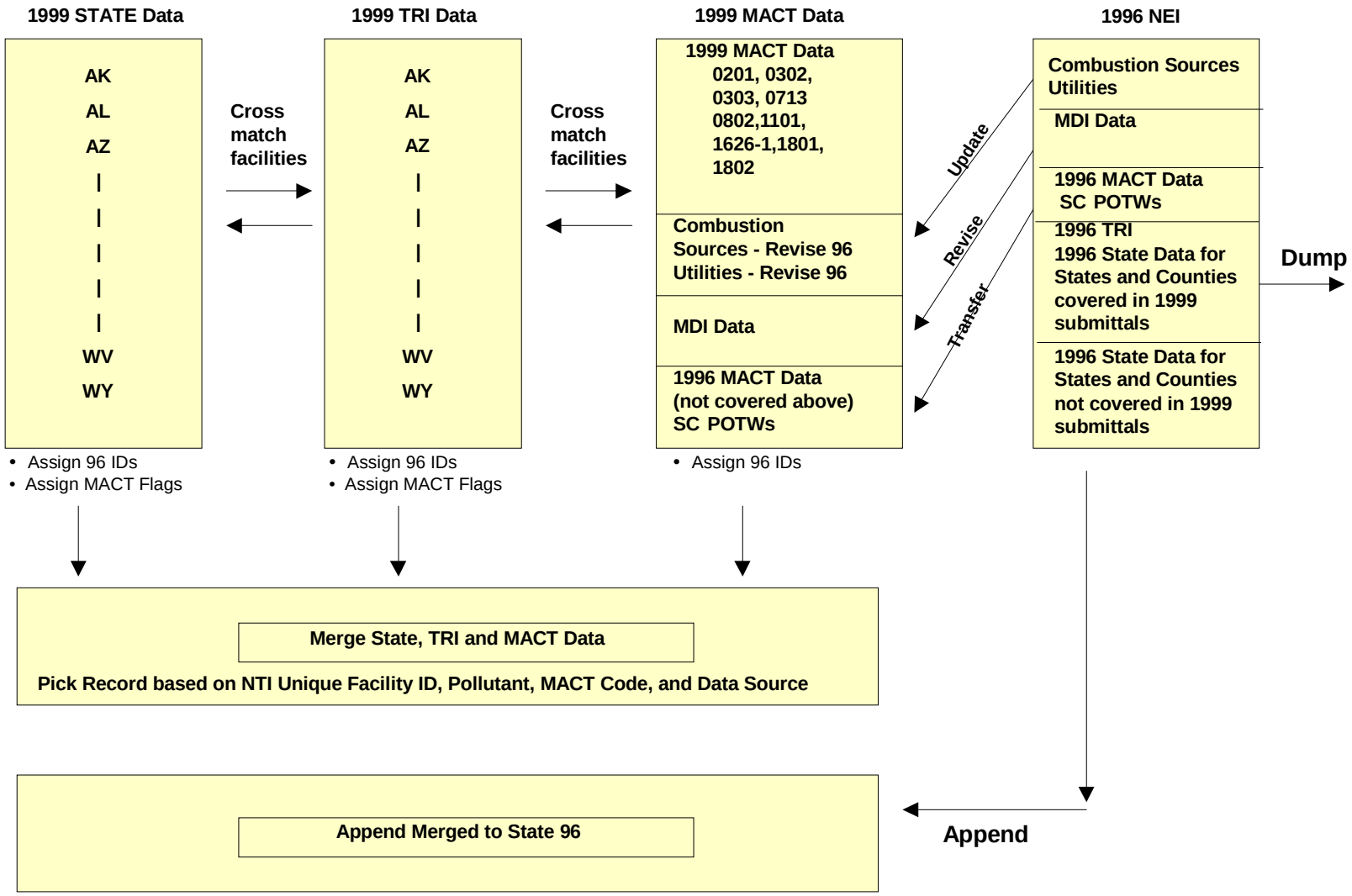
Blending/Merging Steps

- Correct format errors - invalid codes, referential integrity violations, etc.
- Default/Correct latitude/longitude
- Assign MACT codes
  - Emissions data provided by MACT engineer
  - Facility list
  - SIC code or SCC default
- Compare data from 4 data sources - S/L/T, MACT, TRI, 96NTI
  - Identify duplicate facilities using state, county, facility name, address, latitude/longitude, common IDs using automated program
  - Unmatched facilities - run second program to further match facilities
- Select emissions based on hierarchy
  - Goal: emissions of all HAPs emitted from each process within a facility
  - S/L/T data preferred over MACT data except for MWCs and Hg from utility coal
  - MACT data preferred over TRI data
  - TRI data preferred over 96 NTI
- Assign Unique Facility IDs
  - Multiple Site IDs for processes and HAPs supplied by individual data sources
  - Each facility has one Unique Facility ID

Assignment of Site and Facility IDs - Example

County State	NTI Unique Facility ID	Site ID	Facility Name	MACT Code	SIC Code	SCC	HAPs	Data Source
Harris TX	NTI48001	TX011391	BBB Refinery		2911	30600503 Wastewater	Ethylbenzene	State
Harris TX	NTI48001	T$4709	BBB Refinery		2911		Toluene	TRI
Harris TX	NTI48001	480101	BBB Refinery	0502 (Pet Ref: Cat Cracking)	2911	30600201 Cat Cracking	Benzene, Hexane	MACT
Harris TX	NTI48001	480902	BBB Refinery	0503 (Pet Ref: Vents)	2911	40400151 Storage Tanks	Benzene, Formaldehyde	MACT
Harris TX	NTI48001	14455	BBB Refinery	0107 (Industrial Boiler)	2911	10200501 Dist.Oil Boiler	Formaldehyde	MACT

Blending/Merging Process



QC/QA

- **Goals of QC/QA Program**
  - Identify and correct duplicate facility, site, and emissions data
  - Identify and correct erroneous emissions data
  - Verify coordinate data
  - Review and augment stack parameters
  - Identify and correct format errors

- **Types of QC conducted by EPA**
  - Format QC - identify potential errors with format
    www.epa.gov/ttn/chief/nif/
  - QC of Locational Data - verify facility locations in county
  - Content QC - identify emission outliers, duplicate facilities, duplicate emissions

Format QC

- **Based on minimum data requirements - Activities include:**
  - **Checking data integrity (relational data base with required key fields)**
  - **Evaluating code values**
  - **Conducting range checks**

- **Corrections**
  - **Referential integrity violations**
    - **Remove duplicate records and records with null and zero values**
    - **Work with state and local agencies and tribes to correct records by creating needed records or removing records**
  - **Code Values**
    - **Correct invalid codes**
    - **Assign new codes not included in NIF Version 2.0 code table, e.g., additional HAPs reported by an agency that are not included in Pollutant Code Table**
  - **Range Checks**
    - **Contact state and local agencies and tribes to verify that reported value is correct, e.g., stack parameters**
  - **Augment Stack Parameters**

Format QC: Augmentation of Stack Parameters

- **Stack Parameters: height, diameter, temperature, velocity, and flow rate**

- **QC Findings:**
  - **23% of records have all 5 parameters**
  - **11% have only 4 parameters**
  - **1% have only 3 parameters**
  - **<1% have only 2 parameters**
  - **3 % have only 1 parameter**
  - **28% have 0 parameters for stack emissions**
  - **34% have 0 parameters for fugitive emissions**

- **Augmentation Procedures**
  - **Fugitive releases - use lower end of parameter range**
    - **Height - 0.01 ft**
    - **Diameter - 0.01 ft**
    - **Temperature - 50 degrees F**
    - **Velocity - 0.01 ft/sec**
    - **Flow Rate - 0.01 cu ft/sec**
  - **Stack releases**
    - **Calculate missing flow rate, diameter, or velocity if only one of the 3 parameters is missing**
    - **NEI default lookup tables generated by SCC and SIC Codes - if multiple defaults available, use default record with lowest stack height**
    - **If no SCC or SIC code are available, use national default values**
      - **Height - 10 ft**
      - **Diameter - 1 ft**
      - **Temperature - 72 degrees F**
      - **Velocity - 15 ft/sec**
      - **Flow Rate - 12 cu ft/sec**

QC of Locational Data

- **Steps**
  - **Convert UTM coordinates to latitudes/longitudes**
  - **Correct obvious errors in latitudes/longitudes (transcribed coordinates, etc.)**
  - **Examine coordinates for multiple stacks within a facility**
    - **Investigate standard deviations of coordinates within a facility >0.02 degrees**
    - **Correct coordinates of individual stacks by defaulting to coordinates of another stack**
  - **Use a GIS program to plot latitude/longitude for each release point within reported county**
    - **Investigate points not within 5 km of an outside boundary of the county assumed to be valid**

- **QC Findings: 79% of 185, 540 records in emission release point table have valid lat/lons in correct county; approximately 2,000 coordinates were outside the US**

- **Augmentation Procedure**
  - **FRS unavailable**
  - **GEOCODER Program used for 14% of records (26, 378 records)**
    - **Exact Matches (47.7% of GEOCODER matches)**
    - **Near - match to single street block (1.8% of GEOCODER matches)**
    - **Zip code +2 (2.4% of GEOCODER matches)**
    - **Zip code 5 (47% of GEOCODER matches)**
    - **Zip code 3 - match to 3 digit zip code (0.51% of GEOCODER matches)**
    - **Ambig - match to one of possible multiple street segments (0.54% of GEOCODER matches)**
  - **County centroid used for 7% of records (13,014 records)**
  - **691 records do not have lat/lons because they are portable and move from county to county within a state**

Content QC/QA

- **Identify Outliers**
  - Significant changes between 96 and 99 emissions by facility, category, county, and state
  - Significant variation (high std deviation) within 99 draft data for facilities within source categories
  - Individual facility checks
    - Identify top emitters for each HAP nationwide
    - Rank each facility based on emissions nationwide
    - List top emitters for each HAP/MACT category combination
- **Correct Outliers**
  - Investigate outliers by contacting organizations who provided the data
  - Make corrections to emissions

Content QC/QA

Example of Outlier Checks

HAP	Mercury
# Facilities	24660
99 Point Source Emissions (tpy)	5.35E+06
96 Point Source Emissions (tpy)	124
Max Site %	99.9
Maximum Emissions (tpy)	5.35E+06
Minimum Emissions (tpy)	1.18E-012
Standard Deviation	3.4E+04
Range	5.35E+06
Maximum Emitter	Transamerican Waste Central Landfill, Inc.
NTI Facility ID	NTIMS1091185 (State data)

Content QC/QA

Detect Duplicate Facilities, Sites, and Emissions

- **Duplicate facilities with different NTI Unique Facility IDs**
  - **Identify facilities having different IDS, but identical lat/lon**
    - **Similar names - missing coordinate data, addresses, etc.,**
    - **Different names - change in ownership/name, use of subsidiary names, use of abbreviated names, etc.**
  - **List top emitters of each HAP and sum of 188 HAPs**
    - **Identify facilities with same reported HAP emissions**

County State	NTI Unique Facility ID	Facility Name	MACT Code	SIC Code	HAP	Emissions (tpy)	Data Source
Person NC	NTI37037	Progress Energy	1808-1 Coal Utility	4911	Mercury	0.6	MACT
Person NC	NTINC0592	Roxboro CP&L		4911	Mercury	0.6	State

- **Duplicate Sites with same NTI Unique Facility IDs**
  - **List top emitters of each HAP and sum of 188 HAPs by Data Source**

NTI Unique Facility ID	Site ID	HAP	MACT Code	SIC Code	Emissions (tpy)	Data Source
NTI002	AL123	Benzene		291	36	State
NTI002	T$246	Benzene		2911	78	TRI
NTI002	EM819	Benzene	0502		78	MACT

Content QC/QA

Correct Duplicate Facilities, Sites, and Emissions

- Evaluate HAP emissions at potential duplicate facilities and sites
- If all HAPs are duplicated, delete one facility or site.
- If only some HAPs are duplicated, delete only emission records for duplicated HAPs using Blend/Merge hierarchy for data sources

County State	NTI Unique Facility ID	Facility Name	MACT Code	SIC Code	HAP	Data Source	Resolution
Person NC	NTI37037	Progress Energy	1808-1 Coal Utility	4911	Lead, Mercury, Nickel	MACT	Delete Lead; Merge under one NTI Unique Facility ID
Person NC	NTINC0592	Roxboro CP&L		4911	Lead, Mercury, POM	State	Delete Mercury; Merge under one NTI Unique Facility ID

NTI Unique Facility ID	Site ID	HAP	MACT Code	SIC Code	Data Source	Resolution
NTI002	AL123	Benzene, Hexane, Toluene		291	State	Retain all HAPs, tag with MACT code 0502
NTI002	T$246	Benzene, Hexane, Toluene, Xylenes		2911	TRI	Delete all HAPs except Xylenes
NTI002	EM819	Benzene, Butadiene, Hexane, Toluene	0502		MACT	Delete all HAPs except Butadiene

Content QC/QA

- **Identify Incorrect/Missing MACT Codes**

  Example:

  1626: Pulp and Paper Production

  - **Used in 1996 NTI, Replaced by 1626-1 and 1626-2 in 1999**
  - **State supplied data in 1999 that did not include process level information**

  1626-1: Pulp & Paper Production - Non-Combustion

  - **MACT supplied data**
  - **State supplied data if process level details available**

  1626-2: Pulp & Paper Production - Combustion

  - **MACT supplied data**
  - **State supplied data if process level details available**

- **Revise MACT Code Assignments and Delete Duplicate Facilities**

1999 NEI Schedule

Date	Activity	Expected Result
June 1, 2001	State, local and tribal data submitted to EPA	Prescribed data transfer format used
June/July 2001	Initial QC review of submitted data	Confirmation that data file can be read; required minimal data included
October 1, 2001	Draft NEI Version 2.0 published	Draft databases available for external review
February 1, 2002	Comments on draft 1999 NEI Version 2.0 submitted to EPA	Revisions submitted in prescribed format
June 1, 2002	State, local and tribal data submitted to EPA, Version 3.0	Prescribed data transfer format used
October 1, 2002	Draft NEI Version 3.0 published	Draft databases available for external review
February 1, 2003	Comments on draft 1999 NEI Version 3.0 submitted to EPA	Revisions submitted in prescribed format
June 1, 2003	1999 NEI Version 3.0 published	Files available to support CAA and GPRA needs

1999 NEI Resources

- **1999 NEI Data Incorporation Plan**
  - www.epa.gov/ttn/chief/net/nei_plan.pdf
- **1999 NTI Q's & A's - preferences on how to report pollutant compound groups such as Cr**
  - www.epa.gov/ttn/chief/nti/ntiq&a.pdf
- **NIF Version 2.0**
  - www.epa.gov/ttn/chief/eidocs
- **QA/QC NIF Format Tool**
  - www.epa.gov/ttn/chief/nif/
- **NEI QA and Augmentation Memo**
  - www.epa.gov/ttn/chief/emch/invent/
- **Draft 1999 NEI**
  - Data Files in NIF Version 2.0 format
  - Summary Data Files
  - Documentation
  - README
  - ftp://ftp.epa.gov/EmisInventory/
- **1996 Data Summary on AIR Data**
  - www.epa.gov.air/data
- **NEI Users Group and Listserve**
  - http://www.epa.gov/ttn/chief/listserv.html#nei