

Compiling the Point Source 2002 NEI: Whose Data Get Priority?

Anne Pope, US EPA

pope.anne@epa.gov

Stephanie Finn, ERG, Inc.

stephanie.finn@erg.com

Darcy Wilson, ERG, Inc.

darcy.wilson@erg.com



Why is the NEI Important?

We are preparing the 2002 NEI to meet several specific needs, including:

- Key input to regional/national modeling by EPA, RPO's, S/L/T's, etc.
- Basis for National Air Toxics Assessment (NATA) analyses
- Starting point for rule development (residual risk)
- Trends and GPRA tracking
- Public information
- International Reporting

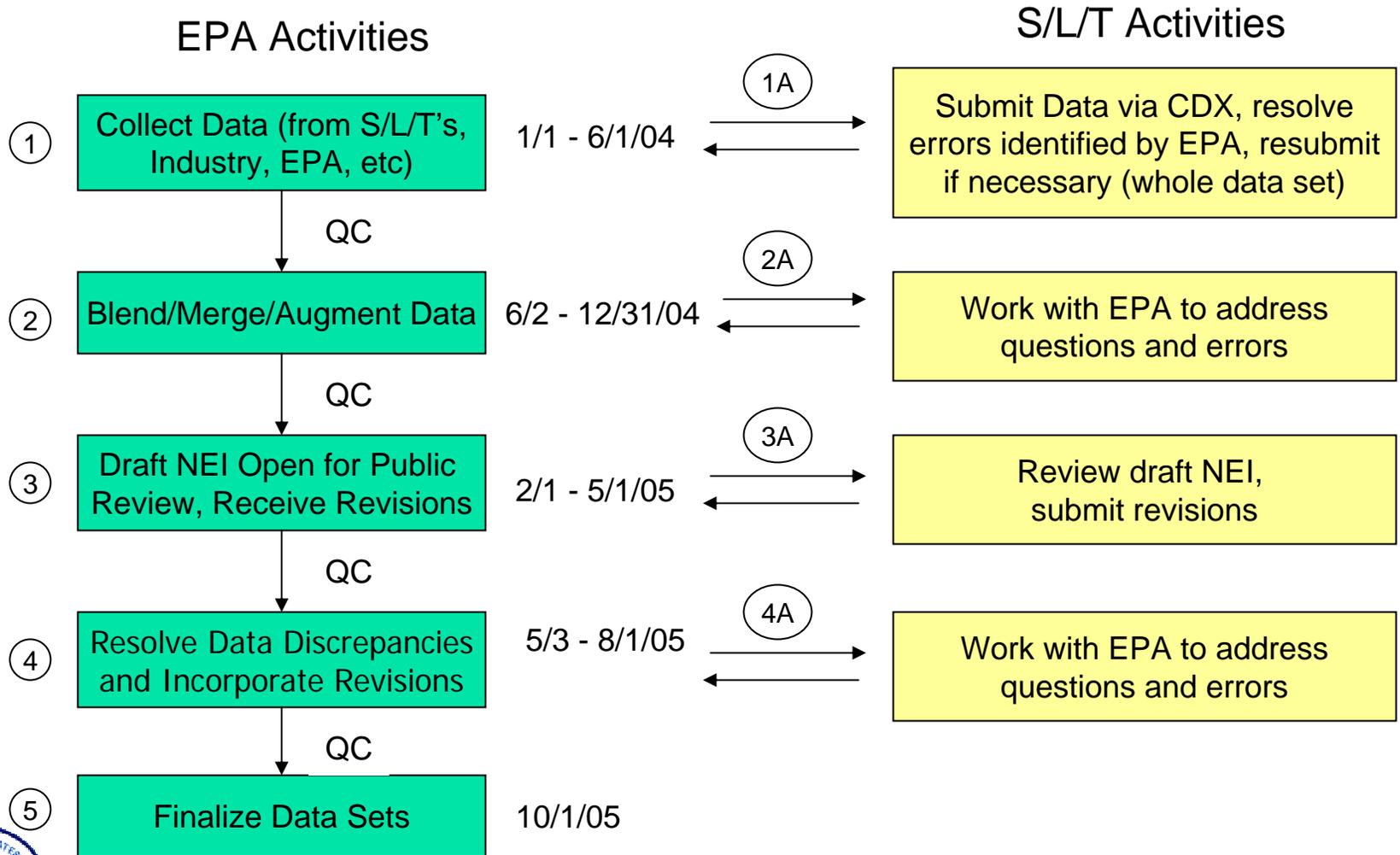


Goals of 2002 NEI

- Make efficient use of multiple data resources
- Integrate HAPs and CAPs data
- Use updated input formats (NIF 3.0 & XML Schema)
- Provide more feedback to S/L/T earlier on quality of data submitted
- Improve quality of data in the 2002 NEI
- Peer Review methodology and final NEI product



2002 NEI Schedule & Activities



Providers of 2002 NEI Point Source Data

- State Agencies (S)
- Local Agencies (L)
- Tribes (B)
- US EPA
 - Emission Standards Division (ESD) MACT, Residual Risk, and new area source data (P, M)
 - Clean Air Markets Division (CAMD) electric generating unit data (E)
 - Toxic Release data (T)
 - Augmented data (A)

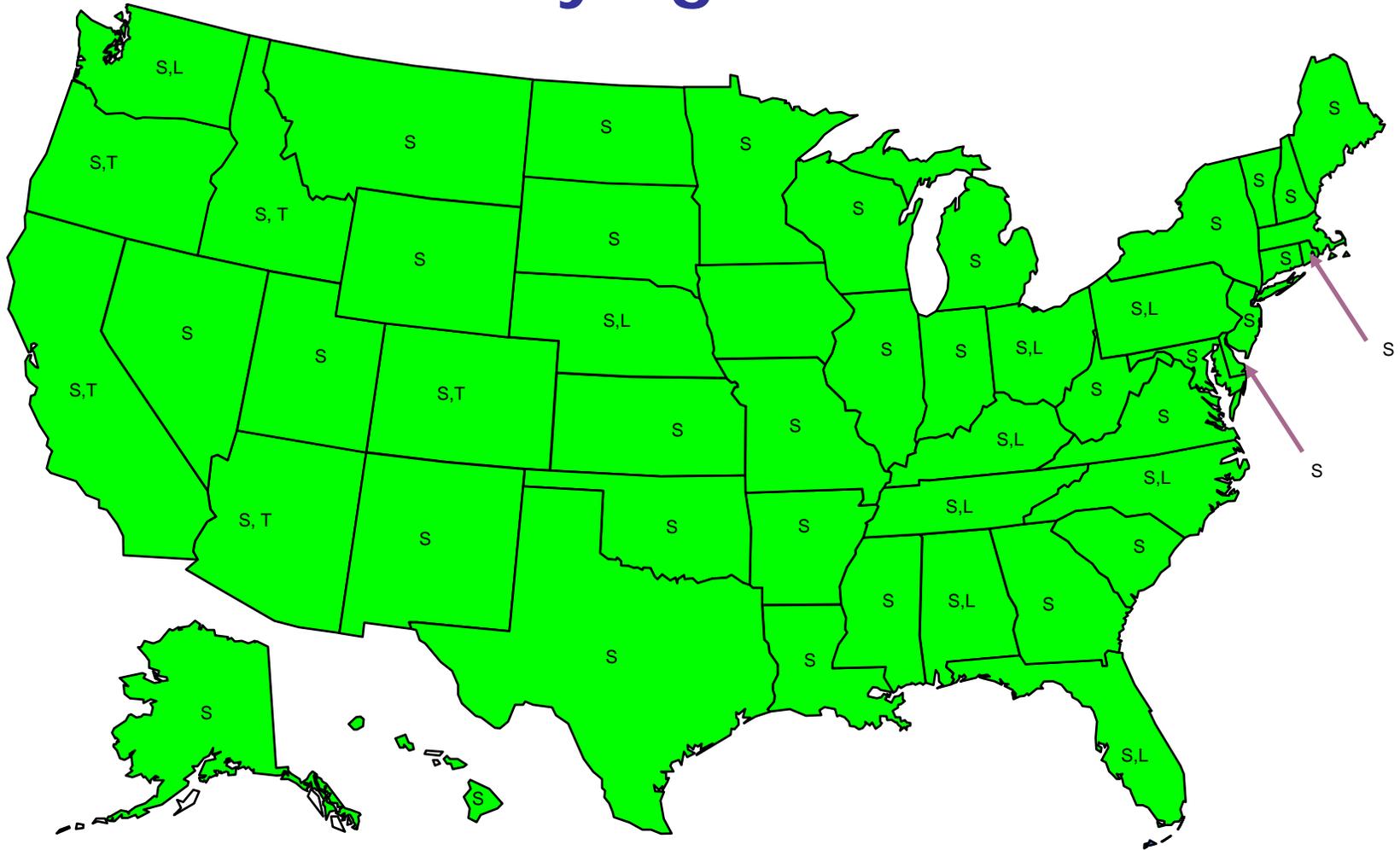


Hierarchy of Data Selected

1. Preferred ESD data (P) for MWCs, MWIs, Coke Ovens, and Brick Manufacturing
CAMD CEM data (E) for NOx and Sulfur Oxides
2. Tribal Agency data (B)
3. Local Agency data (L)
4. State Agency data (S)
5. Other non-preferred ESD data (M) and other non-preferred CAMD data (E)
6. TRI data (T)
7. Augmented data (A) for PM filterable and condensable, HAPs from VOC and PM, and CAP from HAPs



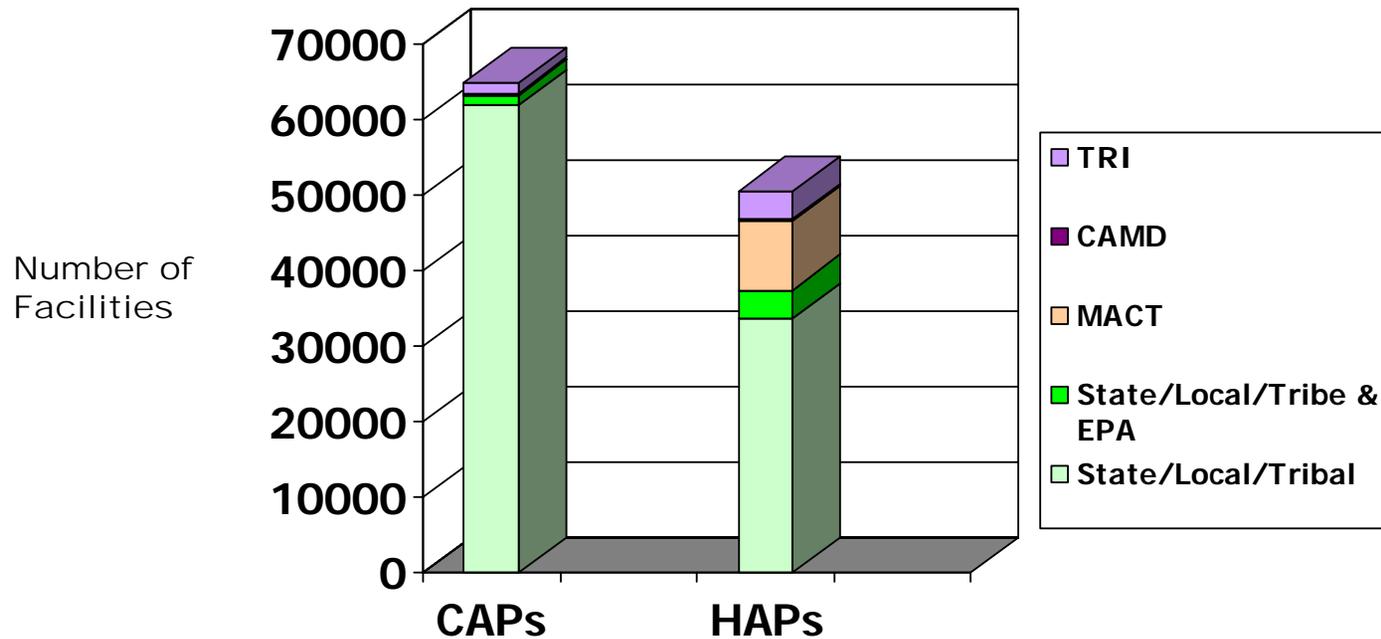
2002 NEI Point Source Data Provided by Agencies



T - tribes
L - local agencies
S - state agencies

HAP data not provided by AK, state of AZ, DC
GA, HI, ND, and NV for the draft.

Data Source of Facilities



CAPs: State/Local/Tribal data only used for 97% of Sources
HAPs: State/Local/Tribal data only used for 62% of Sources

Data Source of Facilities

- **CAPs:**
 - State. Local and Tribal data used for 97% of facilities
 - Mixture of State, Local and Tribal data with EGU and TRI for remaining 3%

- **HAPs:**
 - State. Local and Tribal data only used for 62% of facilities
 - MACT data only used for 17% of facilities
 - TRI data only used for 14% of facilities
 - Mixture of State, Local and Tribal data with MACT, EGU and TRI for remaining 7%



Blend Merge Process in 2002

1. **Match Facilities from multiple data sources**

2. **Data Selection**
 - **Round One: HAP and nonEGU CAP**
 - Selection Pass 1 – pollutant code
 - Selection Pass 2 – pollutant category
 - **Round Two: EGU CAP and HAP**
 - Selection Pass 1 – NO_x and SO₂
 - Selection Pass 2 – all other pollutants



Assignment of NEI Unique Facility ID to Site Records

- **NEI Unique Facility ID assigned to site records submitted by multiple data sources**
 - Integrates HAP and CAP IDs for same facility
 - Integrates multiple site IDs submitted by different data providers for same facility

- **2002 NEI Facility File contains crosswalk of NEI HAP and CAP facilities**
 - Site IDs, NEI Unique Facility ID, FRS ID, TRI ID, ORIS ID
 - Standardized Facility Name and Address
 - Site Latitude and Longitude
 - SIC/NAICS Code
 - Information on whether facility contained 99 HAP, 99 CAP, 2002 HAP or 2002 CAP data

<ftp://ftp.epa.gov/EmisInventory/draftnei2002/point/summaries>



Assignment of NEI Unique Facility ID to Site Records

State County FIPS	Tribal Code	State/Local Facility ID	NEI Unique Facility ID	Facility Name	Data Source
42003	000	4200300157	NEI13183	Orion Power Midwest Dallas Station	State
42003	000	EGU0978	NEI13183	Dallas Power Plant	CAMD
42003	000	T\$1432	NEI13183	Orion Dallas Plant	TRI
42003	000	ESD420031	NEI13183	Orion Power Dallas Plant	Preferred MACT for Utility Coal mercury



Data Selection: Round One HAPs and non-EGU CAPs

- **Selection Pass 1: pollutant code**
 - Uses hierarchy of ranked data sources

Data Source Code	Definition
P	Preferred MACT Data – MWCs, MWIs, Brick and Structural Clay Products, Coke Ovens, Mercury from Utility Coal
B	Tribal Data
L	Local Agency Data
S	State Agency Data
M	MACT Data – non-preferred categories
T	TRI



Data Selection: Round One

HAPs and non-EGU CAPs

- Selection Pass 2: pollutant category
 - **HAPs:** CAA includes 20 compound groups, but speciated compounds are needed for risk assessments. Data providers submit both compound groups and speciated data for same facility
 - **CAPs:** looks at total PM reported by different data sources and does not choose between individual fractions
 - Pass 2 looks for duplication at the pollutant category level at a facility so that only one data source for the pollutant group and its speciated compounds is selected
 - Results of both passes are evaluated and conflicts are resolved



Data Selection: Round One

HAPs and non-EGU CAPs

NEI Unique Facility ID	Pollutant Code	HAP Name	HAP Category	Emissions (TPY)	Pass 1	Pass 2	Result: Pass 1 & 2	Data Source
NEI 18334	195	Lead and Compounds	Lead and Compounds	0.125	select	de-select	de-select	M
NEI 18334	1335257	Lead Oxide	Lead and Compounds	0.0395	select	select	select	S
NEI10585	18540299	Chromium (VI)	Chromium and Compounds	0.005	select	select	select	S
NEI10585	16065831	Chromium III	Chromium and Compounds	0.04	select	de-select	select	M



Data Selection: Round Two

EGU CAPs and HAPs

- Matching for EGUs occurs by pollutant and SCC within a facility, but not at the unit level
- EPA EGU data sources:
 - CAMD CEM data for NO_x and SO₂ and heat input values
 - DOE Form EIA-767 data for fuel use and AP-42 emission factors
- Selection Pass 1: NO_x and SO₂
 - Use CEM data for electric generation, industrial and commercial-institutional fuel combustion
 - SCCs 101*, 102*, 201*. 202*. 103*. And 203*
- Selection Pass 2 – all other CAPs and all HAPs except mercury

Data Selection Rules

- CAPs handled on category basis only
- Hierarchy for utility boilers
 1. State Agency, Local Agency, and Tribal data
 2. EPA EGU data
 3. TRI data
- Non-boiler SCC data for EGU facilities
 - State Agency/Local Agency/Tribal data retained



Data Selection: Round Two

EGU Example

NEI Unique Facility ID	SCC	Pollutant	EPA EGU Emissions (TPY)	S/L/B Emissions (TPY)	EPA EGU Data Selection	S/L/B Data Selection
NEI101	101*	NOx	100	105	Select	De-select
NEI101	101*	SO2	200	203	Select	De-select
NEI101	101*	VOC	50	55	De-select	Select
NEI101	101*	Cadmium	0.002	0.003	De-select	Select
NEI101	301*	NOx	Null	20	N/A	Select
NEI101	301*	PM	Null	20	N/A	Select
NEI101	301*	SO2	Null	25	N/A	select



Data Selection: Benefits

- **Data selection from different sources contributes to more complete inventory.**
- **Benefits of Using NEI Unique Facility ID and Retaining IDs reported by Data Submitters**
 - Easy Identification of data reported by different data providers: Users can identify source of data for each data point because local IDs are retained in the StateFacility ID, Emission Unit ID, Process ID and Emission Release Point ID fields
 - More transparency of Data Reported to EPA: Comparisons with earlier datasets are easier when local IDs are not changed resulting in more transparency in the NEI
 - Reduced Costs: A more complex data selection process involving unit, process and or emission release point level matching is more difficult and would require more resources.



Data Selection: Issues

- **Complete facility estimates are obtained in NEI by using the NEI Unique Facility ID**
 - Emissions are not duplicated
 - Different data sources use different site, unit, process and emission release point IDs
 - Query on NEI Unique Facility ID to obtain facility estimate

- **Data Selection Occurs at the Facility Level**
 - Units appear to have more stacks than exist
 - Emissions are not duplicated
 - This may be an issue for modeling if stacks reported in different data sets have different locations or stack parameters
 - Some non-duplicative HAP estimates are eliminated
 - SCCs are not mandatory for HAPs
 - Some agencies and TRI only report data at the facility level for HAPs
 - Reviewers can add back in HAPs that were eliminated in the draft
 - Inadvertent supplementation of estimates from shutdowns may occur.
 - Data submitters should submit zero emissions for recent or temporary shutdowns
 - Reviewers can delete emissions for shutdowns.



What is Best Way to Review Selection Process in the Draft NEI?

1. **Review Sites Table in state NOF files and the 2002 NEI Facility File to identify duplicate facilities**
 - Look for sites that appear to be identical facilities that do **NOT** have the same NEI Unique Facility ID
 - Mark duplicate sites with a "D" for delete

2. **Review emission estimates for facilities using the NEI Unique Facility ID using the ALL NEI files .**
 - Look at data source code for each record.
 - Review pollutants and emissions for each process.
 - If a unit has shutdown, mark pollutants for deletion.
 - If the emissions need to be revised, provide revised values.

3. **Use the Historical Emissions Report to compare estimates from multiple data sources and to see what value was selected in the draft.**
 - Use the NEI Unique Facility ID to identify facilities
 - Identify pollutants of interest.
 - Compare emissions and data source of Current Data Source Reported to other values provided by different data sources
 - Calculate percent difference for selected pollutants and facilities to identify facilities to investigate further.
 - Provide comments to EPA for proposed revisions.



Historical Emissions Report

State County FIPS	Tribal Code	Facility Name	NEI Unique Facility ID	Pollutant Code	Pollutant Name	Curr Versio n Date	Curr Emiss (TPY)	Curr Data Source	Curr S (TPY)	Curr P (TPY)	Curr M (TPY)	CurrT (TPY)
01001	000	ABC	NEI008	100425	Styrene	2/1/ 2005	2.21	S	2.21	1.82		4.2
01001	000	ABC	NEI008	107028	Acrolein	2/1/ 2005	6.28	P	6.14	6.28		5.3
01001	000	ABC	NEI008	195	Lead and Compounds	2/1/ 2005	0.2	S	0.2		0.8	
01001	000	ABC	NEI008	75070	Acetaldehyde	2/1/ 2005	38.83	T				38.83



TRI Facilities in the NEI

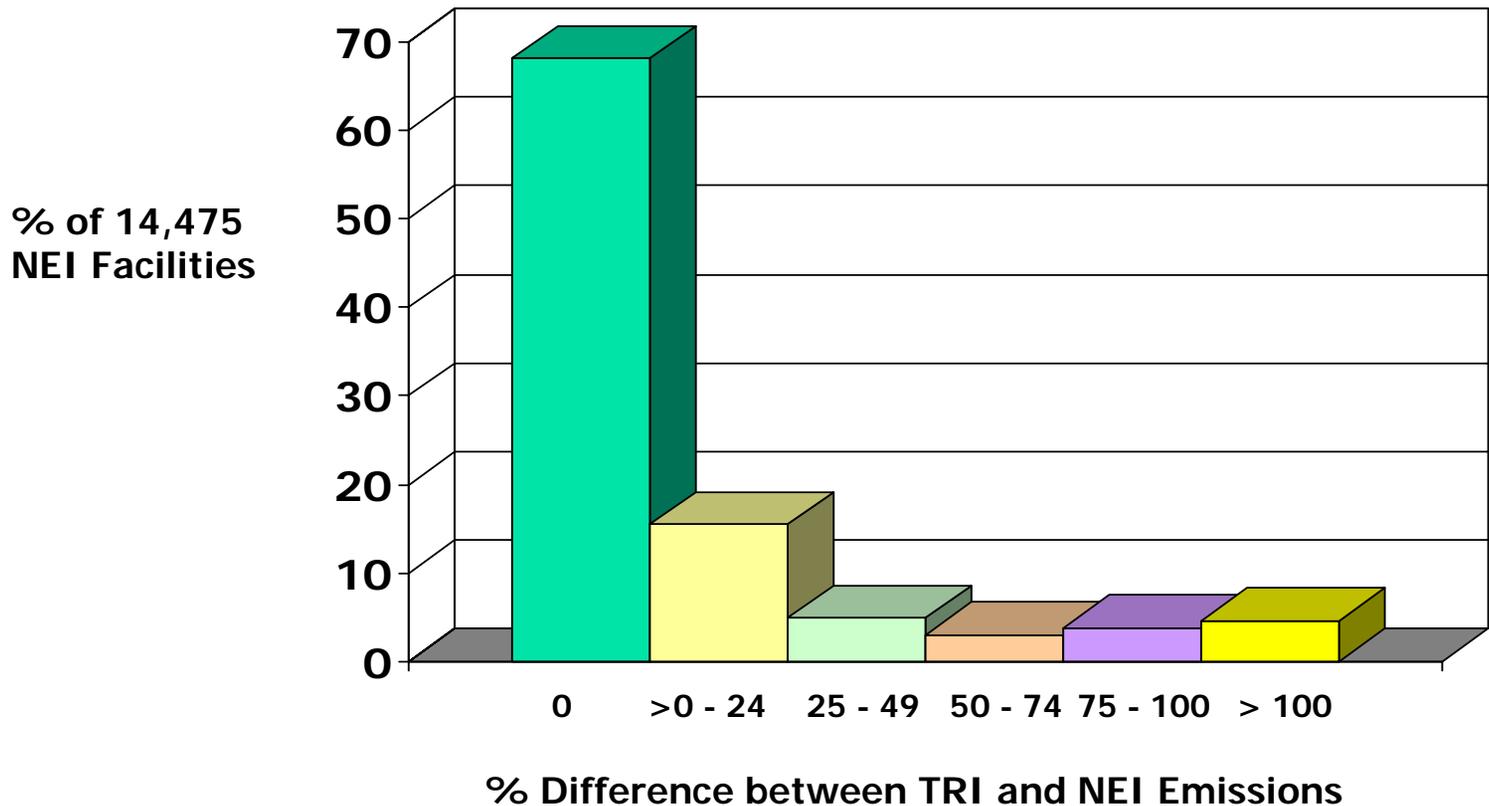
- 25,000 facilities in 2002 TRI have air emissions
- 15,515 TRI facilities emit HAPs or ammonia and are in 2002 NEI
- 10% of facilities in 2002 NEI have TRI as the only data source;
15% of facilities have TRI data

Facilities	Total Number of 2002 Facilities	Number of CAP Facilities	Number of HAP Facilities
NEI Facilities	86,3220	64,438	54,733
TRI Facilities in NEI	15,515	9261	14475
TRI Facilities in the NEI with TRI as only data source		876	7908
TRI Facilities with TRI and other NEI data sources		554	3024



Data Selection: Comparison of TRI Data to NEI HAP Data

- 15,515 TRI Facilities in 2002 NEI
14,475 TRI Facilities in 2002 NEI HAPs



Conclusion: The Main Ingredient of EI Development is Courage

The Zen of Inventories

Inventories are never right,
and never completed.

The more you use an inventory,
the more accurate it
becomes.



Joann Held
NJ DEP
April 15, 2004