Truth or Dare: Data Augmentation in the Point Source 2002 NEI

Anne Pope, US EPA
pope.anne@epa.gov
Madeleine Strum, US EPA
Stephanie Finn, ERG, Inc.

Why is the NEI Important?

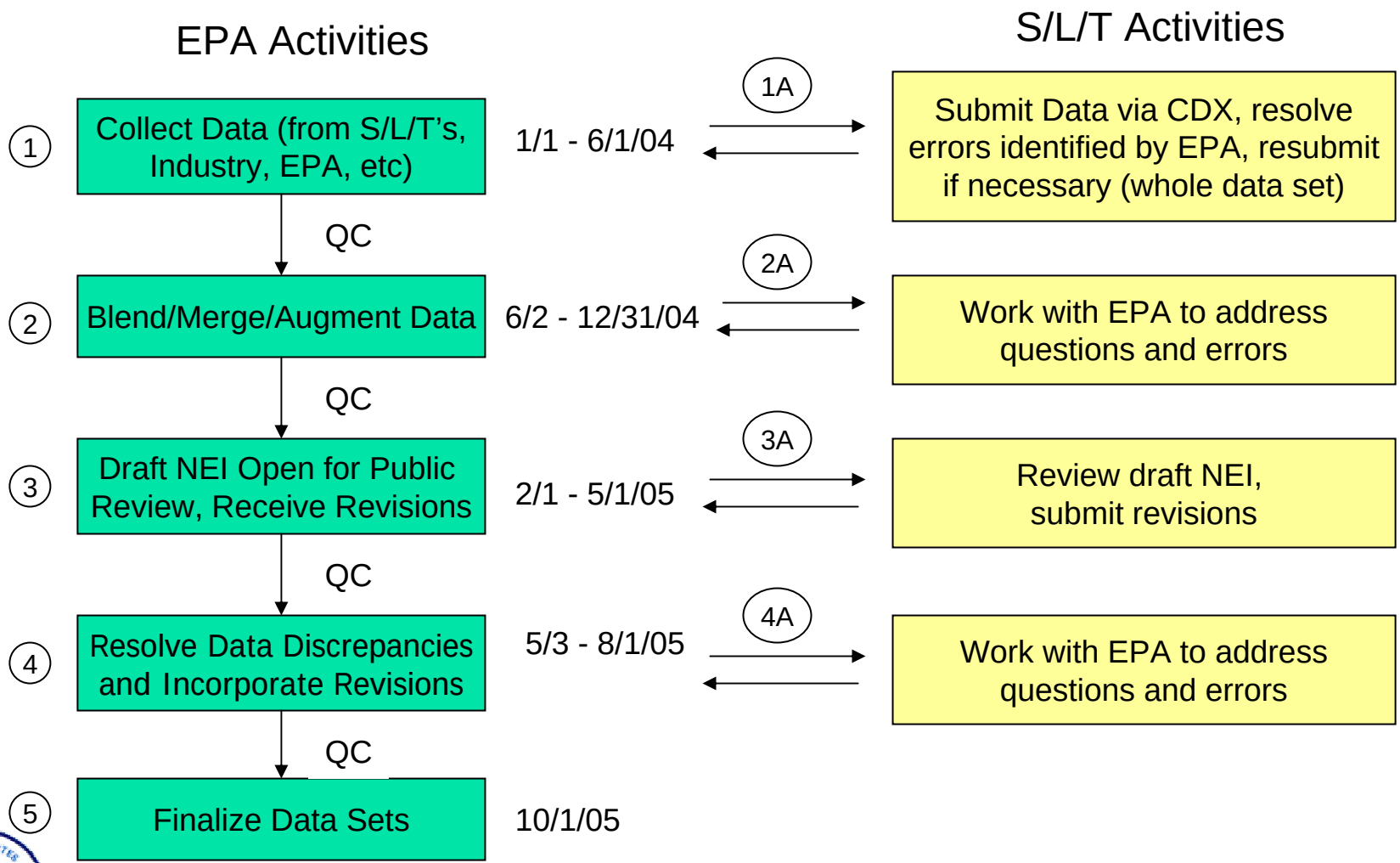
We are preparing the 2002 NEI to meet several specific needs, including:

- 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



Data Augmentation

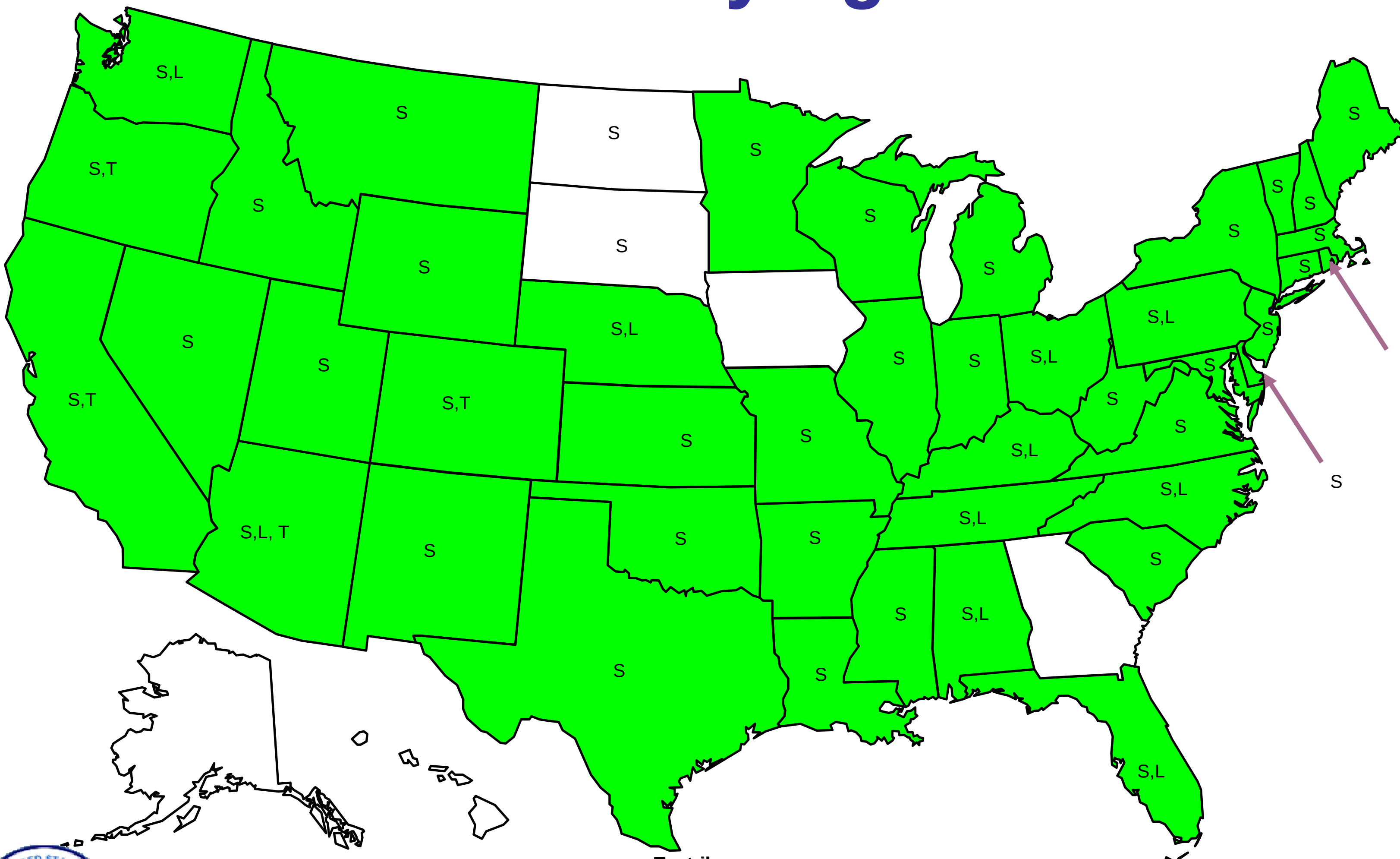
2002 NEI QA and Data Augmentation for Point Sources

www.epa.gov/ttn/chief/emch/invent/

QA and Augmentation Procedures for 115 NIF point source data fields and additional 23 data fields in NEI that are not in NIF

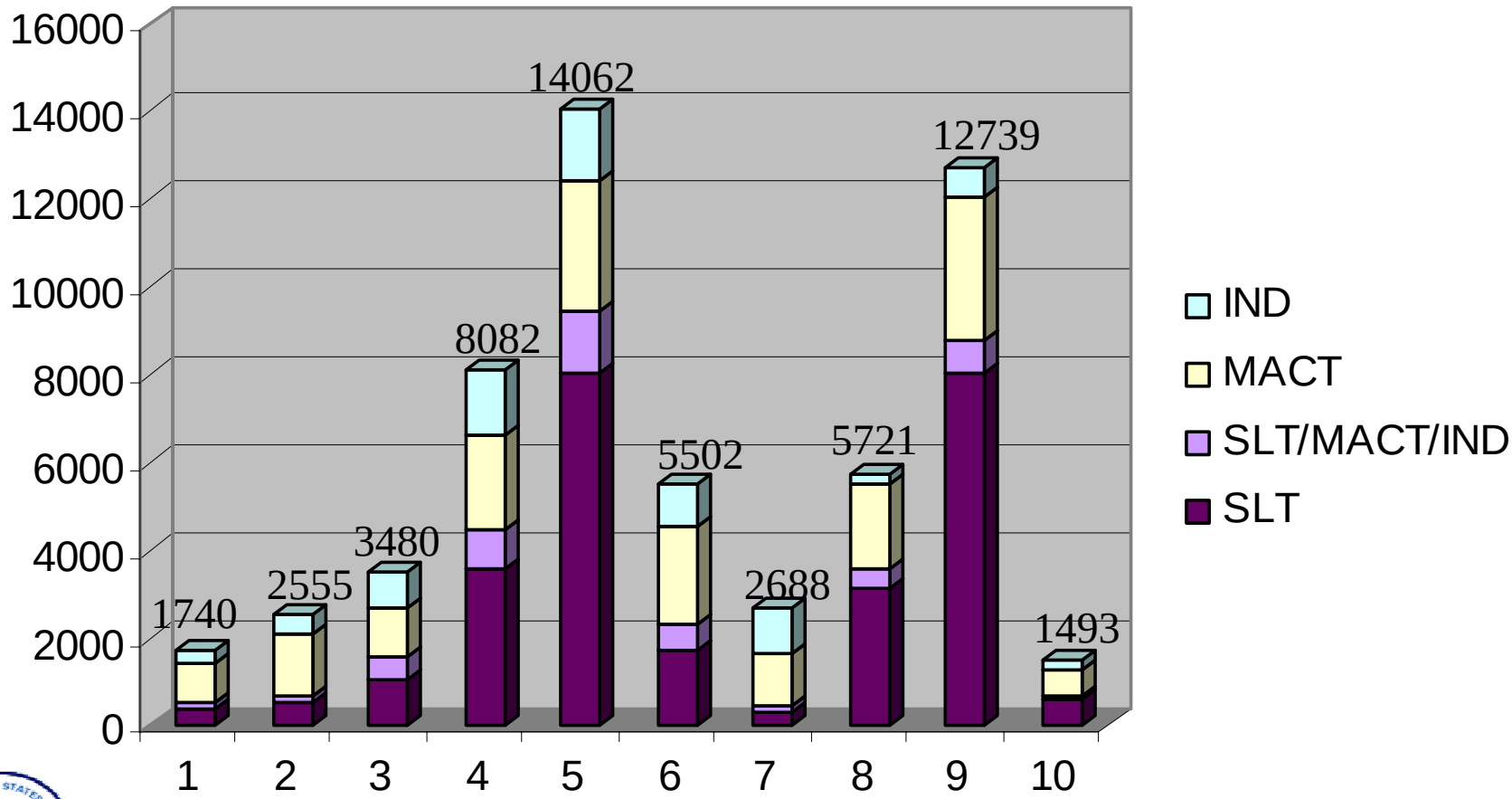
- This presentation focuses on
  - Location Coordinates (ER Table)
  - Stack Parameters (ER Table)
  - HAP Pollutant Codes (EM, CE Tables)

1999 NEI for HAPs Point Source Data Provided by Agencies

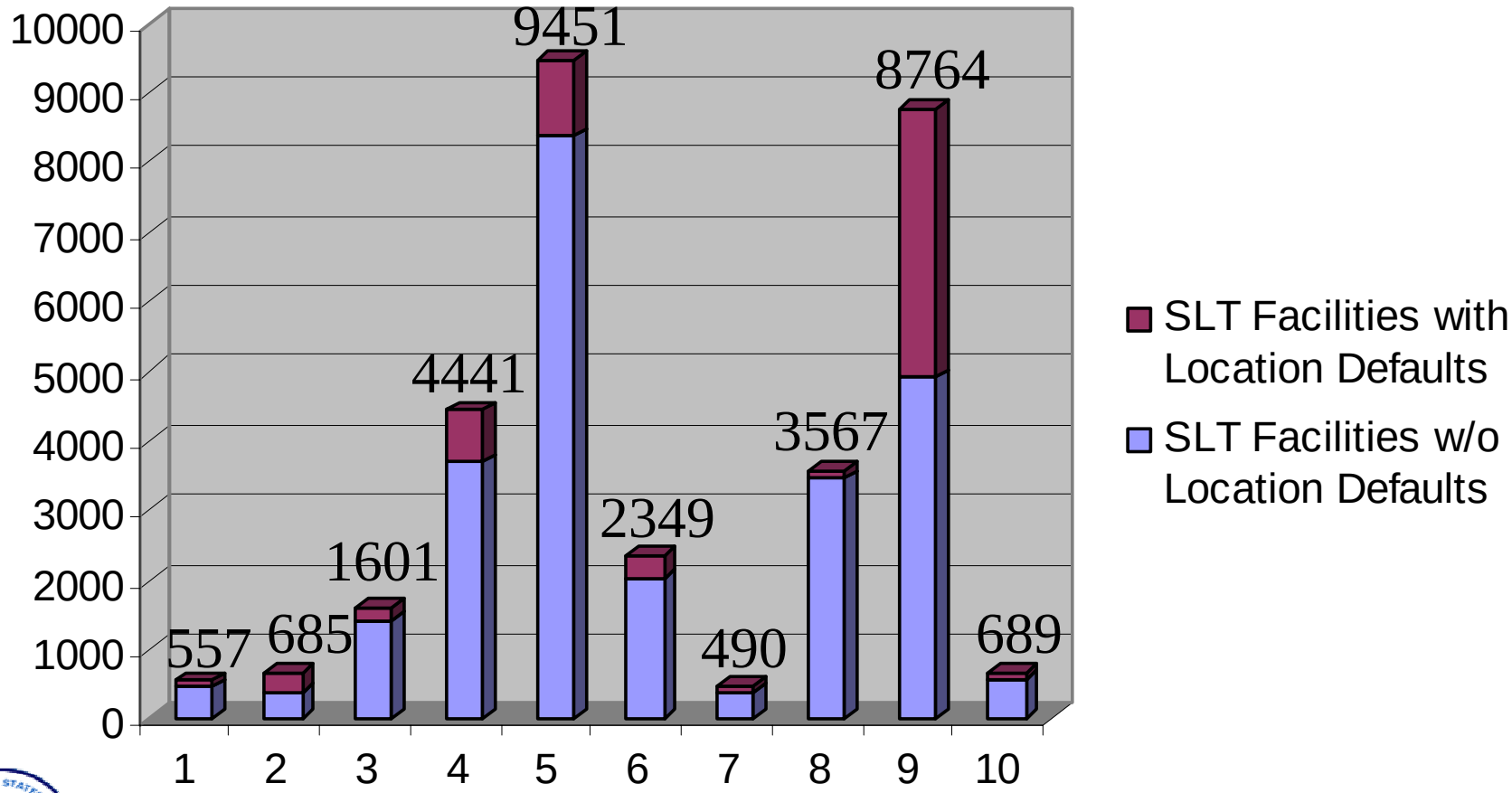


T - tribes
L - local agencies
S - state agencies

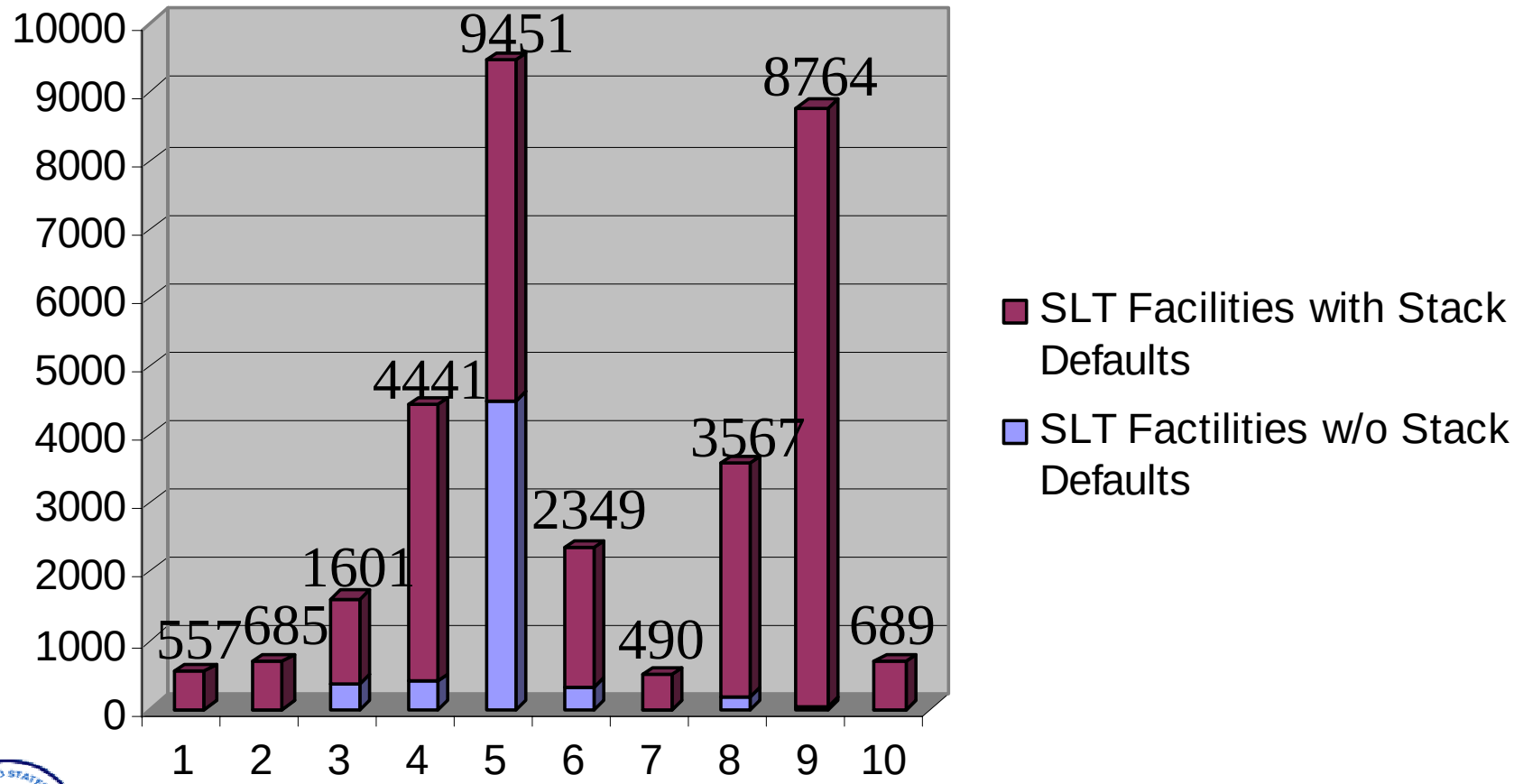
Regions: HAP Facility Count



Regions: SLT HAP Facilities with Location Defaults



Regions: SLT HAP Facilities with Stack Defaults



How Do I QA & Augment Location Coordinates?

1. Convert coordinates to correct units if necessary

2. Verify all emission release points are within 3 km of one another

3. Make sure coordinates are in correct county

4. Replace bad/missing coordinates

5. Assign location coordinate default flags

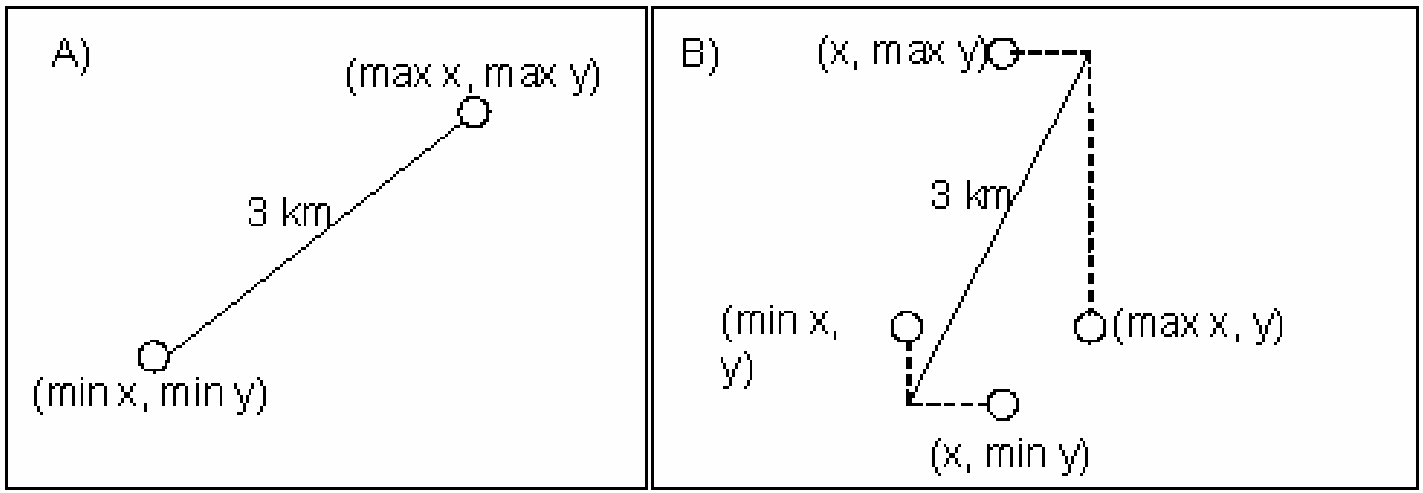
How Do I QA & Augment Location Coordinates?

1. Convert coordinates to correct units if necessary

- QA coordinates reported as UTM and convert to Latitude/Longitude
- QA reported Latitude and Longitude and convert to decimal degrees if reported in degrees-minutes-seconds
- Proceed to Step 2

How Do I QA & Augment Location Coordinates?

2. Verify all emission release points are within 3 km of one another



How Do I QA & Augment Location Coordinates?

2. Verify all emission release points are within 3 km of one another. If not, then:

- Determine if points can be > 3 km for source category; if acceptable proceed to Step 3
- If distance between points should not be > 3km, then:
  - Calculate distances between all points within a facility and identify outliers
  - Use average Site Latitude and Longitude calculated from valid points to replace outliers
  - Proceed to Step 3

How Do I QA & Augment Location Coordinates?

3. Make sure coordinates are in correct county

- Use GIS overlay to plot each latitude/longitude pair
- Verify plotted release point is within 10 km of the county.
- If plotted release point is found inside 10 km of county, process is complete.
- If plotted release point is outside of 10 km of the county or missing, then proceed to Step 4.

How Do I QA & Augment Location Coordinates?

4. Replace bad/missing coordinates using the hierarchy below.

A. Use Facility Specific Data
   - Average Site Latitude and Longitude calculated from valid points

B. Use Geocoding software
   - First QA and correct Zip Codes
   - Then use Geocoder

C. Use NEI Historical Facility Table

D. Assign Site Release Point at County Centroid

How Do I QA & Augment Location Coordinates?

4. Assign location coordinate default flags

Default Code	Description
Exact	Match is to within a unique intersection or within a single side of a single street block.
Near	Match is to a single street block but the correct placement within block is unknown.
Zip code+2	Match to a 5-digit zip code, plus the first two digits of the 4-digit extension.
Zipcode5	Match to a 5-digit zip code
Zipcode3	Match to multiple 3-digit zip codes based on postal service Sectional Center Facility (SCF).
Ambig	Match is to multiple street segments
Cntycent	County centroid.
NEI Fac Table	Coordinate found in the NEI Historical Facility Table
Site-Avg.	Average of accurate coordinates of other emission release points at the same site

Latitude/Longitude EPA Data Standards

- **Latitude/Longitude**
- **Horizontal Collection Method Code -** Method to determine coordinates
  - Default with Geocoded coordinates or nonGeocoded as “027 “ (Unknown)
- **Horizontal Reference Datum Code -** Reference datum for coordinates
  - Only default Geocoded coordinates
- **Horizontal Accuracy Measure -** Measure of accuracy of coordinates
  - Only default Geocoded coordinates
- **Coordinate Data Source Code** – Responsible party for coordinates
  - Default with Geocoded coordinates or nonGeocoded to Source of Data for NIF
- **Reference Point Code** – Place coordinates represent
  - Default to “108” (Points not represented by other codes)
- **Source Map Scale Number** – Map Scale
  - Only default Geocoded coordinates

How Do I QA & Augment Stack Parameters?

1. Evaluate and replace missing and invalid emission release point type

2. Evaluate and replace missing and invalid parameters associated with fugitive emissions release points

3. Evaluate and replace missing, invalid, or inconsistent parameters associated with non-fugitive release points

4. Assign stack parameter default flags

How Do I QA & Augment Stack Parameters?

1. Evaluate and replace missing and invalid emission release point type

- **Emission Release Point Types**
  - Fugitive
  - Non-Fugitive (stack):
    Vertical, Horizontal, Goose Neck,
    Vertical with Rain Cap, Downward Facing vent
- **QA using SCC/Emission Release Point Type Crosswalk**
- **Replace invalid emission release point types**
  - NIF Code 01 for fugitive
  - NIF Code 02 for non-fugitive
- **Proceed to Step 2 if Emission Release Point Type is fugitive and to Step 3 if Emission Release Point Type is non-fugitive**

How Do I QA & Augment Stack Parameters?

2. Evaluate and replace missing and invalid parameters associated with fugitive emissions release points

- **Verify Stack Height is between 0.1 and 100 ft.**
- **If Height is valid, replace remaining parameters with defaults. Proceed to Step 4.**
- **If Height is invalid, replace all parameters with defaults. Proceed to Step 4**

Non-Fugitive Stack Parameter Defaults:

Height:	10 ft
Temperature:	72 ˚F
Diameter:	0.003 ft
Velocity:	0.0003 ft/sec
Flow:	0 ft3/sec

How Do I QA & Augment Stack Parameters?

3. Evaluate and replace missing, invalid, or inconsistent parameters associated with non-fugitive release points

A. Compare parameters to the following min/max values:

Height:	0.1 – 1200 ft
Temperature*:	50 – 1800 ˚F
Diameter:	0.1 – 50 ft
Velocity:	0.1 – 100 ft/sec

Verify Temperature is greater than 250 ˚F for boilers and incinerators using SCCs and MACT Codes

B. Determine if out of range value is acceptable for source category.

C. If all parameters are acceptable, proceed to Step D. If any parameter is missing or out of range, proceed to Step E.

How Do I QA & Augment Stack Parameters?

3. Evaluate and replace missing, invalid, or inconsistent parameters associated with non-fugitive release points

D. Evaluate internal consistency between parameters

- **Stack Height < Stack Diameter.**

  If Stack Height > Stack Diameter, replace all 5 parameters. Proceed to Step E.

  If Stack Diameter< Stack Height, evaluate Flow Rate Consistency.

- **Reported Flow Rate and Calculated Flow Rate are within 10% of each other**

  Flow rate =((Pi)*(Diameter/2))^2 *Velocity

If internal consistency is met, process is complete.
If internal consistency is not met, proceed to Step E.

How Do I QA & Augment Stack Parameters?

3. Evaluate and replace missing, invalid, or inconsistent parameters associated with non-fugitive release points

E. Replace missing, invalid or inconsistent parameters using following hierarchy. Proceed to Step 4.

- **Calculated flow rate or velocity.**
- **SCC default unless SCC is 39999999**
- **MACT Code default**
- **SIC Code default**
- **National default**

Height: 10 ft
Temperature: 72 ˚F
Diameter: 1 ft
Velocity: 15 ft/sec
Flow: 12 ft3/sec

How Do I QA & Augment Stack Parameters?

4. Assign stack parameter default flags

0 = Original value (not a default)
1 = SCC default
2 = SIC Code default
3 = National default
4 = Calculated value
5 = MACT Code default

How Do I QA & Augment HAP Pollutant Codes?

We encourage data providers to report emissions of individual compounds and identify HAP compound by CAS number

1. QA reported Pollutant Codes and contact provider if incorrect
2. Assign Chemical Identification Standard data elements

   www.epa.gov/srs

   - Chemical Abstracts Service Registry Number
   - Chemical Substance Systematic Name
   - EPA Chemical Identifier
   - EPA Chemical Registry Name

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