NEI Input Format (NIF)

Quality Assurance (QA) Software

International Emission Inventory Conference
Training Course
April 15, 2002

Rhonda Thompson

OAQPS - EFIG
What does the QA software do?

FORMAT Checks

CONTENT Checks

MS Access Database
What kind of Format Checks?

NEI Input Format (NIF) Checks

Tables - Existence

Fields - Properties

Referential Integrity - how the tables relate
Tables - Point (8), Area (5) and Mobile (3)

Point Tables:

- Transmittal (TR)
- Site (SI)
- Emission Unit (EU)
- Emission Release Point (ER)
- Emission Process (EP)
- Control Equipment (CE)
- Emission Period (PE)
- Emission (EM)
Field Properties:

Name - EXACT

Data Type - Text, Numeric types

Length - Maximum for each field

Mandatory - All Primary Key (PK) Fields plus
Referential Integrity - How do the tables relate?

Primary Key (PK) Fields - Unique

One to Many Relationship* - Hierarchy

*Most tables have 1-many but there are exceptions
Primary Key Fields - Make Each Record Unique

TR - State, County
There is no more than one record for each county in the TR table.

SI - State, County, Site ID
There is no more than one record for each site in the SI table.
<table>
<thead>
<tr>
<th>Table Code</th>
<th>Primary Key Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>State, County</td>
</tr>
<tr>
<td>SI</td>
<td>State, County, Site ID</td>
</tr>
<tr>
<td>EU</td>
<td>State, County, Site ID, Unit ID</td>
</tr>
<tr>
<td>EP</td>
<td>State, County, Site ID, Unit ID, Process ID</td>
</tr>
<tr>
<td>PE</td>
<td>State, County, Site ID, Unit ID, Process ID, Start and End Dates</td>
</tr>
<tr>
<td>EM</td>
<td>State, County, Site ID, Unit ID, Process ID, Pollutant, EmissRelPtID, Start and End Dates</td>
</tr>
<tr>
<td>ER</td>
<td>State, County, Site ID, EmissRelPt ID</td>
</tr>
<tr>
<td>CE</td>
<td>State, County, Site ID, Unit ID, Process ID, pollutant</td>
</tr>
</tbody>
</table>
One to Many* Relationship - Hierarchy

For every unique record in the parent table, there must be at least one record in the child table.

Most tables have 1-many but there are exceptions.

For every county in the TR table, there must be at least one site in the SI table.

No record is necessary for a county with no plants.

For every site in the SI table, there must be a county record in the TR table.

*Most tables have 1-many but there are exceptions.
One to Many Relationship - Hierarchy

TR
  ↓
SI
  ↓
EU
  ↓
EP
  ↓
PE
  ↓
EM
  ↓
CE*

*Most tables have 1-many but there are exceptions

IF an EM record was controlled, there will be a CE record
Integrity Violations- How do the tables NOT relate?

Duplicates - More than one record for the same PKs even though the other fields have different info

2 records for the same county even though contact info is different.

Orphans - a record in the child table without a corresponding record in the parent table(s)

Emission records in the EM table but no Records for that Site in the SI table

Widows (childless parents) - a record in the parent table with no records in the child table(s)

A County in the TR table with no emissions in the EM table
What are Content Checks?

Acceptable Codes - NIF Code Table Database
- QA/QC Code Table Database

Numeric Ranges - Maximum and Minimum

Geographic Data - Coordinate checks
# Code Table Database

A Table for each coded field

## Tables in the Code Table Database:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>address types</td>
<td>Materials</td>
<td>Unit Codes*</td>
</tr>
<tr>
<td>contact type</td>
<td>Material I/O</td>
<td>XY coordinate codes</td>
</tr>
<tr>
<td>control device types</td>
<td>Max Emission Rate</td>
<td></td>
</tr>
<tr>
<td>county fips</td>
<td>NAICS Codes</td>
<td></td>
</tr>
<tr>
<td>county boundary</td>
<td>Pollutant Codes</td>
<td></td>
</tr>
<tr>
<td>emiss_calculat_meth</td>
<td>Reliability Indicator</td>
<td></td>
</tr>
<tr>
<td>emission release</td>
<td>Rule Effectiveness Method</td>
<td></td>
</tr>
<tr>
<td>emission type</td>
<td>SCC</td>
<td></td>
</tr>
<tr>
<td>emission unit</td>
<td>SIC</td>
<td></td>
</tr>
<tr>
<td>facility category</td>
<td>Source Type</td>
<td></td>
</tr>
<tr>
<td>inventory type</td>
<td>Stack Parameters</td>
<td></td>
</tr>
<tr>
<td>MACT codes</td>
<td>state fips</td>
<td></td>
</tr>
<tr>
<td>MACT Compliance Status</td>
<td>transaction type</td>
<td></td>
</tr>
</tbody>
</table>

* used in several fields
Coded Fields By Table *:

**TR**  
Transaction, Inventory, Source, and Contact Types

**SI**  
Facility Category, Address Type, Site MACT and MACT Compliance

**EU**  
Design Capacity Unit Numerator and Denominator

**ER**  
Emission Release Point and XY Coordinate Types, Fugitive Dim. Unit

**EP**  
SCC, Process MACT and MACT Compliance Codes

**CE**  
Pollutant, Control Device Types

**PE**  
Throughput Unit Numerator, Material, Material I/O

**EM**  
Emission Unit Numerator, Emission Type, Factor Unit (num/den), Material, Material I/O, Emission Calculation Method, EF Reliability Indicator, Rule Effect Meth, Emiss Data Level

*(State and County FIPS in every Table)*
TR Table Coded Fields

State and County FIPS in every Table
(County Code 3 DIGITS; 1 is not equal to 001)

Transaction Type: 00 - Original OR 05 - Replacement

Inventory Type: CRIT, HAP, or CRITHAP
(CRITERIA is not equal to CRIT)

Source Type: POINT, AREA, ON-ROAD MOBILE
NON-ROAD MOBILE, BIOGENIC

Contact Types: 
(There is no 06)

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Facility/Plant</td>
</tr>
<tr>
<td>02</td>
<td>Data Transmittal</td>
</tr>
<tr>
<td>03</td>
<td>Preparer</td>
</tr>
<tr>
<td>04</td>
<td>Receiving Location</td>
</tr>
<tr>
<td>05</td>
<td>Reporting Location</td>
</tr>
</tbody>
</table>
Coded Fields:

MUST BE EXACTLY AS IN CODE TABLE
   EXACT SPELLING
   Leading Zeros
   But not Case Sensitive (sort of)

UNK or UNKNOWN is not in many of the code tables

If you are using some converter program,
your database should have the standards
required or the converter program will not
output your data as you expect it to be
Numeric Range Checks

Emission Release Point Parameters

Annual Emission Values

Emission Period and Process Activity

Other Temporal Fields
Emission Release Point Parameters

Different from Data Augmentation Procedures
EFIG Augments data with ranges specific to SCC

Software checks NATIONAL RANGES not by SCC
with the intention to point out any value above
"normal" though it may very well be real

<table>
<thead>
<tr>
<th>Release Point Parameters</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Release Point Height (feet)</td>
<td>0.01</td>
<td>700</td>
</tr>
<tr>
<td>Release Point Height Fugitive (feet)</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Release Point Diameter (feet)</td>
<td>0.01</td>
<td>50</td>
</tr>
<tr>
<td>Exit Gas Temperature (degrees F)</td>
<td>50</td>
<td>1800</td>
</tr>
<tr>
<td>Exit Gas Velocity (feet/second)</td>
<td>0.01</td>
<td>560</td>
</tr>
<tr>
<td>Exit Gas Flow Rate (cubic feet/second)</td>
<td>0</td>
<td>200,000</td>
</tr>
</tbody>
</table>
Annual Emission Values

Annual: Emission Type = 30 (entire period)
AND the dates must be annual, 20XX0101
AND 20XX1231 AND the Units must be in Tons

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Point Max</th>
<th>Area Max</th>
<th>Mobile Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOC</td>
<td>300</td>
<td>1700</td>
<td>1100</td>
</tr>
<tr>
<td>NOx</td>
<td>1500</td>
<td>1250</td>
<td>1350</td>
</tr>
<tr>
<td>CO</td>
<td>600</td>
<td>5500</td>
<td>10250</td>
</tr>
<tr>
<td>SO2</td>
<td>5000</td>
<td>350</td>
<td>50</td>
</tr>
<tr>
<td>PM-FIL, PM-PRI, PM10-FIL, PM10-PRI</td>
<td>200</td>
<td>5000</td>
<td>50</td>
</tr>
<tr>
<td>PM25-FIL, PM25-PRI, PM-CON</td>
<td>150</td>
<td>1400</td>
<td>30</td>
</tr>
<tr>
<td>NH3</td>
<td>100</td>
<td>1250</td>
<td>50</td>
</tr>
</tbody>
</table>

Checked for the purpose of pointing out any value above "normal" and worth investigating
Emission Period and Process Activity

Number of Days per week ≤ 7

Number of Weeks per Year ≤ 52

Number of Hours per Day ≤ 24

Number of Hours per Year ≤ 8760 non-leap year
  ≤ 8784 leap year

Sum of Seasonal throughputs ≤ 100 Percent
Temporal Fields: Start and End Date and Time

Months from 1 to 12

Hours from 0 to 24

Minutes 0 to 59
Geographic Data - Coordinate checks are made

IF all of the format is correct in the locational fields
(if the state fips is not the correct code for example, the software cannot make the locational checks)

Once the format is verified as correct,
the Check to see if the coordinates fall within the State Boundary can be made

State Boundary - Boxes were drawn around the State to give Min/Max Lat/Long Coordinates which are in the StateFIPS Code Table of the QA/QC codes table database
If the coordinates fall within the State Boundaries, the Check to see if the coordinates fall within the County Boundary can then be made.

County Boundary - Max/Min Lat/Lon Coordinate Box is located in the table Co_bdry which also resides in the QA/QC codes table database.
What does the QA software do?

FORMAT Checks
CONTENT Checks
MS Access Database

What kind of Format Checks?

Tables - Existence
Fields - Properties
Referential Integrity - how the tables relate

What are Content Checks?

Acceptable Codes - Code Table Databases
Numeric Ranges - Maximum and Minimum
Geographic Data - Coordinate checks