Emissions Metadata Needs

Greg Frost

NOAA/ESRL & University of Colorado/CIRES

Boulder, Colorado, USA

Presentation Overview

Emissions Data Definitions and Uses
Emissions Data Providers
Emissions Data Distributors
Inventory Methodology
Emissions Data Dimensions
Emissions Data Considerations

Thanks to Claire Granier, Paulette Middleton, Ann Keane,
Leonor Tarrasón, Rudy Husar, and the ECCAD and CIERA teams

Emissions Information Definitions

Emissions = Inputs of constituents to the atmosphere

Inventory = Quantitative compilation of emissions

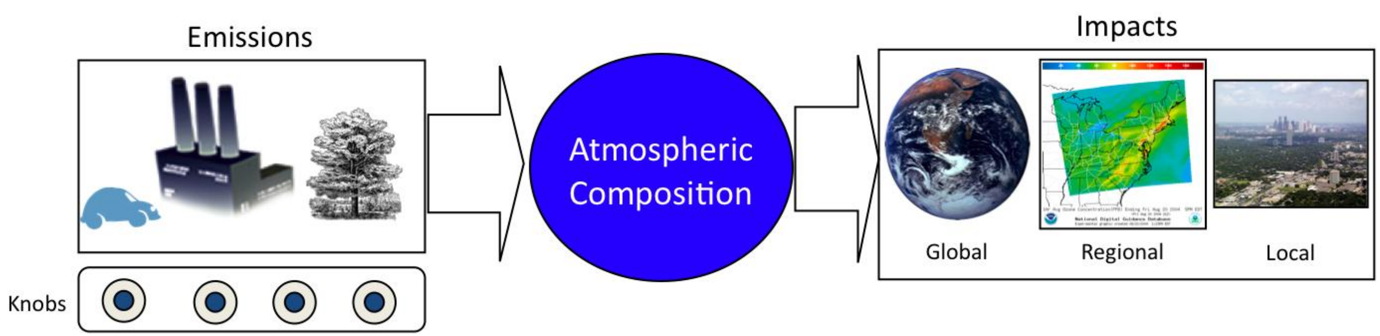
Uses of Emissions Information

Scientific Research

- Emissions affect atmospheric composition
- Impacts of atmospheric composition and its changes: air quality, climate change
- Inventories are input for atmospheric models
- Observations provide constraints on emissions inventories

Policy & Regulatory

- Inventories are regulatory and policy instruments
- Quantifying air quality and climate change
- Understanding impacts of changing emissions
- Actions and decisions about the atmosphere focus on emissions



Emissions Data Providers

"Official" emissions information

- Originates from government regulatory/environmental agencies, often in partnership with private sector

Examples:

UN, IEA

USEPA, USDOE, EMEP, Env. Canada, national environmental ministries

Regional or local environmental agencies

Research emissions information

- Official emissions information often not sufficient for scientific purposes
- Datasets for a variety of scientific purposes are produced by government agencies, academic institutions, and private sector

Examples:

JRC

NASA, USDOE, NOAA, USFS

ACCMIP, HTAP, AQMEII

Emissions Data Distributors

Data providers often provide their data through public portals

- Official distribution points for emissions data
- Data provided can have limitations: access, formats, traceability, timeliness
- Users may need to access several providers for completeness

Data clearinghouses seek to overcome some of these issues

- Aggregate emissions data sets
- Standardize data formats
- Supply data manipulation tools
- Facilitate objective evaluations
- *Examples:*

  GEIA - **G**lobal **E**missions **I**niti**A**tive
  http://www.geiacenter.org/

  ECCAD - **E**missions of atmospheric **C**ompounds & **C**ompilation of **A**ncillary **D**ata
  http://eccad.sedoo.fr

  CIERA - **C**ommunity **I**nitiative for **E**missions **R**esearch and **A**pplications
  http://ciera-air.org/

Inventory Methodology

Inventories are an amalgam of calculations and measurements

Total mass emissions of compound X

$$E_X = \Sigma_S \, [EF_{X,S} \bullet A_S \bullet (1 - CE_{X,S})]$$

Sum up all sources S

Emissions factor = mass of compound X emitted by source S per unit activity

Activity of source S, e.g., amount of fuel burned

Effectiveness of control measures for cmpd X at source S

Accumulate over specific geographic region
Calculate for specific time period
To use inventory in modeling or analysis, may need…

- Spatial allocation
- Temporal variability
- Temporal extrapolation
- Speciation

Emissions Data Dimensions

Dimensions common to earth science data

- Space
- Time
- Data Parameter

Additional dimensions specific to emissions data

- Pollutant
- Source Specification
- Source Aggregation

Emissions Data Considerations – Space

Domain

- Global, regional, local
- 2-D (surface) or 3-D (elevated sources)

Resolution

- Points: Large anthropogenic sources, volcanoes
- Line: Motor vehicles, aircraft, ships
- Area: National, state, county
- Gridded: Coarse (degrees) - fine (km), various projections & grids

Aggregation

- Inconsistent boundary & region definitions

Completeness

- Need to balance domain, resolution

Emissions Data Considerations – Time

Extent

- Days -> years -> decades -> centuries

Resolution

- Hourly -> daily -> monthly -> seasonal -> annual -> decadal

Reference

- Time standard: local vs UT?
- Time stamp: beginning or middle of interval
- Emissions data = accumulated totals for a time period: appropriate time stamp?

Existence

- Time variable may be missing in data/metadata

Emissions Data Considerations – Pollutant

Many possible pollutants

- Trace gases: e.g., CO, NOx, SO_2, VOC, CO_2, …
- Particulates: e.g., dust, PM2.5, BC, OC, sulfate, …
- Other species: CFCs, POPs, allergens (pollen)

Pollutants are application dependent

- AQ and climate species traditionally kept separate

Naming

- Many different names in common use
- Need comprehensive "dictionary"

Speciation

- Use of aggregated & specific pollutant categories: e.g., total VOC *vs.* methane, ethane, propane, …
- Speciation schemes are developer dependent
- General need for speciation cross-walks

Emissions Data Considerations – Source Specification

Type

- Anthropogenic – fossil fuel combustion, human & economic activity
- Biogenic – natural systems (vegetation, dust)
- Mixed - fires

Granularity

- Sector
  - Economic sector level is most commonly used
    e.g., mobile sources, industry, energy production
  - Difficult to connect sector categories between inventories
- Process
  - Detailed description of specific emissions process
    e.g., electricity generation with lignite coal using specific boiler + controls
  - Requires lots of information on emission factors and activities
  - Can aggregate process level information to produce sectors
  - But only used & reported in some inventories

Completeness

- Not all inventories include all categories
- Difficulty comparing different inventory emissions totals
- Difficulty understanding detailed differences between inventories

Emissions Data Considerations – Source Aggregation

Point

- Larger sources often treated individually in many inventories
  - E.g., power plants, factories, cement producers, …
  - Volcanoes: total emissions/eruption

Area

- Most other sources quantified over specific geographic region
  - County, state, nation, world
  - E.g., transport, residential, agriculture, etc.

Gridded

- Need to allocate point/area data to grids for modeling and analysis
- Use spatial proxies to convert point/area data to gridded data
  - E.g., population, road networks, etc.
- Spatial proxies are inconsistent between developers
- Lack of information on proxies for some sectors and regions

Emissions Data Considerations – Data Parameters

Methodologies

- Different approaches used by different developers
- Sector and pollutant dependent calculations
- Is it appropriate to compare estimates from different methods?

Units for emissions data

- Metric vs. other systems: e.g., tonnes vs. short tons or pounds
- Conversion of mass to number (molecular weight) for aggregated species

Multiple values

- Version control

Other Emissions Data Considerations

Documentation

- How emissions were estimated is as important as data themselves!
- Must be captured in metadata
- Inventory development information can be difficult to find

Accessibility

- Differences between providers
- Many data formats
- Public *vs.* proprietary data

Communication between regulatory and scientific community

- Different needs and cultures
- Two-way feedbacks difficult

Evaluation and assessment

- Objective evaluations using observations/models
- Check reliability of methodologies
- Need expert assessments of inventories

Some Closing Thoughts

- Emissions data have unique dimensions and considerations when compared with other earth science data
- There is a lack of consistency in many attributes of emissions data because of different developers, methodologies, and applications for these data sets
- Emissions metadata are not generally compliant with metadata standards for other types of earth science data
- Web services approaches for handling emissions data exist, but are still dataset specific in some cases
- More generic approaches to linking emissions data sets are needed
- Need to learn from other communities working with air quality metadata
- Need to link with other communities developing/using emissions, such as climate modeling groups, satellite observations, assessment/future scenarios, health, etc.