DATA MANAGEMENT, EXISTING TOOLS/SOFTWARE

April/May 2024

U.S. Environmental Protection Agency

DISCUSSION

- Existing approach

- Good practices when using xlsx

- Data management

- Available compilation tools

- Inventory System Management Toolkit

- Other tools
  - FAO Data Collection templates
  - CBIT-GSP NID template (under review) and BTR report template (under development)

EXISTING APPROACH

- Good starting point!
  - Includes some standardization and checks
  - Include documentation of data sources
  - Set up for time series calculations

GOOD PRACTICES FOR ORGANIZING RAW DATA AND PROCESSING

- Retain record of raw unprocessed data as received/downloaded in current inventory files and inventory archive (when inventory is complete)
- Copy/read in raw data into estimation file for further processing/estimation steps (generalized approach)
  - Introduce into own sheet
  - Ensure data source/reference information/links are noted
  - Ensure units are noted
  - Note any data versions if relevant
- Create look up tables for unit conversions, constants, EFs, GWPs – never hardcode
- Calculation steps should be annotated, stepwise for transparency and also to facilitate QA/QC
- Automate QC where feasible, i.e., trend checks, implied EF checks

Discuss – potential to standardize data input sheets to extent possible, for example:
- Year
- Value
- Data type1 (e.g., activity data, emission factors, other factors)
- Data Name (e.g., fuel consumption)
- Units
- Sector
- Category
- Subcategories
- Carbon pool and fuel type/subtype information as relevant
- CRT category/subcategory code
- Data sensitivity (Y/N)

TYPES OF DATA AND OTHER INFORMATION TO MANAGE

- Manage raw input data
- Calculations
- GHG estimates (totals and by gas at various levels)
  - Subcategory (could be most disaggregated level)
  - Category level (could be most disaggregated level)
  - Subsector
  - Sector
  - Subnational (?)
  - National
- Cross-cutting analysis
  - Key category analysis (data for analysis can come from previous list
  - Recalculations (requires easy access to previous estimates)
  - Uncertainty information
- Common Reporting Table data
  - Emissions, activity data, and notation keys/explanations (including use of flexibility)

Support products

National Inventory Document

Common Reporting Tables

Web publication?

AVAILABLE COMPILATION TOOLS

Options

Considerations
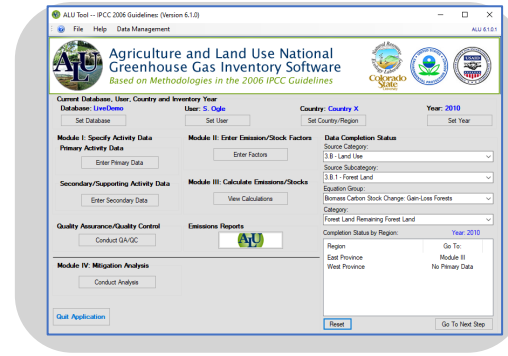
IPCC Inventory Software

- Free, available software, already organized to apply IPCC GL
- Can also import time series data
- Manages all data mentioned on previous slide
- Can aggregate data, generate summary data reports, etc.
- Includes cross-cutting analysis capability (i.e. KCA, uncertainty analysis using approach 1)
- Interoperability with CRTs
- Limitations (i.e., incompatible with Mac OS, usability improvement, other)

Country-specific tools

(i.e., use of xlsx spreadsheets, or other customized application (e.g. database based approach, e.g., R, Python, etc.)

- Develop or build from existing system
- Standardize organization of data inputs/outputs
- XLS easy to use, simple but has limitations for larger data sets, link errors, etc.
- Requires developing system to aggregate and synthesize data
- Potential limitations
  - Database development requires specialized expertise
  - Maintenance over time

Sector-specific tools

(e.g., Agriculture and Land Use National Inventory Software (ALU))

- Free available software, organized to apply IPCC GL
- Facilitates compilation of indicated sector
- Doesn't cover all sectors
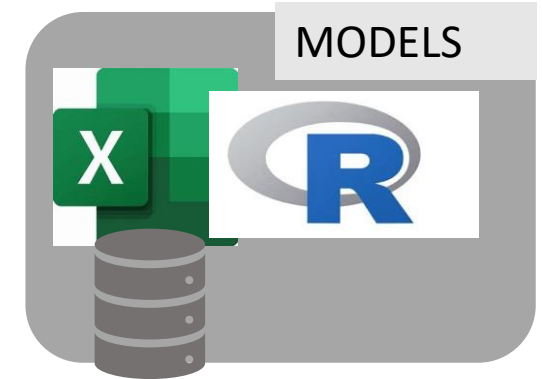
Combination of tools

(IPCC Inventory Software and other tools)

- Use best tool for task
- Lose some standardization across sources
- Potential to import information into IPCC software for aggregation? QC?

POTENTIAL DATA INPUT TEMPLATE STANDARDIZATION

Discuss – potential to standardize data input sheets to extent possible, for example:

- Year
- Value
- Data type1 (e.g., activity data, emission factors, other factors)
- Data Name (e.g., fuel consumption)
- Units
- Sector
- Category
- Subcategories
- Carbon pool and fuel type/subtype information as relevant
- CRT category/subcategory code
- Data sensitivity (Y/N)

	Value	Units	Data Type	Sector	Subsector	Etc.			
Year1									
….									
Year2 5									

AGRICULTURE AND LAND USE GREENHOUSE GAS INVENTORY SOFTWARE (ALU)

- ALU was developed based on experiences in EPA led capacity building projects
  - Developed and housed with Colorado State University
- Guides the compiler through the process of the inventory analysis for the AFOLU sector
- Provides data management capabilities (and documentation) and prevents obvious errors (built in QC checks)
- Provide utilities that encourage *good practice* (e.g., documentation of choices)
- Mitigation analysis
- ALU is difficult to use without having lots of training and direct use experience
- Significant improvements underway addressing usability/user interface, completeness, mitigation analysis, among others.
  - Ideally accessing engineers/programmers who can support development

FAO DATA COLLECTION TEMPLATES

- Excel-based package to facilitate compilation of data inputs

- Includes comprehensive list of activity data and parameters, that must be collected to estimate, mostly at Tier 1, all categories within all sectors

- Link: [The Greenhouse Gas Data Management (GHG-DM) Tool (fao.org)](fao.org)

CBIT-GSP'S NATIONAL INVENTORY DOCUMENT TEMPLATE

- Report template following suggested outline from Glasgow (COP27)
  - User needs to update/elaborate content, but some content is already available

- Coming soon!

EPA Toolkit for Building National GHG Inventory Systems

Customizable Templates to help inventory compilers build and advance GHG inventory management systems that suit their national circumstances.