GEIA: Making Data Relevant

Andres RJ, Boden TA

Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6335 USA

andresrj@ornl.gov

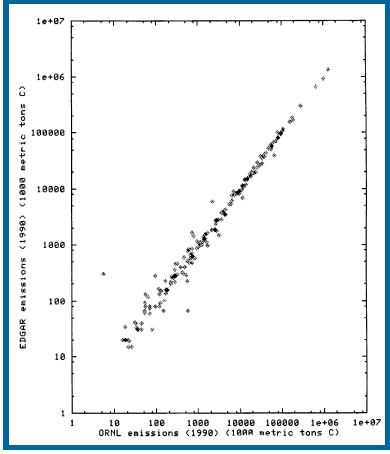
ABSTRACT GEIA is an organization that has been around for almost two decades. It has stayed active and relevant by anticipating and responding to user needs. These needs include high-quality, global inventories of climate-change-relevant chemical (e.g., CFC-11) and/or source-specific (e.g., volcanic sulfur emissions) species. This presentation will focus on two themes: quality assurance/quality control (QA/QC) of data products and anticipating future user community needs.

GEIA already has in place a number of QA/QC procedures, including peer-reviewed publication of resulting inventories, but the real emphasis of QA/QC resides with the efforts of the individual inventory compilers. These compilers are in the best position to assess each and every data point that eventually resides in the inventory. It is paramount that the inventory compilers verify all data used in their inventories as subsequent users of the inventories rarely have the time/resources to check inventory values and instead rely upon the inventory compilers and the GEIA process to ensure that the inventories are as accurate as possible.

Anticipating future user needs is not an easy task, but an important one for GEIA as well as individual inventory compilers. Quality inventories take significant amounts of time and resources to compile. Thus, the use of these inventories by the broader community must not only consider current needs, but also future needs. Inventory characteristics such as spatial scale, temporal scale, species inventoried, data format(s), and data availability all need to be carefully considered. Examples of current as well as future considerations will be given in this presentation.

$$X_{reference} \stackrel{?}{=} X_{inventory}$$

$$\sum \text{spatial/temporal domains} \stackrel{?}{=} \text{global domain}$$



QA/QC One of the fundamental problems for most, if not all, inventories is the lack of independent measurement of the inventoried species in each grid cell of the inventory. Thus, it is difficult to fully assess the uncertainty in the inventory at various spatial and temporal scales. However, the situation is not hopeless. There are several internal and external QA/QC procedures which can be performed on the inventory which can help assess the quality of the inventory and instill confidence for its use. These procedures fall into the categories of internal and external procedures.

Internal procedures include traditional data practices of proper and complete data compilation procedures as well as various sum checks. Since many GEIA inventories are derived from the work of others (as opposed to being generated from field measurements), it is important to document these sources, properly copy and code their inventory values and locations, and retain knowledge of the associated uncertainties. Sum checks involve ensuring internal data consistency over various spatial and temporal domains. For example, the sum of grid cell totals should equal the national totals for a particular specie. Likewise, the sum of national totals should equal the global total for a particular specie.

External procedures include comparisons to previously constructed inventories and comparisons to field measurements. Care must be taken in inventory-inventory comparisons to properly account for changing spatial and temporal domains, category definitions, and data processing procedures and assumptions (e.g., algorithm design). Note that this comparison, by itself, does not judge one inventory better than another. Inventory-field measurement comparisons offer their own unique challenges of spatial and temporal domains. In addition, the level of agreement in one inventory grid cell-field measurement comparison does not necessarily give any information about the quality of the inventory in other grid cells.

Internal and external data inconsistencies often teach us where in the inventory there are problems. This identification then can lead us to the most important areas in need of improvement.

ANTICIPATING FUTURE USER COMMUNITY NEEDS

Species - Major, minor, and trace species are all valuable. Some are climate-relevant (e.g., CFCs) while others are useful to test model dynamics (e.g., radiocarbon).

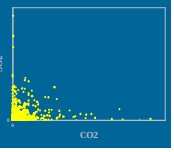
Spatial Scale - GEIA has traditionally used a one degree latitude by one degree longitude format. This was fine when GEIA was started almost 20 years ago. But, the global modeling community has caught up to and regularly utilizes this scale. We need to think about finer spatial scales for future community needs. Local and regional modelers are already utilizing finer spatial scales.

Temporal Scale - GEIA has traditionally used one year as its basic temporal unit. This was appropriate when GEIA was started almost 20 years ago and was then, and for some species, is today the best temporal scale achievable. But, the global modeling community regularly utilizes finer temporal scales (e.g., a 15 minutes temporal unit is not uncommon for atmospheric transport models). We need to think about finer temporal scales for future community needs. Perhaps seasonal, monthly, weekly, daily, hourly, ... inventories could be constructed for some species.

Data Formats - need formats accessible to all, preferably in formats easily utilized by end users (e.g., ASCII, netCDF, ...)

Data Availability - need inventories available from well-known archives for your intended audience (e.g., GEIA/ACCENT, CDIAC, ...)

Documentation - long a hallmark of GEIA inventories. This includes not only inventory preparation, but also data format properties. Documentation should be co-located with the online data. Documentation should also be updated, as needed, as successive versions of the inventory are released.



Visualization Tools - not a traditional GEIA strength. Some inventories have a static map accompanying their data sets. These are useful to users to not only let them see spatial coverage, but also to help them ensure that they are reading in data sets in the proper orientation. However, more dynamic visualization tools could also allow users to see temporal coverage, query specific values in a particular inventory, create inventory versus time trend plots on a global or grid cell basis, and graphically examine covariance between inventories.