

Title: Standardizing Interfaces for External Access to Data and Processing for the NASA Ozone Product Evaluation and Test Element (PEATE)
Authors: Curt A. Tilmes and Albert J. Fleig, NASA Goddard Space Flight Center
Meeting: Geoinformatics 2008, Potsdam, Germany

Abstract:

NASA's traditional science data processing systems have focused on specific missions, and providing data access, processing and services to the funded science teams of those specific missions. Recently NASA has been modifying this stance, changing the focus from Missions to Measurements. Where a specific Mission has a discrete beginning and end, the Measurement considers long term data continuity across multiple missions. Total Column Ozone, a critical measurement of atmospheric composition, has been monitored for decades on a series of Total Ozone Mapping Spectrometer (TOMS) instruments. Some important European missions also monitor ozone, including the Global Ozone Monitoring Experiment (GOME) and SCIAMACHY. With the U.S./European cooperative launch of the Dutch Ozone Monitoring Instrument (OMI) on NASA Aura satellite, and the GOME-2 instrument on MetOp, the ozone monitoring record has been further extended.

In conjunction with the U.S. Department of Defense (DoD) and the National Oceanic and Atmospheric Administration (NOAA), NASA is now preparing to evaluate data and algorithms for the next generation Ozone Mapping and Profiler Suite (OMPS) which will launch on the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP) in 2010. NASA is constructing the Science Data Segment (SDS) which is comprised of several elements to evaluate the various NPP data products and algorithms.

The NPP SDS Ozone Product Evaluation and Test Element (PEATE) will build on the heritage of the TOMS and OMI mission based processing systems. The overall measurement based system that will encompass these efforts is the Atmospheric Composition Processing System (ACPS). We have extended the system to include access to publically available data sets from other instruments where feasible, including non-NASA missions as appropriate. The heritage system was largely monolithic providing a very controlled processing flow from data ingest of satellite data to the ultimate archive of specific operational data products. The ACPS allows more open access with standard protocols including HTTP, SOAP/XML, RSS and various REST incarnations. External entities can be granted access to various modules within the system, including an extended data archive, metadata searching, production planning and processing.

Data access is provided with very fine grained access control. It is possible to easily designate certain datasets as being available to the public, or restricted to groups of researchers, or limited strictly to the originator. This can be used, for example, to release one's best validated data to the public, but restrict the "new version" of data processed with a new, unproven algorithm until it is ready.

Similarly, the system can provide access to algorithms, both as modifiable source code (where possible) and fully integrated executable Algorithm Plugin Packages (APPs). This enables researchers to download publically released versions of the processing algorithms and easily reproduce the processing remotely, while interacting with the ACPS. The algorithms can be modified allowing better experimentation and rapid improvement. The modified algorithms can be easily integrated back into the production system for large scale bulk processing to evaluate improvements.

The system includes complete provenance tracking of algorithms, data and the entire processing *environment*. The *origin of any data or algorithms is recorded and the entire history of the processing chains are stored* such that a researcher can understand the entire data flow. Provenance is captured in a form suitable for the system to guarantee scientific reproducibility of any data product it distributes even in cases where the physical data products themselves have been deleted due to space constraints. We are currently working on Semantic Web ontologies for representing the various provenance information.

A new web site focusing on consolidating information about the measurement, processing system, and data access has been established to encourage interaction with the overall scientific community. We will describe the system, its data processing capabilities, and the methods the community can use to interact with the standard interfaces of the system.