

Final report of the ESDS-RFC-009 Technical Working Group

The ESDS-RFC-009 Technical Working Group recommends that the NASA Earth Science Data Systems Standards Process Group publish ESDS-RFC-009 “A File Format for Satellite Atmospheric Chemistry Data Based on Aura File Format Guidelines” as a Technical Note.

RFC Overview

ESDS-RFC-009, available at <http://www.esdswg.com/spg/rfc/esds-rfc-009>, describes the HDF-EOS5 profile developed for data products from the instruments on the Aura satellite, as described in the excerpt below:

This technical note describes a file format for atmospheric chemistry data, which makes it easier for the end user to read product files from several different instruments. It was adopted by all four instruments on the EOS Aura satellite (HIRDLS, MLS, OMI and TES) and can be readily used by other atmospheric chemistry instruments.

The Aura teams realized that common data and file formats would greatly facilitate the sharing of data. The teams agreed to use the HDF5/HDF-EOS5 data format. In addition, it was important to concur upon a file format to further reduce the need for custom software. The teams agreed upon the names, data types and dimension order of fields. There was also agreement regarding the file-, group- and field-level attributes to include in each product file. A file-naming convention was also adopted.

Future atmospheric chemistry missions can use this technical note to enable their users to benefit from a common data and file format. They can either build upon the guidelines as described here, in which case they will benefit from being able to easily read Aura data files. Or future missions can simply use the techniques described below to adopt a common data and file format for their collection of instruments. There are many advantages to having multiple instruments follow the same data and file format conventions.

Recommendation

The TWG bases its recommendation on technical reviews of the RFC and an analysis of the following factors in a NASA context:

Strengths: The RFC provides comprehensive documentation of the data file format and organization agreed to and implemented by the Aura Instrument Teams, including:

- Major HDF-EOS version (HDF-EOS V2.x and HDF-EOS V5.x are not interchangeable)
- Organization of geolocation and data fields and attributes
- Dimension names
- Geolocation field names, data types and dimension ordering
- Data field names, data types and dimension ordering
- Units for geolocation and data fields
- Attribute names, values and units

Weaknesses: Units specified in the RFC do not conform to the SI Convention for Representation, and in some cases, there are inconsistencies in units among the different

data fields (e.g., squared items are sometimes with a superscript, sometimes with ^ and others inline (vnr2 or K2)). However the RFC reflects the actual data files produced, which will not be changed. Other instrument teams using the Aura File Format Guidelines are urged to consider using the SI Convention.

Applicability and Limitations: If by analogy HDF5 is considered the Database Server, then the Aura Guidelines are the tables, and columns and rows, of a particular, unique database. Other Missions/teams should have the freedom to create their own artifact (i.e., database or "profile") to take advantage of their domain knowledge. What may be missing is the system-level picture of how the databases/profiles are developed for other missions. Is this standard an artifact of Aura that can be re-used? For maximum benefit, should it be re-used or should the process by which it was created be re-used? It's no small accomplishment that this was actually agreed-to by such a large team.

Note that the process by which the Aura guidelines were developed is to be the subject of a forthcoming RFC.