A Summary of Proposed Changes to the Current ICARTT Format Standards and their Implications to Future Airborne Studies

Emily Northup[1], Amanda Benson Early[1], Aubrey Beach[2], Dali Wang[4], John Kusterer[3], Brandi Quam[3], Gao Chen[3]

1. Science Systems and Applications, Inc., Hampton, VA 2. Booz Allen Hamilton, Inc., Norfolk, VA 3. Atmospheric Science Data Center, NASA Langley Research Center, Hampton VA

4. Department of Physics, Computer Science and Engineering, *Christopher Newport University, Newport News, VA*

ASDC Introduction

The Atmospheric Science Data Center (ASDC) at NASA Langley Research Center is responsible for the ingest, archive, and distribution of NASA Earth Science data in the areas of radiation budget, clouds, aerosols, and tropospheric chemistry. The ASDC specializes in atmospheric data that is important to understanding the causes and processes of global climate change and the consequences of human activities on the climate. The ASDC currently supports more than 44 projects and has over 1,700 archived data sets, which increase daily. ASDC customers include scientists, researchers, federal, state, and local governments, academia, industry, and application users, the remote sensing community, and the general public.

Mission, Approach, & Stakeholders

Mission Statement:

The ICARTT Refresh Working Group will work to improve the ICARTT format, the current NASA ESDIS standard for airborne data, in order to enhance the usability, interoperability, and discovery of airborne atmospheric observational data sets while maintaining backward compatibility.

Approach:

1. Identify and collect deficiencies of the ICARTT format user experience from data providers, DAACs, and users
2. Solicit community recommendations to mitigate identified format deficiencies
3. Develop a refreshed draft for ICARTT standards

Stakeholders:

- Airborne campaign sponsoring agencies
- Airborne science data archival centers
- Airborne measurement community
- Atmospheric modeling community
- End users

Special Thanks

ESDIS
ICARTT ESDSWG
TAD

More Information

TAD Toolsets for Airborne Data

https://tad.larc.nasa.gov

Implementation

In order to improve the current ICARTT format, this working group used a four step approach to communicate with appropriate users and providers and formulate recommendations that accurately address the concerns of the community.

- **Step 1**. Compile a list of deficiencies from the current ICARTT format based on user experiences.
  1. Not suitable for 3D (multidimensional data)
  2. Structural/Syntactic inconsistency in header (comma, underscore, etc.)
  3. Date/time standardization
  4. Current metadata do not follow any formal standards
  5. Variable metadata currently not extendable (e.g. no ancillary keywords)
  6. Variable naming convention inconsistency
  7. No consistent number of variables per deployment

- **Step 2**. Conduct a polling survey among working group members to identify top three deficiencies.

Question	Average Score out of 5
Variable naming convention consistency	4.69
Structural/syntactic consistency in header (comma, underscore, etc)	4.54
Date/time standardization	4.38
Consistent number of variables per deployment for the same instrument	4.08
Variable metadata currently not extendable	3.85
Current metadata does not follow any particular set of standards (e.g. ISO 19115)	3.77
Not suitable for 3D	2.62

- **Step 3**. Select top three deficiencies and formulate recommendations. Conduct another survey to vote on best recommendation approach.

Where to put standard variable names?
We agree the need to have standard variable names associated with the PI variable names; however, the issue is where to put them in the header.

Question	Average Score
1. Add a standard variable name entry before the long name in the variable name line, i.e. short name, unit, STANDARD NAME, long name	69%
2. Require the standard variable name to have a separate header line, similar to and just above the short name line	0%
3. Require both 1 and 2	23%
4. Require use of the standard names as part of the short names (e.g. O3_CaRDS_ppbv_O3 – see comments for description)	7%

Where to put ICARTT format version number?
This is to indicate the ICARTT version of the file.

Question	Average Score
1. Add a field to the end of the first line for the version number	93%
2. Add the ICARTT version number to the header (e.g. using a keyword in the Normal Comments Section)	6%
3. Other suggestions	0%

Standardization of the date time variable names

Question	Average Score
1. Require the time variable name to have suffix of "_start", "_mid", and "_stop" to indicate if it represents the start, mid, or stop time	0%
2. Require the standard names of the time variables to be one of e.g. "START_TIME", "MID_TIME", or "STOP_TIME"	100%
3. Require date or Julian day in the data stream (e.g. as a dependent variable	0%

- **Step 4**. Draft recommendations for the ESDIS Standards Office to formalize ICARTT upgrade.

Working Group Composition

- End Users
- Data Modelers
- Data Analysts
- Airborne Measurement Scientists
- Airborne Instrument Scientists
- Data Management Practitioners

Recommendations

Recommendation #1

"Add a required standard variable name entry before the long name in the variable name line"

Current Standard: Variable short name, variable unit, variable long name (optional)

Recommendation: Variable short name, variable unit, variable standard name, variable long name (optional)

Recommendation #2

"Add a field to the end of the first line for the version number"

Current Standard: Number of header lines, FFI

Recommendation: Number of header lines, FFI, version number

Recommendation #3

"Require that the standard names for the date time variables be Time_Start, Time_Stop, Time_Mid"

Current Standard: None

Recommendation: short_name, unit, Time_(Start,Stop,Mid), long_name

Ongoing Activities

- Airborne metadata model formalization
- Development of a standard controlled vocabulary list (aerosols and trace gas)
- Version number formatting recommendation
- File Format Index (FFI) 2110 and 2310 reformatting to handle multidimensional data