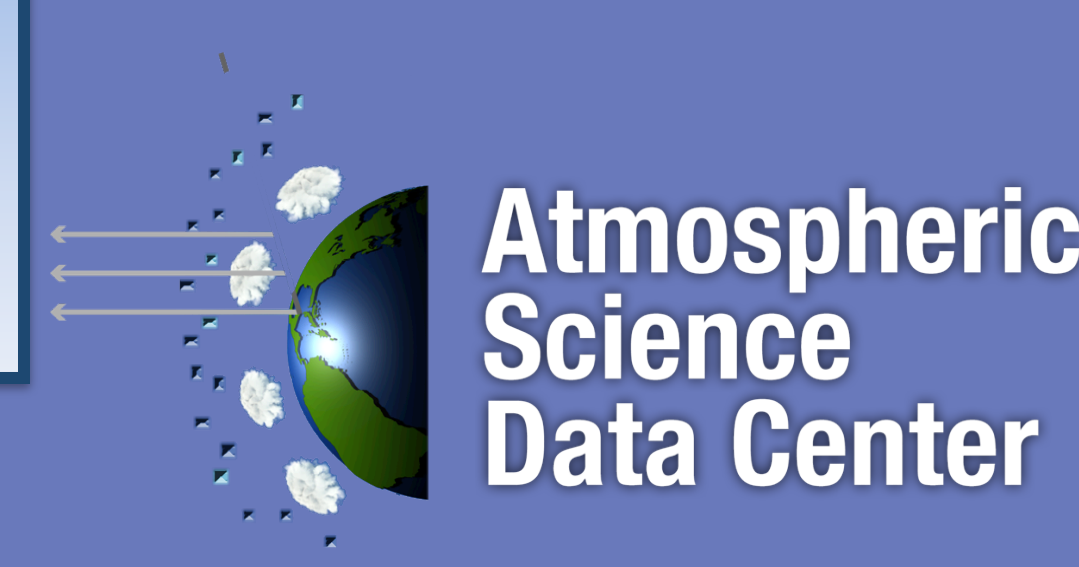


An Overview of the Challenges with and Proposed Solutions for the Ingest and Distribution Processes For Airborne Data Management



Aubrey Beach¹, Emily Northup², Amanda Early², Dali Wang⁴, John Kusterer³, Brandi Quam³, Gao Chen³

1. Booz Allen Hamilton, Inc., Norfolk, VA 2. Science Systems and Applications, Inc., Hampton, 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.

Description of Airborne Campaigns

- Airborne field studies are designed to address specific science objectives by providing a suite of simultaneous measurements at high spatial and temporal resolutions:
 - to investigate atmospheric processes and/or phenomena
 - to assess model performance
 - to validate satellite observations
- May involve 1 or 2+ aircraft platforms, typically last 1 - 3 months, and conduct 10 - 20+ flights. One or more deployment sites
- ~8 hours of science data per flight.
- Consists of measurement scientists, modelers, meteorologists, and project support teams.
- Typically have project data repositories
- Data is open to the public ~1 year after field deployment

Challenges with Data Management

There are considerable differences between satellite and aircraft observations, which contributes to knowledge gaps between users within the airborne field study community and data centers (DAACs).

Metadata vs. Data

- Sampling location is data, not metadata
- Data collections should be created at the project level, since individual PI data is not stand alone

Instrument Collections

- Airborne instruments routinely go through rapid evolution
- New instruments tested via field deployments
- Instrument development process documented through journal publications

ICARTT Data Format

- Lacks a standard variable naming convention and version format identification

Data Processing Cycle

- Processing completes within project funding cycle, normally 3 yrs

Current Airborne Data Management Overview

The current data management practices for NASA airborne field projects have successfully served science team data needs over the past 30 years to achieve project science objectives, however, users have discovered a number of issues in terms of data reporting and format. The ICARTT format, a NASA standard since 2010, is currently the most popular among the airborne measurement community. Although easy for humans to use, the format standard is not sufficiently rigorous to be machine-readable. This makes data use and management tedious and resource intensive, and also create problems in Distributed Active Archive Center (DAAC) data ingest procedures and distribution. Further, most DAACs use metadata models that concentrate on satellite data observations, making them less prepared to deal with airborne data.

Current Process Workflow for Airborne Data Campaigns

A field campaign generally follows the following lifecycle from mission planning, conducting the campaign, data management/archival, distribution, and eventual use for research and applications.



Mission Planning

1) Mission Planning:

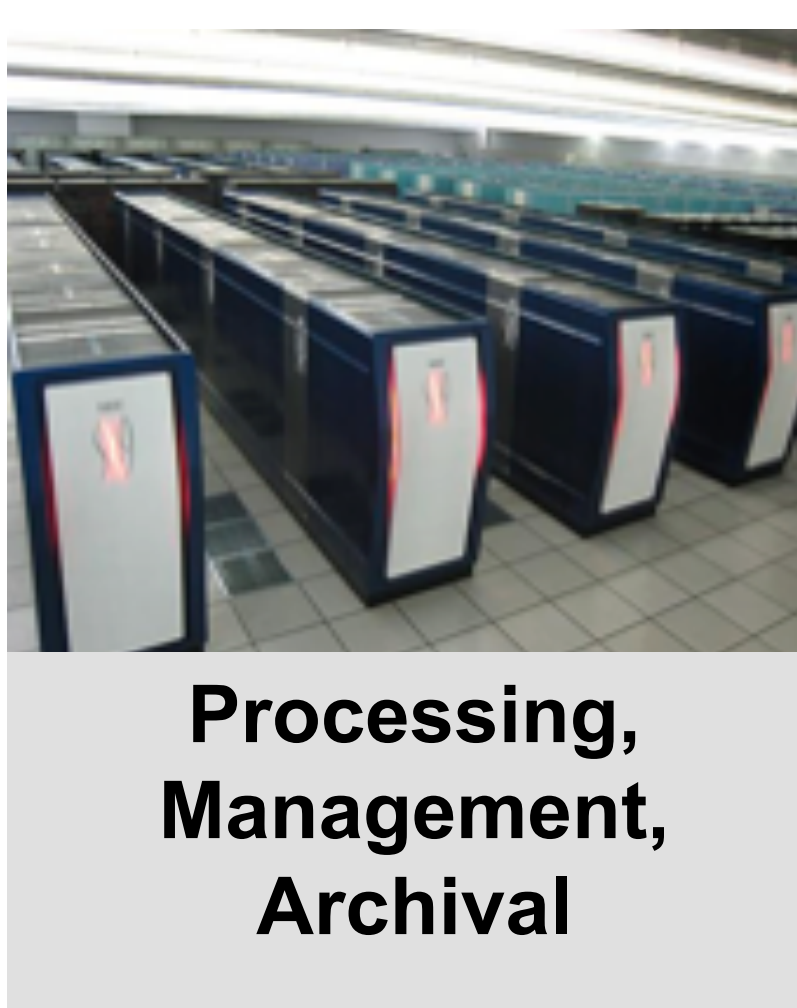
- Formulation science objectives and preliminary mission implementation plan.
- Science team selection through ROSES proposals and peer-review processes.
- Discussion and finalization of the mission implementation



Campaign Operations

2) Campaign Operation:

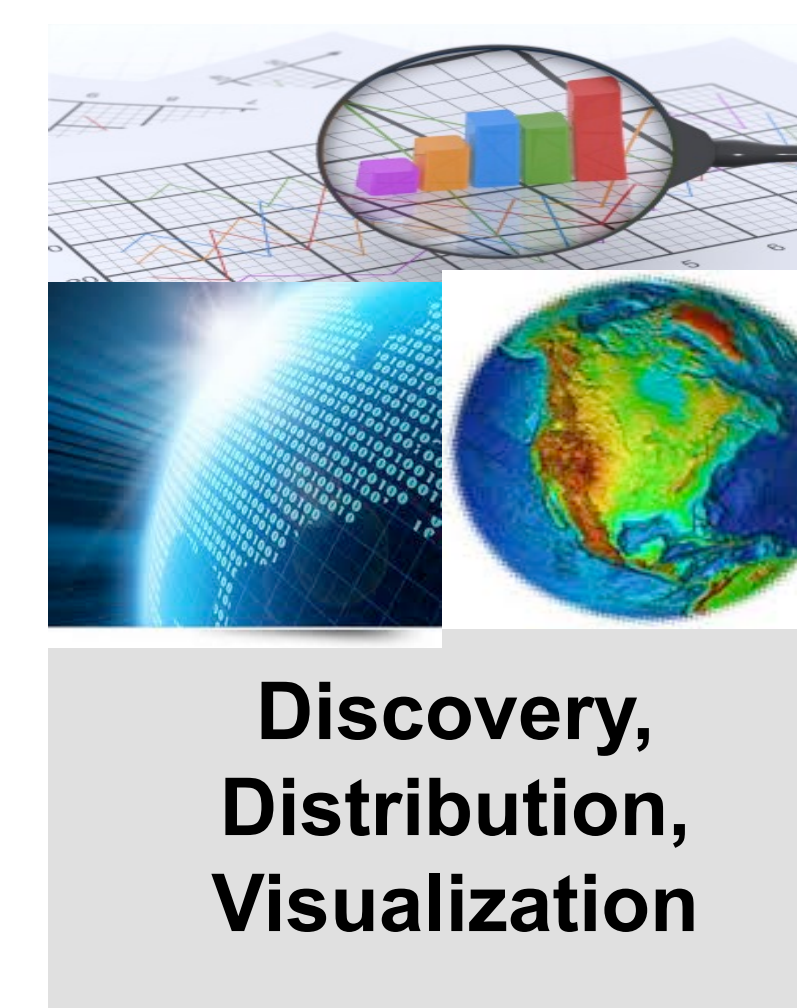
- Daily mission planning meetings.
- Conduct science flights to collect data relevant to one or more science objectives.



Processing, Management, Archival

3) Data Processing, Management, and Archival:

- Three phases of data processing: field, preliminary, and final.
- Only final data is released to public.
- Final data transferred to a NASA DAAC for long term preservation and distribution.



Discovery, Distribution, Visualization

4) Data Discovery, Distribution, Visualization:

- DAAC for ingest, archival, and distribution chosen based on the field campaign area of study.
- Finalization of collection and granule level metadata.
- DAAC may create custom tools for visualization and/or analysis.

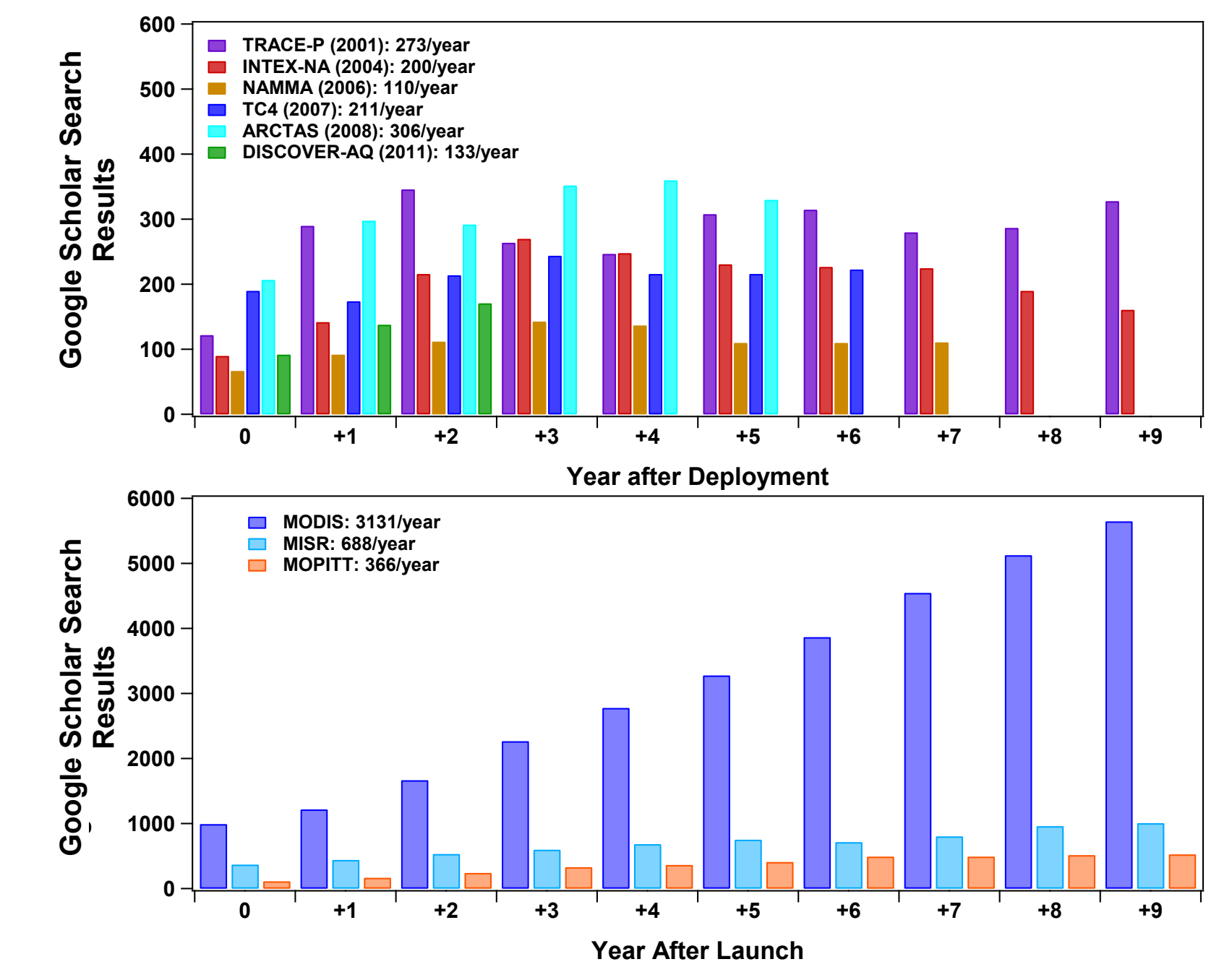


Research, Applications Decision Support

5) Research, Applications, Decision, Support:

- Airborne science data sets are used to address the project science objectives.
- Data are widely used by NASA sponsored investigators through various programs, e.g. ACPMAP and Atmospheric Composition Campaign Data Analysis and Modeling Program

Increasing Usage of Airborne Data



Solutions

- Unified Metadata Model (UMM) for aircraft data to increase metadata quality for aircraft-specific data
- ICARTT Refresh Working Group creating recommendations to improve machine interpretation of metadata, particularly the variable description/definition
- The adoption of a standard variable naming convention to enable greater accessibility and discovery of airborne data sets
- Improved custom distribution tools, e.g. Toolsets for Airborne Data at the ASDC

Toolsets for Airborne Data

- TAD is focused on *in situ* observational data, which represent the majority of the airborne measurements in the Atmospheric Composition Focus Area
- Generates on-the-fly weighted averages of derived value-added products from airborne data holdings.
- Automated parsing tools convert ICARTT files to a more machine-readable format for TAD ingest

More Information



Special Thanks

NOAA ESRL
GCMD
WWW-AIR