

# INTEGRATING AND MANAGING DATA FROM MULTIPLE A-TRAIN SENSORS

**Andrey Savtchenko, Steven Kempler, Peter Smith, Gregory Leptoukh, Graeme Stephens, David Winker**

[Andrey.K.Savtchenko.1@gssc.nasa.gov](mailto:Andrey.K.Savtchenko.1@gssc.nasa.gov)

*NASA's Goddard Space Flight Center (GSFC)/RSIS, Greenbelt, MD 20771*

## Abstract

The succession of US and international satellites that follow each other, seconds to minutes apart, across the local afternoon equator is called the A-Train. The A-Train consists of the following satellites, in order of equator crossing: OCO, Aqua, CloudSat, CALIPSO, PARASOL, GLORY, and Aura, all in the same sun-synchronous orbit. Flying in such formation increases the number of observations in close proximity to each other, spatially and temporally, validates observations, and enables coordination between science observations, resulting in a more complete atmospheric science dataset. However, integrating and managing (e.g., vertical and horizontal co-registering and regridding of datasets) from independently developed missions, to realize the fullest benefits of the A-Train formation flying, can be daunting to some, and repeated by many. Scientists will individually spend much of their time and resources acquiring A-Train datasets of interest residing at various locations, developing algorithms to match up and graph datasets along the A-Train track, and search through large amounts of data for areas and/or phenomena of interest. The aggregate amount of effort that can be expended on repeating pre-science tasks could climb into the tens of millions of dollars. Then science can be performed.

This paper discusses the first ever A-Train virtual data portal/center, the A-Train Data Depot (ATDD), being implemented to integrate and manage datasets generated by all A-Train sensors. The objectives of this paper is to discuss the: Goals of the ATDD; ATDD capabilities; Challenges, solutions, and lessons learned when integrating multi-sensor datasets, valuable to future data integration efforts, and; Expected and unexpected multi-sensor user requests beneficial to the broader community.

The ATDD has been operational for over 18 months, dynamically (i.e., though user initiated requests) serving co-registered data generated by the Cloudsat, Calipso, Aqua AIRS, Aqua MODIS, and Aura OMI instruments. Datasets from PARASOL POLDER and Aqua AMSR-E are next, from a long list, to be added. The ATDD processes, archives, accesses, visualizes, analyzes and correlates atmospheric measurements, acquired from various remote archives. The ATDD enables the free movement of A-Train data so that they are combined to create a consolidated vertical view of the Earth's Atmosphere along the A-Train tracks. Whereas access to data from instruments with relatively narrow fields of view are easily facilitated, the ATDD provides subsets of data from wider field of view (FOV) instruments that only coincide with data from the narrower FOV instruments. Narrowing the data field is generally an essential first step for cross instrument data analysis. In addition, transferring much less data to the user's site is much more manageable. The innovative approach of analyzing and visualizing atmospheric profiles along the platforms track (i.e., time) allow users to discover data of

interest prior to downloading lots of potentially undesired data. Understanding the impacts of spatially different measurement sizes, as well as the specific orbits of the various platforms, makes matching up datasets both challenging and rewarding

Data residing at the ATDD is archived on-line for fast data access, using the Simple, Scalable, Script-based Science Processor for Archive (S4PA) data management system. The MIRADOR data search and access tool enables users to find specific data of interest. The Giovanni exploration tool is used to create vertical and horizontal visualizations of subsetted data and provide data for user download. The ATDD will continue to evolve as A-Train datasets are added and services are implemented in response to science needs.

The integration of multi-sensor data from heterogeneous missions is the next groundbreaking step in maximizing the use and usefulness of NASA Earth science data. Scientists have already begun the laborious tasks of combining pairs of datasets for data validation, inter-comparison, and scientific analysis. ATDD will remove this burden from A-Train data users