SOOT for ACTIVATE: Reaching Aerosol Research Communities
An ASDC Update
Megan Buzanowicz, Morgan Silverman, Sean Leavor, Jennifer Tindell, John Kusterer, Gao Chen
What is the ASDC?

• One of twelve NASA Earth Observing System Data and Information System (EOSDIS) Distributed Active Archive Centers (DAACs)

• ASDC supports four focus areas: Aerosols, Clouds, Tropospheric Chemistry and Radiation Budget
  • Maintain a focus in handling field campaign data

• Sole post project data portal for distribution of data products
  • Responsible for long-term preservation and stewardship

• ACTIVATE data holdings
  • Archive the latest versions of publication quality data, including observational, derived, and value-added data products
  • Contextual information to facilitate data use by research community at large
  • Documentation to maintain reprocessing capability and openness
Data Archival Update

• ASDC has completed ingest, archival, and distribution of all currently available publication quality data
  • If any data products are generated related to manuscript publication, we’d be happy to archive/publish the data at the ASDC

• Data is available through all the ASDC’s search mechanisms (Earthdata Search and SOOT)

• Data updates will be uploaded first through the field repository and ingested periodically to ASDC

• Additional data DOIs can be issued to facilitate future publications
ACTIVATE Data Organization and DOIs

• Project level DOI: 10.5067/SUBORBITAL/ACTIVATE/DATA001

• ACTIVATE_Aerosol_AircraftInSitu_Falcon_Data_1 Collection:
  • In-situ observations of aerosol microphysical, optical, and chemical properties
  • DOI: 10.5067/ASDC/ACTIVATE_Aerosol_AircraftInSitu_Falcon_Data_1

• ACTIVATE_Cloud_AircraftInSitu_Falcon_Data_1 Collection:
  • In-situ observations of cloud physical, optical, and chemical properties
  • DOI: 10.5067/ASDC/ACTIVATE_Cloud_AircraftInSitu_Falcon_Data_1

• ACTIVATE_TraceGas_AircraftInSitu_Falcon_Data_1 Collection:
  • In-situ observations of O3, CH4, CO, and CO2
  • DOI: 10.5067/ASDC/ACTIVATE_TraceGas_AircraftInSitu_Falcon_Data_1
• ACTIVATE_MetNav_AircraftInSitu_Falcon_Data_1 Collection:
  • Navigational parameters, temperature, water vapor, and 3-D winds
  • DOI: 10.5067/ASDC/ACTIVATE_MetNav_AircraftInSitu_Falcon_Data_1

• ACTIVATE_MetNav_AircraftInSitu_KingAir_Data_1 Collection:
  • Navigational Parameters, dropsondes
  • DOI: 10.5067/ASDC/ACTIVATE_Aerosol_AircraftInSitu_Falcon_Data_1

• ACTIVATE_AerosolCloud_AircraftRemoteSensing_Falcon_KingAir_1 Collection:
  • RSP and HSRL remote sensing of cloud and aerosol properties
  • DOI: 10.5067/ASDC/ACTIVATE_AerosolCloud_AircraftRemoteSensing_KingAir_Data_1
• ACTIVATE_Miscellaneous_Data_1 Collection:
  • Camera forward cloud masks
  • DOI: 10.5067/ASDC/SUBORBITAL/ACTIVATE_Miscellaneous_Data_1

• ACTIVATE_Merge_Data_1 Collection:
  • Project generated merge files
  • Merge netCDF files are publicly available
  • DOI: 10.5067/ASDC/ACTIVATE_Merge_Data_1

• ACTIVATE_Model_Data_1 Collection:
  • MERRA-2 Reanalysis Model Data
  • DOI: 10.5067/ASDC/SUBORBITAL/ACTIVATE_Model_Data_1
• ACTIVATE-FLEXPART_1
  • FLEXible PARTicle dispersion model back-trajectories
  • DOI: 10.5067/ASDC/SUBORBITAL/ACTIVATE-FLEXPART_1

• ACTIVATE-MODIS-MERRA2_1
  • Merged MODIS and MERRA-2 dataset
  • DOI: 10.5067/ASDC/SUBORBITAL/ACTIVATE-MODIS-MERRA2_1

• ACTIVATE-Satellite_1
  • GOES-16 satellite data
  • DOI: 10.5067/ASDC/SUBORBITAL/ACTIVATE-Satellite_1
Open-Source Science: Findable and Accessible Data

Earthdata Search

- Search based on temporospatial criteria as well as keywords
- Retrieve data via collections (groupings of files)
- Data DOIs are issued for each collection
A variable centric discovery tool complementary to Earthdata Search

Aims to facilitate research data use
- Expert user interface for users with reasonable understanding of airborne field campaigns and airborne atmospheric composition measurements
- General user interface development is currently on hold

Variable (dataID) based data search - filtered by platform, flights, instruments, principal investigators

Provides field campaign relevant documentation, e.g., flight summary reports

Provides a merge service tool, an enhanced version of the field data repository online merge tool

API under development, which will make data more findable and accessible

Open Source Science: Findable, Accessible, and (Re)usable Data - Sub-Orbital Order Tool (SOOT) (https://asdc.larc.nasa.gov/soot/power-user/)
SOOT Power User Interface – 2 Approaches

Survey study: Few variables for all flights

Process study: All variables for 1 flight
As data are filtered, the number and size of the selected files are shown.

File selections can be made through the filters or the PI selection panes, e.g., PI name, dataID.

Switch between platforms and years to select and review additional files for download.

Once files are selected, click the green 'Review Selected Files' button to download or create merges for the mergeable files.
### Review Selected Files – Choose to Merge or Download

#### Selected 29 files

<table>
<thead>
<tr>
<th>Filename</th>
<th>Size</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAAMES-LARGE-AMS_C130_20151112_R0.ict</td>
<td>83.26 KB</td>
<td>□</td>
</tr>
<tr>
<td>NAAMES-LARGE-AMS_C130_20151114_R0.ict</td>
<td>69.36 KB</td>
<td>□</td>
</tr>
<tr>
<td>NAAMES-LARGE-AMS_C130_20151117_R0.ict</td>
<td>68.23 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-2DS-H-ARM_HU25_20200214_R2.ict</td>
<td>18.79 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-2DS-V-ARM_HU25_20200214_R2.ict</td>
<td>16.8 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-CAMERA-FORWARD_HU25_20200214_R0_L1.mp4</td>
<td>79.76 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-CAMERA-FORWARD_HU25_20200214_R0_THRU20200930_README.docx</td>
<td>16.04 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-DLH-H2O-20HZ_HU25_20200214_R0.ict</td>
<td>5.98 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-DLH-H2O_HU25_20200214_R0.ict</td>
<td>274.96 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-FCDP_HU25_20200214_R0.ict</td>
<td>1.9 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-AMS-CVL_HU25_20200214_R2.ict</td>
<td>4.19 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-AMS_HU25_20200214_R2.ict</td>
<td>40.77 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CAS_HU25_20200214_R0.ict</td>
<td>1.46 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CCN_HU25_20200214_R0.ict</td>
<td>200.02 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CLOUDWATER_HU25_20200214_R0.ict</td>
<td>1.62 MB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CDP_HU25_20200214_R0.ict</td>
<td>12.1 KB</td>
<td>□</td>
</tr>
<tr>
<td>ACTIVATE-LARGE-INSTFLAG_HU25_20200214_R1.ict</td>
<td>130.46 KB</td>
<td>□</td>
</tr>
</tbody>
</table>

**Beta Testing - Do not cite merge results**

- **Merge Files**
- **Click here to provide feedback on merge capability**
- **Download Files**

---

**ACTIVATE**

Support Documentation

**Platforms:**
- **ANALYSIS**
- **MERGE**
- **MODEL**
- **SATELLITE**
- **UC12**

**Principal Investigators:**
- ANDREA CORRAL
- GLENN DISKIN
- RICHARD MOORE
- LEE THORNHILL
- CHRISTIANE VOIGT
- EDDIE WINTSTEIN
- LUKE ZIEMBA

**Data IDs:**
- ACTIVATE-2DS-H-ARM [*variables*
- ACTIVATE-2DS-V-ARM [*variables*
- ACTIVATE-CAMERA-FORWARD
- ACTIVATE-DLH-H2O [*variables*
- ACTIVATE-DLH-H2O-20HZ [*variables*
- ACTIVATE-FCDP [*variables*
- ACTIVATE-LARGE-AMS [*variables*
- ACTIVATE-LARGE-AMS-CVL [*variables*
- ACTIVATE-LARGE-CAS [*variables*
- ACTIVATE-LARGE-CCN [*variables*
- ACTIVATE-LARGE-CLOUDWATER [*variables*
- ACTIVATE-LARGE-CDP [*variables*
- ACTIVATE-LARGE-INSTFLAG [*variables*

**Selected 29 files (152.98 MB)**

- Remove All

---

Clouds cover more than 45% of the ocean surface and exert a large influence on climate. The ACTIVATE project is a five-year project (January 2019-February 2024) focused on improving our understanding of the roles of aerosols, cloud properties, and climate. The project aims to provide improved representations of aerosol and cloud properties in models and to better understand their impact on climate. The project will use data from a range of platforms, including satellites, aircraft, and ground-based instruments, to study the processes that govern the formation and evolution of clouds. The data collected will be used to improve our understanding of the interactions between aerosols, clouds, and climate, and to develop new models that can better predict future climate changes.
SOOT - Merge Service

- Combines data from the same platform and flight onto a common time base.
- Two main steps:
  1. Variable Selection
  2. Time Base Selection
- Currently only merges ICARTT files. In the future will work with other formats.
- Actively seeking users to test merge service and provide feedback. Feedback Form
1. Variable Selection:
   - Expand dataIDs to allow users to select variables
   - All variables are automatically selected.

2. Time Base Selection:
   - Choose a generic time base, e.g., 60 sec
   - Choose a dataID time base, e.g., LegFlags, to merge data based on ACTIVATE flag time base
SOOT Merge Service – Download

Thanks Morgan for using SOOT!

Your download is ready! Click the button below to download.

Download (2.9 GB)

Go back to the previous page
Go to the Home page

- Zipped file with merged and unmergeable files
- Folders within zipped file for each field campaign/platform if chose multiple
SOOT Merge Service – Algorithm

- Weighted time averages of different source data, i.e., individual measurement data, onto to the target (common) time base.
- Target time base can be a continuous time with a constant interval or the time base of an individual measurement.
- The weighting factor calculation is based on the overlap between a source sampling time interval and the target time interval or merge time interval.
- Mathematical calculations handle scalar, vector, wind measurements, data flag (including missing data flags and detection limit flags) variables accordingly.
SOOT - Custom Merge Output

- Currently only outputs .ict files
- PI file header not included
- Traceability to PI files via dataID given in variable long name
- Future output in netCDF
  - CF compliant – more interoperable and usable, especially for modelers
- Capability to better handle metadata, e.g., PI file headers
SOOT - Project Merge

- Project merge from field data repository, made by science team
- Both .ict and .nc files
- ACTIVATE .nc merge files by Sanja Dmitrovic
- Need to check if the data is up to date
SOOT - NetCDF Project Merge

<table>
<thead>
<tr>
<th>Name</th>
<th>Long Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTIVATE-mrg60-HU25-NC...</td>
<td>Local File</td>
<td></td>
</tr>
<tr>
<td>Abs470_total</td>
<td>Dry absorption</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Abs532_total</td>
<td>Dry absorption</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Abs660_total</td>
<td>Dry absorption</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>ACTIVATE-2DS-V-ARM</td>
<td>ACTIVATE-2DS-V-ARM</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-DLH-H2O</td>
<td>ACTIVATE-DLH-H2O</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-FCDP</td>
<td>ACTIVATE-FCDP</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-AMS</td>
<td>ACTIVATE-LARGE-AMS</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CAS</td>
<td>ACTIVATE-LARGE-CAS</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CCN</td>
<td>ACTIVATE-LARGE-CCN</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-CDP</td>
<td>ACTIVATE-LARGE-CDP</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-INLET</td>
<td>ACTIVATE-LARGE-INLET</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-MICROP</td>
<td>ACTIVATE-LARGE-MICROP</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-OPTICAL</td>
<td>ACTIVATE-LARGE-OPTICAL</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-PILS</td>
<td>ACTIVATE-LARGE-PILS</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LARGE-SMP5</td>
<td>ACTIVATE-LARGE-SMP5</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-LEGFLACS</td>
<td>ACTIVATE-LegFlacs</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-SUMARY</td>
<td>ACTIVATE-SUMARY</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-TRACECAS-CH4</td>
<td>ACTIVATE-TraceGas-CH4</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-TRACECAS-CO2</td>
<td>ACTIVATE-TraceGas-CO2</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-TRACECAS-CO3</td>
<td>ACTIVATE-TraceGas-CO3</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-TRACECAS-O3</td>
<td>ACTIVATE-TraceGas-O3</td>
<td></td>
</tr>
<tr>
<td>ACTIVATE-WINDS</td>
<td>ACTIVATE-WINDS</td>
<td></td>
</tr>
<tr>
<td>AAbs_DRY_470to600nm</td>
<td>derived Angstrom</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>AEscat_450to700nm</td>
<td>derived Angstrom Expon</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>AEscatAM_450to700nm</td>
<td>derived AMB安东姆安格...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>alt</td>
<td>average gps MSL altitud...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Ammonium_PILS</td>
<td>Ammonium Mass Concentr...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Bromide_PILS</td>
<td>Bromide Mass Concentr...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Calcium_PILS</td>
<td>Calcium Mass Concentr...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>CCN SS</td>
<td>Water vapor supersatur...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>CH4</td>
<td>Methane dry mole fracti...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Chl_Ave_Isok_STP</td>
<td>Chl Ave Isok STP</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Chloride_PILS</td>
<td>Chloride Mass Concentr...</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Cng10mm</td>
<td>Number concentration</td>
<td>GeoTraj</td>
</tr>
<tr>
<td>Cng10mm_nonvol</td>
<td>Number concentration</td>
<td>GeoTraj</td>
</tr>
</tbody>
</table>

Group "ACTIVATE-2DS-H-ARM"

In file "ACTIVATE-mrg60-HU25-NC_merge_20200214_R3.nc"

- **group attributes:**
  - `dataProvider = "Voigt, Christiane";
  - `dataProviderAffiliation = "DLR e.V.";
  - `DataProviderContact = "Christiane.Voigt@dlr.de";
  - `InstrumentDescription = "Cloud Droplet Information Data from SPEC 2DS Probe Horizontal Direction";
  - `DataInfo = "2DS measurements are reported at ambient temperature and pressure. All measured Data uncertainty = "Depend on instrumental and ambient conditions. Please contact the PI or DM for more information.";
  - `LowerDetectionLimit = "N/A";
  - `UpperDetectionLimit = "N/A";
  - `OtherComments = "Bin Lower Edges (in um) = [28.50, 39.90, 51.30, 62.70, 74.10, 85.50, 96.90, 100."
  - `Revision = "R2";
  - `RevisionNotes = "R2: Final data - log normalization base change to 10. R1: Final data - INC and;
  - `sourceFile = "ACTIVATE-2DS-H-Arm_HU25_20200214_R2.ict";
  - `NotIncludedInMerge = 384; // int
  - `Title = "ACTIVATE-2DS-H-Arm";
  - `NativeTimeResolution = "1 s";
  - `instruments = "2DS";

- **NetCDF project merge from the field data repository**
- **CF compliant: file structure, e.g., variables properly dimensioned and global and variable attribute names**
- **Use standardized units:** [http://codes.wmo.int/wmdr/unit](http://codes.wmo.int/wmdr/unit)
- **Individual file header information captured as group attributes**
• Flight reports
• PI supplied Data Readme files (e.g., ACTIVATE-DROPSONDE_UC12_20200214_R0_20200930_README.pdf)
• Related website links (e.g., ACTIVATE project website)
• General project related data files, not flight specific
• Flight videos
• List of publications
Thank you