

Title: ACCLIP NCAR GV Instrument Data Merges

Merge Version: R0

Author:

Shawn Honomichl

National Center for Atmospheric Research (NCAR), Boulder, CO 80301

Atmospheric Chemistry Observations & Modeling Laboratory (ACOM)

Ph: 303-497-8194

Email: shawnh@ucar.edu

ORCID: <https://orcid.org/0000-0003-4373-1121>

1. Data Description

These data contain all the available NCAR GV instrument data (and interpolated model outputs for applicable variables) collected from the ACCLIP field campaign, merged into one file per research flight (14 research flights in total), and when applicable, all flights merged into a single composite file.

2. Disclaimer

These data merges directly reflect the data quality of NCAR GV instrument data and should only be considered ‘final’ publishable data quality if the version number is R0 or greater. If any errors or anomalies are found in the merges, please email the author so that they can be fixed and/or resolved.

3. Data Processing

Data merges are created by sampling, subsampling, or averaging all the various input data sources onto a uniform time range and time interval so that all the input data sources can be compared independent of the input source’s sampling strategy. The time range and time interval of a particular merge type is determined from a selected input data source, and multiple merge types are created from different input data sources. The input data source that the rest of the merge is timed from is referred to in the rest of this document as the ‘master dataset’. The master dataset may take various forms, such as 1 second data, or data from an instrument that samples for a set amount of time on uniform, or non-uniform time intervals. For NCAR GV ACCLIP merges, there are 7 merge types, and the input data sources utilized as the master dataset originate from core GV data (1, 10, and 60 second), AWAS, GTCIMS, ERICA, and TOGA instrument datasets (see section 4). Additionally, for each merge type (except the 1-second merge) a ‘composite’ of all the flights is appended together into a single file to simplify data analysis.

The way in which a particular set of data is represented in the merge depends on the sampling strategy of a data source relative to the sampling strategy of the master dataset. Most of the data sources for ACCLIP merges are aircraft or instrument data, though model data may also be utilized. Any model data added to the merge is interpolated from its native grid to the aircraft track using linear time-space interpolation. The following two conditions outline the possible combinations of ‘master’ and other instrument types that are applicable for ACCLIP data:

a. 1-Second Master Dataset Merge

- i. Other one second data: directly entered in the merge as it is in the data file where there is a valid UTC time that matches the UTC time in the master data file. If the times are mismatched by a fraction of a second (example: master is in whole seconds, while merged instrument data is recorded at X.5 seconds), the instrument data will get treated as if it is at whole seconds since the software looks to average any data found lying on or between the UTC time of a given datapoint and the UTC time of the next given datapoint in the master dataset.

- ii. Non-1-second data with a start/stop/mid time for each data point: all 1-second UTC times that fall between the start/stop times of each data point are filled with the given instrument value. This creates a repeat of the same values in the 1-second merge between the start/stop UTC times of each instrument datapoint. In a 1-second merge, these data are less useful for data analysis, and are typically more useful for visualizing the data in software such as the NCAR EOL 'AEROS' software (<https://www.eol.ucar.edu/software/aeros>) when the 'trace style' is set to 'sticks'. This could also be of interest if wanting to calculate the variance or standard deviation of an averaged variable from a different merge type. All datasets of this type are assigned an observation number so that the repeating data points can be properly identified.
- b. Non-1-Second Master Dataset Merge with start/stop/mid times
 - i. 1-second data with UTC times that lie between the UTC start/stop time of the individual data points of the master dataset are averaged.
 - ii. Non-1-second data with start/stop time(s) that overlap the start/stop time of the individual data points of the master dataset is also averaged if more than 1 overlapping data point is present. If only 1 data point overlaps, then that value is used. No overlapping data points is set to the missing value for that variable.
- c. 10 and 60-Second Merges with 1-Second Master Dataset
 - i. The 10 and 60-second merges are assigned start/stop/mid times based off the master 1-second dataset, where the times are incremented by 10 or 60 seconds (usually on the top of the minute), and data are then treated the same way as the Non-1-Second Master Dataset with start/stop/mid times.

4. Merge Types

There are currently 7 different NCAR GV merge types available for ACCLIP:

- a. 1 second merge – 'master' data are the 1-second NCAR GV core instrument files.
- b. 10 second merge – 'master' data are the 1-second NCAR GV core instrument files.
- c. 60 second merge - 'master' data are the 1-second NCAR GV core instrument files.
- d. ERICA merge – 'master' data are the ERICA-AMS data that collects a sample once every 10 seconds with a +/- 2.2 second observation window.
- e. TOGA merge – 'master' data are the TOGA data that has a typical observation window of 20 seconds and collects a data point about once every 2 minutes.
- f. AWAS merge – 'master' data are the AWAS data that has a variable observation window depending on the altitude and collects a maximum of 60 samples per flight.
- g. GTCIMS merge – 'master' data are GTCIMS data that samples continuously observation window of approximately 5 seconds with no gap between measurements.

5. Output Formats

There are two file output formats: ICARTT and NetCDF. The ICARTT format is a formatted ascii file while NetCDF is a binary format. NetCDF files contain more metadata information since each variable in the file is allowed to contain variable attributes. The NetCDF files also contain global attributes with information such as a list of the input files used to generate the merge, what the flight date and research flight number are, aviation

waypoints, etc. It should be noted that the merged ICARTT files will not pass the NASA ICARTT file format check because some variables do not contain a ‘proper’ nomenclature. This is intentional since by default, the NCAR EOL AEROS software looks for specific variable names that do not necessarily conform to the most current ICARTT variable naming standards. While AEROS only utilizes the NetCDF files, the variable names between the two formats were chosen to remain the same to maintain as much continuity between the two formats as possible. Additionally, for ‘composite’ merge files that contain campaign data from every flight in one file, there is additional time information included that also does not conform to the latest ICARTT specification.

6. Instrument Inventory

- a. Table 1 shows which research flights NCAR GV instrument data is available and which flights instrument data is not available for the most current version of merge. Refer to the individual instrument ‘READ_ME’ documents located at https://data.eol.ucar.edu/master_lists/generated/acclip/ for more information on missing data.

| Date | 2022 0730 | 2022 0804 | 2022 0806 | 2022 0807 | 2022 0812 | 2022 0815 | 2022 0816 | 2022 0819 | 2022 0822 | 2022 0823 | 2022 0825 | 2022 0826 | 2022 0829 | 2022 0830 |
|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Flight | RF01 | RF02 | RF03 | RF04 | RF05 | RF06 | RF07 | RF08 | RF09 | RF10 | RF11 | RF12 | RF13 | RF14 |
| GV-CORE | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| AWAS | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| ERICA-AMS | No Data | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | No Data | No Data |
| FO3 | R0 | No Data | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| AERO DYNE | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 |
| PICAR RO | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 | R1 |
| GT-CIMS | R0 | R0 | R0 | R0 | No Data | No Data | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| HARP | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| NMAS S | R0 | R0 | No Data | R0 | R0 | R0 | R0 | R0 | No Data | R0 | R0 | R0 | R0 | R0 |
| NO/N O2 | R0 | No Data | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| SP2 | No Data | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| TOGA | R0 | No Data | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 |
| UHSAS | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | R0 | No Data | R0 | R0 |

Table 1. NCAR GV ACCLIP instrument inventory current as of latest merge creation date.

- b. On days where instrument data are not available, the instrument variables are retained in the merge but filled with missing data so that a consistent index of variable names is maintained throughout the merges.
- c. In the NetCDF version of the merges, if an instrument and its associated variables do not have data for a particular flight, the ‘Category’ attribute that lists the instrument’s name will have the text ‘_[MISSING]’ attached to it to identify that the instrument and associated variable does not have any valid data for that flight.