ICARTT File Format Standards V2.0

Status of this Memo

This RFC document describes the International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) file format standards V2.0, with some available implementations and resources.

Distribution of this memo is unlimited.

Change Explanation

This document combines the ICARTT File Format Standards described in ESDS-RFC-019v1.1 and the enhancements described in ESDS-RFC-029.

The resulting ICARTT File Format Standards V2.0 are backwards compatible with the original version.

This document deprecates the original version in ESDS-RFC-019. New implementations should adhere to the version described here.

Updates from V1 to V1.1:

Section 2.1.B Location Information – Added clarification about precision of measurements. "It is recommended that the decimal latitude and longitude be reported at maximum instrument precision. For typical aviation GPS instruments, the latitude and longitude should be reported to at least five decimal places. Altitude is recommended to be reported in meters."

Section 2.6 File Scanning Software – Corrected grammatical error in first sentence.

Updates from V1.1 to V2.0:

In 2015, the NASA Earth Science Data Information Systems (ESDIS) established an Earth Science Data System Working Group (ESDSWG) focused on updating the ICARTT 1.1 standards, which have been actively used for over a decade. The working group consisted of data users, instrument scientists, data system developers, and metadata experts. Drawn from the past experiences of working group participants, the general approach was to identify the major deficiencies and seek solutions, with an overarching objective of making this standard more rigorous and enhancing machine readability, while maintaining backwards compatibility. The updates listed below have been evaluated from all perspectives of data reporting, management, archival, and use.

Section 2.1.1. Character Requirements – The ICARTT format supports UTF-8, which has become the standard encoding for extended characters. However, three parts of the ICARTT file have limits on the characters used, i.e., filename, variable short and standard names, and data section.

Section 2.1.1. Character Requirements – Limit all variable short and standard names to 31 characters, and require that all start with a letter.

Section 2.1.2. Time Information – Require that the standard names for the time variables be Time_Start, Time_Stop, and Time_Mid.

Section 2.1.4.3. Limits of Detection (Multiple ULOD/LLOD Flags or Values) – Explicitly describe how to provide multiple ULOD and LLOD flags and values in FFI 1001 files. The current standards now complement the documentation of FFI 2110 and 2310 files, with changes to improve machine readability of LOD information and to maintain consistency with the rest of the ICARTT header.

Section 2.2. File Names (Revision Identifier) – Clarify the use of the R# data revision identifier. Letter revision identifiers (e.g., “RA”) can be used for field data, while numeric revision identifiers (e.g., “R0”, “R10”) can be used for preliminary and final data files.

Section 2.3.2.1. Format Version Number – Add a field to the end of the first header line for the ICARTT format version number.

Section 2.3.2.9. Independent Variable Definition – Add a required standard variable name entry before the long name in the variable name line.

Section 2.3.2.13. Dependent Variable Definition – Add a required standard variable name entry before the long name in the variable name line.

Section 2.3.2.17 Normal Comments – Clarify the standards for the normal comments section, including required keywords, order of the keywords, keyword/value pairs, and formatting for the keywords.

Copyright Notice

Abstract

This document lists the ICARTT 2.0 standards, which incorporated the ICARTT 2.0 updates into the ICARTT 1.1 document and improved the readability of the original document. The ICARTT 2.0 updates were proposed by members of the airborne data community (i.e., data users, instrument scientists, data system developers, and metadata experts), for formatting and describing ICARTT data, such that the format will support more rigorous standards for metadata and substantial enhancements for machine-readability. Included recommendations address general file format specifications and structure for metadata, such as variable standard name, format version number, data revision identifier, and time information structures. These updates were approved by NASA ESDIS Standards Office in September, 2016.

The ICARTT file format standards were developed to fulfill the data management needs for the **I**nternational **C**onsortium for **A**tmospheric **R**esearch on **T**ransport and **T**ransformation (ICARTT) campaign in 2004. The ICARTT study consisted of eleven highly coordinated individual field experiments with over 300 government-agency and university participants from five countries, i.e., US, Canada, UK, Germany, and France. A common and simple-to-use data file format, ICARTT file format was established for this study to primarily facilitate data exchange and to promote collaborations among the science teams for achieving the ICARTT science objectives. The ICARTT file format is text-based and composed of a header section (metadata) with critical data description information (e.g., data source, uncertainties, contact information, and brief overview of measurement technique), and a data section. Although it was primarily designed for airborne data, the ICARTT format proved to be practical for other mobile and ground-based studies and various data types. Upon the success of the ICARTT study, the ICARTT file format has since been widely accepted in the atmospheric composition field

study community and used in recent major airborne studies sponsored by NASA, NSF, NOAA, and international partners.

TABLE OF CONTENTS

1. Introduction – Origin of the ICARTT file format standard

Since the early 1980s NASA and partner agencies have conducted over 30 major tropospheric airborne field campaigns to investigate atmospheric composition over a wide range of geographical regions. Compared to satellite data, airborne data provides a longer historical perspective, a more extensive suite of observed species/parameters, and higher spatial resolution both horizontally and vertically. Consequently, airborne observations are of unique value for the modeling community to assess its ability to predict future atmospheric composition and its impact on climate change and air quality issues. Furthermore, airborne observations can also be used effectively to develop and/or improve the a priori data used in satellite retrieval algorithms. Therefore it is essential to maintain the standardized file format for airborne observational data, which can serve various research needs.

The ICARTT file format standards were originally developed for the **I**nternational **C**onsortium for **A**tmospheric **R**esearch on **T**ransport and **T**ransformation (ICARTT), an airborne field study which was conducted in summer 2004. The ICARTT study consisted of eleven independent but highly coordinated field experiments, which required exchanging a large number of files among all science team members and collaborating partners. A common data file format, the ICARTT format, was created to accommodate data sharing among the science teams and to fulfill the data management needs for all phases of field study, i.e., field deployment, post-deployment data processing and analysis, and publications.

The ICARTT file format is a text-based, self-describing, and relatively simple to use file structure. The file format was built on two well-established airborne data formats: NASA Ames and GTE. Like its predecessors, the ICARTT file format is designed for handling airborne in-situ measurement data but has limited capability to accommodate data from airborne or ground-based remote sensing (e.g., lidar), ground-based in-situ measurements, and satellite data relevant to a particular flight or an area of interest. The ICARTT format is composed of two sections: a header (metadata) section and a data section. The header section has the instructions for extracting data from the file and the critical information describing the data (e.g., data source, contact information, brief description of measurement technique, measurement uncertainties, and data revision comments) so that a user has sufficient information to either make direct use of the data or contact the measurement PI to get further clarification. The data section contains the observational data in comma-separated format.

Because of the ICARTT field study, the ICARTT data file format was exposed to a broad range of airborne instrument scientists and data users, e.g., modelers. The success of the ICARTT study naturally led to even wider acceptance of the ICARTT file format in later airborne studies. For example, the ICARTT file format was adopted in NASA INTEX-B and NSF MILAGRO field studies in 2006. In 2008, the international polar year POLARCAT field study used the ICARTT file format as the standard for the participating programs sponsored by NASA, NOAA, and international partners in France and Germany. Two of the five NASA Earth Venture Suborbital Program projects adopted the ICARTT format as their data reporting standards in 2011. The growing acceptance and wide use, especially in the airborne in-situ measurement community, has allowed the ICARTT data file format to be recognized as one of the standards for the airborne field study community.

This document provides a complete description of the ICARTT 2.0 standards, by incorporating the updates from ICARTT V1.1 to V2.0 and improving the readability of the original document. The ICARTT file format 2.0 standards are defined in Section 2, which includes format specification, file naming convention, header section specification, and applications to various data types. Several examples are provided for further clarification. Also given here is a brief description of the file scanning software that can be used to test data files for compliance with the ICARTT file format standards.

2. File Format Specifications

Airborne field studies, such as ICARTT 2004, often collect different types of data (e.g., in-situ and remote sensing measurements) using one or more independent variables for data reporting. In-situ observational data sets where all dependent variables are measured simultaneously can be reported as a time series. In this case, only one independent variable, i.e., sampling time, is required. However, remote sensing data sets are often multidimensional, and therefore require more than one independent variable to properly describe the measurements. For example, wind profiler data measures wind speed, wind direction, and temperature as a function of both altitude and time. Such data cannot be represented as a single time series.

Sections 2.1 - 2.2 below outline the ICARTT format for different types of data. They maintain an emphasis on standard time-series types of data, which is typical from in-situ measurements. Section 2.3 is specific for standard time-series types of data. Section 2.4 discusses multidimensional data. Section 2.5 offers guidance for non-standard data.

2.1 File Content Overview

This section describes the standards related to character encoding, measurement time and location reporting, and measurement uncertainties.

2.1.1 Character Requirements

The ICARTT format supports UTF-8, which has become the standard encoding for extended characters. Three parts of the ICARTT file have limits on the characters used:

1. Filename: Uppercase and lowercase ASCII alphanumeric characters (i.e. A-Z, a-z, 0-9), underscore, period, and hyphen. File names can be a maximum 127 characters in length.
2. Variable short names and variable standard names: Uppercase and lowercase ASCII alphanumeric characters and underscores. The first character must be a letter, and the name can be at most 31 characters in length. Note the variable standard names must be selected from a controlled list provided by the data manager and cannot be created by the ICARTT format users.
3. Data section: ASCII numeric characters (including plus and minus signs, decimal point, and scientific notations), commas as delimiters, and spaces for the purpose of visual clarity (alignment) of data.

The end-of-line (EOL) character for text files differs by operating system. Many modern text utilities handle and convert the EOL character automatically, and to the vast majority of users this problem is transparent and a non-issue. It is nonetheless a problem that some users will encounter, and data managers should be aware of it. There are many resources discussing this issue in detail with remedies to overcome it. Potential solutions include using a different (newer) text editor (e.g., notepad++) or using the "ASCII mode" when transferring files via ftp. Some file scanning software, including "FScan" (discussed in Section 2.6), will automatically handle the EOL character when necessary.

2.1.2 Time Information

There are three ways that time information may be reported in ICARTT files, depending on the type and frequency of data being collected.

The default method that will apply to most files is to report both the start and stop times for each measurement sampling period. This is the only unambiguous way to define the measurement integration time. The mid time must be reported and defined when the sampling mid point time is not the center between the start and stop times. For example, the mid time for the whole air sampler may be reported as the time when the canister is half filled. Data reported using this method need not be continuous, nor must the measurement integration time be consistent. This method uses a data interval code of 0 (detailed in Section 2.3.2.8).

For data collected continuously at a constant 1 Hz or higher frequency with a constant measurement sampling interval, measurements may be represented by a single timestamp. Using this method, the reported timeline must be unbroken between the first and last reported measurements; missing data identifiers must be used to account for data gaps due to calibration or other periods of instrument down time. This method uses a data interval code equal to the interval between measurements, in seconds.

Non-standard data (e.g., satellite, trajectory, and ground data) may be represented by a single timestamp, and data gaps do not need to be filled in with missing data identifiers (since long gaps and already large data volume can create excessively large files). That is, each measurement is identified by a single timestamp, but the data timeline is not continuous. These data use a data interval code of -1. Note that this type is rarely used and requires approval by the project scientist and the data manager to accommodate special research needs. See Section 2.5 for more information about non-standard datasets.

For all time variables, time is reported as seconds UTC from the start of the date (i.e., midnight) on which measurements began. This date appears in both the file header and filename. The reported time must be monotonically increasing even when crossing over to a second day. Variable standard names must be one of [Time_Start, Time_Stop, Time_Mid], as appropriate.

2.1.3 Location Information

All data points need to have an associated sampling location, i.e., latitude, longitude, and altitude. Latitude and longitude must be reported in decimal degrees with south latitudes and west longitudes represented as negative numbers (i.e., no N, E, W, S identifiers). It is recommended that the decimal latitude and longitude be reported at maximum instrument precision. For typical aviation GPS instruments the latitude and longitude should be reported to at least five decimal places. Altitude is recommended to be reported in meters. Altitudes must be explicitly defined since many types of altitude measurements are in use (e.g., pressure altitude, GPS altitude, geopotential altitude, radar altitude, etc.).

By default, latitude, longitude, and altitude must be included as dependent variables in the data section. However, it is often advantageous and necessary to report location (sometimes along with basic meteorology) information in an independent file (e.g., an aircraft parameter file). In that case, it is not necessary to report this information in the data files for each instrument on the platform. Instead, they may simply refer to the sampling location parameter file name or dataID in the LOCATION keyword of the file header. This option is specified below in Section 2.3.2. If the sampling location does not change over the course of the file (e.g., for a ground site), the fixed latitude, longitude, and altitude may be listed in the LOCATION keyword of the file header rather than providing separate variables.

2.1.4 Measurements

In general, each file contains data of one parameter species. Multiple variables per file are allowed only if all were reported on exactly the same time base. For example, data from a single instrument measuring many species at the same sampling time or data from several synchronized instruments reporting related parameters at the same time stamps may be placed in a single file. The numeric value of a variable must be reported in the units listed on the corresponding header line. The ICARTT format contains the provision for a data scaling factor; however, it is recommended that all scale factors be 1 unless it is grossly inconvenient to do so. If very large or very small numbers are required, then they can be represented with exponential notation, e.g., 1.01E9 or 5.23E-6.

2.1.4.1 Uncertainties

Uncertainty is inherently associated with every measurement. The ICARTT data format requires reporting the TOTAL uncertainty (including all systematic and random effects). If available, the measurement accuracy and precision should be reported individually. If uncertainty estimates are available for each measurement period, the uncertainties can be reported in a separate column immediately following the associated data column in the file. This requirement can be relaxed if the uncertainty data can be reproduced by information in the header of the file. For example, if all uncertainties can be calculated by a function that has any given data point as input, then the formula can be included as header information. It is imperative that the sigma confidence interval (e.g., 1 sigma or 2 sigma) be reported with the uncertainty. Equally important, the units for the uncertainty must be explicitly reported in the file header. When absolute uncertainty is reported, the same unit should be used for the uncertainty as the associated measurement. For relative uncertainty, the value should be reported in percentage (e.g., 30% or 10%).

2.1.4.2 Missing Data

Missing data occurs when an instrument was not taking data due to calibration or instrument problems. Missing data are represented by negative numbers large enough to never be construed as actual data. For the ICARTT file format the value is -9999 (or -99999, etc.).

2.1.4.3 Limits of Detection

An instrument typically has an operational detection range, which is defined by the lower limit of detection (LLOD) and upper limit of detection (ULOD). When sampling conditions are beyond the instrument detection range and the measurements are not quantifiable, limit of detection (LOD) flags can be used to convey semi-quantitative information. The flag for data values GREATER THAN the ULOD is -7777 (or -77777, etc.), and the flag for data value LESS THAN the LLOD is -8888 (or -88888, etc.). These flags and the values of the upper and lower LOD (if used) are documented at specific locations in the header file (see below). If LLOD or ULOD values vary from point to point, they must be given in a separate column of data. It is noted that some investigators, instead of using LLOD flags, choose to tabulate all of their quantifiable data, including negative values for concentration. This allows the data to be further averaged, as the lower limit of detection decreases with longer integration time.

The value for the LLOD_FLAG keyword must consist of either a single flag value that applies to all dependent variables in the file or a comma-separated list of flag values with one flag value per dependent variable (where the order of the flag values corresponds to the order of the dependent variables). Each flag value must either be “N/A” or -888[8...], with enough 8s so that the flag value cannot be misconstrued as real data. The flag value must have at least three 8s and be at least one order of magnitude more negative than the most negative real data value in the corresponding dependent variable column (or file, if only one flag value is provided for the entire file).

The value for the ULOD_FLAG keyword must consist of either a single flag value that applies to all dependent variables in the file or a comma-separated list of flag values with one flag value per dependent variable (where the order of the flag values corresponds to the order of the dependent variables). Each flag value must be either "N/A" or -777[7...], with enough 7s so that the flag value cannot be misconstrued as real data. The flag value must have at least three 7s and be at least one order of magnitude more negative than the most negative real data value in the corresponding dependent variable column (or file, if only one flag value is provided for the entire file).

The value for the ULOD_VALUE/LLOD_VALUE keywords must consist of a single LOD value that applies to all dependent variables in the file or a comma-separated list of LOD values with one LOD value per dependent variable (where the order of the LOD values corresponds to the order of the dependent variables). Each LOD value must be one of the following: "N/A", a numeric value (i.e., the LOD is constant throughout the file), or the short name of a dependent variable column (i.e., the LOD changes over time and is recorded in the indicated dependent variable column).

Note that the ULOD_VALUE/LLOD_VALUE for Time_Stop and Time_Mid variables must be "N/A", as they have no limit of detection.

For multidimensional data formats: When identifying LOD flags or values using the comma-separated list method, the values for auxiliary variables must be provided first, followed by the values for primary variables. The order of the variables is maintained within each group. Examples of implementation are given in Section 2.4.2.

2.1.4.4 Data Delimiter Characters

Commas are used to delimit data fields within records (lines) of data in a file.

2.2 File Names

Features of different file naming conventions have been adapted here. File names for the ICARTT data format, limited to 127 characters or less, are defined as follows:

dataID_locationID_YYYYMMDD[hh[mm[ss]]]_R#[_L#][_V#][_comments].ict

The only allowed characters are: A-Z, a-z, 0-9 (that is, uppercase and lowercase alphanumeric characters), underscore, period, and hyphen. All fields not in square brackets are required. Fields are described as follows:

dataID: short description of measured parameter/species, instrument, or model (e.g., O3, RH, VOC, PTRMS, MM5), prefixed frequently by project acronym (e.g., KORUSAQ-O3). The dataID format shall comply with data management requirements issued by the archiving data center.
locationID: short description of site, station, platform, laboratory, or institute. The locationID shall comply with data management requirements issued by the archiving data center.
YYYY: four-digit year
MM: two-digit month
DD: two-digit day
hh: optional two-digit hour
mm: optional two-digit minute
ss: optional two-digit second
R: revision identifier of data
L: optional launch number
V: optional volume number
comments: optional additional information
extension: **ict** file extension, always "ict"

The underscore is used ONLY to separate the different fields of the file name as it has special significance for file-checking software (see Section 2.6). To separate characters within a field for readability, use lower and upper case letters. The square brackets, "[]", enclose elements of the filename that are optional; the brackets themselves are only for this indication and do not appear in the actual file names. Dates and times in file names are always UTC. The date and time in the file name give the date/time at which the data within the file begin (data files), or date/time at which the image applies (image files). For aircraft and sonde data files, the date always refers to the UTC date of launch.

The **dataID** is a short string of characters used to identify the parameters in the file. For files that contain one or two variables, those variable names can be used in the file name. For files in which many variables are represented, it may be best to indicate in the file name a class of compounds (e.g., "Hskping" for aircraft housekeeping data or "NOxyO3" for nitrogen oxide and ozone data) or an abbreviation of the instrument used to make the measurements (e.g., PTRMS). Each dataID needs to be unique within an airborne field study. It is a common practice that the dataID is prefixed by the project acronym. For example, the dataID for KORUS-AQ NOx, NOy, and O3 measurements is KORUSAQ-NOxyO3.

The **locationID** is used to identify the measurement platform, site, station, or source (laboratory or institute) of the information within a data file. Some examples could be: DC8, BAE146, RHBrown, GOME (satellite), IoS (Appledore Island site), ChebPt (Chebogue Point site), etc. It may be useful to have a standardized set of abbreviations used for a given field mission. These should be decided upon by the mission Science Team.

The **R** parameter is required in the ICARTT data format. One must specify a data revision identifier that tracks updates to the data. This also requires documentation of those updates (e.g., new calibration, timing error, etc.) to be recorded in the file header (see Section 2.3.2). For this we specify a revision identifier "_R#" where the underscore is required element to separate the fields (this is needed for certain file checking software). The revision identifier, "R#", must match the revision identifier specified in the Normal Comments section of the file header (see Section 2.3.2).

If filenames for field data need to be distinguished, campaigns may use revision identifiers starting from letter "A", e.g., RA, RB, ..., etc. The field data revision identifier must be a single letter and should be capitalized. Otherwise, revision identifiers must be numbers starting from "0", e.g., R0, R1,..., R10,..., etc. Numerical revision identifiers must be at most two digits. This practice is typically outlined in project data management plans.

The optional **L** parameter may be needed in some special cases. If the contents of the file pertain to a day on which multiple launches or aircraft flights occur, then a launch counter, "_L#" (i.e., L1, L2, etc.), must appear after the "R" identifier but before a volume counter (if present, see below). The L parameter is only to be used when there are multiple launches or aircraft flights in a single day.

The optional **V** parameter may be needed in some special cases. If a data file is one volume of a multi-volume dataset (i.e., one that requires more than one file per day), then a volume counter, "_V#" (i.e., V1, V2, V3, etc.), must appear after the "R" parameter (and the "L" parameter, if present) separated by an underscore from the rest of the identifier. The volume number, "#", must match the volume number in the file header. The V parameter is only to be used when part of a multi-volume dataset.

The optional comments parameter is for additional information required by the PI (or data manager) to identify the file contents but that does not fit into the other fields of the file name. This should be used sparingly.

2.3 File Format Specification for ICARTT Data Files

The ICARTT file format can handle different types of data, e.g., time series and arrays. The File Format Index (FFI) is defined to denote each type. The FFI can have 3 predefined values: 1001, 2110, and 2310. The time series file is associated with FFI value of **1001**, which mimics the Ames file format. The FFI values of 2110 and 2310 are for two types of array data files, which have two independent variables, e.g., time and altitude. Details of the array data file structures are given in Section 2.4.

The remainder of this section will discuss the details of the time series data file format.

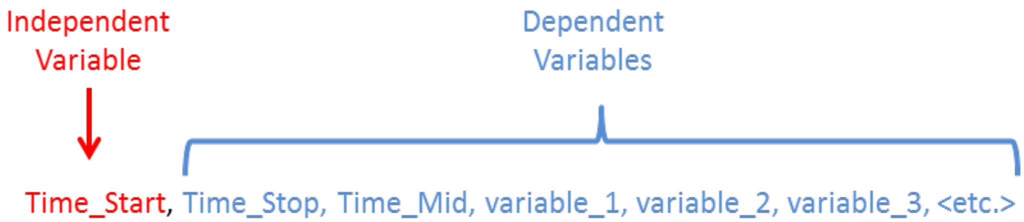
2.3.1 Structure

The ICARTT time series data file format is structured to mimic the Ames file format, **File Format Index (FFI) = 1001**. The definition of FFI in the Ames format is as follows: *The File Format Index (FFI) is used to uniquely define the exchange file format. By reference to pre-defined format options, the value of the FFI determines the number of INDEPENDENT variables, whether the values of the INDEPENDENT and dependent variables are numeric or character string, the format of the file header; and the format of the data records.*

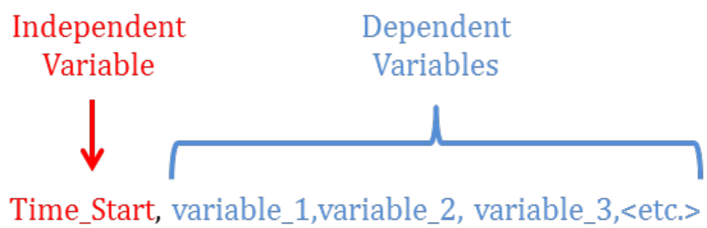
We recommend that, whenever possible, ICARTT time series data files conform to the file format **FFI = 1001**.

FFI = 1001: one real, unbounded independent variable; primary variables are real; no auxiliary variables; independent and primary variables are recorded in the same record

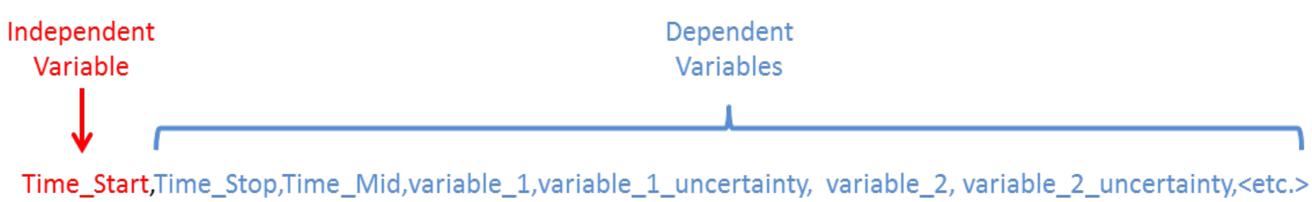
This indicates that there is one independent variable and that all other data depend on the independent variable. In the typical case, the independent variable is the start time of the measurement. Others can be defined as in the following example, where the variable names refer to a column in the data file:



This format accounts for most time series data measured any time, over any arbitrary integration period, and at any place on or above the planet. If measurements are reported continuously at 1 second intervals or less, the format can be condensed so that a single reporting time may be used.



If the sampling location parameters are reported in the same data file, latitude, longitude, and altitude/elevation shall be reported immediately after the timestamp variable(s). As described in Section 2.1.3, if the sampling location is constant throughout the file or is recorded in a separate file, then this information can be described in the header, and the latitude, longitude, and altitude need not be recorded as dependent variables. Similarly, if uncertainty is defined as some function that is the same for all data points, then that function can be included in the header information, and the user can then calculate uncertainties. Otherwise, the uncertainty shall be reported for each sampling interval as a separate data variable immediately after the associated data variable.



2.3.2 File Header Information

For the ICARTT data format, metadata information is required and included in the comments section. A typical header is shown below as an example. Delimiters to separate fields (items) are commas only. For delimiters to separate text within an item, use underscores. The order in which data appears in the header is listed below. The text below describes file header sections: in each numbered section in the file header, the items within each file header section are shown in italicized text, and the items are explained in the text that follows. Example headers are provided following these sections.

2.3.2.1 File Format Information

Number of lines in header, file format index, format version number
The number of lines in the header for FFI = 1001 files is defined in the following way: number of header lines = 14 + (# of dependent variables, given in line 10) + (# lines of special comments) + (# lines of normal comments).

File format index is described in Sections 2.3.1 and 2.4.1.

To clearly identify the version of ICARTT format, a version number is used to distinguish the specific standards aligned with that version. The version number is represented by “V##_YYYY”, where “##” denotes the version number and “YYYY” indicates the year this version was approved. The current ICARTT version number is V02_2016

2.3.2.2 PI Name

Last name, first name/initial

2.3.2.3 PI Affiliation

One-line statement of organization or affiliation of PI

2.3.2.4 Data Source Description

One-line, brief description of instrument name/platform name/model name, etc.

2.3.2.5 Mission Name

Mission acronym and deployment year

2.3.2.6 File Volume Information

File volume number, total number of file volumes
These integer values are used when the data requires more than one file per day. For data that require only one file, these values will both be 1. A single volume should be used unless grossly inconvenient for data users.

2.3.2.7 Data Collection and Revision Dates

YYYY, MM, DD (of collection), YYYY, MM, DD (of revision)
The first three fields of this line are the UTC year, month, and day when data collection began. The second three fields are the UTC year, month, and day of the most recent data revision.

2.3.2.8 Data Interval Code

Code for time spacing in seconds between consecutive records
This code describes the interval between values of the independent variable, as described in Section 2.1.2. For time series data, the data interval code is set to 1 for 1 Hz data and 0.1 for 10 Hz data. All intervals longer than 1 second must be reported with start and stop times, and the data interval code is set to 0. Mid time must be reported if it is not the exact midpoint of the interval. For additional information on non-standard data intervals, see Section 2.5.

2.3.2.9 Independent Variable Definition

Variable short name, variable unit, variable standard name, [optional variable long name]
Variable short name is the name chosen for the independent variable (usually measurement start time, i.e., Time_Start). Variable short name is limited to 31 characters, as described in Section 2.1.1. For measurement time reporting, units shall be the number of seconds since midnight UTC of the day on which the measurements began. The independent variable shall monotonically increase even when crossing over to a second day. Variable standard name for date/time variables must be one of [Time_Start, Time_Stop, Time_Mid], as appropriate. The optional variable long name is a more complete description of the independent variable without the limitations imposed on the variable short name or standard name. The variable long name can be as long as necessary and contain any UTF-8 character.

2.3.2.10 Number of Dependent Variables

Number of dependent variables
Integer value denoting the number of dependent variables. The total number of columns of data must be this number plus 1.

2.3.2.11 Scale Factors of Dependent Variables

Variable 1 scale factor, variable 2 scale factor, ...
Scale factors should be 1 unless grossly inconvenient for data users. For more information, see Section 2.1.4.

2.3.2.12 Missing Data Flags of Dependent Variables

Variable 1 missing data flag, variable 2 missing data flag, ...
These values denote the indicator for any missing data condition, with one per dependent variable. These values take the form of -9999 (-99999, etc.) and must be negative enough to never be construed as actual data. There is no missing data flag for the independent variable because it must never be missing. For more information on missing data, see Section 2.1.4.2.

2.3.2.13 Dependent Variable Definitions

Variable short name, variable unit, variable standard name, [optional variable long name]
Variable short name, units, and standard name are required, in that order. An optional long descriptive name can follow the variable standard name. All values are separated by commas. Variable short names are limited to 31 characters, as described in Section 2.1.1. The variable long name can be as long as necessary and contain any UTF-8 character. If the variable is unitless, enter "none" for its units. There must be exactly one line for each dependent variable. The variable short name must correspond exactly to the name used for that variable as a column header (i.e., the last header line prior to the start of data). In an effort to improve usability, standardization, and machine-readability, a standard variable name (from a controlled list) must be specified prior to the optional variable long name.

For data interval code of 0, the Time_Stop shall be the first dependent variable, followed by Time_Mid if necessary. When the Time_Mid is required, a definition of what the Time_Mid variable represents must be given in the long name entry (see discussion in Section 2.1.2).

2.3.2.14 Number of Special Comment Lines

Number of special comment lines
Integer value indicating the number of lines of special comments, NOT including this line.

2.3.2.15 Special Comments

Special comments are notes of problems or special circumstances unique to the containing file. An example would be comments/problems associated with a particular flight.

2.3.2.16 Number of Normal Comment Lines

Number of normal comment lines
Integer value indicating the number of additional lines of supporting information, NOT including this line.

2.3.2.17 Normal Comments

Normal comments are the place for investigators to describe the data and measurements. The normal comments section consists of three parts: an optional *free-form text section*, a required *"KEYWORD: value" pairs block*, and *the variable short name list* (last line), including both the independent and dependent variables.

The *free-form text section* consists of the lines between the beginning of the normal comments section and the first required keyword. This section may include, for example, user-defined "KEYWORD: value" pairs or information in unformatted text form regarding the dataset that does not fall under any of the required keywords. The user may choose not to include this free-form text section, in which case the normal comments section will begin with the first required keyword. It is noted that the user-defined keyword should be used only if the information cannot be adequately represented in the required keyword entries.

The required *"KEYWORD: value" pairs block* starts with the line that begins with the first required keyword and must include all required "KEYWORD: value" pairs in the order listed in the ICARTT documentation. For clarity, the required keywords and order are listed in Table 1 below. Each keyword must appear exactly once. Within this block, the value of a keyword continues up to the line containing the next keyword; that is, values spanning multiple lines are explicitly allowed. Where no value exists, "N/A" should be used.

Throughout the normal comments section, any keyword (user-defined or ICARTT-required) must appear at the start of a line, with no preceding spaces. Keywords must consist exclusively of capital letters and underscores, and must be followed by a colon and a space.

The last line of the normal comment section must always be the comma-separated list of variable short names.

Table 1: ICARTT-required keywords and keyword order

Order	KEYWORD	Description	N/A allowed?
1	**PI_CONTACT_INFO**	Phone number, mailing address, email address, etc.	**Y**
2	**PLATFORM**	Platform or site information	**Y**
3	**LOCATION***	Lat/lon/elevation for fixed ground site, the location data file or dataID for the mobile platform, or description of lat/lon/altitude variables in the current file	**Y**
4	**ASSOCIATED_DATA**	File names with associated data: location data, aircraft parameters, ship data, etc.	**Y**
5	**INSTRUMENT_INFO**	Instrument description, sampling technique and peculiarities, literature reference, etc.	**Y**
6	**DATA_INFO**	Information regarding data use, e.g., data reported in STP or dry mixing ratio	**Y**
7	**UNCERTAINTY**	Uncertainty information, whether a constant value or a function, if the uncertainty is not given as separate variables	**N**
8	**ULOD_FLAG****	-7777 (Upper LOD flag, always -7s)	**Y**
9	**ULOD_VALUE****	Upper LOD value (or function) corresponding to the -7777s flag in the data records	**Y**
10	**LLOD_FLAG****	-8888 (Lower LOD flag, always -8s)	**Y**
11	**LLOD_VALUE****	Lower LOD value (or function) corresponding to the -8888s flag in the data records	**Y**
12	**DM_CONTACT_INFO**	Data Manager for instrument group – Name, affiliation, phone number, mailing address, email address, etc.	**Y**
13	**PROJECT_INFO**	Study start & stop dates, web links, etc.	**Y**
14	**STIPULATIONS_ON_USE**	(self-explanatory)	**Y**
15	**OTHER_COMMENTS**	Any other relevant information	**Y**
16	**REVISION*****	Revision identifier; see file names section	**N**
17	**R# (current revision identifier)**	Comments specific to this data revision; correspond to the revision date on line 7	**N**
As needed	***All previous revision identifiers in descending order***	*Revision identifiers and associated comments are cumulative*	***N***

***See Section 2.1.3 for more details**
****See Section 2.1.4.3 for more details**
*****See Section 2.2 for more details**

2.3.3 Examples

Below are three examples of (similar) time series data using different forms of header information. Be aware that the automatic word-wrap feature in word processing programs gives the appearance that there are more lines of text than are really there. In these examples any continuation of lines from directly above has been indented for clarity.

Note that all standard variable names shown are temporary placeholders. These will be chosen from the controlled list, once available.

Example 1. Uncertainty data exists as a separate data column

File name: SEAC4RS-PTRMS-acetaldehyde_DC8_20130806_R1.ict

37, 1001, V02_2016
Wisthaler, Armin
University of Innsbruck
PTR-MS instrument, Acetaldehyde mixing ratios, A. Wisthaler
SEAC4RS
1, 1
2013, 08, 21, 2014, 10, 23
0
Start_UTC, seconds, Time_Start, number of seconds from 00:00 UTC
4
1, 1, 1, 1
-9999, -9999, -9999, -9999
Stop_UTC, seconds, Time_Stop, number of seconds from 00:00 UTC
Mid_UTC, seconds, Time_Mid, number of seconds from 00:00 UTC
Acetaldehyde_ppbv, ppbv, Acetaldehyde, volume mixing ratio
Acetaldehyde_uncertainty_ppbv, ppbv, UncertaintyData, volume mixing ratio
0
19
PI_CONTACT_INFO: Institut fuer Ionenphysik und Angewandte Physik,
   Technikerstrasse 25, 6020 Innsbruck, AUSTRIA
PLATFORM: NASA DC-8 Aircraft
LOCATION: Latitude, longitude and elevation data are included the corresponding
   “seac4rs-dc8hskping” files
ASSOCIATED_DATA: N/A

INSTRUMENT_INFO: Proton Transfer Reaction Mass Spectrometer
DATA_INFO: N/A
UNCERTAINTY: Precision as reported in the data file; accuracy of +/- 15 %.
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A
DM_CONTACT_INFO: Armin Wisthaler, Institut fuer Ionenphysik und Angewandte Physik, Technikerstrasse 25, 6020 Innsbruck, AUSTRIA; email: armin.wisthaler@uibk.ac.at
PROJECT_INFO: SEAC4RS
STIPULATIONS_ON_USE: The users are strongly encouraged to consult with the PI for proper data use.
OTHER_COMMENTS: Data have not been synchronized with the DLH water signal because of DLH timing issues. If exact synchronization is critical for the analysis, the user must develop and apply own method of synchronization.
REVISION: R1
R1: Final data.
R0: Preliminary data.
Start_UTC, Stop_UTC, Mid_UTC, Acetaldehyde_ppbv, Acetaldehyde_uncertainty_ppbv
64752.41, 64753.41, 64752.91, 0.289, 0.057
64768.17, 64769.17, 64768.67, 0.124, 0.057

Example 2. All required data columns are shown explicitly.

File name: DISCOVERAQ-NOXYO3_P3B_20140720_R0.ict

47, 1001, V02_2016
Weinheimer, A.J.; Montzka, D.D.
National Center for Atmospheric Research
P3-B in situ NO, NO2, NOy, O3
DISCOVER-AQ
1, 1
2014, 07, 20, 2015, 03, 11
0.0
StartTime_UTsec, seconds, Time_Start, Start Time in UT seconds
6
1.0, 1.0, 1.0, 1.0, 1.0, 1.0

-999999.9, -999999.9, -999999.9, -999999.9, -999999.9, -999999.9
StopTime_UTsec, seconds, Time_Stop, Stop Time in UT seconds
MidTime_UTsec, seconds, Time_Mid, Midpoint Time in UT seconds
NO_pptv, pptv, NO, Nitric Oxide Mixing Ratio
NOy_pptv, pptv, NOy_NO, Total Reactive Nitrogen Mixing Ratio
NO2_pptv, pptv, NO2, Nitrogen Dioxide Mixing Ratio
O3_ppbv, ppbv, O3, Ozone Mixing Ratio
0
27
PI_CONTACT_INFO: wein@ucar.edu, 303-497-1444, NCAR, 1850 Table Mesa Dr., Boulder, CO 80305
PLATFORM: NASA P3-B Aircraft
LOCATION: Latitude, Longitude, and Altitude of the P3 are included in REVEAL or PDS archive files
ASSOCIATED_DATA: N/A
INSTRUMENT_INFO: 4-channel chemiluminescence instrument
DATA_INFO: N/A
UNCERTAINTY: 10 pptv + 10%, 50 pptv + 20%, 20 pptv + 10%, 0.1 ppbv + 5%
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A
DM_CONTACT_INFO: Gao Chen, gao.chen@nasa.gov
PROJECT_INFO: N/A
STIPULATIONS_ON_USE: No restrictions on use, though users are strongly encouraged to contact PI.
OTHER_COMMENTS: N/A
REVISION: R0
R0: Final data.
No data for 20140720 (problem with main power to instrument). No NOy for 20140729 thru 20140810 (poor conversion).
Time sync to DLH data. Used version R1 for all except 20140807 for which R2 was used.
Water correction uses DLH data. Used version R1 for all except 20140807 for which R2 was used. For 2nd flight on 20140723, use empirical fit to Palt.
RB: Time adjustments and conversion to 2-file format for 4 flight days.
For 20140717, time was adjusted to match DLH (-730 s).
For 20140721, time was adjusted to match DLH (-730 s).

For 20140722, convert to two-file format; time was adjusted to match DLH (-42 s).
For 20140723, convert to two-file format; no change to time.
RA: Initial (~24-hr) archival, quick-look. Approximate calibration factors.
StartTime_UTsec, StopTime_UTsec, MidTime_UTsec, NO_pptv, NOy_pptv, NO2_pptv, O3_ppbv
51199.5, 51200.5, 51200.0, -999999.9, -999999.9, -999999.9, -999999.9
51200.5, 51201.5, 51201.0, -999999.9, -999999.9, -999999.9, -999999.9

Example 3: Latitude, longitude, and altitude are included

Filename: discoveraq-CO2_p3b_20140721_R0.ict

37, 1001, V02_2016
Yang, Melissa
NASA/LaRC
Non-dispersive IR Spectrometer measurements of CO2
NASA DISCOVER-AQ MISSION 2013
1, 1
2014, 07, 21, 2015, 01, 28
1.0
UTC, seconds, Time_Start, UTC time
4
1, 1, 1, 1
-9999, -9999, -9999, -9999
Lat, Degs, AircraftLatitude, Latitude
Lon, Degs, AircraftLongitude, Longitude
Alt, Feet, AircraftAltitude, Altitude
CO2_ppmv, ppmv, CO2, Carbon dioxide mixing ratio
1
FINAL Data
18
PI_CONTACT_INFO: NASA LaRC, MS 483, Hampton, VA 23681; 757-864-6943 (voice), 757-864-7790 (fax) e-mail: melissa.yang@nasa.gov
PLATFORM: NASA P3-B Aircraft
LOCATION: Latitude, Longitude, and Altitude included in data records
ASSOCIATED_DATA: N/A
INSTRUMENT_INFO: LI-COR 6252
DATA_INFO: N/A

UNCERTAINTY: +/- 0.25 ppmv
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A
DM_CONTACT_INFO: ali.a.aknan@nasa.gov
PROJECT_INFO: N/A
STIPULATIONS_ON_USE: Users of these data are expected to abide by the
DISCOVER-AQ Data Policy.
OTHER_COMMENTS: N/A
REVISION: R0
R0: Data time offset has been adjusted to provide maximum temporal registration with
DLH water vapor data.
UTC, Lat, Lon, Alt, CO2_ppmv
50428,39.91,-105.117,5381,424.935
50429,39.91,-105.118,5381,424.363

2.4 File Format Specification for ICARTT Multidimensional Data Files

2.4.1 Structure

ICARTT multidimensional data file formats are designed based on Ames standard file formats **FFI = 2110** and **FFI = 2310**. We recommend using these FFIs for exchange of most multidimensional data files. The FFI descriptor is:

FFI 2110: two real independent variables, one unbounded and one bounded, with its values recorded in the data records; primary variables are real; the first auxiliary variable is NX(m,1) (or, primary variables' array dimension), all other auxiliary variables are real.

FFI 2310: two real independent variables, one unbounded and one bounded, with its number of constant increment values, base value, and increment defined in the auxiliary variable list; primary variables are real; auxiliary variables are real.

For more details on these file types, please see the following documents:

Appendix B - Amended FFI 2110

Appendix C - Amended FFI 2310

2.4.2 Examples

Below are two examples of types FFI 2110 and FFI 2310

Example FFI 2110 using V02_2016 version

File name: PAVE-AR_DC8_20050203_R0.ict
55, 2110, V02_2016
PI Last Name, First Name
Code 916, Goddard Space Flight Center, Greenbelt, MD 20771
AROTAL
PAVE Mission
1, 1
2005, 02, 03, 2006, 01, 18
0, 1
Altitude[], meters, Altitude, Altitude_array
UTC, seconds, Time_Start, number of seconds from 00:00 UTC
7
0.1, 0.0001, 0.1, 0.01, 0.0001, 0.1, 0.0001
-999999, -999999, -999999, -999999, -999999, -999999, -999999
TempK[], K, Temperature, Temperature_array
Log10_NumDensity[], part/cc, Log10_NumDensity, Log10_NumDensity_array
TempK_Err[], K, Temperature_Error, Temperature_error_array
AerKlet[], Klet, Aerosol, Aerosol_array
Log10_O3NumDensity[], part/cc, Log10_O3NumDensity,
Log10_Ozone_NumDensity_array
O3_MR[], ppb, Ozone_mixing_ratio, Ozone_mixing_ratio_array
Log10_O3NumDensity_Err[], part/cc, Log10_O3NumDensity_Error,
Log10_NumDensity_error_array
11
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
-9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999
NumAlts, #, Number_of_altitudes, Number_of_altitudes_reported
Year, yyyy, Year_UTC, Year_UTC
Month, mm, UTC, Month_UTC, Month_UTC
Day, dd, UTC, Day_UTC, Day_UTC

AvgTime, minutes, Averaging_time, Averaging_time_of_presented_data xxx.x_minutes
Lat, degrees, Latitude
Lon, degrees, Longitude
PAlt, meters, pressure_altitude
GPSAlt, meters, GPS_altitude
SAT, K, Static_air_temperature
SZA, degrees, Sun_Zenith_Angle
1
Enter any special comments specific to this flight here
18
PI_CONTACT_INFO: Enter PI Address here
PLATFORM: NASA DC8
LOCATION: Lat, Lon, and Alt included in the data records
ASSOCIATED_DATA: N/A
INSTRUMENT_INFO:N/A
DATA_INFO:N/A
UNCERTAINTY: Contact PI
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A
DM_CONTACT_INFO: Enter Data Manager Info here
PROJECT_INFO: PAVE MISSION: Jan-Feb 2005
STIPULATIONS_ON_USE: Use of these data should be done in consultation with the PI
OTHER_COMMENTS: N/A
REVISION: R0
R0: Version 2005-0: AROTAL T & O3 Rayleigh Retrievals.
UTC, NumAlts, Year, Month, Day, AvgTime, Lat, Lon, PAlt, GpsAlt, SAT, SZA,
  Altitude[], TempK[], Log10_NumDensity[], TempK_Err[], AerKlet[],
  Log10_O3NumDensity[], O3_MR[], Log10_O3NumDensity_Err[]
54000, 9, 2005, 2, 3, 0, 42.308, -70.582, 6910, 6979, 242.5, 65.5
  9154, -999999, -999999, -999999, -999999, 113178, 212, -999999
  9304, -999999, -999999, -999999, -999999, 123353, 2250, -999999
  9454, -999999, -999999, -999999, -999999, 123008, 2116, -999999
  9604, -999999, -999999, -999999, -999999, 120933, 1337, -999999
  9754, -999999, -999999, -999999, -999999, 119675, 1019, -999999
  9904, -999999, -999999, -999999, -999999, 122655, 2061, -999999
  10054, -999999, -999999, -999999, -999999, 124384, 3126, -999999

10204, -999999, -999999, -999999, -999999, 124632, 3371, -999999
10354, -999999, -999999, -999999, -999999, 121341, 1609, -999999
54001, 8, 2005, 02, 03, 0, 42.278, -70.613, 6978, 7043, 241.7, 65.5
10118, -999999, -999999, -999999, -999999, 124458, 3205, -999999
10268, -999999, -999999, -999999, -999999, 123160, 2421, -999999
10418, -999999, -999999, -999999, -999999, 121221, 1582, -999999
10568, -999999, -999999, -999999, -999999, 120950, 1523, -999999
10718, -999999, -999999, -999999, -999999, 117339, 680, -999999
10868, -999999, -999999, -999999, -999999, 122751, 2423, -999999
11018, -999999, -999999, -999999, -999999, 124230, 3491, -999999
11168, -999999, -999999, -999999, -999999, 124039, 3424, -999999

{Note the use of scale factors in this example.}
Example 2310 using V02_2016 version

File name: ICARTT-LIDARO3_WP3_20040830_R0.ict

46, 2310, V02_2016
Williams, Eric
NOAA/Earth System Research Laboratory
Ozone number density profile from WP3 aircraft LIDAR
ICARTT_ITCT
1, 1
2004, 08, 30, 2009, 09, 04
1
Geo_Alt, meters, Geometric_altitude, Geometric_altitude_of_observation
UT_TIME, seconds, Time_Start, Elapsed_time_from_0_hours_on_day_given_by_date
1
1.0e9
-9999
O3_NumDensity[], molecules/cc, Ozone_NumDensity, Ozone_NumDensity_Array
9
1, 1, 1, 1, 1, 1, 1, 1, 1
-9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999
Num_Altitudes, #, Number_of_altitudes, number_of_altitudes_at_current_time_mark
Geo_Alt_Begin, meters, Geometric_altitude_Start,
geometric_altitude_at_which_data_begin

Alt_Increment, meters, Geometric_altitude_increment,
altitude_increment_between_observations
Geo_Alt_Aircraft, meters, Geometric_altitude_aircraft, geometric_altitude_of_aircraft
UT_hour, hours, Hour_UTC
UT_min, minutes, Min_UTC
UT_sec, seconds, Sec_UTC
Lon_aircraft, degrees_E, Longitude
Lat_aircraft, degrees_N, Latitude
0
18
PI_CONTACT_INFO: 325 Broadway, Boulder, CO 80305; 303-497-3226;
eric.j.williams@noaa.gov
PLATFORM: NOAA WP3
LOCATION: Lat, Lon, and Alt data included in the data records
ASSOCIATED_DATA: N/A
INSTRUMENT_INFO: Differential absorption LIDAR. See Williams et al., BigScience,
42, p. 50-51, 2001
DATA_INFO: The units are number density (#/cc). The vertical averaging interval is 975
m at 1-7 km above the aircraft and 2025 m > 7 km above the aircraft. Horizontal
averaging interval: 60 km.
UNCERTAINTY: Contact PI
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A
DM_CONTACT_INFO: Contact PI
PROJECT_INFO: ICARTT study; 1 July-15 August 2004
STIPULATIONS_ON_USE: Use of these data requires PRIOR OK from the PI
OTHER_COMMENTS: N/A
REVISION: R0
R0: No comments for this revision.
UT_TIME, Num_Altitudes, Geo_Alt_Begin, Alt_Increment, Geo_Alt_Aircraft,
UT_hour, UT_min, UT_sec, Lon_aircraft, Lat_aircraft, O3_NumDensity[]
30335, 26, 12819, 75, 10389, 8, 25, 35, -133.24, -9.45
1340, 1519, 1660, 1779, 1868, 1939, 1973, 1992, 1989, 1955, 1934, 1897, 1817, 1721,
1619, 1514, 1434, 1343, 1258, 1203, 1140, 1088, 1037, 956, 892, 878
30336, 22, 12819, 75, 10383, 8, 26, 0, -133.22, -9.93

1351, 1523, 1658, 1774, 1860,1930, 1962, 1974, 1966, 1932, 1909, 1877, 1803, 1706, 1600, 1493, 1407, 1310, -9999, -9999, 1094, 1045

{Note the use of scale factors in this example.}

2.5 File Formats for Non-Standard Data

Some non-standard data do not easily conform to the standard ICARTT time-series format. For these cases the datasets should either use the ICARTT format with a data interval code of -1 or utilize a different data format option that is a better fit for the dataset. This decision shall be made in consultation with the data manager to best meet the needs of the science team and to represent the data accurately.

2.5.1 Data Interval Code of -1

The data interval code of -1 has been used in past airborne studies for satellite, trajectory, and ground site data with large data gaps, as these data types are not conveniently incorporated into the ICARTT format. These datasets typically have the following properties: they are recorded on a nominally constant data interval, there may be significant gaps in the data, and the concept of start and stop time is not applicable (e.g., because integration times are so short). In these cases, the data files report a data interval code of -1 on line 8 of the file header. This signifies that each data record is identified by a single timestamp, but the actual timeline is discontinuous. Other uses of a data interval code of -1 for handling non-standard data reporting are discouraged and will be allowed strictly at the discretion of the data manager with prior consultation of the project scientist to ensure that research needs are met.

2.5.2 Other Data Format Options

The data management team shall highly encourage science team members to consider, on a case-by-case basis, using standards common to their data's user community. For example, some instruments provide data as image files, usually in standard formats such as GIF or JPEG. Also, any modeling data sets store data in NetCDF (Network Common Data Form) files, and lidar data is often stored in HDF (Hierarchical Data Format) files. For multidimensional data such as models and remote sensors, the ICARTT formatting described in Section 2.4 can accommodate these data sets, and therefore we leave it as an optional format. In any case, not all software for reading and writing formats other than

ICARTT allow additional text information (e.g., as a header), so the file names for these files must be defined to include as much information as possible. If necessary, the data management team should work with these PIs to achieve a solution to ensure that the data products can be adequately preserved, distributed, and used.

2.6 File Scanning Software

A software package called “FScan” has been developed for scanning data files and verifying if the files are in compliance with the ICARTT format standards. The scanning function does a thorough examination on the file to ensure compliance; the file is checked line-by-line, value-by-value, and, in some cases, letter-by-letter. A detailed report is generated displaying error messages, if any, along with line numbers and reasons. The FScan offers both online and standalone versions (see URL below). Further details on FScan is given at:
http://www-air.larc.nasa.gov/missions/etc/helpFscan.html.

There are two versions available to scan ICARTT formatted files:
1. Web-based: http://www-air.larc.nasa.gov/cgi-bin/fscan
2. Standalone Version (Windows only): http://www-air.larc.nasa.gov/missions/etc/wFscan.htm

Additionally, another online ICARTT format checker is maintained at ESRL/NOAA/CSD:
http://esrl.noaa.gov/csd/groups/csd7/measurements/icartt_file_check/
(access available upon request).

3. References

Informative References

[1] NASA Ames Format Specification for Data Exchange:
http://espoarchive.nasa.gov/archive/docs/formatspec_2_0.html

[2] NASA GTE Data Archive Format:
http://www-gte.larc.nasa.gov/trace/TP_APP-E.htm

[3] Official ICARTT Data Format Document:
https://earthdata.nasa.gov/standards/icartt-file-format

[4] International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) campaign:
http://www.esrl.noaa.gov/csd/ICARTT/

[5] ARCTAS Data Policy and Management Plan:
http://www-air.larc.nasa.gov/missions/arctas/docs/arctas_data_plan.pdf

[6] MILAGRO Data Policy and Management Plan:
http://www.eol.ucar.edu/projects/milagro/data/MILAGRO_DataPolicy.html

4. Author Information

ICARTT Refresh ESDSWG:

Emily Northup (Chair), emily.a.northup@nasa.gov, NASA LaRC
Kenneth Aikin (Co-Chair), kenneth.c.aikin@noaa.gov, NOAA
Christopher Webster (Co-Chair), cjw@ucar.edu, NCAR Aviation Facility
Gao Chen, (Co-Chair), gao.chen@nasa.gov, NASA LaRC

Ali Aknan, ali.a.aknan@nasa.gov
Aubrey Beach, aubrey.l.beach@nasa.gov
Andreas Beyersdorf, andreas.beyersdorf@csusb.edu
Christopher Cantrell, christopher.cantrell@colorado.edu
Joshua DiGangi, joshua.p.digangi@nasa.gov
Katja Drdla, katja.drdla@nasa.gov
Amanda Early, amanda.l.benson@nasa.gov
Keith Evans, evans@umbc.edu
Melanie Follette-Cook, melanie.cook@nasa.gov
Clare Flynn, claref@uci.edu
Mike Gangl, mike.gangl@gmail.com
Qing Liang, qing.liang@nasa.gov

Michael Shook, michael.shook@nasa.gov
Morgan Silverman, morgan.l.silverman@nasa.gov

The authors would like to acknowledge that this work was supported by ESDIS through the ESDSWG ICARTT Refresh Working Group. The authors would like to thank all of the ESDSWG members listed above, and provide a special thank you to Michael Shook, Amanda Early, and Ali Aknan for their considerable contribution and continuous support to the development of this document. This team was able to work together in seeking community needs and to compile a list of recommendations based on extensive community discussion regarding the ICARTT file format. These recommendations have been implemented into this ICARTT 2.0 document.

Appendix A - Glossary of Acronyms

Acronym	Description
ARCTAS	Arctic Research of the Composition of the Troposphere from Aircraft and Satellites
ESDIS	Earth Science Data Information Systems
ESDSWG	Earth Science Data System Working Groups
EU	European Union
GPS	Global Positioning System
GTE	Global Tropospheric Experiment
Hz	Hertz
KORUS-AQ	An International Cooperative Air Quality Field Study in Korea
ICARTT	International Consortium for Atmospheric Research on Transport and Transformation
INTEX-B	Intercontinental Chemical Transport Experiment – Phase B
INTEX-NA	Intercontinental Chemical Transport Experiment – North America
ITCT	Intercontinental Transport and Chemical Transformation
ITOP	Intercontinental Transport of Ozone and Precursors
LaRC	Langley Research Center
LIDAR	Light Detection and Ranging
MILAGRO	Megacity Initiative: Local and Global Research Observations
N/A	Not Applicable
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NEAQS	New England Air Quality Study
NOAA	National Oceanic and Atmospheric Administration
NSF	National Science Foundation
PI	Principal Investigator
POLARCAT	POLar study using Aircraft, Remote sensing, surface measurements and modelling of Climate, chemistry, Aerosols and Transport
QA/QC	Quality Assurance / Quality Control
UTC	Universal Time Coordinated

Appendix B - Amended FFI 2110

Amended FFI 2110

Ali Aknan; NASA LaRC

(November 2016)

<u>EXAMPLE – Start/Stop/Mid-point Intervals Sampling</u>

File Format Specification for ICARTT Multidimensional Data Files Using FFI 2110:

Structure

The descriptor for FFI 2110: two real independent variables, one unbounded and one bounded with its values recorded in the data records. Primary variables are real; Auxiliary variables are real.

The file structure (pseudocode) described by the Ames format document looks like this:

NLHEAD 2110
ONAME
ORG
SNAME
MNAME
IVOL NVOL
DATE RDATE
DX(1) DX(2)
[XNAME(s)] s=1,2
NV
[VSCAL(n), n=1,NV]
[VMISS(n), n=1,NV]
[VNAME(n)] n=1,NV
NAUXV ....The first Auxiliary variable is NX(m,1)
[ASCAL(a), a=1,NAUXV]
[AMISS(a), a=1,NAUXV]
[ANAME(a)] a=1,NAUXV
NSCOML
[SCOM(k)] k=1,NSCOML

NNCOML
[NCOM(k)] k=1,NNCOML
[X(m,2) NX(m,1) (A(m,a), a=2,NAUXV)]
[X(i,m,1) (V(i,m,n), n=1,NV)] i=1,NX(m,1)

Let's break it down line-by-line:

The first 7 lines are identical to FFI 1001 file type and described in Section 2.3.2.

DX(1): this is the sampling interval for the bounded independent variable (e.g., GeoAltitude). Enter "0" for its value if the sampling interval is non-constant.

DX(2): this is the sampling interval (see Section 2.3.2.8) for the unbounded independent variable (e.g., UTC); the same details/restrictions apply as they do to the sampling interval in FFI 1001 file type.

[XNAME(s)] s=1,2: here you list the names for the 2 independent variables starting with the bounded variable (e.g., GeoAltitude), then the unbounded variable (e.g., UTC).

NV: represents the total number of Primary variables. ALL Primary variables are arrays.

[VSCAL(n), n=1,NV] : analogous to FFI 1001 file type (Primary variables scale factors).
[VMISS(n), n=1,NV] : analogous to FFI 1001 file type (Primary variables missing data indicators).
[VNAME(n)] n=1,NV: analogous to FFI 1001 file type (Primary variables names).

NAUXV .... total number of Auxiliary variables.

The first Auxiliary variable is NX(m,1). Subsequent Auxiliary variables will include the Stop-time and Mid-point variables (i.e., will be the second and third Auxiliary variables) if they are being provided, then followed by the rest of the Auxiliary variables; where,
- NX(m,1) = Dimension; total number of values for Primary variables (i.e., array size)

[ASCAL(a), a=1,NAUXV]: analogous to FFI 1001 file type (Auxiliary variables scale factors).
[AMISS(a), a=1,NAUXV]: analogous to FFI 1001 file type (Auxiliary variables missing data indicators).
[ANAME(a)] a=1,NAUXV: analogous to FFI 1001 file type (Auxiliary variables names).

NSCOML .... Number of SPECIAL comment lines.
[SCOM(k)] k=1,NSCOML: analogous to FFI 1001 file type (SPECIAL comments text).

NNCOML .... Number of NORMAL comment lines.
[NCOM(k)] k=1,NNCOML: analogous to FFI 1001 file type (NORMAL comments text).

The Data Section:
[X(m,2) NX(m,1) (A(m,a), a=2,NAUXV)]
[X(i,m,1) (V(i,m,n), n=1,NV)] i=1,NX(m,1)

Each data record contains UTC, Auxiliary and Primary variables:

- Each data record always begins with UTC, the next value will be the Primary variable(s)' dimension, then the values for the rest of the Auxiliary variables, including the Stop/Mid-point Sampling Intervals, if provided. All of these values (UTC and all Auxiliary variables) are listed on one line.

The next "NX(m,1)" lines (i.e., 1 thru ArraySize) display the bounded independent variable (e.g., GeoAltitude) and all Primary variables' values in "a column format" starting with bounded independent variable:

- The first line in this section will display the first value of the bounded independent variable and the first value for each Primary variable.
- The second line will display the second value of the bounded independent variable and the second value for each Primary variable.

... and so forth until all "NX(m,1)" lines are listed. The array size must be the same for ALL Primary variables per each data record.

Note, in the **Header NORMAL Section**, please include all KEYWORD/VALUE pairs.

If LODs are to be used, then try to use one ULOD_FLAG and one LLOD_FLAG for all parameters. Also remember that LODs DO NOT apply to Independent variables.

For example:

ULOD_FLAG: -7777 *(i.e., ALL parameters will use –7777 for their ULOD_FLAG)*
ULOD_VALUE: ulod_Vi, ulod_Vi+1, ... ulod_Vn

[where,
ulod_Vi = ULOD_arraySize (normally this will be N/A)
ulod_Vi+1 = ULOD_NextAuxVar
...
ulod_Vn = ULOD_Last_PRIMARY_VAR
]

Use similar procedure for LLODs.

LODs VALUE should be N/A for Stop/Mid-point variables and for variables that do not use LODs.
Below is an example for clarification. Continuous lines have been indented for clarity.

EXAMPLE *[Start/Stop/Mid-point Intervals Sampling]*

File name: AROTAL-RAY_DC8_20040715_R1.ict

68, 2110, V02_2016
MAHONEY, MJ
M/S 169-237; Jet Propulsion Laboratory; Pasadena, CA 91109-8099
MTP - Microwave Temperature Profiler
SEAC4RS
1, 1
2013, 08, 01, 2013, 08, 18
0, 0
Palt[], meters, Pressure_Altitude, Pressure_Altitude_array
Start_UTC, seconds, Time_Start, elapsed time from 0000 UTC
4
1.0, 1.0, 1.0, 1E+21
-99999, -99999, -99999, -99999
Temperature[], K, Temperature, Temperature_array
Temperture_SE[], K, Temperature_Standard_Error, Temperature_Standard_Error_array
Geometric_altitude[], meters, Geometric_Altitude, Geometric_Altitude_array
Molecular_air_density[], number_per_cubic_meter, Molecular_air_density,
Molecular_air_density_array
17
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0
-9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -
9999, -9999, -9999, -9999, -9999

NZ, none, number_of_altitudes_in_data_record
Stop_UTC, seconds, Time_Stop, end of scan
Mid_UTC, seconds, Time_Mid, horizon location of scan
Zp, km, Pressure_altitude
Pitch, deg, Aircraft_Pitch
Roll, deg, Aircraft_Roll
OAT, K, Outside_air_temperature
Zt1, km, Tropopause_1_pressure_altitude
Zt2, km, Tropopause_2_pressure_altitude
PTt1, K, Potential_Temperature_tropopause_1
PTt2, K, Potential_Temperature_tropopause_2
Latitude, deg, Aircraft_Latitude
Longitude, deg, Aircraft_Longitude
LRac, K/km, Lapse_Rate_at_Flight_Level
MRI, none, Retrieval_quality_metric
Tcp, K, Cold_point_temperature
Zcp, km, Cold_point_pressure_altitude
1
*** Preliminary Data Preliminary Data Preliminary Data Preliminary Data Preliminary Data ***
28
Here's a brief free-form tutorial on how to decipher the MTP data:
Data groups consist of the following group of lines per 15-second observing cycle.
First line is: UTstart, number of retrieval levels, UTend, UTmid, Pressure Altitude, Pitch, Roll, Outside air temp (K), tropopause altitude #1 (km), tropopause altitude #2 (km) [if present], potential temperatures of tropopause #1 and #2, latitude, longitude, & lapse rate near flight level.
Remaining set of lines for each cycle consist of 5 columns: col#1 is pressure altitude (meters), col#2 is temperature from MTP (Kelvin), col#3 is temperature error estimate (K), col#4 is geometric altitude (meters), based on GPS altitude (meters), and col#5 is molecular air density [1E+21/m3].
PI_CONTACT_INFO: M/S 246-102; Jet Propulsion Laboratory; Pasadena, CA 91109-8099; Michael.J.Mahoney@jpl.nasa.gov
PLATFORM: NASA Global Hawk 872
LOCATION: see http://espoarchive.nasa.gov/archive/info
ASSOCIATED_DATA: See http://espoarchive.nasa.gov/archive/browse/attrex
INSTRUMENT_INFO: See http://mtp.jpl.nasa.gov/
DATA_INFO: See text above

UNCERTAINTY: Contact PI
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A
DM_CONTACT_INFO: M/S 169-237; Jet Propulsion Laboratory; Pasadena, CA 91109-8099; Michael.J.Mahoney@jpl.nasa.gov
PROJECT_INFO: SEAC4RS mission 01 August - 30 September 2013; DAOF and Ellington Field, Houston, Texas
STIPULATIONS_ON_USE: Use of these data should be done in consultation with the PI
OTHER_COMMENTS:
REVISION: R0
R0: Preliminary Data
Start_UTC, NZ, Stop_UTC, Mid_UTC, Zp, Pitch, Roll, OAT, Zt1, Zt2, PTt1, PTt2, Latitude, Longitude, LRac, MRI, Tcp, Zcp, Palt[], Temperature[], Temperture_SE[], Geometric_altitude[], Molecular_air_density[]
77381, 0, 77394, 77386, 6.358, 15.8, -0.2, 261.9, -9999, -9999, -9999, -9999, 34.779, -118.165, -9999, 0.25, 200.66, 15.36
77394, 0, 77407, 77399, 6.618, 11.9, -0.6, 259.7, -9999, -9999, -9999, -9999, 34.792, -118.171, -9999, 0.17, 202.05, 15.62
77407, 13, 77420, 77412, 6.790, 6.1, 0.2, 258.2, 15.2, 99.9, 373.2, 999.9, 34.805, -118.177, -6.7, 0.21, 201.61, 15.79
13690, 208.0, 1.4, 14295, 5150
10790, 229.1, 1.0, 11364, 7387
9290, 241.3, 0.8, 9775, 8827
8390, 248.4, 0.7, 8816, 9799
7790, 252.3, 0.5, 8176, 10526
7290, 255.6, 0.4, 7643, 11160
6940, 257.9, 0.2, 7270, 11623
6640, 259.7, 0.3, 6950, 12035
6290, 262.3, 0.5, 6577, 12512
5790, 265.6, 0.8, 6045, 13239
5211, 269.0, 1.0, 5429, 14142
4343, 274.6, 1.0, 4506, 15559
1834, 288.6, 9.1, 1839, 20460
77621, 14, 77634, 77626, 10.279, 6.1, -0.1, 230.2, 15.0, 99.9, 386.7, 999.9, 35.058, -118.258, -5.1, 0.32, 210.34, 15.48
17179, 211.0, 1.3, 17650, 2929

```
14279, 213.0, 1.4, 14814,  4582
12779, 218.4, 1.1, 13314,  5662
11879, 221.8, 0.8, 12397,  6425
11279, 225.0, 0.5, 11776,  6963
10779, 228.1, 0.3, 11252,  7433
10429, 229.9, 0.2, 10885,  7789
10129, 231.3, 0.2, 10571,  8109
 9779, 233.2, 0.3, 10206,  8485
 9279, 236.2, 0.4,  9686,  9034
 8679, 240.0, 0.6,  9062,  9720
 7779, 246.8, 0.8,  8125, 10777
 6279, 259.0, 0.9,  6554, 12691
 3779, 279.4, 1.2,  3910, 16472
```

Appendix C - Amended FFI 2310

Amended FFI 2310

Ali Aknan; NASA LaRC

(November 2016)

<u>EXAMPLE – Equal Intervals Sampling</u>

<u>EXAMPLE – Start/Stop/Mid-point Intervals Sampling</u>

File Format Specification for ICARTT Multidimensional Data Files Using FFI 2310:

Structure

The descriptor for FFI 2310: two real independent variables, one unbounded and one bounded with its number of values, base value, and constant increment defined in the Auxiliary variable list; primary variables are real; Auxiliary variables are real.

The file structure (pseudocode) described by the Ames format document looks like this:

NLHEAD 2310
ONAME
ORG
SNAME
MNAME
IVOL NVOL
DATE RDATE
DX(2)
[XNAME(s)] s=1,2
NV
[VSCAL(n), n=1,NV]
[VMISS(n), n=1,NV]
[VNAME(n)] n=1,NV
NAUXV .... The first 3 Auxiliary variable are NX(m,1),X(1,m,1),DX(m,1)
[ASCAL(a), a=1,NAUXV]
[AMISS(a), a=1,NAUXV]

[ANAME(a)] a=1,NAUXV
NSCOML
[SCOM(k)] k=1,NSCOML
NNCOML
[NCOM(k)] k=1,NNCOML
[X(m,2) NX(m,1) X(1,m,1) DX(m,1) (A(m,a), a=4, NAUXV)]
[V(i,m,n), i=1,NX(m,1)] n=1,NV

Let's break it down line-by-line:

The first 7 lines are identical to FFI 1001 file type and described in Section 2.3.2.

DX(2): this is the sampling interval (see Section 2.3.2.8) for the unbounded independent variable (e.g., UTC); the same details/restrictions apply as they do to the sampling interval in FFI 1001 file type.

[XNAME(s)] s=1,2: here you list the names for the 2 independent variables starting with the bounded variable (e.g., GeoAltitude), then the unbounded variable (e.g., UTC).

NV: represents the total number of Primary variables. ALL Primary variables are arrays.

[VSCAL(n), n=1,NV] : analogous to FFI 1001 file type (Primary variables scale factors).
[VMISS(n), n=1,NV] : analogous to FFI 1001 file type (Primary variables missing data indicators).
[VNAME(n)] n=1,NV: analogous to FFI 1001 file type (Primary variables names).

NAUXV .... total number of Auxiliary variables.

The first 3 Auxiliary variable are NX(m,1),X(1,m,1),DX(m,1). Subsequent Auxiliary variables will include the Stop-time and Mid-point variables (i.e., will be the fourth and fifth Auxiliary variables) if they are being provided, then followed by the rest of the Auxiliary variables; where,

- NX(m,1) = Dimension; total number of values for Primary variables (i.e., array size)
- X(1,m,1) = Base value of the bounded variable.
- DX(m,1) = Increment value of the bounded variable.

Therefore, we can now construct ALL bounded variable's values using:
X(i,m,1) = X(1,m,1) + (i-1) * DX(m,1) for i=1,NX(m,1). Simplified,
Value[i] = baseValue + (i * increment), where i=0 to arraySize-1

Note: if the bounded variable's increment is non-constant, then FFI 2110 should be used instead of FFI 2310.

[ASCAL(a), a=1,NAUXV]: analogous to FFI 1001 file type (Auxiliary variables scale factors).
[AMISS(a), a=1,NAUXV]: analogous to FFI 1001 file type (Auxiliary variables missing data indicators).
[ANAME(a)] a=1,NAUXV: analogous to FFI 1001 file type (Auxiliary variables names).

NSCOML .... Number of SPECIAL comment lines.
[SCOM(k)] k=1,NSCOML: analogous to FFI 1001 file type (SPECIAL comments text).

NNCOML .... Number of NORMAL comment lines.
[NCOM(k)] k=1,NNCOML: analogous to FFI 1001 file type (NORMAL comments text).

The Data Section:
[X(m,2) NX(m,1) X(1,m,1) DX(m,1) (A(m,a), a=4, NAUXV)]
[V(i,m,n), i=1,NX(m,1)] n=1,NV

Each data record contains UTC, Auxiliary and Primary variables:

- Each data record always begins with UTC, the next values are the Primary variable(s)' dimension, the bounded variable base value, and the bounded variable increment value, respectively. If Stop-time and Mid-point Sampling Intervals are provided then these 2 values are listed next, followed by the rest of the Auxiliary variables. All of these values (UTC and all Auxiliary variables) are listed on one line.

The next "NV" lines (i.e., 1 thru NV) display all Primary variables' values in "a row format":

- The first line in this section lists ALL values (in array) for the first Primary variable.
- The line after that lists ALL values for the second Primary variable (in array); and so forth until ALL "NV" Primary variables are listed. Please remember there is no limit on the number of char's per line; so **DO NOT** break arrays into multiple lines. The array size must be the same for ALL Primary variables per each data record.

Note, in the **Header Normal Section**, please include all KEYWORD/VALUE pairs.

If LODs are to be used, then try to use one ULOD_FLAG and one LLOD_FLAG for all parameters. Also remember that LODs DO NOT apply to Independent variables.

For example:

ULOD_FLAG: -7777 *(i.e., ALL parameters will use –7777 for their ULOD_FLAG)*
ULOD_VALUE: ulod_Vi, ulod_Vi+1, ... ulod_Vn

[where,
ulod_Vi = ULOD_arraySize (normally this will be N/A)
ulod_Vi+1 = ULOD_NextAuxVar
...
ulod_Vn = ULOD_Last_PRIMARY_VAR
]

Use similar procedure for LLODs.

LODs VALUE should be N/A for Stop/Mid-point variables and for variables that do not use LODs.
Below are 2 examples for clarification. Continuous lines have been indented for clarity.

EXAMPLE – Equal Intervals Sampling

File name: SOLVE-II-AD_DC8_20030202_R1.ict

53, 2310
Hostetler, Chris
NASA Langley Research Center
DC-8 GSFC/LaRC Lidar aerosol scattering ratios and depolarization
SOLVE II
1, 1
2003, 02, 02, 2004, 01, 16
1
Geo_Alt, meters, Geometric_altitude, Geometric altitude of observation (km)
UT_TIME, seconds, Time_Start, Time (UT seconds) from 00 hours on flight date
6
1.E-4, 1.E-3, 1.E-4, 1.E-9, 1.E-9, 1.E-4

-9999, -9999, -9999, -9999, -9999, -9999
TS_532[], #, Total Scattering Ratio, 532nm=((aerosol+molecular backscatter)/(molecular backscatter))
TS_1064[], #, Total Scattering Ratio, 1064nm=((aerosol+molecular backscatter)/(molecular backscatter))
AD_532[], #, Aerosol Depolarization Ratio, 532nm=((perpendicular aerosol backscatter)/(parallel aerosol backscatter))
ABCoef_532[], 1/km-sr, Aerosol Backscatter Coefficient at 532 nm (1/km-sr)
ABCoef_1064[], 1/km-sr, Aerosol Backscatter Coefficient at 1064 nm (1/km-sr)
Log10_mol_den[], #/cc, Log base 10 of molecular density used in analysis, #/cc
9
1.0, 1.E-3, 1.E-3, 1.0, 1.0, 0.01, 1.0, 0.01, 1.E-4
-9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999, -9999
Num_Altitudes, #, NumAltitudes, Number of altitude levels
Geo_Alt_Begin, km, GeoAltitude, Geometric altitude (km) at which data begins
Alt_Increment, km, AltIncrement, Altitude increment (km)
ProfileNumber, #, ProfileNumber, Profile number starting at 0
Lat_Deg, degrees_N, Latitude, Latitude Deg Aircraft
Lat_Min, degrees, Latitude, Latitude Min Aircraft
Lon_Deg, degrees_E, Longitude, Longitude Deg Aircraft
Lon_Min, degrees, Longitude, Longitude Min Aircraft
dpolM, #, Molecular depolarization Ratio, dpolM=(perpendicular aerosol backscatter)/(parallel aerosol backscatter)
1
Data products were interpolated onto a uniform vertical grid based on GPS geometric altitude. The vertical resolution is 75 meters.
19
PI_CONTACT_INFO: Address: MS 401a NASA/LaRC, Hampton, VA 23681;email:johnathan.w.hair@nasa.gov; 757-864-1406
PLATFORM: NASA DC8
LOCATION: Lat, Lon, and Alt included in the data records
ASSOCIATED_DATA: N/A
INSTRUMENT_INFO: Differential absorption lidar.
DATA_INFO: DIAL ozone mixing ratios in ppbv (online= 291.5nm & offline=299.6nm)
UNCERTAINTY: Provided as an array following the ozone array
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888

LLOD_VALUE: N/A
DM_CONTACT_INFO: c.f.butler@larc.nasa.gov
PROJECT_INFO: ARCTAS, March 20 - April 19, 2008, June 18 - July 13, 2008
STIPULATIONS_ON_USE: Use of these data should be done in consultation with PI
OTHER_COMMENTS: Only MidTime (midpoint of 3 minute average) is valid
REVISION: R1, R0
R1: Insertion of DC-8 in situ ozone at a/c altitude, and modeled ozone values inserted in data gaps.
R0: No comments
UT_TIME, Num_Altitudes, Geo_Alt_Begin, Alt_Increment, ProfileNumber, Lat_Deg, Lat_Min, Lon_Deg, Lon_Min, dpolM, TS_532[], TS_1064[], AD_532[], ABCoef_532[], ABCoef_1064[], Log10_mol_den[]
34997, 17, 11325, 075, 0, 69, 1440, 16, 4665, 155
10288, 10327, 10298, 10217, 10306, 10308, 10298, 12594, 13544, 13376, 13451, 12625, 13110, 12164, 12055, 13507, 12374
1089, 1104, 1128, 1143, 1130, 1118, 1142, 1396, 1359, 1056, 983, 1095, 1146, 1179, 991, 1310, 1251
133, 134, 135, 139, 139, 139, 143, 1951, 2242, 1802, 1861, 1906, 2044, 1508, 1713, 2323, 1787
10934, 12292, 11058, 7932, 11059, 11008, 10528, 4944, 6688, 6286, 6361, 4769, 5592, 3848, 3591, 6084, 4066
2039, 2374, 2863, 3172, 2844, 2544, 3048, 457, 411, 64, -18, 105, 158, 194, -10, 326, 259
187961, 187909, 187856, 187804, 187751, 187699, 187646, 174972, 174920, 174867, 174815, 174762, 174710, 174657, 174604, 174552, 174499
34998, 17, 11325, 075, 1, 69, 1950, 16, 3525, 155
10439, 10415, 10377, 10340, 10298, 10322, 10337, 11944, 12551, 12390, 12001, 12156, 12372, 12230, 11805, 11588, 11875
1336, 1329, 1342, 1350, 1347, 1378, 1372, 1551, 1284, 1314, 1549, 1540, 1458, 1581, 1622, 1599, 1882
137, 136, 137, 137, 138, 138, 139, 1863, 1652, 1750, 1645, 1907, 2036, 2003, 1588, 1222, 1374
16684, 15554, 13988, 12468, 10770, 11497, 11888, 3555, 3841, 2504, 2433, 3896, 2842, 3200, 3065, 2266, 3323
3676, 3914, 4256, 3954, 3161, 2747, 3203, 636, 324, 353, 611, 594, 497, 624, 660, 627, 915
187959, 187906, 187854, 187801, 187749, 187696, 187644, 174965, 174912, 174860, 174807, 174755, 174702, 174650, 174597, 174545, 174492

34999, 17, 11325, 075, 2, 72, 4301, -8, -3138, 155
10676, 10679, 10683, 10665, 10649, 10641, 10640, 10525, 10804, 10637, 10322, 10396, 10544, 10587, 10751, 10889, 10815
1394, 1394, 1399, 1398, 1401, 1403, 1407, 1220, 1295, 1511, 1360, 1065, 693, 1034, 958, 1177, 1184
-9999, -9999, -9999, -9999, 139, 140, 143, 156, 388, 138, 150, 144, 299, 360, 221, 246, 258
25436, 25234, 25058, 24125, 23233, 22684, 22362, 980, 1481, 1161, 580, 704, 956, 1020, 1289, 1508, 1367
8975, 8877, 8865, 8741, 8686, 8634, 8619, 249, 330, 564, 393, 70, -325, 34, -43, 182, 187
187918, 187864, 187811, 187758, 187704, 187651, 187598, 174874, 174822, 174770, 174718, 174666, 174615, 174563, 174511, 174459, 174408

EXAMPLE – Start/Stop/Mid-point Intervals Sampling

File name: SOLVE-II-AD_J31_20040129_r0.ict

61, 2310
PI LASTNAME, FIRSTNAME
NASA Langley Research Center
DC-8 GSFC/LaRC Lidar aerosol scattering ratios and depolarization
SOLVE II-File For Testing
1, 1
2004, 1, 29, 2004, 2, 16
0
GeoAlt, km, Geometric_altitude_of_observation
UTC, seconds, UT_time_from_00_hours_on_flight_date
6
0.0001, 0.001, 0.0001, 1.E-9, 1.E-9, 0.0001
-9999999, -9999999, -9999999, -9999999, -9999999, -9999999
TScatRatio532[], #, Total_Scattering_Ratio_{532nm=((aerosol+molecular-backscatter)/(molecular-backscatter))}
TScatRatio1064[], #, Total_Scattering_Ratio_{1064nm=((aerosol+molecular-backscatter)/(molecular-backscatter))}
AerDepRatio532nm[], #, Aerosol_Depolarization_Ratio_{532nm=((perpendicular-aerosol-backscatter)/(parallel-aerosol-backscatter))}
AerBkScatCoef532[], (1/km-sr), Aerosol_Backscatter_Coefficient_at_532nm
AerBkScatCoef1064[], (1/km-sr), Aerosol_Backscatter_Coefficient_at_1064nm

Log10_MolDensity[], #/cc, Log10_of_molecular_density_used_in_analysis
11
1.0, 1.0, 1.0, 0.001, 0.001, 1.0, 1.0, 0.01, 1.0, 0.01, 0.0001
-999, -999, -999, -99999, -999, -9999, -999, -9999, -999, -9999, -999999
NumAlt, number, Number_of_altitude_levels
GeoAltAC, km, Geometric_altitude_at_which_data_begins
AltIncre, km, Altitude_increment
StopUTC, seconds
MidUTC, seconds
ProfileNum, Number, Profile_numbe_starting_at_0
N_Lat, degrees, latitude_North
LatMin, minutes, Latitude
E_lon, degrees, Longitude_East
LonMin, minutes, Longitude
MolDepRatio, #, Molecular_depolarization_ratio_{dpolM=(perpendicular-aerosol-backscatter)/(parallel-aerosol-backscatter))}
8
Data products were interpolated onto a uniform vertical grid based on GPS
geometric altitude. The vertical resolution is 75 meters.
Molecular density profiles derived from the DAO data sets were used in
retrieval of the lidar data products.
Reanalysis of the data set for improvements in quality and resolution will
occur as funding permits. Researchers interested in using this data should
contact Chris Hostetler for the latest update (Chris.A.Hostetler@NASA.gov).

18
PI_CONTACT_INFO: Address: 325 Broadway, Boulder, CO 80305; email: Chris.A.Hostetler@NASA.gov
PLATFORM: NASA DC8
LOCATION: Lat, Lon, and Alt included in the data records
ASSOCIATED_DATA: N/A
INSTRUMENT_INFO: Molecular depolarization ratio
DATA_INFO: dpolM=(perpendicular aerosol backscatter)/(parallel aerosol backscatter)
UNCERTAINTY: Enter uncertainty description here
ULOD_FLAG: -7777
ULOD_VALUE: N/A
LLOD_FLAG: -8888
LLOD_VALUE: N/A

DM_CONTACT_INFO: N/A
PROJECT_INFO: Test File; April 15 2004
STIPULATIONS_ON_USE: Use of these data requires PRIOR OK from the PI
OTHER_COMMENTS: N/A
REVISION: R0
R0: No comments for this revision.
UTC, NumAlt, GeoAltAC, AltIncre, StopUTC, MidUTC, ProfileNum, N_Lat, LatMin,
 E_Lon, LonMin, MolDepRatio, TScatRatio532[], TScatRatio1064[],
 AerDepRatio532nm[], AerBkScatCoef532[], AerBkScatCoef1064[],
 Log10_MolDensity[]
32385, 10, 11325, 075, 32440, 32400, 0, 69, 229, 5, 1140, 156
 10871, 10868, 10879, 10902, 10907, 10876, 10854, 10860, 10843, 10839
 1166, 1174, 1182, 1183, 1179, 1175, 1175, 1178, 1186, 1189
 -9999999, -9999999, -9999999, -9999999, -9999999, -9999999, -9999999, -9999999,
159, 164
 33611, 33085, 33116, 33579, 33373, 31866, 30698, 30549, 29585, 29128
 3871, 4026, 4161, 4121, 3990, 3856, 3803, 3824, 3947, 3981
 188028, 187977, 187926, 187875, 187824, 187772, 187721, 187670, 187619, 187568
32442, 15, 11325, 075, 32505, 32475, 1, 69, 703, 4, 3777, 156
 10835, 10839, 10861, 10870, 10886, 10872, 10851, 10845, 10829, 10822, 10799,
10783, 10774, 10782, 10780
 1201, 1210, 1217, 1218, 1214, 1213, 1219, 1219, 1217, 1220, 1212, 1212, 1212, 1213,
1220
 -9999999, -9999999, -9999999, -9999999, -9999999, -9999999, -9999999, -9999999,
159, 163, 168, 173, 178, 181, 190
 32207, 31984, 32447, 32391, 32584, 31715, 30597, 30011, 29093, 28513, 27395,
26527, 25939, 25875, 25521
 4687, 4858, 4941, 4911, 4767, 4695, 4772, 4704, 4624, 4614, 4409, 4345, 4303, 4261,
4365
 188026, 187975, 187924, 187873, 187822, 187771, 187720, 187669, 187619, 187568,
187517, 187466, 187415, 187364, 187313
32510, 13, 11325, 075, 32560, 32525, 2, 69, 1027, 4, 1834, 156
 10835, 10839, 10861, 10870, 10886, 10872, 10851, 10845, 10829, 10822, 10799,
10783, 10774
 1201, 1210, 1217, 1218, 1214, 1213, 1219, 1219, 1217, 1220, 1212, 1212, 1212
 -9999999, -9999999, -9999999, -9999999, -9999999, -9999999, -9999999, -9999999,
159, 163, 168, 173, 178
 32207, 31984, 32447, 32391, 32584, 31715, 30597, 30011, 29093, 28513, 27395,

26527, 25939

4687, 4858, 4941, 4911, 4767, 4695, 4772, 4704, 4624, 4614, 4409, 4345, 4303

188026, 187975, 187924, 187873, 187822, 187771, 187720, 187669, 187619, 187568, 187517, 187466, 187415