

ICARTT v1.1

Offline ICARTT Merge Configuration Specification

23 February 2016

ICARTT Merge Configuration File

An ICARTT Merge Configuration file will be in an XML format, containing all of the information required to perform a series of ICARTT Merges. This document serves as an overview of the structure and format of a configuration file.

Scope

A single ICARTT Merge Configuration file will contain information to merge all of the data files from a single platform and a single mission. For example, an ICARTT Merge Configuration file generated for the `\TEXAS_2013\P3B_AIRCRAFT\` data directory will contain information to perform 12 separate ICARTT data merges. This is because the TEXAS 2013 mission used the P3B platform during 12 days.

The ICARTT Merge Configuration file will provide all information relevant to merging a series of ICARTT data files. As a result, the offline ICARTT Merge Tool will not need to read information from ICARTT file headers. Instead, variables will be referenced by their column number for each individual, unmerged ICARTT data file. The variables will also have a specified output column. If a variable does not exist during a particular mission day, then the output merged data for that mission, platform, day, variable will be reported as missing data.

Changelog

Updated XML structure and contents.

- Changed time-base specification.
 - Removed timebase specification from Generic Metadata
 - Added a Data Line attribute to every file reference
 - Added a time-base specification to each merge
 - Included the start and stop times for each merge
- Updates to the Variables section
- Changed language to reflect usage

Generic MetaData Section

The Generic MetaData section will contain the metadata that will be printed at the top of the resulting ICARTT files. This section will contain the following information.

1. Directory in which to store resulting merged data files

2. Name of the resulting merged data files
3. A list of lines containing the header that will be printed at the top of all of the

These fields should support replaceable character sequences. These sequences will start and end with the \$ character. In between these tags will be the key value of that character sequence in all capital letters.

Below is an example of this section. Please note the presence of character sequences that will be replaced during the generation of merged data files. These sequences have been bolded. Lines that end with ...</Line> have been shortened to fit the width of this document. A majority of the variables have been omitted to reduce the page count.

```
<GenericMetadata>

<FilePath>C:\ICARTT_Merge_Temp_Folder\TEXAS_2013\P3B_AIRCRAFT\MERGES\</FilePath>
<FileName>discoveraq-mrg60-p3b_merge_<b>$YYYY$</b><b>$MM$</b><b>$DD$</b>_R3.ict</FileName>

<GenericHeader>
  <Line>175, 1001</Line>
  <Line>Shook, Michael</Line>
  <Line>NASA Atmospheric Composition Branch, NASA Langley Research Center
  (...</Line>
  <Line>Merged data file for DISCOVERAQ-Texas_<b>$YYYY$</b>, Flight <b>$FLIGHTNUM$</b>
  ($...</Line>
  <Line>NASA DISCOVER-AQ Texas Mission 2015</Line>
  <Line>1, 1</Line>
  <Line><b>$YYYY$</b>, <b>$MM$</b>, <b>$DD$</b>, 2015, 07, 23</Line>
  <Line>60</Line>
  <Line>UTC,s</Line>
  <Line>128</Line>
  <Line>1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
  1...</Line>
  <Line>-999999, -999999, -999999, -999999, -999999, -999999, -999999,
  -999...</Line>
  <Line>JDAY, days</Line>
  <Line>INDEX, none</Line>
  <Line>FLIGHT, none</Line>
  <Line>LOCAL_SUN_TIME, hour</Line>
  <Line>LATITUDE, degs</Line>
  <Line>LONGITUDE, degs</Line>
  <Line>ALTP, km</Line>
  <Line>PRESSURE, hPa</Line>
  <Line>TEMPERATURE, K</Line>
  <Line>THETA, K</Line>

  ...

  <Line>MVK-MACR-crotonaldehyde_MixingRatio, ppbv</Line>
```

```

<Line>MEK-butanal_MixingRatio, ppbv</Line>
<Line>Benzene_MixingRatio, ppbv</Line>
<Line>Toluene_MixingRatio, ppbv</Line>
<Line>C8-alkylbenzenes_MixingRatio, ppbv</Line>
<Line>C9-alkylbenzenes_MixingRatio, ppbv</Line>
<Line>Monoterpenes_MixingRatio, ppbv</Line>
<Line>0</Line>
<Line>33</Line>
<Line>PI_CONTACT_INFO: michael.shook@nasa.gov,jennifer.r.olson@nasa.gov,
...</Line>
<Line>PLATFORM: NASA P3-B Aircraft</Line>
<Line>LOCATION: Latitude, Longitude, and Altitude included in data
records</Line>
<Line>ASSOCIATED_DATA: N/A</Line>
<Line>INSTRUMENT_INFO: N/A</Line>
<Line>DATA_INFO: Please see PI data file for explanation of ULOD and
LLOD.</Line>
<Line>UNCERTAINTY: See individual files for uncertainty information.</Line>
<Line>ULOD_FLAG: -777777</Line>
<Line>ULOD_VALUE: N/A</Line>
<Line>LLOD_FLAG: -888888</Line>
<Line>LLOD_VALUE: N/A</Line>
<Line>DM_CONTACT_INFO: gao.chen@nasa.gov</Line>
<Line>PROJECT_INFO: N/A</Line>
<Line>STIPULATIONS_ON_USE: N/A</Line>
<Line>OTHER_COMMENTS: This merge was created using data in the NASA
DISCO...</Line>
<Line>In most cases, variable names have been kept identical to those
sub...</Line>
<Line>However, in some cases, names have been changed (e.g., to
eliminate...</Line>
<Line>Units have been standardized throughout the merge.</Line>
<Line>See the mrg60 p3 readme.txt file for additional information.</Line>

<Line>-----...</Line>
<Line>ProfileNumber- This six digit number starts with a 1, followed by
t...</Line>
<Line>SiteSeqNumber- This six digit number starts with a 7, followed by
t...</Line>
<Line>MissedApproachNumber- This six digit number starts with three
digit...</Line>
<Line>SiteIDs- Galveston (01), Manvel Croix (02), Smith Point (03), Deer
...</Line>
<Line>AirportIDs- Galveston - Scholes International Airport (801),
Manvel...</Line>
<Line>Note that unless the corresponding raw data flags are set for the
e...</Line>

<Line>-----...</Line>
<Line>REVISION: R3</Line>

```

```

    <Line>R3: Updated to include data submitted as of 07/23/2015.</Line>
    <Line>R2: Updated to include data submitted as of 10/27/2014.</Line>
    <Line>R1: Updated to include data submitted as of 09/30/2014.</Line>
    <Line>R0: Updated to include data submitted as of 04/21/2014.</Line>
    <Line>UTC, JDAY, INDEX, FLIGHT, LOCAL_SUN_TIME, LATITUDE, LONGITUDE,
ALTP...</Line>
  </GenericHeader>
</GenericMetadata>

```

Variables Section

This section will contain the specification of all variables that will be included in each merged data file. Every variable in this section will have

1. a unique output column number, to specify destination in the output file,
2. an output name, a unique identifier used to reference variables,
3. the variable's type listed in capital letters, and
4. an attributes list, specified as key-value pairs.

The required contents of the attributes list will be determined by the variable's type. Below are some of the known variable types.

Scalar Variables are averaged normally. Scalar variables require no additional attributes.

Vector Magnitude Variables are averaged in tandem with a direction variable. Its attributes list will require the name of a Vector Direction Variable.

Vector Direction Variables are averaged in tandem with a magnitude variable. Its attributes list will require the name of a Vector Magnitude Variable.

Ratio Variables are averaged with both the AAAR weighting and a weight based on some combination of variables. Its attributes list will require the names of a certain number of other variables, depending on the ratio function used.

MAX Variables are variables whose maximum value over an interval is used instead of the AAAR weighting. No additional attributes will be required.

MIN Variables are variables whose minimum value over an interval is used instead of the AAAR weighting. No additional attributes will be required.

Potential Temperature is a variable that is calculated based on two or more other columns. (https://en.wikipedia.org/wiki/Potential_temperature) Its attribute list may require additional relational information.

Ozone Column Data is a variable that is pulled from a table based on the platform's latitude and longitude measurements. Its attributes section will require the names of these variables.

```

<Variables>
  <Variable>
    <Name>Scalar_Variable_1</Name>

```

```

    <Column>4</Column>
    <Type>SCALAR</Type>
    <Attributes></Attributes>
</Variable>
<Variable>
    <Name>VectMagVar_1</Name>
    <Column>7</Column>
    <Type>VECTOR_MAGNITUDE</Type>
    <Attributes>
        <Attribute>
            <Key>VectorDirectionVariable</Key><Value>VectDirVar_1</Value>
        </Attribute>
    </Attributes>
</Variable>
<Variable>
    <Name>VectDirVar_1</Name>
    <Column>8</Column>
    <Type>VECTOR_DIRECTION</Type>
    <Attributes>
        <Attribute>
            <Key>VectorMagnitudeVariable</Key><Value>VectMagVar_1</Value>
        </Attribute>
    </Attributes>
</Variable>
</Variables>

```

Merges Section

This section will contain as many merges as there are distinct dates in the unmerged data. In the case of */TEXAS_2013/P3B_AIRCRAFT/* the Merges section will contain 12 Merge subsections. Each merge subsection will contain a section for . Each merge subsection will also contain a Replaceable Header Text section and an Input Data section.

Replaceable Header Text Section

This section contains a list of maps of key, value pairs. For example, the Replaceable Header Text section for the 20130926 data from the TEXAS 2013 P3B directory will contain the following.

```

<ReplaceableHeaderText>
    <Map>
        <Key>YYYY</Key>
        <Value>2013</Value>
    </Map>
    <Map>
        <Key>MM</Key>

```

```

    <Value>09</Value>
  </Map>
  <Map>
    <Key>DD</Key>
    <Value>26</Value>
  </Map>
  <Map>
    <Key>FLIGHTNUM</Key>
    <Value>09</Value>
  </Map>
</ReplaceableHeaderText>

```

Merge Base Section

This section contains the information that allows all files in a merge to be related to a single time-base. This section contains the following fields. An example for `\TEXAS_2013\P3B_AIRCRAFT\` on has been included below.

- The start time of the merge
- The end time of the merge
- The interval of the time base
- Information relating to a relational data ID, such as a NAV File, including
 - the Data ID of the relational file,
 - the filename of the relational file, and
 - the filepath of the relational file.

```

<MergeBase>
  <StartTime>82500</StartTime>
  <StopTime>90000</StopTime>
  <Interval>60</Interval>
  <RelationalDataID>
    <DataID>DISCOVERAQ-REVEAL</DataID>

  <FilePath>C:\Users\...\TEXAS_2013\P3B_AIRCRAFT\VANGILST.DAVID</FilePath>
    <FileName>discoveraq-reveal_p3b_20130906_r1.ict</FileName>
  </RelationalDataID>
</MergeBase>

```

Files Section

This section will contain information related to the ICARTT files that will be included in the merge. The immediate children of the Files section will all be named File. Each file will have the following information.

1. Absolute file path to the ICARTT file from the root directory

2. File name of the ICARTT file
3. All 4 possible time-base variables found in the file (-1 if not available)
 - a. Data Interval
 - b. Start Time Column
 - c. Midpoint Time Column
 - d. End Time Column
4. A line number corresponding to the first line of data
5. A list of the variables found in this ICARTT file

The list of variables will only contain children named Variable. Each file-specific variable will contain the following information.

1. the variable's output name, used to reference an entry from the previous Variables section
2. the column number that contains the variable in the unmerged ICARTT file
3. the variable's scale factor
4. the variable's missing data indicator
5. the variable's ULOD flag
6. the variable's LLOD flag

High Level View

Below is an overview of the ICARTT Merge Configuration file schema. This is a composition of all the previous sections. ++ has been placed after tags that can exist 0 or more times under their parent section. A group of four dashes indicates an indent level and shows the nesting of different sections.

```
<ICARTTMergeConfiguration>
---- <GenericMetadata>
---- ---- <FilePath>
---- ---- <FileName>
---- ---- <GenericHeader>
---- ---- ---- <Line> ++
---- <Variables>
---- ---- <Variable>
---- ---- ---- <Name>
---- ---- ---- <Column>
---- ---- ---- <Attributes>
---- ---- ---- ---- <Attribute> ++
---- ---- ---- ---- <Key>
---- ---- ---- ---- <Value>
---- <Merges>
---- ---- <Merge> ++
```

```

---- <MergeBase>
---- <StartTime>
---- <StopTime>
---- <Interval>
---- <RelationalDataID>
---- <DataID>
---- <FilePath>
---- <FileName>
---- <ReplaceableHeaderText>
---- <Map> ++
---- <Key>
---- <Value>
---- <Files>
---- <File> ++
---- <FilePath>
---- <FileName>
---- <DataLine>
---- <TimeVariables>
---- <DataInterval>
---- <StartTimeColumn>
---- <MidPointTimeColumn>
---- <StopTimeColumn>
---- <Variables>
---- <Variable> ++
---- <Name>
---- <Column>
---- <Scale Factor>
---- <Missing Data Flag>
---- <ULOD Flag>
---- <LLOD Flag>

```