GTE - Data Archive Format Document

GTE Project Office
Mail Stop 483
National Aeronautics and Space Administration
Hampton, VA 23681-0001

September, 1994

Abstract:

This document describes the GTE Data Archive Format, the native format of all Global Tropospheric Experiment datasets. The information contained in this document is necessary to correctly decode the GTE Archive Datasets.

Table of Contents:

Figures and Tables:

I. General Instructions and Rationale

The data format described herein is designed to provide a file structure that effectively utilizes magnetic storage media for the exchange of data for scientific analysis. Recognizing that the variety of GTE measurement techniques span a large dynamic data range, every attempt has been made to meet the needs of scientists that generate either large or small data sets. The GTE Data Archive Format is intended to satisfy both the data archive and data exchange requirements of the GTE field expeditions.

I.A. Suitable Media:

Files must be in ASCII. All data should be numeric, separated by commas. Each aircraft PI shall submit a minimum of one file per flight. Whenever possible, data shall be submitted to the GTE Data Archive via electronic transfer. Suitable media include File Transfer Protocol (FTP), electronic mail (E-mail), and IBM-compatible diskettes (3.5" High Density preferable). The file size should be compatible with the submitting media limitations. Therefore, if the ASCII file size is greater than 1.44 mb, it should either be compressed before being submitted on 3.5" disks, or it should be submitted via FTP or E-mail. If data are compressed, either the un-compressing programs must be included with the data, or the data should be in executable form (.EXE file extension). Large files can be most easily accommodated via electronic submissions directly into the GTE Archives (Refer to Section V. FTP Protocol for GTE Data Exchange). As mediums of data exchange are constantly being upgraded and amended, and as new mediums of exchange introduced, the GTE Data Archive Office will continue to upgrade and change its data exchange processes when feasible. Any medium of exchange differing from those mentioned herein should be approved by the GTE Data Archive Office prior to data submittal.

I.B. Common File Structure:

A GTE Archive file is composed of two sections: a HEADER SECTION, which contains all the information with which to process the file; and the DATA SECTION, which contains the data in the prescribed GTE Data Archive format. The file naming convention, header and data format descriptions are detailed in Section II. of this document.

In order to facilitate the exchange of data among the GTE investigators and the scientific community at large, the GTE requires that all datasets be referenced by time. Time data is composed of the Julian day of the year and the seconds of the day, both referenced to Greenwich Mean Time (GMT). When a reporting period extends past GMT midnight (86,400 seconds), increment the Julian day of year and reset the time to zero. See Section II.C. Dataset Types for an explanation of the different time reporting options.

Aircraft PI's should submit a file for each aircraft mission. If no data are submitted to the archive for a particular mission, the PI shall in any case submit a file containing the Header section of the GTE Data Archive format. The PI shall use comment lines to explain the reason no data are provided for the mission.

II. Header Record Format

This section contains a detailed description of the format for each header record. The header records should supply all information needed to read and interpret the data records. Even if the actual data are not included in the file, the header records should include the parameters as noted and information needed to identify and locate the data files.

II.A. Definition of Header Records:

Each line of the header section is reserved for specific information. These headers will be scanned by computer programs; therefore, all data must be in the format as described in order to be correctly interpreted. The following sections define each line of the Header Record Format.

Line 1. Number of Lines in the Header (NH)

The first line in the header contains the number of lines in the header section (NH). The first line of data = NH + 1.

Line 2. File Name

The second line in the header contains the name of the dataset file. The data file naming convention for the GTE is as follows:

PPTTTLXX.EXP where:

PP = 2-character code identifying PI. See Table 1. for appropriate code.

TTT	=	3-character code identifying data set or instrument technique.
L	=	Location code: 'D' for DC-8 aircraft, 'E' for Electra aircraft. Codes for other locations will be provided by the GTE Project Office.
XX	=	2-digit numeric mission descriptor.For a flight, XX = flight number; for ground-based data, this may be a sample number or day identifier.
EXP	=	Expedition Code. See Table 2. for appropriate code.

NOTE: For aircraft investigators, each landing constitutes a new flight. Flights are named by integer numbers only (i.e. there can not be a flight named 21A or 21.1). If there are several flights on a day, each flight will receive a separate numeric identifier.

Examples follow:

DU_SUD01.PWB	=	Drexel University sulfur data for DC-8 Flight 1, for the PEM West-B Expedition (Replace any unused code letters with an underline "_")
POM10D05. TRA	=	Project Navigational/Meteorological 10-second data for DC-8 Flight 5, for the TRACE-A Expedition

Line 3. Principal Investigator

Line 3 contains the Experimenter Last Name, First Name, and Institution. Always separate items by commas.

Line 4. Species Measured / Technique Used

Line 4 contains a brief summary of the species measured and/or the method employed for this sampling. For example, non methane hydrocarbons should be entered here rather than listing each species. Use only one line.

Line 5. Expedition Name

The Official name for this expedition. Please refer to Table 2. for Expedition Names and their abbreviations.

Line 6. Data Start Date, Date Last Revised (YY,MM,DD,YY,MM,DD)

Line 6 contains two dates. Report the start date for this dataset, followed by the date of last revision, both in the format YY, MM,DD

Line 7. Number for this Flight, or Sequential Dataset Number (NF)

Line 7 contains a unique number for this dataset (NF). For aircraft datasets, this number is the number for this flight. For ground datasets, this is usually a sequential number starting at "01" for the first dataset submitted.

Line 8. Number of Variables Reported in this Dataset (NV)

For most standard datasets, the number of variables remains constant for each line in the dataset (Dataset Types 1, 2,4 and 6). In those cases, NV = number of variables for each record of data submitted. Several dataset types are defined which include a matrix of data whose dimensions may vary between dataset records. For these datasets, the matrix variable is considered as a single variable and is always the last variable reported in the list. The variable or variables immediately preceding the dimensioned variable will contain the number of elements in the matrix. In these cases, NV = number of unique variables for each record in the dataset, with the matrix being considered as a single variable. Refer to Figure 3 and Figure 5 for a complete explanation and examples of this variable.

Line 9. Number of Comments (NC)

Line 9 contains the number of comment lines included in the header for this dataset. Comment lines should include any pertinent information needed to help the scientific community correctly interpret the dataset.

Line 10. Dataset Type (DT)

The GTE Data Archive Format currently supports six (6) Standard Dataset Types, numbered from 1 to 6, with 0 = non-standard dataset type. The six Standard Dataset Types are:

Data Type = 1: Standard reporting of time as the midpoint of a standard averaging period.

Data Type = 2: Standard reporting of time as start, stop, and midpoint of a non-standard averaging period.

Data Type = 3: Vertical Column Sample. Non-standard sampling using a variable number of sampling points in the sample.

Data Type = 4: Vertical Column Sample. Standard submission for sonde data.

Data Type = 5: Gridded Dataset. Submission standards for mapping products such as the AVHRR Fire Counts maps.

Data Type = 6: Trajectory Dataset. Trajectory data follows a particular air parcel over a finite period of time and records values at regular intervals over the period.

Refer to Section II.C. Dataset Types for a complete description and examples of the six dataset types.

Line 11. Data Averaging Period

Report the Data Averaging Period, in SECONDS, in the 11th header record. For a 10-second averaging period, report data at 5-second time ticks; i.e., 5, 15, 25,... GMT seconds. Thus, the data value reported at 15 seconds represents the average value in the interval equal to or greater than 10 seconds and less than 20 seconds.

For 60-second averaged data, the reporting times should be at 30-second time ticks; i.e., 30, 90, 150,... GMT seconds. Thus, the data value reported at 90 seconds represents the average value in the interval equal to or greater than 60 seconds and less than 120 seconds.

For other constant time intervals, time ticks of the data reporting interval should be consistent with the 10-sec or 60-sec format. For a non-standard averaging period, enter "0" for this line.

Line 12. Data Sampling Frequency (Hertz)

Enter the instrument sampling frequency, in hertz. Data recorded at 1 sample/second = 1 Hz; 5 samples/second = 5 Hz; 1 sample every 5 seconds = .2 Hz. ; 1 sample / minute = 1/60 = .0167 Hz. Enter "0" here for instruments having non-standard sampling periods. For Data Type =6, replace this variable with the number of Samples per Trajectory.

Line 13. Definition of First Variable in the Dataset: Var(1)

This variable must be the Julian Day of the Year (GMT) that the sample was taken. (Refer to Section II.C. for further explanation.) A minimum of eight parameters are necessary for each Variable Definition Line. They are:

- Variable Name
- Units
- Scale
- Offset
- Minimum
- Maximum

- Null Value
- LOD Code

If Limits of Detection (LOD's) are used in the dataset, then an additional four parameters are added to the end of the above definition line:

- Lower LOD Code
- Lower LOD
- Upper LOD Code
- Upper LOD

Refer to Section II.B. Data Variable Definition for a complete explanation of each element of the variable definition.

Line 14. Definition of Second Variable in the Dataset: Var(2)

The second variable must be a time in seconds, GMT. For Data Type 1, this is the median time of the sampling period. For all other Data Types, this is the Start Time of the sample. (The start time will be null (-999) in the case of Data Type 5). Refer to Section II.B. for further explanation.

Line 15. Definition of Third Variable in the Dataset: Var(3)

Refer to Figure 1., Figure 2., Figure 3., Figure 4., Figure 5., and Figure 6. for any standard definitions for this record.

Line 13+NV-1: Definition of the last variable in the dataset, Var(NV)

Line 13+NV: First line of Comments

This line contains the first line of comments for this header. If NC = 0, then there are no comment lines in the dataset.

Line 13+NV+NC-1: Last line of Comments

This is the last line of comments for this header.

Line 13+NV+NC: First line of Data

This is the first line of data for this dataset. This location is the same as Number of Lines in the Header (NH)+1 (Header line 1 defines NH)

II.B. Data Variable Definition:

Each variable is defined on a separate line in the header beginning with Header Record 13. Each variable definition is composed of a minimum of eight (8) parameters, separated by commas. The eight required parameters are:

- Variable Name
- Units
- Scale Offset
- Minimum
- Maximum
- Null Value
- LOD Code

If Limits of Detection (LOD's) are used in the dataset, then an additional four parameters are added to the end of the above definition line:

- Lower LOD Code
- Lower LOD

- Upper LOD Code
- Upper LOD

Each item in the Variable Definition line is explained below.

Variable Definition in the Header:

Item 1. Name of Data Variable. (Time, Day, Parameter, Chemical Species, etc.)

Item 2. Units of Data Variable.

Item 3. Scale Factor of Data Variable.

Item 4. Offset of Data Variable.

NOTE: Scale factors and offsets should be submitted with all data, such that:

Reported DataValue * ScaleFactor + Offset = Engineering Units

Scale Factors and Offsets must be applied to ALL data submitted to the archives, including data summary values reported in the variable definition section of the header. This includes Minima, Maxima, and Limits of Detection for the data. It is preferable that data be reported in engineering units such that Scale Factors = 1 and Offsets = 0.

Item 5. Minimum value for this variable in this dataset.

The minimum value that is found for this variable for **THIS** dataset, not an absolute minimum value that could be experienced for this variable. For TIME, use the first sample time reported in the dataset. Minimum value is reported using the Scale Factor and Offsets given for the variable, such that:

ReportedMinimumValue * ScaleFactor + Offset = Minimum Value (Engineering Units)

Item 6. Maximum value for this variable in this dataset.

The maximum value for a variable that is found in **THIS** dataset. For TIME, use the last sample time reported in the dataset. Maximum value is reported using the Scale Factor and Offsets given for the variable, such that:

ReportedMaximumValue * ScaleFactor + Offset = Maximum Value (Engineering Units)

Item 7. Code for missing or bad data.

This code is composed of a series of negative nines which will be outside the range of actual data for this variable. No scale factoring or offsets are applied.to this code. The code reported here should be to the same number of significant digits as the actual Reported Data Value.

Example 1. Scale Factor = 1, Offset = 0, Reported Maximum Value = 59.9, Reported Minimum Value = -125.4: Missing Data Code = -999.9.

Example 2. Scale Factor = 10, Offset = 0, Reported Maximum Value = 5.99, Reported Minimum Value = -12.54: Missing Data Code = -999.9 (NOTE: Actual Data Maximum = 5.99*10 + 0 = 59.9: Actual Data Minimum = -12.54 * 10 + 0 = -125.4).

Item 8. Limit of Detection code (LOD Code):

Acceptable values are 0, 1, or 2.

0: There are no Limits of Detection given for this variable.

1: The Limits of Detection for this variable (Items 10 and 12) are pointers to a variable in the dataset which contains the actual LOD. See Items 9-12 for further explanation of this code.

2: The Limits of Detection for this Variable (Items 10 and 12) contain the actual LOD's for this dataset.

Items 9-12 are used only when Limits of Detection are given (LODCode > 0).

Item 9. Lower Limit of Detection Code (LLOD Code):

This code is composed of a series of negative eights (-8's) which will be outside the range of actual data for this variable. No scale factoring or offsets are applied to this code. The code reported here should be in the same format as the actual Reported Data Value.

Item 10. Lower Limit of Detection (LLOD) (Value or Variable column number)

If LOD Code = 1, LLOD = variable column number containing the actual LOD for this sample. If LOD Code = 2, LLOD = value of the LLOD for this dataset.

LLOD value is reported using the Scale Factor and Offsets given for the variable, such that :

Reported LLOD Value * ScaleFactor + Offset = LLOD Value (Engineering Units)

Item 11. Upper Limit of Detection Code (ULOD Code):

This code is composed of a series of negative sevens (-7's) which will be outside the range of actual data for this variable. No scale factoring or offsets are applied to this code. The code reported here should be to the same number of significant digits as the actual Reported Data Value.

Item 12. Upper Limit of Detection (ULOD) (Value or Variable column number).

If LOD Code = 1, ULOD = variable column number containing the actual LOD for this sample.

If LOD Code = 2, ULOD = value of the ULOD for this dataset.

ULOD value is reported using the Scale Factor and Offsets given for the variable, such that :

Reported ULOD Value * ScaleFactor + Offset = ULOD Value (Engineering Units)

Example of LOD Codes:

LOD Code = 0

No LOD's for this variable.

LOD Code = 1

This code assumes the Limits of Detection for this flight were changing while in flight. Assume this is Var(7) definition. (Based on 1 to NV variables). Data reported has one significant digit, with a Min = 2.1 and a Max = 155.8.

LLOD Code = -8.8 (beyond the normal data range using -8's.)

ULOD Code = -7.7

LLOD = 8. Var(8) contains the LLOD when the LLOD Code = -8.8.

ULOD = 8. Var(8) contains the ULOD when the ULOD Code = -7.7.

Thus, if Var(7) contains -8.8, instrument has detected data below its lower limits of detection. Refer to Var(8) for that limit for this time period. If Var(7) contains -7.7, observed data are above the upper LOD of the instrument. Refer to Var(8) for the Upper LOD for this time period.

NOTE: When Var(7) contains real data, Var(8) will contain the Code for Missing Data (-9.9)

LOD Code = 2

This code is used for defining constant LOD's for the entire dataset. Refer to the above section **(LOD Code = 1)** for general information on the Limits of Detection. In this case, when Var(7) contains LLOD Code (-8.8), the actual Lower LOD for this dataset is the LLOD from the Data Definition line of the header. Likewise, if Var(7) contains ULOD Code (-7.7), ULOD contains the actual Upper LOD.

Like all other data values reported to the archive, scale factors and offsets must be applied to the reported value to arrive at the true data value, in engineering units.

II.C. Dataset Types (DT):

Sample Description 1. (DT = 1): Standard Dataset, Constant time increment, Time reported as the midpoint of sampling time. (See Constant Interval Time Reporting for a discussion of midpoint reporting.) Most instruments using a constant time interval between samples can utilize this dataset type. The first two variables must be:

Var(1): Julian Day of the Year (GMT)

Var(2): Time (Seconds GMT) - Midpoint of sample. See Constant Interval Time Format section.

Refer to Figure 1 for an example of this dataset type.

Sample Description 2. (DT=2): Standard Grab Sample, Times are non-standard. This dataset type should be used by experimenters who take samples at non-constant sampling times and periods. The first four variables must be:

Var(1): Julian Day of the Year (GMT) of the Sample Midpoint

Var(2): Sample Start Time (Seconds of the Day, GMT)

Var(3): Sample Stop Time (Seconds of the Day, GMT)

Var(4): Sample Midpoint (Seconds of the Day, GMT)

Refer to Figure 2 for an example of this dataset type.

Sample Description 3. (DT = 3): Vertical Profile Sample, with a variable number of points in the sample. Variable "nv" defines the number of variables, with the last variable being an array dimensioned Var(nv). For example, if nv = 8, then variable 8 contains the number to dimension the profile array variable. The first two variables must be:

Var(1): Julian Day of the Year (GMT) of the Sample Start Time

Var(2): Sample Start Time (Seconds of the Day, GMT)

The final two variables must be:

Var(nv-1): Number of profile data values for this sample

Var(nv): Profile array, dimensioned Var(nv-1)

Refer to Figure 3 for an example of this dataset type.

Sample Description 4. (DT = 4): Vertical Profile Sample, Standard Sonde Data. This dataset type follows the example of DT 1, except that time is not constant. The first two variables must be:

Var(1): Julian Day of the Year (GMT) of this altitude

Var(2): Time of this altitude (Seconds of the Day, GMT)

Refer to Figure 4 for an example of this dataset type.

Sample Description 5. (DT = 5): Standard Gridded Data, Gridded Map Data. This dataset type is constructed to enable investigators to submit satellite data to the Archive, for example, AVHRR Fire Count data maps. It is assumed that these datasets are geographically rectified to a given latitude and longitude at a given time. This data type requires the following variables:

Var(1): Start Julian Day of the Year (GMT)

Var(2): Start Time of the Scan (Seconds, GMT)

Var(3): Stop Julian Day of the Year (GMT)

Var(4): Stop Time of the Scan (Seconds, GMT)

Var(5): Start Latitude of the scan, Degrees North = positive

Var(6): Start Longitude of the scan, Degrees East = positive

Var(7): Latitude Increment , Degrees North . Add the Latitude Increment to the Start Latitude to find the second row of the scan. If the Latitude Increment is positive, scan is going from south to north; negative is north to south.

Var(8): Longitude Increment, Degrees East. Add the Longitude Increment to the Start Longitude to find the second column of the scan. If the Longitude Increment is positive, the scan is going from west to east; negative is east to west.

Var(9): Number of Rows of Latitude (NRows).

Var(10): Number of Columns of Longitude (NCols).

Var(11): Pixel Data (NRows, NCols). The data are gridded by Latitude (NRows) and Longitude (NCols) and is presented as scan lines of longitude, one data record for each Latitude Increment.

A simple decode program flow for the data scans follows:

```
                         DIM PixelData(NRows,NCols)
                         FOR i = 1 to NRows
                                                 ' read in a line of data for a
constant Latitude
                                                 For j = 1 to NCols
                                                         INPUT filenumber, PixelData
(i , j)
                                                 NEXT j
                         NEXT i
```

The dimensioned variable PixelData will now contain the entire matrix of data for this image. Geographically rectify the image using the Start Latitude and Start Longitude and incrementing each correctly. Refer to Figure 5 for an example of this type of dataset.

Sample Description 6. (DT = 6) : Standard Trajectory Dataset. Trajectory data report samples from a particular air parcel over a finite period of time. Each trajectory in the dataset must cotain the same number of data points, defined by Header Line 12. The first data line contains data from the trajectory end point, corresponding to a selected aircraft time and position. Each following line is the trajectory's position on the prvious time step. Time steps prior to the start time are all normalized according to the Data Averaging Period defined in Header Line 11. Refer to Figure 6. for an example of this type of dataset.

III. Data Record Format

The Data Record section begins on line NH+1 (one line after the header). For each record, data should be reported in the sequence as stated in the Header Record. Each reported value in the dataset utilizes the scale factor and offset for that variable so that

Reported DataValue * ScaleFactor + Offset = Engineering Units

Only numeric values are acceptable as data for the standard archive data types. Please contact the GTE Data Management Office if your data cannot meet this criteria. It is preferable that data be reported in engineering units such that Scale Factors = 1 and Offsets = 0. If data is missing, or is above or below the Limits of Detection (LOD), use the appropriate code as stated in the variable definition for that variable in the header. These special codes should be entered exactly as stated in the Variable Definition Section so that computer programs can correctly determine validity. For example, if the null value is defined as -9.99 for a variable in the Variable Definition Section, that value should be used in the data for the null value. Scale factors and offsets are **NOT** applied to any codes in the data.

Note that uncertainties associated with a given measurement may be included in the archive as a variable.

III.A. Time Reference:

All data should be referenced to day of the year (Julian Day) and Greenwich Mean Time (GMT seconds of the day). When a reporting period extends past GMT midnight (86,400 seconds), increment the Julian day of year and reset the time to zero. Please refer to Section II.C. Dataset Types for a complete list of required variables for each dataset.

III.A.1 Constant Interval Time Averaging Reporting:

Report the Data Averaging Period in the 11th header record. For a 10-second reporting interval, report data at 5-second time ticks; i.e., 5, 15, 25,... GMT seconds. Thus, the data value reported at 15 seconds represents the average value in the interval equal to or greater than 10 seconds and less than 20 seconds.

For 60-second data (1-minute averages), the reporting times should be at 30-second time ticks; i.e., 30, 90, 150,... GMT seconds. Thus, the data value reported at 90 seconds represents the average value in the interval equal to or greater than 60 seconds and less than 120 seconds. Data reported in 60-second time intervals would be reported in the same format as for 10-second reporting, except the data reporting interval in header record 12 would be 60.

For other constant time intervals, time ticks of the data reporting interval should be consistent with the 10-sec or 60-sec format.

III.A.2 Variable Interval Time Reporting:

Report "0" for the Data Averaging Period in the 11th header record. When reporting intervals are irregular, it is necessary to report the sample start time, end time, and the midpoint time for each record.

IV. Sample Data Files

Header Sample Description 1

Figure 1: Standard Dataset: Time at Midpoint

Record	Explanation	Sample Data Record
1	NH	61
2	FileName	GTXY2D03.TRA
3	Experimenter Name, Instit	John Bradshaw & Scott Sandholm, GA. INST. OF TECH.
4	Species	NxOy/PF-LIF
5	Expedition	TRACE-A
6	Flt Date, Rev Date	92, 09, 21, 93, 06, 01
7	NF	3
8	NV	17
9	NC	32
10	DT	1
11	AvgPeriod	90
12	Samp Freq	1
13	Var(1)	Day, Julian (GMT), 1, 0,265, 265, -999, 0
13+1	Var(2)	Time, Sec (GMT), 1, 0, 58139, 72883, -999, 0
13+2	Var(3)	[NO], (pptv), 1, 0,,9.0, 595.0, -999.9, 1, -888.8, 9, -777.7, 9
13+3	Var(4)	sigma_NO, (pptv), 1, 0, 2.7, 18.0, -999.9, 1, -888.8, 10, -777.7,10
13+4	Var(5)	[NO2], (pptv), 1, 0, 34.5, 139.0, -999.9, 1, -888.8, 11, -777.7,11
13+8	Var(9)	LV_[NO], (pptv), 1, 0, -999.9, -999.9, -999.9, 0
13+9	Var(10)	LV_NO_sigma, (pptv), 1, 0, -999.9, -999.9, -999.9, 0
13+10	Var(11)	LV_[NO2], (pptv), 1, 0, 30.4, 78.3, -999.9, 0
		***
13+nv-1	Var(nv)	NOy_com_code, , 1, 0, 0, 6, -9,0
13+nv	Comment 1	The variable names that start with a "LV" are limiting values, either an upper
	Comment 2	or lower limit, (see the coding in the column for that molecule for details).
	Comment 3	The reported time is the center point of the integration period. The data is
	Comment 4	recorded at 30 second, the values reported are for 90 seconds signal
	Comment 5	integration periods. Calibration uncertainty (accuracy) is estimated to be
	Comment 6	approzimately +/-15 0.000000or [NO], +/-18 0.000000or [NO2], and +/-20 0.000000or [NOy] at
	Comment 7	the 95onfidence limit and should be treated as a random additive error
	Comment 8	term. Sigma values represent measurement precision estimates based on
	Comment 9	photon 95onfidence limit and should be treated as a random additive
	Comment 10	error term.
13+nv+nc-1	Comment nc	COMMENT CODE 12 = Lower limit estimate based on {NO} LOD value
13+nv+nc or nh+1	Data Rec 1	265,63463,114,7,49.2,15.2,1220,46,-999.9,-999.9,-999.9, -999.9, -999.9, -999.9,0,0,0
nh+2	Data Rec 2	265,63553,126,7.3,-888.8,-888.8,1160,45,-999.9,-999.9,35.8,17.9,-999.9,-999.9,0,2,0
nh+3	Data Rec 3	265,63643,132,7.7,-888.8,-888.8,1340,50,-999.9,-999.9,37.2,18.6,-999.9,-999.9,0,2,0

Header Sample Description 2

Figure 2: Standard Grab Sample: Start/Stop/Midpoint Time

Record	Explanation	Sample Data Record
1	NH	30
2	FileName	NHATGD03.TRA
3	Experimenter Name, Instit	Talbot, Robert, University of New Hampshire
4	Species	ACIDIC TRACE GASSES/MIST CHAMBER
5	Expedition	TRACE-A
6	Flt Date, Rev Date	92, 9, 21, 93, 4, 30
7	NF	3
8	NV	7
9	NC	11
10	DT	2
11	AvgPeriod	0
12	Samp Freq	1
13	Var(1)	Day, Julian (GMT), 1, 0, 265, 265, -999, 0
13+1	Var(2)	Start Time, Sec (GMT), 1, 0, 56490, 72450, -999, 0
13+2	Var(3)	Stop Time, Sec (GMT), 1, 0, 57303, 72935, -999, 0
13+3	Var(4)	Sample Midpoint, Sec (GMT), 1, 0, 56897, 72693, -999, 0
13+4	Var(5)	HNO3, pptv, 1, 0, 26, 195, -999,2, -888, 5, -777, -999
13+5	Var(6)	HCOOH, pptv, 1, 0,649, 2668, -99, 2, -88, 10, -77, -99
13+nv-1	Var(nv)	CH3COOH, pptv, 1, 0,202, 649, -99, 2, -888, 15, -777, -999
13+nv	Comment 1	Acidic gas data are stated in mixing ratios (molar ratio in
	Comment 2	part per trillion by volume, pptv). Mixing ratios below the
	Comment 3	limit of detection are indicated as -888. Estimated mean detection
	Comment 4	limits are as follows: 5 pptv HNO3, 10 pptv HCOOH, 15 pptv CH3COOH.
	Comment 5	Overall uncertainty in mixing ratios are 15-20 0.000000or HNO3;
	Comment 6	15 0.000000or HCOOH and; 20 0.000000or CH3COOH.
	Comment 7	CAUTION: Do not use these data on time scales shorter than
	Comment 8	those reported here. Direct inquires about these data to:
	Comment 9	Robert W. Talbot, Institute for the Study of Earth, Oceans, and
	Comment 10	Space, Morse Hall, University of New Hampshire, Nurham, NH 03824
13+nv+nc-1	Comment nc	Phone: 603-862-1546, Fax: 603-862-0188, E-mail: R_TALBOT@UNHH.UNH.EDU
13+nv+nc	Data Rec 1	265, 56490, 57303, 56897, -888, 649, 280
	Data Rec 2	265, 57569, 58410, 57990, 46, 776, 381

Header Sample Description 3

Figure 3: Vertical Column Sample: Variable Number Sampling Points (DIAL Datasets, etc.)

Record	Explanation	Sample Data Record
1	NH	27

2	FileName	BEABZD03.TRA
3	Experimenter Name, Instit	Browell, Dr. Edward V., NASA Langley Research Center
4	Species	DC-8 IR Zenith Aerosol Relative Backscatter
5	Expedition	GTE/TRACE-A
6	Flt Date, Rev Date	92, 9, 21, 93, 6, 4
7	NF	3
8	NV	9 (NOTE: This dataset type has NV-1 regular variables, and one array variable. The array variable is dimensioned by variable NV-1)
9	NC	6
10	DT	3
11	AvgPeriod	0
12	Samp Freq	0.0168 (Data recorded at 59.5 seconds)
13	Var(1)	Day, Julian (GMT), 1, 0, 265, 266, -999, 0
13+1	Var(2)	Start Time, Sec (GMT), 1, 0, 67386, 86400, -999, 0
13+2	Var(3)	Geometric altitude of aircraft, (m), 1, 0, 9398, 12939, -9999, 0
13+3	Var(4)	Geometric altitude at which data begins, (m), 1, 0, 9998,12576, -9999, 0
13+4	Var(5)	Altitude increment, (m), 1, 0,450, 450, -9999, 0
13+5	Var(6)	Latitude, (+N degrees), .01, 0, -90, 90, -999.999,0
	Var(7)	Longitude (+E degrees), .01, 0, -180, 180, -999.999,0
	Var(nv-1)	Number of data values, #, 1, 0, 0, 100, -99,0
13+nv-1	Var(nv)	Relative aerosol backscatter profile, , 1, 0, -50000,50000, -99999,0
13+nv	Comment 1	Final Reduced Resolution Archive of IR Aeroson Lidar Data - Vertical sampling interval is
	Comment 2	450 meters. Horizontal sampling interval is 59.5 seconds (apporximately 14 kilometers).
	Comment 3	Number of lines per reporting interval variable due to nature of data. Number of data
	Comment 4	points in profile located in variable NV-1 (Variable 8). Read through variable 8, then read
13+nv+nc-1	Comment nc	variable 9 is an array with 1 to (variable 8) elements.
13+nv+nc	Data Rec 1	265, 67386, 9397, 9997, 450, 2935, -9342, 0
13+nh+1	Data Rec 2	265, 67454, 9398, 9998, 450, 2934, -9325, 44 2050, 2547, 2580, 2511, 2340, 1804, 9494, 2580, 1625, 2602, 1775, 1809, 1789, 2864, 1853, 7281, 9107, 13225, 12607, 11518, 11493, 8504, 3184, 1067, 5842, 7404, 2391, 6119, -528, 66, 5435, 3422, 927, -3807, -7678, -3190, -9880, 13053, -1852m 1109, 5371, -9696, 5140, -3300
13+nh+2	Data Rec 3	265, 67515, 9400, 10000, 450, 2934, -9310, 44, 2034, 2620, 2499, 2482, 2194, 2087, 17791, 2389, 1372, 2117, 1273, 2036, 2883, 1611, 3915, 5565, 7568, 11035, 9493, 14609, 7538, 7979, 7112, 6753, 8772, 4974, 6670, 3325, 2622, 2405, 2014, -113, 3130, 2660, 5135, -7359, -1037, 5070, -3471, 3462, 12667, -1378, 17117, -5156

Header Sample Description 4

Figure 4: Vertical Column Sample: Standard Sonde Data (Ozonesondes, Rawinsondes, etc.)

Record	Explanation	Sample Data Record
1	NH	31
2	FileName	FJACPS01.TRA
3	Experimenter Name, Instit	FISHMAN, J., NASA LaRC
4	Species	OZONE/ECC4 SONDE
5	Expedition	PRE-TRACE-A
6	Flt Date, Rev Date	90, 07, 28, 93, 05, 28
7	NF	1
8	NV	11
9	NC	8
10	DT	4
11	AvgPeriod	0
12	Samp Freq	0
13	Var(1)	Day, Julian (GMT), 1, 0, 209, 209, -999, 0
13+1	Var(2)	Start Time, Sec (GMT), 1, 0, 67920, 74640, -999, 0
13+2	Var(3)	PRESSURE, (hPA), 1, 0, 8.2, 1009.9,-999.9,0
13+3	Var(4)	ALTITUDE, (m), 1, 0, 91, 32925, -9999,0
13+4	Var(5)	PARTIAL PRESSURE OZONE, (nb), 1, 0, 10.3, 151.8,-9.9,0
13+5	Var(6)	CUMULATIVE INTEGRATED OZONE, (at-cm), 1, 0, 0, .2248,-9.99999,0
	Var(7)	TEMPERATURE, (deg K), 1, 0, 119.4, 297.5, -999.9, 0
	Var(8)	OZONE NUMBERS DENSITY, , 1, 0, 3.6e11, 4.8e12, -999.9, 0
		DEW POINT TEMPERATURE, (DEG k) 1, 0, 195.8, 289.4, -999.99, 0
		OZONE, (ppbv), 1, 0, 35.87, 7904.73, -999.99, 0
13+nv-1	Var(nv)	RELATIVE HUMIDITY, (%), 1, 0, 12.13, 87.66, -999.99, 0
13+nv	Comment 1	DATA REPORTED AT VARIABLE TIME INTERVALS. 47 PRE-
	Comment 2	TRACE-A SONDES COVERING THE PERIOD OF JULY 1990 TO
	Comment 3	AUGUST 1992 AND 20 SONDES DURING TRACE-A COVERING
	Comment 4	THE PERIOD SEPTEMBER-OCTOVER 1992 ARE REPORTED.
	Comment 5	67 TOTAL ECC4 SONDES LAUNCHED AT ASCENSION ISLAND.
	Comment 6	LAT/LONG: 8 DEGREES SOUTH, 15 DEGREES WEST.
	Comment 7	STATION HEIGHT: 91 METERS.
13+nv+nc-1	Comment nc	ECC4 SONDE SERIAL NUMBER: 4A4685
13+nv+nc	Data Rec 1	209, 67920, 1009.9, 91, 36.2, 0.00000, 297.5, 8.8178E+11, 288.7, 35.87, 57.98
13+nh+1	Data Rec 2	209, 67980, 969.0, 452, 35.0, 0.00116, 295.2, 8.5951E+11, 289.2, 36.15, 68.69
13+nh+2	Data Rec 3	209, 68040, 931.0, 799, 34.8, 0.00227, 291.9, 8.6323E+11, 289.4, 37.37, 85.31

Header Sample Description 5

Figure 5: Gridded Dataset (Fires Gridded Data)

Record	Explanation	Sample Data Record
1	NH	31
2	FileName	SABFTI09.TRA

3	Experimenter Name, Instit	FISHMAN, J., NASA LaRC
4	Species	WEEKLY TOTAL PIXEL FIRE COUNTS/AVHRR
5	Expedition	TRACE-A
6	Flt Date, Rev Date	92, 10, 30, 93, 06, 06
7	NF	09
8	NV	11
9	NC	8
10	DT	5
11	AvgPeriod	0
12	Samp Freq	0
13	Var(1)	Start Day, Julian (GMT), 1, 0, 283, 283, -99, 0
13+1	Var(2)	Start Time, Seconds (GMT), 1, 0, -99, -99, -99, 0
13+2	Var(3)	Stop Day, Julian (GMT), 1, 0, 289, 289, -99, 0
13+3	Var(4)	Stop Time, Seconds (GMT), 1, 0, -99, -99, -99, 0
13+4	Var(5)	Start Latitude, Deg North, 1, 0, -44.5, -44.5, -999.999,0
13+5	Var(6)	Start Longitude, Deg East, 1, 0, -89.75, -89.75, -999.999,0
13+6	Var(7)	Latitude Increment, Deg North, 1. 0. 0.5, 0.5, -999.9, 0
13+7	Var(8)	Longitude Increment, Deg East, 1, 0, 1.0, 1.0, -999.9, 0
	Var(9)	Number of Rows of Latitude, #, 1, 0, 60, 60, -999, 0
	Var(10)	Number of Columns of Longitude, #, 1, 0, 140, 140, -999, 0
13+nv-1	Var(nv)	Pixel Count, Number, 1 0, -40, 521, -120, 0
13+nv	Comment 1	THIS DATA FILE OF, SATELLITE CHANNEL 3 AVHRR BASED, WEEKLY TOTAL
	Comment 2	PIXEL FIRE COUNTS ARE BASED ON DATA FROM OCTOBER 9-15, 1992.
13+nv+nc-1	Comment nc	READ "FIRECNT1.RME" for CAPABILITIES AND LIMITATIONS
13+nv+nc	Data Rec 1	283, -99, 289, -99, -44.5, -89.75, .5, 1.0, 60, 140,
13+nh+1	Data Rec 2	-120., -40., -8., -40., -40., 2., 48., 1., 44., 132., 178., 70., 30., 33., 71., 64., 188., 1., 500., 121., 90., 43., 24., -40., 17., 479., 37., 63., 120., 115., 199., 217., 56.., 46., 51., 24., 55., 107., 41., 13., 11., 28., 1., 3., 103., 92., 63., 35., 44., 67., 34., 3., 40., 6., -40., 5., 2., 2., 6., -40., -40., -40., -40., -40., -40., -120.,

13+nh+2	Data Rec 3	-120., -40., -40., -40., 13., -40., -40., -40., -40., 3., -40., 10., 4., 95., 143., 471., 3., 60., 299., 380., 36., 48., 31., 555., 31., 4., 60., 7., -40., 5., 1., 48., 37., 44., 23., 31., 32., 9., 2., 12., 11., 27., 7., 3., 4., -40., 5., -40., -40., -40., -40., -40., -40., -40., -40., -40., -40., -40., -120.,

Header Sample Description 6

Figure 6: Trajectory Dataset (Isentropic and Isobaric Trajectory Data)

Record	Explanation	Sample Data Record
1	NH	26
2	FileName	MJ101d04.PWA
3	Experimenter Name, Instit	Merrill, John, University of Rhode Island
4	Species	Isentropic Air Mass Trajectory
5	Expedition	PEM West-A
6	Flt Date, Rev Date	91, 09, 16, 94, 12, 29
7	NF	4
8	NV	7
9	NC	7
10	DT	6
11	AvgPeriod	43200
12	Samp Freq	21
13	Var(1)	Day, Julian (GMT), 1, 0,259, 260, -999, 0
13+1	Var(2)	Time, Sec (GMT), 1, 0, 64740, 22883, -999, 0
13+2	Var(3)	Potential Temperature, Kelvin, 1, 0, 302, 343, -999, 0
13+3	Var(4)	Latitude, Degrees North, 1, 0, 14.2, 66.2, -99.9, 0
13+4	Var(5)	Longitude, Degrees East, 1, 0, -179.8, 179.7, -999.9, 0
13+5	Var(6)	Pressure, hPa (same as millibars), 1, 0, 213, 864, -999, 0
13+6	Var(7)	Height, meters above sea level, 1, 0, 1445, 11452, -9999, 0
13+nv	Var(nv)	NOy_com_code, , 1, 0, 0, 6, -9,0
13+nv	Comment 1	This and every trajectory is described in 21 data lines. The first line
13+nv+1	Comment 2	is the trajectory end point, corresponding to a selected aircraft time
13+nv+2	Comment 3	and position. Each following line is the trajectory's position on the previous
13+nv+3	Comment 4	time step. Time steps prior to the start time are all at 0000 and 1200 GMT.
13+nv+nc-1	Comment nc	For each trajectory read 21 data lines.
13+nv+nc or nh+1	Data Rec 1	259, 64740, 335, 42.4, -126.0, 285, 10076 *(NOTE: Time here is the actual time during the flight, referred to as the trajectory end point)
nh+2	Data Rec 2	259, 43200, 335, 43.1, -127.8, 279, 10217 *(NOTE: This time is the next even increment of the Averaging Period in Header Line 11)

nh+3	Data Rec 3	259, 0, 335, 40.7, -131.7, 285, 10046
nh+4	Data Rec 4	258, 43200, 335, 32.9, -134.1, 279, 10086
nh+5	Data Rec 5	258, 0, 335, 24.1, -137.8, 330, 8991
nh+6	Data Rec 6	257, 43200, 335, 20.1, -143.0, 411, 7358
nh+7	Data Rec 7	257, 0, 335, 19.3, -146.6, 442, 6847
nh+8	Data Rec 8	256, 43200, 335, 20.1, -150.5, 435, 6962
nh+9	Data Rec 9	259, 0, 335, 22.1, -153.8, 422, 7205
	...	...
nh+21	Data Rec 21	250, 0, 335, 33.6, -145.2, 305, 9581
nh+22	Data Rec 22	259, 67680, 334, 46.6, -131.5, 285, 10039* (NOTE: This record is the second trajectory end point. 20 more data lines follow.)

V. FTP Protocol for GTE Data Exchange

DATA ARCHIVE MEMO: FTP File Transfer to/from the GTE Archives

Some GTE Data Archive datasets are now located on ANONYMOUS FTP at the TYPHOON machine at NASA Langley and may be accessed using standard FTP protocol. Type the quoted (" ") commands to logon and access the various on-line files. Note that you can now submit your data files via FTP. Follow the instructions given below to logon and transfer files to/from the GTE Archives. NOTE: The proprietary archives (those not yet released to the public) are protected from being read by the public. Please notify the Archive Facility prior to accessing these files, so that we may give you a time period for data access.

LOGGING ON TO THE GTE ARCHIVE

1. Access via FTP the GTE Archives computer "typhoon.larc.nasa.gov". If this name is not on your host table, address the computer using the IP address of 128.155.17.246.
2. Logon using "anonymous" as your name.
3. Enter your E:mail name as your password (e.g. "d.w.owen").
4. "cd pub" to change directories to the public directory. (UNIX is case sensitive - use lower case for all commands. Our GTE main directories are in all caps, although subdirectories may be in either.

LOCATING THE GTE ARCHIVE FILES YOU WISH TO TRANSFER TO YOUR COMPUTER

5. Change to the directory of the GTE Expedition archive files of interest. For example, "cd PEMWESTA" or "cd ABLE3A"
6. "ls" to list the files and/or sub-directories in this directory. NOTE: The GTE Archives have the following long directory structure:
   ftp/pub/Expedition/Sub-category/Investigator.Institution/Species/Files
   PLEASE NOTE: not all Expeditions have "Sub-categories" and most Investigators do not have a "Species" sub-directory.
7. Assuming you changed directories to PEMWESTA, you can now change to the subdirectory of interest. For example, "cd dc-8/heikes.uri" will put you in the dc-8 aircraft subdirectory, and then to the investigator heikes.uri subdirectory. You can now access files in that subdirectory. FTP only allows you to retrieve files in one subdirectory at a time, thus type "cd .." to return to the parent directory if you desire to view the subdirectory containing files of another investigator.

TRANSFER COPIES OF THE FILES TO YOUR COMPUTER

8. Select the local drive and directory on your machine to receive the data. Useful commands:
   "drive x:" to change the local drive to "x:",
   "lcd xxxxx" to change the local directory to "xxxxx",
   "ldir" to get a directory listing of the local drive.

"help" to get a complete listing of the FTP commands on your machine.
Your actual commands may differ slightly from these.
9. The command "mget *.*" will get all files in the subdirectory.
10. "cd .." to transfer to the parent directory: return to (7) to repeat the process.

SENDING FILES TO THE GTE ARCHIVE USING FTP

11. Change directories to GTE-PUT (path = ftp/pub/GTE-PUT). Useful change directory commands include: "cd .." to change to the parent directory and "cd /" to change to the root directory.
12. Change to the appropriate subdirectory: "cd TRACEA" to send GTE TRACE-A files.
13. Create a subdirectory for your files if one does not already exist. Use your name and institution to name the subdirectory. For example, "mkdir bradshaw.git" will create a subdirectory for files from the John Bradshaw investigation team of Georgia Tech.
14. Use the commands of (8) to go to your local drive and directory where the files to be transferred are stored.
15. Transfer the files to the TYPHOON computer with one of the following commands: "put filename" To transfer one file from your machine to TYPHOON. You will be prompted to enter the foreign file name. "mput *.*" To transfer all files in your host directory and keep same names. "mput gt*.*" To transfer all files beginning with "gt" in your host directory.
16. Logoff the system. "quit"
17. Contact the GTE Data Management Office (use e-mail if possible) and describe the action you have taken. Provide a list of files transferred and the directory path on the TYPHOON that was used:

GTE Data Management Office
Dennis W. Owen
e-mail: d.w.owen@larc.nasa.gov
voice: (757) 864-5837
fax: (757) 864-5841

***** END OF INFO ON FTP TRANSFER *****

Table 1. Principal Investigator Codes

Principal Investigator	Institution	PI Code
Akimoto, H.	Nat. Inst. for Environmental Studies	AH
Anderson, B.E.	NASA Langley Research Center	AB
Arimoto, R.	University of Rhode Island	AR
Bandy, A.R.	Drexel University	DU
Barrick, J.D.W.	NASA Langley Research Center	PO
Bradshaw, J.	Georgia Institute of Technology	BJ
Browell, E.V.	NASA Langley Research Center	BE
Carmichael, G.R.	University of Iowa	UI
Chameides, W.L.	Georgia Institute of Technology	CW
Chatfield, R.	NASA Ames Research Center	CR
Davis, D.	Georgia Institute of Technology	DD
Gregory, G.L.	NASA Langley Research Center	GG
GTE Project Office	NASA Langley Research Center	PO
Heikes, B.	University of Rhode Island	HB
Kelly, K.	NOAA Aeronomy Laboratory	KK

Kitada, T.	Toyohashi University, Japan	TU
Kondo, Y.	Nagoya University, Japan	NG
Lui, C.M.	National Taiwan University	LC
Lui, S.C.	NOAA Aeronomy Laboratory	LS
Merill, J.	University of Rhode Island	MJ
Mission Manager Logs	NASA Ames Research Center	MM
Park, J.K.	Korean Institute of Science and Technology	KI
Prospero, J.M.	University of Miami	UM
Pueschel, R.	NASA Ames Research Center	PR
Ridley, B.	NCAR	NR
Rodriguez, J.	Atmos. & Environ. Research, Inc.	RJ
Rowland, F.S.	Univ. California - Irvine (UCI)	UC
Sachse, G.W.	NASA Langley Research Center	SG
Sakamaki, F.	Nat. Inst. for Environmental Studies, Japan	SF
Singh, H.B.	NASA Ames Research Center	SH
Talbot, R.W.	University of New Hampshire	NH
Xiuji, Z.	Academy of Meteorological Science, Peoples Republic of China	AM

Table 1: Principal Investigator Codes

Table 2. GTE Expedition Codes

GTE Expedition		Expedition Location	3-Digit Code
CITE 1	1983	Wallops Island, VA	CTW
CITE 1	1983	Hawaii	CTH
CITE 1	1984	Eastern North Pacific - off the California coast	CTA
ABLE 1	1984	Barbados, French Guyana	AB1
ABLE 2A	1985	Amazon Basin	A2A
CITE 2	1986	Western USA	CT2
ABLE 2B	1987	Amazon Basin	A2B
ABLE 3A	1988	Alaska - Barrow, Bethel	A3A
CITE 3	1989	Western North Atlantic - off the Virginia coast Western South Atlantic - off the Brazil coast	CT3
ABLE 3B	1990	North Bay, Ontario - Goose Bay, Labrador	A3B
PEM-West A	1991	Western North Pacific Rim	PWA
TRACE A	1992	Brazil, South Atlantic, Southwest Africa	TRA
PEM-West B	1994	Western North Pacific Rim	PWB

Table 2: GTE Expedition Codes