

TOLNet Technical Document (TD-2): TOLNet Surface Data Format Description

Document History:

2013/07/16 Document version 1.0, released for official use by TOLNet

Introduction

The TOLNet surface data file format was discussed during the last quarter of 2012. Its format structure follows closely that of the TOLNet profile data format described in a separate document. The present documentation describes the surface data format and makes frequent references to the description of the profile data format. It is therefore strongly recommended that users first read the Profile Data Format Documentation (TD-1). **The TOLNet surface data format version at the time this documentation was last updated is v1.0.** It will be incremented each time a “physical” format change is implemented and updated documentation will be provided accordingly.

General concept

The agreed format was built upon past experience and achievements from various atmospheric observation networks or projects, principally NDACC (Network for the Detection of Atmospheric Composition Change), and ICARTT (International Consortium for Atmospheric Research on Transport and Transformation). It was agreed to conserve a high degree of consistency between the TOLNet Profile Data Format and the TOLNet Surface Data Format. Consequently, it was decided that an ASCII format containing a unique general header followed by one or multiple blocks, each of which having a similar structure and reporting data and metadata for a specific measurement time window, was the most appropriate choice for TOLNet.

Filename Structure

The TOLNet filename structure of the TOLNet data files is strongly inspired from that of the ICARTT data files. It is of the form: TOLNet-O3Surface *T_S_YYYYMMDD_Rv[_c].XXX*, and is broken up as follows:

TOLNet-O3Surface = Prescribed 16-character string for all TOLNet surface instruments

T = Prescribed free-length string describing the measuring technique (e.g.: “Photometer”)

S = Free length string for the TOLNet site/station name (PI’s choice)

YYYYMMDD = 8-character string for the date of the first profile in file (year, month, day)

R = Prescribed 1-character string referring to the next character (the revision number)

v = 1- or 2-character string for the data revision number (0-99)

c = Optional, free length string used by PI to report special (e.g., non-routine) data

XXX = extension depending on file content type: “dat” for data, “png” or “jpg” for images, etc...

Delimiters

Commas are used to delimit data in the file header and data section. At the end of each line in the header, a semicolon is used, followed by a brief text describing the content on that line (see examples).

Surface data file general structure

The TOLNet surface data format follows closely that of the TOLNet profile data described in the main documentation. The following formatting rules apply to all TOLNet surface data files:

- 1) There may be one and only one data file per day (recorded in Universal Time) for a given TOLNet instrument.
- 2) There may be one continuous or several measurement time windows (referred to as “records” thereafter) in each data file
- 3) Each data file must be built of 4 data/metadata blocks, the structure of which is being prescribed by the TOLNet Data Format version. For version 1.0, these blocks are:
 - a. The general header
 - b. The general comments section
 - c. The record header
 - d. The data section
- 4) Each file must contain a unique “general header” followed by a unique “general comments section”, both located before all the actual data
- 5) The general header contains the following information:
 - a. The file structure (incl. format version and other formatting details)
 - b. The number of measurement records reported in the file
- 6) The general comments section contains the following information:
 - a. General meta-data information (e.g., PI/contact info, site name, geo-location, etc.)
 - b. Data revision information, applying to entire file (i.e., not record-specific)
- 7) Each record is then reported sequentially in the form of a “record header” followed by the 2-dimensional (time x variables) data (the “data section”)
- 8) Each record header contains the following information:
 - a. Number of data lines in record
 - b. Measurement and analysis conditions
 - c. Any meta-data information pertaining to the record and deemed critical by the TOLNet community (see details in record header description)
- 9) The last line of each record header serves as data column labels, and contains all data variables’ short names separated by commas
- 10) Each record header is preceded by a separator line starting with: “#BEGIN RECORD”
- 11) Each record header is immediately followed by the record data, reported in the form of multiple lines of floating-point numbers separated by the comma character “,”
- 12) The number of columns in the data section is constrained by the TOLNet format version
- 13) The number of lines and columns in the data section must always be consistent with the values reported in the record header and general header respectively.

Structure and information content between line 1 and first data line

	Description	Type	Dim.	Prescription	Ref.
Line 1	Number of lines to follow in general header, <u>excluding</u> general comments	Integer	Scalar	Fixed by TOLNet format version	<i>ngh</i>
Line 2	TOLNet format version	String	Scalar	Fixed by TOLNet format version	
Line 3	Number of records in file	Integer	Scalar	Set by data originator	<i>nrec</i>
Line 4	Number of data columns for all profiles	Integer	Scalar	Fixed by TOLNet format version	<i>ncol</i>
Lines 5 to 5+ <i>ncol</i> -1	Description of the <i>ncol</i> data variables (see list in TOLNet format version description)	Strings	Scalars	Fixed by TOLNet format version	
Line 5+ <i>ncol</i> (also 1+ <i>ngh</i>)	Missing values used for each column, separated by commas	String	Scalar	Fixed by TOLNet format version	
Line 2+ <i>ngh</i>	Number of lines to follow in general comments section	Integer	Scalar	Fixed by TOLNet format version	<i>ngc</i>
Lines 3+ <i>ngh</i> to 6+ <i>ngh</i>	First 4 general comments lines (see list in TOLNet format version description)	Strings Integers Floats	Scalars	Fixed by TOLNet format version	
Line 7+ <i>ngh</i> 1 st character	Letter "R" ("R" for "revision")	String	Scalar	Fixed by TOLNet format version	
Line 7+ <i>ngh</i> next character(s)	Revision number ("0" if first time data)	Integer	Scalar	Set by data originator	
Lines 8+ <i>ngh</i> to 2+ <i>ngh</i> + <i>ngc</i>	Revision comments lines (mandatory if rev number is not "0")	Strings	Scalars	Set by data originator	
Line 3+ <i>ngh</i> + <i>ngc</i>	Profile separator: must contain "#BEGIN PROFILE"	String	Scalar	Fixed by TOLNet format version	
Line 4+ <i>ngh</i> + <i>ngc</i>	Number of lines to follow in record header <u>including</u> all comments	Integer	Scalar	TOLNet format and PI	<i>nph</i>
Lines 5+ <i>ngh</i> + <i>ngc</i> to 5+ <i>ngh</i> + <i>ngc</i> + <i>nph</i> -2	Record header lines (see list of metadata in Record Header description)	Strings Integers Floats	Scalars	Set by data originator	
Line 5+ <i>ngh</i> + <i>ngc</i> + <i>nph</i> -1	Data variables short names for easy visual identification of data columns (separated by commas)	Strings	Scalars	Fixed by TOLNet format version	
Line 5+ <i>ngh</i> + <i>ngc</i> + <i>nph</i>	First data line (<i>ncol</i> scalars separated by commas)	Floats	Scalars	Set by data originator	

Prescribed surface data variables for TOLNet Format version v1.0

Number of data columns: 11

	Description	Unit	Short name	Nb. of significant decimals	Missing value
Column 1	Time at start of measurement, elapsed since 00:00:00 UT same day	s	StartTime	<i>i5</i>	-9999
Column 2	Time at end of measurement, elapsed since 00:00:00 UT same day	s	EndTime	<i>i5</i>	-9999
Column 3	Weighted average time of measurement, elapsed since 00:00:00 UT same day	s	MeanTime	<i>i5</i>	-9999
Column 4	Surface Ozone mixing ratio (measured)	ppbv	O3MR	<i>f0.2</i>	-9999
Column 5	Surface Ozone mixing ratio uncertainty	ppbv	O3MRUncert	<i>f0.2</i>	-9999
Column 6	Measurement Precision	%	Precision	<i>f0.2</i>	-9999
Column 7	Surface Pressure	hPa	Press	<i>e0.3</i>	-9999
Column 8	Surface Temperature	K	Temp	<i>f0.2</i>	-9999
Column 9	Surface Relative Humidity with respect to water	%RH	RH	<i>f0.1</i>	-9999
Column 10	Surface Wind Speed	m.s-1	WndSpd	<i>f0.2</i>	-9999
Column 11	Surface Wind Direction, clockwise degrees from north	deg	WndDir	<i>f0.1</i>	-9999

Prescribed meta-data in General Comments Section for TOLNet Format version v1.0

Number of metadata lines: 5

	Description	Unit	Type
Line <i>ngh</i> +3	Instrument name	N/A	String
Line <i>ngh</i> +4	PI and contact info (separated by commas)	N/A	Strings
Line <i>ngh</i> +5	Site name	N/A	String
Line <i>ngh</i> +6	Longitude, latitude, altitude (separated by commas)	deg. East deg. North m	Floats
Line <i>ngh</i> +7 character(s) following "R"	Revision number	Undimensional	Integer

Prescribed meta-data in Record Header for TOLNet Format version v1.0

Number of metadata lines: 10

	Description	Unit	Type	Ref.
Line <i>ngh+ngc+5</i>	Number of data lines in record	Undimensional	Integer	<i>ntime</i>
Line <i>ngh+ngc+6</i>	Data processing date and time	YYYY-MM-DD, HH:MM:SS	Strings	
Line <i>ngh+ngc+7</i>	Data processing software name and version	N/A	String	
Line <i>ngh+ngc+8</i>	Result quality: must be NOMINAL, FAIR or GOOD	N/A	String	
Line <i>ngh+ngc+9</i>	Start date/time of profile	YYYY-MM-DD, HH:MM:SS	Strings	
Line <i>ngh+ngc+10</i>	End date/time of profile	YYYY-MM-DD, HH:MM:SS	Strings	
Line <i>ngh+ngc+11</i>	Source of surface pressure, temperature, humidity and wind	N/A	String	

Optional comments for TOLNet Format version v1.0

Optional comments may be added at two locations:

- 1) Data revision comments may be added at the end of the general comments section, immediately following line *ngh+7*. If data is of revision number “0”, there should be no added comment (and therefore no added line). Otherwise, each increment in revision number should be associated with at least one additional comment line, most recent comments being on top of older comments. These comments pertain to the description of the revisions that were made and led to an update of the data files
- 2) At the end of a record header, comments may be added just before the line listing the short names (line $5+ngh+ngc+nph-1$). These comments pertain to each individual record being reported

What is similar to the “Profile data” format?

- 1) The formatting rules (including the use separators) are identical
- 2) The structure and meta-data of the general header are identical
- 3) The structure of the “record headers” is identical to that in the “profile headers”

What is different from the “Profile data” format?

- 1) The number and names of data variables (data columns) are different
- 2) The number of meta-data in the “record headers” is different from that in the “profile headers”
- 3) The “independent” variable in the data section is no longer altitude, but time (seconds elapsed from 00 UT)