The WOUDC extended Comma Separated Values (extCSV) Format File Header Description

by

E.W. Hare and E.J. Carty

-

Released by:

WMO/WOUDC, Toronto, January, 2000

-

Available at :

http://www.tor.ec.gc.ca/woudc/publications_reports/O3_guidev2.html

-

A s a m p l e

for creation / coding of extCSV files / reports of

Dobson Total Ozone Observations

Adjusteded by: Karel Vanicek, as an introduction for observers, May, 2000

What is the WOUDC Header?

The WOUDC header presents a set of variables that uniquely describes the data to follow in that particular file, and that are common to all data sets independent of the specific WOUDC data. The uniqueness of each field within the header is required such that if a duplicate data file was received, one or more of the header variables would have to have changed in order for this file to be accepted in to the data archive.

-

-

-

General structure of a extCSV Report / File

Metadata Header Format Rules

1. The order of the static metadata tables is fixed.
2. A table name is in UPPERCASE and denoted by the pound symbol (#).
3. A table is comprised of three sections: a **table name**, field (column) names with units (where applicable) and a **data record(s)**.
4. Columns are delimited by commas within each record which includes each field (column) name and datum.
5. Comments may appear anywhere within the file and are denoted by an asterisk (*).

-

-

-

Metadata Header Content Rules

1. Only ONE instrument may be represented in a file.

2. To report changes in time and space, more than one #LOCATION and #TIMESTAMP table is permitted.
3. Since most data submitted are considered from a stationary source, a single #LOCATION table is all that is required.
4. Further information about site conditions, meteorology, sky images etc. may be included as comments or by other tables generated at the discretion of the data originator(s). For example, if a data originator wishes to indicate an occurrence of a special note ("instrument was bumped while sampling") than a comment is probably reasonable.
5. Dynamic variations occurring instantaneously such as meteorology: eg. sky or wind conditions) or the reporting of calibration data; these examples would be represented in a table.

```
*Example 2

#SITE METEOROLOGY

Date,Time,Temperature,Pressure,WindDirection,Wi
nd Speed
```

-

-

Description of the Tables and Fields

Static metadata

#CONTENT

Class - The general type of data to follow. WOUDC would be used here.

Category - Sub group of data submitted. For example: OzoneSonde or TotalOzone

Level - Data level. The level refers to the data product. Raw data would be Level 0, processed data Level 1 etc..

Form - The version of the data format to follow. The specification for the Form (initial value=1) refers to the table descriptions for that index value. If a data table is already defined (form=1) with a specific number of columns, then the addition of one or more columns to that table would require the Form value to be indexed by 1 (form=1+i, where i=1) for every change.

#DATA_ GENERATION

Date - Date the file was processed and/or generated. Date is represented in the **ISO 8160** format standard, i.e. **yyyy-mm-dd**. Example 1998-09-21.

Agency - The acronym of the submitting agency.

Version - Data version specified by the submitting agency. These versions have the form **major.minor** (eg. 3.2) where major values are incremented with changes to the processing algorithm and minor values are incremented when the characterisation or calibration values have changed.. Note, minor values are reset to zero with changes to the processing algorithm.

ScientificAuthority - The ScientificAuthority is the person(s) responsible for the data quality

#PLATFORM

Type - Type of observing platform Stationary (STN) or Moving (Examples: Airborne (FLT) Ship borne, (SHP) etc.). The default is. This field, for example, will accommodate the future inclusion of an onboard GPS on sonde flights which records exact location throughout a flight. Thus, FLT would be selected instead of STN.

ID - Unique station or flight ID assigned by the WOUDC. The station number is a pointer to other information about the station such as the observation programs, the different instruments in operation and related site information like the proximity to an urban center and other background effects.

Name - Station Name

Country - The country where the station is located. Country is the 3-letter **ISO-3166** code. For example, Finland is FIN, Switzerland is CHE

GAW_ID - Applicable only to stations with the 5-digit WMO number.

#INSTRUMENT

Name - Common name of instrument. For example, the Dobson spectrophotometer is called the "Dobson"

Model - Model ID where applicable. For the Dobson would be Beck or Japanese..

Number - Serial number of the instrument

Description of the Tables and Fields

Dynamic metadata

-

#LOCATION

Latitude - Benchmark latitude of the instrument. Typically an instrument is located at a fixed station location and shares the station latitude. The format is decimal degrees (deg.).

Longitude - Benchmark longitude of the instrument. Typically an instrument is located at a fixed station location and shares the station longitude. The format is decimal degrees (deg.).

Height - Benchmark height of the instrument. Typically an instrument is located at a fixed station location and shares the station elevation. The format is meters above sea level (m.a.s.l.)

-

#TIMESTAMP

UTCOffset - The UTCOffset is the amount of time SUBTRACTED from the local time in order to obtain UTC time (where East is positive and Local Time as defined by the data originator). This is equivalent to UTC=time-UTCOffset. UTCOffset is represented in the **ISO-8601** standard of **hh:mm:ss.** The default is to report time in UTC, thus UTCOffset=+00:00:00.

Date - The Date corresponding to the time code represented in the **ISO 8160** standard. E.g. for reports with Dobson total ozone data:

- for single OBSERVATIONS report the Date = date of the particular day on which the observations were taken
- for DAILY and MONTHLY reports the Date = date of the first day of a month in which daily representative values or monthly averages were calculated

-

-

Description of the Tables and Fields

Total ozone data - single observations

-

#OBSERVATIONS

Time,WLcode,ObsCode,Airmass,ColumnO3,StdDevO3,

Time - hh:mm:ss - time of observation

WLCode - Code to designate the wavelength pair(s) used for total ozone measurement

ObsCode - Code to designate the type of total ozone measurement

Airmass - Relative slant plant through atmosphere

ColumnO3 - Discrete total column ozone (O3) amount (in Dobson units, i.e., m atm-cm) measured at the time of observation.

StdDevO3 - Standard Deviation of total column ozone measurement.

Example: **A report with single observations from Hradec Kralove, 11,02. 2000:**

#CONTENT

Class,Category,Level,Form

WODUC,TotalOzone,1.0,1

#DATA_GENERATION

Date,Agency,Version,ScientificAuthority

2000-02-17,CHMI-HK,0.0,Martin Stanek

#PLATFORM

Type,ID,Name,Country,GAW_ID

STN,096,Hradec Kralove,CZE,11649

#INSTRUMENT

Name,Model,Number

Dobson,Beck,74

#LOCATION

Latitude,Longitude,Height

50.180,15.833,290

#TIMESTAMP

UTCOffset,Date,Time

+01:00:00,2000-02-11,

#OBSERVATIONS

Time,WLCode,ObsCode,Airmass,ColumnO3,StdDevO3,ColumnSO2,StdDevSO2

11:04:00,0,2,2.422,357

11:07:00,0,2,2.409,360

11:10:00,0,2,2.397,360

11:30:00,0,0,2.331,353

11:33:00,0,0,2.323,352

13:02:00,0,0,2.357,356

13:05:00,0,3,2.367,359

13:08:00,0,3,2.378,359

Description of the Tables and Fields

Total ozone data - Daily averages

#DAILY

Date,WLCode,ObsCode,ColumnO3,StdDevO3,UTC_Begin,UTC_End,UTC_Mean, nObs,mMu,ColumnSO2

Date - yyyy-mm-dd

WLCode - Code to designate the wavelength pair(s) used for total ozone measurement

ObsCode - Code to designate the type of total ozone measurement

ColumnO3 - Daily value of total column ozone amount (in Dobson units, i.e., m atm-cm) defined as the "best representative value" in order of DS, ZS and FM.

StdDevO3 - Is the estimated population standard deviation of the total column ozone measurements used for the daily value.

UTC_Begin - The starting time of observations (in decimal hours, UTC).

UTC_End - The ending time of observations (in decimal hours, UTC).

UTC_ Mean - The mean time of observations (in decimal hours, UTC).

nObs - Number of observations used to calculate the total column ozone value.

mMu - The harmonic mean of the relative slant path at 22Km (μ 1) for each of the observations used to compute the daily value. It is a useful statistic in relation to effects of uncertainties in the zero airmasss extrapolations and in the generation of simulated data from independent data. $mMu=NN/\Sigma (1/\mu_i)$

ColumnSO2 - The mean daily total column of sulphur dioxide (SO2) - not from Dobson stations

Example: **A monthly report of daily averages, Hradec Kralove, January, 2000:**

#CONTENT

Class,Category,Level,Form

WOUDC,TotalOzone,1.0,1

#DATA_GENERATION

Date,Agency,Version,ScientificAuthority

2000-02-16,CHMI,1.0,Martin Stanek

#PLATFORM

Type,ID,Name,Country,GAW_ID

STN,096,Hradec Kralove,CZE,11649

#INSTRUMENT

Name,Model,Number

Dobson,Beck,74

#LOCATION

Latitude,Longitude,Height

50.180,15.833,290

#TIMESTAMP

UTCOffset,Date,Time

+01:00:00,2000-01-01,

#DAILY

Date,WLCode,ObsCode,ColumnO3,StdDevO3,UTC_Begin,UTC_End,UTC_Mean,NObs,mMu,ColumnSO2

2000-01-01,0,4,314,1.1,11.67,12.02,11.80,3,3.36,

2000-01-02,

2000-01-03,

2000-01-04,

2000-01-05,0,0,318,2.3,11.40,12.77,12.13,6,3.36,

2000-01-06,0,0,300,0.9,11.72,13.18,12.45,4,3.40,

2000-01-07,

2000-01-08,

2000-01-09,0,5,326,1.0,11.72,12.85,12.28,2,3.27,

2000-01-10,

2000-01-11,0,4,323,2.2,11.58,12.97,12.28,6,3.24,

2000-01-12,0,0,287,3.6,11.35,12.75,12.05,8,3.20,

2000-01-13,

2000-01-14,

2000-01-15,

2000-01-16,0,0,283,1.6,10.97,12.35,11.44,3,3.18,

2000-01-17,

2000-01-18,0,0,334,1.4,13.00,13.05,13.02,2,3.11,

2000-01-19,0,2,357,3.4,12.83,12.93,12.88,3,3.04,

2000-01-20,

2000-01-21,0,2,343,0.5,11.78,11.88,11.83,3,2.89,

2000-01-22,

2000-01-23,0,0,373,7.7,12.92,13.07,12.99,4,2.94,

2000-01-24,2,0,404,11.3,11.98,12.92,12.17,9,2.82,

2000-01-25,0,4,295,1.3,13.02,13.12,13.07,3,2.90,

2000-01-26,

2000-01-27,

2000-01-28,0,0,317,4.1,10.62,12.70,11.41,12,2.83,

2000-01-29,

2000-01-30,0,0,316,1.8,10.92,11.02,10.97,3,2.83,

-

-

-

-

Description of the Tables and Fields

Total ozone data - Monthly averages

#MONTHLY

Date,ColumnO3,StdDevO3,Npts

Date - yyyy-mm-dd

ColumnO3 - Daily value of total column ozone amount (in Dobson units, i.e., m atm-cm) defined as the "best representative value" in order of DS, ZS and FM.

StdDevO3 - Standard Deviation of daily total column ozone measurement

Npts - The number of points (typically this is the number of daily averages) used to determine the monthly mean ozone value.

Example: **A report of monthly average, Hradec Kralove, January, 2000:**

#CONTENT

Class,Category,Level,Form

WOUDC,TotalOzone,1.0,1

#DATA_GENERATION

Date,Agency,Version,ScientificAuthority

2000-02-16,CHMI,1.0,Martin Stanek

#PLATFORM

Type,ID,Name,Country,GAW_ID

STN,096,Hradec Kralove,CZE,11649

#INSTRUMENT

Name,Model,Number

Dobson,Beck,74

#LOCATION

Latitude,Longitude,Height

50.180,15.833,290

#TIMESTAMP

UTCOffset,Date,Time

+01:00:00,2000-01-01

#MONTHLY

Date,ColumnO3,StdDevO3,Npts

2000-01-01,326,32.7,15