

CRU TS version 4 data availability and file formats explained

Where to find all of the CRU TS version 4 data and metadata files:

All of the CRU TS v4 data (i.e. climate variables) and metadata files (i.e. stations and observations) are available from the CEDA Archive at:

http://badc.nerc.ac.uk/browse/badc/cru/data/cru_ts/

This directory contains all the CRU TS versions and sub directories of monthly gridded variables are within the "data" directory.

All the data files (ASCII ".dat" and netcdf ".nc") are compressed (.gz extension). From version 3.21 onwards for each variable a single NetCDF file that spans the full time period of the version is made available in an uncompressed format for use with the CEDA OpenDAP service.

Note: The files are >2GB when unzipped, which may cause problems on the computer systems of some users.

CRU TS v4 Data and Metadata File Formats explained:

The CRU TS v4 data are stored in both ASCII and NetCDF formats:

- ASCII data: The 360-lat x 720-long grid is presented exactly as that, with 720 columns, and 360 rows per timestep.
The first row in each grid is the southernmost (centred on 89.75S).
The first column is the westernmost (centred on 179.75W).
There are the scaling factors in use in the data files (see Table below). For example: One gets the whole global grid for the first time step, then the whole grid for the second, and so on. So the first 360 rows show the data for Jan 1901, the next 360 rows show the data for Feb 1901, the next 360 rows after show the data for March 1901 and so on.
- In the ASCII text files, missing values are stored as '-999'.
- NetCDF data: There are no scale factors in the NetCDF files because the data is FLOAT instead of INT.
 - Please see the CEDA [NetCDF](#) help page for more information on this data format if you are not familiar with it.

How to read the CRU TS v4 data:

The CRU TS v4 data files contain:

Label	Variable	Units (Multiplying factor for ASCII data ONLY)	Comments
cld	Cloud Cover	percentage (x10)	
dtr	Diurnal Temperature Range	Degrees Celcius (x10)	the diurnal temperature range is the difference between the daily minimum and maximum temperatures

frs	Frost Day Frequency	Days (x100)	Frost days are constructed synthetically from monthly TMN. The process is described in: Representing Twentieth-Century Space-Time Climate Variability. Part II: Development of 1901-96 Monthly Grids of Terrestrial Surface Climate; New et al (2000). A frost day is a period of 24 hours in which the minimum temperature falls below 0°C. If the temperature stays below zero all day, that's an 'ice day'.
pet	Potential Evapo-Transpiration (PET)	Millimetres (x10)	The method used is the FAO (Food and Agricultural Organization) grass reference evapotranspiration equation (Ekstrom et al., 2007, which is based on Allen et al., 1994). It is a variant of the Penman Monteith method using the gridded TMP, TMN, TMX, VAP and CLD. Note that PET values are mean mm/day for each month (with a scaling factor of 10 applied to the PET ascii (*.dat) files, but NOT the PET netcdf files (*.nc). The pet values in the datafiles therefore need to be multiplied by the number of days for each month to get the mean pet for that month.
pre	Precipitation	Millimetres (x10)	
tmp	Monthly average daily mean temperature	Degrees Celcius (x10)	The daily 'mean' temperature is the mid-point (median) between the daily minimum and maximum temperatures.
tmn	Monthly average daily minimum temperature	Degrees Celcius (x10)	
tmx	Monthly average daily maximum temperature	Degrees Celcius (x10)	
vap	Vapour pressure	Hecta-Pascals (x10)	
wet	Wet Day Frequency (rain days per month)	Days (x100)	

To read the CRU TS v4 ASCII data, users have so far been writing their own scripts as these are fairly easy to parse. If you would like to share your script to read the CRU TS ASCII data with other users, then please email CEDA Support. The ASCII data should be read using free-format.

To read the CRU TS v4 NetCDF data, you may use any NetCDF enabled software, such as Xconv, CDAT or FERRET, Matlab or Python). Please also see the [CRU data user guide](#) for more information.

How to read the CRU TS v4 Station data (metadata):

The CRU TS v4 station files are available from the CEDA Archive for each CRU TS version under the subdirectory named “station”.

The ‘.stn.’ files: the values in these represent, for each cell and timestep, the number of stations that could have influenced the data value for that cell and timestep. The sphere of influence is the Correlation Decay Distance, which is 450 km for precipitation, 750 km for diurnal temperature range, and 1200 km for mean temperature (New et al, 2000).

Station data files are available for the following variables: cld, dtr, frs, pre, tmn, tmp, tmx, vap and wet. All the station data files are compressed (.gz extension).

The NetCDF files in the data folder contain an additional variable, ‘stn’, which provides, for each datum in the main variable, a count (between 0 and 8) of the number of stations used in that interpolation.