CRU TS 3.00 - Data availability and File Formats explained

Where to find the CRU TS 3.00 data and metadata files:

The CRU TS 3.00 data (i.e. climate variables) and metadata files (i.e. stations) are available from the BADC Archive at:

http://badc.nerc.ac.uk/browse/badc/cru/data/cru_ts/cru_ts_3.00/data/

where the 'data' directory contains the CRU TS 3.00 monthly gridded variables.

All the data files (ASCII ".dat" and netcdf ".nc") are compressed (.gz extension).

CRU TS 3.00 Data and Metadata File Formats explained:

The CRU TS 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). 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, next 360 rows the data for Feb 1901, next 360 rows for March 1901 and so on. Missing values are stored as -999.

- **NetCDF data:** CRU3.00 - This is experimental, not all attributes are correctly set yet. The data themselves should be identical to those in the ASCII versions.

  NetCDF missing value issue. The headers of the netcdf tmp files state that the missing value should be -9999, however, the values stored are -999, which could be confused with actual values after the scale_factor is applied.

  If you have NCO, then something like this command works to fix files:

  ```
  ncatted -a missing_value,tmp,o,i,-999 -a _FillValue,tmp,o,i,-999 cru_ts_*.nc
  ```

  (thanks to Cathy Smith of NOAA for providing the fix)

  Please see the CEDA NetCDF pages at http://www.ceda.ac.uk/help/users-guide/file-formats/netcdf/ for more information.

How to read the CRU TS 3.00 data:

The CRU TS 3.00 data files contain:

Label	Variable	Units (Multiplying factor - ASCII only)	Comments
cld	Cloud Cover	percentage (x10)	This dataset is unchanged from CRU TS 2.10 for the period 1901-2002. The additional data is a combination of converted observations of sun hours, and synthetic cloud derived from its relationship with DTR. A NetCDF version of this dataset is not yet available. Please note that, due to a processing error, the data format changes in this file. For 1901-2002, it is (720i6), that is, each datum has six character spaces. After that point, it is (720i8), with each datum occupying eight spaces. If the data is read in 'free format', there will be no problems.
dtr	Diurnal Temperature Range	Degrees Celcius (x10)	
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'.
pre	Precipitation	Millimetres (x10)	
tmp	Daily mean temperature	Degrees Celcius (x10)	
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)	This dataset is constructed from observations, backed with synthetic data derived from precipitation. However, from 1990 onwards it is derived only from synthetic data, owing to problems with the observed database.

How to read the CRU TS 3.00 Station data (metadata):

The CRU TS 3.00 station files are available from the BADC Archive at:

http://badc.nerc.ac.uk/browse/badc/cru/data/cru_ts/cru_ts_3.00/station/

There are now two kinds of station files. Both types contain one value for every value in the data file:

- **Regular '.stn.' files**. The values in these represent, for each cell and timestep, the number of station 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).

- **Cell station '.cstn.' files**. These new files give the actual number of stations reporting in that cell at that timestep.

Station data files are available for the following variables: pre, dtr and tmp. There are no station data files available for cld, frs, tmn, tmx, vap or wet.

There is also an elevation file available in the data directory (halfdesg.elv.grid.data.gz)

All the station data files (ASCII ".dat" and netcdf ".nc") are compressed (.gz extension).

Reference

Mitchell and Jones, 2005 can be used as background information for TS 3.00. The major difference between the CRU TS 2.1 and CRU TS 3.00 processes is that no new homogenization is explicitly performed in the latter. Existing homogenizations in the underlying datasets, and homogenizations performed by national meteorological agencies prior to releasing their station data, are incorporated.