How to extract and handle CLM data from CERA

Diego Rybski[1], Markus Böttle[1], and Micha Steinhäuser[2]

[1] Potsdam Institute for Climate Impact Research (PIK)
BaltCICA Project: www.baltcica.org

[2] University of Oldenburg
Chameleon Project: www.climate-chameleon.de

April 14, 2010 (version 4)

Contents

1 Introduction

This tutorial is based on Unix/Linux. The download tool "jblob", Sec. 2, also runs under windows. This is probably not the case for "cdo", Sec. 3. In Sec. 4 you find some hints on how to read netcdf files directly in matlab. Frequently asked questions are tried to be answered in Sec. B.

2 Download your data

Use the download tool "jblob" to download data from CERA.

- It can be found here: http://cera-www.dkrz.de/CERA/jblob/
- Download data with the following command (in one line)
  `jblob -dataset DATASET -username USERNAME -password PASSWORD -dir DESTINATION -tmin BTIME -tmax ETIME`

– DATASET is the dataset you want to download, replace with the corresponding name, such as CLM_C20_1_D3_dm_T_2M_AV.
Go to
http://cera-www.dkrz.de/WDCC/ui/BrowseExperiments.jsp
and choose CLM_regional_climate_model_runs
(note, we agreed to use **data stream 3**!)

– USERNAME is your username, put your username

– PASSWORD is your password, put your password

– DESTINATION determines where the file is being saved, put your directory

– BTIME (optional) is the beginning of the time frame, such as 1971-01-01

– ETIME (optional) is the end of the time frame, such as 1980-12-31

In order to obtain USERNAME and PASSWORD you need to register at cera. BTIME and ETIME specify the time frame you want to download.

- Example (type in one line):
`jblob -dataset CLM_C20_1_D3_dm_T_2M_AV -username ??? -password ???`
`-dir /work/cclm/manual -tmin 1971-01-01 -tmax 1980-12-31`
You should get an output such as this:

```
jblob V1.4.22, 22/06/2009 (ceraget build 27/10/2006)
Dataset CLM_C20_1_D3_dm_T_2M_AV
rmin = 4019, rmax = 7671
Writing to /home/diego/work/cclm/manual/CLM_C20_1_D3_dm_T_2M_AV_4019-7671.nc
```

3 Cut and transform your data

Use "cdo" to cut and transform your netcdf files.

- It can be found here: http://www.mpimet.mpg.de/fileadmin/software/cdo/

See Sec. A.

3.1 Get Information

- Type
`cdo sinfo FILENAME`
to obtain short data set information, whereas FILENAME is your file, such as
CLM_C20_1_D3_dm_T_2M_AV_4019-7671.nc

- You should get something like this

```
File format: netCDF
-1 : Institut Source Table Code Time Typ Grid Size Num Levels Num
1 : unknown unknown 0 -1 variable F32 42126 1 1 1
Horizontal grids :
1 : lonlat > size : dim = 42126 nlon = 238 nlat = 177
lon : first = -10.6 last = 36.8 inc = 0.2 degrees_east
lat : first = 34.6 last = 69.8 inc = 0.2 degrees_north
Vertical grids :
1 : height m : 2
Time axis : 3653 steps
RefTime = 1955-01-01 00:00 Units = minutes Calendar = STANDARD
YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm
```

```
1971-01-01 23:59 1971-01-02 23:59 1971-01-03 23:59 1971-01-04 23:59
1971-01-05 23:59 1971-01-06 23:59 1971-01-07 23:59 1971-01-08 23:59
1971-01-09 23:59 1971-01-10 23:59 1971-01-11 23:59 1971-01-12 23:59
...
1980-12-23 23:59 1980-12-24 23:59 1980-12-25 23:59 1980-12-26 23:59
1980-12-27 23:59 1980-12-28 23:59 1980-12-29 23:59 1980-12-30 23:59
1980-12-31 23:59
```

3.2 Cut a window

- Type
  `cdo sellonlatbox,LON1,LON2,LAT1,LAT2 INAME ONAME`
  to cut an area defined by
  - `LON1` and `LON2` are the lower and upper longitude of your window
  - `LAT1` and `LAT2` the lower and upper latitiude of your window
  - `INAME` is your input filename
  - `OUTPUT` is your output filename
- Example (type in one line):
  `cdo sellonlatbox,-10.6,-10.5,34.6,34.7 CLM_C20_1_D3_dm_T_2M_AV_4019-7671.nc CLM_C20_1_D3_dm_T_2M_AV_4019-7671_sellonlatbox.nc`
  should return something like
  ```
  cdo sellonlatbox:  Processed 153886278 values from 1 variable over 3653 timesteps.  ( 3.28s )
  ```
- Check with the command:
  `cdo sinfo CLM_C20_1_D3_dm_T_2M_AV_4019-7671_sellonlatbox.nc | less`
  which should return
  ```
  File format:  netCDF
  -1 :  Institut Source Table Code Time Typ Grid Size Num Levels Num
  1 :  unknown unknown 0 -1 variable F32 1 1 1 1
  Horizontal grids :
  1 :  lonlat > size :  dim = 1 nlon = 1 nlat = 1
  lon :  first = 349.4 last = 349.4 degrees_east
  lat :  first = 34.6 last = 34.6 degrees_north
  Vertical grids :
  1 :  height m :  2
  Time axis :  3653 steps
  RefTime = 1955-01-01 00:00 Units = minutes Calendar = STANDARD
  YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm
  1971-01-01 23:59 1971-01-02 23:59 1971-01-03 23:59 1971-01-04 23:59
  1971-01-05 23:59 1971-01-06 23:59 1971-01-07 23:59 1971-01-08 23:59
  1971-01-09 23:59 1971-01-10 23:59 1971-01-11 23:59 1971-01-12 23:59
  ...
  ```

3.3 Cut a time window

- Type
  `cdo seldate,DATE INAME ONAME`
  to cut a time window, where
  - `DATE` is the date, such as `1971-01-01`

  - `INAME` is your input filename
  - `OUTPUT` is yout output filename

- Example (type in one line):
  `cdo seldate,1971-01-01 CLM_C20_1_D3_dm_T_2M_AV_4019-7671.nc CLM_C20_1_D3_dm_T_2M_AV_4019-7671_seldate.nc`
  should return something like
  ```
  cdo seldate:  Processed 42126 values from 1 variable over 3653 timesteps.  ( 0.25s )
  ```

- Check with the command:
  `cdo sinfo CLM_C20_1_D3_dm_T_2M_AV_4019-7671_seldate.nc`
  which should return
  ```
  File format:  netCDF
  -1 :  Institut Source Table Code Time Typ Grid Size Num Levels Num
  1 :  unknown unknown 0 -1 variable F32 42126 1 1 1
  Horizontal grids :
  1 :  lonlat > size :  dim = 42126 nlon = 238 nlat = 177
  lon :  first = -10.6 last = 36.8 inc = 0.2 degrees_east
  lat :  first = 34.6 last = 69.8 inc = 0.2 degrees_north
  Vertical grids :
  1 :  height m :  2
  Time axis :  1 step
  RefTime = 1955-01-01 00:00 Units = minutes Calendar = STANDARD
  YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm YYYY-MM-DD hh:mm
  1971-01-01 23:59
  ```

3.4 Transform to ascii

- Type
  `cdo output NNAME > ANAME`
  in order to transform your data to ascii, whereas
  - `NNAME` is the name of your netcdf file
  - `ANAME` is the name of your ascii file

- Examples (type in one line each):
  - `cdo output CLM_C20_1_D3_dm_T_2M_AV_4019-7671_sellonlatbox.nc > CLM_C20_1_D3_dm_T_2M_AV_4019-7671_sellonlatbox.ascii`
  - `cdo output CLM_C20_1_D3_dm_T_2M_AV_4019-7671_seldate.nc > CLM_C20_1_D3_dm_T_2M_AV_4019-7671_seldate.ascii`

If you followed the above steps, the first ascii-file should contain a time series of the grid point 349.4,34.6 and the second one a snapshot of the grid-points for the date 1971-01-01.

The values of the gridded data are arranged according to (lon,lat,time), i.e. first, all values for fixed latitude and time are run through, then the latitiude is incremented by one (while time is unchanged) and again all longitudes are run through, and so on.

Another useful command is "ncdump" which comes with netcdf.

4 Read netcdf with matlab

- First you need to get some matlab toolboxes, which are contained in the following zip file: https://sourceforge.net/projects/mexcdf/files/R2008b/r2784/mexcdf.m.r2784.zip/download
- After extraction, set a matlab path to the downloaded directories `mexnc` and `snctools` (it seems like, you don't need `netcdf_toolbox`, which is also contained in the zip file).
- Now you should be able to read `.nc` files in matlab via the command `nc_varget`.
- There is a nice and comprehensive tutorial available on http://mexcdf.sourceforge.net/tutorial

This works fine with matlab R2009a and should work with R2008b as well.

A Configure "cdo"

- install netcdf from the repository (e.g. libnetcdf-dev for debian based distributions). Now there should be the file netcdf.h on your system (e.g. in /usr/include/ for ubuntu)
- download cdo from http://www.mpimet.mpg.de/fileadmin/software/cdo/
- unpack the tar-file
- change to the directory of the unpacked file
- Type
  `./configure --prefix=/usr --with-netcdf=/usr/include` (This links cdo to the file netcdf.h! If necessary, change /usr/include to the location of netcdf.h)
  `make`
  `sudo make install`

B FAQ

1. With which GCM is CLM driven?
   ECHAM5-MPIOM
2. What is the difference between the CLM realizations?
   The different realizations of CLM (3 runs for the 20th Century, 2 runs each for A1B and B1) are based on different realizations of the GCM. The GCM realizations differ in the moment when the forcing (such as CO2 concentration) starts relative to the spin-up time. This way, one tries to take into account modes of internal variability, such es ENSO.