CARRA-based forcing files for PISM

July 5, 2023

1 General information on forcing files

The following is a short, definately not complete, list of general issues that must be taken into consideration when making a forcing file for PISM. Focus here is on forcings for runs with pdd schemes. For other run types, such as runs driven by surface mass balance data, it may be necessary with some modifications. Consult the PISM manual on the requirements for forcing files for different different setups.

1.1 Naming conventions

When making forcing files, it is important to give the various fields pism-recognisable names. PISM will search in the attributes of any given field for 'approved' short names or long names. For instance, the short names for bedrock topography must be 'topg' and ice thickness 'thk', respectively. After selecting the type of forcing for a given run, consult the PISM manual for the chosen forcing type and its options and note the variable names listed in the manual under the header 'variables'. Make sure that these variable names are the ones used in the forcing file.

1.2 Choice of units

Similar to the choice of variable names, PISM has strict expectations in terms of the units of the input fields. These may be found in the documentation concerning the various types of forcing as well.

1.2.1 Precipitation

In the current setup of pdd-type forcings, precipitation data is needed. The variable name should be 'precipitation' and the unit $\frac{\mathrm{kg}}{\mathrm{m}^2 \cdot \mathrm{s}}$.

1.2.2 Surface elevation

In order for PISM to be able to perform lapse rate corrections to the precipitation field, the surface elevation on which the precipitation field is calculated should be included, so PISM may correct the precipitation field depending on the difference in PISM's surface topography and the surface elevation of the model which made the precipitation field. This field must be called 'surface_altitude' and have units of 'm'.

1.2.3 Temperature

In the pdd setup, the temperature field should be called 'air_temp' and have the unit of degrees Kelvin (K).

1.2.4 Time

The time variable should be called 'time' and is given as the number of seconds since an initial cut-off date. This can either be a real date as in the case of the CARRA date: 'seconds since 1-1-1991 00:00:00' or some imaginary 'dummy date': 'seconds since 1-1-0000 00:00:00'.

1.3 Time bounds

Along with the time variable itself, PISM needs a variable of time bounds, called 'time_bnds'. PISM interprets temperature data as constant for the duration of the time interval between two time steps and needs the time bounds to state the start and end time of each time interval. The time bounds can be constructed by making a two-dimensional array of the time variable, the first column indicating the start point of the time intervals, the second the end points. There are snippets of scripts available in the PISM documentation in the section 'Adding time bounds', which do this.

2 CARRA input

The input files for the current setup are not raw CARRA output files, rather files that have already been processed by Jason Box. The files contain total precipitation, 2m temperature and the standard deviation of the 2m temperature. Two sets of files are currently available: A set of multi-year monthly means for the period 1990-2021, indicative of an average yearly cycle and a set of monthly data, with the full 1990-2021 time series. Currently, only the multi-year monthly mean data sets contain 2m temperature standard deviation data.

2.1 Location of data

The CARRA data is located here:

/mnt/data/syhsv/Data/PISMIn/CARRA

In this directory the folder *GridsAndMasks* contains basic files with grid, elevation and mask information. The folder *MultiAnnualMeans* contains the multi-year monthly mean time series and the folder *MonthlyMeans* holds the full time series.

3 Processing scripts

Producing PISM-readable forcing files from the CARRA data are done in two steps; first, the CARRA data is unpacked and the individual files are concatenated to make a single time series of the available data, see sec.(3.1). Next, the CARRA time series is interpolated onto the PISM computational grid, given PISM-readable names and appropriate units after which bedrock topography and geothermal heat flux fields are added, see sec.(3.2).

3.1 Unpacking CARRA files

The CARRA files from Jason Box are saved as either .npy or .npz files. The scripts

/home/syhsv/Projects/PISMForcings/CARRA/read_CARRA_loop.py

/home/syhsv/Projects/PISMForcings/CARRA/read_CARRA_loop_multiyear.py

reads the CARRA files and produces a NetCDF file with a time series of the temperature and precipitation data (along with temperature standard deviation if it is available). *read_CARRA_loop.py* processes the multi-year monthly means and *read_CARRA_loop_multiyear.py* reads the monthly mean files.

3.2 Making forcing files for PISM

After unpacking the .npy and .npz files in the script described in sec.(3.1), the data is used to produce forcing files for PISM. This is done in the scripts

/home/syhsv/Projects/PISMForcings/CARRA/preprocess_CARRA_mean.sh

/home/syhsv/Projects/PISMForcings/CARRA/preprocess_CARRA_multiyear.sh

The script *preprocess_CARRA_mean.sh* makes forcing files based on the multi-year monthly means and *preprocess_CARRA_multiyear.sh* makes forcing files based on the full time series of monthly means.
In the scripts, the various input fields are interpolated onto the PISM grid and given PISM-readable names and units. First, the ice thickness, bed topography and mask information of the Bedmachine (v5) data set are interpolated. These interpolations (as well as ensuing interpoaltions of the geothermal heat flux fields) are done with gdalwarp using bilinear interpolation. However, using gdalwarp for the CARRA data proved cumbersome since gdalwarp transforms each timestep in the file as a separate band, calling for quite a bit of tedious renaming and concatenation after interpolating. Instead, the CARRA data is interpolated using **cdo's remapbil** routine, which performs a bilinear transformation. In order for this routine to work, grid description files for both the source and target grids are needed. The cdo routine **cdo griddes** makes grid desription

files based on the metadata and coordinate info inthe files. The script produces one such file based on a raw CARRA data file with the CARRA land-sea mask using **cdo griddes**. Note that this did not work when using the NetCDF version of the CARRA data! Somehow, the CARRA postprocessing conversion of the original .grb files to NetCDF format corrupts something which makes it impossible to use the NetCDF version and **cdo griddes** results in an error message. However, using the .grb version of the land-sea mask file works without any trouble.
The interpolated bedrock and climate data are saved in a base file, whereupon a loop in number of available geothermal heat data sets interpolates the various geothermal heat flux fields and adds this field to the base file, thereby producing a separate forcing file for each geothermal heat flux field.
Finally, a mean geothermal heat flux field is produced by using the cdo routine **cdo ensmean** to produce an ensemble mean. This is then used to produce a forcing file with the mean geothermal heat flux.
Strictly speaking, the forcing files need not contain the topography and geothermal heat flux files. These could be kept in a separate file (initfile) and the climate forcings (forcingfile) in a different file. In a PISM run script, one would then use different file names (initfile and forcingfile, respectively) for the specification of initial file and forcing file.

3.3 Location of CARRA forcing files for PISM

The forcing files are located in

/mnt/data/syhsv/Data/PISMIn/CARRA/PISMForcings

In this directory, two types of files may be found:
pism_CARRA_30yMean_BHF_*NAME*.nc
pism_CARRA_Monthly_BHF_*NAME*.nc

where *NAME* is the name of the geothermal heat flux data set. The first file is the multi-year monthly mean, the second file the full, monthly time series.