Regridding E3SM Data

Release 1.0

Sterling Baldwin Jill Chengzhu Zhang
Renata McCoy Tony Bartoletti Charlie Zender

Jun 04, 2020

CONTENTS

CHAPTER

ONE

REGRIDDING GUIDE

This guide will cover regridding native E3SM model output from each of its components native grids, to any regular grid. These regridding examples are specific for the E3SMv1.x model, and may not work with later versions. For detailed up-to-date information on the E3SM model activities, refer to the E3SM organizational website.

1.1 Environment Setup

You will need to first create an anaconda environment with the depedencies and install the netCDF Operators (NCO). Follow this link for the full set of NCO documentation. Although there are several ways to intall nco, the recommended method is to use the conda package manager.

1.1.1 Conda

```
conda install -c conda-forge nco
```

To create a new conda environment named "regrid" with just nco use:

```
conda create --name regrid -c conda-forge nco
```

1.1.2 Pre-build Executible

Pre-build executibles are available on a variety of platforms.

1.2 Map files

All of these methods assume you have a mapping file on hand to convert between the relevant grids. This is a partial list of mapping files for commonly used E3SM resolutions.

- High resolution MPAS ocean (18km to 6km) to 1/4 degree lon lat
- High resolution EAM/CAM atmospheric data (ne120) to 1/4 degree lon lat
- Standard resolution MPAS ocean (60km to 30km) to 1x1 degree lon lat
- Standard resolution EAM/CAM atmospheric data (ne30) to 1x1 degree lon lat
- Standard resolution ELM/CLM atmospheric data to 1x1 degree lon lat

A set of commonly used mapfiles can be found here. The mapfiles follow the naming convention of "map_<source-grid>_to_<destination-grid>_<regridding-algorithm>". This allows you to mix-and-match whichever raw grid was used to whichever destination grid is desired. The "ne30" raw grid corresponds to roughly 1 by 1 degree longitude/latitude, and is the "standard" resolution for the E3SM model. Higher resolutions are ne120 which corresponds to roughly 0.25 degree by 0.25 degree lon/lat. The two most commond regridding algorithns are bi-linear ("bilin"/"blin") and Area Averaged ("aave"). For bi-linear each of the grid points in the dest grid contains the average of the nearest four grid points in the source grid, weighted by their distance from the destination grid point. If any of the four surrounding input grid points contain missing data, the interpolated value will be flagged as missing. Bilin is a good default regridding method. The aave method is the area average with latitude weighting, and gives better results when going from a high resolution grid to a lower resolution. A spatial average of the source data is calculated in the area outlined by each grid box in the destination grid. If a grid box in the source grid partially overlaps the area of a destination grid cell, its contribution to the spatial average is weighted by the fraction of area within the destintion grid cell domain. See here for additional information.

1.3 Data files

1.3.1 EAM/CAM

Raw E3SM model output can be obtained from ESGF here. The ESGF search interface allows faceted search of the published E3SM model data. On the left side of the interface, each of the E3SM search facets is represented by a drop down menu. For example, the "Campaign" facet allows for the selection of data from a specific simulation campaign. Selecting the "DECK-v1" option and pressing "search" will narrow down the datasets displayed to only those experiments that participated in the DECK experimental campaign. Similarly, the "experiment" facet can be used to display all datasets belonging to a single simulation. The facets can be added together in an AND relationship, for example selecting the experiment=1pctCO2 AND realm=atmos will display all datasets that have both belong to the 1pctCO2 experiment, and are atmospheric data files.

The atmospheric files take the form of <experiment_case_name>.cam.h0.<year>-<month>.nc, for example 20180215.DECKv1b_1pctCO2.ne30_oEC.edison.cam.h0.0001-01.nc. The E3SM atmospheric model is a heavily modified decendent of the UCAR Community Atmospheric Model, and so include "cam" in their name. Note that future versions of the E3SM model will include a name change to the "E3SM Atmosphere Model" or eam. The EAM/CAM data files come in 4 temporal frequencies, monthly (cam.h0), daily (cam.h1), 6 hourly (cam.h3), and 3hr (cam.h4).

1.3.2 ELM/CLM

Similarly, the E3SM Land Model decends from the Community Land Model, and so contain the substring "clm2". Future versions of the E3SM output will replace this with "elm" for the E3SM Land Model. All land model data are monthly averages.

1.3.3 MPAS

The Ocean/Ice components of E3SM come from the Model for Prediction Across Scales (MPAS), a modern state-of-the-art ocean model developed specifically for the E3SM project. All MPAS data are monthly averages.

CHAPTER

TWO

ATMOSPHERIC REGRIDDING

2.1 Regridding EAM/CAM atmospheric data files

2.1.1 Low example

This example will regrid all files in a directory of ne30 cam.h0 files. This assumes that the mapfile is already in place in the current directory and that a set of cam.h0 files are inplace under a directory named "cam.h0". All the variables inside the file will be regridded.

```
#!/bin/bash
mapfile=map_ne30np4_to_cmip6_180x360_aave.20181001.nc
drc_in=./cam.h0/
drc_out=./180x360/

ncremap -m ${mapfile} -I ${src_in} -O ${drc_out}
```

2.1.2 Medium Complexity

In this example, the same regridding operation as above is performed, but only on the PRECT and FSDS variables. Running with specific variables selected significantly speeds up the regridding process, as the rest of the variables are ignored. Output file size will also be reduced, as only the selected variables will be present.

```
#!/bin/bash
mapfile=map_ne30np4_to_cmip6_180x360_aave.20181001.nc
drc_in=./cam.h0/
drc_out=./180x360/
vars="PRECT,FSDS"

ncremap -v ${vars} -m ${mapfile} -I ${src_in} -O ${drc_out}
```

2.1.3 High Complexity

In this example, the PRECT and FSDS variables will be regridded and extracted into single-variable-per-file time-series files. The output will be compressed and deflated. The model data starts at 1850-01 and ends at 2014-12. For high temporal frequency files (i.e. higher frequency then monthly), add the *–clm_md=hgh_frq* argument to the ncclimo command.

```
#!/bin/bash

drc_in=./cam.h0/
drc_out=./180x360/
mapfile=map_ne30np4_to_cmip6_180x360_aave.20181001.nc
vars="PRECT,FSDS"              # comma separated variable list
start=1850                     # the first year of data
end=2014                       # the last year of data
flags="-7 --dfl_lvl=1"  # compression and deflation flags

cd ${drc_in}
ls | ncclimo \
  ${flags} \
  --var=${vars} \
  --yr_srt=${start} \
  --yr_end=${end} \
  -O=${dir_out} \
  --map=${mapfile}
```

CHAPTER

THREE

LAND REGRIDDING

3.1 Regridding ELM/CLM land data files

3.1.1 Low Complexity

In this example, a directory of ne30 clm2.h0 files will be regridded. This will regrid all variables in the input files, using sub-grid-scale regridding, which uses the land fraction around coastal areas to better represent the values around complex coastal geometry.

```
#!/bin/bash
mapfile=map_ne30np4_to_cmip6_180x360_aave.20181001.nc
drc_in=./clm2.h0/
drc_out=./180x360/
land_file=<input file path> # this is the path to a single land file to pull the␣
↪landfrac variable from

ncremap -m ${mapfile} -I ${src_in} -O ${drc_out} --sgs_frac=${land_file}/landfrac
```

3.1.2 Medium Complexity

In this example, the same regridding operation as above is performed, but only on the DEADSTEMC and CDWC variables. Running with specific variables significantly speeds up the run and reduces output files size, as the rest of the variables are ignored.

```
#!/bin/bash
mapfile=map_ne30np4_to_cmip6_180x360_aave.20181001.nc
drc_in=./clm2.h0/
drc_out=./180x360/
vars="DEADSTEMC,CDWC"             # the variables must be in a comma sepparated␣
↪list with no spaces
land_file=<input file path>        # this is the path to a single land file to pull␣
↪the landfrac variable from

ncremap -v ${vars} -m ${mapfile} -I ${src_in} -O ${drc_out} --sgs_frac=${land_file}/
↪landfrac
```

3.1.3 High Complexity

In example, the regridding operation is performed and the selected variables are output in single-variable-per-file time-series files. The output will be compressed and deflated. The model data starts at 1850-01 and ends at 2014-12.

```bash
#!/bin/bash
mapfile=map_ne30np4_to_cmip6_180x360_aave.20181001.nc
drc_in=./clm2.h0/
drc_out=./180x360/
vars="DEADSTEMC,CDWC"               # the variables must be in a comma sepparated list with no spaces
start=1850                          # the first year of model data
end=2014                            # the last year of model data
land_file=<input file path>         # this is the path to a single land file to pull the landfrac variable from
flags="-7 --dfl_lvl=1"       # compression and deflation

cd ${drc_in}
ls | ncclimo \
  ${flags} \
  --var=${vars} \
  --yr_srt=${start} \
  --yr_end=${end} \
  -O=${dir_out} \
  --map=${mapfile} \
  --sgs_frac=${land_file}/landfrac
```

CHAPTER

FOUR

MPAS REGRIDDING

4.1 Regridding MPAS ocean/sea-ice data files

4.1.1 Low Complexity

In this example, a directory of mpaso.hist.am.timeSeriesStatsMonthly files will be regridded. For mpassi files, add the flag "–sgs_frc=timeMonthly_avg_iceAreaCell" to turn on sub-grid-cell regridding.

```
#!/bin/bash
mapfile=map_oEC60to30v3_to_cmip6_180x360_aave.20181001.nc # map from the MPAS 60km-to-30km mesh to the 1x1 degree grid
drc_in=./mpaso/
drc_out=./180x360/
flags="-P mpas --d2f"  # This invokes the mpas regridder, and converts output from double precision to single

ncremap -m ${mapfile} -I ${drc_in} -O ${drc_out} {flags}
```

4.1.2 High Complexity

In this example, two variables will be extracted into single-variable-per-file time-series files. These will be compressed and deflated, and the data will be converted from double to single precision. For mpas sea-ice files, add the flag "–sgs_frc=timeMonthly_avg_iceAreaCell" to turn on sub-grid-cell regridding.

```
#!/bin/bash
mapfile=map_oEC60to30v3_to_cmip6_180x360_aave.20181001.nc # map from the MPAS 60km-to-30km mesh to the 1x1 degree grid
drc_in=./mpaso/
drc_out=./180x360/
vars="timeMonthly_avg_seaSurfaceSalinity,timeMonthly_avg_seaSurfaceTemperature"
start=1850                                   # first year of model data
end=2014                                     # last year of model data
flags="-7 --dfl_lvl=1 -m mpas --d2f"         # compression level 7, deflate level 1, using ncclimo mode=mpas, converting from double to float

cd ${drc_in}
ls | ncclimo \
  ${flags} \
  --var=${vars} \
  --yr_srt=${start} \
  --yr_end=${end} \
```

(continues on next page)

(continued from previous page)

```
-O=${dir_out} \
--map=${mapfile}
```