

Regridding meteorological data

There are two options for nudging available in CESM, generally referred to as 'specified dynamics' and 'physics-side nudging'. To use physics-side nudging in your simulation, see the CAM User's Guide, [Section 8.6](#). All configuration is set in the namelist (user_nl_cam), and numerous options are available. CESM2.2 has a compset, FCnudged, with a standard configuration (except, it is recommended to use `model_times_per_day = 48`, instead of the default `model_times_per_day = 24`).

For physics-side nudging, available reanalysis data is pre-processed prior to model use. This preparation step interpolates the desired reanalysis data onto a given model grid, adjusts values to account for topographical differences between the reanalysis model and the current topography of CESM, and saves the results in individual files with one time slice per file. The scripts that carry out this processing must be configured for each combination of input reanalysis dataset and output model grid. The scripts are currently available for the FV and SE dycore grids.

The details below are for regridding GEOS and MERRA2 files, particularly on the NCAR HPC. Users wishing to use alternate reanalysis datasets should consult the README.

The processing scripts are slightly different for the two available dycores, so they are contained in separate directories in the IPT repository https://github.com/NCAR/IPT/tree/master/Meteorological_Reanalysis_Data. In each directory, there is a README that provides general guidance for setting the script variables for a desired reanalysis product.

```
Finite_volume_dycore/
  README
  Gen_Data_f09/
Spectral_element_dycore/
  README
  Gen_Data_ne30/
  Gen_Data_ne30np3/
  Gen_Data_ne0CONUSne30x8/
  Gen_Data_NEWGRID/
```

For the FV dycore, the directory `Gen_Data_f09` has processing scripts set up for a 1 degree grid. Within this directory, there are two csh scripts, one configured for processing MERRA2 data and the other for processing ERA-Interim data. If a user desires data from either of these reanalysis products, then all that is needed is to adjust the dates for processing, the paths for input/output files, and any changes in vertical grid or topography values. The NCL scripts are pre-configured for a number of available reanalysis products (see `makeIC_extract_analysis_info.ncl`), so converting to one of those will just involve editing a csh script to set the paths and processing options consistent with the product. See the README for guidance. If, on the other hand, a user has a newer or other reanalysis product, then a template must be added in the NCL programs to define the structure of the input dataset. For this case, the user should use a known and familiar existing template in the NCL programs as a guide.

In the `Spectral_element_dycore` directory, there are sub-directories containing scripts that are configured for a particular SE grid. The `Gen_Data_ne30` directory has scripts set up for a uniform 1 degree spectral element grid, `Gen_Data_ne30pg3` is set up for a physics grid, and `Gen_Data_ne0CONUSne30x8` is configured for the variable resolution CONUS grid. In each of these there are `Gen_*.csh` scripts set up for MERRA2 and other reanalysis products. As for the FV dycore, the README provides some guidance in making modifications to these scripts. There is also a directory that can be used as a template when tailoring the scripts for a user defined VR grid.

One substantial difference with the FV processing is in the horizontal interpolation. When the scripts for the SE grids were being developed, there were significant errors in the ESMF processing at the poles and for wrap around gridpoints. Complicating the development, was the fact that each update to the ESMF routines within NCL would result in failures in the processing scripts. For this reason, a version of the ESMF processing routines that contain fixes for these errors is included with the reanalysis processing scripts in the `Gen_Data` directory (`ESMF_regridding.ncl`). From the user stand point, this means that for each horizontal SE grid, the user must edit `makeIC_se_002.ncl` and set a hardcoded path to the SCRIP file for that grid. Note that the ESMF processing routines have been substantially improved since this time, but it is not known if the pole problems have been resolved. So a user may wish to revise the scripts to eliminate the ESMF code and use the up to date ESMF routines in NCL instead. The plan moving forward is for this processing step to be carried out at run time directly from the reanalysis archives (making these scripts unnecessary), so for the time being these scripts will continue as is without further development.

WRAPIT

The NCL scripts depend on FORTRAN subroutines which must be pre-compiled to create a shared library (`MAKEIC.so`) using the `WRAPIT` command. On casper (July 2024), `WRAPIT` is available in `/glade/u/apps/opt/ncl/6.6.2/bin/WRAPIT` after running: `'module load ncl'`.

Running multiple streams of processing

The other problem is for users that need to process a large amount of data. The processing scripts can take quite a long time to run, so to speed up the process a `RUNNUM` variable was added to the script so that multiple copies can be run at the same time. Since the processing is in a common directory, the `WRAPIT` command from one instance clobbers the shared library used by all. This would result in total failure. Users who need to run in this manner must comment out the `WRAPIT` command in the script and run it to create the library prior to submitting the processing scripts.

Example: Modifying the Scripts For a Newly Created Grid to run on CASPER (Jan 2024)

As an example of the steps needed to process data for a new variable resolution SE grid, suppose you have a new grid **ne0np4.SAM01.ne30x4** for a uniform ne30 SE resolution with a regional refinement to ne120 over South America and you plan to use MERRA2 reanalysis data. Begin by making a copy of the NEWGRID directory for your new grid:

```
casper% cp -r Gen_Data_NEWGRID Gen_Data_ne0np4.SAM01.ne30x4

casper% cd Gen_Data_ne0np4.SAM01.ne30x4

casper% mv Gen_MERRA2_NEWGRID.csh Gen_MERRA2_ne0np4.SAM01.ne30x4.csh
```

1. Edit your new script, and set the PBS commands:

```
#!/bin/csh
#PBS -N Gen_MERRA2_ne0np4.SAM01.ne30x4
#PBS -A Pxxxxxxx
#PBS -l select=2:ncpus=4:mpiprocs=4:mem=20GB
#PBS -l walltime=12:00:00
#PBS -q casper
#PBS -o Log.Gen_MERRA2_SAM01.err
#PBS -e Log.Gen_MERRA2_SAM01.out

module load ncl
```

2. In the Configuration section, **set the reference date** corresponding to the first day of data you desire, then number of days of data to process from that date, and the path where you wish to have the data stored:

```
#=====

# CONFIGURATION SECTION:

#=====

# Set a REFERENCE (Starting) Date and the number of days to process

#-----

set RUNNUM=01

set REF_DATE='20121201'

set NUM_DAYS=400


# Set INPUT/OUTPUT/TMP directories

#-----

set NAMELIST='./Config/Config_makeIC-'$RUNNUM'.nl'

set MYLOGDIR='./LOG/LOG_002.'$RUNNUM'/'

set MYTMPDIR='./TMP/TMP_002.'$RUNNUM'/'

set MYOUTDIR='/path/to/my/repo/ne0np4.SAM01.ne30x4/nudging/MERRA2_ne0np4.SAM01.ne30x4_L32/'

set INPUTDIR='/glade/collections/rda/data/ds313.3/orig_res/'


# Set ESMF options

#-----

set ESMF_interp='conserve'

set ESMF_pole='none'

set ESMF_clean='False'

set ESMF_clean='True'
```

3. For the processing options, **set the CASE name**. This is the root filename for your nudging data files. Note that some reanalysis datasets store winds in the form of vorticity and divergence values rather than U,V. It is important that the VORT_DIV_TO_UV flag is set to True for these datasets, this is a common source of processing errors. Finally, set the fname_grid_info value to point to a file containing the desired output grid, and set fname_phis_output to point to a file containing the model topography.

```
# Set Processing options
#-----

set CASE = 'MERRA2_ne0np4.SAM01.ne30x4_L32'

set DYCORE = 'se'

set PRECISION = 'float'

set VORT_DIV_TO_UV = 'False'

set SST_MASK = 'False'

set ICE_MASK = 'False'

set OUTPUT_PHIS = 'True'

set REGRID_ALL = 'False'

set ADJUST_STATE_FROM_TOPO = 'True'

set MASS_FIX = 'True'


# Set files containing OUTPUT Grid structure and topography
#-----

set fname_grid_info = '/path/to/my/repo/ne0np4.SAM01.ne30x4/inic/cami-mam4_0000-01-01_ne0np4.SAM01.
ne30x4_L32_c200309.nc'

set fname_phis_output = '/path/to/my/repo/ne0np4.SAM01.ne30x4/topo/topo_ne30np4.SAM01.ne30x4_blin_200309.nc'

set ftype_phis_output = 'SE_TOPOGRAPHY'
```

4. Update the path to WRAPIT: /glade/u/apps/opt/ncl/6.6.2/bin/WRAPIT MAKEIC.stub MAKEIC.f90

5. Edit the file **makeIC_se_002.ncl**. At about line 430, the dstGridName variable has to be set to use the SCRIP file for your new grid:

```
dstGridName="/path/to/my/repo/ne0np4.SAM01.ne30x4/grids/SAM01_ne30x4_SCRIP.nc"
```

After making the changes above you can submit this job on casper:

casper% qsub Gen_MERRA2_ne0np4.SAM01.ne30x4.csh

With these changes, the script can be submitted to generate the desired 400 days of data. The user must always check the Log and output file to verify that the dataset was processed correctly. Selecting some files and browsing variables for valid values using ncdump can save a lot of time and effort if there were problems during the processing.

A script for regridding to FV on casper (functioning as of 7/19/2024) is: /glade/u/home/emmons/my_IPT/Meteorological_Reanalysis_Data/Finite_volume_dycore/Gen_Data_f09_83L/Gen_MERRA2_fv09_83L_001.csh