Comparing GCM and RCM output

Cyberinfrastructure: Climate Modelling
Feb 16, 2011

In this exercise we'll compare the output from global and regional climate models. We've selected the GFDL model as an example for today.

Here are a few relevant links:

- Overview of the GFDL model
- NARCCAP website

Objective

In this session you will work with GCM and RCM data. First, we'll download monthly data from the GFDL GCM output using the IRI data library that you learned about last week. Then we download and process *daily* RCM output from NARCCAP and compare them.

Instructions

1. Download mean annual temperature and precipitation from the INGRID site:
    a. From the IRI/LDEO website: http://iridl.ldeo.columbia.edu/, click on the "expert" button on the left-hand side of the screen.
    b. For *future* monthly data for precip and temperature copy and paste the following script for the GFDL CM2.0 sresA2 run (2038-2070) into the expert box and click "OK":

```
 expert
SOURCES .WCRP .CMIP3
 (ipcc4/sresa2) @@
 (ipcc4/sresa2/gfdl_cm2_0) @@
 .pcmdi.ipcc4.gfdl_cm2_0.sresa2.run1.atm.mo.xml .tas
 time (2038) (2070) RANGEEDGES
 lat (40N) (49N) RANGEEDGES
 lon (77W) (65W) RANGEEDGES
 SOURCES .WCRP .CMIP3
  (ipcc4/sresa2) @@
  (ipcc4/sresa2/gfdl_cm2_0) @@
  .pcmdi.ipcc4.gfdl_cm2_0.sresa2.run1.atm.mo.xml .pr
  c: 0.001 (m3 kg-1) :c
     mul
  c: 1000 (mm m-1) :c
     mul
```

```
c: 86400 (s day-1) :c
   mul
time (2038) (2070) RANGEEDGES
lat (40N) (49N) RANGEEDGES
lon (77W) (65W) RANGEEDGES
{tmean ptot}ds
```

c. Do you understand what's happing there? We're asking for two sources (temp -tas- and precip -pr-) and subsetting both to a region and time like we did last week. In addition, the lines that look like " **`c:0.001(m3 kg-1):c`**" convert the units from kg m-2 day-1 to mm/day (a more familiar unit). The new units (mm/day) will be documented in the data you download (nice!).
d. Next click on the red "**Data Downloads & Files**" button next to the views. Then, scroll down and download the link that says "**netCDF file**." Save the file with the name: **`GFDL_Future.nc`**
e. For the *past* monthly data for precip and temperature copy and paste the following script for the GFDL CM2.0 20c run (1968-2000) into the expert box and click "OK":

```
expert
SOURCES .WCRP .CMIP3
 (ipcc4/20c3m) @@
 (ipcc4/20c3m/gfdl_cm2_0) @@
 .pcmdi.ipcc4.gfdl_cm2_0.20c3m.run2.atm.mo.xml .tas
 time (1968) (2000) RANGEEDGES
 lat (40N) (49N) RANGEEDGES
 lon (77W) (65W) RANGEEDGES
 SOURCES .WCRP .CMIP3
  (ipcc4/20c3m) @@
  (ipcc4/20c3m/gfdl_cm2_0) @@
  .pcmdi.ipcc4.gfdl_cm2_0.20c3m.run2.atm.mo.xml .pr
     c: 0.001 (m3 kg-1) :c
           mul
     c: 1000 (mm m-1) :c
           mul
     c: 86400 (s day-1) :c
           mul
  time (1968) (2000) RANGEEDGES
  lat (40N) (49N) RANGEEDGES
  lon (77W) (65W) RANGEEDGES
  {tmean ptot}ds
```

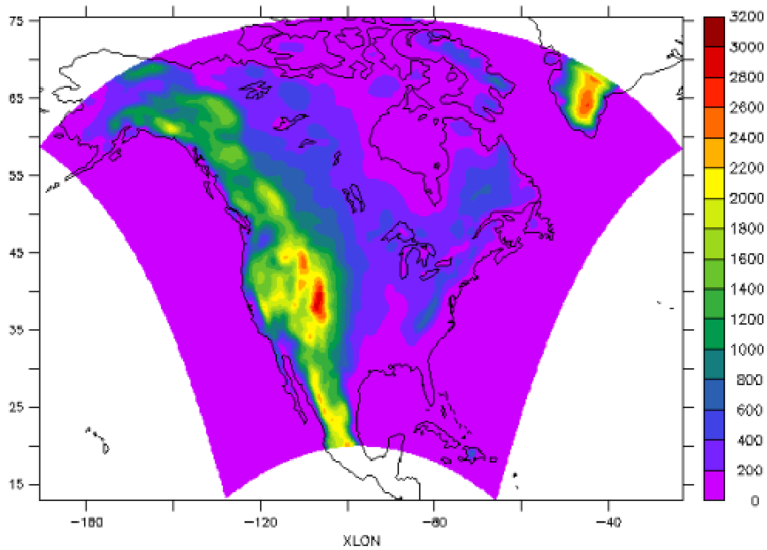
f. Next click on the red "**Data Downloads & Files**" button next to the views. Then, scroll down and download the link that says "**netCDF file**." Save the file with the name: **`GFDL_Current.nc`**
g. OPTIONAL: You could also calculate the different (future-current) directly on the IRI webpage if you wanted, but we're going to do it later using a new tool explained below. But, for the curious, here is how you would do it (you don't actually have to run this, we'll use the data you already downloaded for the rest

of the exercise). For the future minus present data (in a single step) repeat step-a, and in the “expert” box, copy and paste the following script for the GFDL CM2.0 sresA2 run (yrs 2038-2070) minus the 20c run (yrs 1968-2000) monthly mean precipitation. Again, this is optional, you don’t need to run this for the rest of the exercise.

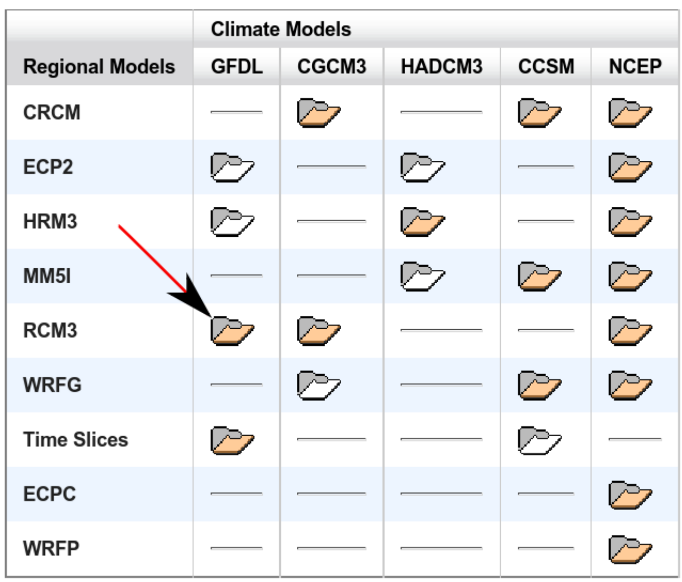
```
expert
SOURCES .WCRP .CMIP3
 (ipcc4/sresa2) @@
 (ipcc4/sresa2/gfdl_cm2_0) @@
 .pcmdi.ipcc4.gfdl_cm2_0.sresa2.run1.atm.mo.xml .pr
 time 0.0 monthlyAverage
      c: 0.001 (m3 kg-1) :c
            mul
      c:    1000 (mm m-1) :c
            mul
      c:    86400 (s day-1) :c
            mul
 time (Jan 2038) (Dec 2070) RANGE
 lon (85W) (65W) RANGEEDGES
 lat (30N) (50N) RANGEEDGES

 SOURCES .WCRP .CMIP3
 (ipcc4/20c3m) @@
 (ipcc4/20c3m/gfdl_cm2_0) @@
 .pcmdi.ipcc4.gfdl_cm2_0.20c3m.run1.atm.mo.xml .pr

 time 0.0 monthlyAverage

 time (Jan 1968) (Dec 2000) RANGE
 lon (85W) (65W) RANGEEDGES
 lat (30N) (50N) RANGEEDGES
 time 12 splitstreamgrid
 [time2]average
 sub
```

i. Then hit “ok”. Next click on the red “Data Files” button next to the views. Then, at the bottom of the page click on red “netCDF” .
ii. Run the script again and download a netCDF for temperature change but replace “.pr” with “.tas”

2. Now we’ll work with output from a regional climate model (RCM) that was *forced* by the same GCM. We’ll get these data from the North American Regional Climate Change Assessment Program (http://www.narccap.ucar.edu/). Explore the site a little to see what they offer. Here’s a figure of their “region of interest.”

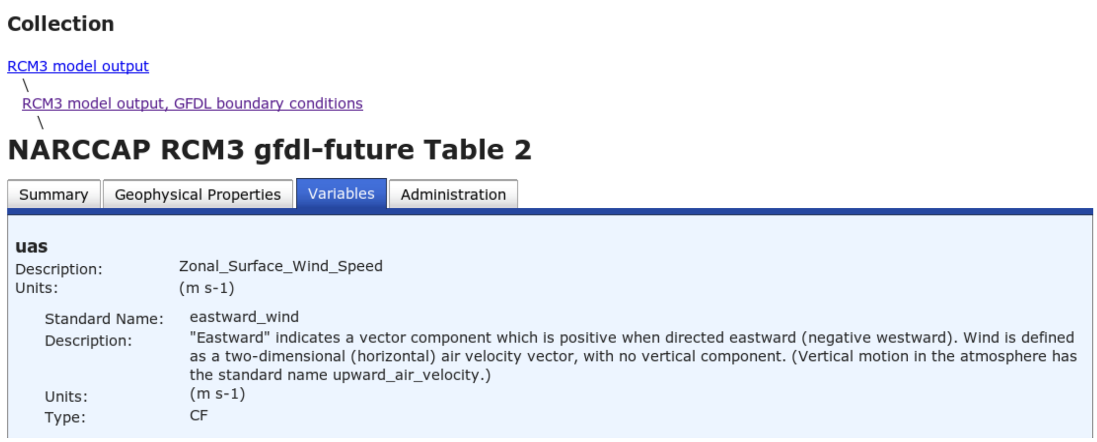


a. I've already downloaded the data from NARCCAP you need for this exercise and put it on the website in the zipped "Class3_Data" file on the website, but it will be good to explore the site and see how the data are organized. When you are ready to look at data, click on the "output data catalog" in the left column.
b. Choose the folder of data from the RCM3 regional climate model embedded in the GFDL global climate model.



	Climate Models				
Regional Models	GFDL	CGCM3	HADCM3	CCSM	NCEP
CRCM	—		—		
ECP2		—		—	
HRM3		—		—	
MM5I	—	—			
RCM3			—	—	
WRFG	—		—		
Time Slices		—	—		—
ECPC	—	—	—	—	
WRFP	—	—	—	—	

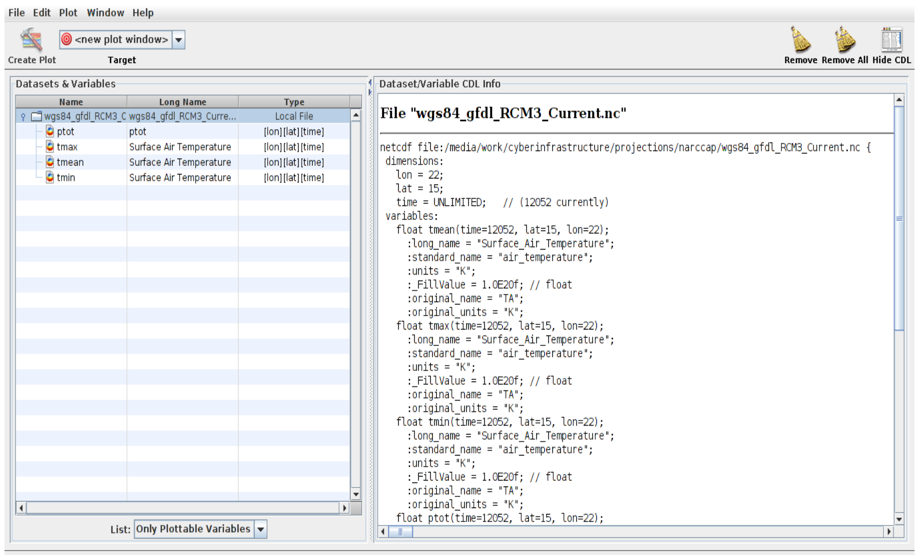
c. Select [NARCCAP RCM3 gfdl-future Table 2]
d. Explore the "Variables" tab to see what variables are available:



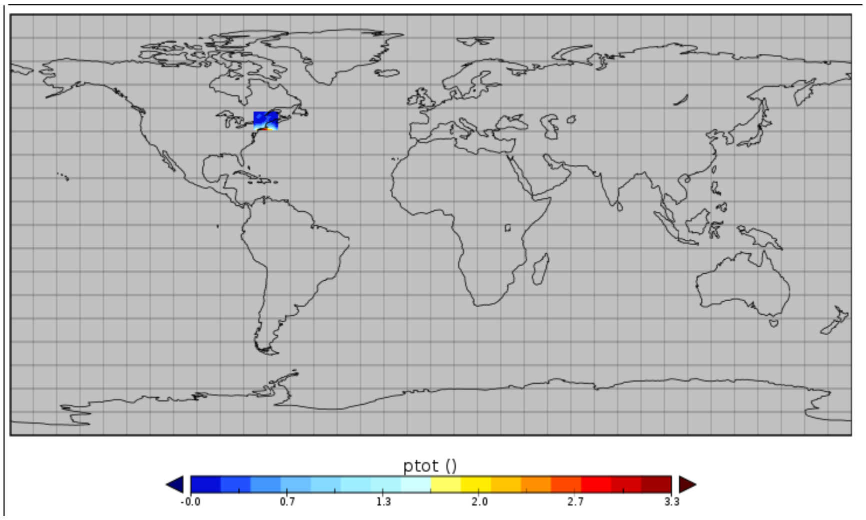
e. On the "Summary" tab, click: "Download files for this collection"
f. login with (They know I'm sharing my password with the class and I'll change it after, so you will have to make your own account if you want to work with this site later).
   user: wilsonadam
   pass: Cyber1
g. To select a specific variable, check the box on the left and click "Sub-Select"
   You do NOT need to actually download any data from here. These are the 3-hourly data, note the size of the files and time periods available. Do you see why we didn't want all of you to download a dozen or more files from here at the same time? NARCCAP has not set up a mechanism to subset the data (like what is available via IRI Data Library) so the entire file has to be downloaded even if only a small portion is required. To avoid everyone downloading the full dataset, I've already downloaded the complete 3-hourly precipitation and temperature data from the NARCCAP site for several GCM-RCM combinations for the 1) historic (1968-2000) and 2) future (2038-2070) time periods. I also downloaded the RCMs that were forced by the 'observed' data in the NCEP reanalysis dataset (1) observed (1979-2004). I then subsetted all of them them to just the New England region and reprojected them to wgs84 latitude-longitude grid. I then took used the 3-hourly data to generate the daily total precipitation and min/mean/max temperature for each day. This reduced the size of the data from ~221GB to 1.2GB. I've posted the code I used to do this on the website in the Supplementary materials section if you are interested...
h. Today we're going to work with just one GCM-RCM combination (GFDL-RCM3). You can find these files (122MB) on the course website.
i. Download and unzip the Data folder from the website to your working directory. You'll find the following files:
   i. GFDL_RCM3_Current.nc - this is the 1968-2000 daily data from the RCM3 RCM forced by the GFDL GCM for the historical period from NARCCAP.
   ii. GFDL_RCM3_Future.nc - this is the 2038-2070 daily data from the RCM3 RCM forced by the GFDL GCM for the future period (A2 scenario) from NARCCAP.

        iii. NewEngland.shp (and other files) - This is a vector shapefile (map) of the New England states. We'll use it to make some plots later.
        iv. CDO_Process.R - This is an R script that you will work though, much like we did last week.
        v. You will also see the two files that you should have downloaded from the IRI data library (GFDL_Future.nc and GFDL_Current.nc). We put them here in case the website did not cooperate during class.

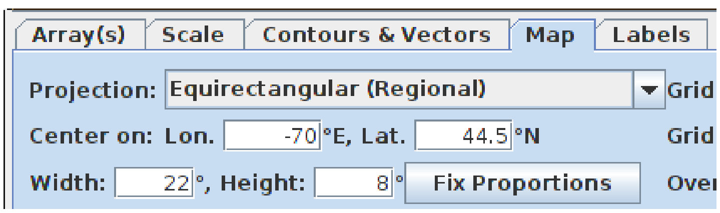
2. Exploration with [Panoply]
    a. Panoply is a useful program for 'browsing' netcdf files developed by NASA. You can view all the metadata (units, dimensions, variable names, and other attributes) and look at simple plots. This section will use Panoply to explore the data you just downloaded.
    b. Open Panoply (icon on desktop or in "All Programs" menu), then open one of the netcdf (.nc) files.



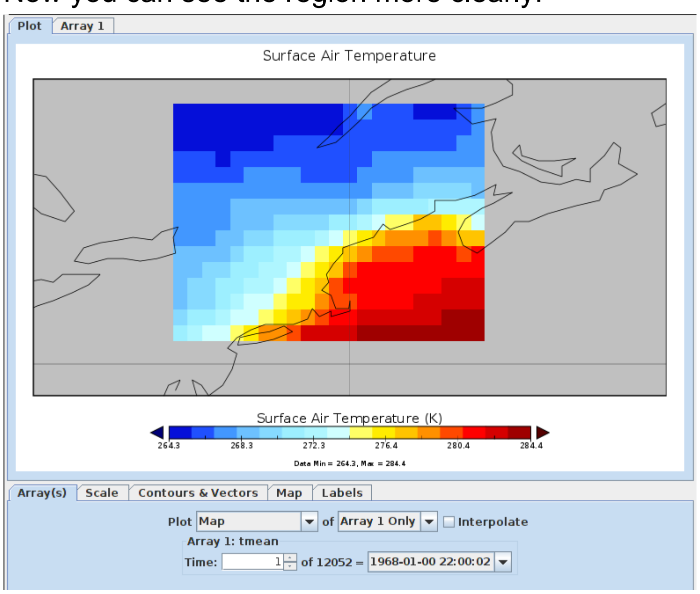
    c. What variables are present? How are the data structured? How many latitude grid cells are there, how many time steps? What are the units of the variables?
    d. Make a plot:
        i. double click on a variable and choose "lon-lat" plot
        ii. The default display shows the entire globe, but these data are only for New England:



So we need to change the view parameters in the “Map” tab below the image. Feel free to choose your own, or enter the following:



iii. Now you can see the region more clearly:

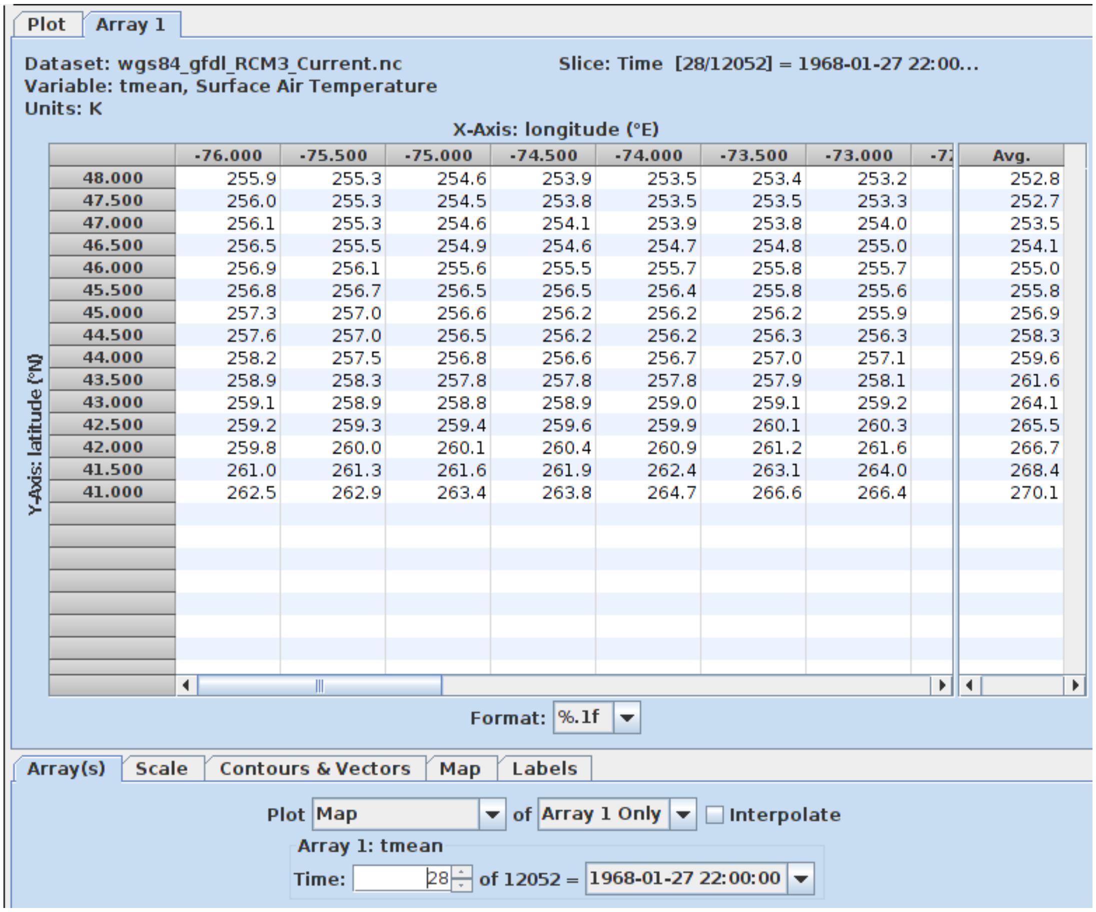


if you want this to be the default region the next time you open Panoply, change the parameters in the preferences as well (edit-> preferences-> lonlat tab).

iv. What does the “Interpolate” checkbox (near the bottom on the right) do? Is this useful?

e. Note the time/date at the bottom of the “Array” tab. To navigate to different times, change the time index or choose a date in the list to the right. You can put the cursor in the time box and use the up and down arrow keys to animate the display (note that there may be a delay between when you select a different time and when the display updates).

f. Now look directly at the data being displayed by clicking on the “Array 1” tab at the top:



	-76.000	-75.500	-75.000	-74.500	-74.000	-73.500	-73.000	-7	Avg.
48.000	255.9	255.3	254.6	253.9	253.5	253.4	253.2		252.8
47.500	256.0	255.3	254.5	253.8	253.5	253.5	253.3		252.7
47.000	256.1	255.3	254.6	254.1	253.9	253.8	254.0		253.5
46.500	256.5	255.5	254.9	254.6	254.7	254.8	255.0		254.1
46.000	256.9	256.1	255.6	255.5	255.7	255.8	255.7		255.0
45.500	256.8	256.7	256.5	256.5	256.4	255.8	255.6		255.8
45.000	257.3	257.0	256.6	256.2	256.2	256.2	255.9		256.9
44.500	257.6	257.0	256.5	256.2	256.2	256.3	256.3		258.3
44.000	258.2	257.5	256.8	256.6	256.7	257.0	257.1		259.6
43.500	258.9	258.3	257.8	257.8	257.8	257.9	258.1		261.6
43.000	259.1	258.9	258.8	258.9	259.0	259.1	259.2		264.1
42.500	259.2	259.3	259.4	259.6	259.9	260.1	260.3		265.5
42.000	259.8	260.0	260.1	260.4	260.9	261.2	261.6		266.7
41.500	261.0	261.3	261.6	261.9	262.4	263.1	264.0		268.4
41.000	262.5	262.9	263.4	263.8	264.7	266.6	266.4		270.1

You can browse these data in the same way as the image (the only difference is the first panel assigns a color to each number and overlays a map of the coastline).

g. Saving subsets of data or images
    i. If you want to save an image of your current view, you can choose file->Save Image.
    ii. You can also export the data as a text file or a KML file for viewing in Google Earth (or any program that supports kml).

h. Do you notice any patterns as you look at different days? Do you see differences over the land and water? How would you use daily data in analysis?

3. Process NARCCAP daily data with CDO software.
    a. In the last class, we used “pure R” to process the monthly decadal mean temperature and precipitation into various summaries. That means we used

the built-in R functions like `apply()` to do most of the calculations. While extremely flexible and powerful, it can also be quite complicated to work with 4-dimensional arrays that way and it's possible that errors will go unnoticed if you aren't experienced with R. Today we're going to explore another tool that is designed to do exactly those sorts of summaries (and faster too!) on this type of data. We're going to still use R to 'drive' the process, because R scripts are wonderful ways to organize and store your analytical workflow (starting with your original data and resulting in your figures, tables, and other results).

b. Now you have downloaded two netCDF files with more than 12,000 days of temperature and precipitation data. How do you summarize this into something useful? There are many ways to work with these data, but one of the easiest and most powerful are the the CDO tools. Check out the CDO tools documentation (https://code.zmaw.de/embedded/cdo/1.4.7/cdo.html) to get an idea of what this program can do.
c. Open the R script called "CDO_Process.R" that you downloaded in the data directory in R. Like Monday's exercise, it will guide you through an analysis and there are a few exercises at the end. Start working on through that script now and come back to this document if you finish early.
d. If you finish early, explore the other functions available through the CDO tools (https://code.zmaw.de/embedded/cdo/1.4.7/cdo.html) and practice stringing together commands to generate the data you are interested in. In a later class we'll explore some of the climate indices available through the CDO tools. You can check them out now here.