MileStone2 – Data Analysis
METEO473 Spring 2016

We have been working quite a bit with NetCDF files and now it is time to extend this to the next level. If you're not yet familiar, to review basic information on NetCDF files look at lesson 14 at this link: http://snowball.millersville.edu/~adecaria/ESCI386P/. MileStone2 will further advance the idea of working with NetCDF files, but in addition we will explore some basic calculations using the extracted data.

Goals: In this project, we will gather monthly averaged precipitation data over 1979-2016 from Earth System Research Laboratory (ESRL) for the whole world. Then we will isolate the precipitation data for a country of choice with the following requirements: (a) you cannot choose the USA and (b) it must have an area greater than 50,000 km^2. You will then run a *regression model* at each grid point corresponding to this country to estimate trends in seasonal precipitation over the years 1979-2015 (using meteorological seasons, for more info click here: https://www.ncdc.noaa.gov/news/meteorological-versus-astronomical-summer—what's-difference. Once the trend is calculated at each grid point, **four sub maps on a first figure will be created depicting seasonal spatial average precipitation over the 1979-2015 for the country of choice. Four other sub maps, on the second figure, will show spatial trend of the precipitation for each season over the whole country, the third figure will contain p-values at each grid point corresponding to the trends in the figure 2** and finally **four sub plots on the fourth figure will illustrate time series of averaged precipitation over all of the grid points in the country for each season**. There will be total of 4 figures with 4 subplots each.

Step 1) Make a directory called:

/home/meteo/your_username/meteo473Sp2016/MileStones/MileStone2

and change its permissions to 700. Within this directory make a script called

MileStone2_PrecipitationAnalysis_LastNameFirstName_LastNameFirstName_
LastNameFirstName_ LastNameFirstName.py

where the first "LastNameFirstName" is the name of the partner whose last name occurs first alphabetically, the second "LastNameFirstName" is the name of the partner whose last name occurs second alphabetically, and so on. Do not use nicknames in your file names and remove apostrophes. For example, Bill O'Brian would be listed as OBrianWilliam in the file name. The two instructors would have a Python file named:

MileStone2_ PrecipitationAnalysis _ForestChris_BalashovNikolay.py

Step 2) Go to http://www.esrl.noaa.gov/psd/. Find the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) precipitation dataset - *monthly and pentad global gridded precipitation means. It includes a standard and enhanced version (with NCEP Reanalysis) from 1979 to near the present.* Use OpenDAP to read the standard version of `precip.mon.mean.nc` inside of your program. Use the link below to help you with this task.

https://publicwiki.deltares.nl/display/OET/Reading+data+from+OpenDAP+using+python

Step 3) Once your program can access the precipitation data file, you are in a position to begin extracting the appropriate precipitation data. First, choose a country for which you would like to carry out precipitation trend analysis. Then write a code that extracts appropriate data for the country of your choice. ***Your code must get monthly time series for precipitation on all grid points inside of the country of choice.*** From these monthly time series at each grid point, you will derive seasonal time series. As review, there are four meteorological seasons: winter (December-January-February), spring (March-April-May), summer (June-July-August), and autumn (September-October-November). You can average 3 appropriate monthly values to get a seasonal average. When this is accomplished you will use the `scipy` module to calculate a trend in precipitation at each of the grid points for each season. This will produce the corresponding basic statistical parameters: slope of the regression line, intercept of the regression line, correlation coefficient, two-sided p-value for a hypothesis test whose null hypothesis is that the slope is zero, standard error of the estimate and coefficient of determination. You will need to search the web for how to use linear regression from the `scipy` module. Python provides many powerful tools, but **the most important skill is to learn how to use these tools**. This milestone gives you an opportunity to practice this skill.

Deliverables:

The code must produce 4 figures each with 4 subplots:

Figure 1: Four maps for each season with spatially averaged precipitation over the 1979-2015 for the country of your choice.

Figure 2: Four maps for each season with spatial trend of the precipitation over the 1979-2015 for the country of your choice.

Figure 3: Four maps for each season with p-values at each grid point corresponding to the trends in the Figure 2.

Figure 4: Four subplots for each season illustrating the time series of averaged precipitation over all of the grid points in the country. On each subplot, a corresponding linear trend line and a text box with the following parameters are required: *slope of the regression line*, *intercept of the regression line*, *correlation coefficient*, *two-sided p-value for a hypothesis test whose null hypothesis is that the slope is zero, standard error of the estimate,* and *coefficient of determination.*

Step 4) Create a simple website that contains your Python scripts and graphics so that they can be accessed and downloaded. Your web site should be well-organized and contain text that describes the project, what the scripts do, and explains your group's findings as illustrated in the graphics. Well-written and to-the-point sentences are more valuable than more numerous poorly written sentences. *(Hint: You should develop the website on the local system first (i.e., ulteosrv1), and then "publish" to the final location.)* When your group is done with MileStone2, send an email to the Instructor with the link to your web site. The Instructor will use this link to download your scripts and images, assigning a grade to them based on the quality of the code, the correctness of the results, and the quality of the web site.

You can follow these directions (http://www.met.psu.edu/facilities/computing/faqs-and-how-tos/personal-web-space) to get a web space for your website or you can use any other web space that you might prefer to use. You are allowed to use other development tools (e.g., WordPress or SquareSpace) if you are familiar with these.

Comments:

Here is the list of the modules you may want to use for this assignment:

```
netCDF4
ncdump
numpy
scipy
pylab
datetime
matplotlib.pyplot
mpl_toolkits.basemap
```