Assignment 3 (50 pts)

You may complete the assignment using your programming language of choice. Feel free to use built-in functions but make sure you have read the documentation about these functions and are confident they are indeed conducting the calculations you intend. Please submit your assignment and the code used to generate any results by uploading the files to Canvas by the assignment due date. I should be able to re-create all of your results given the scripts you have provided me, this includes any data import commands. For your data loading, please ensure that your path is a relative path in the form '../../data/filename.ncl'. Your code should be well commented so that others can easily understand what has been done and marks may be removed from your assignment if this is not the case.

In this assignment, we are going explore the use of linear regression for removing linear trends and removing the variance in a time series that is linearly associated with a particular phenomenon (ENSO). Each of the questions below will build upon this goal.

1. (5 pt) Data preparation

a) In the class Box folder you will find two new datasets:
1) `nina34.data.climatedatacenter.txt` a table of monthly ENSO index values where the first row is the year and the subsequent rows are the 12 months from January to December. This data is the Nino 3.4 index (1948-present) ESRL/NOAA available at the UCAR climate data guide website `https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni`.
2) NCEP/NCAR reanalysis 1 monthly surface air temperature data `air.mon.mean.1948-2020.nc`. This was obtained from NOAA at `https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.surface.html`. This data is 3-D with dimensions (lon, lat, time) and has already been subset spatially to cover the region 15-65N and 150-60W.

b) All of our analysis will be conducted for the month of January, for years 1950-2020. Reduce the ENSO index and NCEP reanalysis data to include only January means from 1950-2020. The time variable in the netcdf file is in units of hours since 1800-01-01 00:00:0.0. This is a fairly typical time format but one that is difficult to use. *Coding Tip: Creating a short function that can convert this type of time variable into a more usable yyyymmddHHMM format will be beneficial now and in the future.*

2. (10 pt) In this question, we will focus on the ENSO index time series.

a) Plot the January ENSO index values as a function years. Calculate the linear trend (linear regression with years as the predictor) of the ENSO index and add this trend line to the plot.

b) Detrend the ENSO index timeseries (ie. remove the portion of the ENSO index that is linearly associated with the year) and normalize the time series (Note: you can consider the trend removal as removing the mean). Plot the normalized ENSO time series and include a line at the value 0 and two dotted lines denoting ENSO values of -1σ and +1σ.

Any years where the ENSO index lies below -1σ will be consider La Nina years and any years where the ENSO index lies above 1σ will be considered El Nino years. For the remainder of the assignment you will be using this detrended and standardized time series for ENSO.

3. (10 pt) In this question, you will work on detrending the surface air temperature, much like what was done for the ENSO index but at each grid point in the surface air temperature data.

a) Compute the linear trend of surface air temperature at each grid point. Convert your trend values into units of C/century. Plot a map of the linear trend at each location and indicate significant values (alpha = 0.05). Significant values can either be shown using stippling or by masking out insignificant values, however if you use the masking out method you must be able to clearly distinguish between values close to 0 and values that are insignificant. *Coding Tip: This is not the only regression you will be conducting in this assignment, it will be useful to make a function for your regression calculations. Similarly, you will be making several maps and so creating a plotting script or re-using your plotting script from last assignment will be helpful.*

b) Remove the linear trend from each grid point. For the remainder of the assignment you will be using this detrended data.

4. (25 pt) In this question we are going to use our detrended and standardized ENSO index as well as our detrended air temperatures to investigate the impact of removing variance that is linearly associated with the ENSO index.

a) Compute the composite surface air temperature for La Nina events (years with ENSO index <= 1σ). Use a two-sided student's t-test to compute the statistical significance. Plot a map of the composited surface air temperature values and indicate significant anomalies. Either in your figure or in your figure caption, include a list of the years used for the composite (ex. the La Nina years). Briefly describe your results.

b) Compute the linear regression of surface air temperature and the ENSO index where ENSO index is the predictor. Similar to the trend map you plotted in question 2, plot the correlation value and indicate with stippling, whether the correlations are significant. Note: recall the regression coefficient (b) is related to the correlation coefficient (r), here you are plotting the correlation coefficient. Conduct a filtering using the linear regression to obtain surface air temperature anomalies that are uncorrelated with the ENSO index, like the detrending conducted in question 2. Briefly describe your results.

c) Re-compute your composites but with your new surface air temperatures that are linearly uncorrelated with the ENSO index. Plot the new composites. What did you expect this new composite to look like? What differences are there as compared to your expectations and can you provide some ideas as to why this might occur?

d) Repeat your analysis steps 4a and 4c for El Nino years. *Coding Tip: Instead of creating a second code for the case of El Nino years, create a switch at the top of the code to choose which case to analyze. This switch should help you to modify your ENSO index criteria, filenames for saving, and titles.*