Assignment 4 (35 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 power spectral analysis applied to a particular phenomenon (ENSO) using the same data source as Assignment 3. The El Nino Southern Oscillation is known to exhibit variability on 3-7 year timescales.

1. (5 pt) Data preparation

a) The same ENSO index dataset employed in Assignment 3 will be used for this assignment. Namely, in the class Box folder: `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`.

b) In this assignment we will use data from all months but only for years 1950-2020. Reduce the ENSO index data to include only years 1950-2020. Re-organize your data into a time series that is in time order.

c) Plot the ENSO index values as a function time and include this plot of the raw data in your report.

d) Detrend the ENSO index timeseries (ie. remove the portion of the ENSO index that is linearly associated with the year) and subsequently remove the mean. Plot the resulting time series and include it in your report. This is the time series (f_t) that will be used for computing your power spectrum.

2. (28pt) In this question create the scripts necessary to produce a series of power spectrum plots. Inclusion of figures and discussion will be left to question 4.

a) (2pt) Create the following variables to be used to produce your power spectrum and include a list of their values in your report: period (T), number of time steps (N), time interval length (Δt), fundamental frequency (ω_0), and Nyquist frequency (ω_N).

b) (4pt) Compute the fast Fourier transform of f_t to obtain F_k. Create the Power spectrum estimated as: $\hat{P}_k = \frac{1}{N}|F_k|^2$.

c) (5pt) Create two plots that will always be produced when running your code. The first will be a power vs. frequency plot. For better visualization you can limit your x axis to f between 0 and 0.2. Be sure to include an informative title, x-label, and y-label. Specify the tick locations in your x-axis to be the frequencies that correspond to the following τ values: 50, 20, 10, 5, 2, 1, 0.5 years. Modify the ticklabels to be these τ values in years. *Practical Note: we want these displayed in years but time is currently measured in months.* Create a second plot of f*Power vs. ln(f). As with the first plot, be sure to include an informative title, x-label, and y-label and use the same tick locations and labels. *Practical Note: This is a semi-log plot in the x-axis of f*Power*

d) (4pt) Add a new functionality to your code of applying a window of your choice (ex. Hann). This should be included as an option at the top of your code so that it may be run with and without a window.

e) (5pt) Add a new functionality to your code to employ chunking, where adjacent F_k's are summed produce an estimate of the power over a frequency band. Your code should be designed so that one factor can be changed (at the top of the code) in order to accomplish this chunking. You will test the following number of chunks (M_{ch}): $M_{ch} = N/1$ (no chunking), $M_{ch} = N/3$, and $M_{ch} = N/6$. This chunk choices divides the time series equally into spectral bands, which is not guaranteed for any chosen M_{ch}. *Practical Note: In this assignment, I will not ask you to create a programming solution for the situation where $M_{sp} = M_{ch}/2$ is a non-integer but it would be useful for you to make a note of this potential problem in your code or add an error message should this occur.*

f) (4pt) Fit an AR(1) model to your ENSO index input data (f). Use this fit to produce an estimate for the Fourier transform of the red-noise time spectrum of the form:

$$S(h) = \frac{1-\rho^2}{(1-2\rho cos(\frac{h\pi}{Msp}) + \rho^2)}. \tag{1}$$

Scale this S by multiplying by the total power in the observed spectrum (P_T) and dividing by the total power in the theoretical AR(1) spectrum ($S(h)$). Add this as a new line to your plots and include a legend.

g) (4pt) Using S computed above the apriori and aposteriori estimated as follows:

$$\frac{S(h)}{\nu}\chi^2_\nu(1-\alpha) \tag{2}$$

where $\alpha = 0.05$ for the apriori estimate and $\alpha = 1-(1-\alpha^*)^{1/M_{sp}}$ with $\alpha^* = 0.05$ for the aposteriori estimate. Plot these two new lines on both plots. Power spectra above these lines are considered significant at the 0.05 aprior and aposteriori levels (respectively).

3. (14 pt) This final question involves including plots created using the code developed in question 2 and describing the resulting power spectra.

a) (2pt) Parseval's theorem states that total power (P_T) can be evaluated equally in the time and frequency domains. Check the validity of this theorem for your data by computing the total power (in discrete form) in the time and frequency domains separately. Include each of these estimations of total power in your report. Does your analysis support Parseval's theorem?

b) (2pt) Include in your report two plots of your final power spectrum one with and the other without the window applied. Both spectra should use N chunks (no chunking applied) and include the red-noise (AR(1)) power spectra for determining significance. Describe the impact the use of a window has on your power spectrum. Did this improve your results, if so how?

c) (4pt) Include three figures of your final power spectra with the following three number of chunks(M_{ch}): $M_{ch} = N/1$ (no chunking), $M_{ch} = N/3$, and $M_{ch} = N/6$. The spectra should be computed with a window applied and include the red-noise (AR(1)) power spectra for determining significance. Describe the impact that chunking has on your interpretation important time scales of variability. Which figure would you choose to show in a manuscript and why?

d) (4pt) From your figures above and your knowledge of the El Nino Southern Oscillation, which time scales would you say are a significant sources of SST variability in this region. In making these statements are you using apriori or aposteriori estimates of significance? Justify your choice.