ADVANCED ENVIRONMENTAL DATA ANALYSIS

HOMEWORK #5

Q1. Download the following timeseries: HW5_Timeseries1.txt and HW5_Timeseries2.txt. Assume that $\Delta t = 1$ month. Determine the relationships between the two time series using the following two methods:

(a) ***Method 1:***

- Calculate the lag correlation between the two time series. Also, calculate and plot the 95% (or 99% or both) significance level(s) for the correlations. Do you see any coherent and/or significant relationship at any particular lag?
  - To assess significance, use the Student t-test. To calculate the t-score for each lag, use the formula: $t = r\dfrac{\sqrt{N^* - 2}}{\sqrt{1 - r^2}}$, where r is the correlation coefficient, $N^* = N\dfrac{1 - r_1 r_2}{1 + r_1 r_2}$ is the *effective* degrees of freedom, N is the total sample size, r_1 is the lag-1 autocorrelation of time series 1, and r_2 is the lag-1 autocorrelation of time series 2.
  - Look up the t-score from a table either online or in Hartmann's notes. **Note:** To assess the 95% significance level using the *two-tailed* Student t-test, you need to use $t_{0.975}$ on tables that present the t-score for *1-tailed* tests. (Confusing, I know...).
- Calculate the power spectra of the two time series separately. To increase the degrees of freedom, subdivide the data into 10 segments and average the separate spectra. Use a Hanning window on each subdivide. Plot the spectra, the corresponding red-noise fits, and 95% significance curves (one spectrum plot per figure). Are there any significant peaks in the data?

(b)***Method 2:***

- Calculate the cross spectrum between the two time series. Subdivide the data 10 times and average the spectral estimates together. Plot the cospectrum and also the coherence squared (coh^2), both as a function of either frequency or period. **Note:** First calculate the cospectrum for the 10 subsets of the timeseries, then average them together. Use the *mean* cospectrum and power spectra in your calculation of coh^2.

- Plot the 95% and 99% significance level on your coh^2 plot (see Hartmann's notes, Section 6.3, for significance levels). At which frequencies is coh^2 significant at that level? How do these frequencies compare to where you have peaks in the cospectrum? If there are coh^2 values that pass both significance levels, what is the phase relationship between the time series at those frequencies?

Compare and contrast the conclusions you came to using both methods. In particular, look at the frequencies where coh^2 is significant and compare that to where the significant peaks are in each of the power spectra you did in part (a).

What are the advantages and disadvantages of each method of analysis? In your opinion, which method would be more meaningful to present, say, in a journal article to explain the relationship between the two time series? Explain.