ATS655 HW4

Tyler Barbero

11 March 2022

0. Time Management

Problemset Code: Link
Estimate of Time to Completion: 10 hrs
Maximum Allotted Time to Completion: 15 hrs
Actual Time to Completion: 9 hrs
Collaborators: James Larson, Andrey Marsavin, Anindita Chakraborty, Ivy Glade, Nico Gordillo, Amanda Bowen, Spencer Jones

Problem 1

(a) First, you will compute the power spectrum. To do this, apply a Hanning window to your data , and split the data into chunks (typically, 256 or 512 days is used). Use an overlapping window technique (WOSA) allowing the windows to overlap by half each time (the standard approach). Plot the resulting normalized (fraction of variance) power spectrum as a function of frequency.

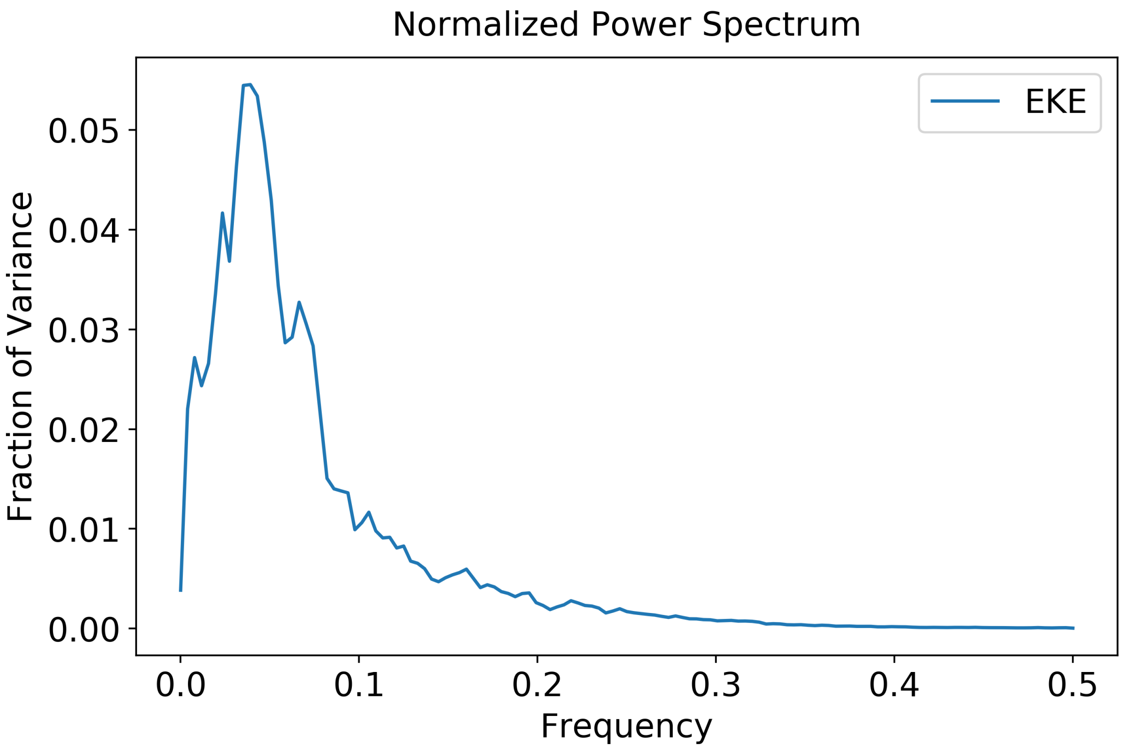

Figure 1: Normalized power spectrum for eddy kinetic energy anomalies in the Southern Hemisphere with seasonal cycle removed. Hanning window of chunkLength 256 used and overlap of 128.

(b) Assess the significance of the spectral peaks using a red-noise null-hypothesis. Plot the rednoise null hypothesis with the power spectrum of the data.
Looking at Figure 2 below, we see several peaks at lower frequencies passing the red-noise null-hypothesis (red line).

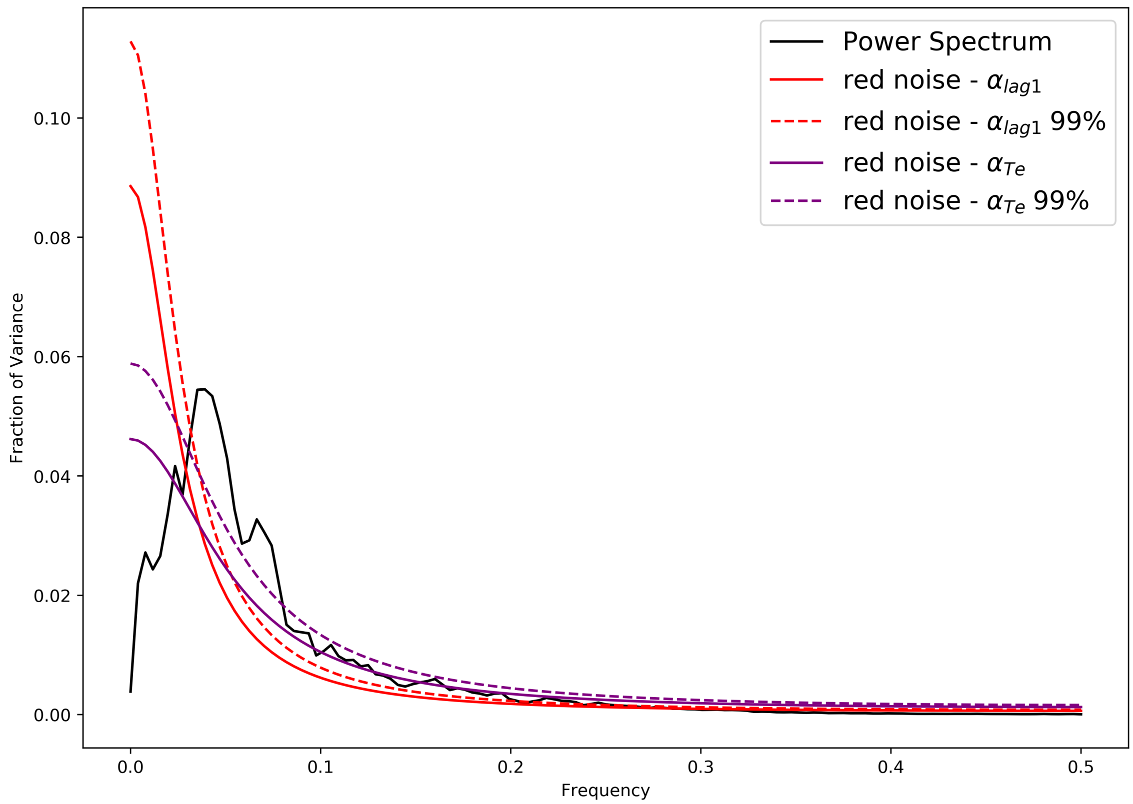

Figure 2: Power spectrum for Problem 1 with red noise lag-1 autocorrelation alpha red noise e-folding time autocorrelation alpha. Also plotted is the 99% confidence interval for each of the red-noise spectra for comparison in part C.

(c) Comment on how your red-noise test differs if you define the memory of the red-noise fit based on the lag-1 autocorrelation versus the e-folding time of the autocorrelation. For the degrees of freedom you can neglect any smoothing from the window.

The alpha for lag-1 autocorrelation = 0.84 while the alpha for e-folding time (which I found was approximately 3 days) was 0.72. This means less memory when using the e-folding time to calculate alpha and more memory for the lag-1 autocorrelation definition. However, we see red noise spectrum with lag-1 drop off faster than the red noise spectrum with e-folding time scale alpha.
(d) Write a report summarizing your findings and the physical meaning behind the peaks (or lack of peaks) you found.

In Figure 2, we see the power spectrum has a higher fraction of variance near the lower frequencies with a peak fraction of variance (peak power) at 0.0390 frequency. This peak is greater than both the red-noise null-hypothesis at both alpha definitions, which suggests the is an important cycle going on at this frequency. I calculated the number of cycles at this frequency since $f = \#cycles/chunklength$, therefore # cycles = 10. I interpret this as some signal at frequency 0.0390 is dominant in the EKE time series (surpasses the red-noise null-hypothesis) and repeats 10 times within each chunkLength. Since I chose 256 chunkLength arbitrarily, the 10.0 cycles per chunkLength may have no physical meaning, but is a consequence of the window width I chose. One thing I also notice is many of the normalized power peaks of EKE easily surpass the lag-1 red noise autocorrelation spectrum (red) but do not pass the e-folding time red noise spectrum (purple). I think this makes sense because the high the alpha (memory), the high the persistence of the data, and the easier the EKE spectrum will be able to pass the lag-1 red-noise null-hypothesis. Similarly, since there is less memory in the Te red-noise null-hypothesis, the harder it is for the EKE spectrum to pass this curve.

The BAM

I recognize the normalized power spectrum of EKE. Although in the plot in the paper, the power goes below the start value of normalized power at a frequency of zero as compared to my plot in Problem 1.

Problem 2

Identify a data series (call it X(t)) of interest to you. Calculate and plot the power spectrum of X(t). Write-up a detailed explanation of the data, why you chose the particular window/chunk length that you did, what the spectrum tells you, which peaks (if any) are significantly different from a red-noise null hypothesis if such a test is appropriate, and the degrees of freedom used to determine this.

My dataset is ERA5 reanalysis zonal and meridional wind speed on the 300, 400, 500, 600, and 700 hPa levels at hourly temporal resolution from 2018 to 2021. I selected a rectangular grid from 10N to 20N and 10W to 30E. With this data, I get the total wind by combining u and v wind components, and average these across all levels and across the whole grid for each time step to get a time series of 35,064 hourly wind values. Lastly, I subtracted the mean to get the anomalies. I plan to look at variations in the African Easterly Jet (AEJ), which lies generally over the domain and levels I specified.

At first I performed the same analysis in #1 on the hourly-resolution anomaly time series. I realized since I am looking at a jet, the memory in the time series, especially for hourly data, is bound to be extremely high and indeed it was. I calculated alpha = 0.999. Thus my power spectrum and red noise spectrum were not smooth (Figure 5). It was difficult to understand what was going on with my time series of such high temporal resolution so I purposely reduced the resolution by averaging for the purpose of this homework assignment.

I reduced the resolution of my data and calculated wind speed averages in 7-day consecutive chunks. So now my dataset is 208 in length, with each value being a 7-day average of wind speed. I chose a chunkLength of 52, which corresponds to yearly chunks, and chose this because of the seasonal cycle of the AEJ (see Figure 4). The AEJ peaks once per year in the NH wintertime when the temperature gradient is highest to due zenith angle of the sun. Now looking at Figure 6, the red-noise spectrum is smooth and the peaks are clearer. I observe a max peak at a frequency of 0.03846 hz. Since frequency = cycles/$\triangle t$, I can solve for number of cycles corresponding to a frequency. The # of cycles corresponding to the frequency with max fraction of variance (highest spectral peak) is = $\triangle$ * frequency = 52*0.03846 = 2. I think this makes sense because the AEJ peaks during the winter and the window chunks kind of split the jet peak in half (Figure 3/4), so for each chunk we see 2 cycles, one peak at the later end of the winter season and one at the beginning (when the cycle repeats). There are several other peaks at higher frequencies which pass the red-noise null-hypothesis. These reside at frequencies: [0.0385, 0.0962, 0.25, 0.2692, 0.3077, 0.3269, 0.3462, 0.3654, 0.3846, 0.4038, 0.4231, 0.4808]. More research is required to understand these what causes these peaks to pass the red-noise null-hypothesis.

The degrees of freedom for the hanning window and half overlap is:

$$DOF = f_w * N/(chunkLength/2) = 1.2 * 208/(52/2) = 9.6$$

degrees of freedom. Since I am averaging, I am purposely decreasing the resolution of the data, and thus decreasing N.

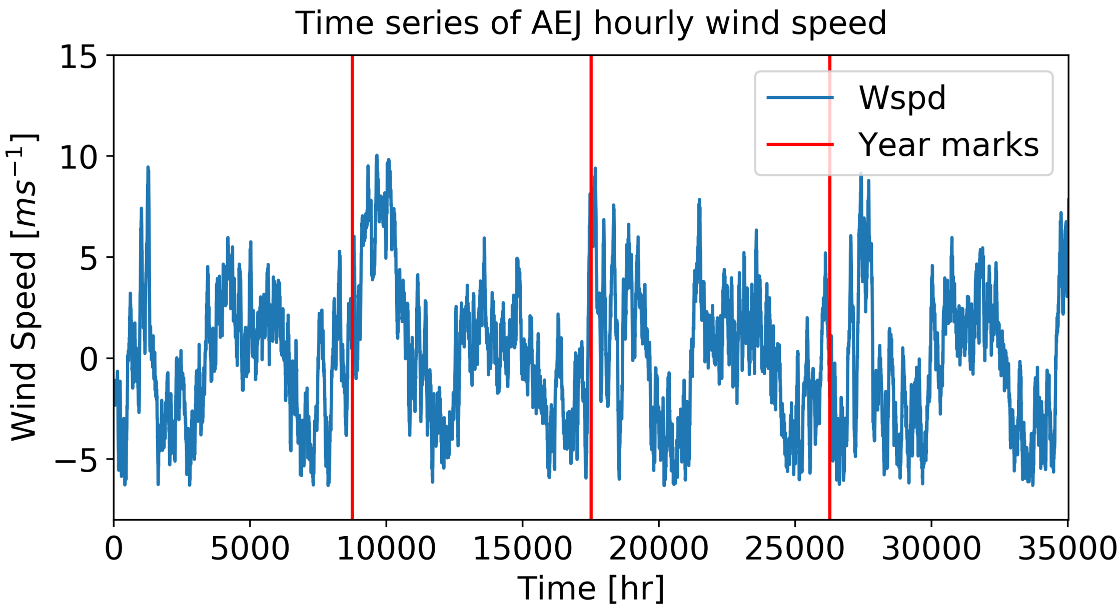

Figure 3: Time series of hourly wind speed data. Zonal and meridional wind is averaged on the 300, 400, 500, 600, and 700 hPa levels over a rectangular grid from 10N to 20N and 10W to 30E from 2018 to 2021.

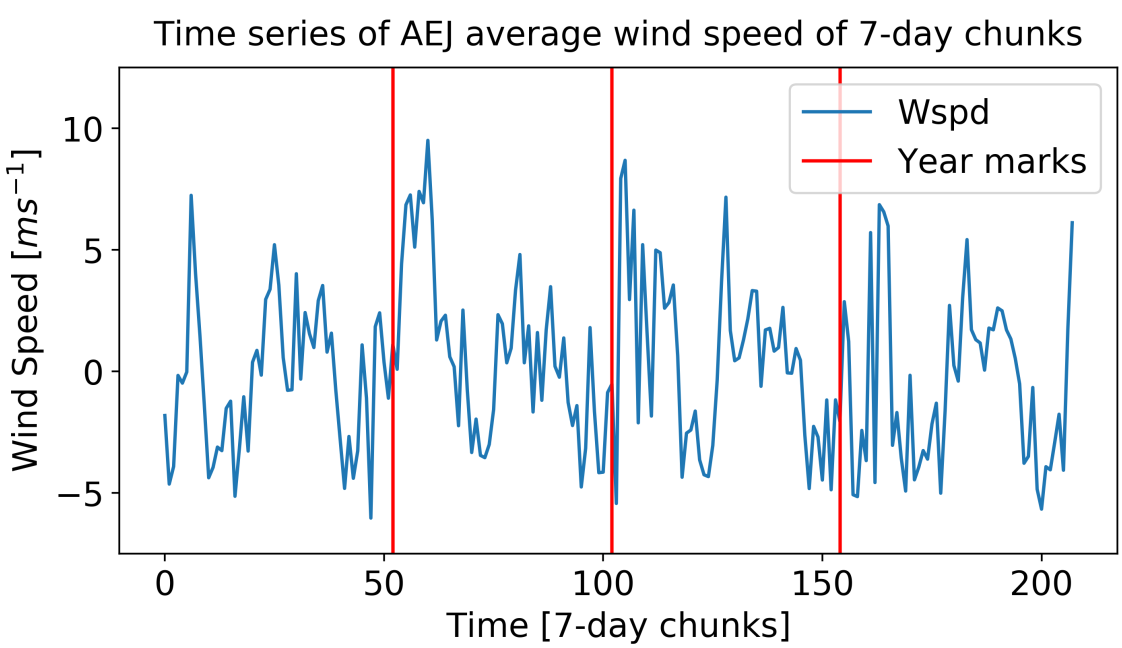

Figure 4: Time series of average wind speed in 7-day consecutive chunks of the hourly time series.

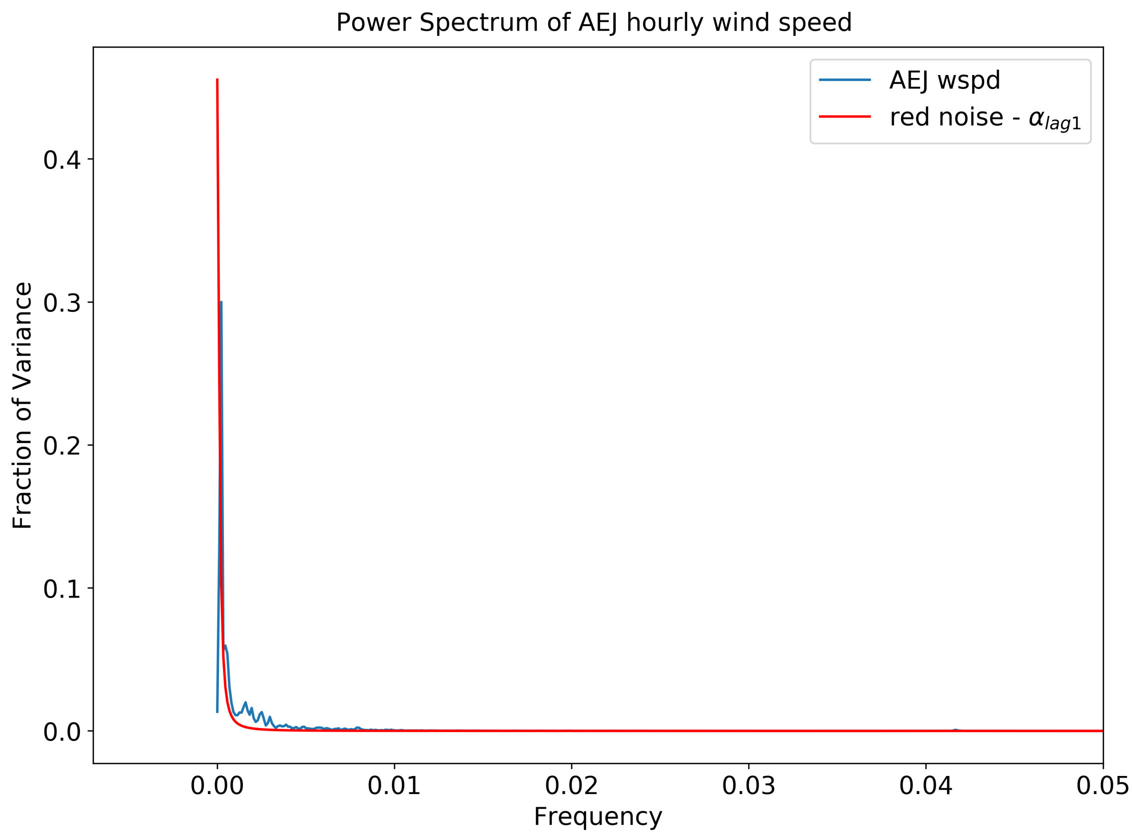

Figure 5: Power spectrum with red-noise null-hypothesis for hourly wind speed data

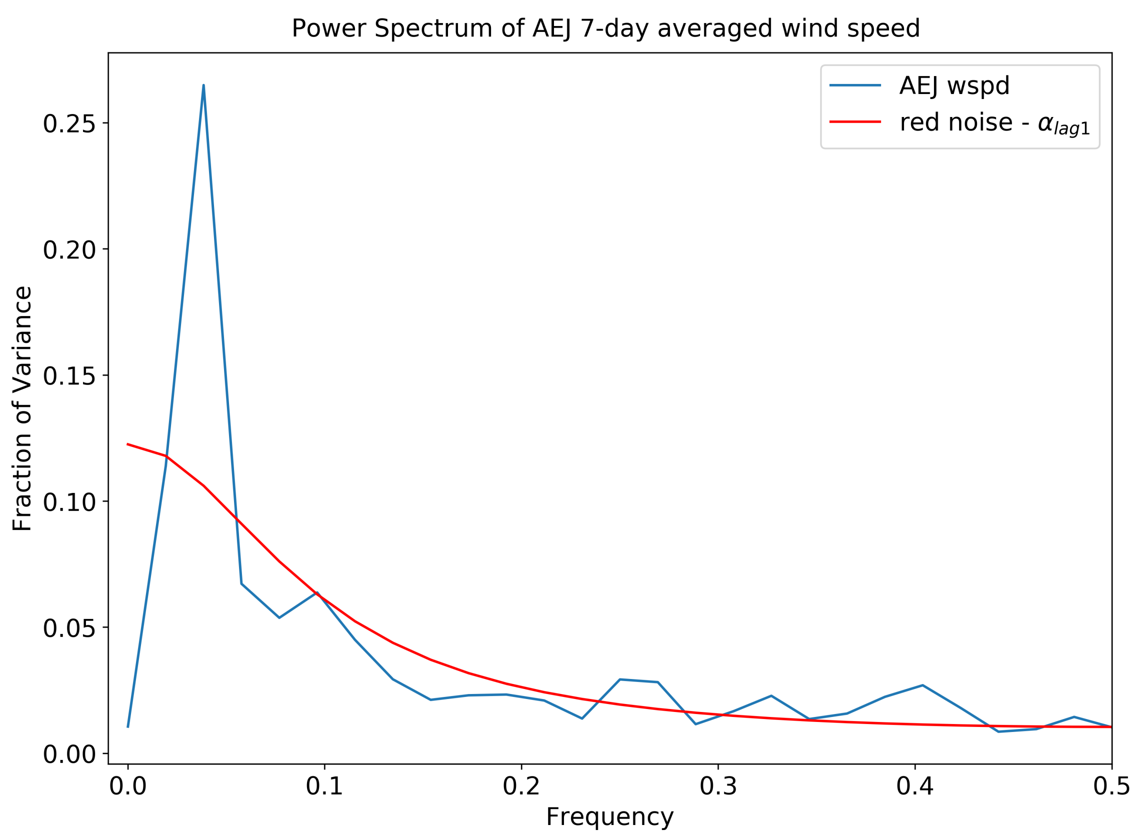

Figure 6: Power spectrum with red-noise null-hypothesis for 7-day averaged wind speed data