**ATOC5860 – Application Lab #4**

**Spectral Analysis of Timeseries**

**Notebook #1 – Spectral analysis of hourly surface air temperatures from Fort Collins, Colorado at Christman Field**

**Questions to guide your analysis of Notebook #1:**

1) Look at your data. What are the autocorrelation and e-folding time of your data? What spectral peaks do you expect to find in your analysis and how much power do you think they will have?

The lag-1 autocorrelation of the data is 0.99 and the e-folding time is 100.92 hours. This suggests the data is very red and has a lot of memory. I expect to find a spectral peak around 24 hours, representing the diurnal cycle. I also expect to find a peak at 365 days, representing the annual/seasonal cycle.

2) Calculate the power spectra using the Numpy method, which assumes a Boxcar window that is the length of your entire dataset. Graph the power spectra, the red noise fit to the data, and the 99% confidence interval. What statistically significant spectral peaks did you find? What do they represent? How did you assess the statistical significance (what is the null hypothesis that you are trying to reject)? Compare back to Barnes and Hartman notes to make sure all of the equations and functions in the notebook are working as you expect them too.

A picture containing chart

Description automatically generated

There are significant spectral peaks at the following frequencies:

365.0 days – yearly cycle

1.0 day – daily cycle/diurnal cycle

0.997 – daily cycle

0.5 days – night/day cycles? Atmospheric tide? Unsure…

The 365.0 and 1.0 day peaks are large whereas the 0.997 and 0.5 day peaks are small.

The null hypothesis is that the data has the same power spectrum as red noise. We assessed statistical significance using an f-test.

Note: The plot above has the red noise power spectrum as a solid red line and 99% confidence interval as a dotted red line so any peaks above the dotted red line are statistically significant. For example, the first peak at 365 days is well above the 99% confidence limit, thus we can reject the null hypothesis that this peak is red noise.

3) Calculate the power spectra using the scipy method. Check that you get the same result as you got using the Numpy method. Next – compare the power spectra obtained using both a Boxcar window and a Hanning window. Assume a window length that is the entire length of the dataset. Do you get the same statistically significant peaks when applying the Hanning window and the Boxcar window? How do they differ? Can you explain why?

**Table

Description automatically generated with medium confidence**

The scipy method gives the same result as the numpy method.

When we use both a Hanning and a Boxcar window, they give the same statistically significant peaks at the same frequencies but the Boxcar window gives higher peaks (more power) that are steeper. This makes sense since the window is skinnier.

*4) If time – take a look at other surface meteorological variables in the dataset. Do you obtain similar spectral peaks?*

Question: Are you seeing power at 12-hour frequencies when looking at temperature? Maybe it is atmospheric tides? Or is it some kind of spectral ringing artifact? Unsolved mysteries of ATOC7500 Objective Data Analysis…

**Notebook #2 – FFT analysis using Dome-C Ice Core Data**

**Questions to guide your analysis of Notebook #2:**

1) Look at your data and pre-process for FFT analysis: Power spectra analysis assumes that input data are on an evenly spaced grid. The Dome-C temperature data are not uniformly sampled in time. Regrid the Dome-C temperature data to a uniform temporal grid in time. Plot the data before and after re-gridding to make sure the re-gridding worked as expected.

Data before re-gridding:

**Chart

Description automatically generated**

Data after re-gridding:

Chart

Description automatically generated

From the plots above, we can see that the re-gridding worked as expected. The layers in the ice core are more compacted towards the bottom so you can’t make out an annual cycle deeper in the ice core because the signal gets smeared.

2) Signal and Noise: What is the autocorrelation and e-folding time of your data? What spectral peaks do you expect to find in your analysis and how much power do you think they will have? *Hint: Think back to the Petit 1999 Vostok ice core dataset discussed in class.*

The lag-1 autocorrelation is 0.96 and the e-folding time is 25 time steps which is about 25,000 years. I expect to find peaks that correspond to the frequencies of the Milankovitch cycles, namely every 100,000 (eccentricity), 41,000 (obliquity), and 21,000 (precession) years.

3) Use Boxcar Window to calculate power spectra: Calculate the power spectra using the Numpy method, which assumes a Boxcar window that is the length of your entire dataset. Graph the power spectrum, the red noise fit to the data, and the 99% confidence interval. What statistically significant spectral peaks did you find? What do they represent?

Chart, line chart

Description automatically generated

We find statistically significant peaks at frequencies 0.01, 0.025, 0.0425, and 0.04375 which correspond to 100328, 40131, 23607, and 22932 years, respectively. However, from the plot, it looks like the peak at a frequency of 0.04375 is a part of the peak at 0.0425.

These different frequencies represent the Milankovitch cycles, which are a result of Earth’s orbit around the sun (eccentricity), Earth’s tilt (obliquity), and Earth’s rotation about its tilt (precession). A frequency of once every 100328 years represents the cycle of ice ages which matches the Earth’s eccentricity cycle (how circular Earth’s orbit is). A frequency of once every 40131 years is about the period of the Earth’s obliquity (how tilted Earth’s axis is). A frequency of once every 22932 - 23607 years is about the period of Earth’s precession (the direction Earth’s axis points in). The Milankovitch cycles affect climate by changing the distance and angle to the sun which affects the strength and length of the seasons. Thus, they should definitely show up in this ice core that dates back ~800,000 years.

4) Compare Boxcar Window vs. Hanning Window: Calculate the power spectra using the SciPy method. Compare the results obtained using a Boxcar window that is the length of your entire dataset to those obtained using a Hanning window that is the length of your entire dataset. Graph the power spectrum, the red noise fit to the data, and the 99% confidence interval. What statistically significant spectral peaks did you find? What do they represent? What are the differences between the results obtained using the Boxcar window and the Hanning window? Is the intuition that you gained by looking at Fort Collins temperatures the same as what you are seeing here with Dome-C temperature records? Why or Why not?

Chart, line chart

Description automatically generated

Using the Hanning window, we find peaks at frequencies of 0.01, 0.01125, 0.02375, 0.025, and 0.0425 which correspond to 100328, 89181, 42244, 40131, and 23607 years, respectively. The peak at 100328 years is the same as the Boxcar window peak and corresponds to the eccentricity cycle. I would guess that the peak at 89181 years is also part of this peak. The peak at 40131 years is the same as the Boxcar peak at 40131 years and likely represents the obliquity cycle. I would also guess that the peak at 42244 is a part of this peak. As for the peak at 23607 years, this is the same as the Boxcar peak at 23607 years which is likely the cycle of Earth’s precession. Thus, it looks like we get the same peaks but with a little more smear/a little wider than before.

From the plot above, we see that we get peaks at about the same frequencies when we use the Hanning window compared to the Boxcar window. However, the Hanning window leads to a slightly lower 99% confidence level and lower peaks (lower normalized power at each peak). The Hanning peaks also look wider than the Boxcar peaks. This makes sense because the boxcar window is skinner whereas the Hanning window is wider. This is the same general pattern that we found with the Fort Collins temperatures which makes sense – the shape of the window affects the shape of the peaks.

5) Hanning Window with different window lengths: Using the SciPy method, compare the power spectra obtained using Hanning window with different window lengths. Graph the power spectra, the red noise fit to the data, and the 99% confidence interval. Did you find any statistically significant spectral peaks? How does decreasing the window length affect the temporal precision of the spectral peaks and their statistical significance? Did you find the classic tradeoff between 1) high spectral/temporal resolution but low quality statistics, and 2) high quality statistics but low spectral/temporal resolution?

Chart, line chart

Description automatically generated

A Hanning window of length 200 years leads to 9.6 degrees of freedom and peaks at frequencies of around 0.025, 0.04, and 0.045. A Hanning window of length 800 years leads to 2 degrees of freedom and peaks at frequencies of around 0.01, 0.012, 0.025, and 0.0425. Therefore, it looks like decreasing the window length makes the data chunks too short to see low frequency oscillations, getting rid of the lower peaks. We see this because the peak at 0.01 is no longer statistically significant with a window of 200 years. Conversely, the normalized power of the peak at 0.025 has increased a lot. Thus, we do see the tradeoff between high spectral/temporal resolution but low-quality statistics and high statistics but low spectral/temporal resolution.

5) Add WOSA (Welch Overlapping Segment Averaging): Having found what you think is a good balance between precision in the identification of the spectral peaks and statistical significance – Try applying WOSA (Welch Overlapping Segment Averaging) in addition to using the Hanning Window with different window lengths. How does this change your results?

Chart, line chart

Description automatically generated

Using Welch’s Overlapping Segment Averaging (WOSA) and a Hanning window of length 200 years leads to 8 degrees of freedom and the plot shown above. In the code, we have the windows set up to overlap by 100 years. We can see that now we get all the main Milankovitch cycle peaks back (around 0.01, 0.025, and 0.045) and they are statistically significant, some more so than others. The peaks are still wider from the Hanning window but they have a stronger normalized power.