Alice Ren

Data I

November 21, 2017

HW 7

Midterm: “I certify that this represents my own work and that I have not worked with classmates or other individuals to complete this assignment.”

1. The air temperature and sea surface (1 m) temperature data were downloaded from the Stratus XIV mooring. The data did not have gaps, and there were no missing values. There were a number of spikes in the data, especially the sea surface temperature data. However, further inspection of the sea temperature data showed that each spike was composed of many data points, more than 15, and data after each spike seemed to return to normal. Thus, I did not treat any of the data (from either air or sea temperature) as bad data. In the austral winter, the sea surface temperature is warmer than the air temperature. In the austral summer months, the sea surface temperature and air temperature match pretty well. Over all time, the air temperature has larger daily variability than the sea temperature. These observations are what we would expect based on what we know about the ocean and atmosphere. The data were collected every one minute, and the record is over one year long. The full record contains 615,727 data points in time each of air temperature and sea temperature.
2. To compute a spectrum to look at the variability in frequency space, I first chose to compute a spectrum of the entire time-series. The resulting spectrum was well resolved in the low frequencies, but the high frequencies had large uncertainties and large variability. To follow a best practice for obtaining better resolution of high frequency signals and a smaller error estimate, I decided to segment the data using 50% overlapping segments and a Hanning window filter. My choice of segment length was to divide the data roughly to 9 non-overlapping segments. This produced individual segments of around 47 days, which would still be long enough to resolve phenomena with a period of a month. I had a total of 9+8=17 segments. Also, to make uniform-length segments, I did not use the last data point. To each segment, I de-trended, applied a Hanning window, and took the Fourier transform.
   1. Parseval’s theorem is verified for the spectrum without any Hanning windowing. Windowing distorts the normalization and preservation of variance.
   2. The x-axis is cycles per day.
   3. The y-axis is [°C] 2/cycles per day.
   4. The uncertainty estimate is provided on the spectra using the segmenting method with degrees of freedom = 17\*2 = 34.
   5. The Nyquist frequency is 1/(2\*sampling interval) which is 1/(2\*1 minute) = 0.5 minutes or 6.94x10-4 days.
   6. The frequency resolution is 1/(N\*Δt) = 0.0210 cycles per day for the segmented case (roughly 1/47 cycles per day).
   7. The largest spectral peaks are at 1 and 2 cycles per day. It appears that there is a peak at 3 cycles per day. The peaks at 1 and 2 cycles per day are tidal frequencies. I am not sure what the phenomenon is that has a period of around 8 hours.
   8. The spectral slope of SST is close to -2. The low frequency half of the air temperature spectrum appears to have a spectral slope of -4/3. At the high frequencies, the air temperature spectrum has a steeper slope and then a shallower slope. The high frequencies of the air temperature spectrum may be more impacted by white noise or instrumental error.
   9. The sea surface temperature spectrum has wider peaks at 1 and 2 cycles per day. I would guess that the ocean mixes heat more slowly than the atmosphere, so patterns of temperature over time might not follow the period of the forcing phenomenon exactly. The air temperature spectrum has interesting features over the last one-third of highest frequencies, as mentioned above. The air temperature spectrum generally has more energy than the SST spectrum. Perhaps this is because the record had more time in colder months than warmer months; the record was over one year and extended to a second winter.
3. To plot a variance preserving spectra, the energy was multiplied by the frequency. Further, the error bars changed with frequency as well.
   1. The x-axis is plotted in log-space with unit cycles per day.
   2. The y-axis is plotted in linear-space with unit [°C] 2 .
   3. The error bar is plotted in a thin black line both above and below the spectrum.
   4. The variance preserving spectrum highlights the peaks at 1, 2, and 3 cycles per day, and emphasizes that the peak at 1 cycle per day has the greatest energy, with the peak at 2 cycles per day coming in second. The variance preserving spectrum diminishes the features at the high frequencies greatly. The error at high frequencies also basically vanishes as well.