Homework 4

Masoud Akbarzadeh

March 15, 2024

Estimated time to Completion: 8 hours
Maximum Allocated Time: 12 hours
Actual Time to Completion: 12 hours
Collaborators: Nari, Juan, Amel, Brandon, Jennifer, Kat,Jesse, Delian

Problem 1

1.a

For computing the power spectrum of the time series, I used the 256 chunk size and 50% overlap.

1.b

The significance of the spectral peaks was calculated using the red-noise null-hypothesis. There are two distinct peaks at around 0.45 and 0.7 which are significant at the 99% confidence level.

1.c

The lag-1 autocorrelation and the e-folding time of the autocorrelation were computed and plotted. Their values are considerably different in this example. In this case the lag-1 autocorrelation is 0.84 and the e-folding autocorrelation is 0.72. The reason for this difference might because the autocorrelation function may not decay at a constant exponential rate, which means the e-folding time, which assumes an exponential decay, might not capture the true decay behavior of the autocorrelation function On the other hand, the rate of decay of the autocorrelation function changes dramatically after the first lag, which means the lag-1 autocorrelation is not perfect in capturing of the true decay behavior of the autocorrelation function.

1.d

Using eaiher the lag-1 autocorrelation or the e-folding time of the autocorrelation, the power spectrum of the time series was compared to the red-noise null-hypothesis and plotted. The most significant peak when comparing the power spectrum of the time series to the red-noise null-hypothesis is at the frequency of 0.045 and 0.07 Both of these peaks are significant at the 99% confidence level if we use either the lag-1 autocorrelation or the e-folding time of the autocorrelation. Since the time step is 1 day, these peaks are representative of phenomena that have 22 and 14 days time scales.
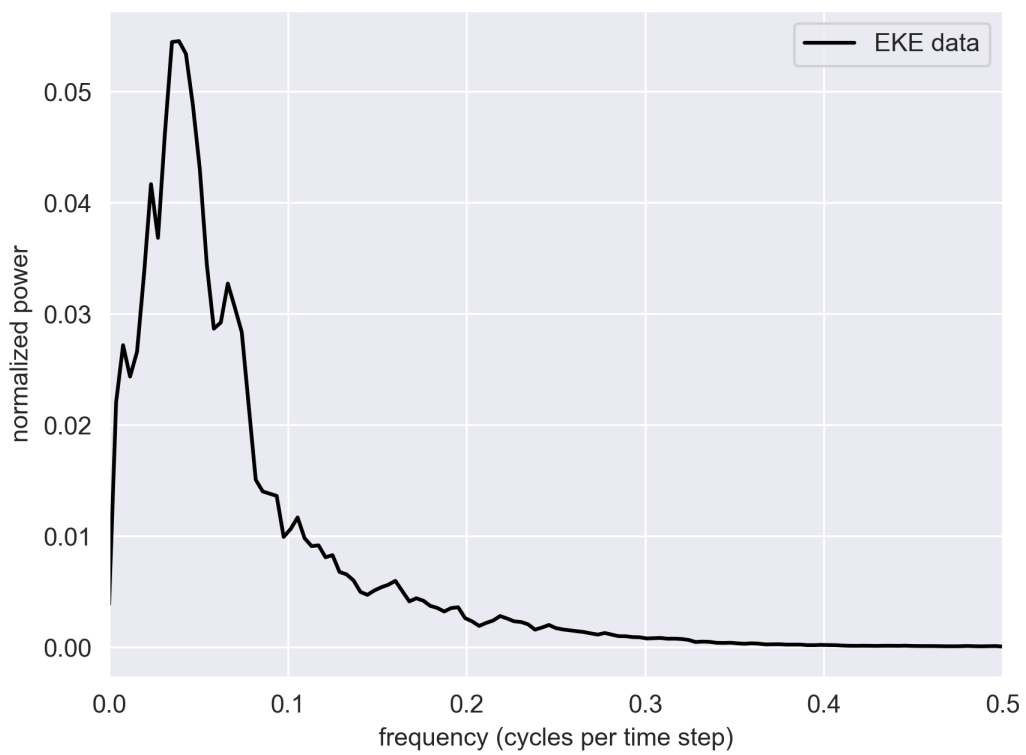
Figure 1: Problem 1 - Plot of normalized power spectrum of EKE time series as a function of frequency
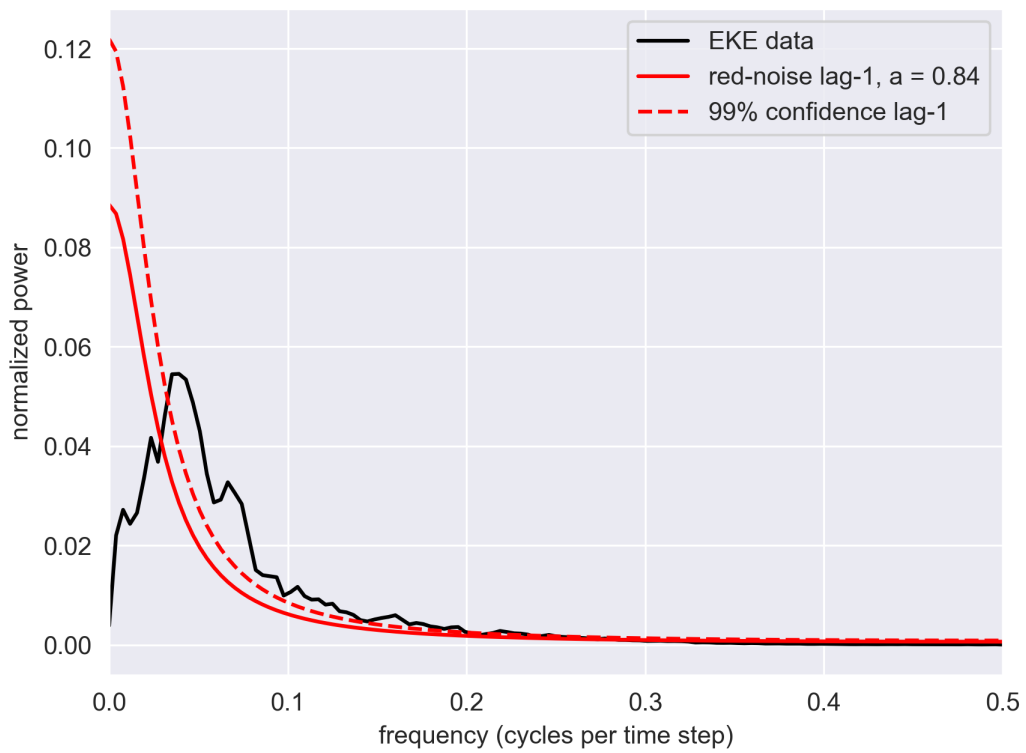
Figure 2: Problem 1 - Plot of normalized power spectrum and red-noise null-hypothesis as a function of frequency
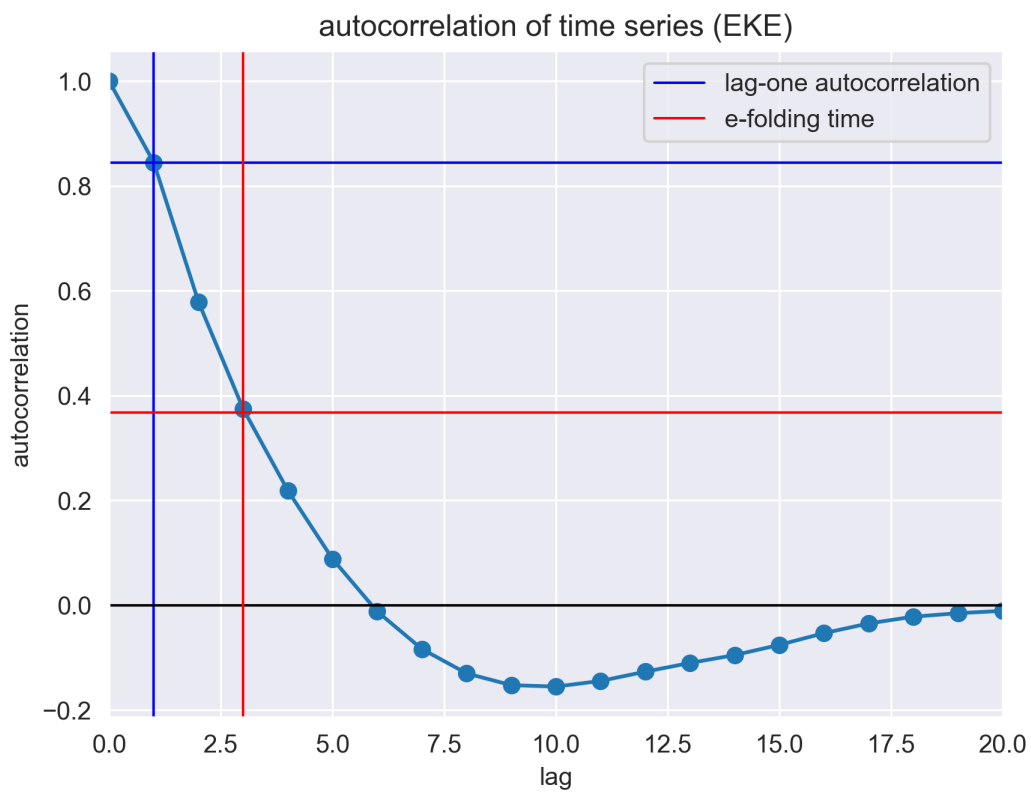
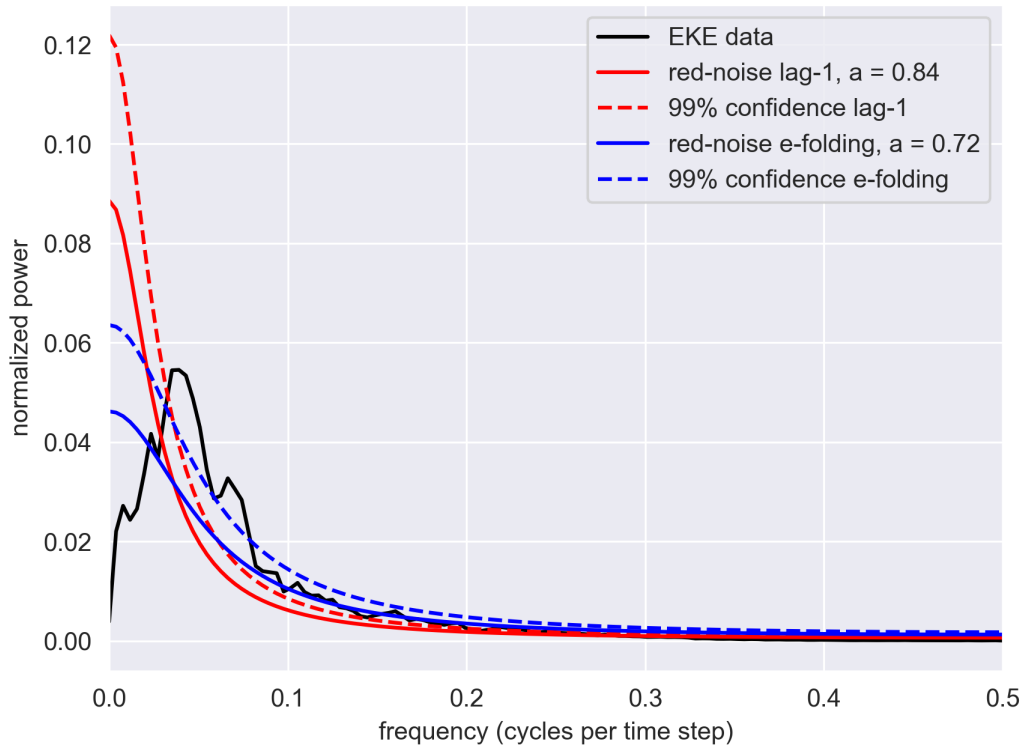Figure 3: Problem 1 - Plot of autocorrelation vs lag for EKE time series

Figure 4: Problem 1 - Plot of normalized power spectrum and red-noise null-hypothesis based on lag-1 autocorrelation and e-folding time of autocoreealtion as a function of frequency
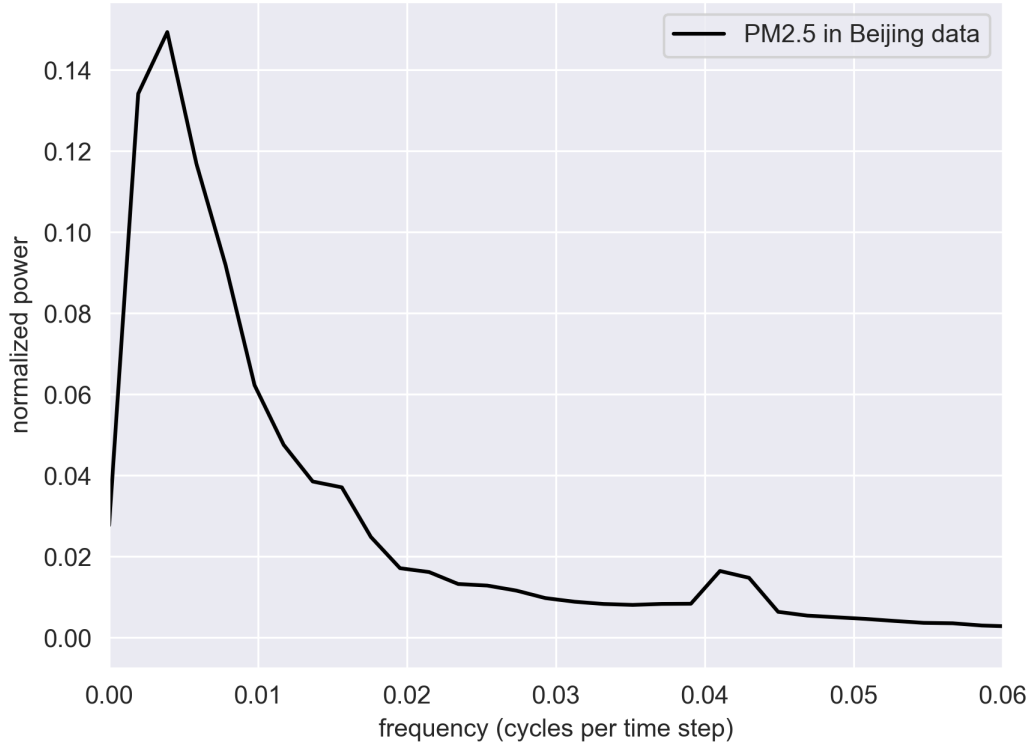
Figure 5: Problem 2 - Plot of normalized power spectrum of PM2.5 time series as a function of frequency

Problem 2

The PM2.5 concentration in one of the stations in the city of Beijing was used as the time series for this problem. The time series was divided into 256 chunks with 50% overlap. Because my data had more than 50,000 data points I had a range of options. I used 256 samples chunk as it gives a suitable degree of freedom and also considers around 10 days chunks. The power spectrum of the time series was computed and plotted. Also, the degree of freedom for this analysis was computed as 198.

The lag-1 autocorrelation and the e-folding time of the autocorrelation were also computed and plotted. Their values are very close to each other in this case. The lag-1 autocorrelation is 0.97 and the e-folding autocorrelation is 0.96. So I used the lag-1 autocorrelation for the red-noise null-hypothesis.

The power spectrum of the time series was compared to the red-noise null-hypothesis and plotted. The most significant peak when comparing the power spectrum of the time series to the red-noise null-hypothesis is at the frequency of 0.041. Since the time step is 1 hour, the period of this peak is 24.4 hours. This must be related to the daily cycle of the PM2.5 concentration in Beijing.
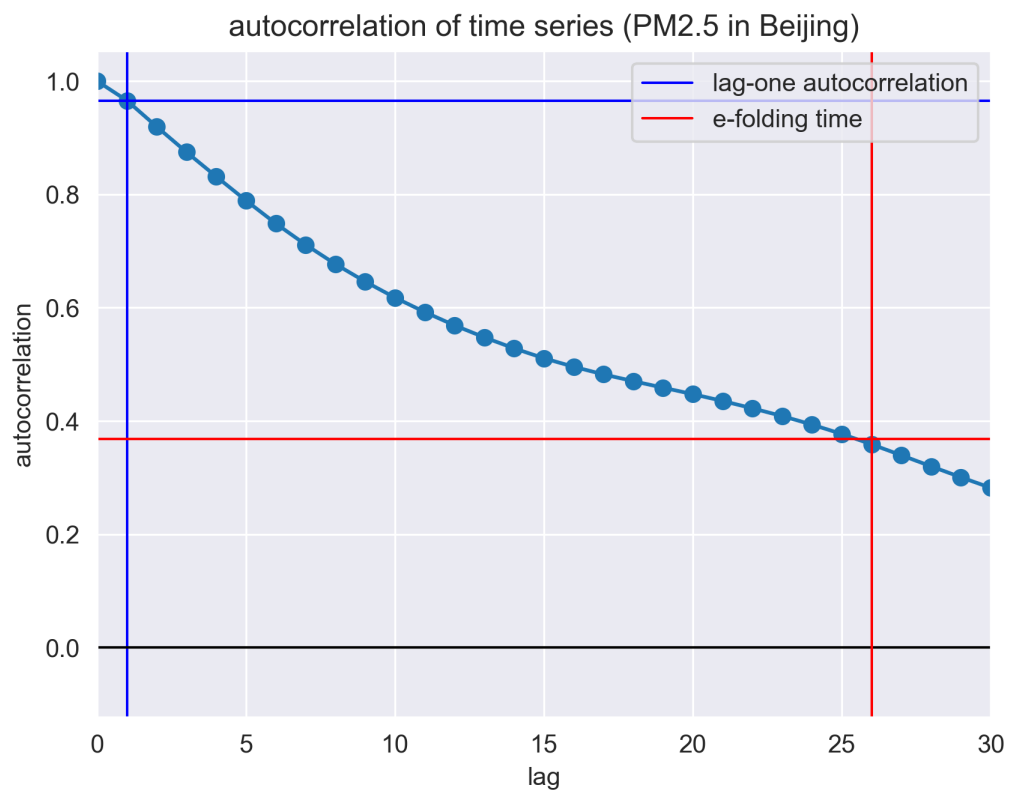
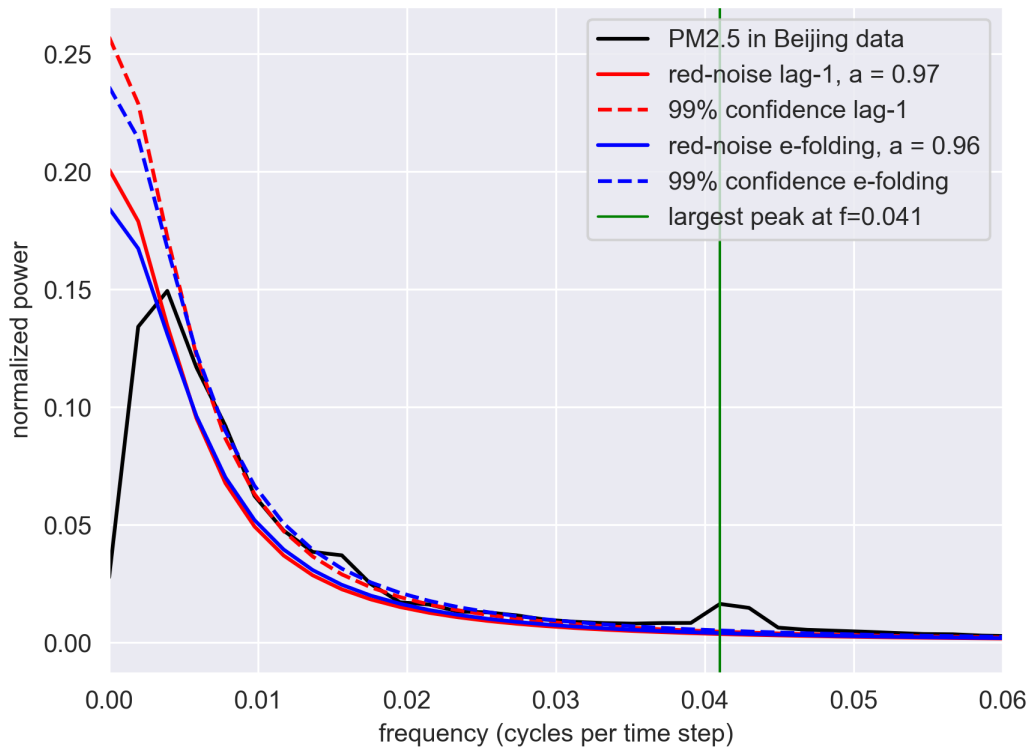Figure 6: Problem 2 - Plot of autocorrelation vs lag for PM2.5 in Biejing

Figure 7: Problem 2 - Plot of normalized power spectrum of PM2.5 time series and red-noise null-hypothesis as a function of frequency