

Comparison of Interpolation Methods: Saliva Subject A

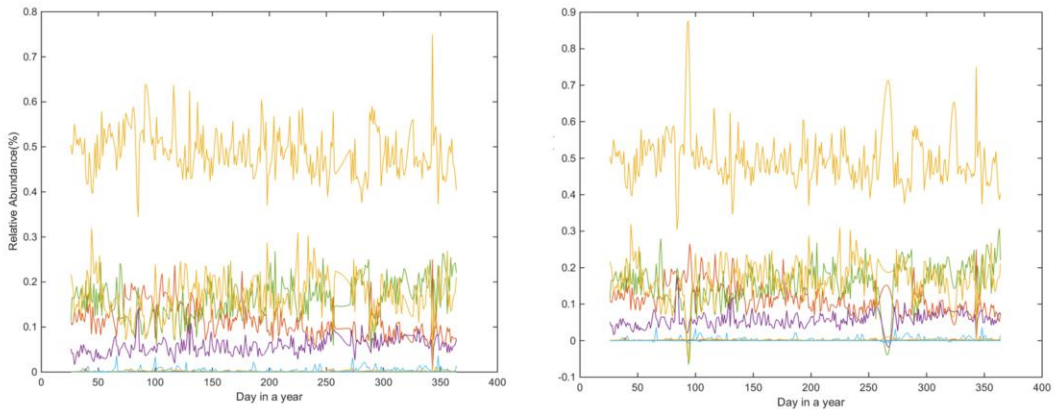


Figure 1. The plots above both show time series of 27 phylums across time span of nearly a year, however the missing values were interpolated with different methods. The plot on the left were interpolated by cubic method, while the one on the right by spline method. Apparently spline method introduces unrealistic negative values, which may become artifacts in the following investigation. The x axis is day in a year, the y axis is relative abundance of a phylum in one day's sample.

Comparison of Interpolation Methods: Stool Subject A

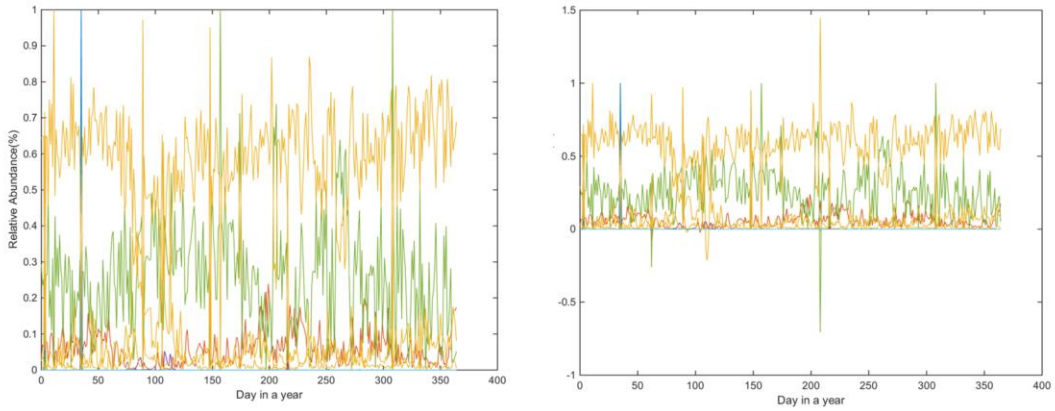


Figure 2. Format is same as figure 1, except the data is from stool samples from subject A.

Comparison of Interpolation Methods: Stool Subject B

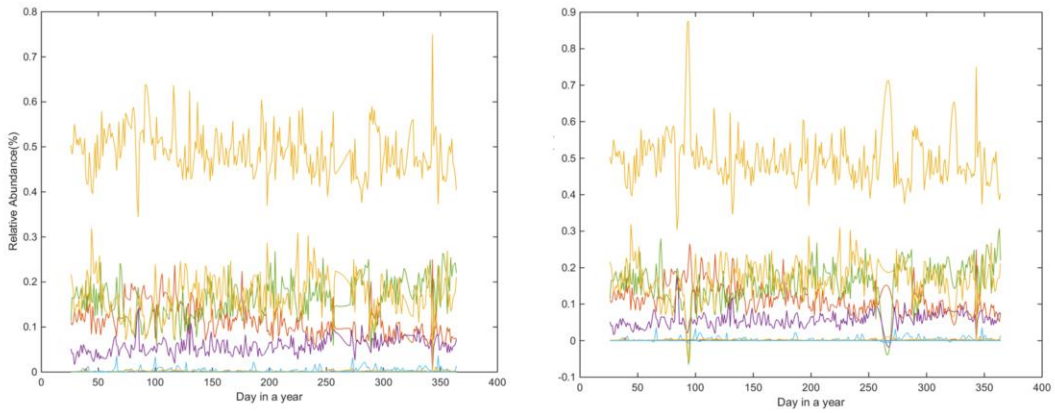


Figure 3. Format is same as figure 1 and 2, except the data is from stool samples from subject B.

Method---Autocorrelogram

Saliva A: top 5 abundant phylums and autocorrelogram

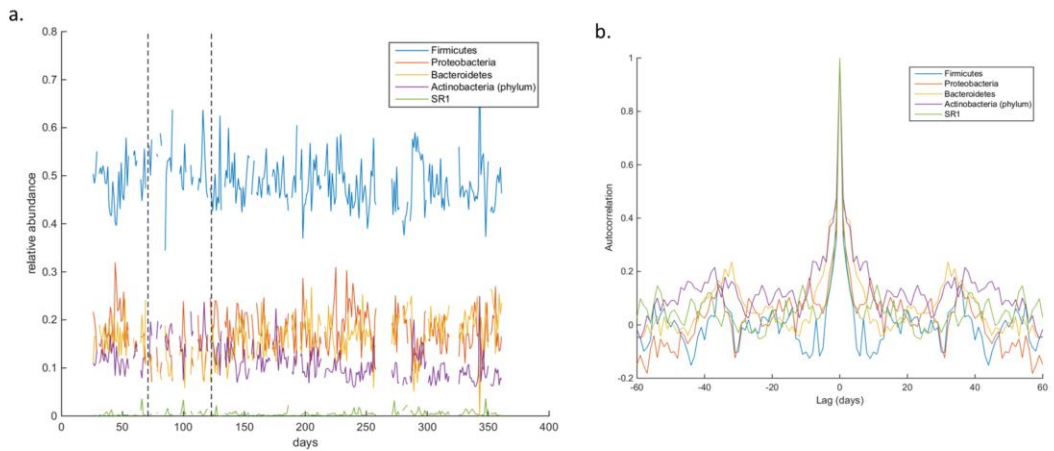


Figure 4. Time series (a) and autocorrelograms (b) for the top 5 abundant phylums in saliva samples from subject A.

Method---Power Spectral Density(PSD) Comparing FFT and Lomb-Scargle Periodograms

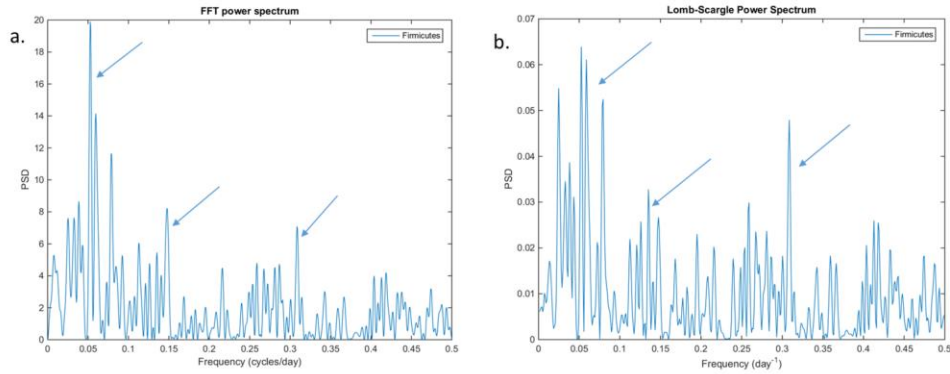


Figure 5. A comparison of periodograms generated by FFT (a) and Lomb-Scargle method (b). The data is Firmicutes relative abundance time series from saliva of subject A, and the time period is from day 123 to day 364. The major difference is input time series data for FFT is interpolated for the missing values, while for lomb-scargle method the input is raw with missing values. The result shows similar representation of power spectral density, although the background noise is higher in (b), they show similar dominant frequency.

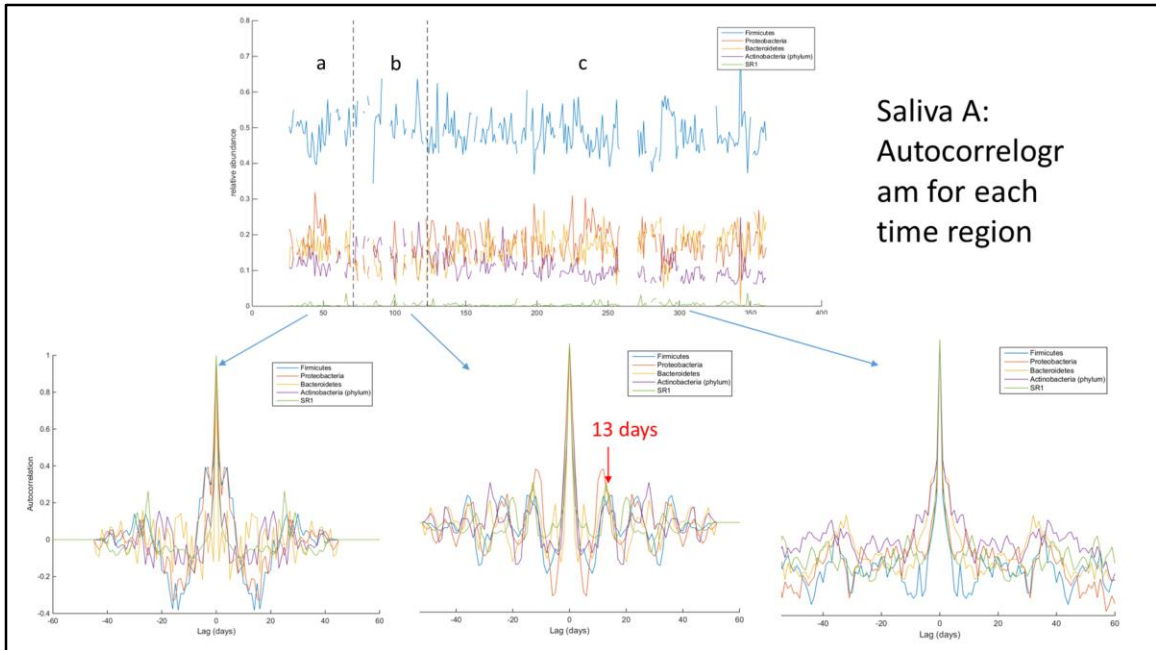


Figure 6. Autocorrelograms for each of the time region a, b and c. This subject was relocating from a major American metropolitan area to the capital of a developing nation in Southeast Asia between days 71 and 122 of the study. This subject was exposed to a novel diet and environment while traveling and had diarrhea on days 80 to 85 and 104 to 113 (during which most of data points are missing). The autocorrelograms shows very distinctive patterns during different time region, indicating drastic changes in phylum abundance periodicity in response to life style changes. One thing to notice is, during the traveling period(b), the top 5 phylums showed consistent periodicity at 13 days, which does not exist before or after the travel.

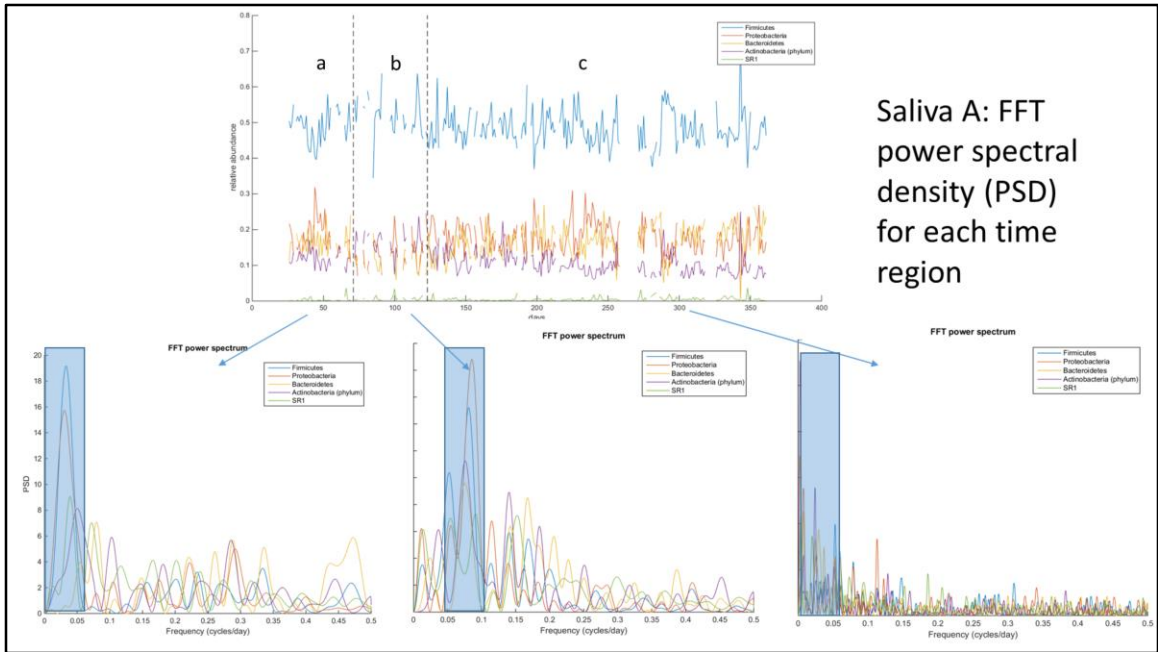


Figure 7. FFT PSD for each of the time region a, b and c. The time events are the same as in figure 6. The FFT PSD shows the dominant frequency shifted from round 30 days (freq = 0.034 cycles/day) to 13 days (freq = 0.079 cycles/day) and then back to 30 days or more.

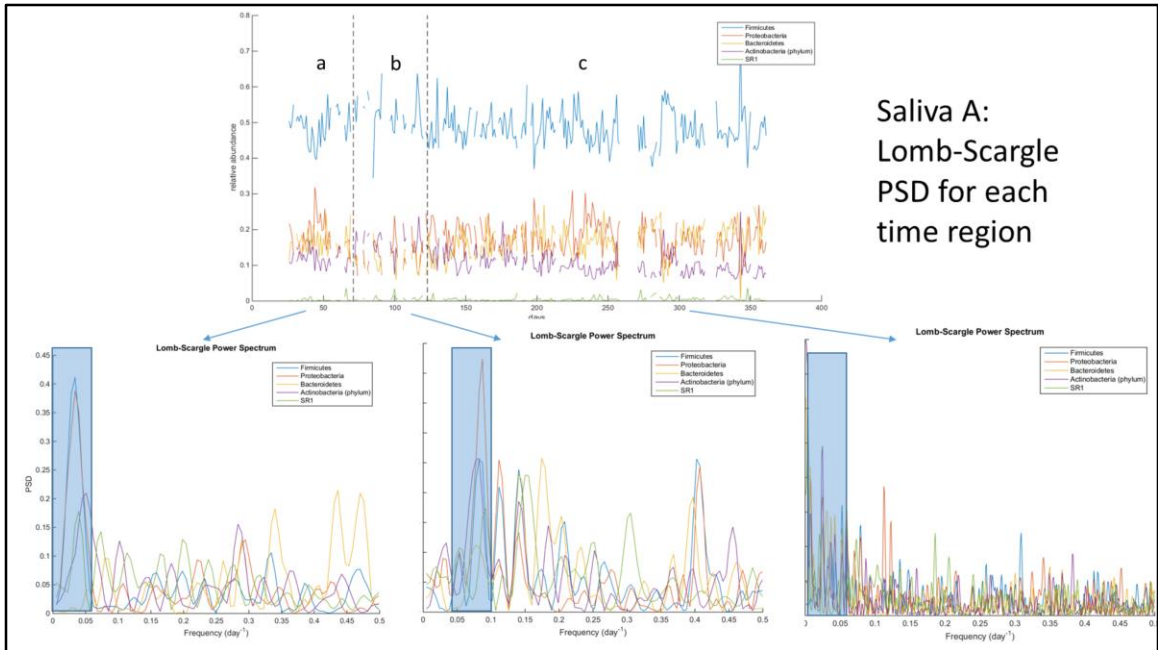


Figure 8. Lomb-Scargle PSD for each of the time region a, b and c. The time events are the same as in figure 6. The Lomb-Scargle PSD shows similar trend as FFT PSD, where the dominant frequency shifted from round 30 days (freq = 0.034 cycles/day) to 13 days (freq = 0.079 cycles/day) and then back to 30 days or more.