Homework #7 (Midterm)

Due March 18, 2022, 11:59pm

Problem 7.1.

The file `atacama-2012-sample.csv` contains hourly measurements from various sensors from the Atacama desert in Chile. The sensors `c3` and `c4` measure the CO_2 concentration in part per million (ppm), uncorrected for the high altitude (Atacama is at 16,000 ft, and the air pressure is about half of the sea-level one). The columns `t5` and `t6` are the outside temperature from two sensors in ^oC. The time is displayed in different granularities (hours from the beginning of the experiment, hours within each day (`dhours`), days from the beginning of the experiment. There is a glitch in the CO_2 sensor values on day 70, ignore those values (set them to zero).

The expression below defines the cross-correlation between two different time-series *a* and *b*.

$$C_{ab}(\tau) = \frac{1}{N}\sum_t \Big(a(t) - \langle a \rangle\Big)\Big(b(t+\tau) - \langle b \rangle\Big)$$

Here *N* is the number of measurements included in the sum, <*a*> and <*b*> are the averages of the two series.

a. Consider the time series of the two temperature sensors. Break these into daily vectors, and compute the top 3 principal components. Guess the meaning of each component. Expand each vector on the basis of the top 3 components. Estimate the fraction of variance contained in the three components. Estimate the truncation error due to using three components only. Display the amplitudes of the components as a function of time during the observations.
b. Repeat the above with 5 components and compare.
c. Compute the temporal autocorrelation function of both the average temperature and the average CO_2 concentration, out to 48 hours. Interpret the result.
d. Compute the temporal cross-correlation function between the average temperature and the average CO2 concentration, out to 48 hours. Discuss the meaning of the result.

Hint: look out for missing or erroneous values in the data, often marked with NaN (not a number).

Problem 7.2.

Calculate the frequency spectrum of the following signals, sampled at 44.1kHz:

a. Guitar sounds are given in the two files `nylon16.wav` and `strat16.wav`. Determine the frequency of the fundamental for both sounds.
b. Calculate the relative energy in each of the first 3 harmonics (exclude the fundamental)
c. Estimate the decay time of the third harmonic (in sec)

Hint:

(i) The energy in a signal is the square of the absolute value of the Fourier transform.
(ii) The frequency spectrum is the energy as a function of frequency

Problem 7.3.

The file `DS.wav` contains a recording of the mysterious 'cricket' sound hurting US diplomats in Cuba. Determine the frequency spectrum of the signal. Do you notice anything unusual? Break the sound into 1 sec long overlapping segments, stepping in 0.5 sec intervals and calculate the power spectrum and plot the frequency spectrum as a function of time. What is the accuracy if the frequency measurement for these spectra?