The Lomb-Scargle periodogram: analysis of incomplete and unequally spaced time series in R.

Thomas Ruf
Research Institute of Wildlife Ecology, University of Veterinary Medicine, Savoyenstrasse 1, A-1160 Vienna, Austria.

R provides several methods to estimate the spectral density of time series, which is a key method in biological rhythm research. For these methods, time series must be equally spaced. Unfortunately however, unevenly spaced time series, or series with missing data are not uncommon in biomedical studies. For instance, telemetric and observational data from free-living animals almost inevitably lack data points due to animal movements, weak transmitter reception, or poor weather and lighting conditions that hinder observations. Periodogram methods for evenly spaced data can still be employed in many of those cases after placing of data on an even-spaced grid, and interpolation or simply 'clamping' of missing values at the last valid point. However, at least in certain cases of long and possibly regularly spaced data gaps these interpolation techniques can perform poorly and may produce false periodogram peaks.

There exists a method to estimate spectral densities in unequally sampled time series, the Lomb-Scargle periodogram. This method is derived from, but not identical with the classical Fourier spectrum analysis and was developed to detect weak rhythms in noisy data obtained from astronomical observations. The Lomb-Scargle method also has the benefit of providing an easy way to calculate 'false alarm probabilities', i.e., levels of statistical significance of peaks in the periodogram. I have implemented this method in R, together with several functions to summarize and plot results.

I have tested the Lomb-Scargle method using both real and simulated time-series with even and uneven sampling, and compared it to a standard method in biomedical rhythm research, the Chi-square periodogram. Results indicate that the Lomb-Scargle algorithm shows a clearly better detection efficiency and accuracy in the presence of noise, and avoids possible bias or erroneous results that may arise from replacement of missing data by interpolation techniques. Hence, the Lomb-Scargle algorithm represents an all-purpose method than may be used for both complete and incomplete evenly sampled time series, as well as entirely unequally spaced data.