Timing Analysis

CONCLUSION

Timing Analysis

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data.

Some Questions That We'd Like to Answer:

Does My Source Vary?
On What Time Scales Does it Vary?
Are the Variations Periodic or Aperiodic?
How Do Different Energy Bands Relate to One Another?

Timing Analysis

Some Questions That We'd Like to Answer:

Does My Source Vary?

On What Time Scales Does it Vary?

Are the Variations Periodic or Aperiodic?

How Do Different Energy Bands Relate to One Another?

Timing Analysis

Some Questions That We'd Like to Answer:

Does My Source Vary?

- ✓ On What Time Scales Does it Vary?
- ✓ Are the Variations Periodic or Aperiodic?
- ✓ How Do Different Energy Bands Relate to One Another?

Variability Test: Kolmogorov-Smirnov

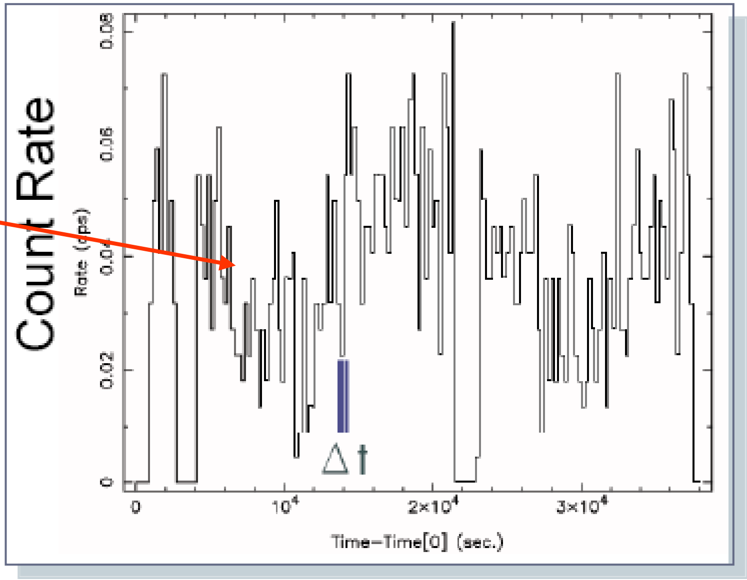
The first Questions That We'd Like to Answer:

Does My Source Vary?

Does My Source Vary?

- Simplest Measure of Variability:
  The root-mean-square variability – r.m.s.
  (the same as standard deviation):

$$\text{r.m.s.} = \sqrt{\frac{1}{N}\sum_i (\text{RATE}_i - \langle \text{RATE} \rangle)^2}$$



- Also, it is common to quote
  the fractional r.m.s., r.m.s./<RATE>

- **Limit:** the above def. is bin-size dependent
  (i.e. we miss any variations smaller than our time bin size)

Variability Test: Chi-Square Test

- ***Hypothesis:*** *the source is intrinsically constant*

  Can I reject this hypothesis?

  Chi-square statistic

$$\chi^2 = \sum_i \left(\frac{\mathrm{RATE}_i - <\mathrm{RATE}>}{\mathrm{ERROR}_i}\right)^2$$

- If measurements are gaussian (!), the statistic should have a chi-square distribution with *(N-1)* degrees of freedom.
- We can calculate the statistic, compare to tabulated values, and compute confidence in our hypothesis.

Does My Source Vary?

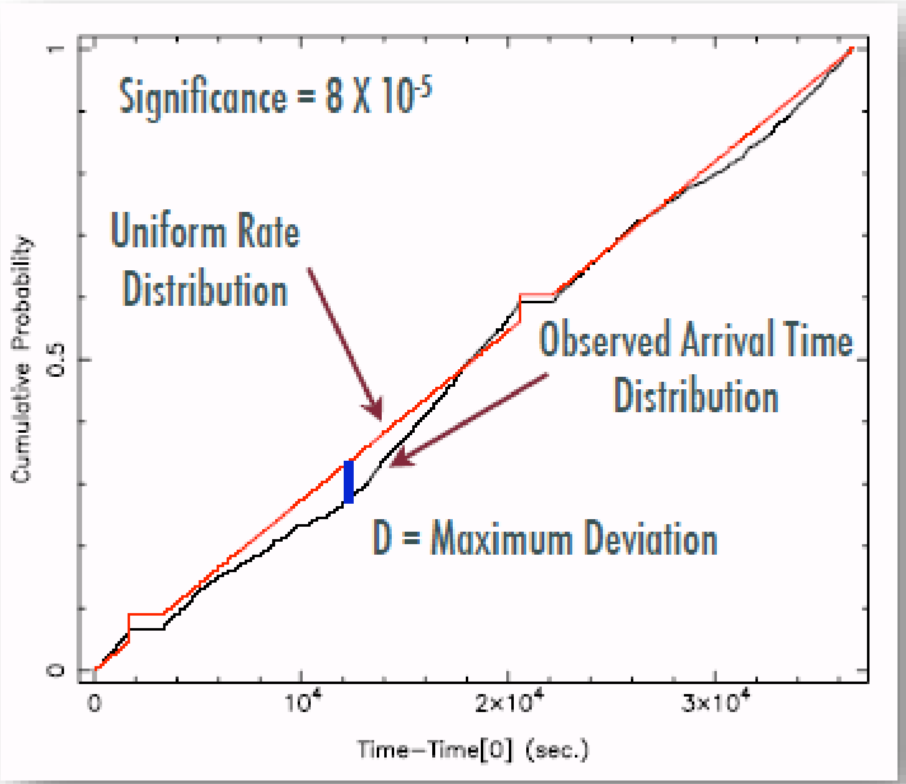
- X-ray Astronomy **Warning**:
  Because we are counting individual photons, the relevant statistics are *Poisson*, not *Gaussian*, so we EXPECT some variation even if the source has a constant intrinsic intensity.
- The assumption of Gaussian statistics eventually fails, when the number of counts per bin is less than ~10, and Chi-Square Test is no longer useful.

Variability Test: Kolmogorov-Smirnov

- Is cumulative arrival time consistent with constant rate?

  -or-

- Is distribution of times in-between events consistent?



Variability Test: The H-test

- The H-test is a test for **sparse** ("arrival time") data that is good at testing for general periodicities.
- Let $t_1 \ldots t_N$ be the set of arrival times of your data. Using the assumed period, calculate the phase φ_i of each event.
- Define: $\alpha_k = \frac{1}{N}\sum_{i=1}^{N} \cos k\varphi_i$ and $\beta_k = \frac{1}{N}\sum_{i=1}^{N} \sin k\varphi_i$
- Now define:

$$Z_m^2 = 2N \sum_{k=1}^{m} (\alpha_k^2 + \beta_k^2)$$

- Finally define H as:

$$H \equiv \max_{1 \le m \le 20} (Z_m^2 - 4m + 4)$$

Variability Test: The H-test

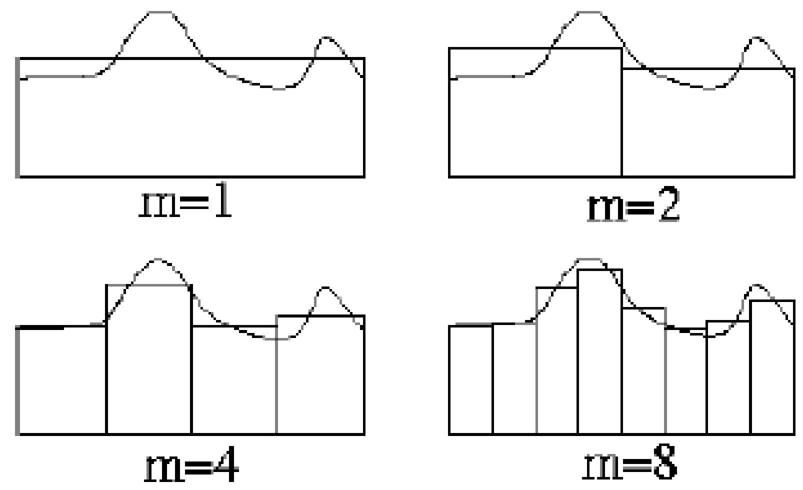
- Under the null hypothesis, up to about H of 23, H has the distribution

$$\mathrm{Prob}(H>h)=\exp[0.398h]$$

- The H-test basically considers fitting the phase diagram with varying number of Fourier components, finds a "best fit" of sorts, and tests on that. This means it's good for anything between broad pulses and very narrow ones (up to about $1/20^{\mathrm{th}}$ of the width of the phase peak).
- As with other periodicity statistics, it can be applied in a scan across a frequency range if you account for the trials factor associated with testing many, not entirely independent, frequencies. Often this is best done by Monte Carlo.

Variability Test: Bayesian analyses

- At any given period, parameterize the light curve by a variable number of bins.
- For m bins, there are the following free parameters: frequency/period, phase, value in each of the m bins
- Use a Bayesian analysis to compare the probability of m=1 vs the probability of m=2, m=3, m=4, etc. Usually m=12 bins is an adequate number.

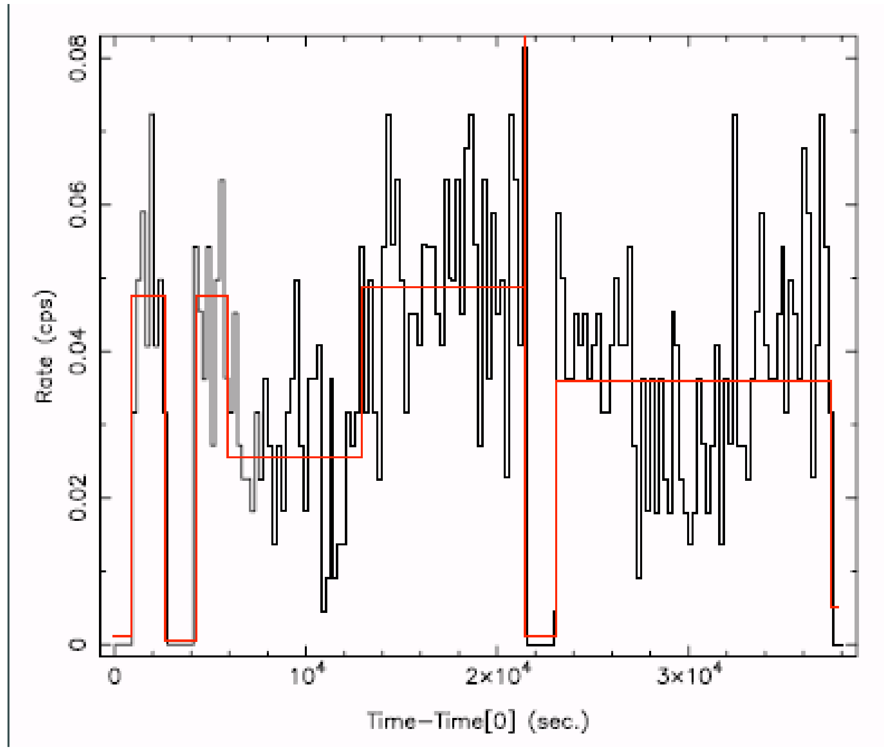


Some epoch-folded ”light curves”.

This approach is ideal when you have no idea what the light curve should look like.

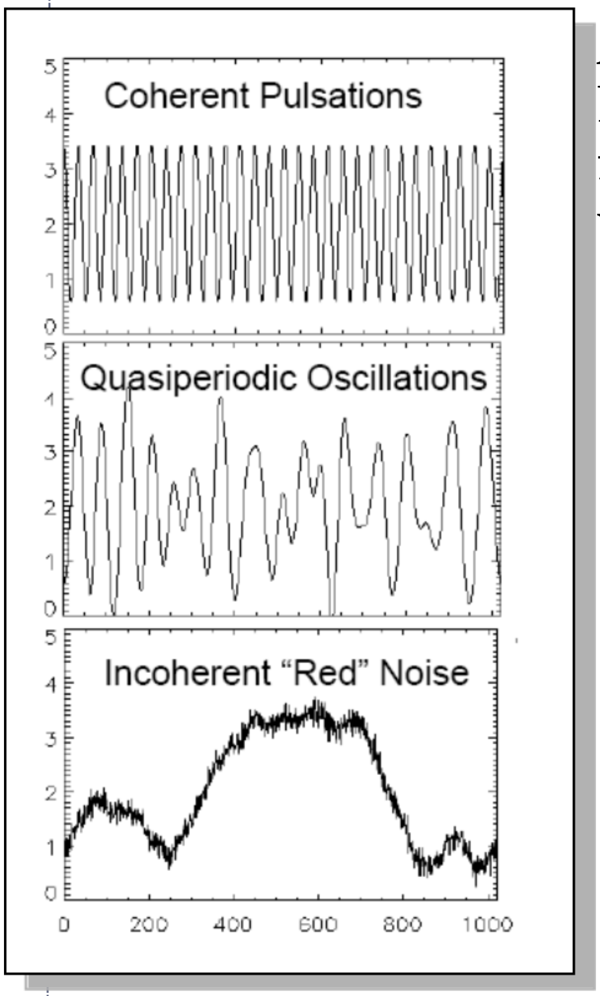
Variability Test: Bayesian analyses

- Occam's factors penalize models with more bins, so the m=1 (no periodicity) model is automatically preferred unless the data demands it.
- Gregory & Loredo (1992, ApJ, 398, p. 146) - Determines Optimal Uniform Binning.
- Bayesian Blocks (J. Scargle) -Determines Optimal Non-uniform Binning.



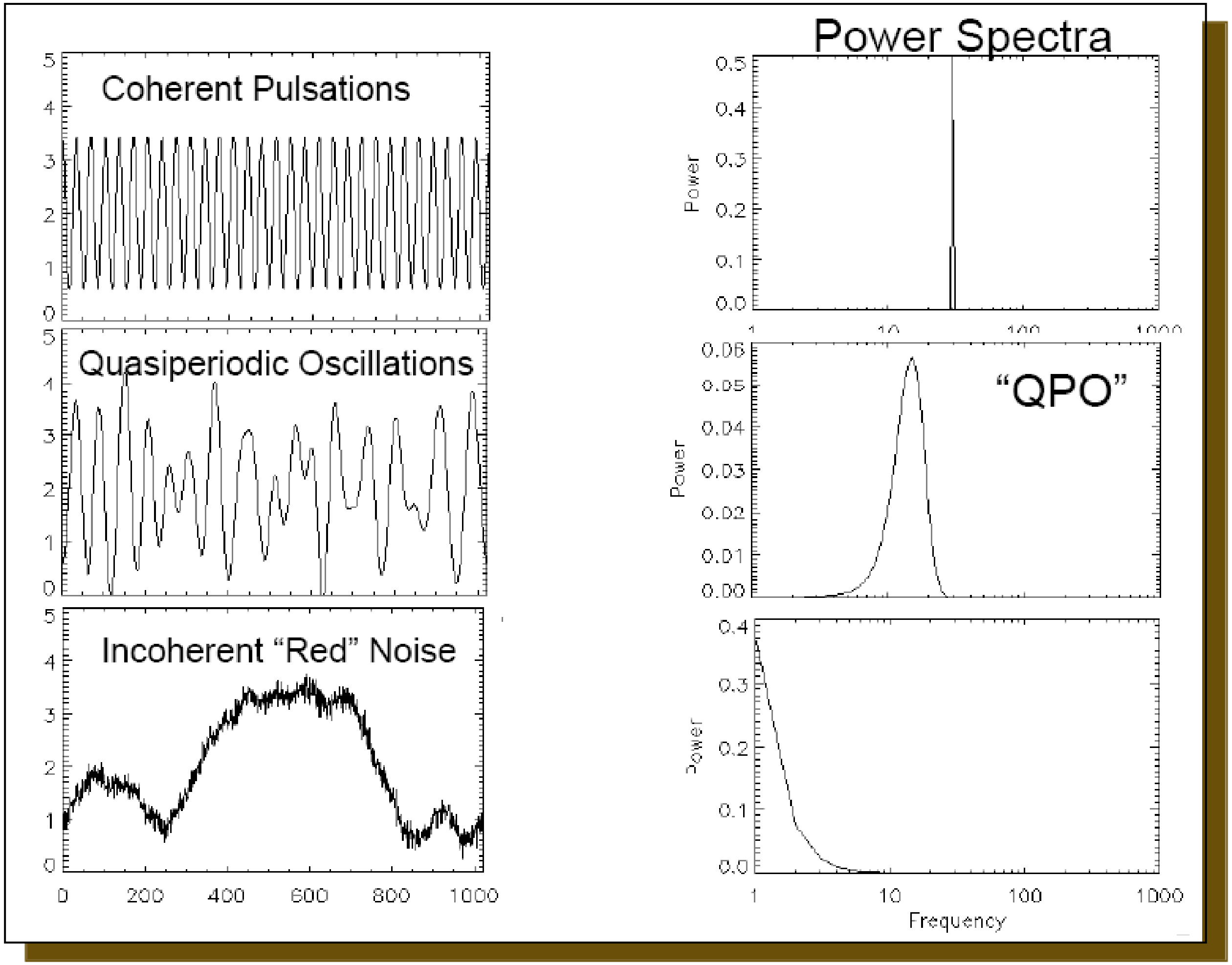
Does My Source Vary?

- **Limits:**
  - This method cannot isolate particular timescales of interest.
  - If we are interested in faster time scales (higher frequencies), we must make a light curve with smaller time bins
- So far, our analysis has focused on the total variability in a light curve.
- TOTAL variability (r.m.s.) does not capture the full info. Its time-scale is important as well.



All light curves have 50% fractional r.m.s. variability

Power Spectrum



Power Spectrum

- **Know Your Normalization!!!**
  - "One-sided" Leahy (mean power = 2): $P_j = 2|a_j|^2/N_{ph}$
  - "One-sided" (RMS/mean)²/Hz: $P_j = 2|a_j|^2/(N_{ph} \times \langle \text{Rate} \rangle)$

- **Uniform sampling:**
  - The highest frequency is the Nyquist frequency, $\nu_{\mathrm{Ny}} = \frac{N}{2T}$
  - Time step, $\delta t = T/N$
  - Frequency step, $\delta\nu = 1/T$

- **Advantages of non-uniform sampling:**
  - Nyquist cutoff is not a hard limit anymore!

What when we don't know the shape of the signal?

- The "Phase Dispersion Minimization" (**PDM**) and similar techniques: bin the data modulo the period. The result is a phase diagram. Then test a chi-square test for flat!

- If you did know the shape of the light curve, you should instead fit the light curve to the phase diagram and test whether the amplitude is consistent with zero.

Detrending

- *Trend* in a time series is a slow, gradual change in some property of the series over the whole interval under investigation.

- Trend is sometimes defined as a long term change in the mean, but can also refer to change in other statistical properties.

- *Detrending* is the statistical or mathematical operation of removing trend from the series.

Detrending

- Identification of trend in a time series is subjective because trend cannot be unequivocally distinguished from low frequency fluctuations. What looks like trend in a short segment of a time series segment often proves to be a low-frequency fluctuation – perhaps part of a cycle – in the longer series.

- We can view the entire observed time series as a segment of an unknown infinitely long series, and cannot be sure that an observed change in mean over that segment is not part of some low-frequency fluctuation imparted by a stationary process.

Decoherence

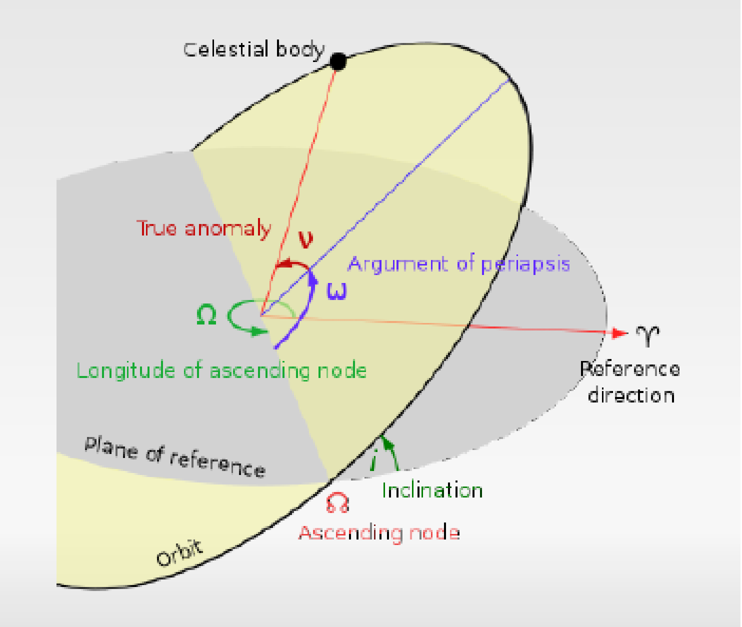
- DFT and simple epoch-folding searches require the single be coherent throughout the interval being considered
- But, this might not be the case because of:
  - Orbital Doppler shifts from a binary system
  - Intrinsic period derivative of the source
  - Satellite or Earth motion that isn't fully compensated for
- Searching still possible with several techniques:
  - Wavelets and Sliding Periodograms!

Decoherence

- Orbital Motion: Doppler effect in a circular orbit: (eccentricity terms have to be added in the real life analysis!)
- Orbital motion easily washes out a fast signal in a close orbital system:

$$\frac{\delta \nu}{\nu} = -\frac{v_{\parallel}}{c} = -\frac{a_1 \sin i}{c} \cdot \frac{2\pi}{P_{orb}} \cdot \cos\left[\frac{2\pi}{P_{orb}}(t - T_{asc})\right]$$

- Use shorter time intervals in this case!



Decoherence

Test: artificial time series consists of seven simulated signals:

(i) The simple sinusoid F_1 with constant semi-amplitude of 0.05 and frequency of $\nu_1 = 0.002$ Hz.

(ii) The sinusoid F_2 with constant semi-amplitude of 0.05, which frequency changes three times by a saltation (0.0145, 0.0150 and 0.0155 Hz) at regular time intervals.

(iii) Two amplitude-modulated sinusoids F_3 and F_4 with frequencies ν_3 and ν_4 of 0.0035 and 0.012 Hz, respectively, and with frequency of amplitude modulations of $\nu_{orb} = 1/P_{orb} = 1.94 \times 10^{-4}$ Hz. The amplitude of F_3 varies sinusoidally from 0 up to 0.05, while the amplitude of F_4 from 0.03 up to 0.06 (Fig. A1):

$$F_3 = \frac{0.05}{4}(1 + \sin 2\pi \nu_{orb} t)(1 + \sin 2\pi \nu_4 t) \qquad \text{((A1))}$$

$$F_4 = \frac{1}{2}(0.045 + 0.015 \sin 2\pi \nu_{orb} t)(1 + \sin 2\pi \nu_5 t). \qquad \text{((A2))}$$

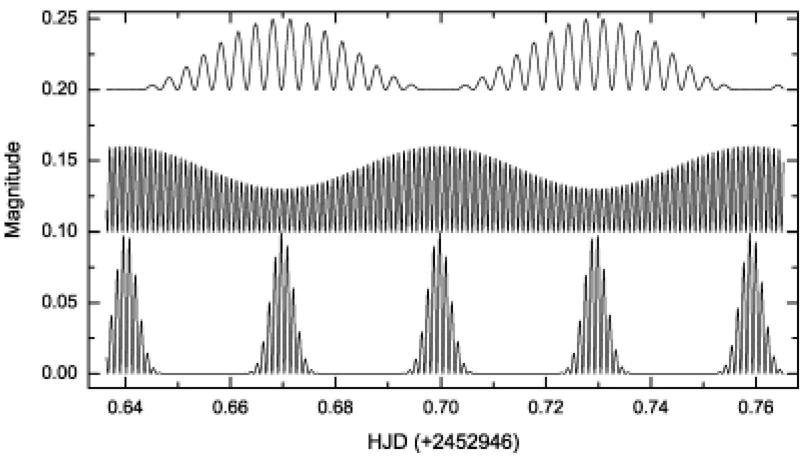
(iv) One more amplitude-modulated sinusoid F_5 with frequency ν_5 of 0.010 Hz and with an amplitude modulation frequency of $2\nu_{orb} = 2/P_{orb} = 3.89 \times 10^{-4}$ Hz. Unlike the previous functions, this sinusoid has visible oscillations (from 0 to 0.10) for about 33 per cent of the time (Fig. A1).We calculated it using the following formula:

$$F_5 = 0.05(\sin^{20} 2\pi \nu_{orb} t)(1 + \sin 2\pi \nu_6 t). \qquad \text{((A3))}$$

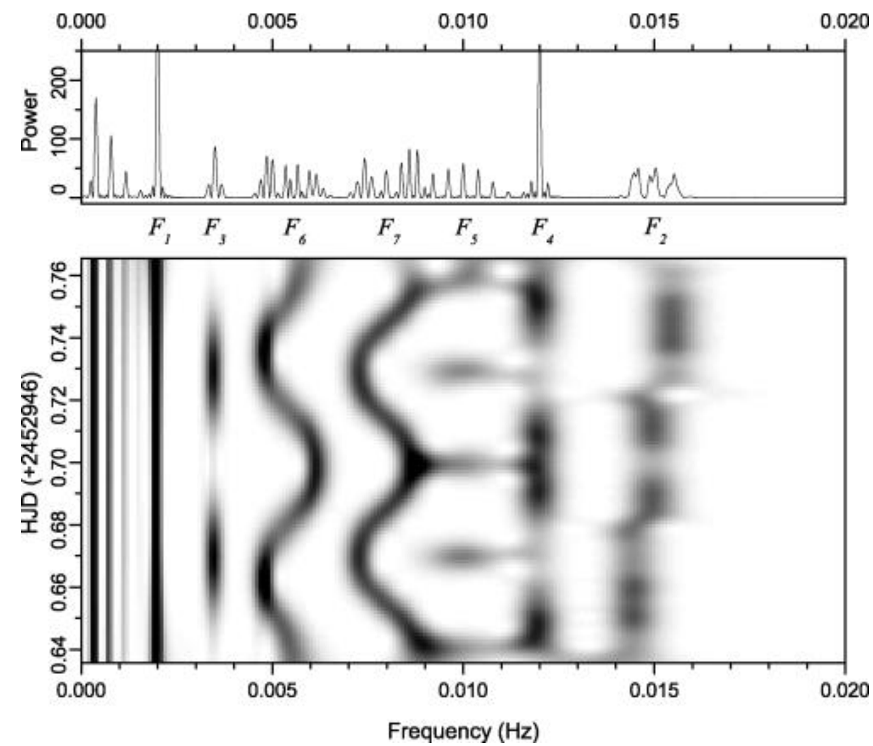
(v) Two frequency-modulated sinusoids F_6 and F_7 with carrier frequencies ν_c of 0.0055 Hz and 0.0080 Hz respectively, and with modulating frequency ν_m of 1.62×10^{-4} Hz = $2/P_{phot}$ and 1.94×10^{-4} Hz = $1/P_{orb}$, respectively. The modulation index δ has been chosen to obtain a maximum frequency deviation of 0.0008 Hz. The amplitude of both sinusoids is 0.05:

$$F_{6,7} = 0.05 \sin(2\pi \nu_c t + \delta \sin 2\pi \nu_m t). \qquad \text{((A4))}$$

Decoherence

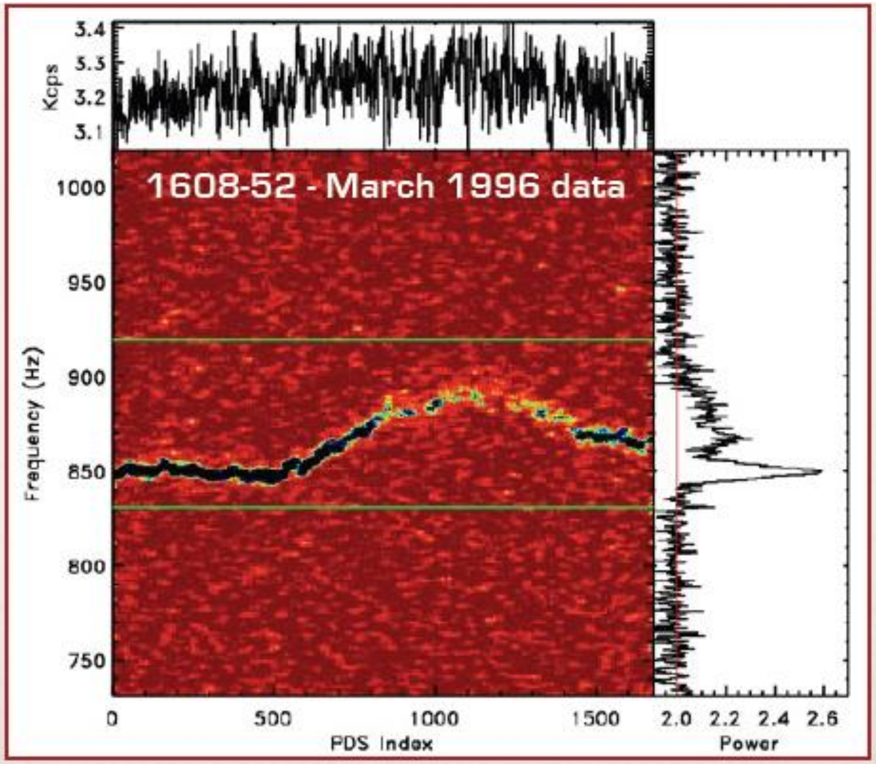


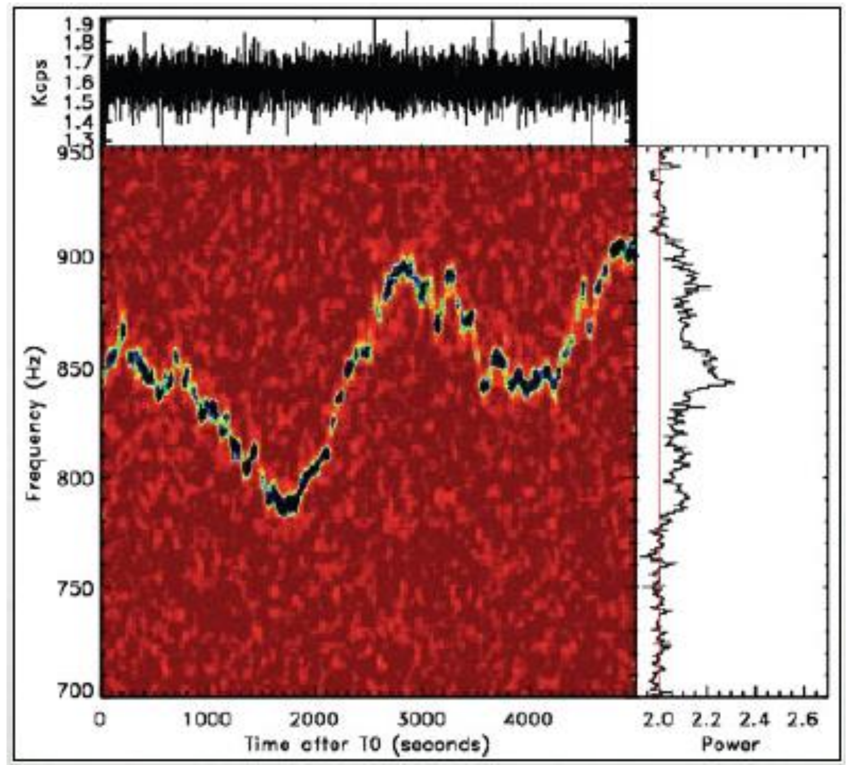
Plots of the test functions F3 (upper), F4 (middle) and F5 (bottom) shifted by 0.1 in the y-direction.



The Lomb–Scargle power spectrum (upper panel) and the scalogram (bottom panel) of the artificial time series. Frequency position of test functions is shown by F_1, ..., F_7.

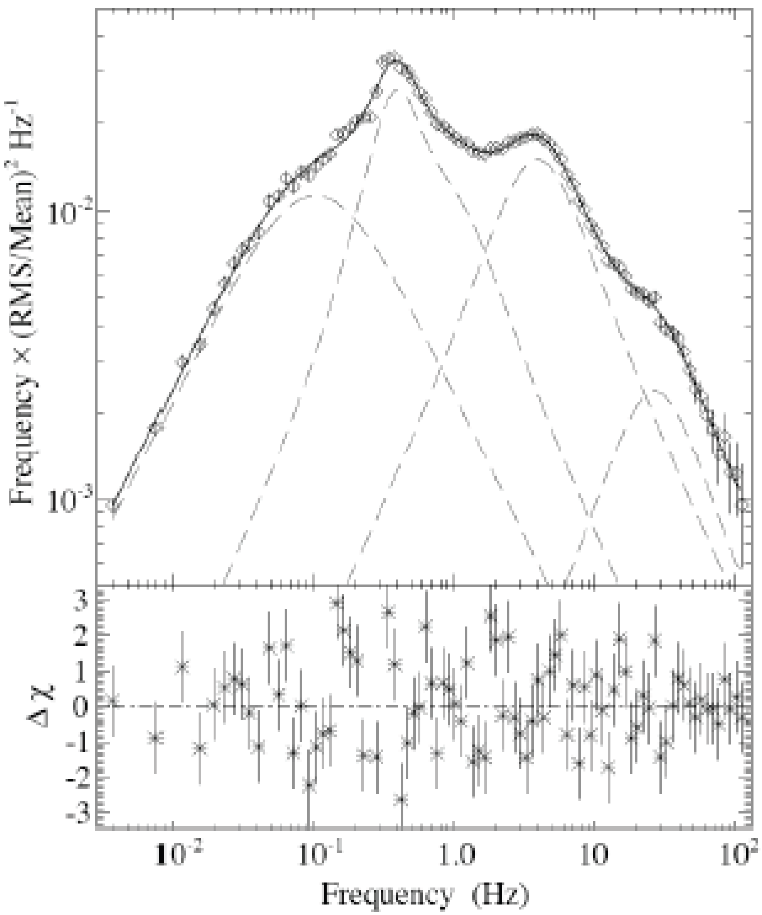
Decoherence



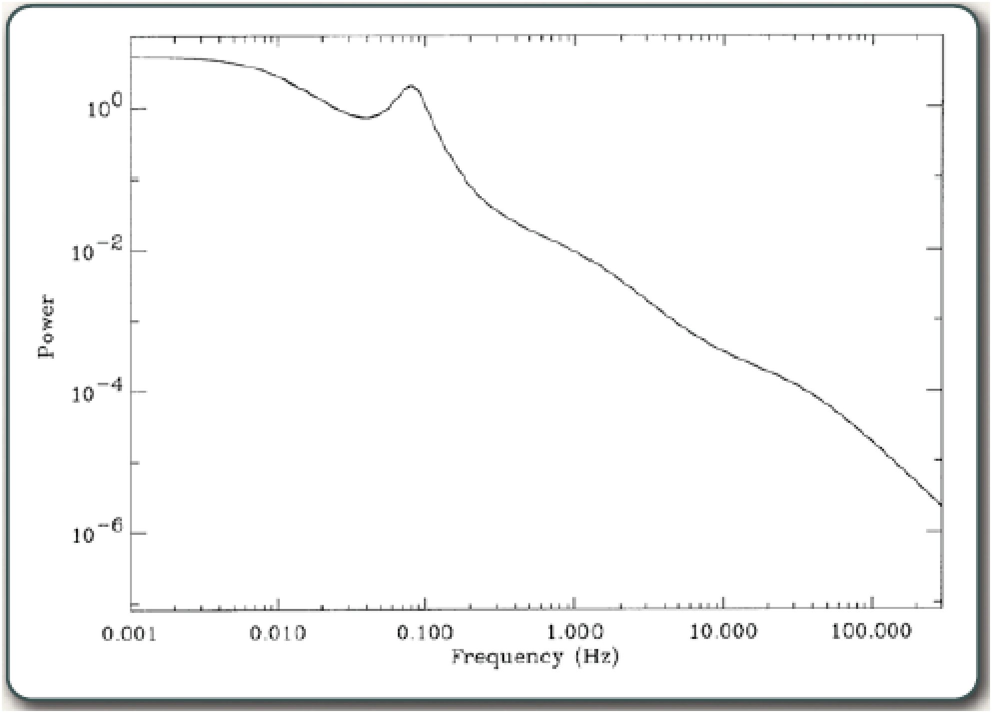


PSD Model Fitting

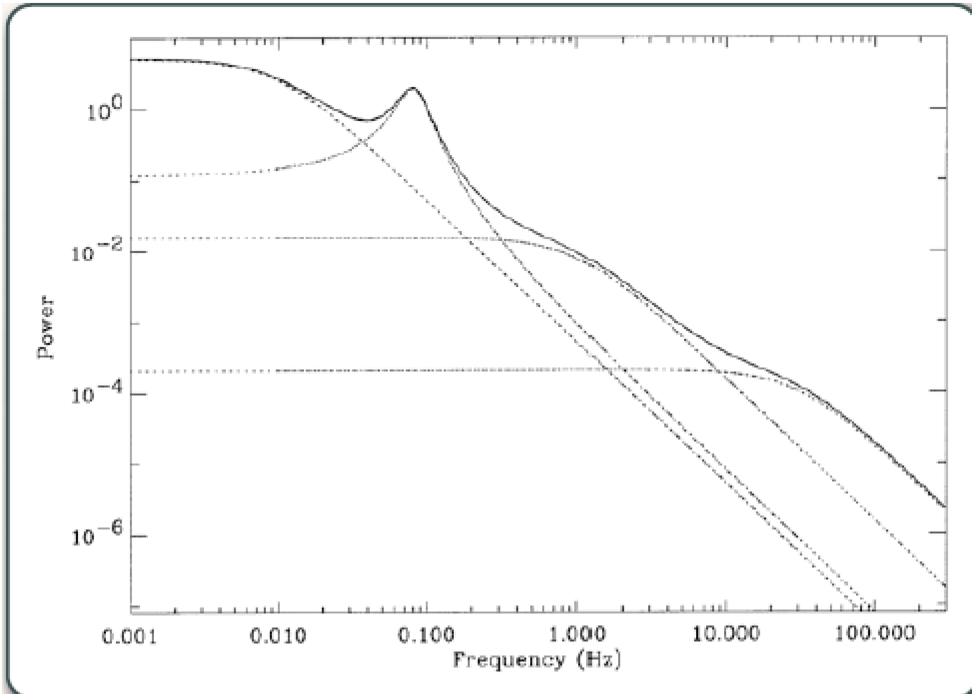
- After subtracting Poisson level, you can fit models.
- Popular choice currently is a sum of Lorentzians



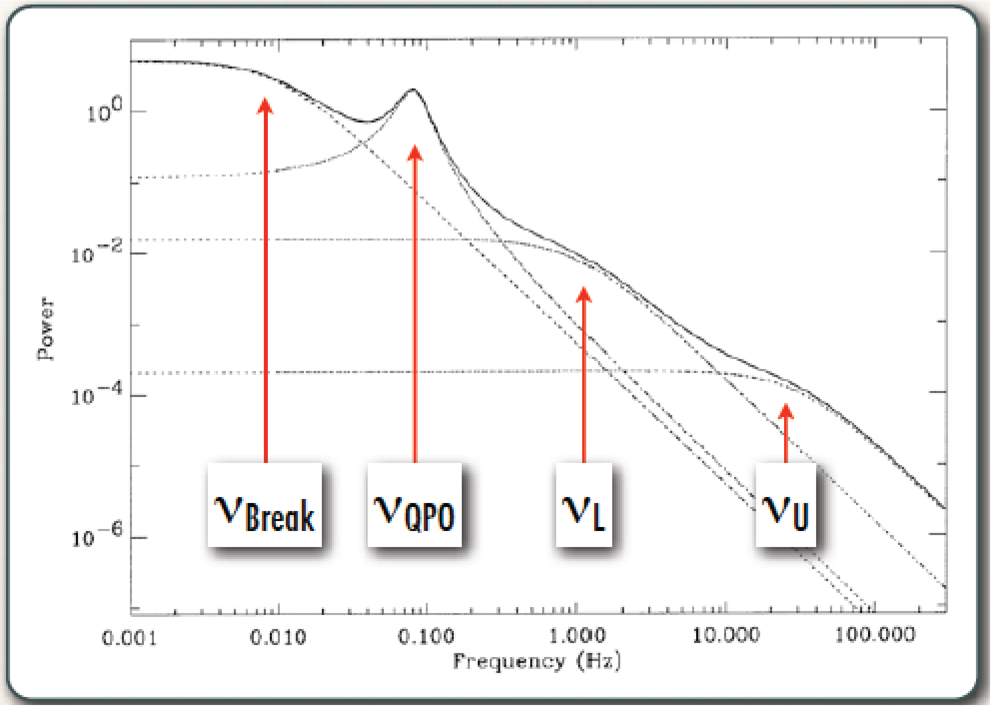
PSD Model Fitting



PSD Model Fitting



PSD Model Fitting



The END