A Comparison of Various Algorithms for the Spectral Analysis of Unevenly Spaced Data Series

Michael R. Craymer and Petr Vaníček

Geodetic Research Laboratory, Dept. of Surveying Engineering
University of New Brunswick, Fredericton, New Brunswick, E3B 5A3

CISM/CGU'90, Ottawa, 22-25 May 1990

The Problem

Interpolation Approach

Least Squares Approach

Hybrid Approach

Comparisons

Conclusions

The Problem - Spectral Analysis

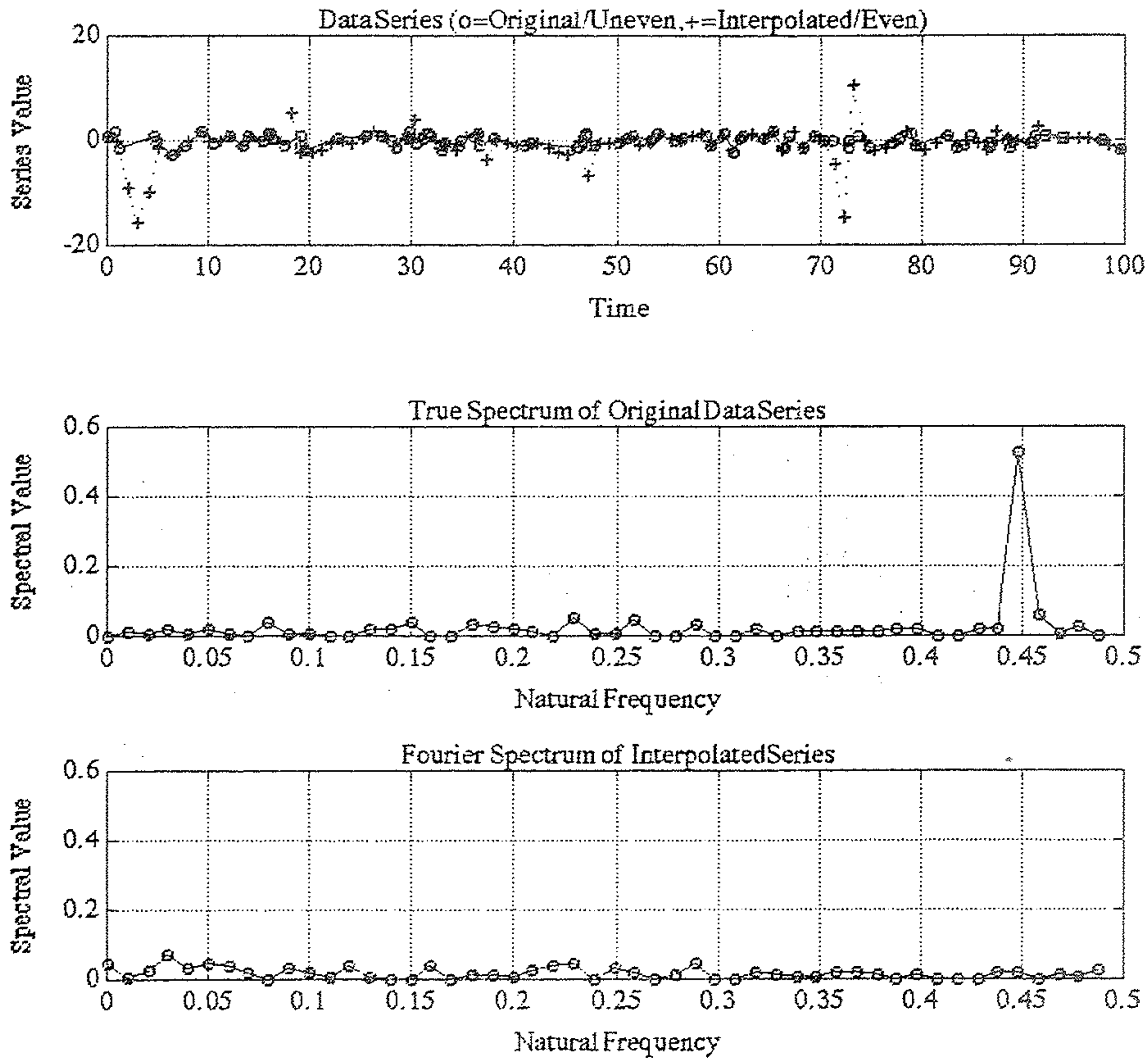
- Uneven data spacing
  - » Missing data points (gaps)
  - » Randomly spaced series
- Presence of signals
  - » Datum biases, linear trends, known periodic trends (tides)
  - » Should be estimated along with spectral analysis
- Fourier spectral analysis algorithms require
  - » Evenly spaced data series
  - » Integer multiples of the fundamental frequency (discrete)
  - » 2^n data points (FFT algorithm only)
  - » No interaction (correlation) between estimated signals and spectral values
  - » Doesn't account for finiteness of data (assumes infinite length series)

Interpolation Approach

- Interpolation to evenly spaced series (creates new data series)
- FFT of interpolated (evenly spaced) data series
- Accuracy depends on
  - » Form of interpolation function
  - » Smoothness of original series
  - » Presence of large data gaps

Problems

- Spectrum for interpolated data series, <u>not original series</u>
- Interpolation
  - » Smooths high frequency part of data series
  - » No way of assessing accuracy
  - » Very poor in presence of data gaps
- FFT
  - » Most algorithms limited to 2^n data points
  - » Only "integer frequencies" (discrete spec.)
  - » Can't zoom in to precisely locate new peaks


DataSeries (o=Original/Uneven,+=Interpolated/Even)
Series Value
20
0
-20
0
10
20
30
40
50
60
70
80
90
100
Time
True Spectrum of Original DataSeries
Spectral Value
0.6
0.4
0.2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Natural Frequency
Fourier Spectrum of InterpolatedSeries
Spectral Value
0.6
0.4
0.2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Natural Frequency

Least Squares Spectral Analysis (Vaníček)

- Direct least squares fitting of Fourier series terms to data
- Spectral values = relative variance explained by fitting each frequency (percent of total variance explained)
- Provides spectral values at any frequency - *continuous*
- Correlation among observations - include a weight matrix
- Presence of known signals - include additional functions when fitting each frequency
- Fourier spectrum just a special case
    - » Evenly spaced data
    - » "Integer" frequencies (discrete spectrum)

- Simultaneous evaluation of all Fourier terms (all at once)

  » Can result in singularities for some frequencies
  » Requires A LOT OF RAM
  » Very slow

- Separate evaluation of Fourier terms (one at a time) - LSSA

  » Avoids singularities
  » Little memory required
  » Faster
  » Significant spectral frequencies estimated simultaneously as signals to be removed

Advantages

- Correct spectrum - not affected by finiteness of data and uneven spacing
- Spectral values for any frequencies (continuous spectrum) - => Can zoom in on peaks
- Accounts for interaction (correlation) of known signals with spectral frequencies

LSSA Algorithms

- Based on different numerical procedures

- Direct inversion of normal equations (Vaníček, 1969) - LSSA
  - » Most general
  - » Accounts for presence of estimated signals

- Orthogonalization via time shifting (Lomb, 1975)
  - » Shift time so that normal eqns are diagonal
  - » Different time shift for each frequency [redacted].
  - » Doesn't account for estimated signals
  - » Slower than direct inversion

- Gram-Schmidt orthogonalization (Ferraz-Mello, 1981)
  - » Doesn't account for estimated signals

- All are mathematically equivalent under same assumptions (no estimated signals)

Hybrid FFT-LSSA Approach (Press & Rybicki)

- Based on "extirpolation" & FFT - FASPER (very fast)
- Extirpolation
  - » Find an equally spaced series that gives the original unequally spaced one after interpolation (reverse interpolation)
- FFT
  - » Used to evaluate evenly spaced (extirpolated) sine and cosine summations
  - » Applied to time-shifting algorithm
  - » Original algorithm requires two complex FFTs
  - » Can be reduced to only one FFT using a trigonometric identity

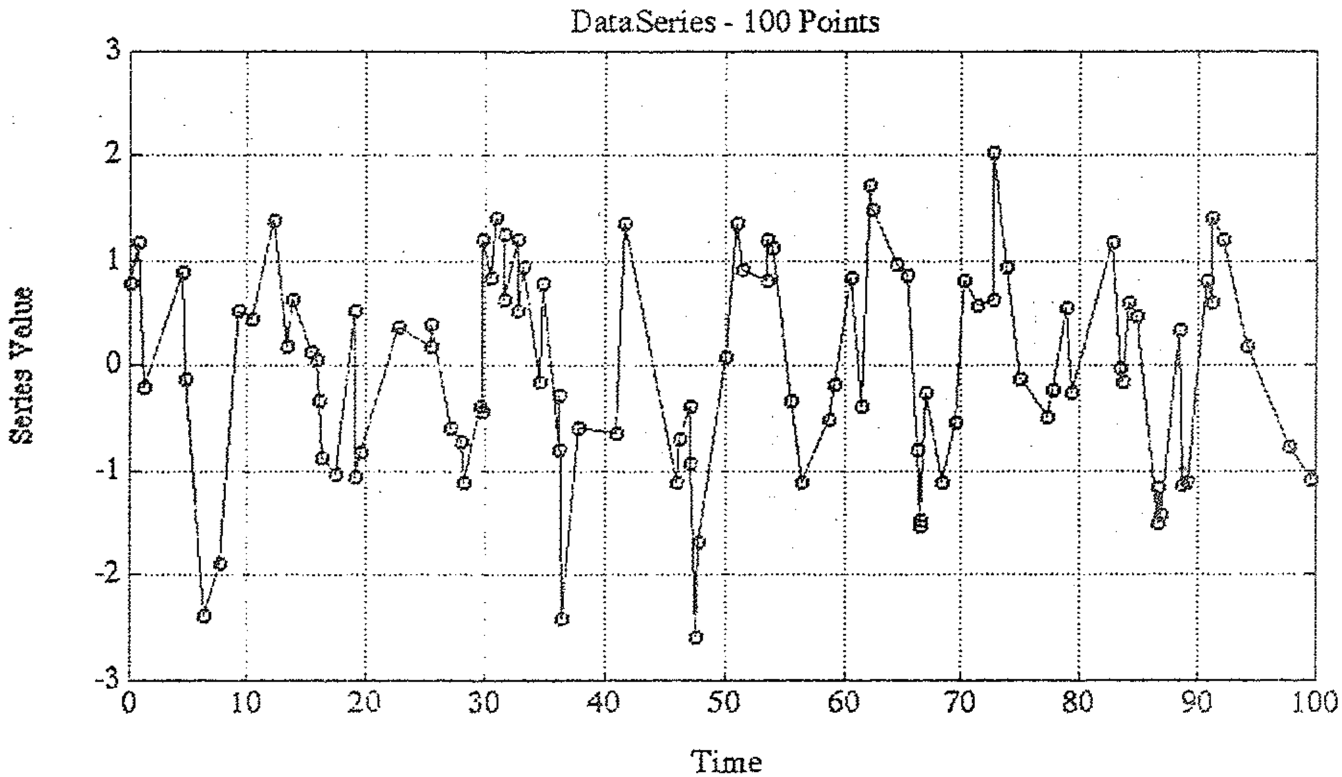
Disadvantages

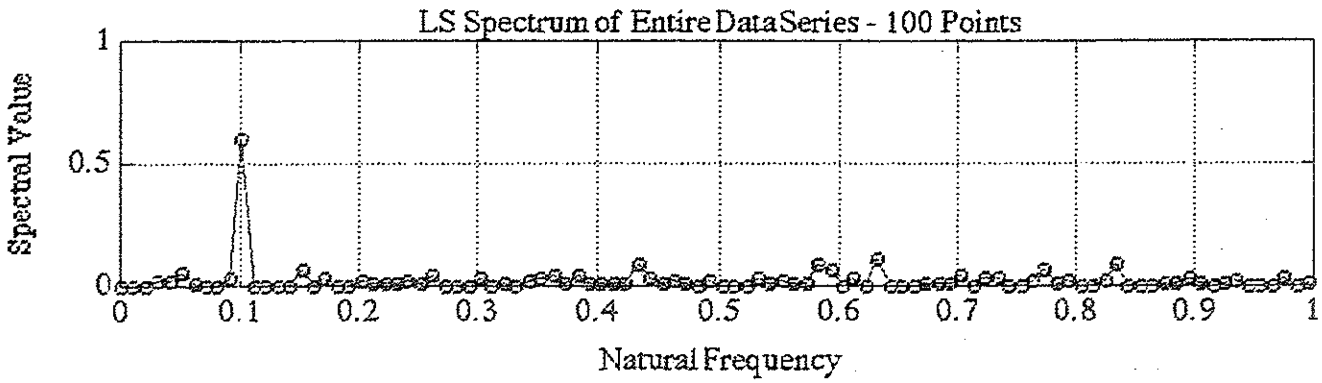
- » Extirpolation accuracy unknown
- » Same as FFT (2^n data points, "integer" frequencies)
- » Doesn't account for estimated signals

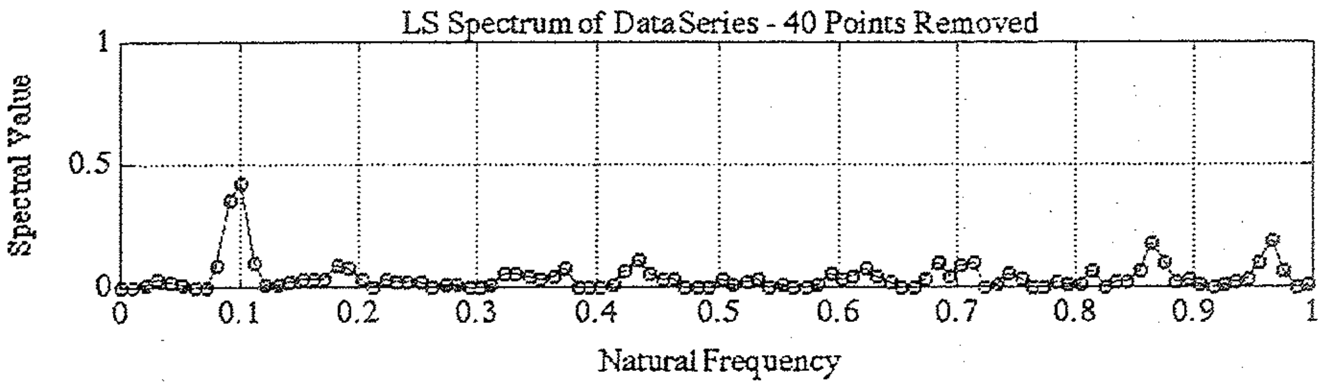
Example

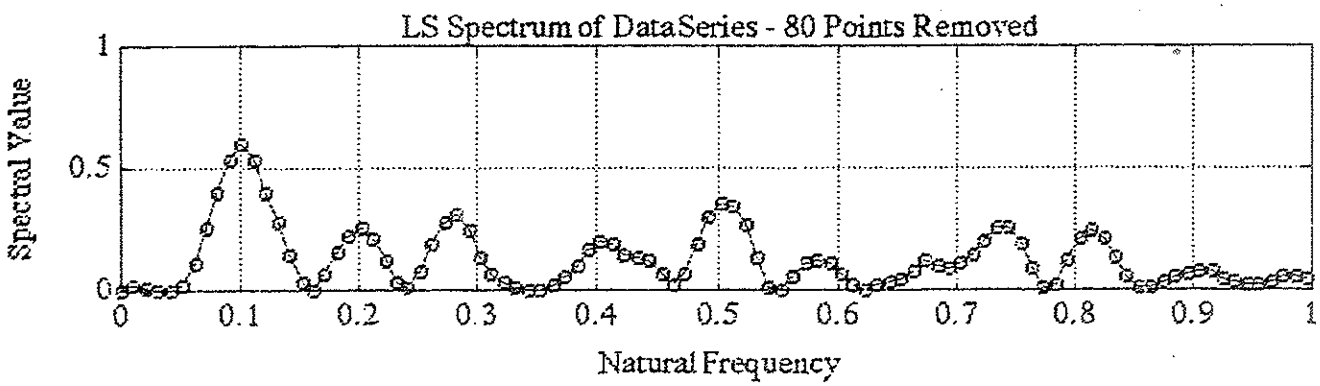
Periodic signal
- » Period = 10
- » Magnitude = ±1

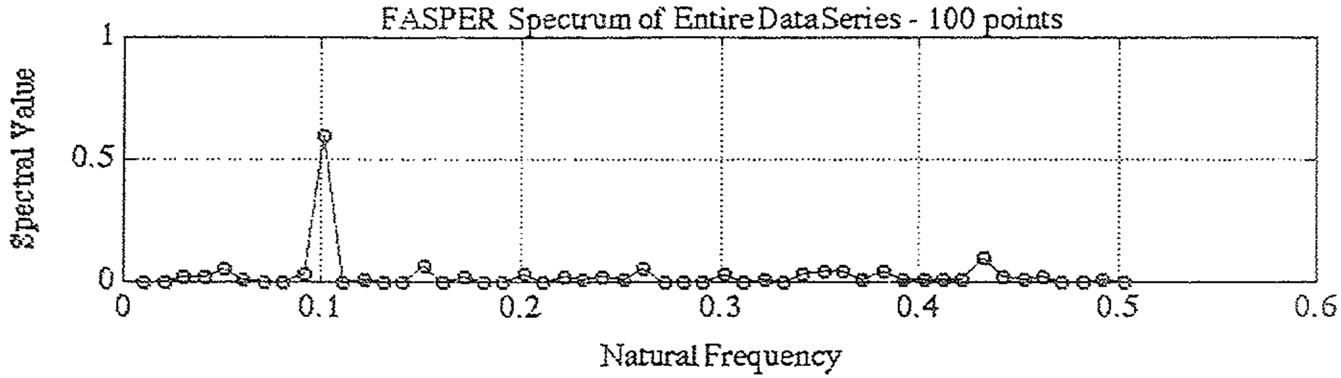
Pseudo-random noise
- » Mean = 0
- » Standard deviation = 2/3

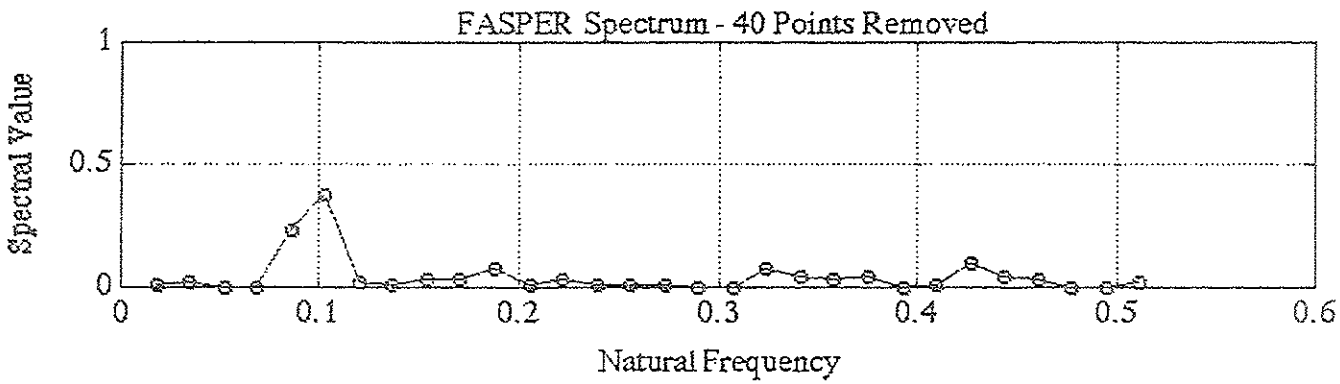


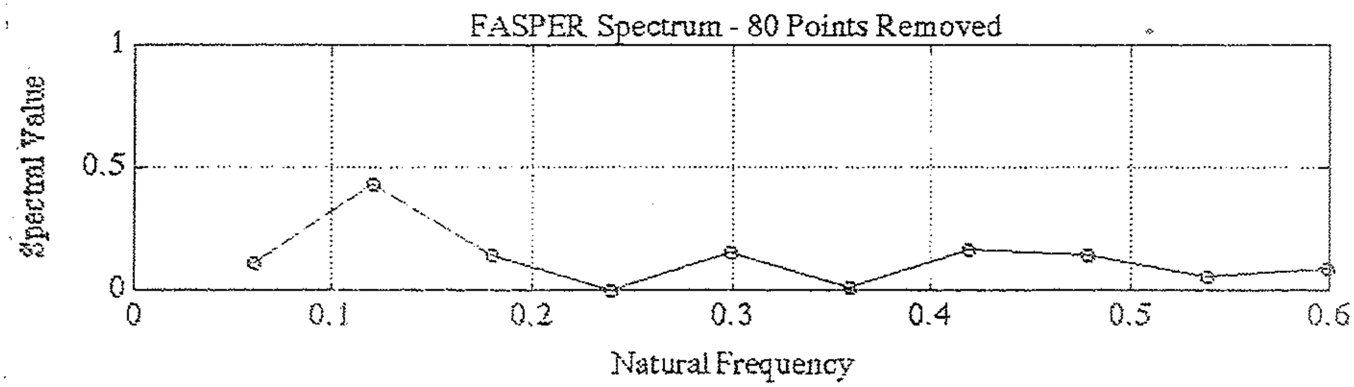

LS Spectrum of Entire DataSeries - 100 Points
Spectral Value
1
0.5
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Natural Frequency


LS Spectrum of DataSeries - 40 Points Removed
Spectral Value
1
0.5
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Natural Frequency


LS Spectrum of DataSeries - 80 Points Removed
Spectral Value
1
0.5
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Natural Frequency


FASPER Spectrum of Entire Data Series - 100 points
Spectral Value
1
0.5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
Natural Frequency


FASPER Spectrum - 40 Points Removed
Spectral Value
1
0.5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
Natural Frequency


FASPER Spectrum - 80 Points Removed
Spectral Value
1
0.5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
Natural Frequency

Conclusions

- LSSA
  - » Best results - most reliable
  - » Provides a *continuous spectrum*
  - » Accounts for presence of estimated signals
  - » Relative slow
- FASPER
  - » Very fast (especially with extra trig. identity)
  - » Must be very careful - accuracy of extirpolation unknown
  - » Provides only a *discrete spectrum*
  - » More research needed to ascertain limitations
- Interpolation - Very poor => Forget it!

Recommend

- » Use LSSA whenever possible
- » Exercise great caution when using FASPER
- » At least use LSSA as a check on FASPER

Software

- LSSA

  » Wells, Vaníček and Pagiatakas (1985). Least Squares Spectral Analysis Revisited Surveying Engineering Technical Report No. 84.

  » Available for Macintosh, IBM PC and mainframe

- FASPER

  » Press and Rybicki (1989). Fast algorithm for spectral analysis of unevenly sampled data. The Astronomical Journal, Vol. 338, pp. 277-280.

  » Paper contains FASPER routine

  » Uses some Numerical Recipes routines