The Periodogram and its Statistical Properties

S.K. Avery and S.E. Palo
University of Colorado, Boulder

I. Periodogram

II. Spectral Estimation Problem

III. Probability Distributions

IV. Confidence Intervals and Significance Levels

V. Summary

Periodogram

Objective: To determine the spectral content of a random process based on a finite set of observations from that process.

A. Schuster, 1897: Lunar and solar periodicities of earthquakes, Proc. Roy. Soc.

Definitions.

Power Spectral Density

$$PSD(f) = \lim_{M\to\infty} E\left[\frac{1}{2M+1} \left| \sum_{n=-M}^{M} x(n) \exp(-j2\pi fn)\right|^2\right]$$

continuous and periodic in frequency with period $f_s = \frac{1}{\Delta t}$

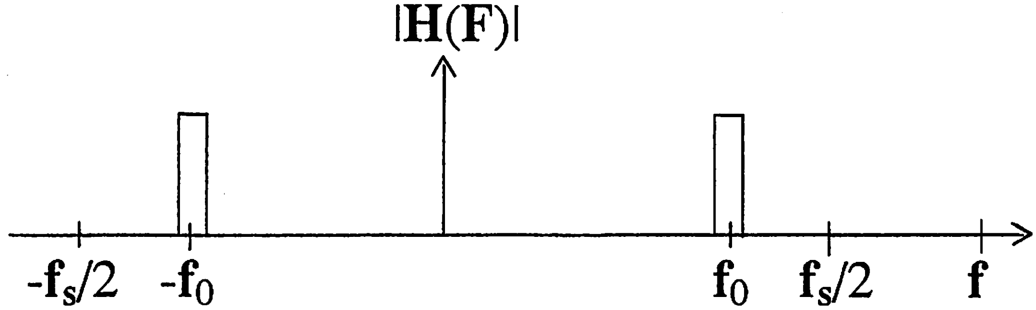
Periodogram

$$P(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) \exp(-j2\pi fn)\right|^2$$

Periodogram is continuous function and is proportional to magnitude squared of DTFT over range $-f_s/2 \le f \le f_s/2$ **divided and normalized by number of independent frequencies, N.**

Interpretation

$x(n) \rightarrow$ [$H(f)$] $\rightarrow y(n)$



Bandpass filter centered at $f = f_0$ with narrow filter bandwidth. Measure output and compute magnitude squared. Divide by filter bandwidth to obtain density. Repeat for all $-f_s/2 \leq f_0 \leq f_s/2$ Note filter bandwidth can't be smaller than $\frac{1}{N\Delta t}$

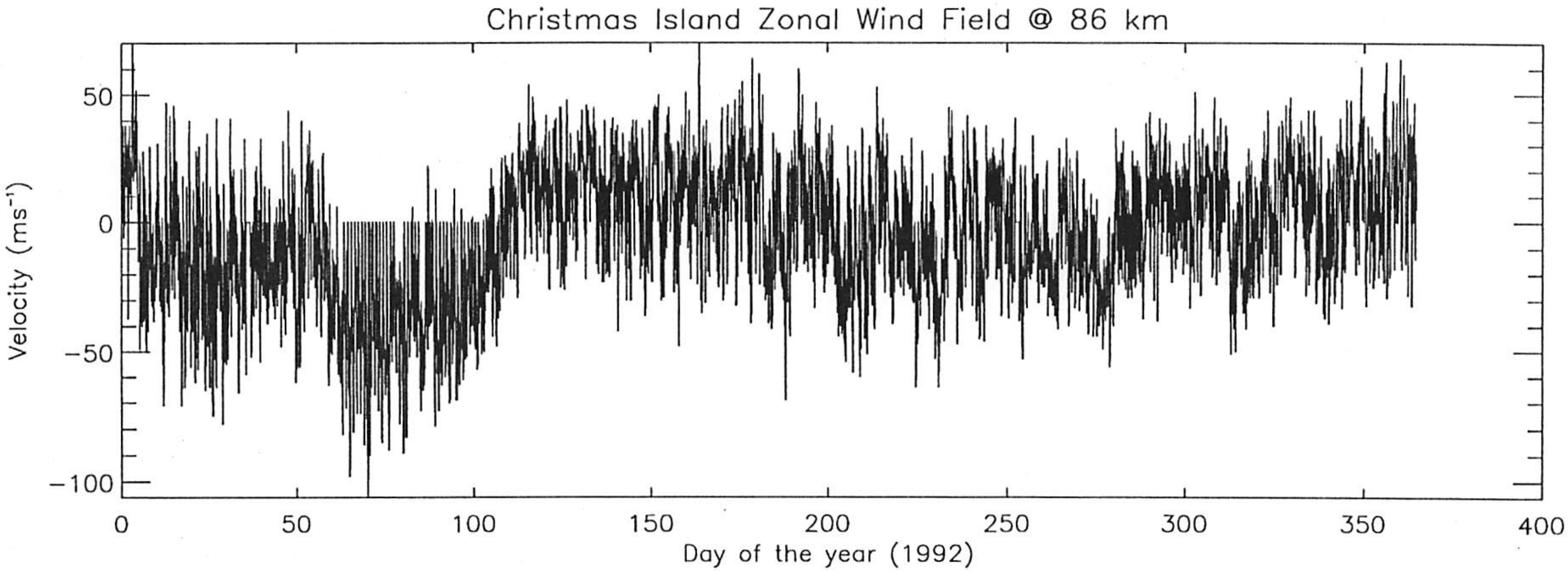
Note that in reality have DFT i.e.

$$X(k) = \sum_{n=0}^{N-1} x(n)\, e^{-j(2\pi/N)kn}$$

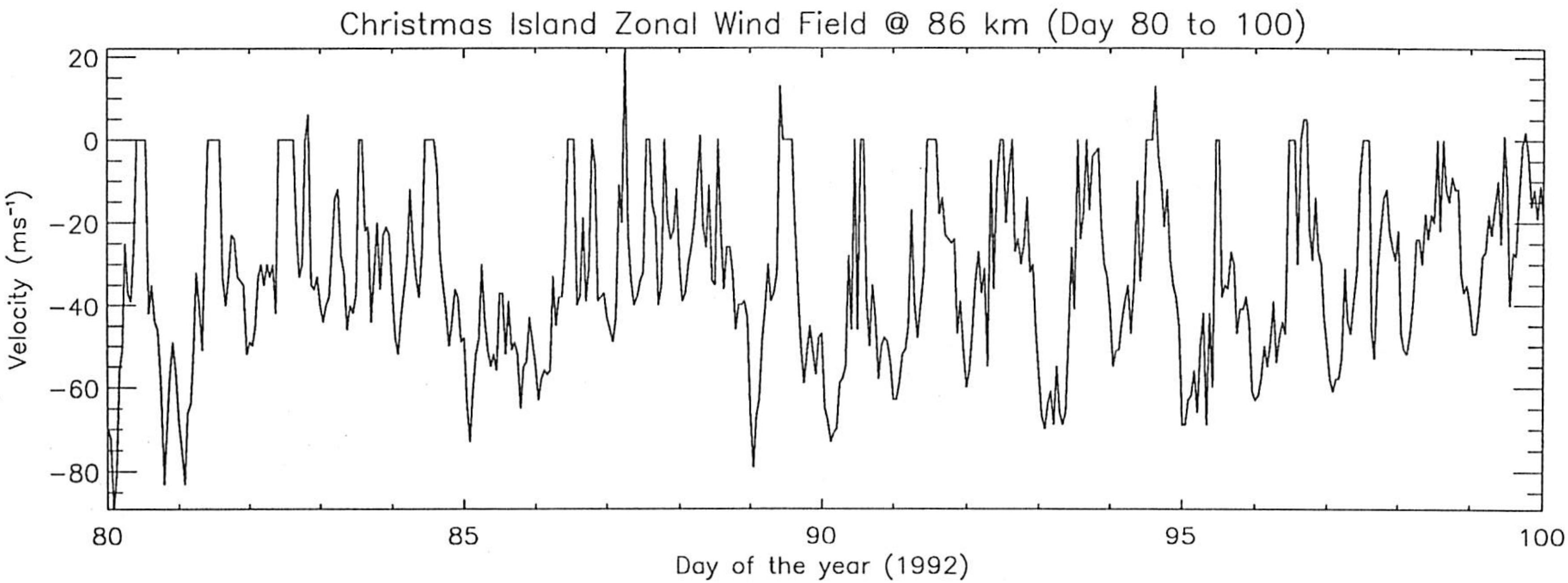
$$P(k) = \frac{1}{N} |X(k)|^2$$

i.e. samples of power spectral density at intervals $\frac{1}{N\Delta t}$ but interpreted as continuous function.

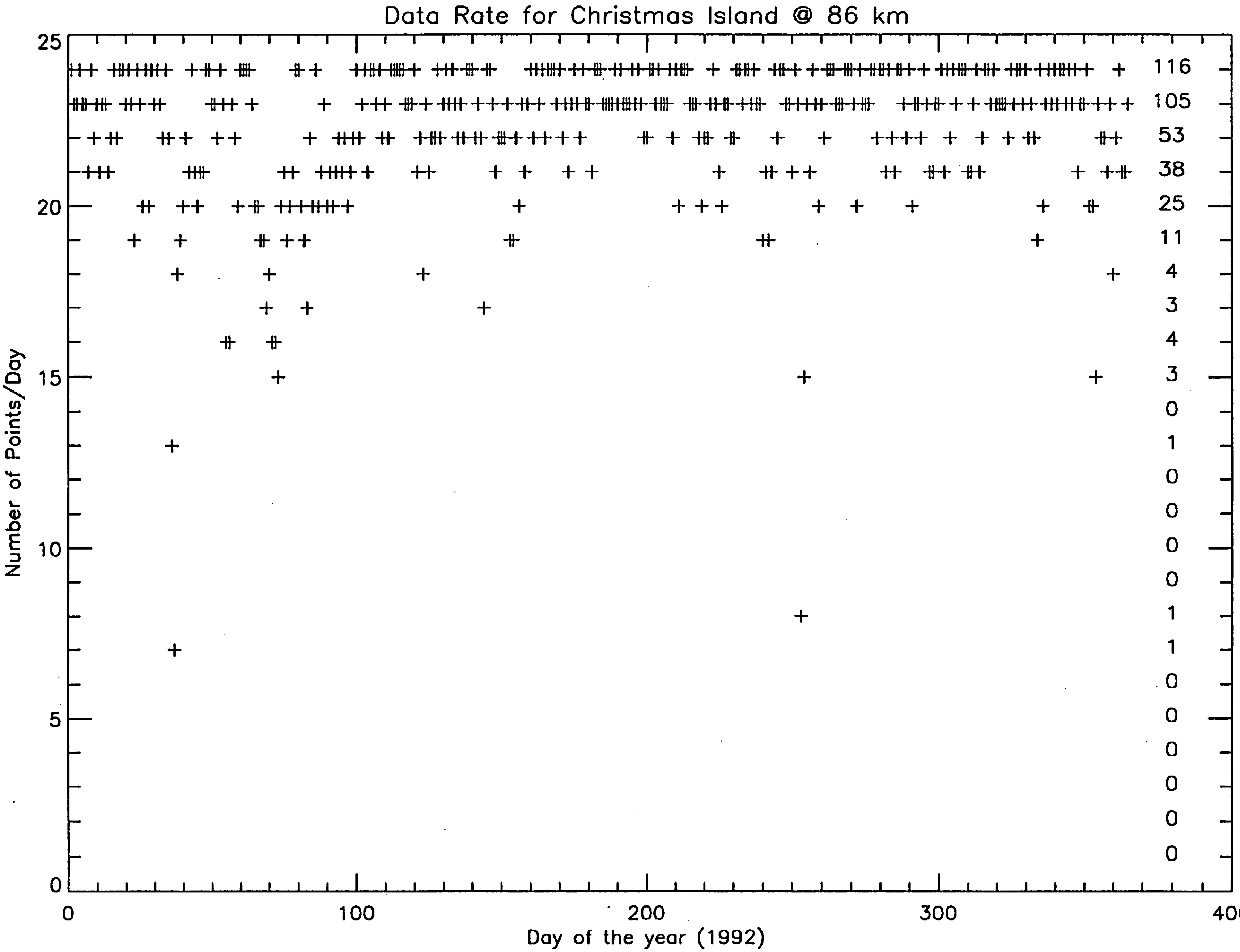
As $N \rightarrow \infty$ mean of P(k) $\rightarrow$ mean of power spectral density. However because have only one sample from narrowband filter to obtain power need to average many periodograms in order for variance of the periodogram $\rightarrow$ 0.



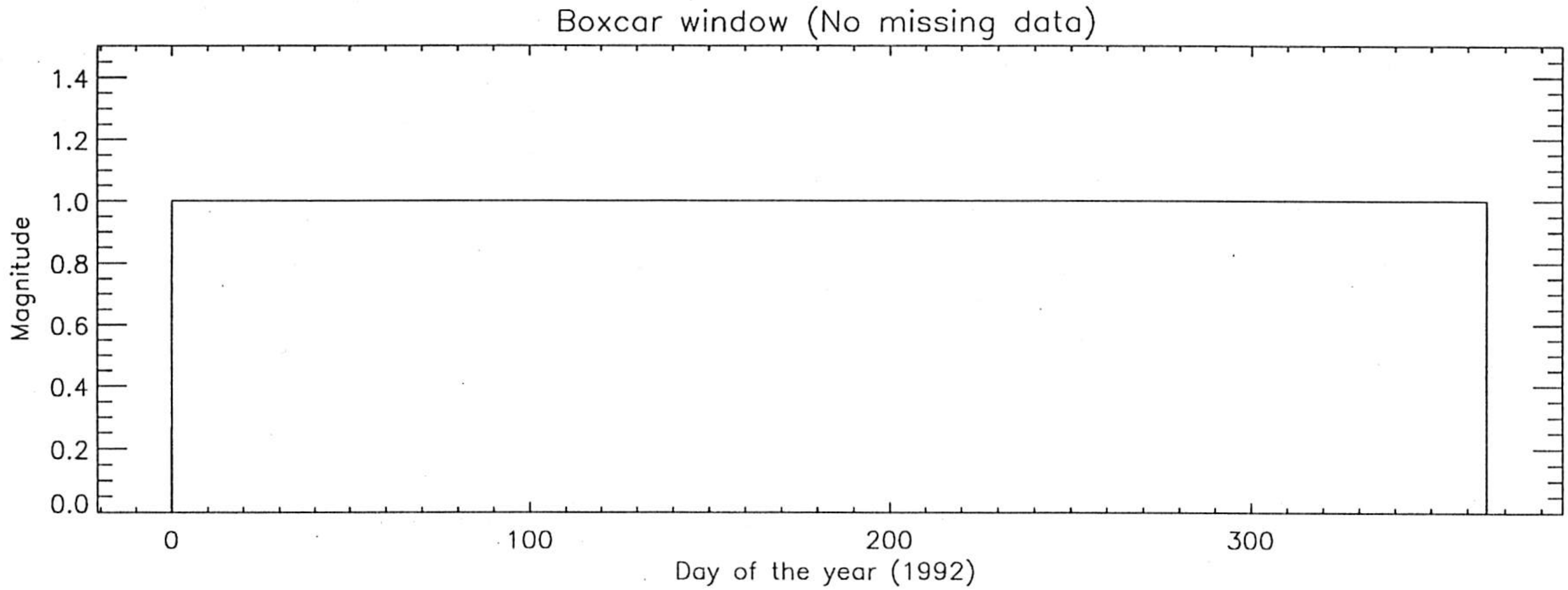
Christmas Island Zonal Wind Field @ 86 km
Velocity (ms⁻¹)
50
0
−50
−100
0
50
100
150
200
250
300
350
400
Day of the year (1992)

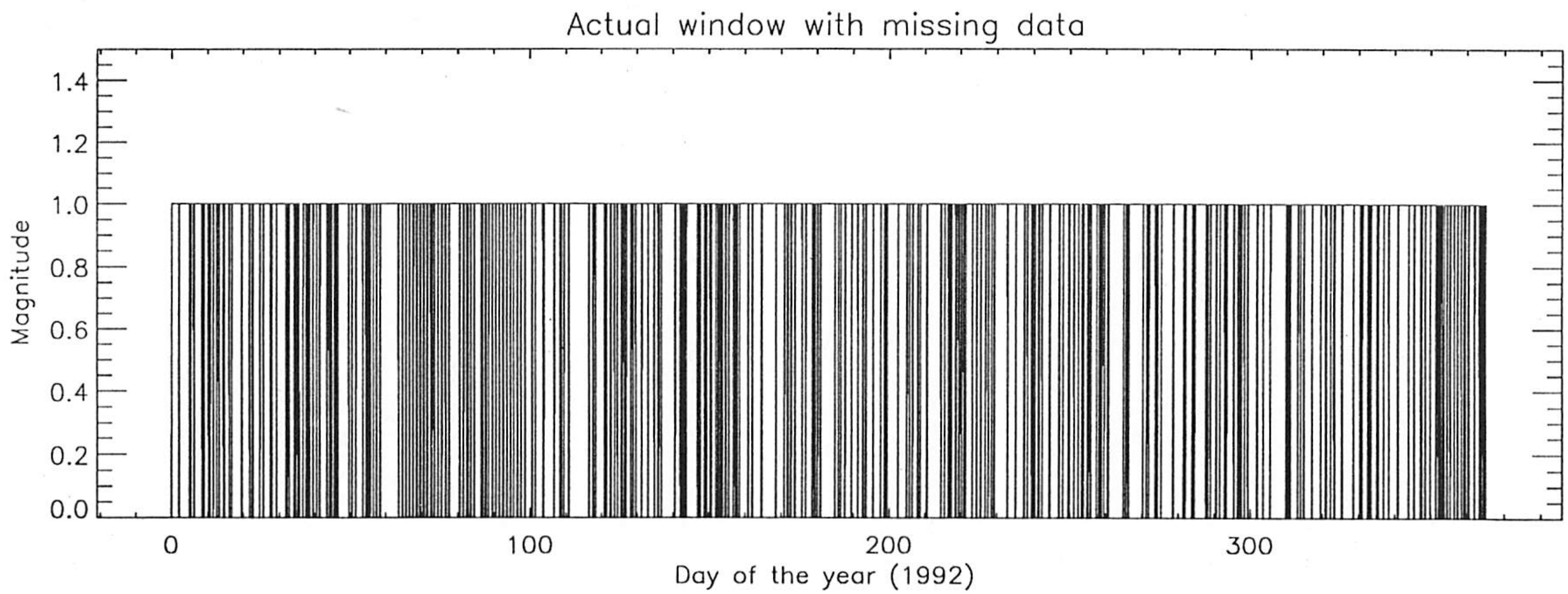


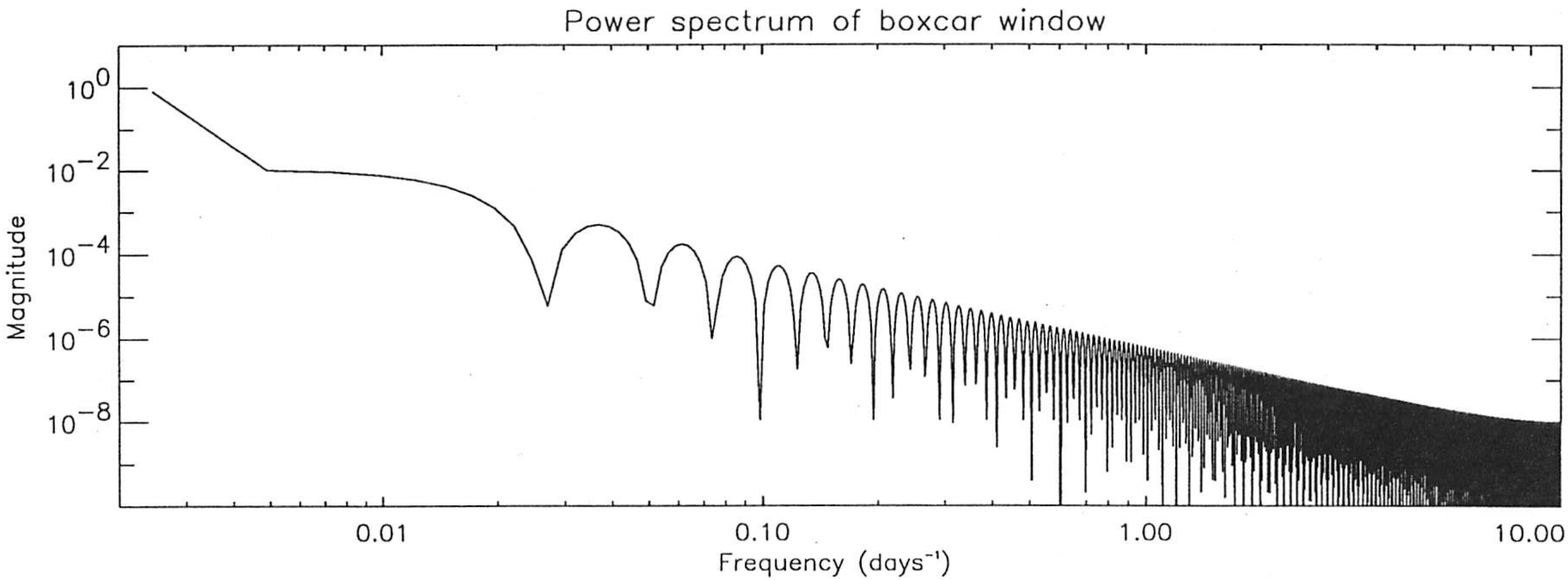
Christmas Island Zonal Wind Field @ 86 km (Day 80 to 100)
Velocity (ms⁻¹)
20
0
−20
−40
−60
−80
80
85
90
95
100
Day of the year (1992)



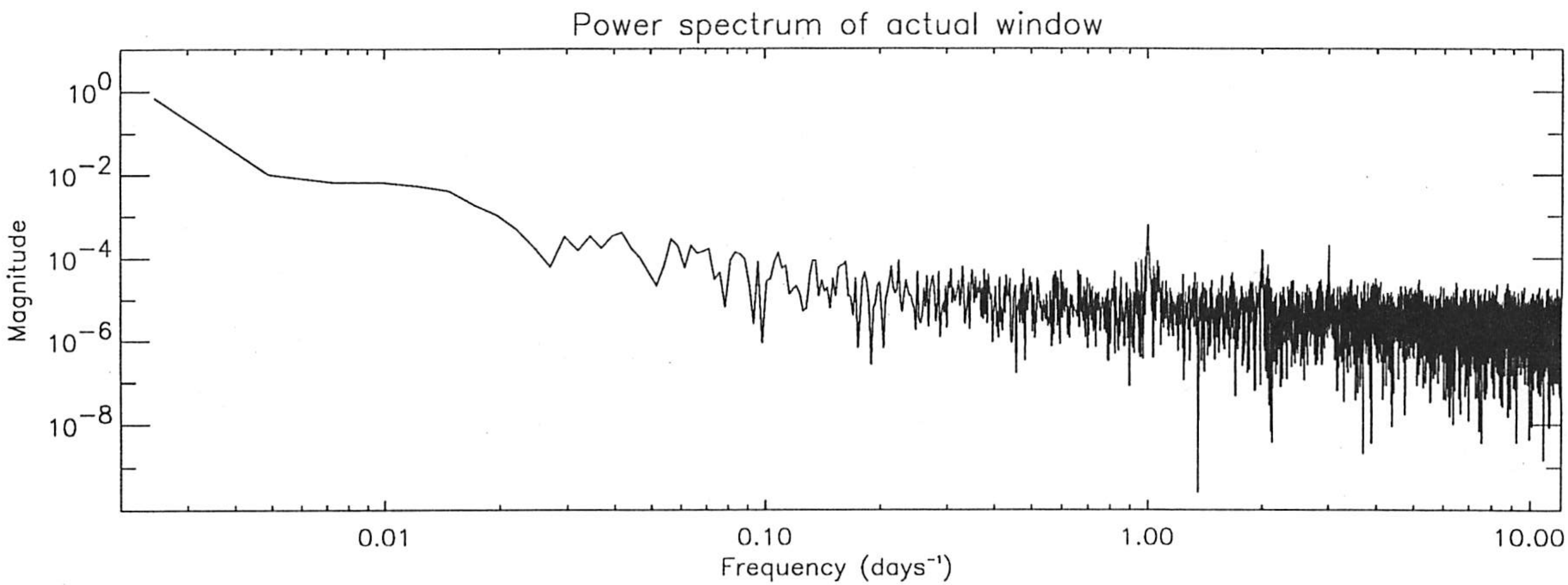
Data Rate for Christmas Island @ 86 km
Number of Points/Day
Day of the year (1992)
0
100
200
300
400
5
10
15
20
25
116
105
53
38
25
11
4
3
4
3
0
1
0
0
0
0
1
1
0
0
0
0
0
0


Boxcar window (No missing data)
Magnitude
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
Day of the year (1992)

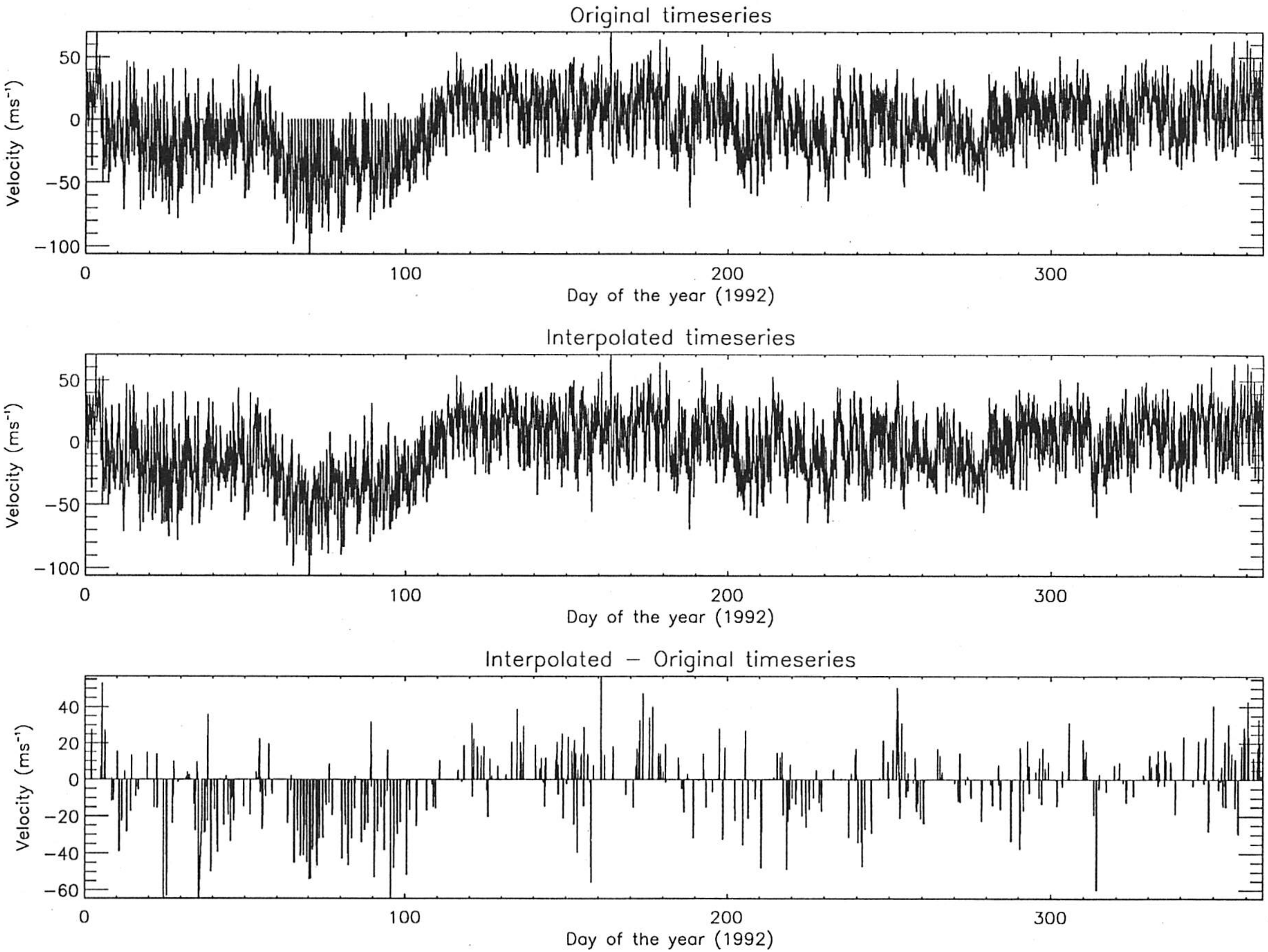

Actual window with missing data
Magnitude
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
Day of the year (1992)



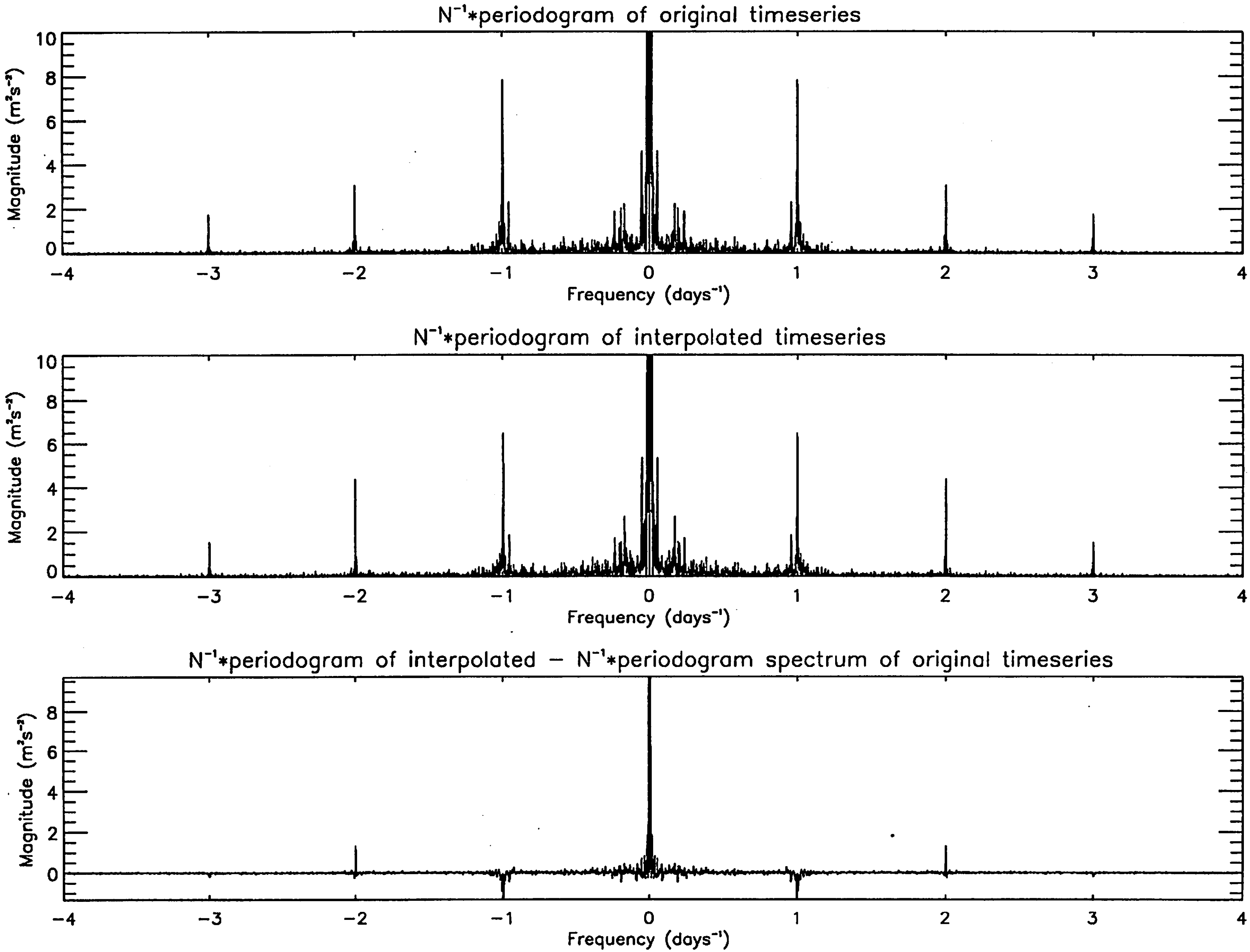
Power spectrum of boxcar window
Magnitude
Frequency (days⁻¹)
0.01
0.10
1.00
10.00



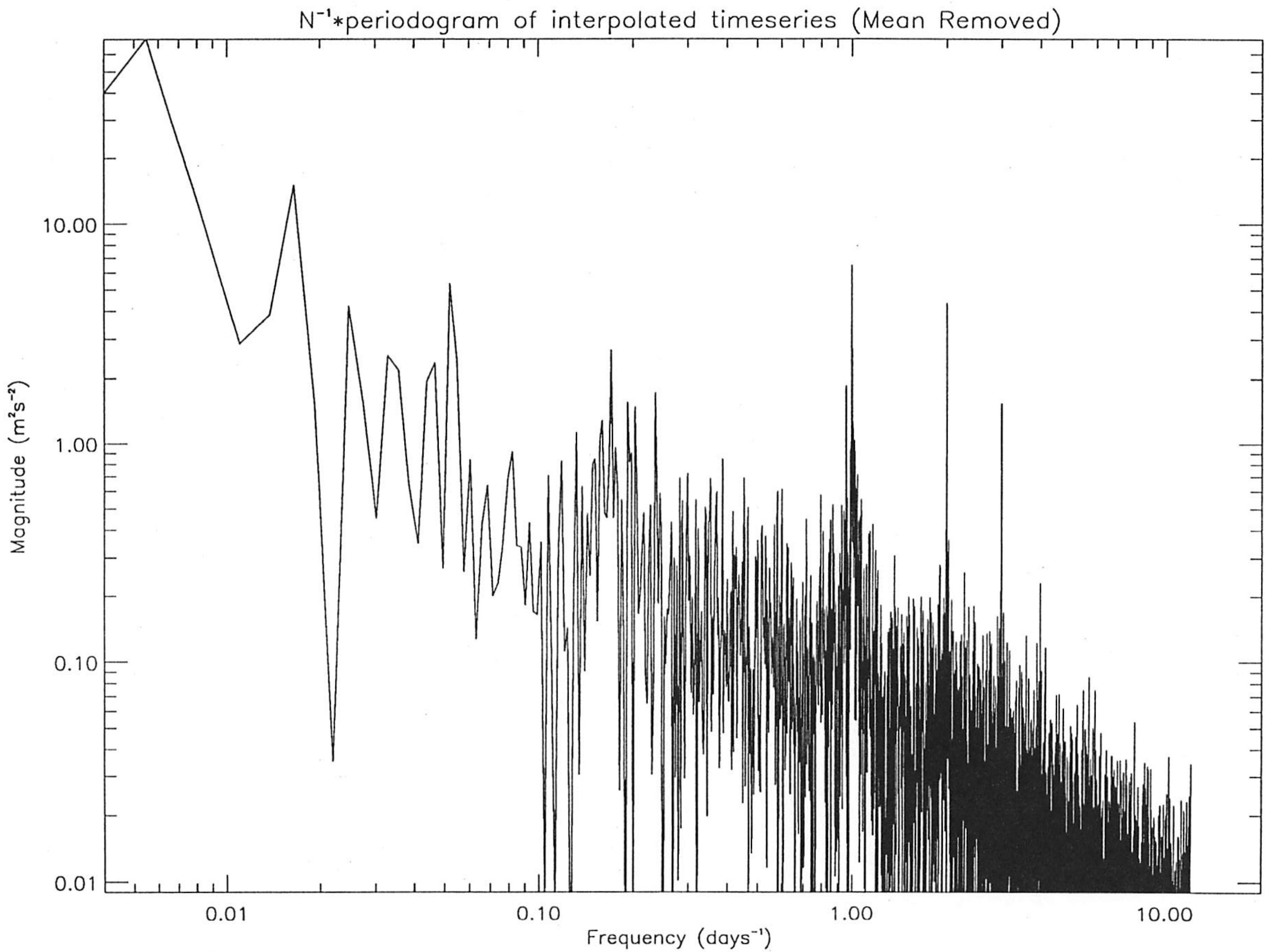
Power spectrum of actual window
Magnitude
Frequency (days⁻¹)
0.01
0.10
1.00
10.00



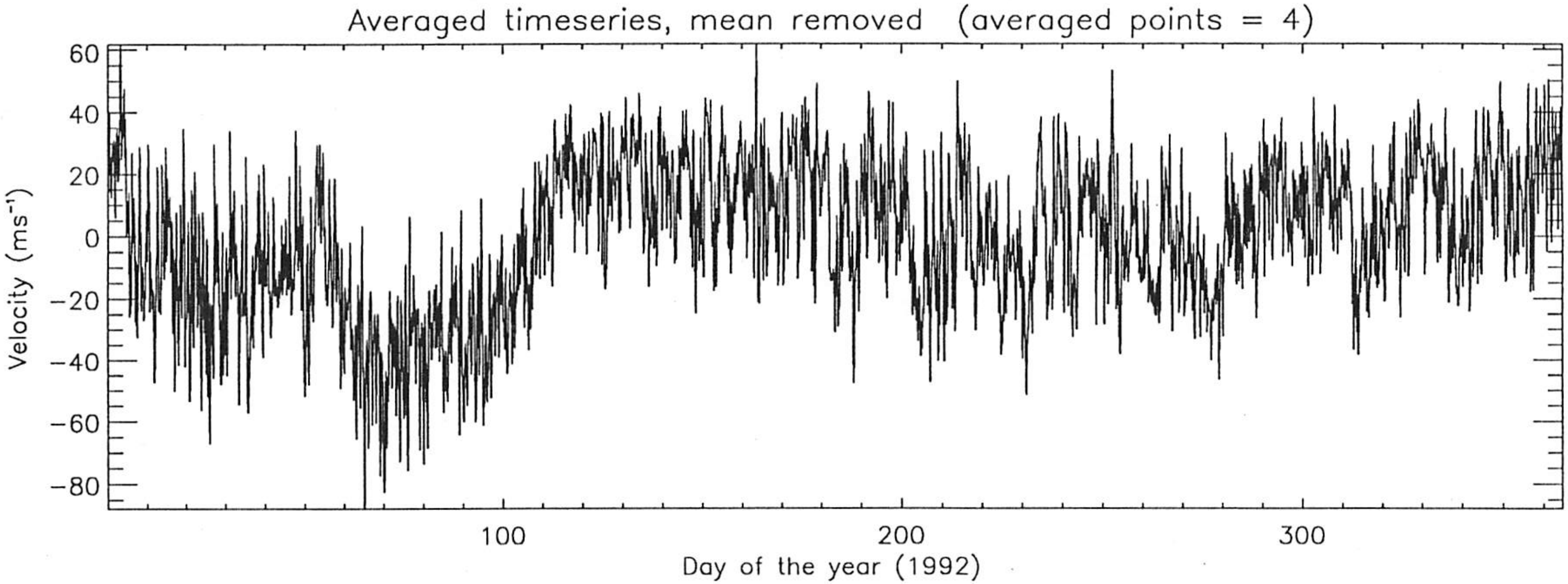
Original timeseries
Velocity (ms⁻¹)
50
0
−50
−100
0
100
200
300
Day of the year (1992)
Interpolated timeseries
Velocity (ms⁻¹)
50
0
−50
−100
0
100
200
300
Day of the year (1992)
Interpolated − Original timeseries
Velocity (ms⁻¹)
40
20
0
−20
−40
−60
0
100
200
300
Day of the year (1992)



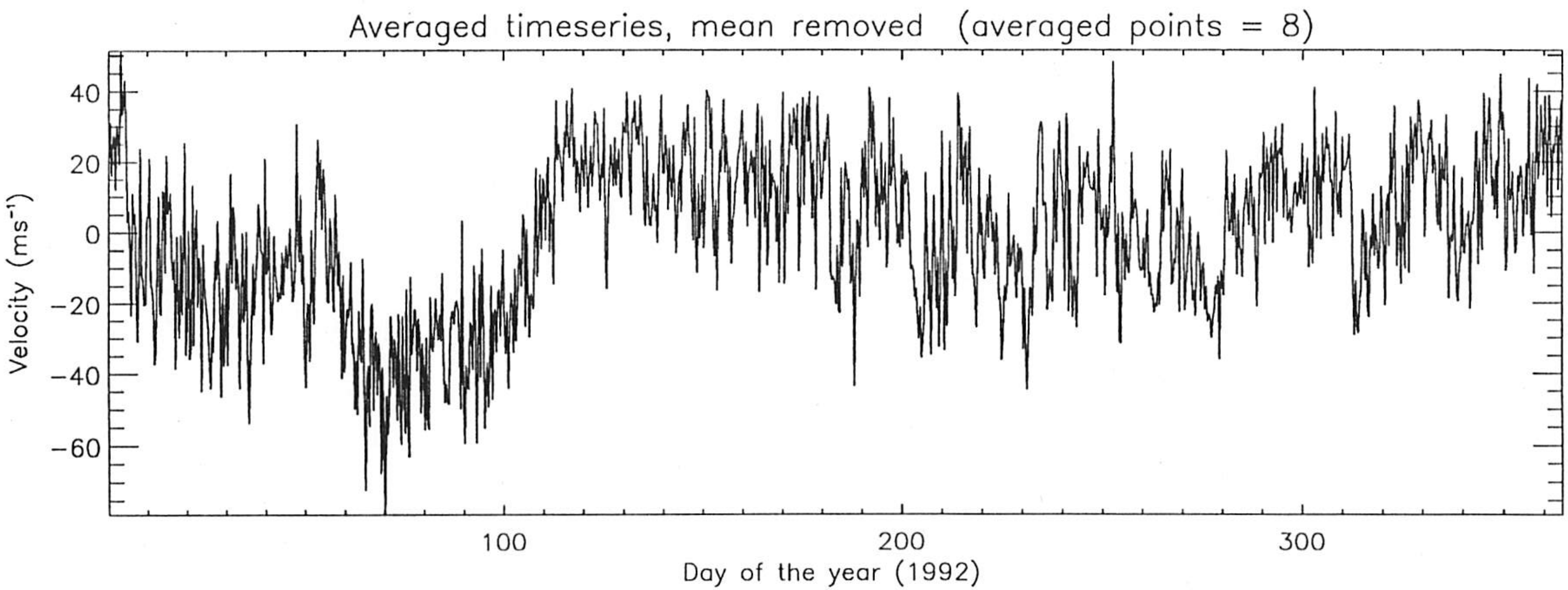
N^{-1}*periodogram of original timeseries
Magnitude (m^{2}s^{-2})
Frequency (days^{-1})
N^{-1}*periodogram of interpolated timeseries
Magnitude (m^{2}s^{-2})
Frequency (days^{-1})
N^{-1}*periodogram of interpolated − N^{-1}*periodogram spectrum of original timeseries
Magnitude (m^{2}s^{-2})
Frequency (days^{-1})



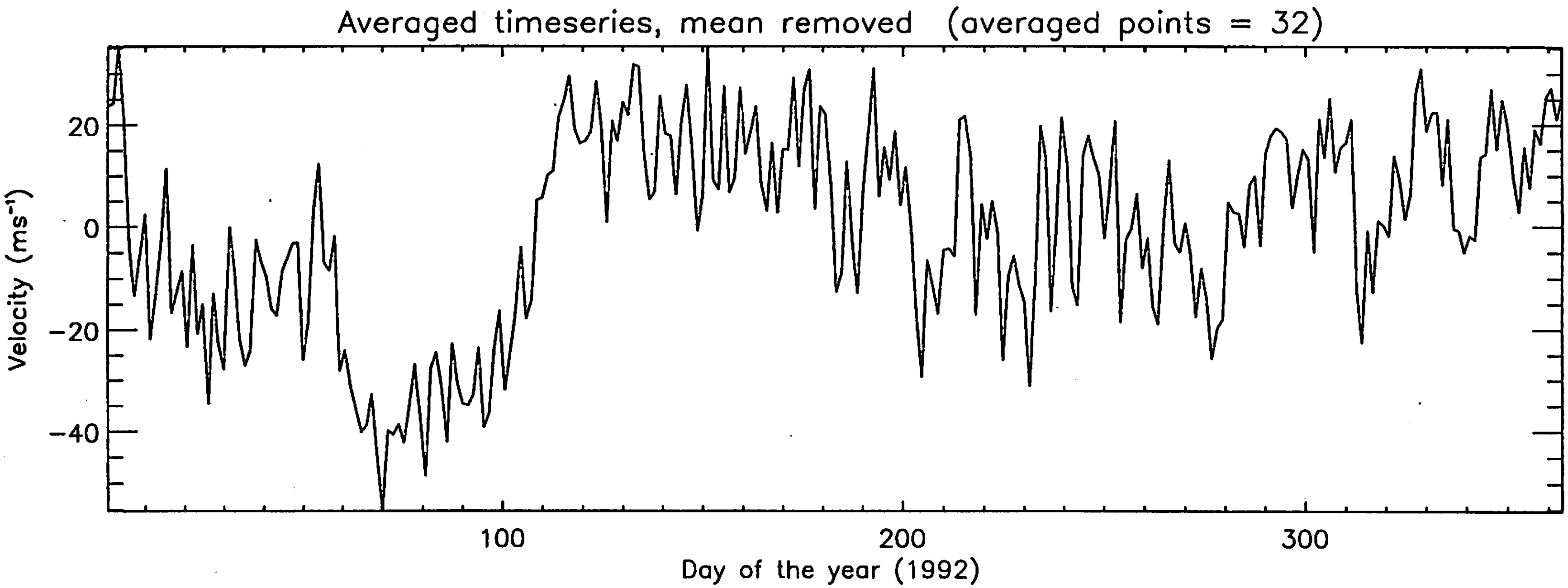
N⁻¹*periodogram of interpolated timeseries (Mean Removed)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)

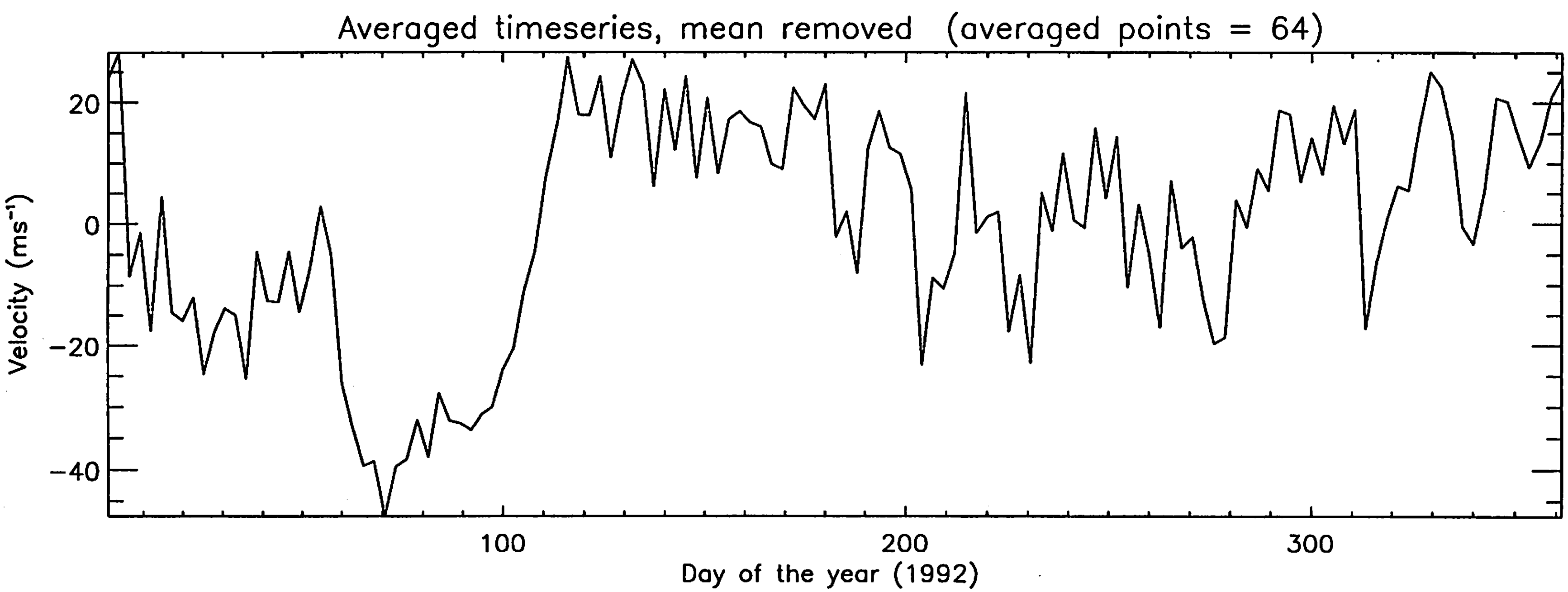


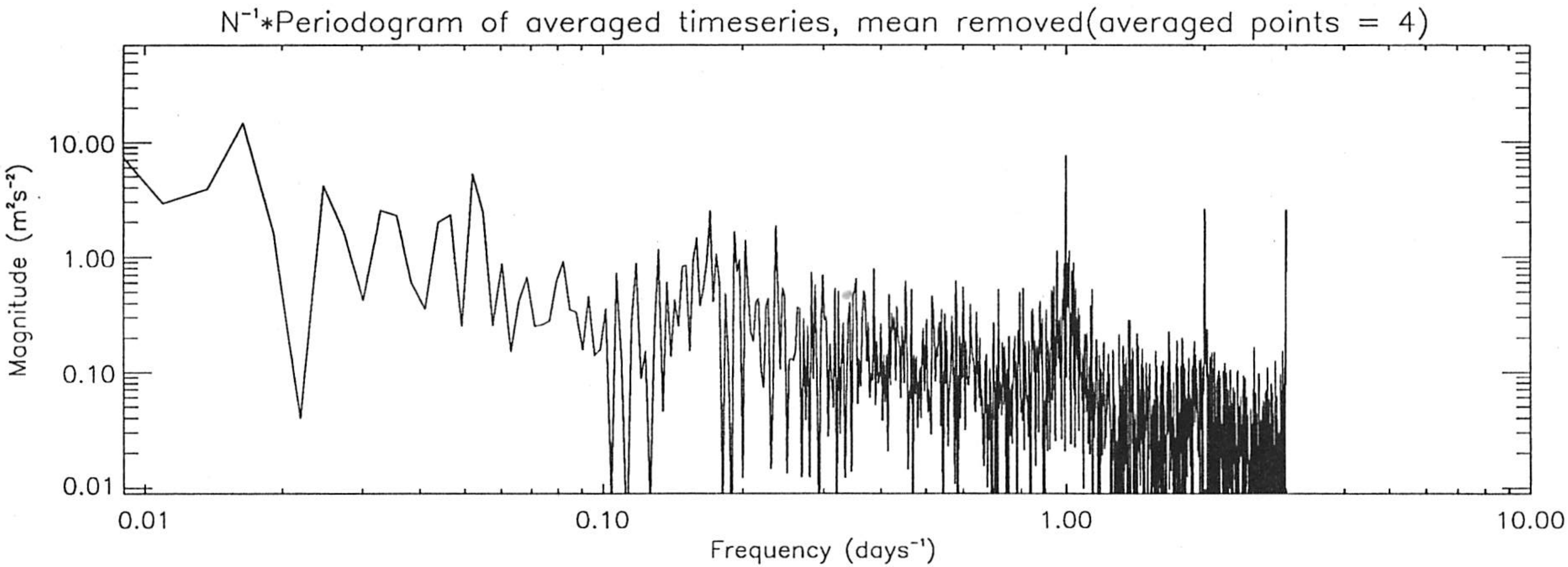
Averaged timeseries, mean removed (averaged points = 4)
60
40
20
0
−20
−40
−60
−80
Velocity (ms⁻¹)
100
200
300
Day of the year (1992)

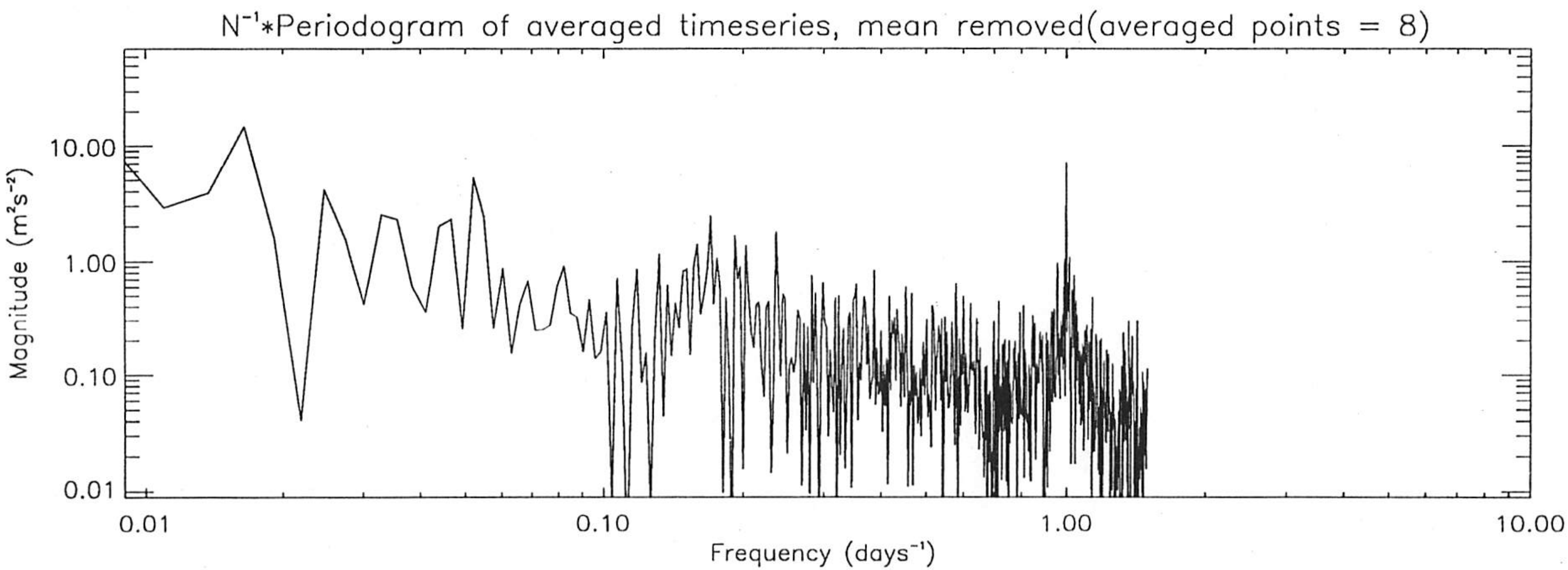


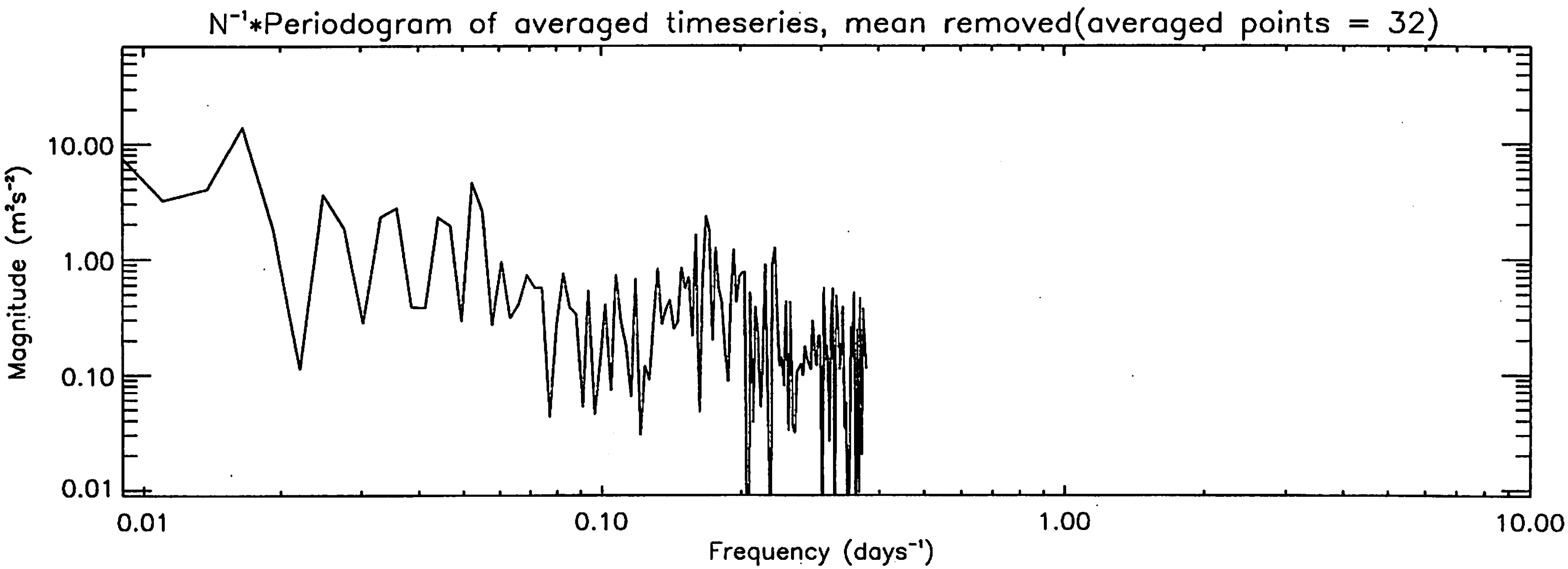
Averaged timeseries, mean removed (averaged points = 8)
40
20
0
−20
−40
−60
Velocity (ms⁻¹)
100
200
300
Day of the year (1992)


Averaged timeseries, mean removed (averaged points = 32)
Velocity (ms⁻¹)
20
0
−20
−40
100
200
300
Day of the year (1992)

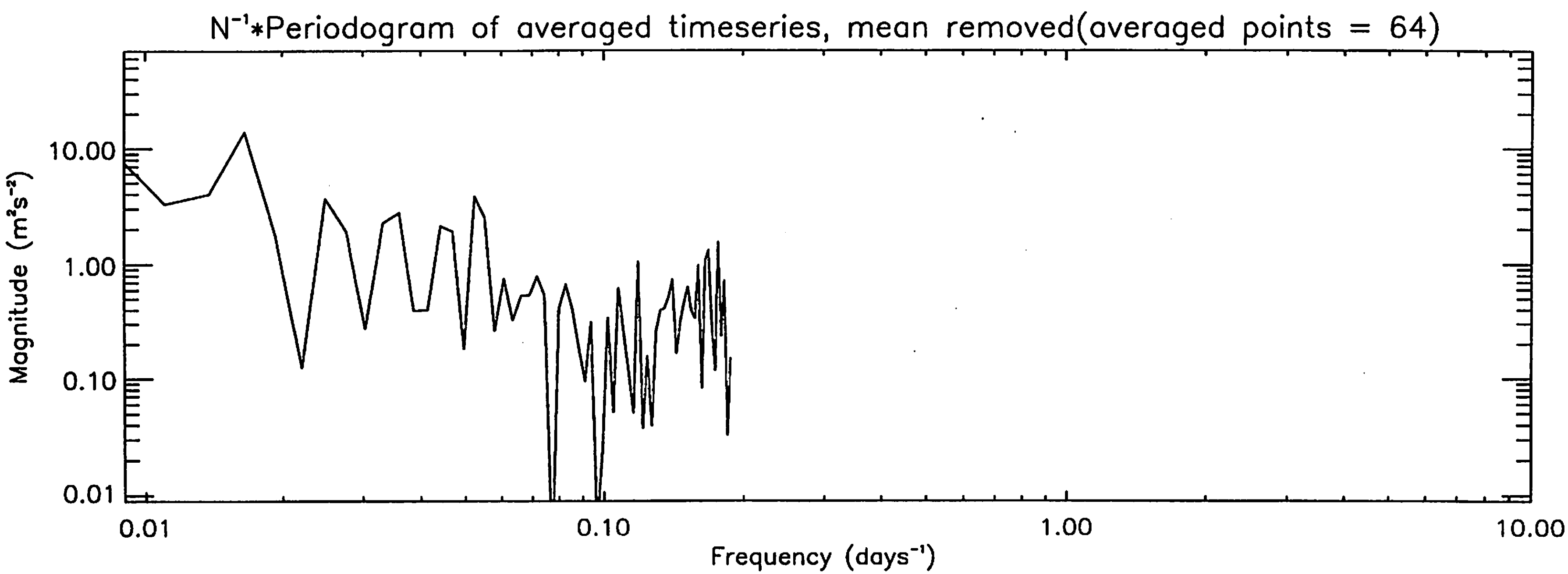

Averaged timeseries, mean removed (averaged points = 64)
Velocity (ms⁻¹)
20
0
−20
−40
100
200
300
Day of the year (1992)


N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 4)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)

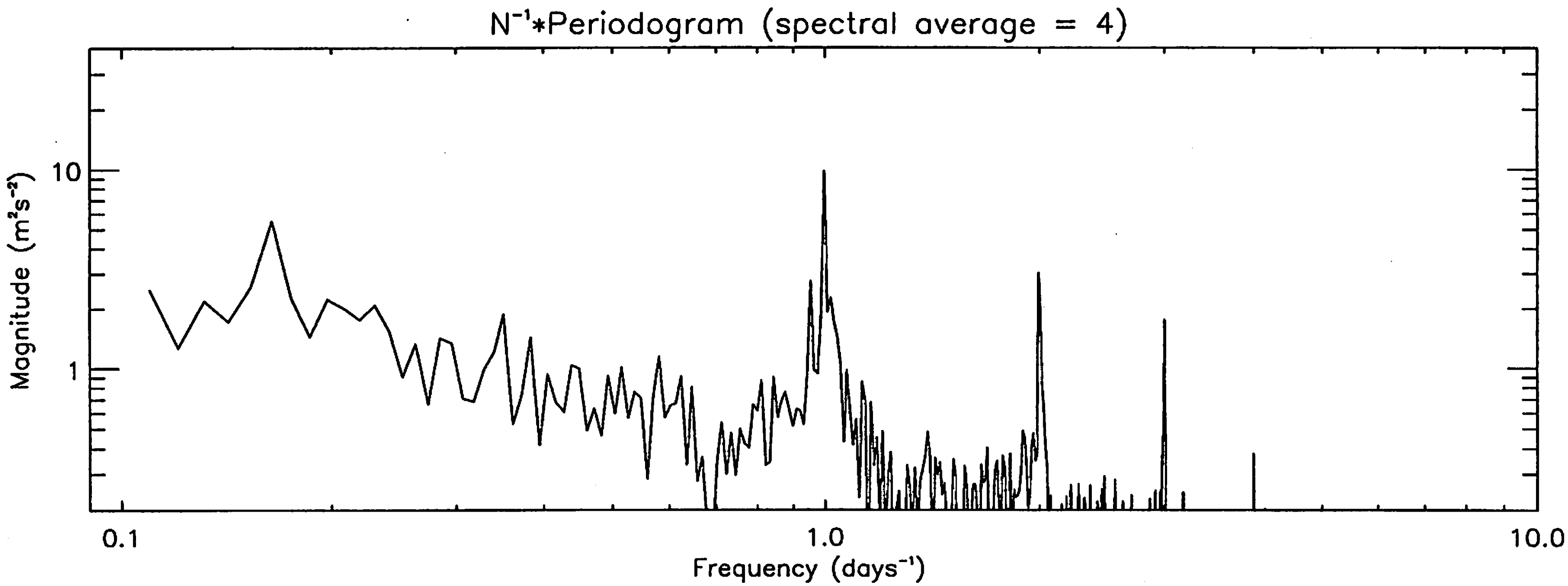

N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 8)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)

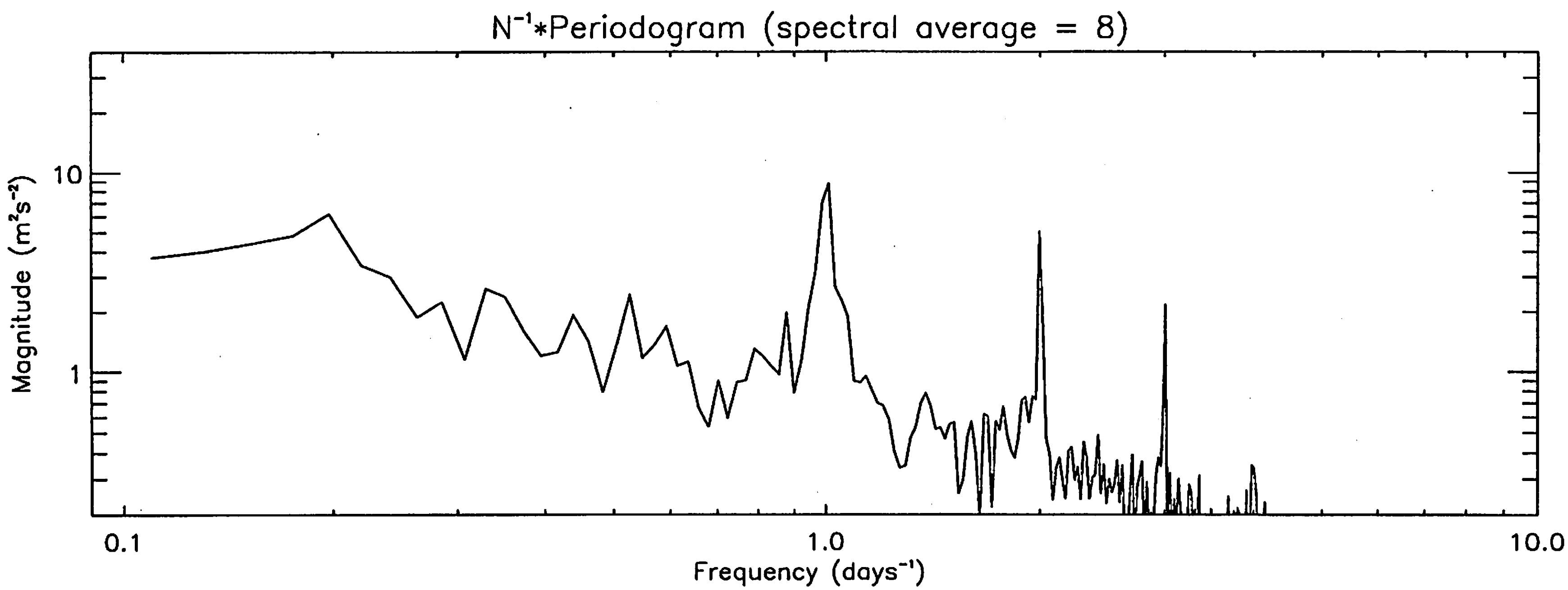


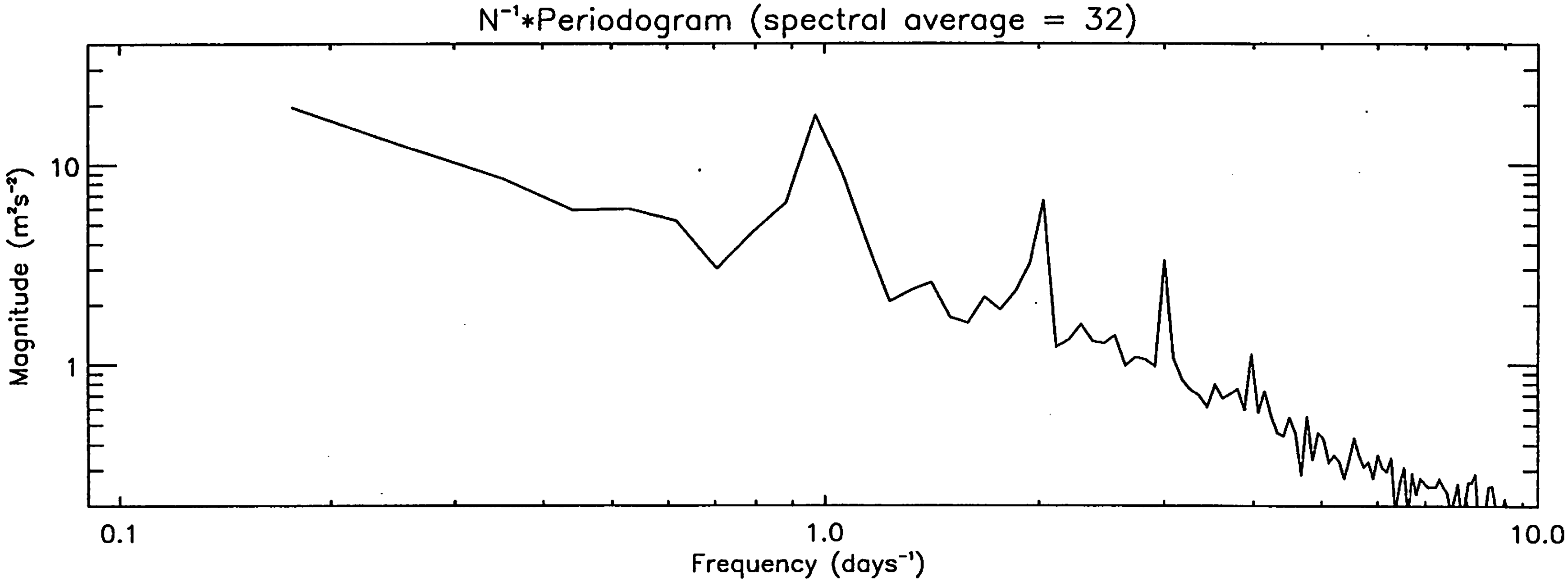
N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 32)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)

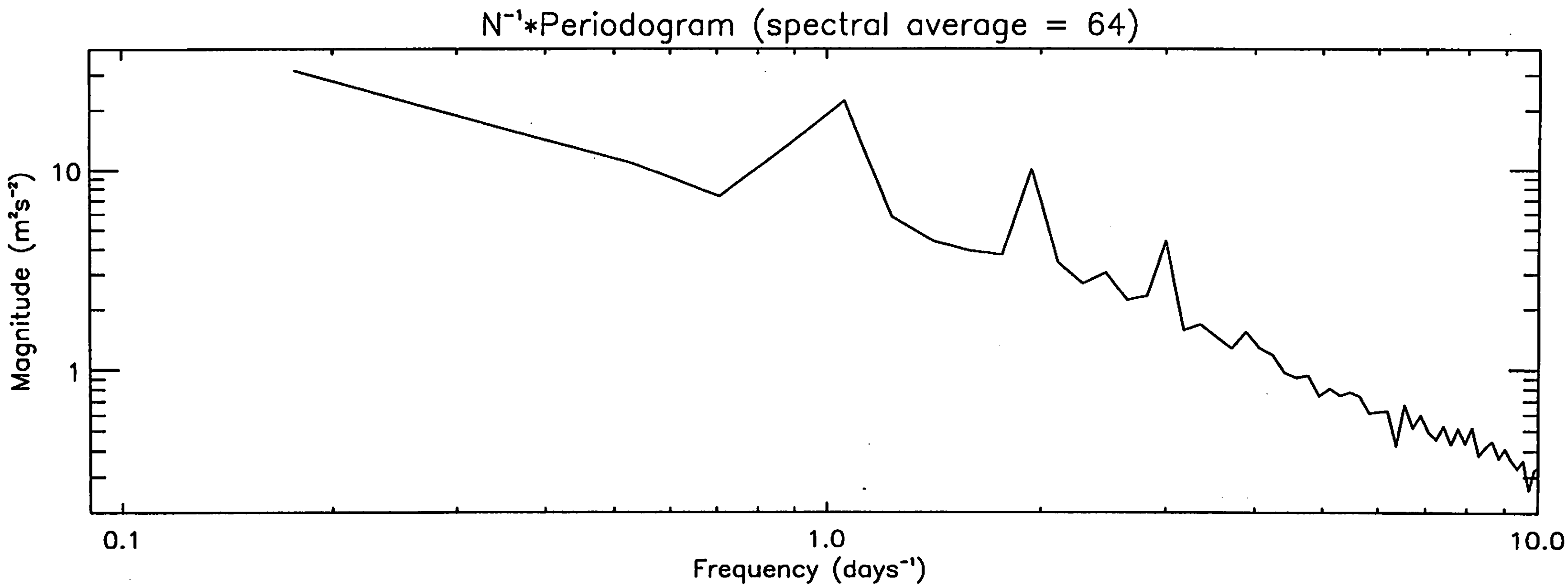


N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 64)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)


N⁻¹*Periodogram (spectral average = 4)
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)


N⁻¹*Periodogram (spectral average = 8)
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)


N⁻¹*Periodogram (spectral average = 32)
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)


N⁻¹*Periodogram (spectral average = 64)
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)

Spectral Estimation Problem

Issues
1. Finite set of observations
2. Random (stochastic process)
3. Non-stationarity of data

Random process:

Repeated occurrences of process does not give same result, i.e. measurement is repeated and the results of the measurement are not identical. (Random errors)

Leads to random fluctuations in data.

Examples:

(a) Variations due to fluctuations in environmental conditions i.e. temperature, humidity, etc.
(b) Mechanical vibrations, electrical noise, cosmic noise, etc.

Systematic errors:

Usually constant. Can be determined and removed or apply correction.

Example: Calibration error of instrument.

Small random errors $\Rightarrow$ high precision
Small systematic errors $\Rightarrow$ high accuracy

Random process v.s. Deterministic process

Model for data : $\sum_i A_i \cos(\omega_i t+\phi_i)$ **If know** A_i, ω_i, ϕ_i

exactly and can reconstruct the signal for all time $\{-\infty,\infty\}$, **then process is deterministic.**

However true model for data contains signal and noise i.e. $\sum_i A_i \cos(\omega_i t+\phi_i)$ **+ Noise. Noise due to random errors. Therefore we have a random process and the problem becomes one of spectral <u>estimation</u> and providing a <u>confidence</u> <u>measure</u> of that estimate.**

Description of random process includes assigning probabilites to the possible values of the process.

Probability Distributions

Probablity: A number that measures the likelihood of a certain outcome when experiment is done.

Intuitive connection between frequency of occurrence and probability

Probability is the number of times the outcome is observed when the experiment is performed a large number of times (n/N, N large). Ensemble average

Note: We are taking time or spectral averages from one experiment. The assumption is that we have an ergodic process, i.e. time averaging from one experiment ≡ ensemble average.

Probability Density Functions (p(x) v.s. x)

Probability of x being in a specified interval of values

Normalized: $0 \leq p(x) \leq 1$

$$\int_x p(x)dx = 1$$

Definition: Mean and variance

x is a random variable

The nth moment is $M_n(x) = \int_x x^n p(x)dx$

mean ≡ expected value $\equiv \mu \equiv E(x) = M_1(x)$

variance of $x = \sigma^2 = E((x-\mu)^2)$

$= E(x^2) - \mu^2$

$= M_2(x) - (M_1(x))^2$

Important PDF's for our use:

1. Normal (Gaussian)
2. Chi-Square

1. **Normal PDF**

   Proposed by Gauss (1777-1855) as a model for frequency of occurrence of errors such as measurement error. Superposition of several random processors yields a normal PDF.

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty$$

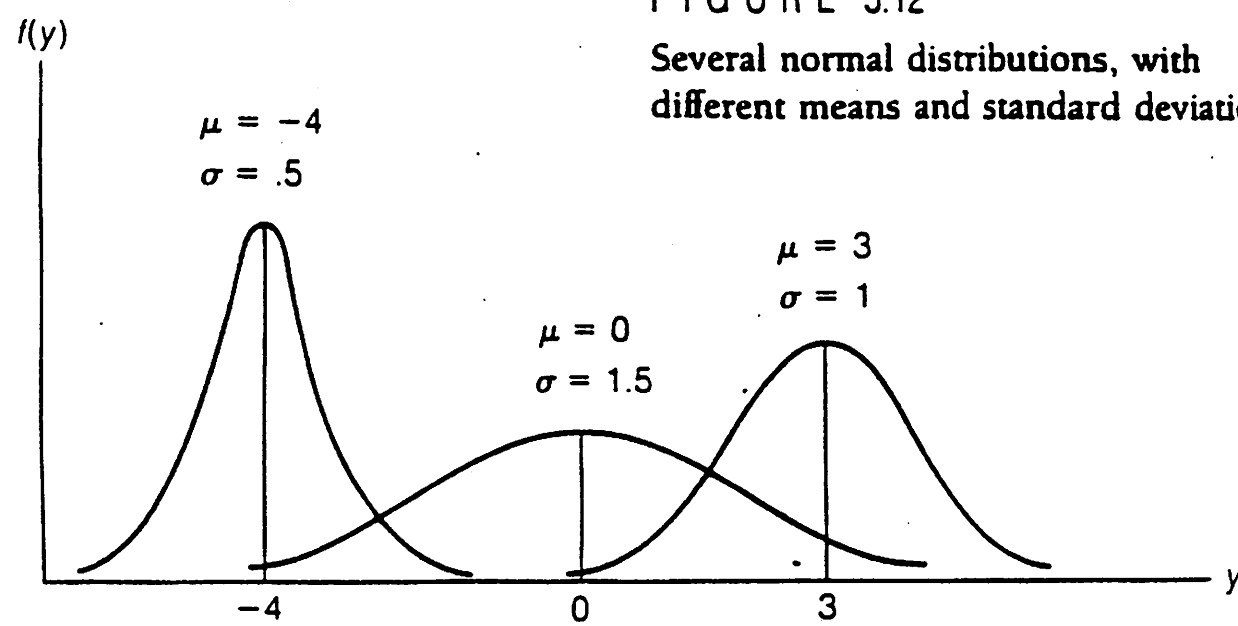


FIGURE 5.12

Several normal distributions, with different means and standard deviations

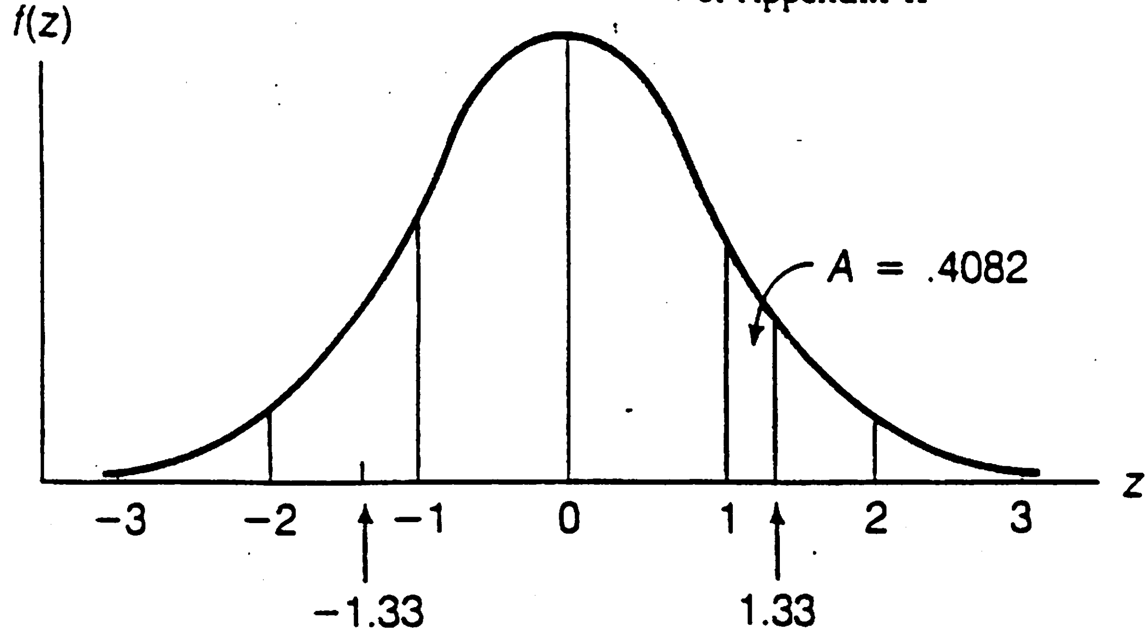
Tabulations of standard normal PDF

p(z) v.s. z where $z = \frac{x-\mu}{\sigma}$

z has zero mean and variance of one. Can relate to different means and variances.

FIGURE 5.13

Standard normal density function showing the tabulated areas given in Table 4 of Appendix II



2. Chi-Square PDF

Part of class of random variables that can only have non-negative values. Useful for hypothesis testing and confidence intervals.

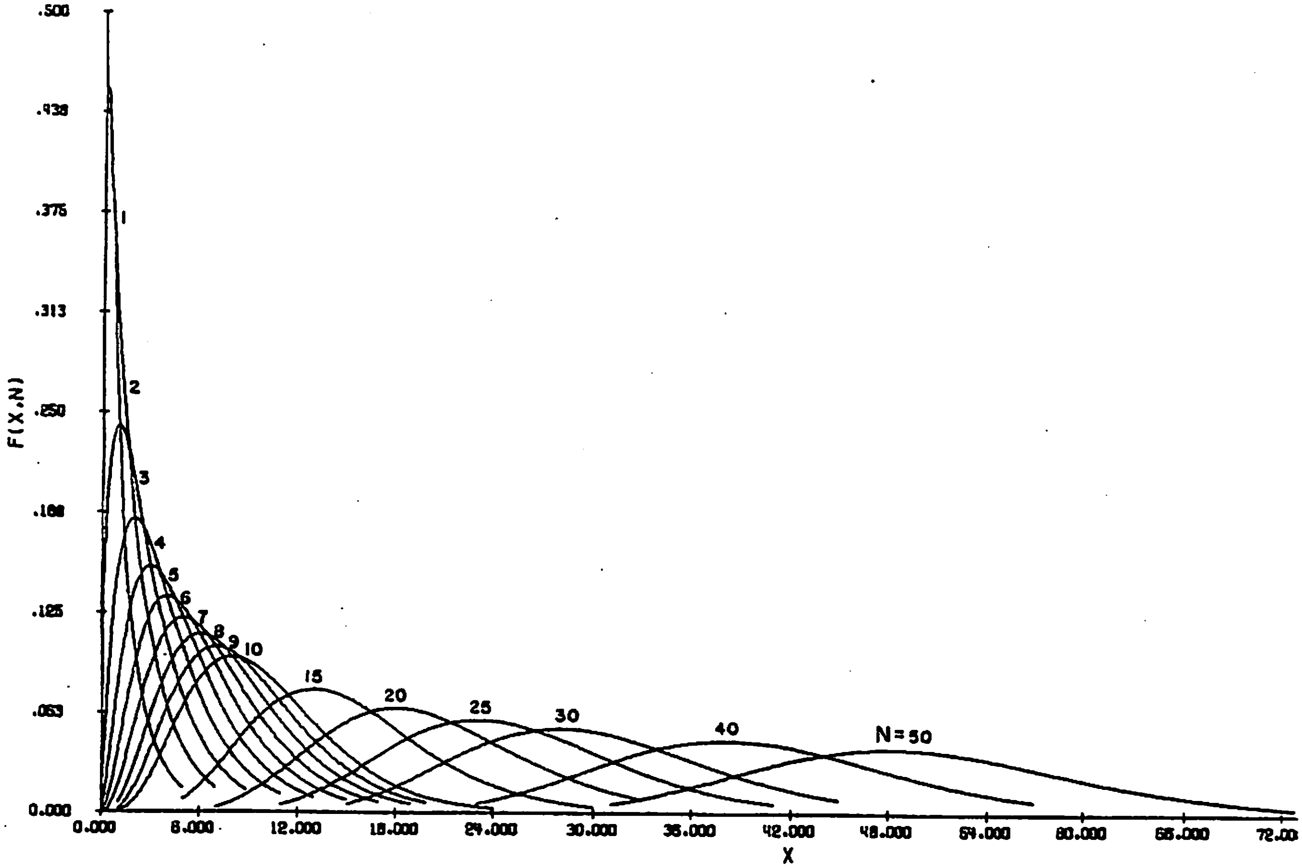


Figure 26.1
The probability density function of the chi-square distribution plotted as $F(X,N)$ versus X for different values of N, the number of independent degrees of freedom. X = value of χ^2.

$$P(\chi^2) = \frac{(\chi^2)^{1/2(n-2)}}{\Gamma(\frac{n}{2})2^{n/2}} e^{-\chi^2/2}$$

n = number of degrees of freedom for χ^2 distribution

mean = n;

variance = 2n

Confidence Intervals and Significance Levels

Confidence interval: The interval that will contain the true power spectral density with a high degree of confidence.

Significance level: The probability that a signal is present or not. Use hypothesis testing.

In order to determine these quantities, we need to examine the statistical distribution of the power spectrum. Look at behavior of noise.

Given $x(n) \sim N[0,\sigma^2]$; $x(n)$ are independent.

$$X(k) = \sum_{n=0}^{N-1} x(n)\, e^{-j\, 2\pi/N\, kn}$$

Note that if x_i are independent normal random variables then $\sum_{i=0}^{N-1} x_i$ is a normal random variable.

Expected value

$$E[X(k)] = E\left[\sum_{n-0}^{N-1} x(n)\, e(\frac{-j2\pi kn}{N})\right]$$

$$= \sum_{n-0}^{N-1} E[x(n)]\, e(\frac{-j2\pi kn}{N}) = 0$$

Variance

$$\mathrm{Var}[X(k)] = E[|X(k)|^2] = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E[x(n)x^*(m)]e^{-j2\pi kn/N}\ e^{j2\pi km/N}$$

$$= \sum_{\substack{n=0 \\ n \neq m}}^{N-1} \sum_{m=0}^{N-1} E[x(n)x^*(m)]\ e^{-j2\pi kn/N}\ e^{j2\pi km/N}$$

$$+ \sum_{n=0}^{N-1} E[|x(n)|^2]$$

$$= \sum_{n=0}^{N-1} \sigma^2 = N\sigma^2$$

Now recall that $\sum_{i=0}^{N-1} z_i^2 = \sum_{i=0}^{N-1} (\frac{x_i - \mu}{\sigma})^2$ **is** χ_N^2

i.e. chi-square with N degrees of freedom.

In our case then $\frac{|X(k)|^2}{N\sigma^2} \sim \chi_1^2$

But $P(k) = \frac{1}{N} |X(k)|^2$, **so** $\frac{P(k)}{\sigma^2} \sim \chi_1^2$

If only noise is present, $\frac{P(k)}{\sigma^2}$ **is chi-square with 1 degree of freedom.**

Confidence Interval

Recall confidence interval is the probability that true power spectrum lies within interval about the estimated power spectrum.

For an infinite series of white noise, $N[0,\sigma^2]$ **the true power spectrum, is**

$$S(f) = \sigma^2 \qquad \textbf{for all f.}$$

Given a finite number of measurements of white noise then the expected value of the power spectrum estimator is:

$$E[P(k)] = S(f) \; * \; N\,\text{sinc}^2\left(\frac{Nf}{2}\right)$$

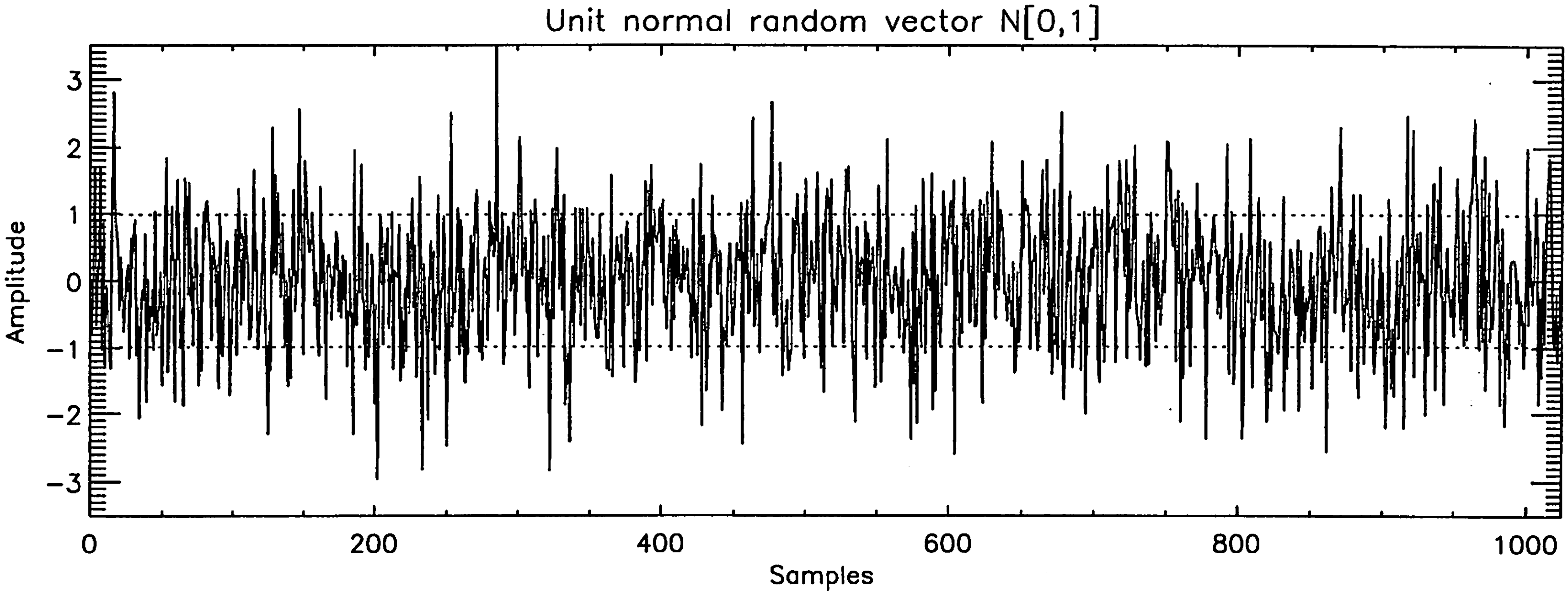
$$\text{Bias} = E[P(k)] - S(f)$$

$$\lim_{N\to\infty} \text{Bias} \to 0 \;\; \text{since} \;\; N\,\text{sinc}^2\left(\frac{Nf}{2}\right) \to \delta(f)$$

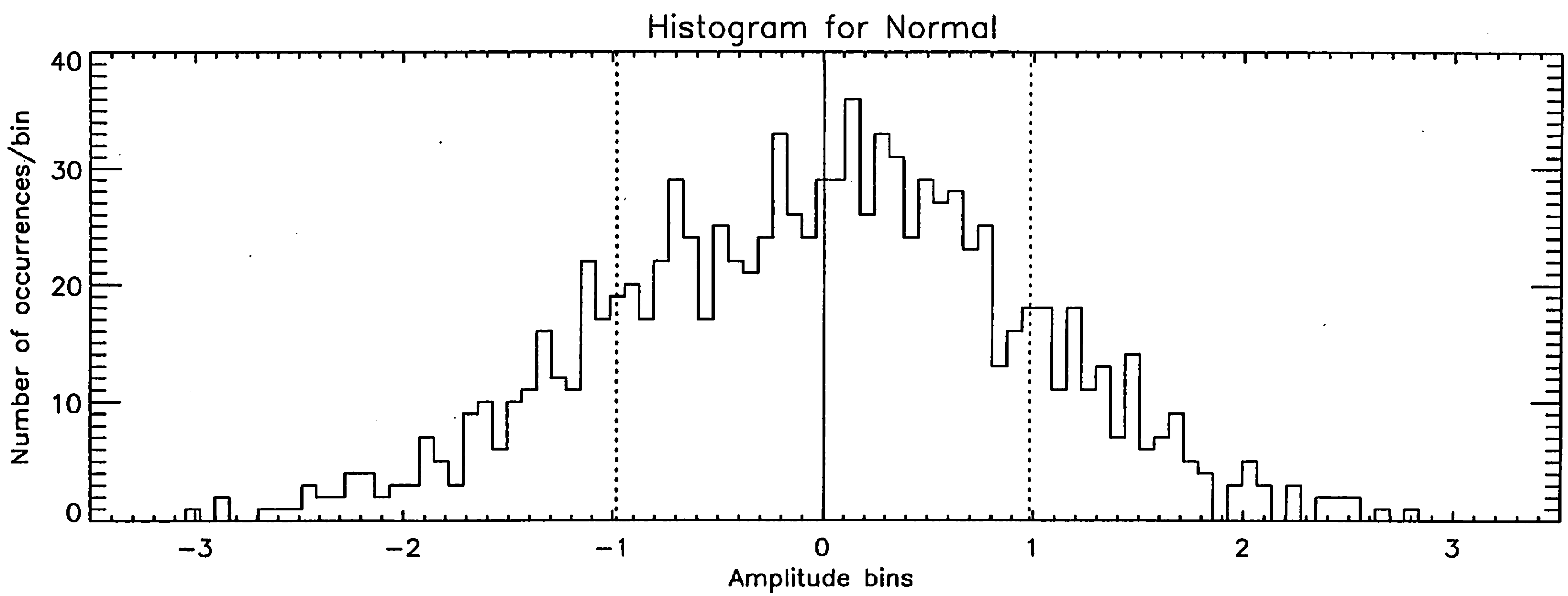
However,

$$\text{var}\,(P(k)) \sim S^2(f) \qquad , \;\; \textbf{i.e.} \;\; \sigma^4$$

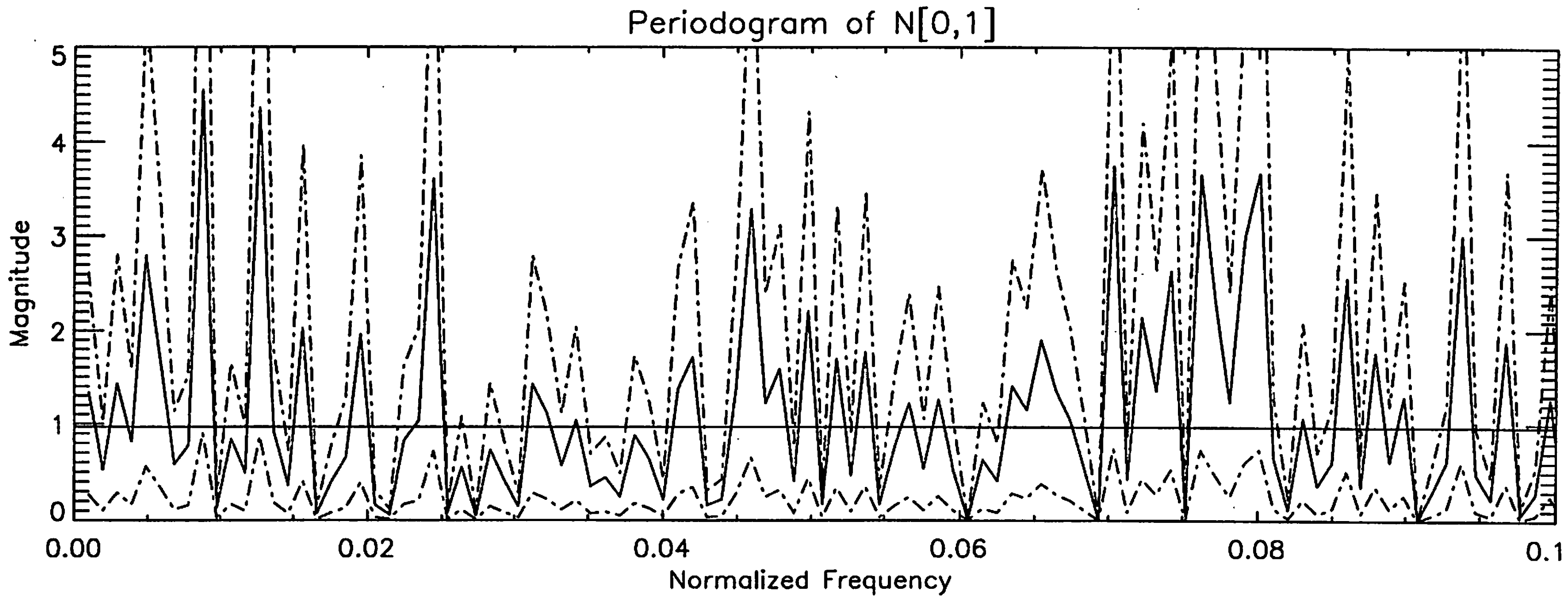
hence increasing N will not decrease the variance of the estimator. In fact the variance of the estimator is on the order of the square of the estimator. Time averaging will decrease σ^2**.**



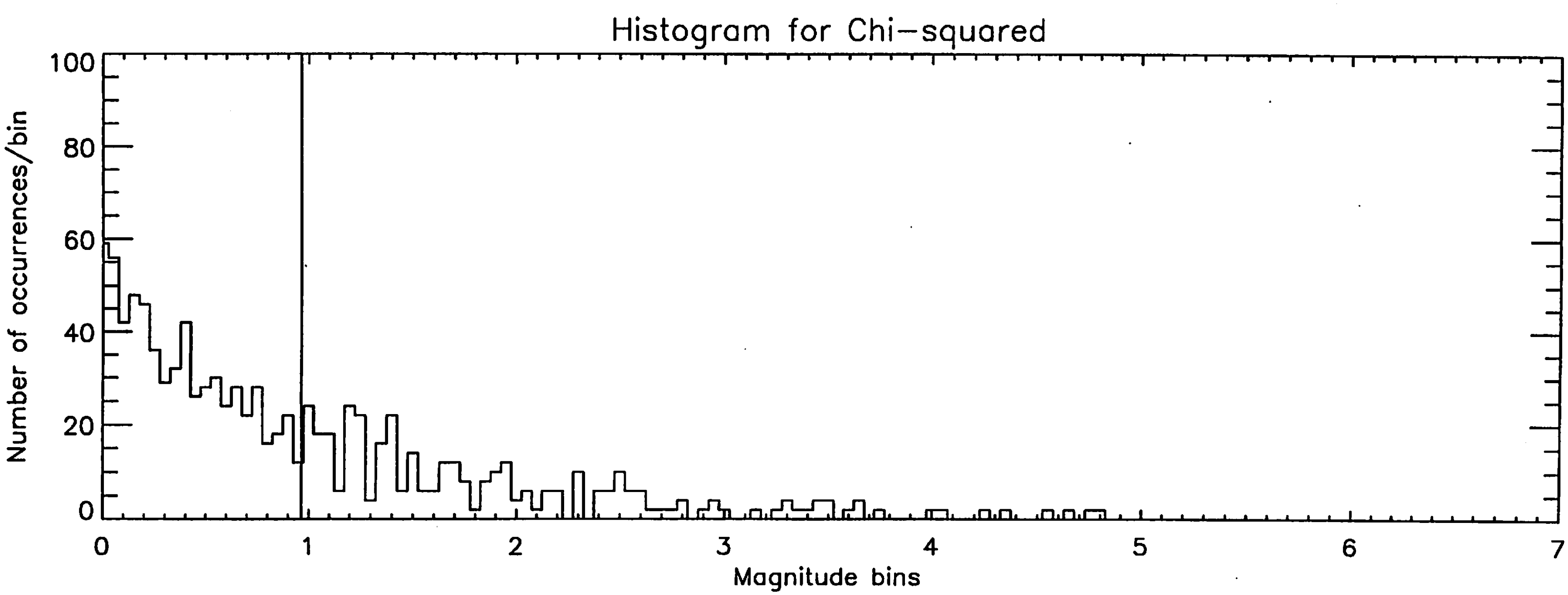
Unit normal random vector N[0,1]
Amplitude
3
2
1
0
−1
−2
−3
0
200
400
600
800
1000
Samples



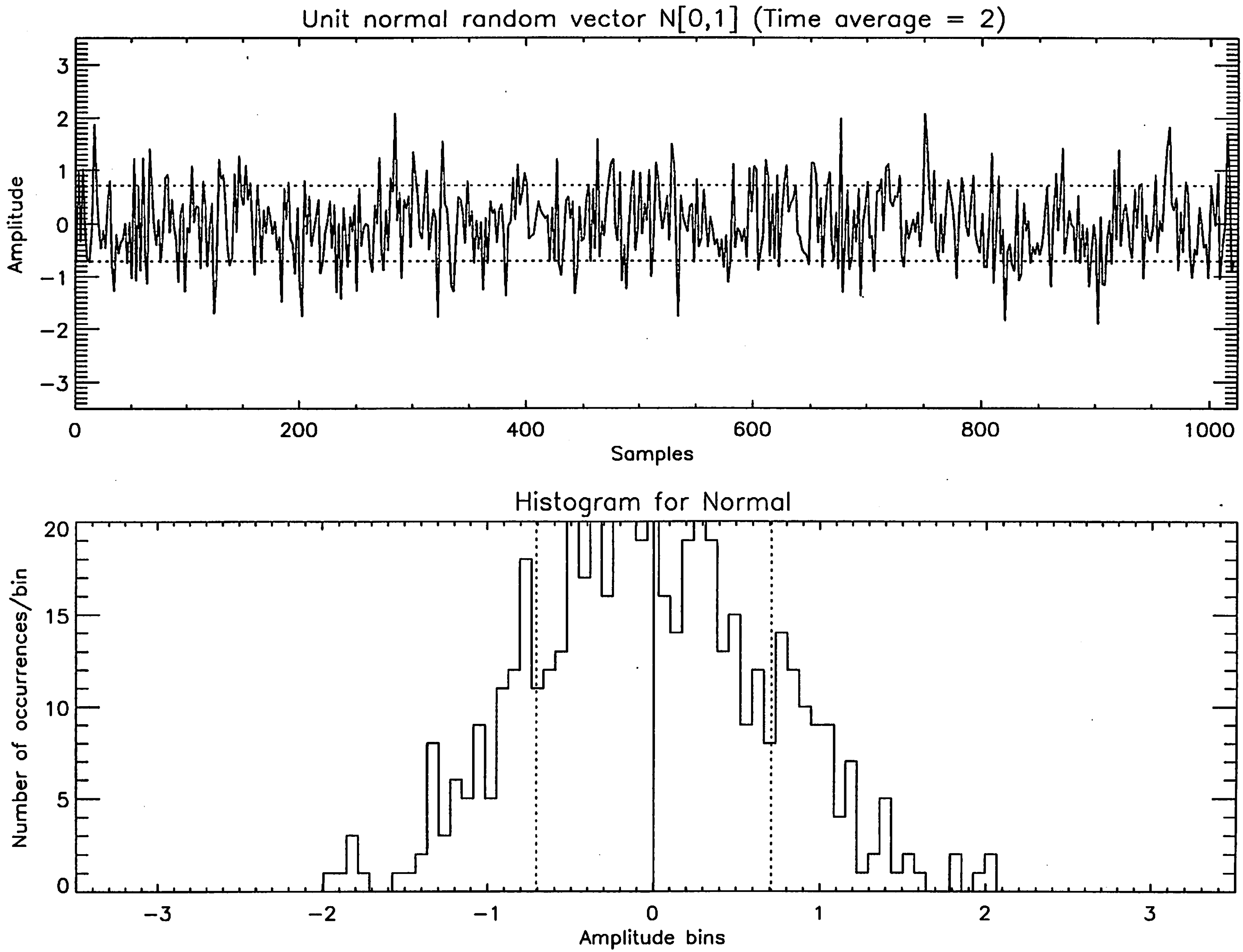
Histogram for Normal
Number of occurrences/bin
40
30
20
10
0
−3
−2
−1
0
1
2
3
Amplitude bins

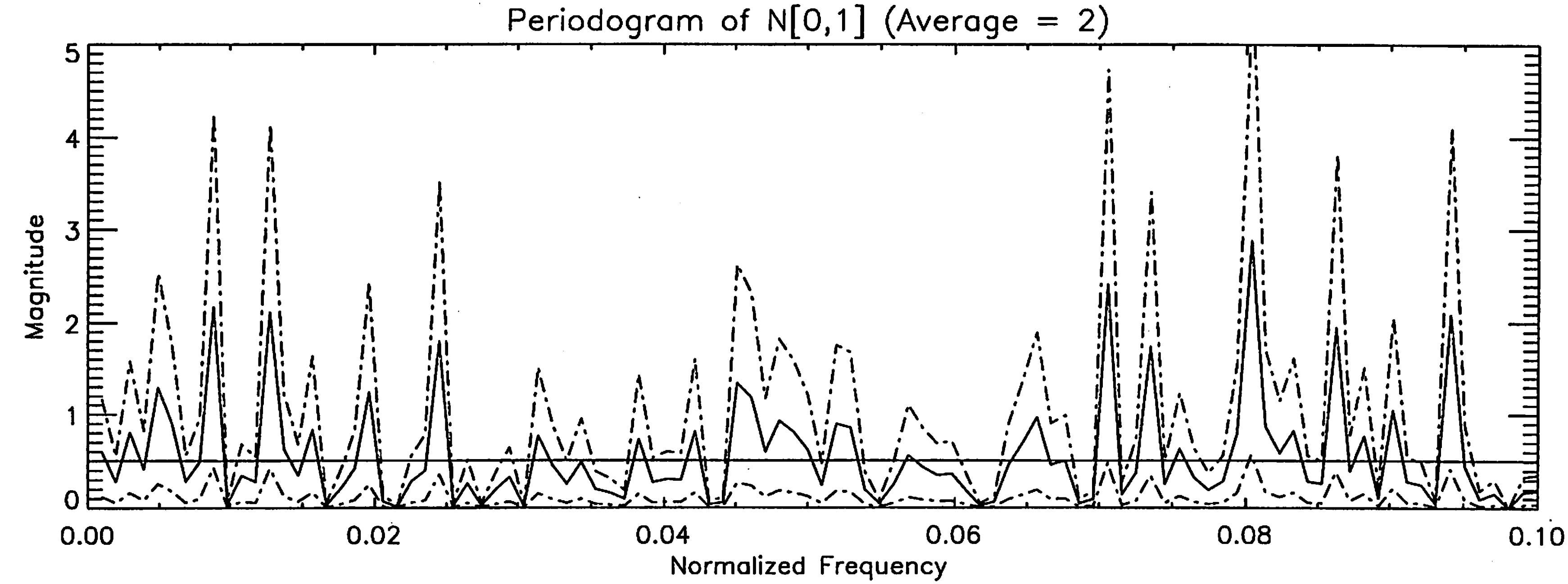


Periodogram of N[0,1]
Magnitude
0
1
2
3
4
5
0.00
0.02
0.04
0.06
0.08
0.10
Normalized Frequency

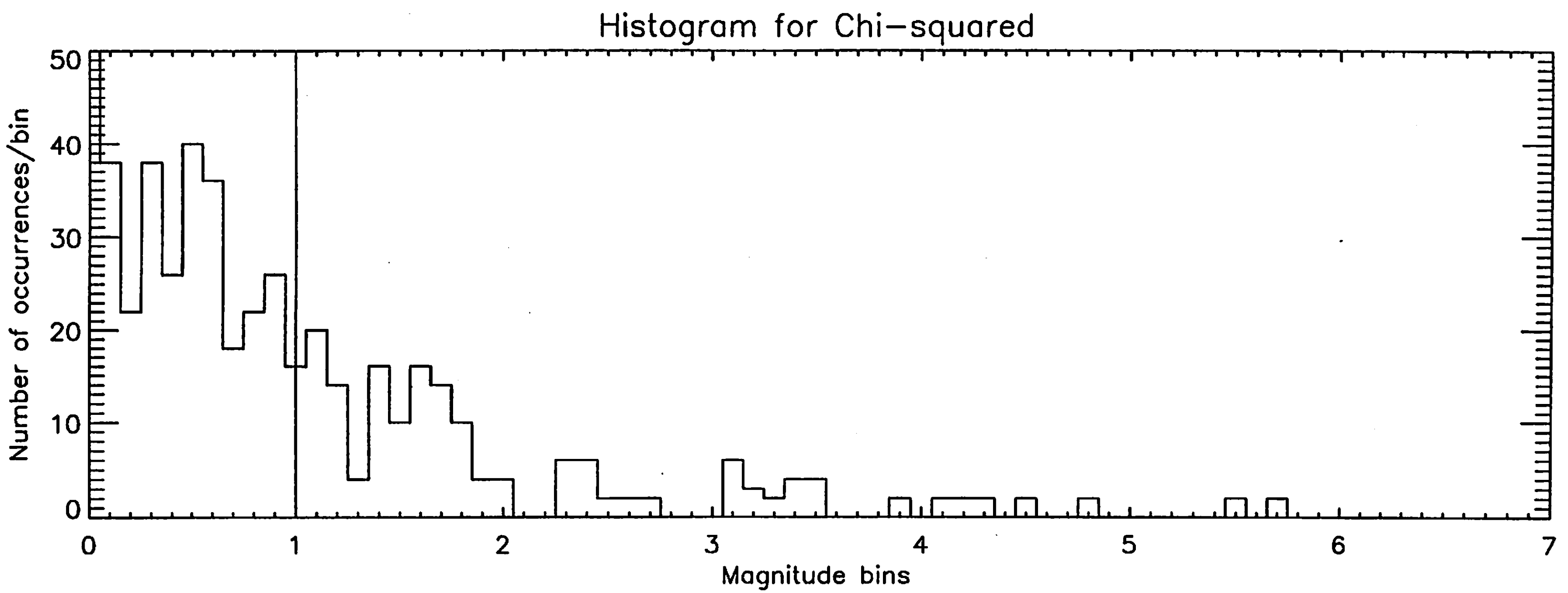


Histogram for Chi-squared
Number of occurrences/bin
0
20
40
60
80
100
0
1
2
3
4
5
6
7
Magnitude bins


Unit normal random vector N[0,1] (Time average = 2)
Amplitude
Samples
Histogram for Normal
Number of occurrences/bin
Amplitude bins



Periodogram of N[0,1] (Average = 2)
Magnitude
Normalized Frequency
0.00
0.02
0.04
0.06
0.08
0.10
0
1
2
3
4
5



Histogram for Chi-squared
Number of occurrences/bin
Magnitude bins
0
1
2
3
4
5
6
7
0
10
20
30
40
50

Spectral averaging

If time series x(n) is parsed into M segments of length L such that ML = N, then P(k,m) can be computed for each segment and

$$P(k)_{avg} = \frac{1}{M}\sum_{m=0}^{M-1} P(k,m)\,.$$

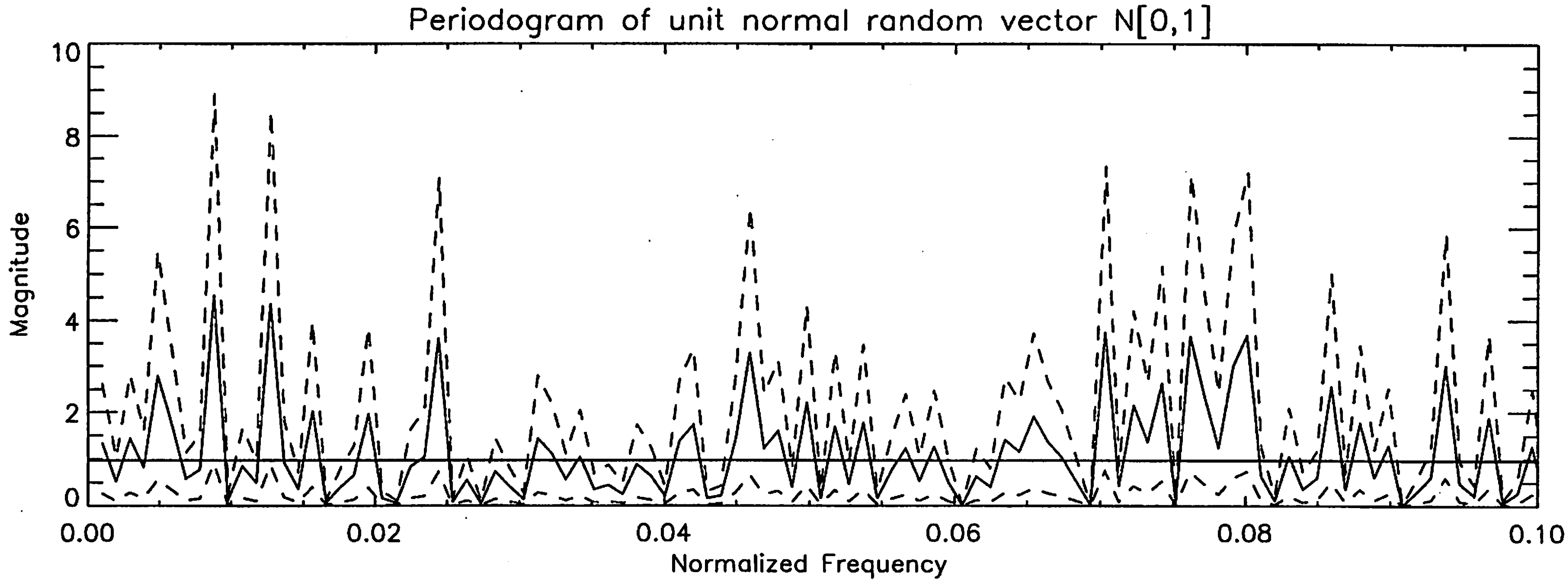
the variance of $P(k)_{avg}$ **is:**

$$\mathrm{Var}(P(k)_{avg}) = \frac{1}{M^2}\mathrm{Var}\left[\sum_{m=0}^{M-1} P(k,m)\right]$$

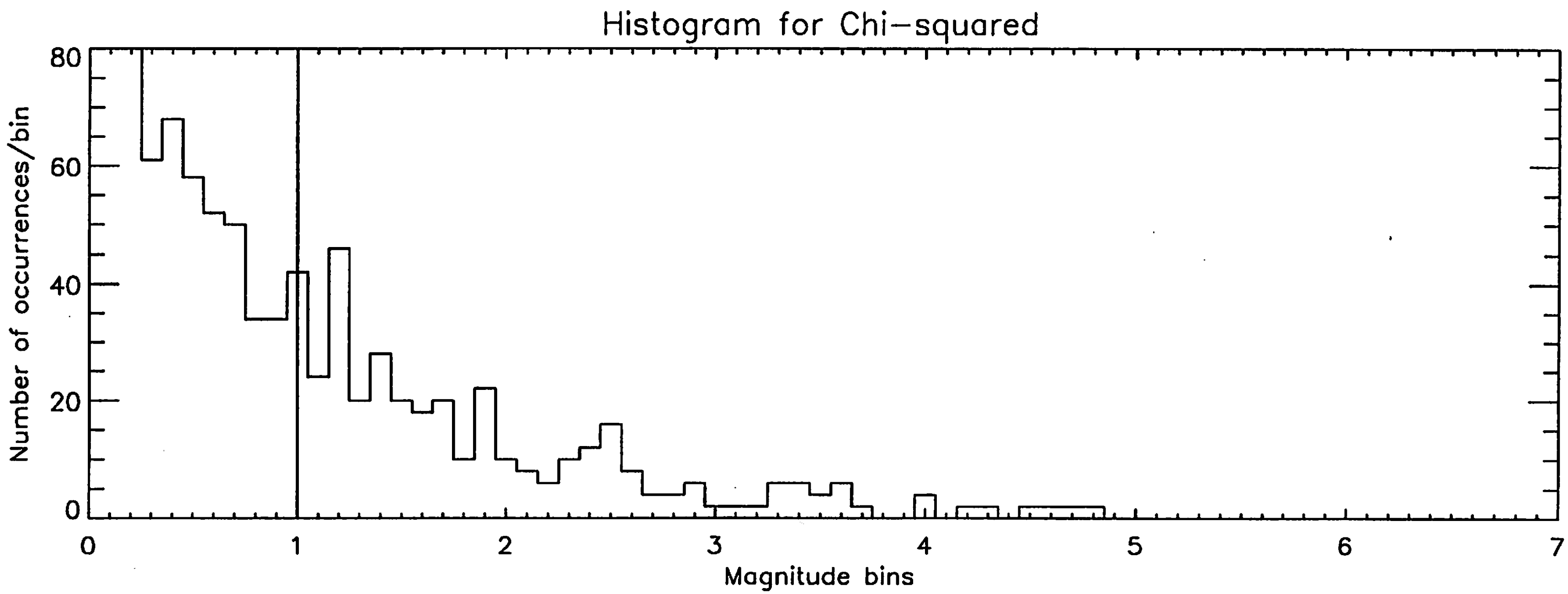
$$= \frac{1}{M}\mathrm{Var}\,[P(k,m)]$$

Thus, spectral averaging reduces the variance at the cost of resolution because $N \rightarrow L$ **.**

α = 0.5

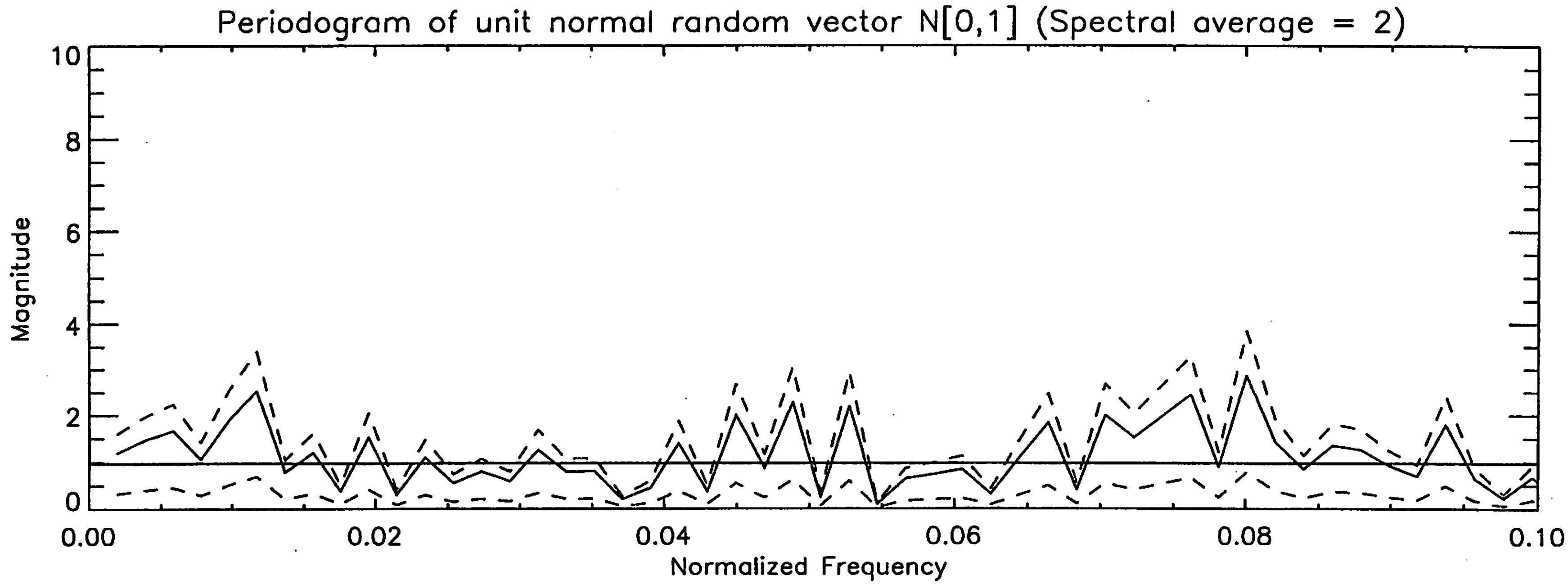


Periodogram of unit normal random vector N[0,1]
Magnitude
0
2
4
6
8
10
0.00
0.02
0.04
0.06
0.08
0.10
Normalized Frequency

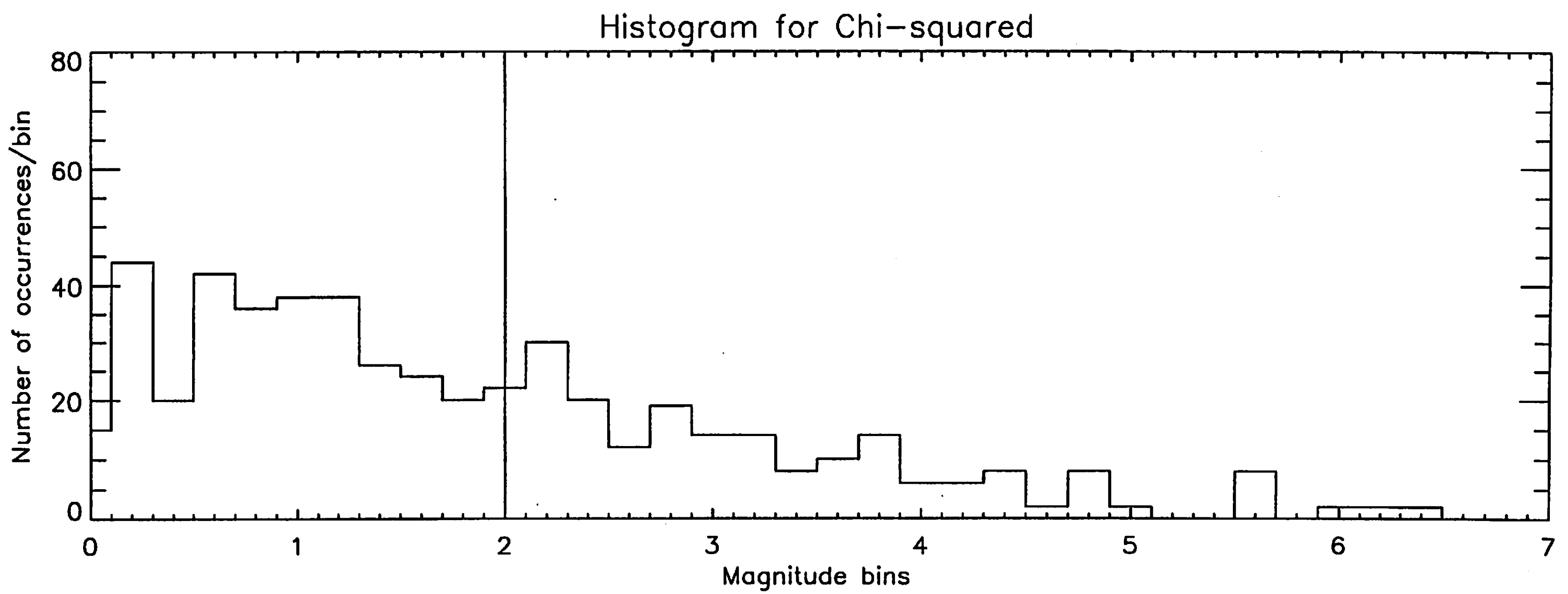


Histogram for Chi-squared
Number of occurrences/bin
0
20
40
60
80
0
1
2
3
4
5
6
7
Magnitude bins

α= 0.5



Periodogram of unit normal random vector N[0,1] (Spectral average = 2)
Magnitude
0
2
4
6
8
10
0.00
0.02
0.04
0.06
0.08
0.10
Normalized Frequency



Histogram for Chi-squared
Number of occurrences/bin
0
20
40
60
80
0
1
2
3
4
5
6
7
Magnitude bins

Confidence interval

Recall $\frac{P(k)}{\sigma^2} \sim \chi_1^2$, i.e. $\frac{P(k)}{S(f)} \sim \chi_1^2$,

For spectral averaging and rectangular window

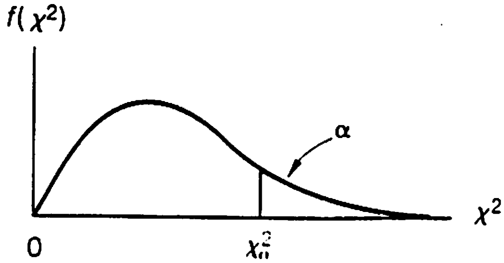
$$\frac{MP(k)_{avg}}{S(f)} \sim \chi_M^2$$

and the confidence interval for each spectral estimate becomes

$$\frac{MP(k)_{avg}}{\chi_M^2(\alpha/2)} \leq S(f) \leq \frac{MP(k)_{avg}}{\chi_M^2(1-\alpha/2)}$$

Therefore to obtain the tightest confidence interval about our estimate, we should perform as many spectral averages that can be tolerated.

TABLE 7 Critical Values of χ^2



Example: 95% Confidence Interval

$\alpha = .05$ $\alpha/2 = .025$

$1-\alpha = .95$ $1-\alpha/2 = .975$

$0.199 P(k) \le S(f) \le 1018\ P(k)$ (H=

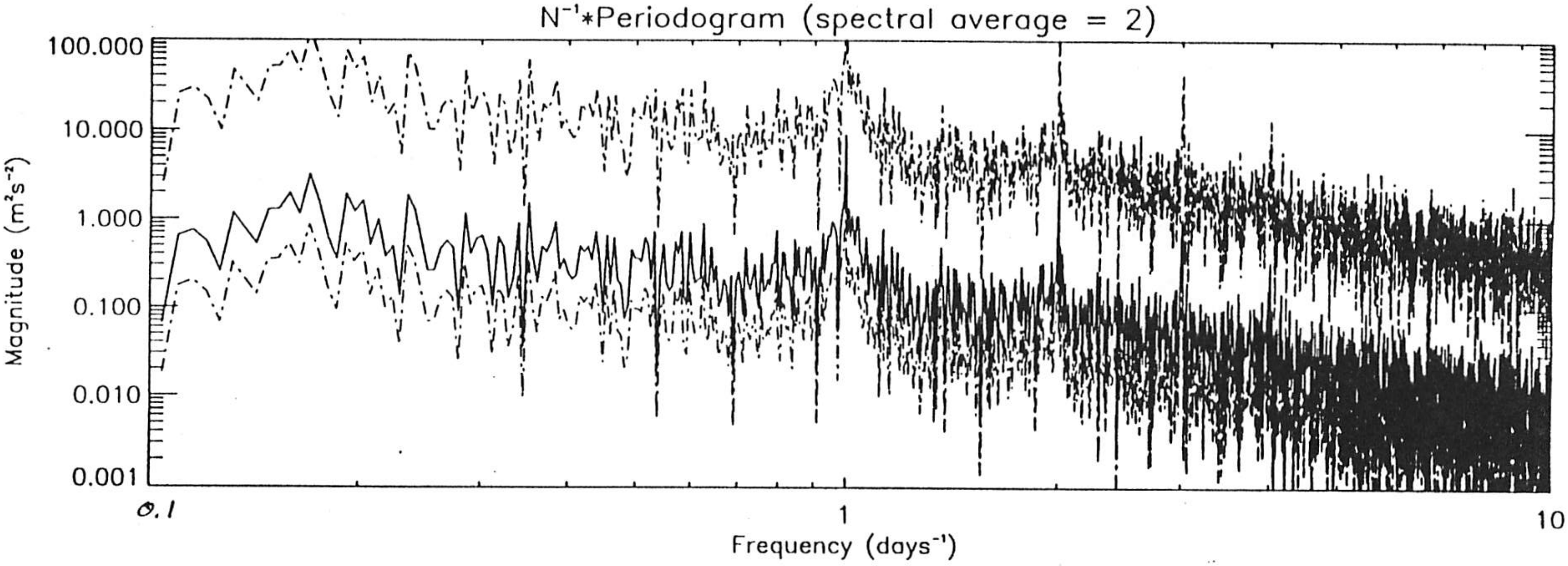
$0.5547 P(k) \le S(f) \le 2.316\ P(k)$ (H=

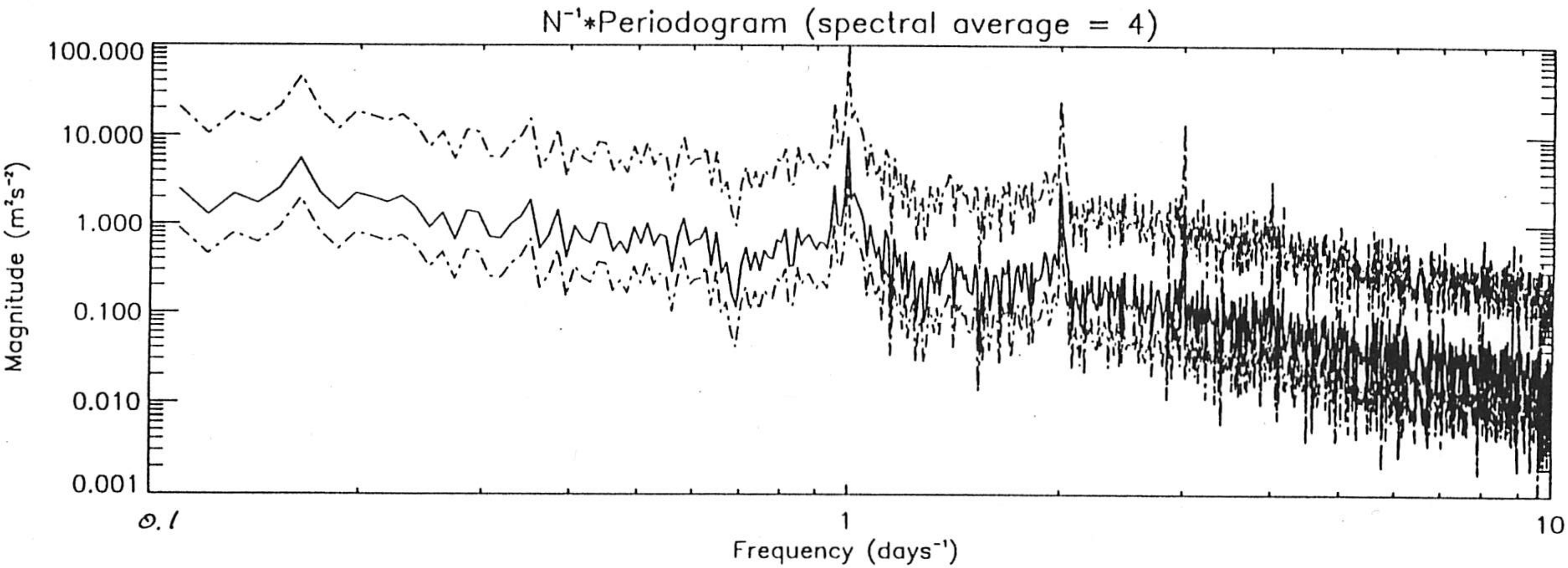
DEGREES OF FREEDOM	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$
1	.0000393	.0001571	.0009821	.0039321	.0157908
2	.0100251	.0201007	.0506356	.102587	.210720
3	.0717212	.114832	.215795	.351846	.584375
4	.206990	.297110	.484419	.710721	1.063623
5	.411740	.554300	.831211	1.145476	1.61031
6	.675727	.872085	1.237347	1.63539	2.20413
7	.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100	67.3276	70.0648	74.2219	77.9295	82.3581

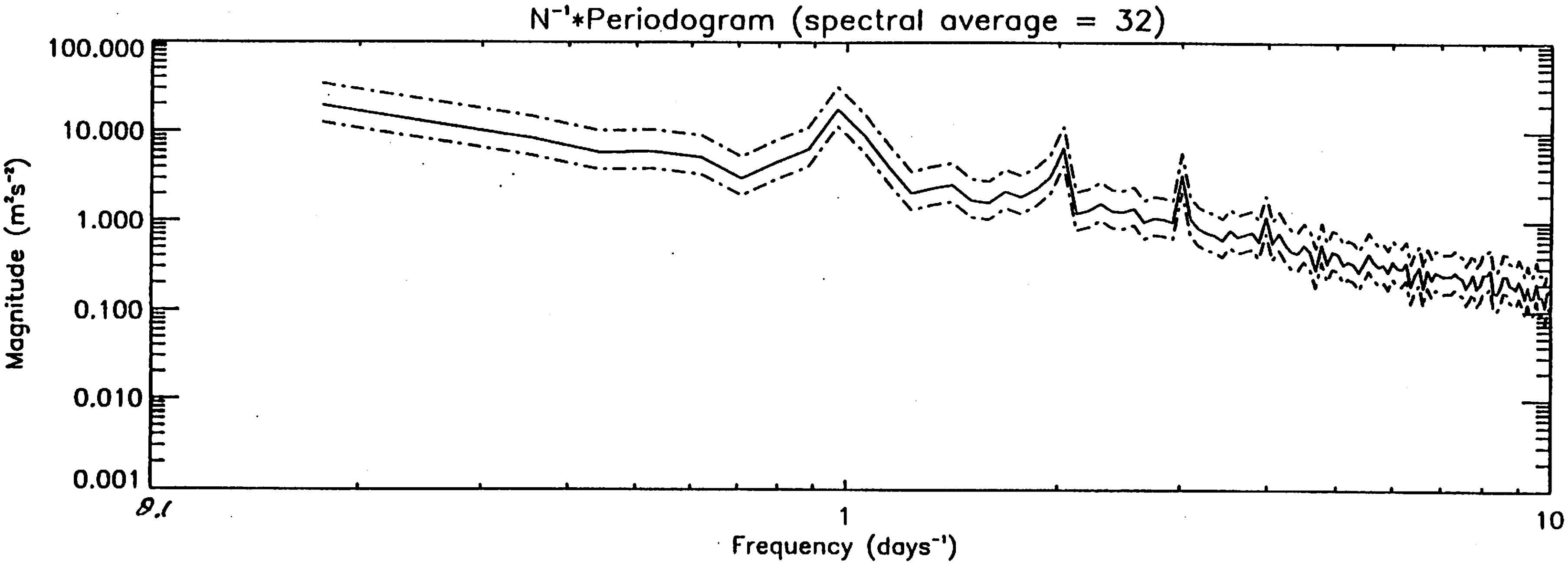
Source: From C. M. Thompson, "Tables of the Percentage Points of the χ^2-Distribution," *Biometrika*, 1941, 32, 188–189. Reproduced by permission of the *Biometrika* Trustees.

DEGREES OF FREEDOM	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	2.70554	3.84146	5.02389	6.63490	7.87944
2	4.60517	5.99147	7.37776	9.21034	10.5966
3	6.25139	7.81473	9.34840	11.3449	12.8381
4	7.77944	9.48773	11.1433	13.2767	14.8602
5	9.23635	11.0705	12.8325	15.0863	16.7496
6	10.6446	12.5916	14.4494	16.8119	18.5476
7	12.0170	14.0671	16.0128	18.4753	20.2777
8	13.3616	15.5073	17.5346	20.0902	21.9550
9	14.6837	16.9190	19.0228	21.6660	23.5893
10	15.9871	18.3070	20.4831	23.2093	25.1882
11	17.2750	19.6751	21.9200	24.7250	26.7569
12	18.5494	21.0261	23.3367	26.2170	28.2995
13	19.8119	22.3621	24.7356	27.6883	29.8194
14	21.0642	23.6848	26.1190	29.1413	31.3193
15	22.3072	24.9958	27.4884	30.5779	32.8013
16	23.5418	26.2962	28.8454	31.9999	34.2672
17	24.7690	27.5871	30.1910	33.4087	35.7185
18	25.9894	28.8693	31.5264	34.8053	37.1564
19	27.2036	30.1435	32.8523	36.1908	38.5822
20	28.4120	31.4104	34.1696	37.5662	39.9968
21	29.6151	32.6705	35.4789	38.9321	41.4010
22	30.8133	33.9244	36.7807	40.2894	42.7956
23	32.0069	35.1725	38.0757	41.6384	44.1813
24	33.1963	36.4151	39.3641	42.9798	45.5585
25	34.3816	37.6525	40.6465	44.3141	46.9278
26	35.5631	38.8852	41.9232	45.6417	48.2899
27	36.7412	40.1133	43.1944	46.9630	49.6449
28	37.9159	41.3372	44.4607	48.2782	50.9933
29	39.0875	42.5569	45.7222	49.5879	52.3356
30	40.2560	43.7729	46.9792	50.8922	53.6720
40	51.8050	55.7585	59.3417	63.6907	66.7659
50	63.1671	67.5048	71.4202	76.1539	79.4900
60	74.3970	79.0819	83.2976	88.3794	91.9517
70	85.5271	90.5312	95.0231	100.425	104.215
80	96.5782	101.879	106.629	112.329	116.321
90	107.565	113.145	118.136	124.116	128.299
100	118.498	124.342	129.561	135.807	140.169

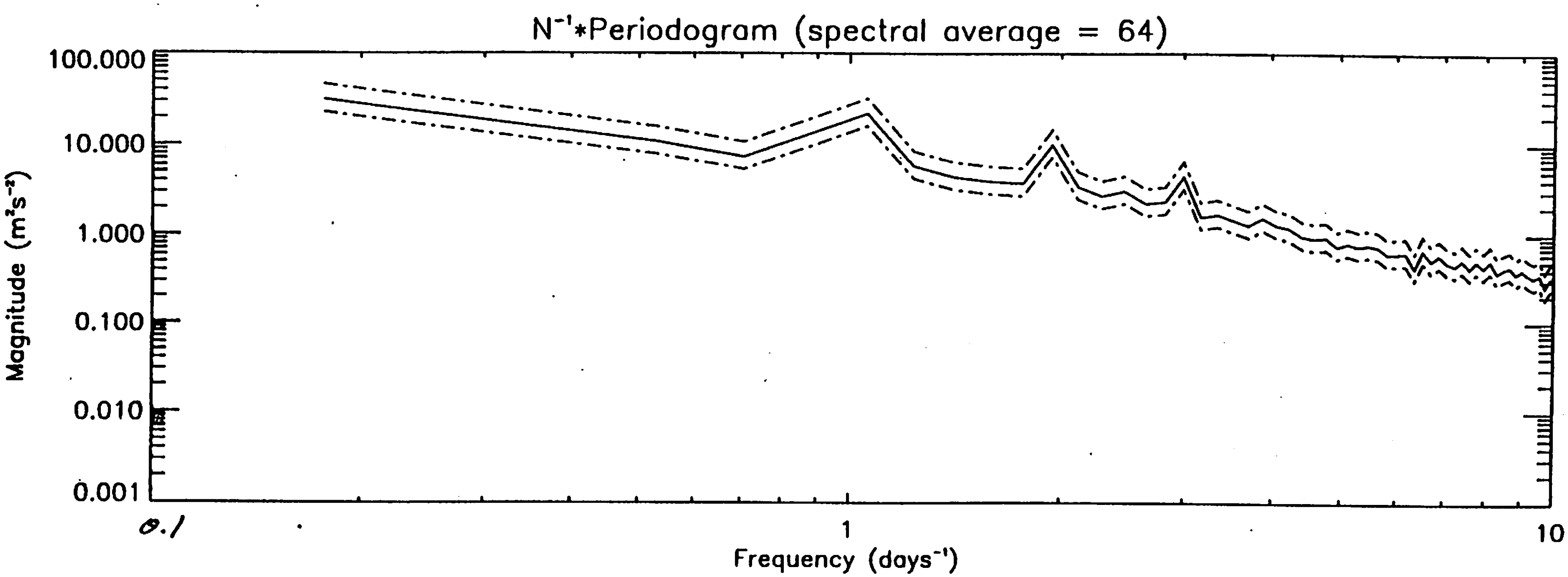
95% Confidence Intervals.



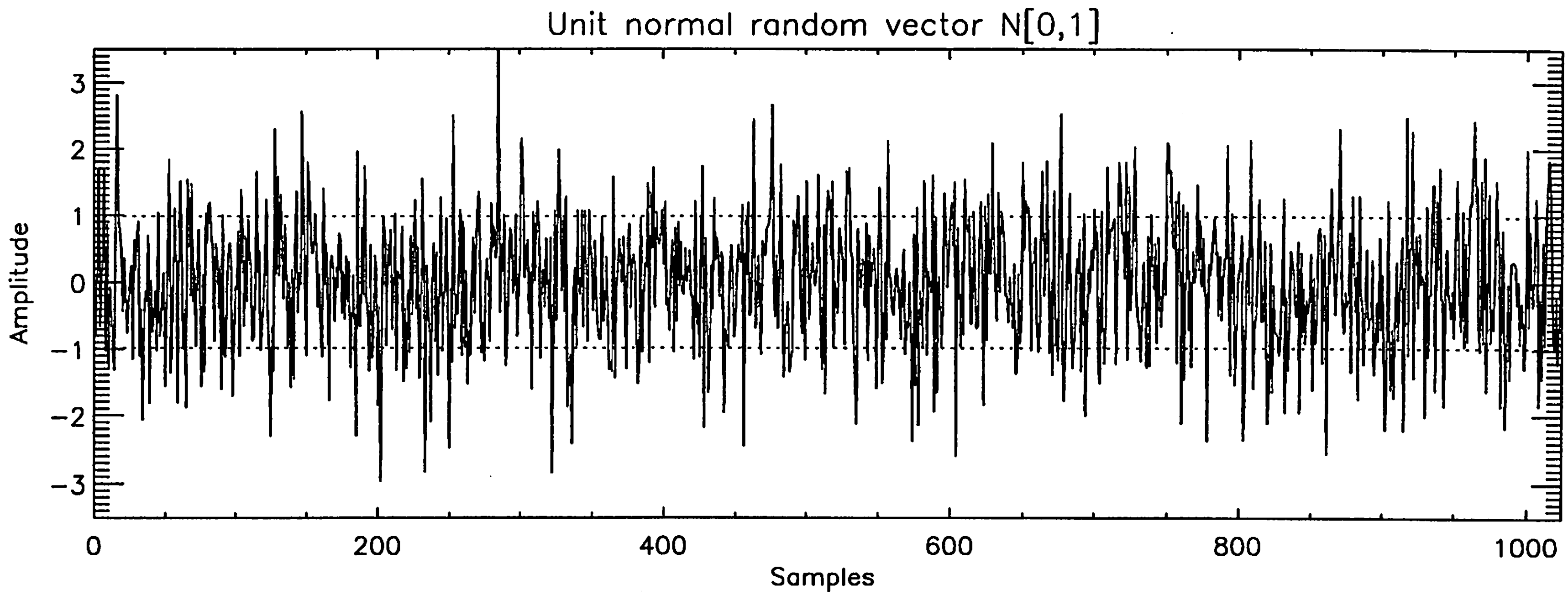


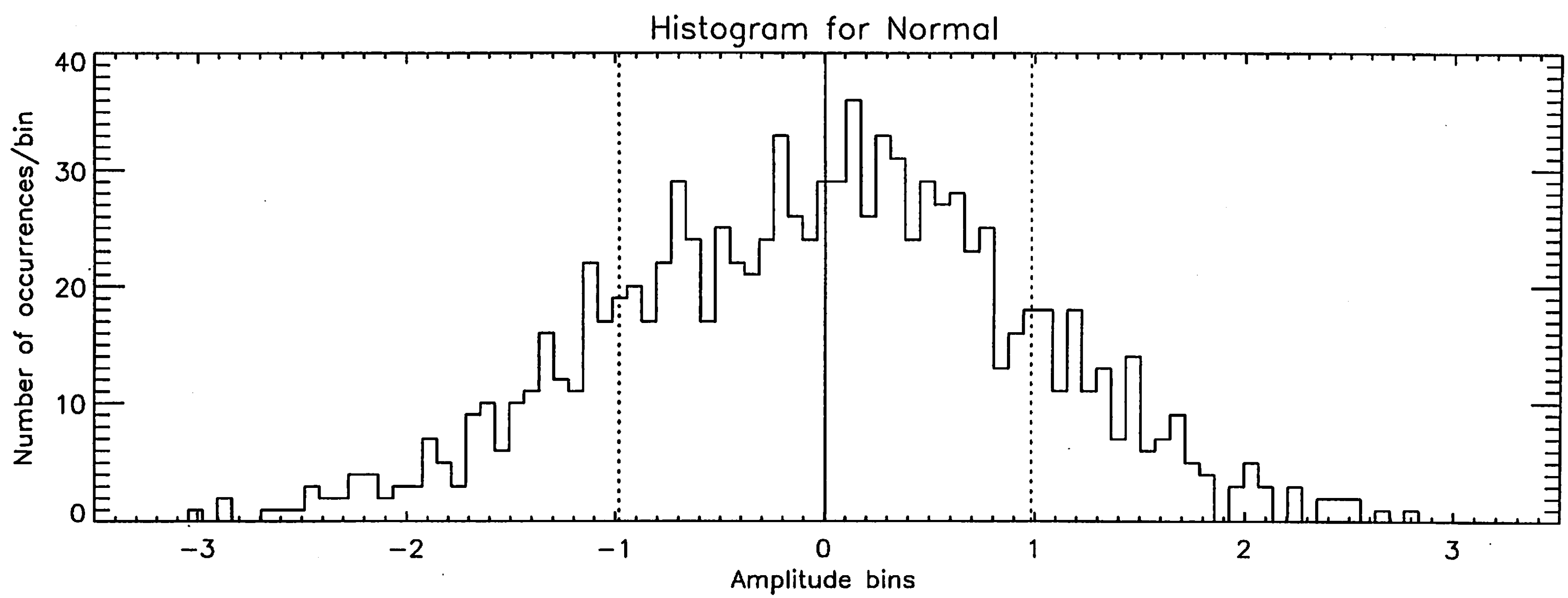


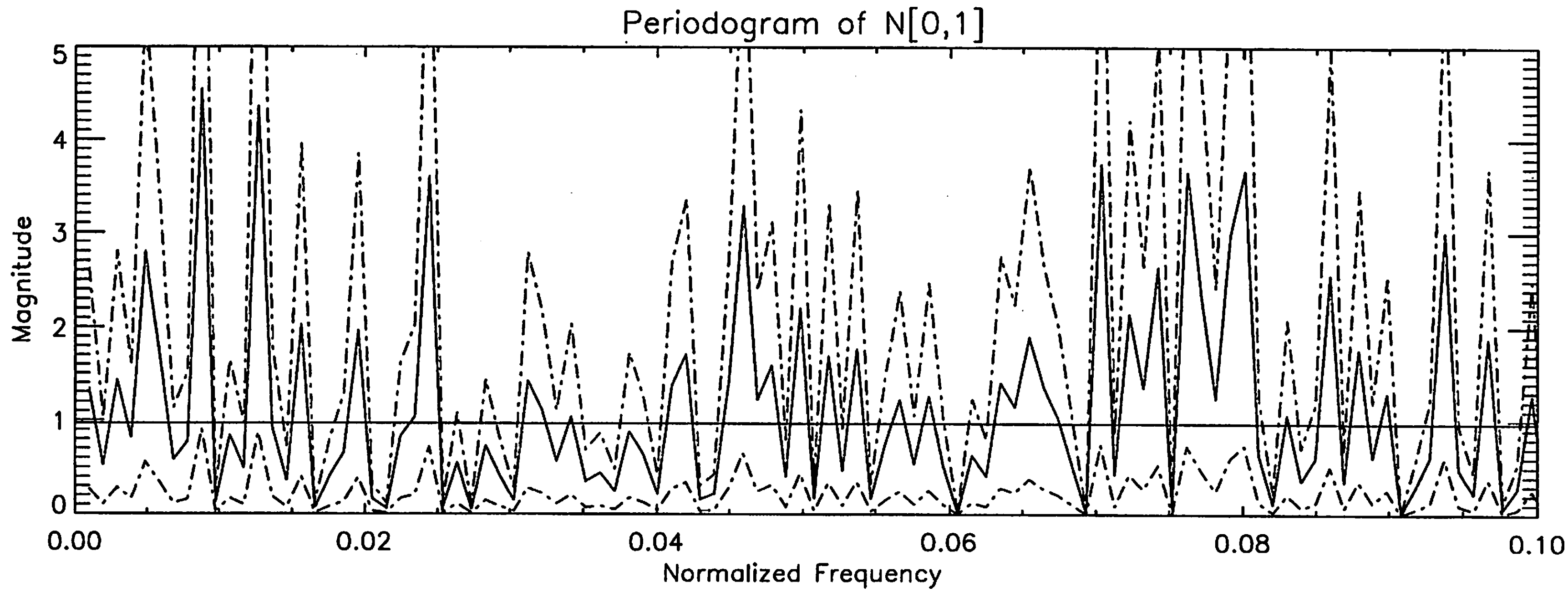
N⁻¹*Periodogram (spectral average = 32)
100.000
10.000
1.000
0.100
0.010
0.001
Magnitude (m²s⁻²)
0.1
1
10
Frequency (days⁻¹)



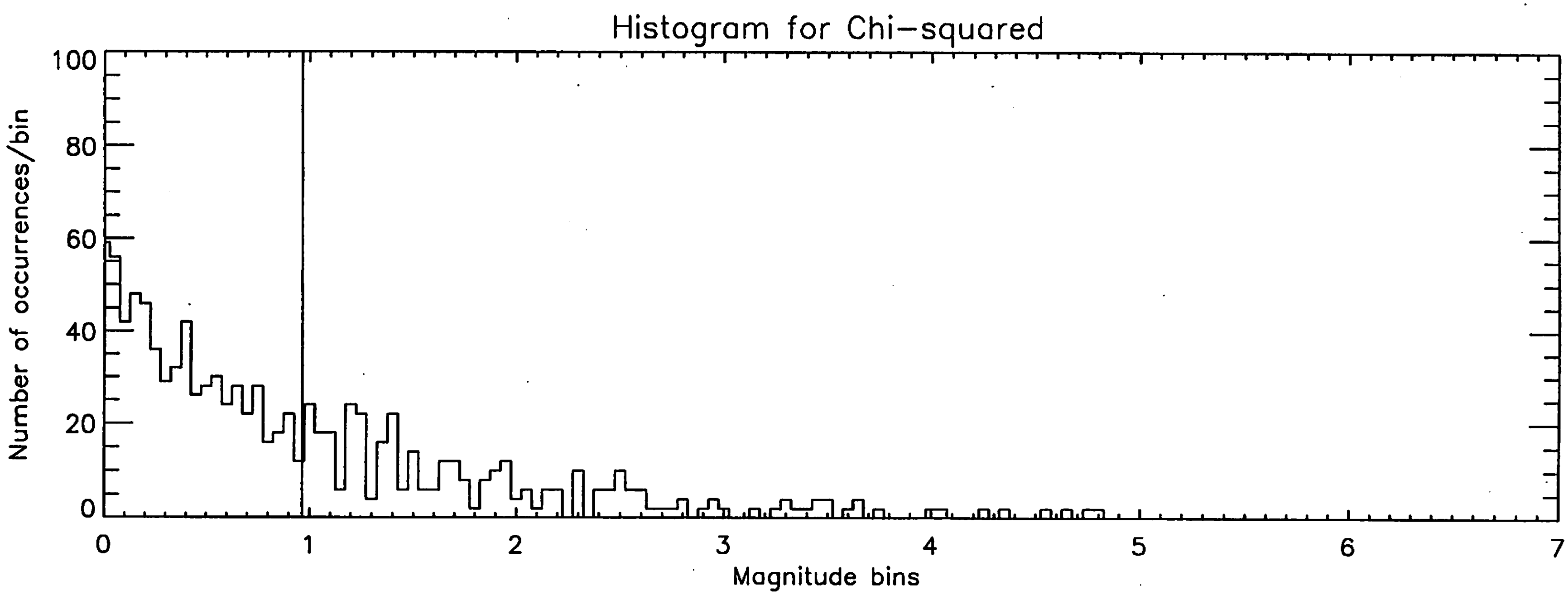
N⁻¹*Periodogram (spectral average = 64)
100.000
10.000
1.000
0.100
0.010
0.001
Magnitude (m²s⁻²)
0.1
1
10
Frequency (days⁻¹)


Unit normal random vector N[0,1]
Amplitude
Samples
3
2
1
0
−1
−2
−3
0
200
400
600
800
1000

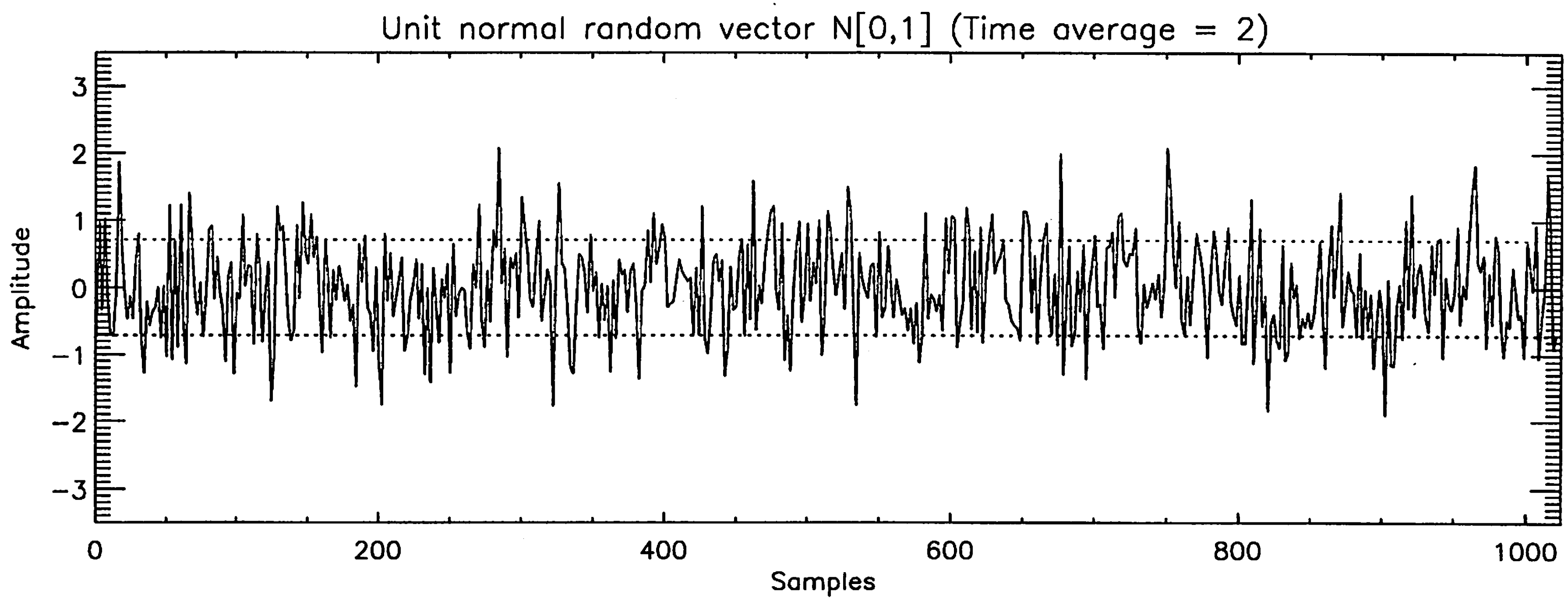

Histogram for Normal
Number of occurrences/bin
Amplitude bins
40
30
20
10
0
−3
−2
−1
0
1
2
3



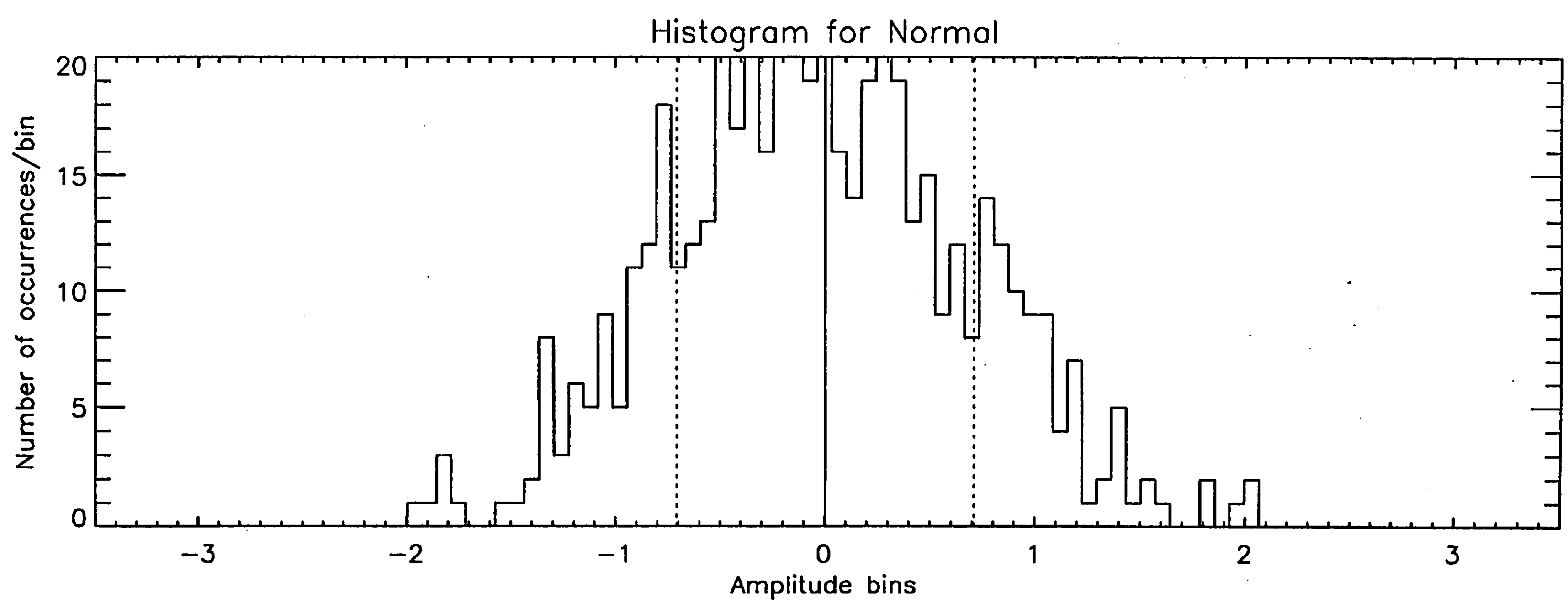
Periodogram of N[0,1]
Magnitude
Normalized Frequency
0.00
0.02
0.04
0.06
0.08
0.10
0
1
2
3
4
5



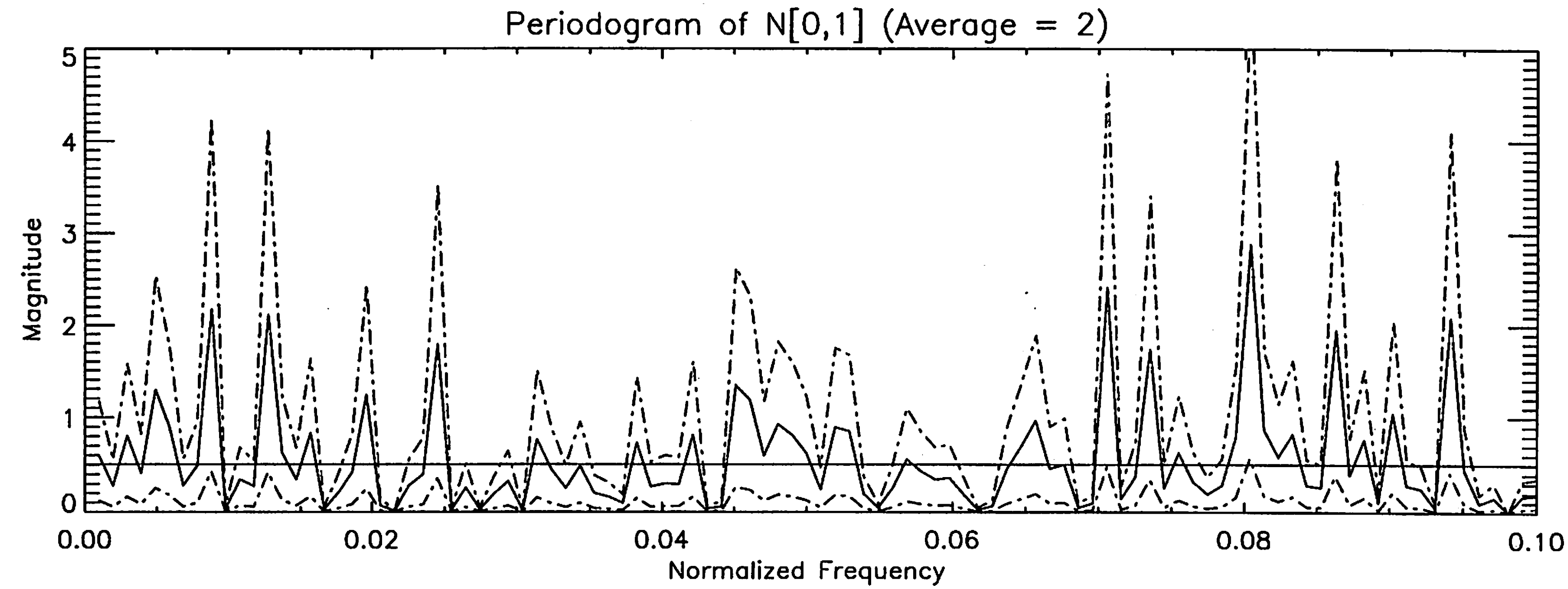
Histogram for Chi-squared
Number of occurrences/bin
Magnitude bins
0
20
40
60
80
100
0
1
2
3
4
5
6
7



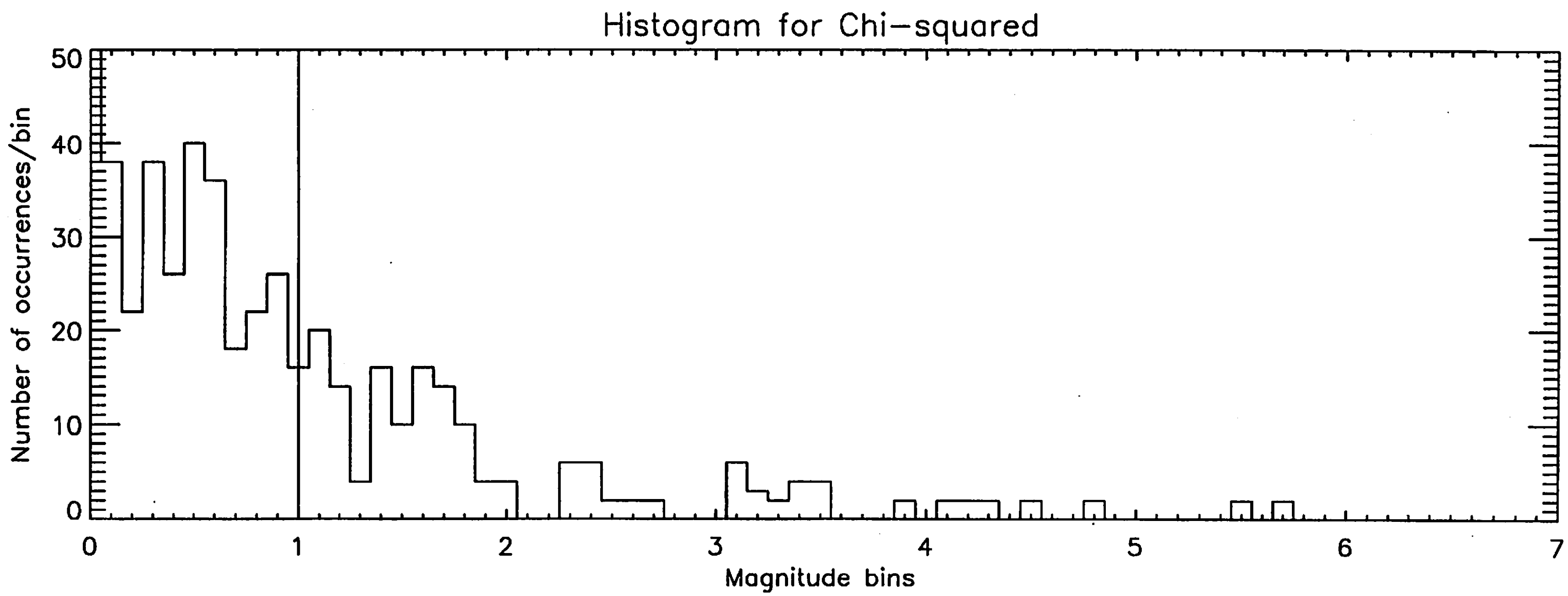
Unit normal random vector N[0,1] (Time average = 2)
Amplitude
3
2
1
0
−1
−2
−3
0
200
400
600
800
1000
Samples



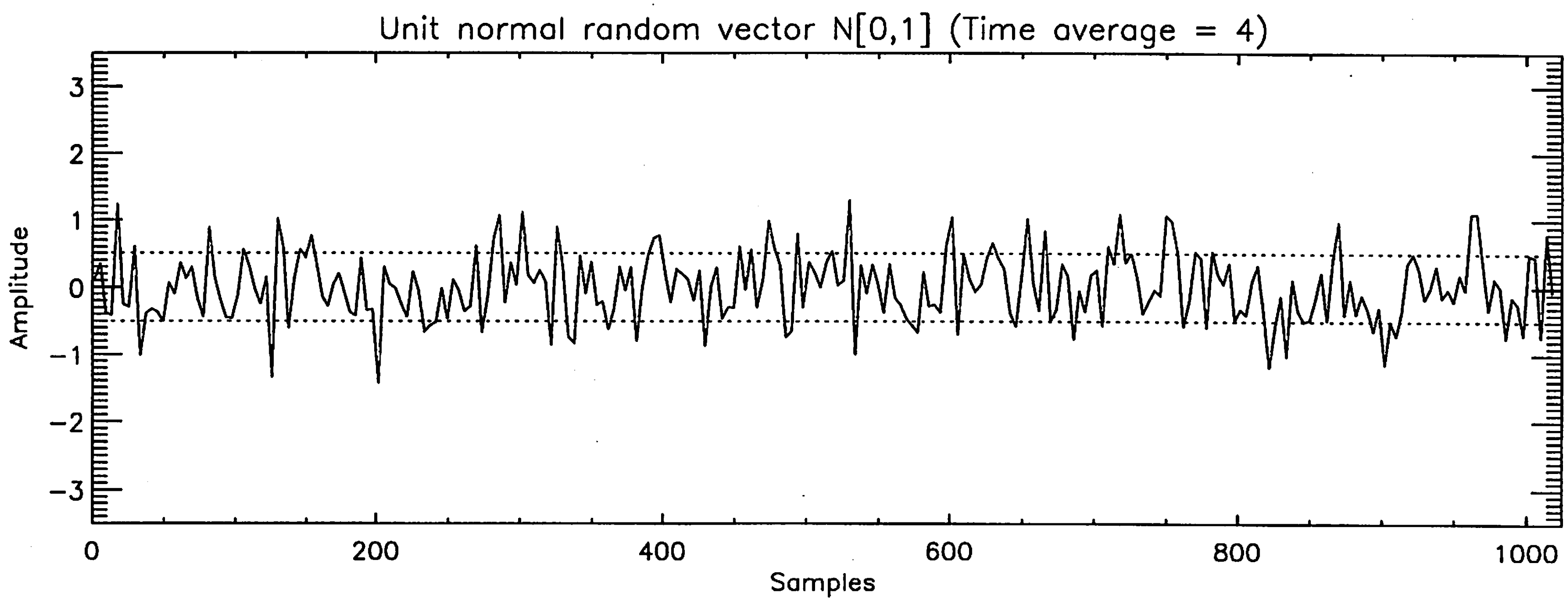
Histogram for Normal
Number of occurrences/bin
20
15
10
5
0
−3
−2
−1
0
1
2
3
Amplitude bins



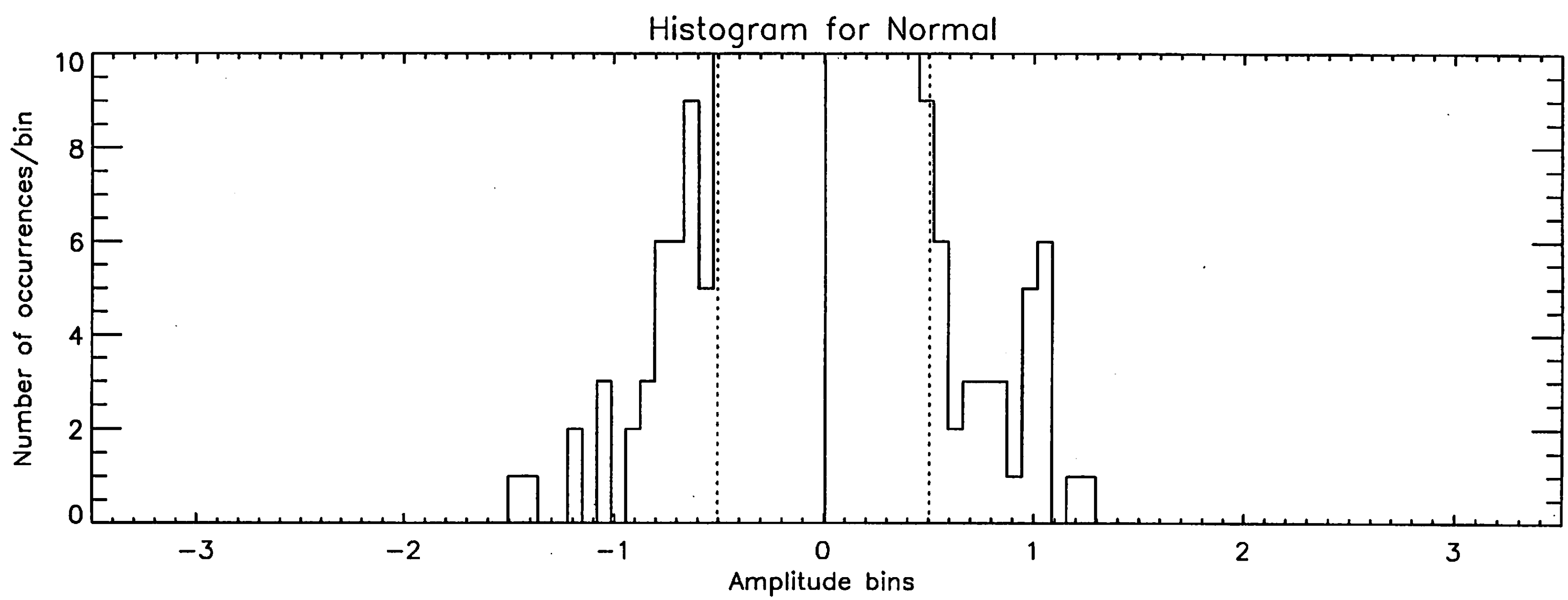
Periodogram of N[0,1] (Average = 2)
Magnitude
Normalized Frequency
0
1
2
3
4
5
0.00
0.02
0.04
0.06
0.08
0.10



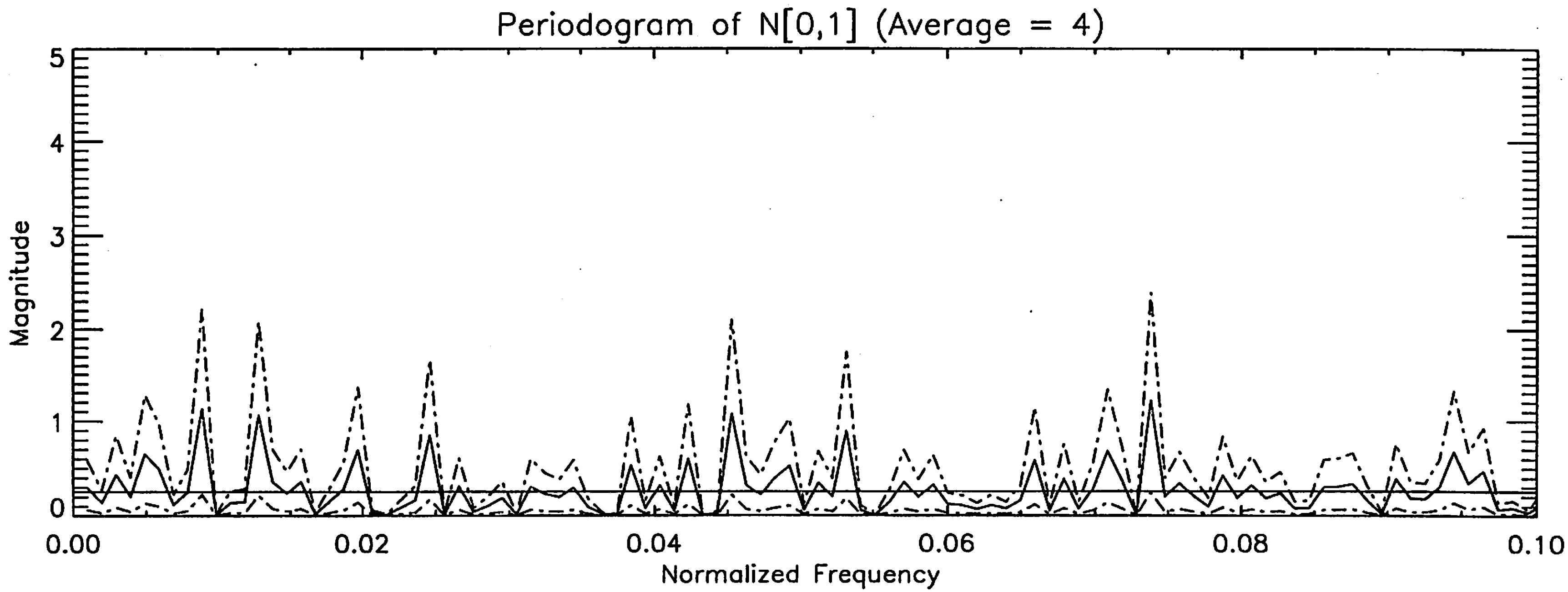
Histogram for Chi-squared
Number of occurrences/bin
Magnitude bins
0
10
20
30
40
50
0
1
2
3
4
5
6
7



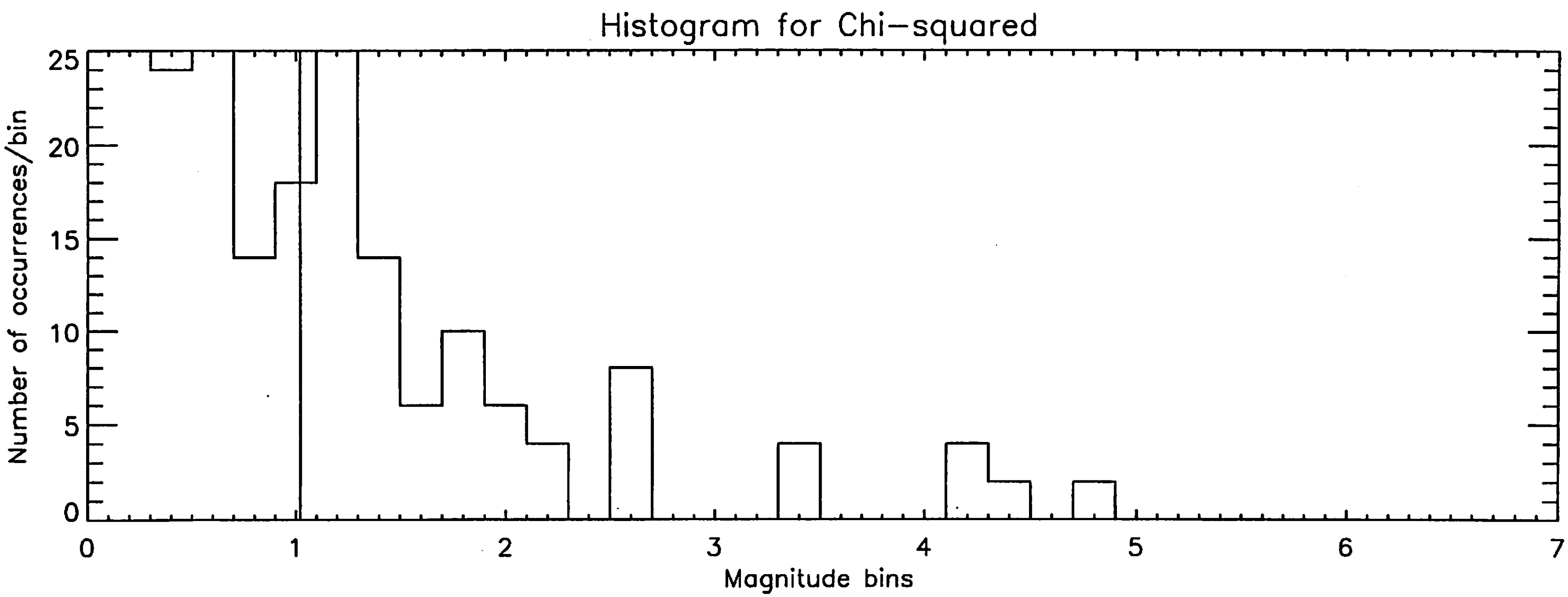
Unit normal random vector N[0,1] (Time average = 4)
Amplitude
Samples
3
2
1
0
−1
−2
−3
0
200
400
600
800
1000



Histogram for Normal
Number of occurrences/bin
Amplitude bins
10
8
6
4
2
0
−3
−2
−1
0
1
2
3

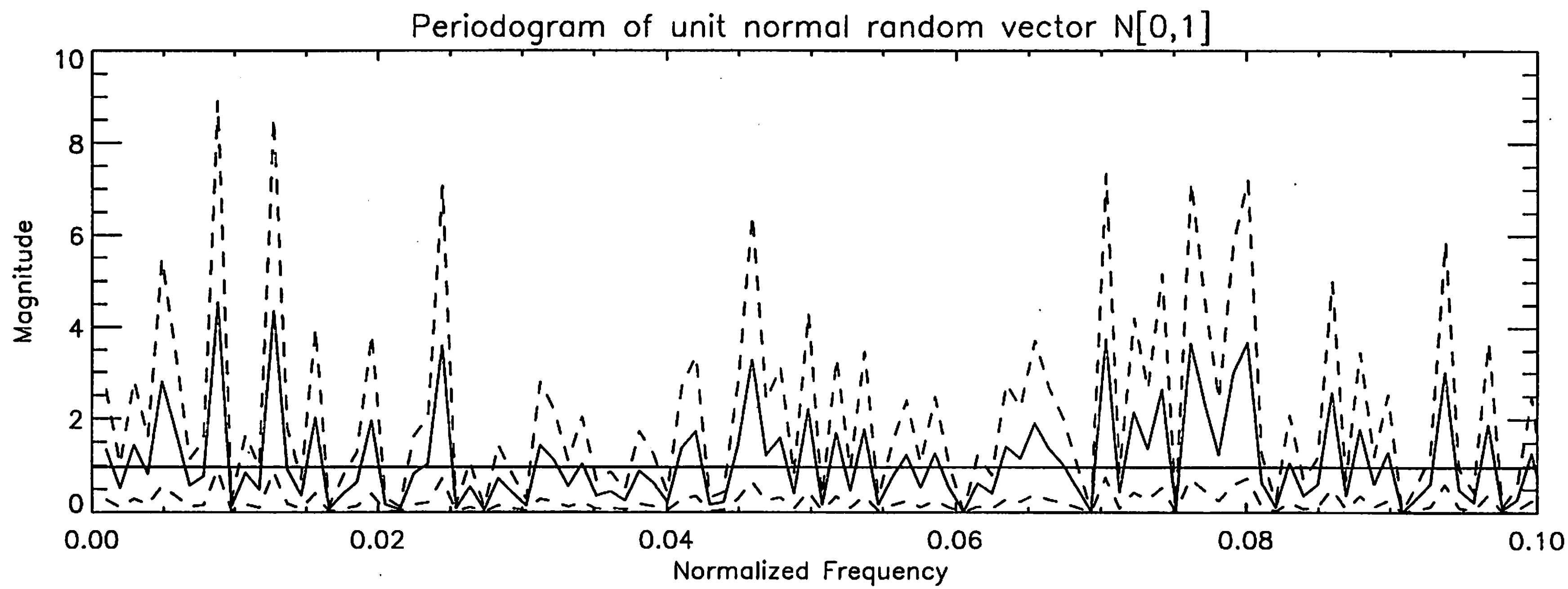


Periodogram of N[0,1] (Average = 4)
Magnitude
0
1
2
3
4
5
0.00
0.02
0.04
0.06
0.08
0.10
Normalized Frequency

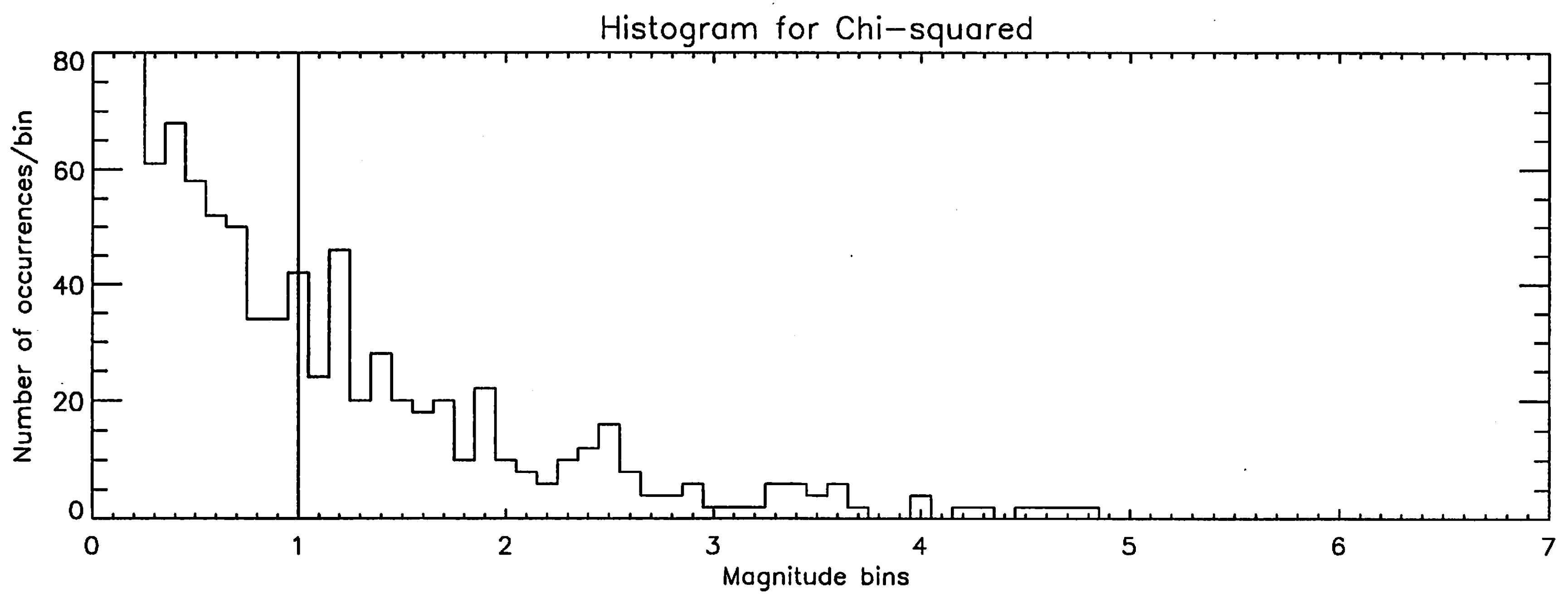


Histogram for Chi-squared
Number of occurrences/bin
0
5
10
15
20
25
0
1
2
3
4
5
6
7
Magnitude bins

α = 0.5

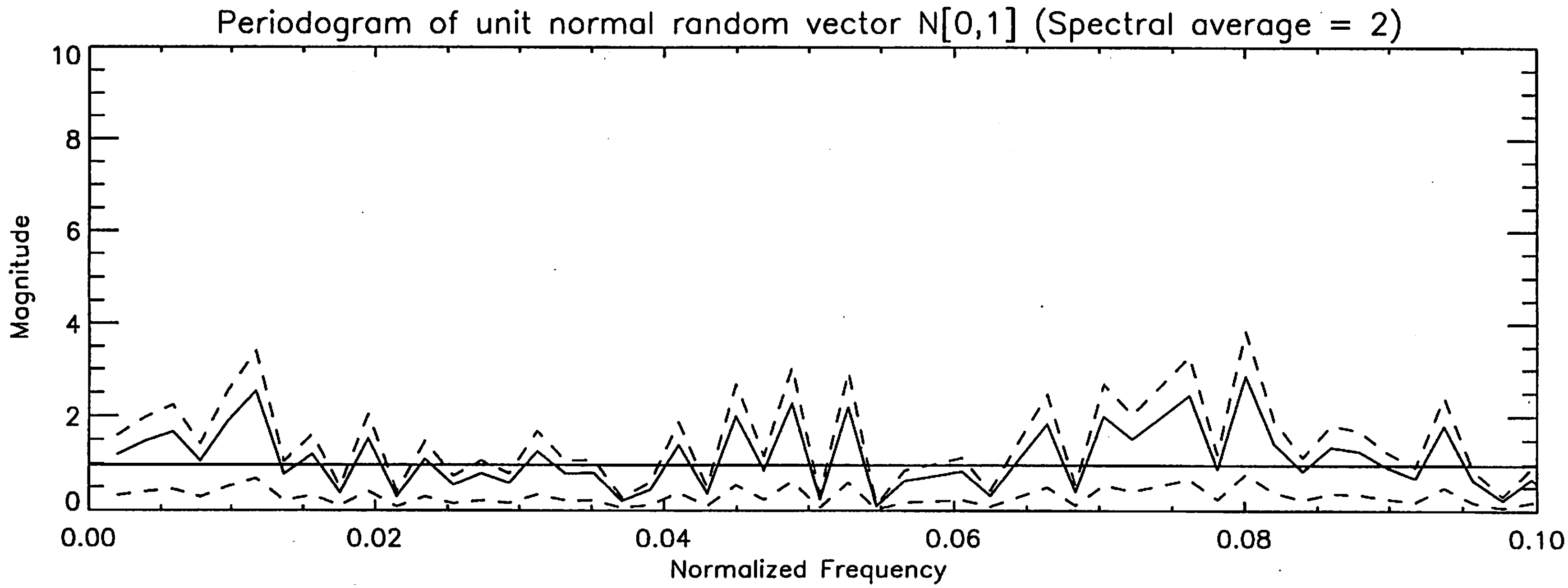


Periodogram of unit normal random vector N[0,1]
Magnitude
Normalized Frequency
0.00
0.02
0.04
0.06
0.08
0.10
0
2
4
6
8
10

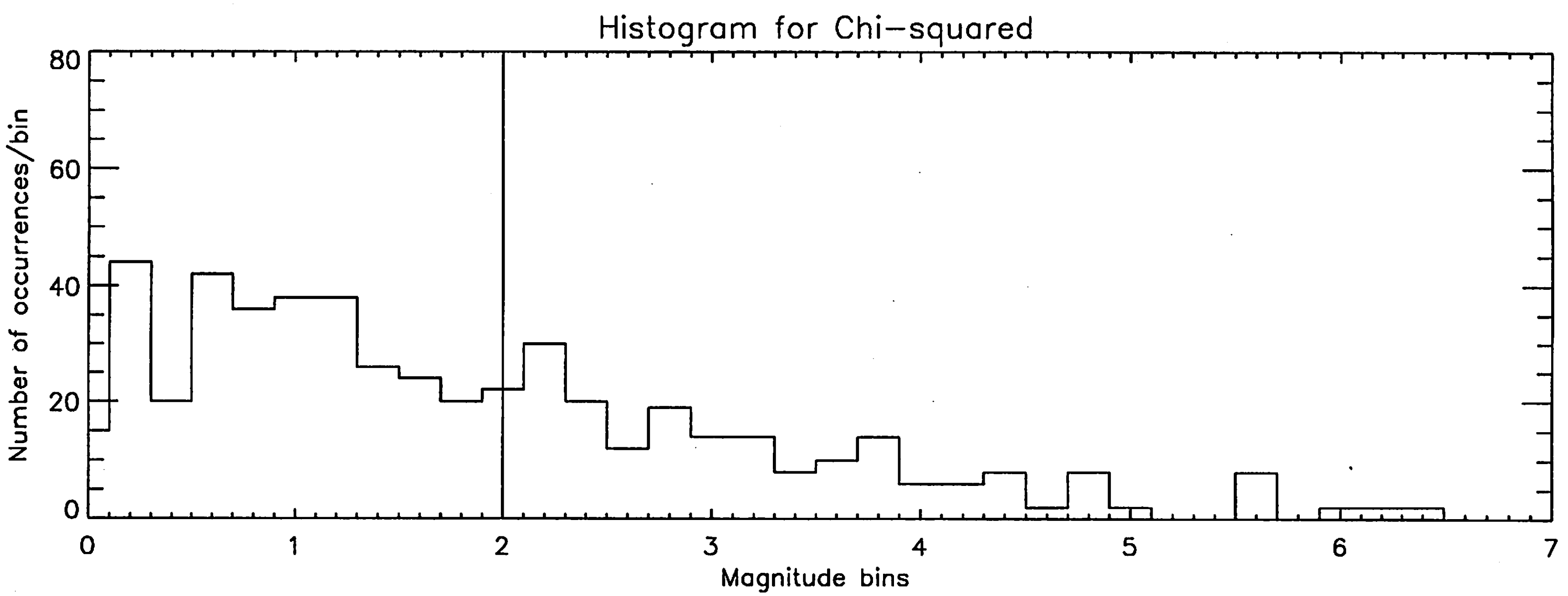


Histogram for Chi-squared
Number of occurrences/bin
Magnitude bins
0
1
2
3
4
5
6
7
0
20
40
60
80

α= 0.5

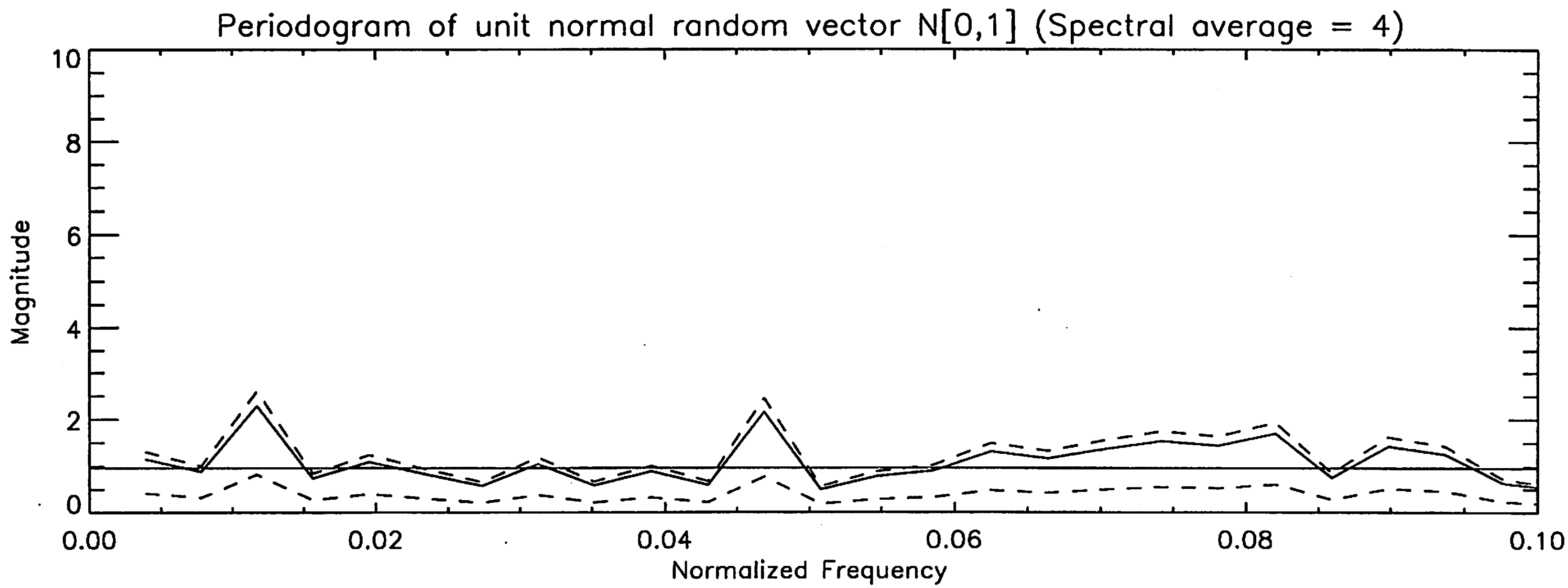


Periodogram of unit normal random vector N[0,1] (Spectral average = 2)
Magnitude
10
8
6
4
2
0
0.00
0.02
0.04
0.06
0.08
0.10
Normalized Frequency

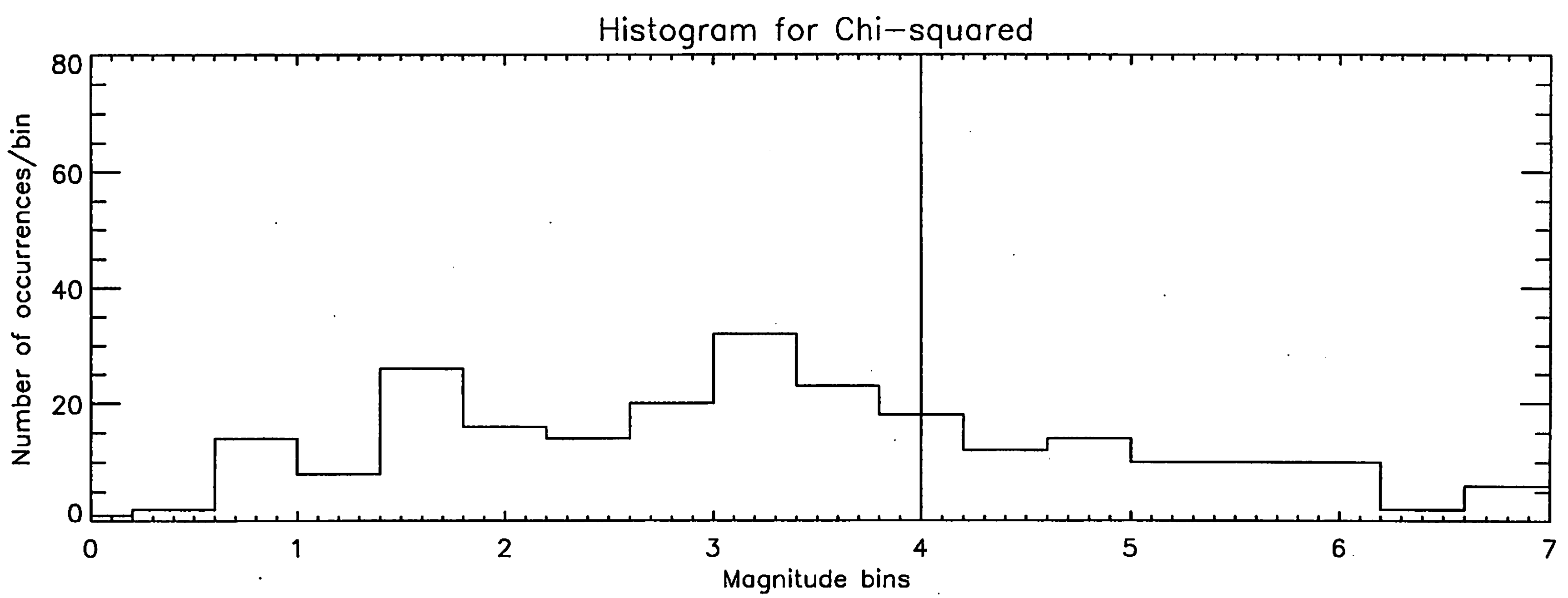


Histogram for Chi-squared
Number of occurrences/bin
80
60
40
20
0
0
1
2
3
4
5
6
7
Magnitude bins

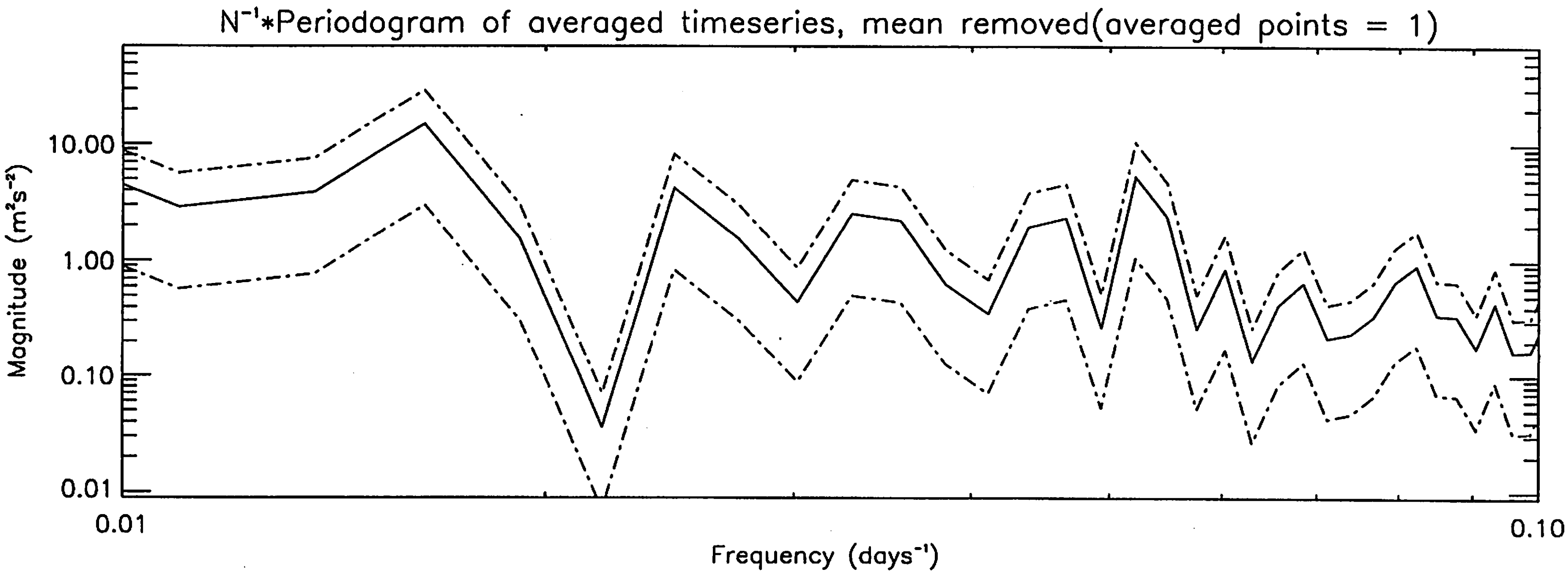
α = 0.5



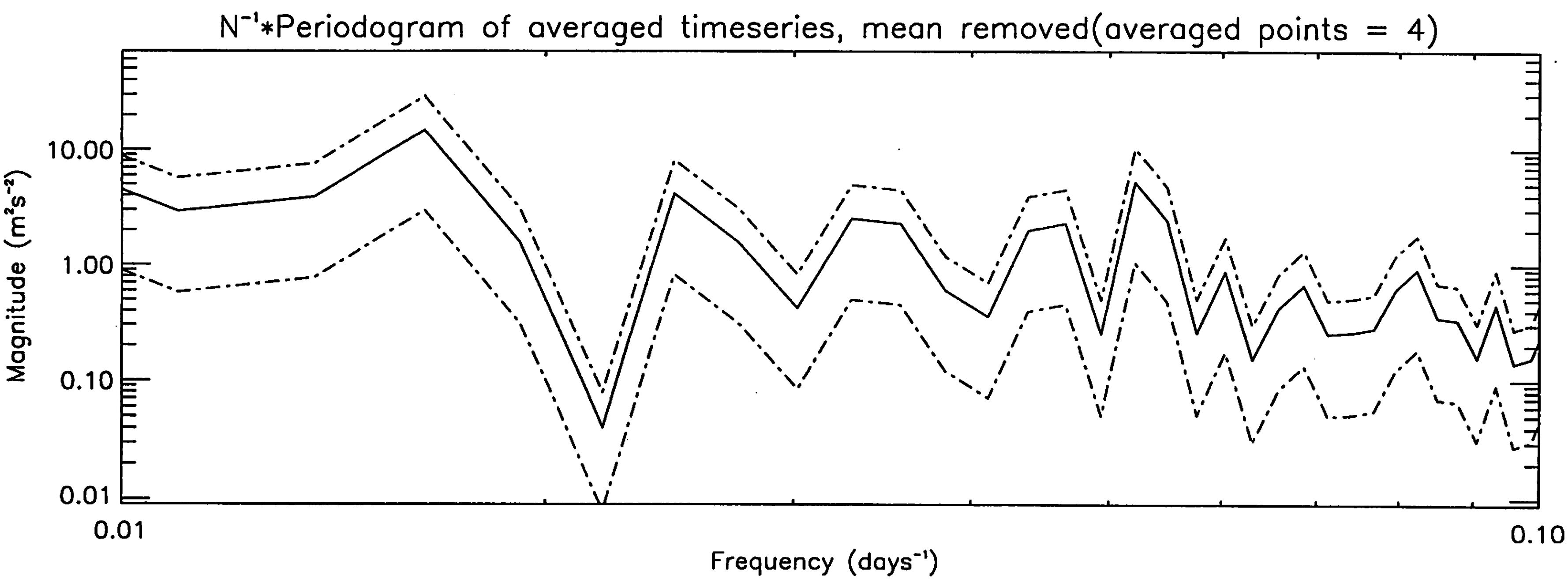
Periodogram of unit normal random vector N[0,1] (Spectral average = 4)
Magnitude
10
8
6
4
2
0
0.00
0.02
0.04
0.06
0.08
0.10
Normalized Frequency



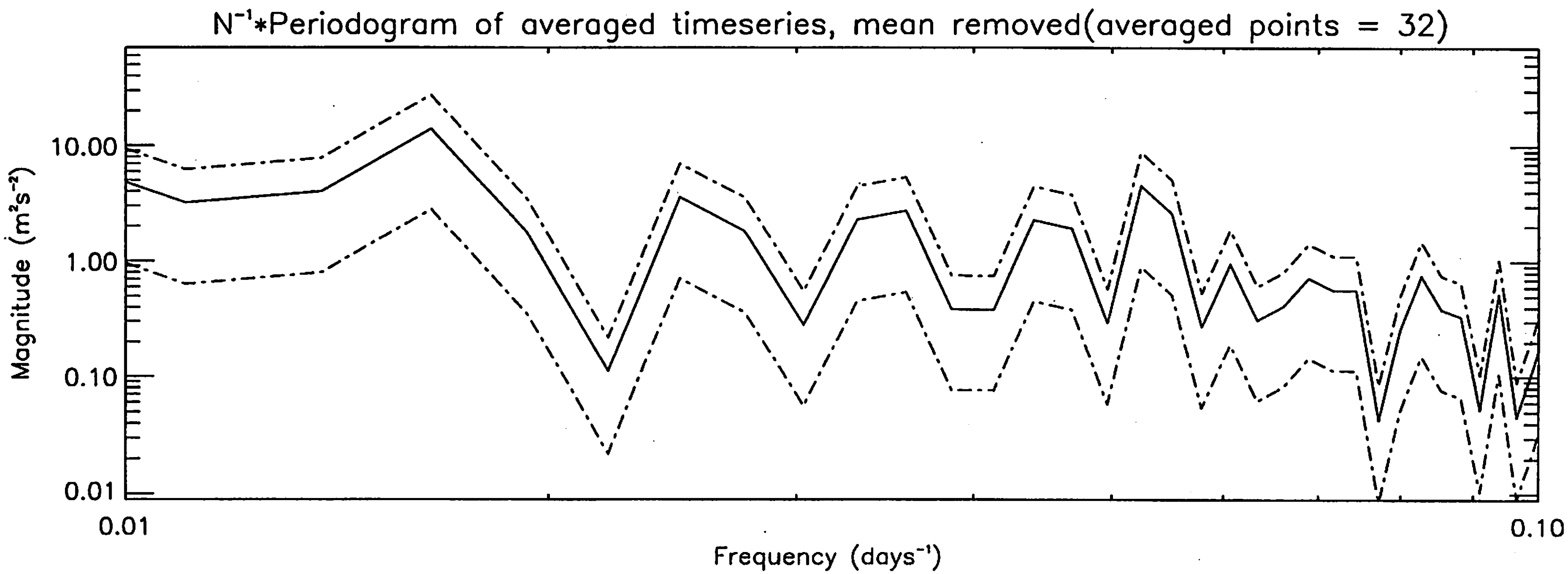
Histogram for Chi-squared
Number of occurrences/bin
80
60
40
20
0
0
1
2
3
4
5
6
7
Magnitude bins

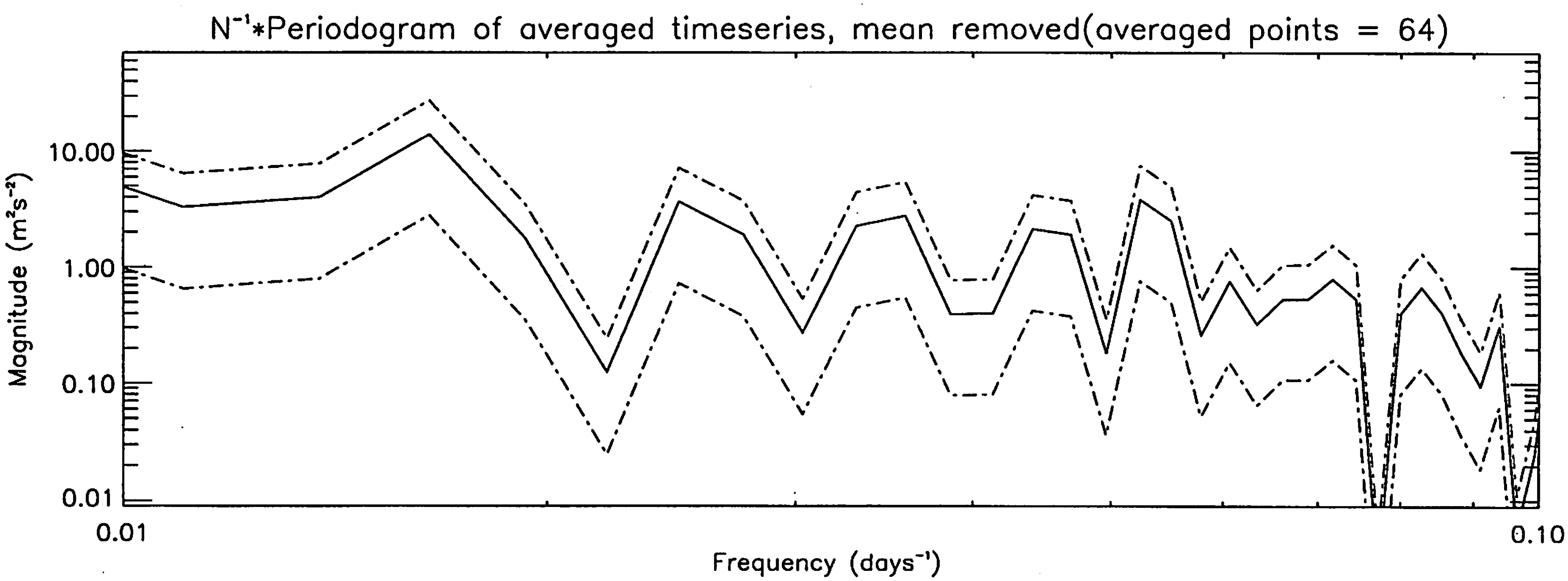


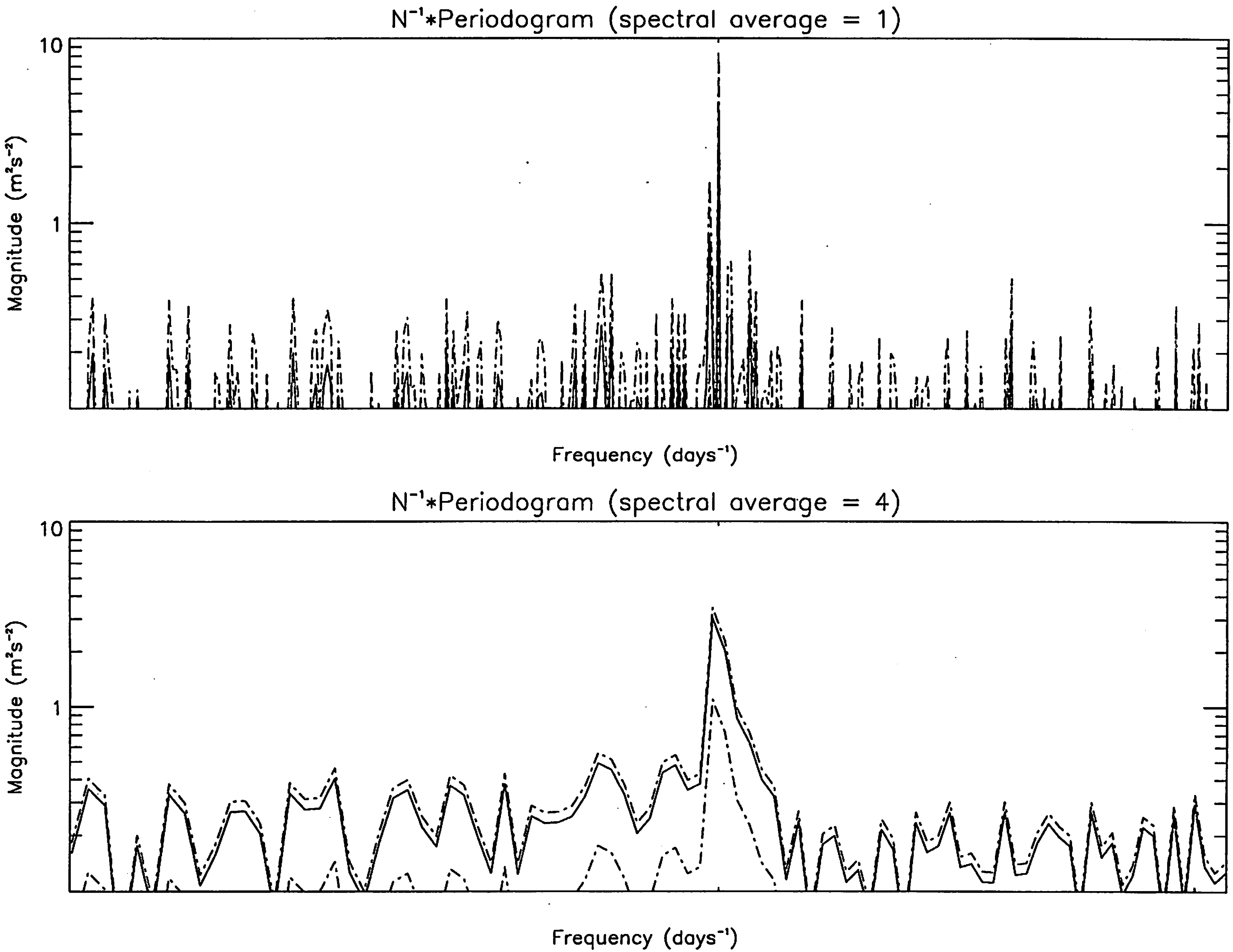
N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 1)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
Frequency (days⁻¹)



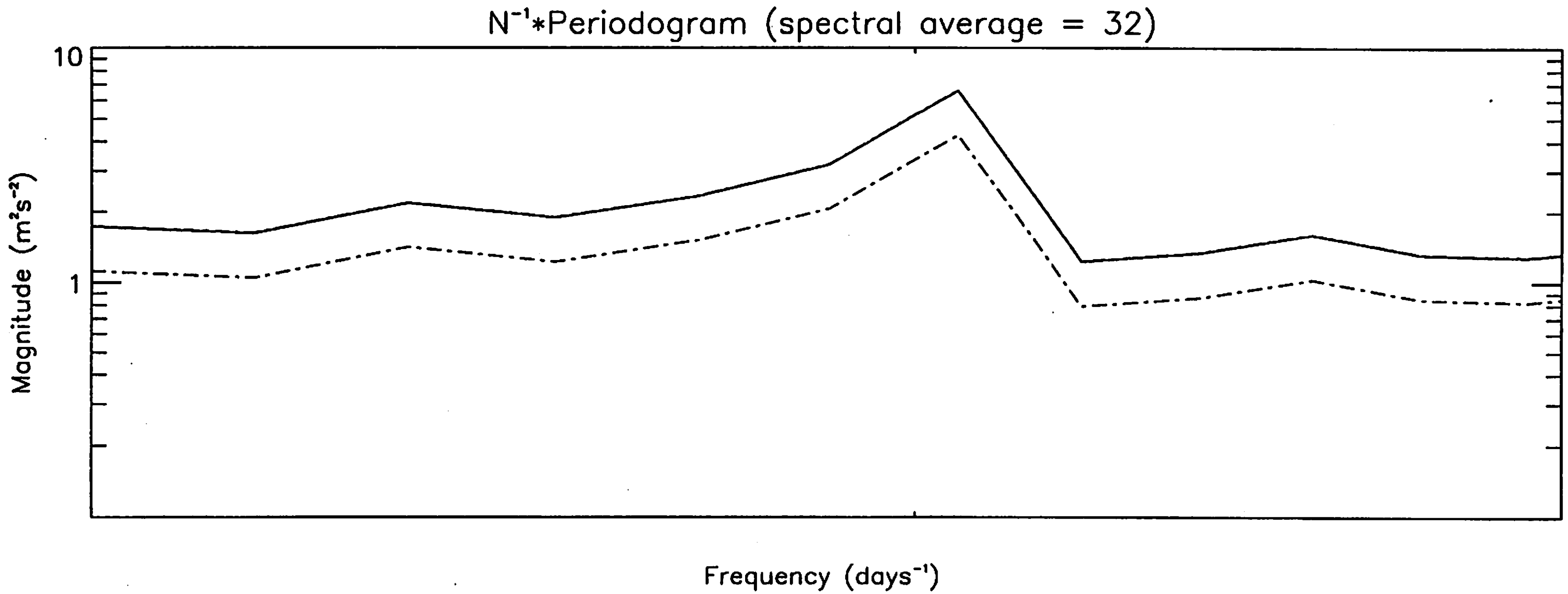
N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 4)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
Frequency (days⁻¹)

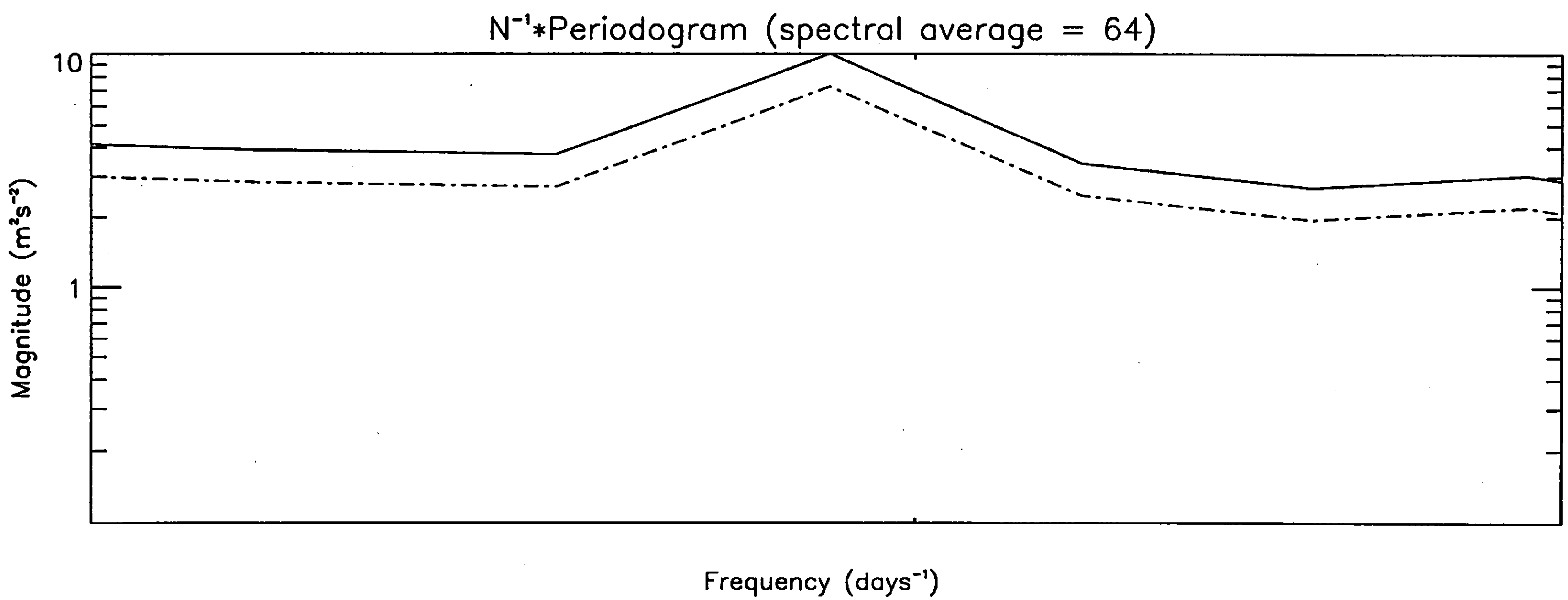

N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 32)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
Frequency (days⁻¹)


N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 64)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
Frequency (days⁻¹)



N⁻¹*Periodogram (spectral average = 1)
10
1
Magnitude (m²s⁻²)
Frequency (days⁻¹)
N⁻¹*Periodogram (spectral average = 4)
10
1
Magnitude (m²s⁻²)
Frequency (days⁻¹)


N⁻¹*Periodogram (spectral average = 32)
10
1
Magnitude (m^2s^{-2})
Frequency ($days^{-1}$)


N⁻¹*Periodogram (spectral average = 64)
10
1
Magnitude (m^2s^{-2})
Frequency ($days^{-1}$)

Significance Level:

Trying to determine if signal is present. A 90% significance level is saying that given only noise, 10% of the spectral points on average will exceed this level.

Know now that $\sigma^2 P(k) \sim \chi_1^2$ **with 1 degree of freedom.**

However, σ^2 **is unknown. Can (1) estimate** σ^2 **and apply** χ^2 **test or (2) use another test statistic.**

F distribution:

Let $\chi_{n_1}^2$ **and** $\chi_{n_2}^2$ **be chi-square random variables with** n_1 **and** n_2 **degrees of freedom. Then if** $\chi_{n_1}^2$ **and** $\chi_{n_2}^2$ **are independent**

$$F = \frac{\chi_1^2/n_1}{\chi_2^2/n_2}$$ **is said to have an F distribution with** n_1

numerator degrees of freedom and n_2 **denominator degrees of freedom.**

F test - check kth component of periodogram relative to other spectral components.

$$\frac{\frac{1}{\sigma^2} P(k)}{\frac{1}{\sigma^2} \sum_{\substack{j=0 \\ j \neq k}}^{N-1} P(j)/(N-1)} \sim \frac{\chi_1^2}{\chi_{N-1}^2/N-1} \sim F(1,N-1)$$

Test statistic becomes (after rewriting)

$$F = \frac{(N-1)\, P(k)}{\sum_{j=0}^{n-1} P(j) - P(k)} \geq F_0$$

Note F is invariant to σ^2.

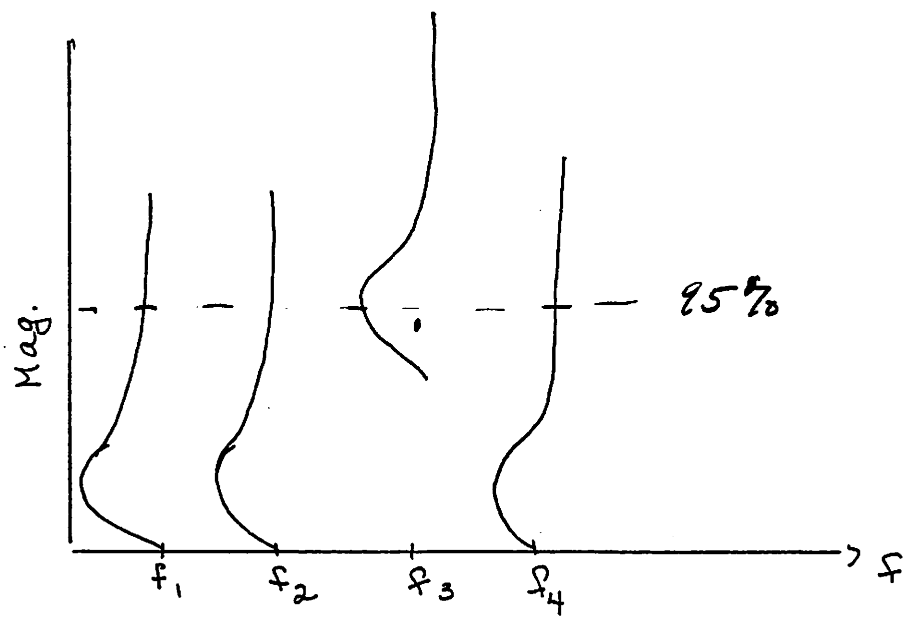
Look up F_0 in F table for given significance level and compare to test statistic.

If $F > F_0$ component is significant
$F < F_0$ component is not significant

Can transform test such that periodogram is test statistic.

$$P(k) \geq \frac{F_0}{N-1+F_0} \sum_{j=0}^{N-1} |X(j)|^2$$

If a signal is present then $\frac{P(k)}{\sigma^2}$ will be a chi-square random variable with nonzero non-centrality parameter (shift).



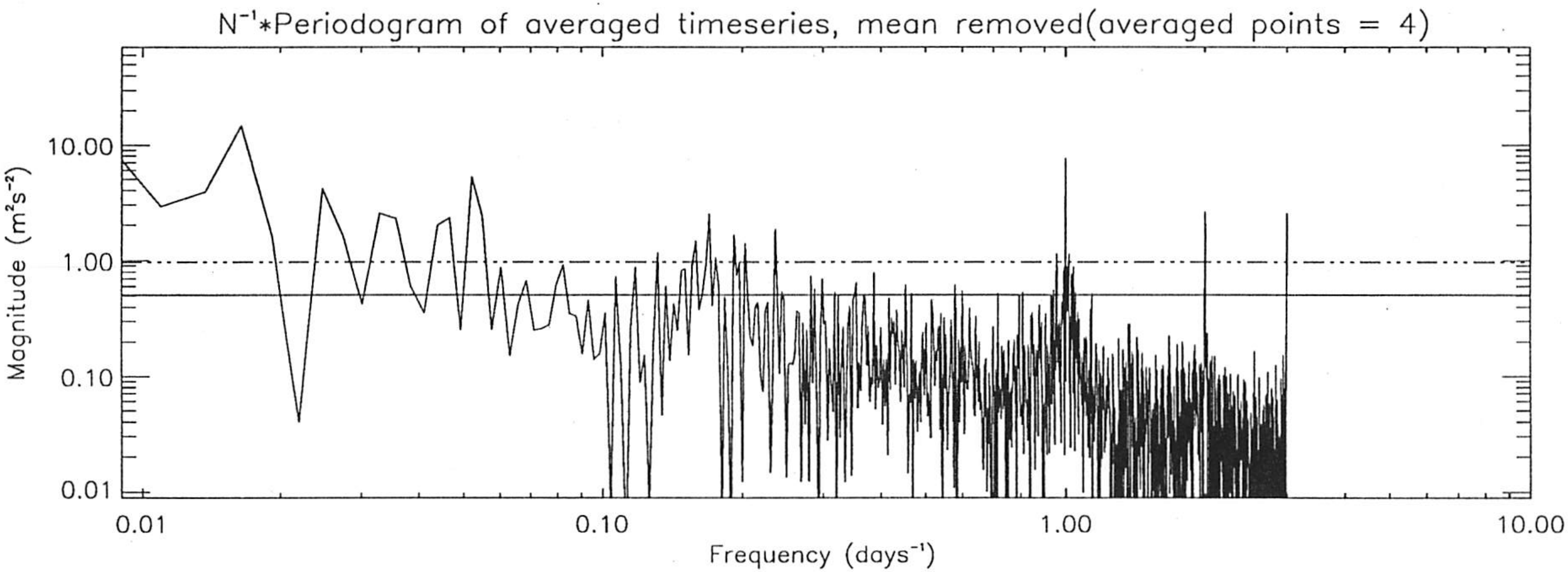
Percentage Points of the F Distributions
$\alpha = .05 \Rightarrow$ 95% significance level
$F_0 = 3.84$

TABLE 7 (*Continued*)

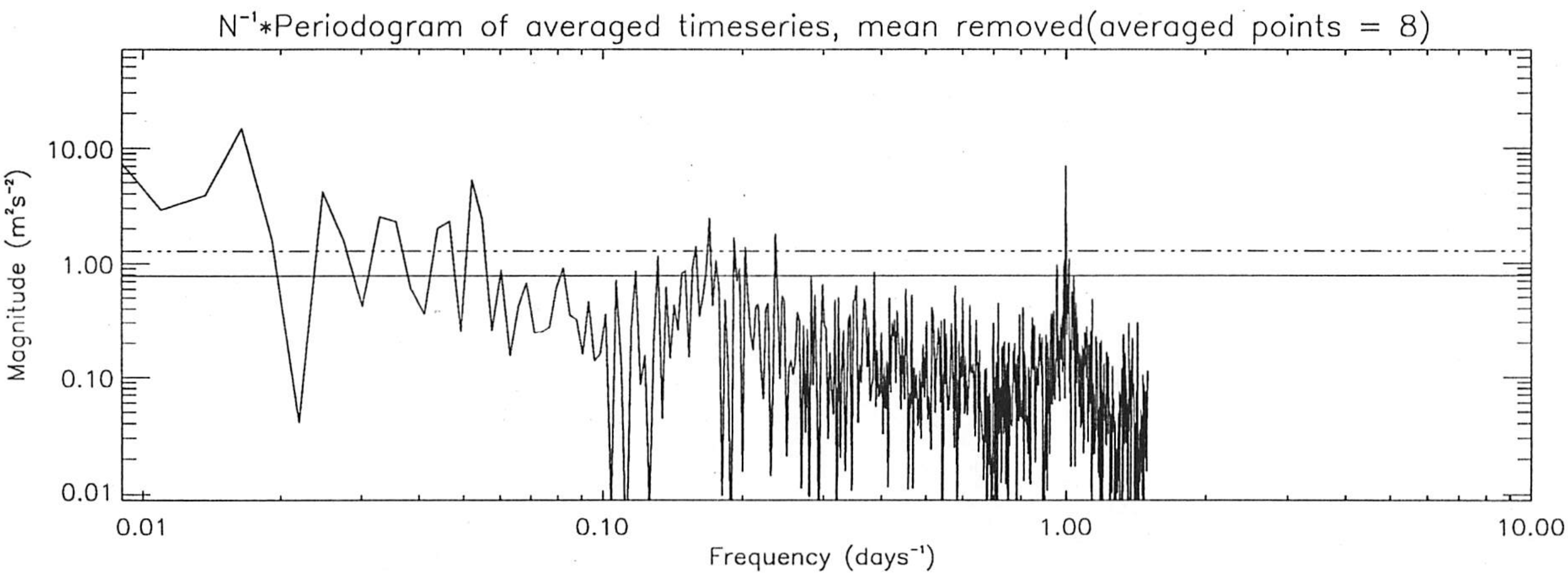
F_α

Denominator df	α	Numerator df 1	2	3	4	5	6	7	8	9
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.005	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.005	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.005	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.005	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
120	.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.050	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.005	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
∞	.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.050	3.84 (F_0)	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	.025	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.005	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62

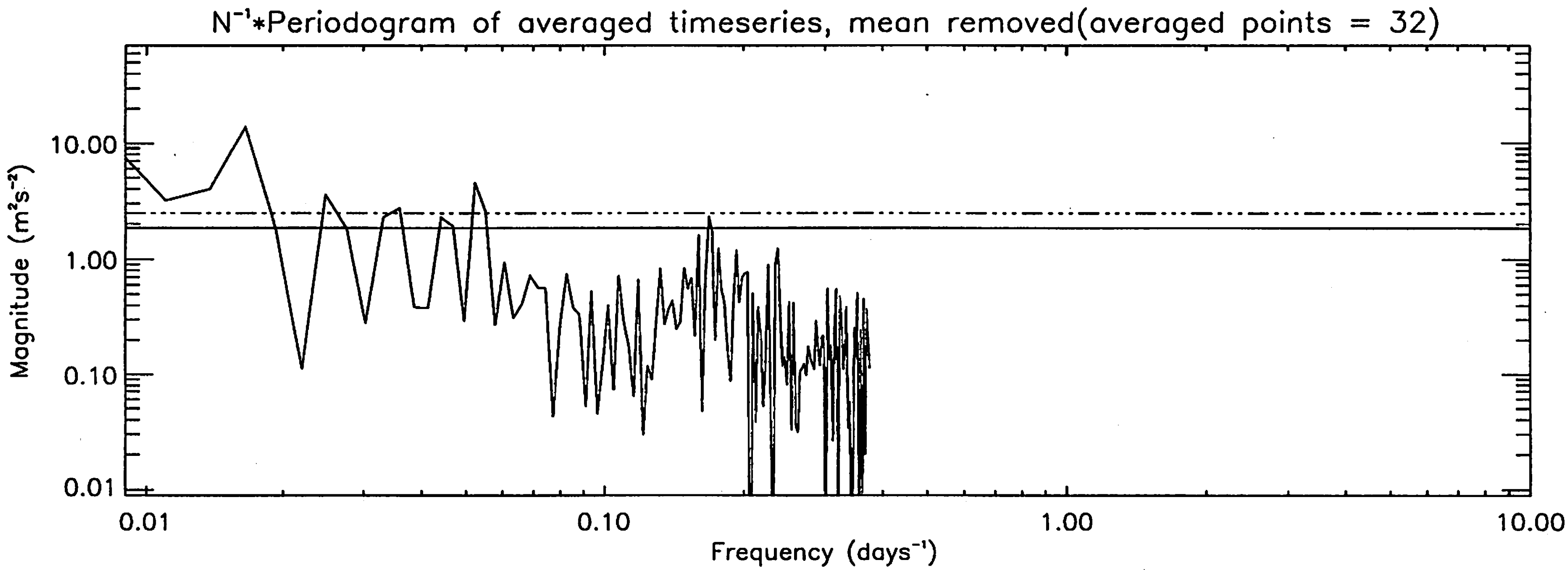
From "Tables of percentage points of the inverted beta (*F*) distribution." *Biometrika*, Vol. 33 (1943) by M. Merrington and C.M. Thompson and from Table 18 of *Biometrika Tables for Statisticians*, Vol. 1, Cambridge University Press, 1954, edited by E.S. Pearson and H.O. Hartley. Reproduced with permission of the authors, editors, and *Biometrika* trustees.



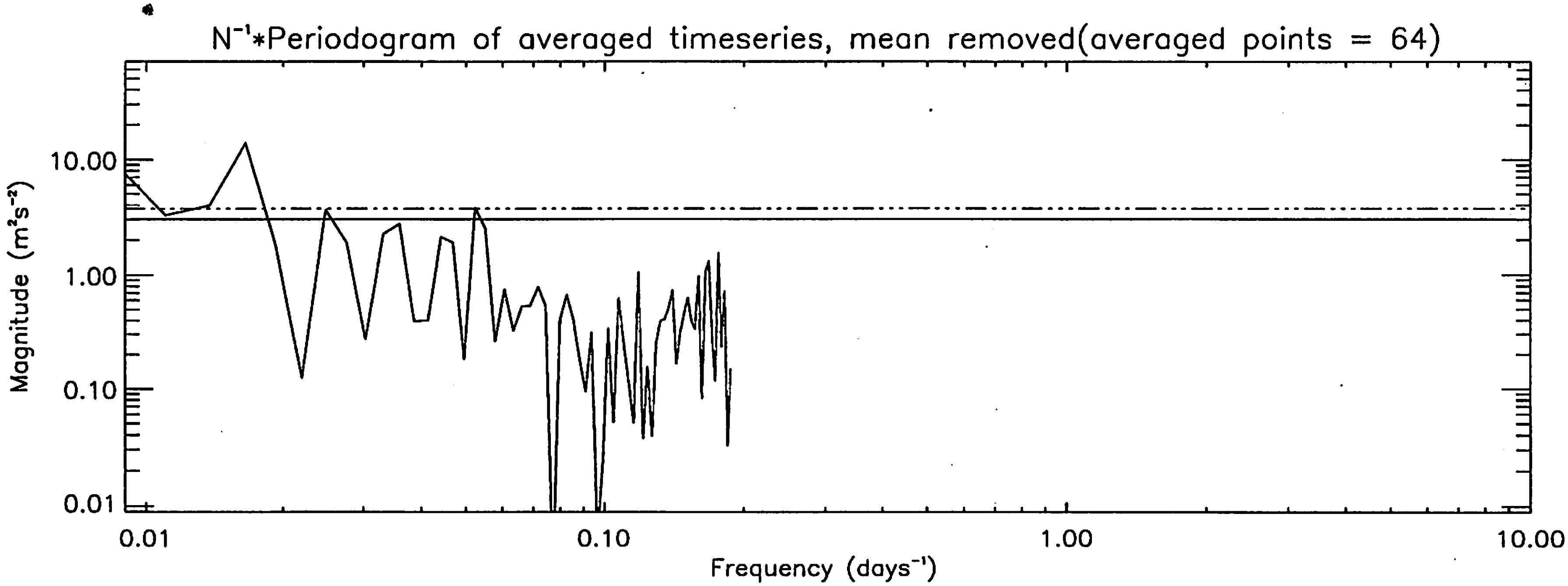
N^{-1}*Periodogram of averaged timeseries, mean removed(averaged points = 4)
Magnitude (m^{2}s^{-2})
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days^{-1})



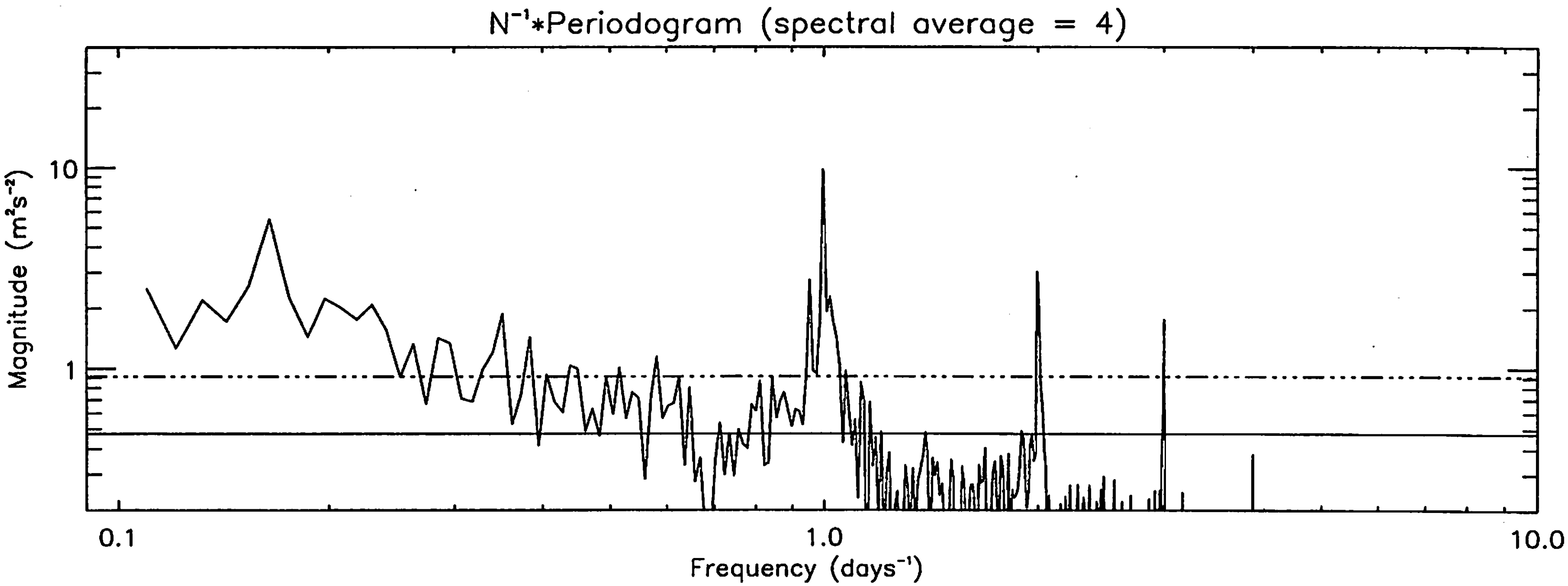
N^{-1}*Periodogram of averaged timeseries, mean removed(averaged points = 8)
Magnitude (m^{2}s^{-2})
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days^{-1})



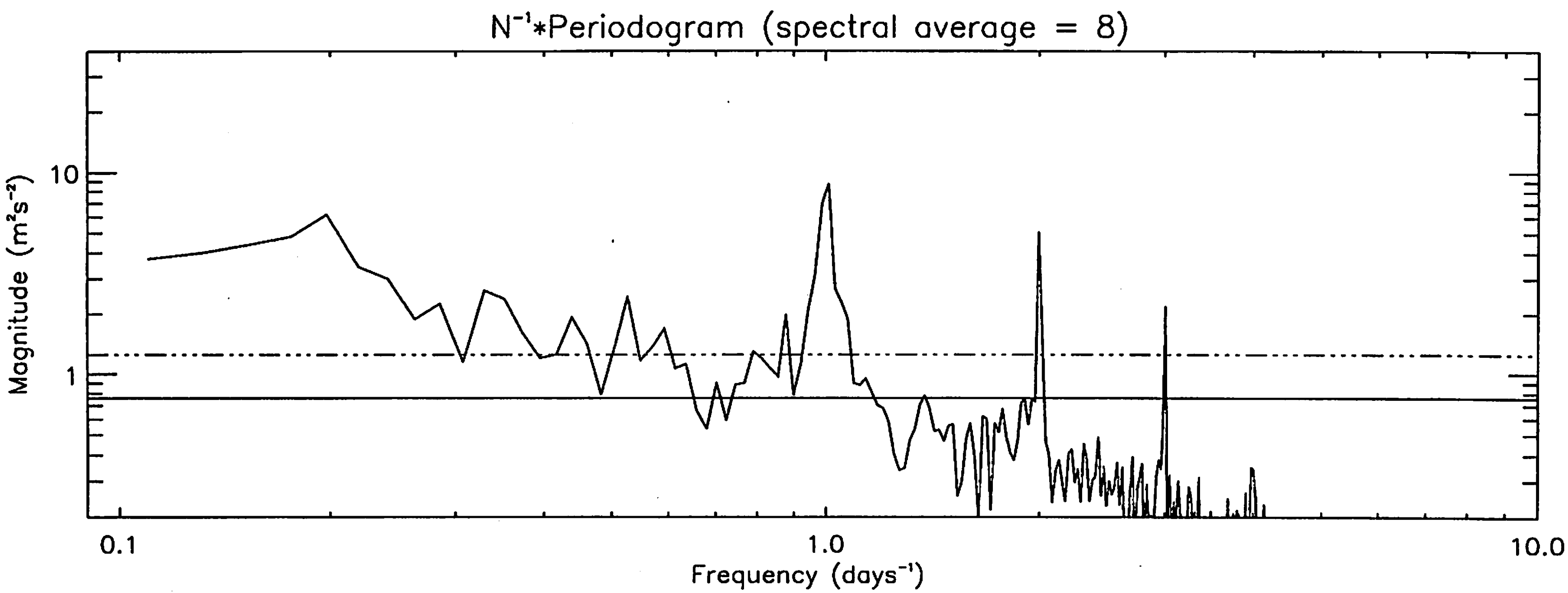
N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 32)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)



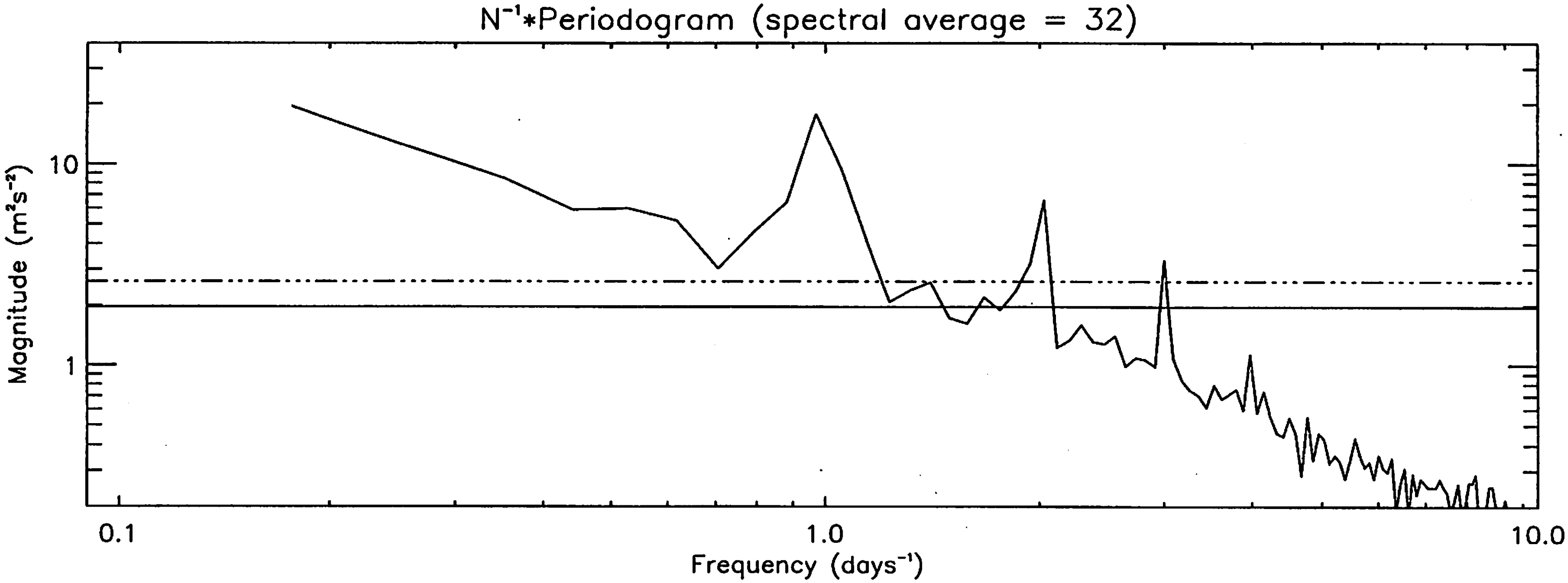
N⁻¹*Periodogram of averaged timeseries, mean removed(averaged points = 64)
Magnitude (m²s⁻²)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
10.00
Frequency (days⁻¹)

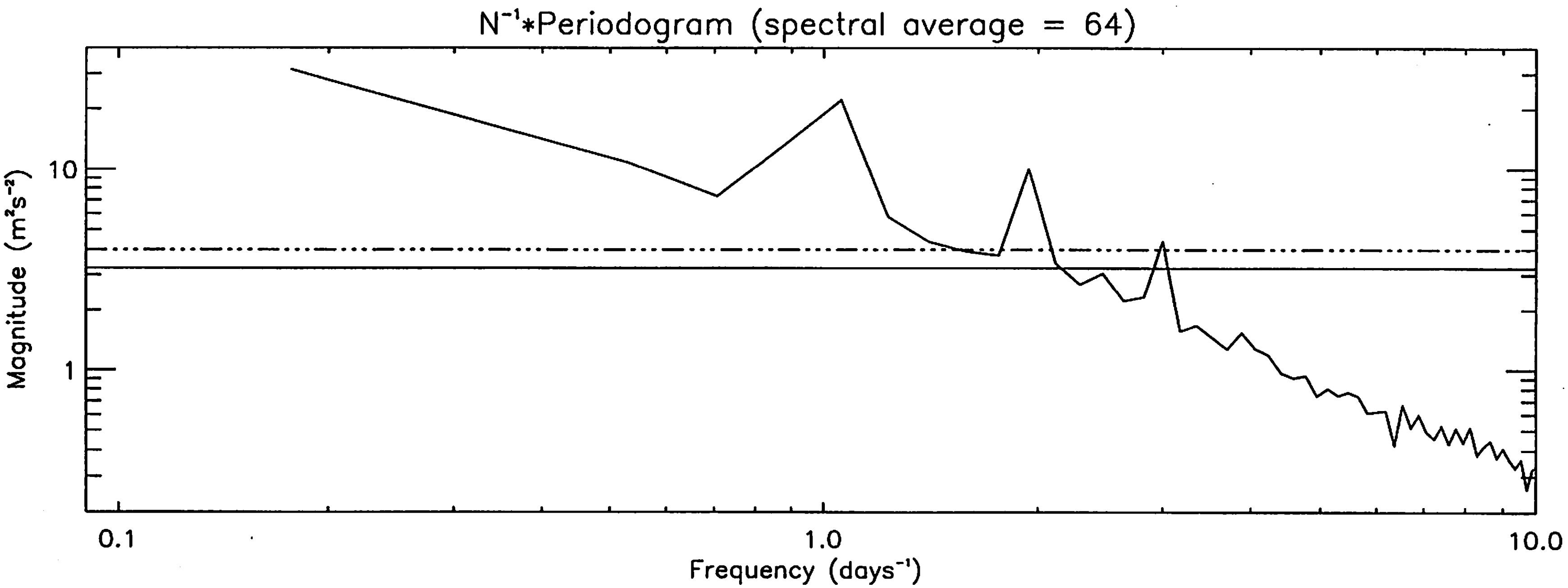


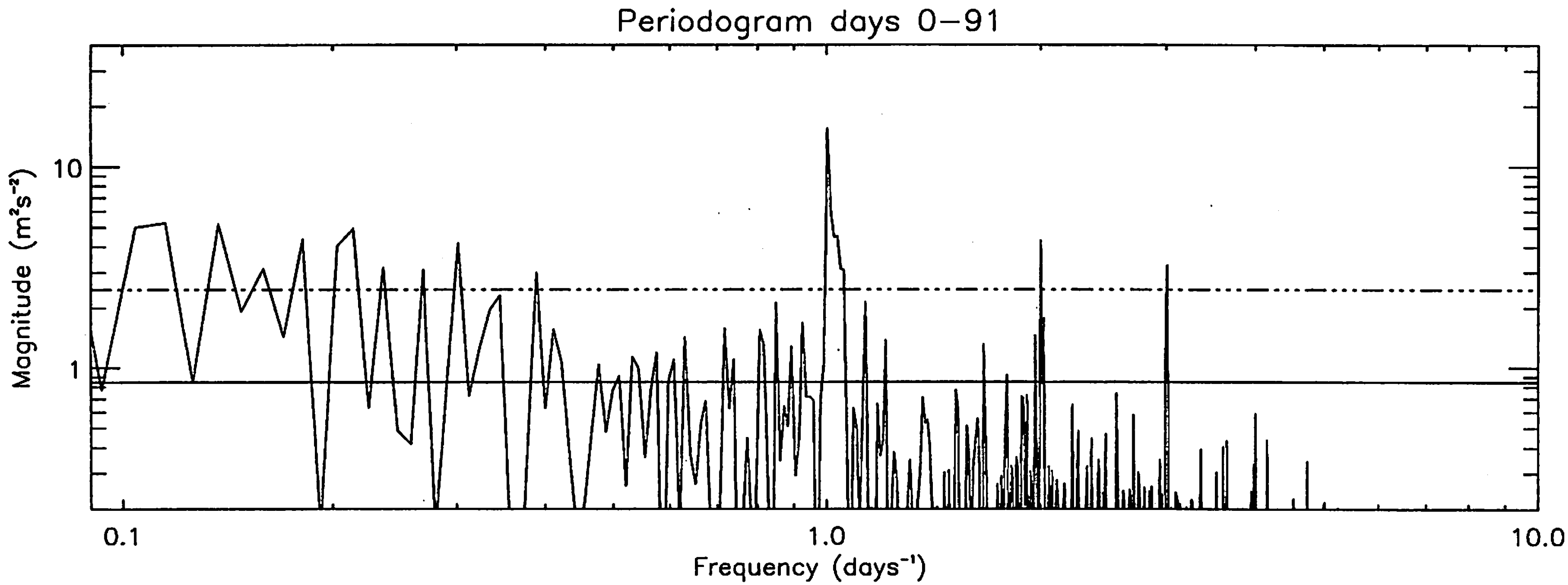
N⁻¹*Periodogram (spectral average = 4)
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)



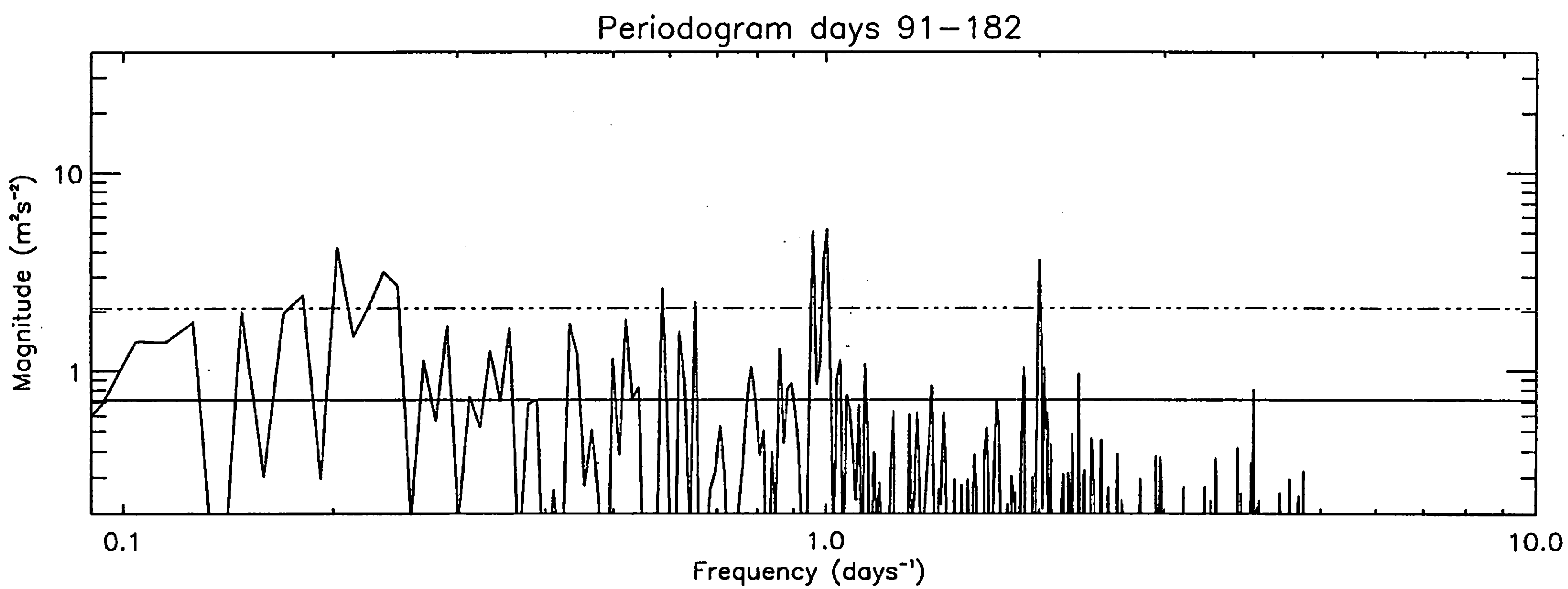
N⁻¹*Periodogram (spectral average = 8)
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)


N^{-1}*Periodogram (spectral average = 32)
Magnitude (m^2s^{-2})
10
1
0.1
1.0
10.0
Frequency ($days^{-1}$)

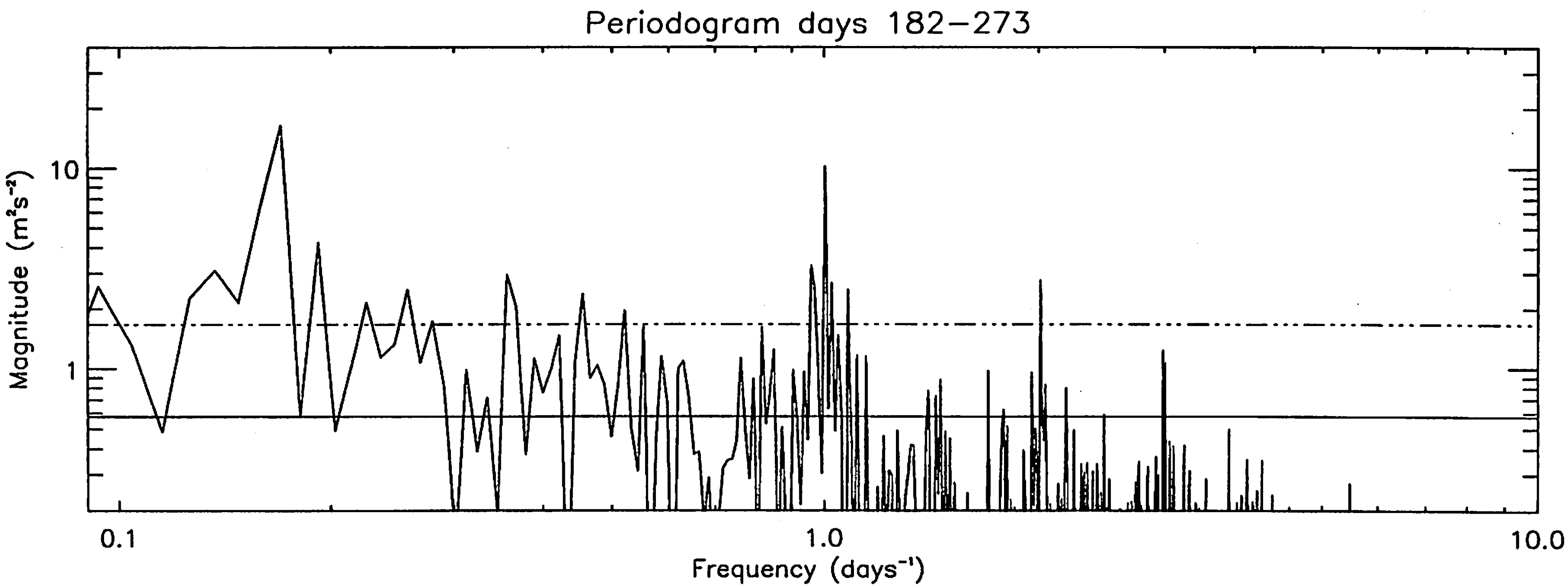

N^{-1}*Periodogram (spectral average = 64)
Magnitude (m^2s^{-2})
10
1
0.1
1.0
10.0
Frequency ($days^{-1}$)

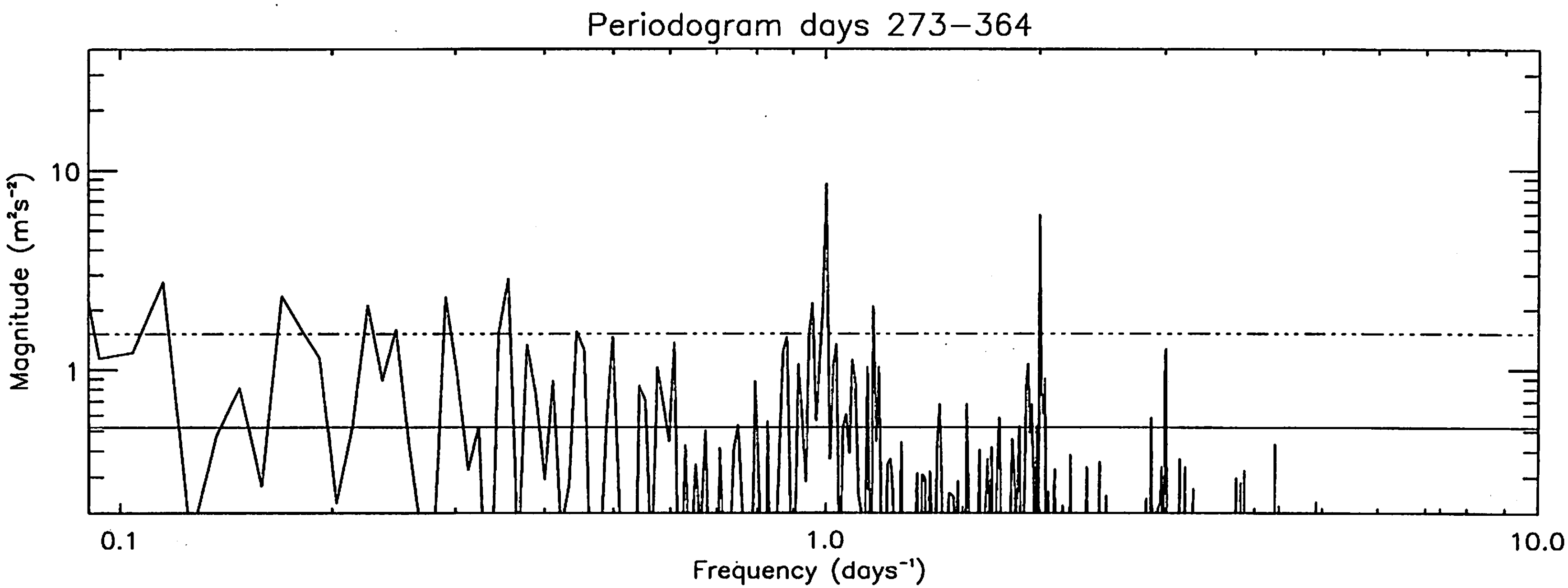


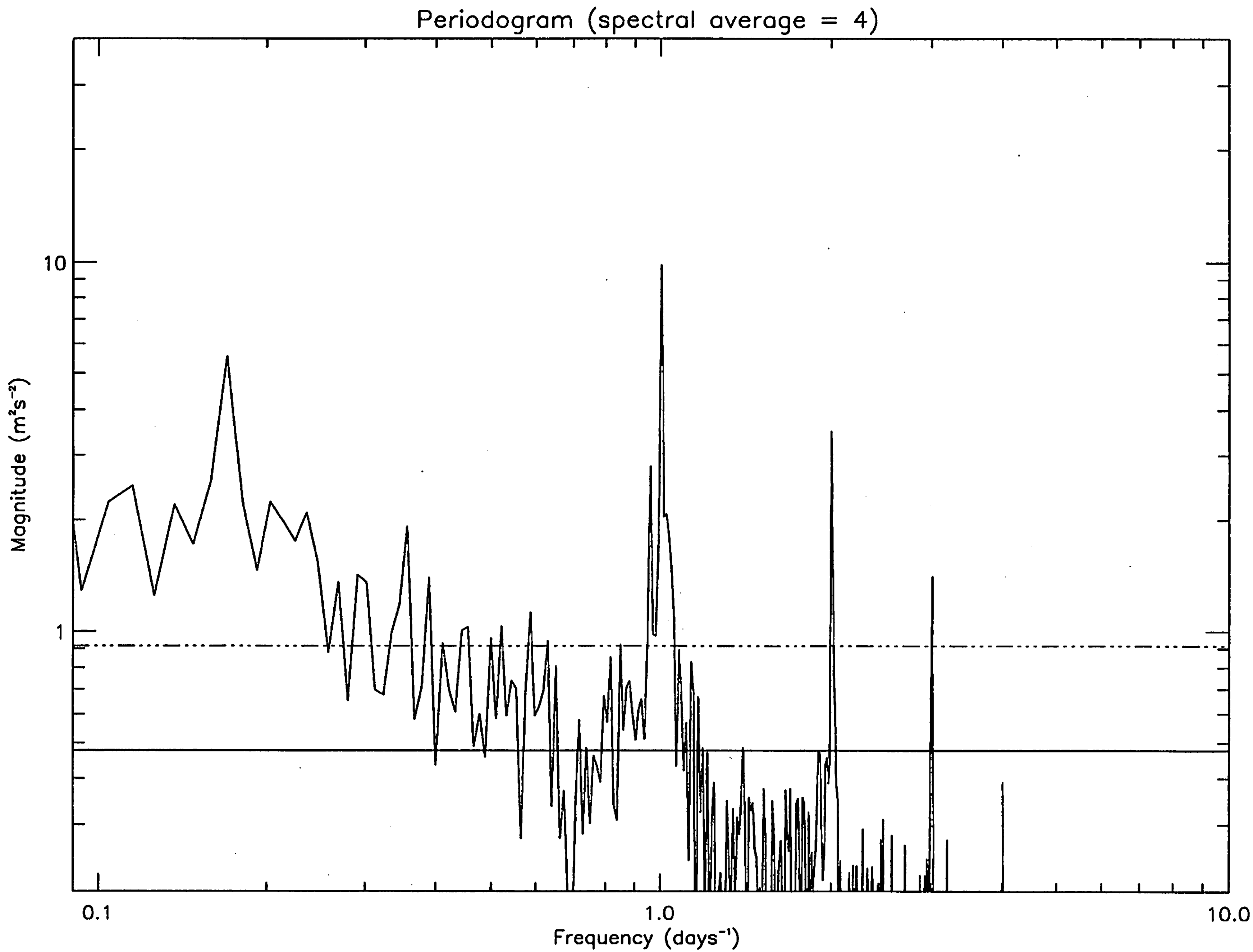
Periodogram days 0–91
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)



Periodogram days 91–182
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)

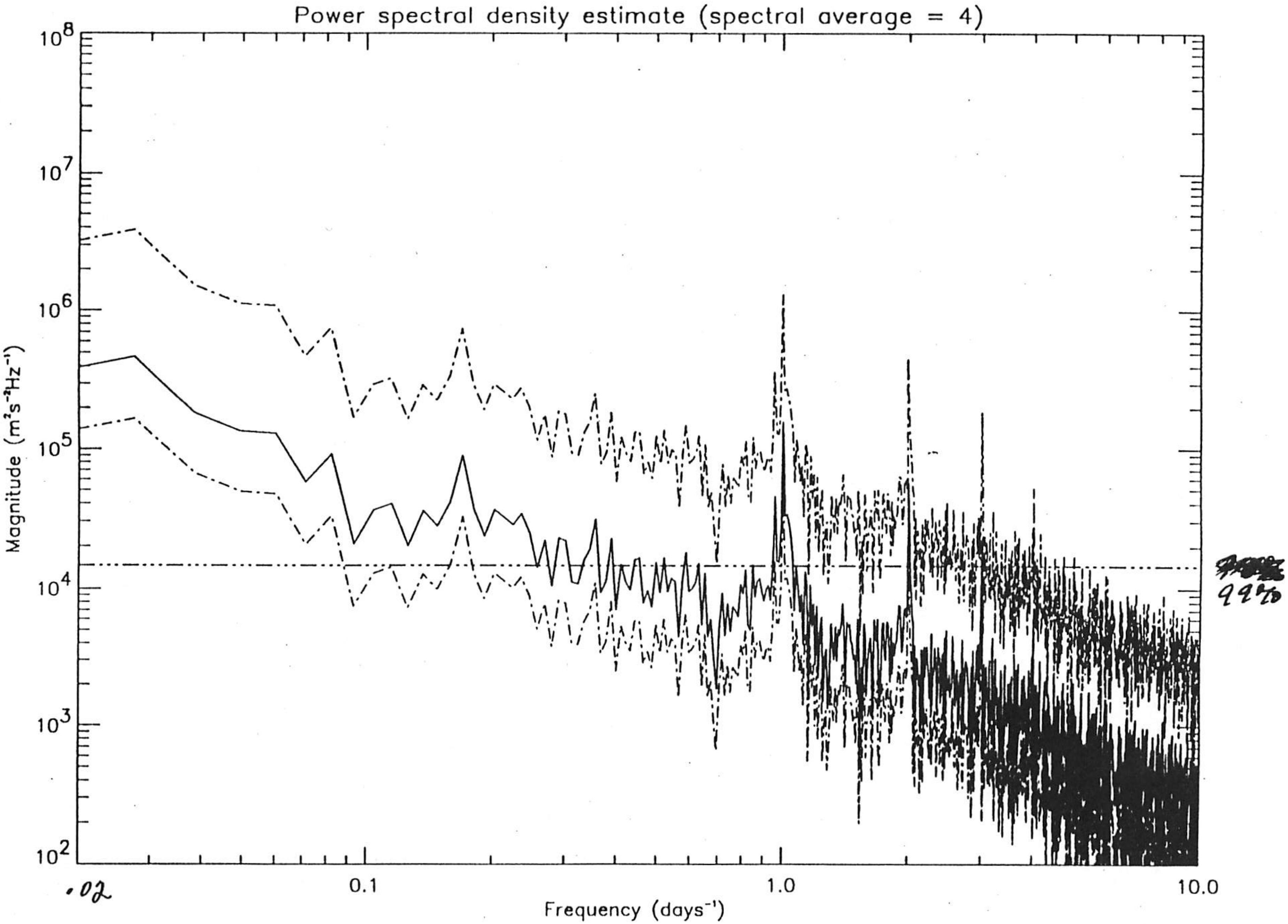

Periodogram days 182–273
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)


Periodogram days 273–364
Magnitude (m²s⁻²)
10
1
0.1
1.0
10.0
Frequency (days⁻¹)



Periodogram (spectral average = 4)
10
1
Magnitude (m^2s^{-2})
0.1
1.0
10.0
Frequency ($days^{-1}$)

95% Confidence Interval



Summary

1. Objective is to estimate the spectral content of a random process based on a finite set of observations from that process.

2. Spectral estimation requires a measure of the confidence of the estimate.

3. The periodogram assumes a model for the data which is a sum of sinusoids in the presence of Gaussian noise.

4. Spectral estimation via the periodogram assumes the data is stationary over the time window.

5. Spectral estimation is a tool that gives insight on further analysis of the data.

References

1. Bloomfield, Peter: Fourier Analysis of Time Series: An Introduction, John Wiley & Sons, 1976.

2. Kay, Steven: Modern Spectral Estimation: Theory and Application, Prentice Hall, 1988.

3. Jenkins, Gwilym and Donald Watts: Spectral Analysis and its Applications, 2nd ed., Holden-Day, 1969.

4. Mendenhall, Scheaffer, Wackerly: Mathematical Statistics with Applications, 3rd ed., PWS Publishers, 1986.

5. Mendenhall, William and Terry Sincick: Statistics for Engineering and the Sciences, 3rd ed., Macmillan Publishing Co., 1992.

6. Meyer, Stuart: Data Analysis for Scientists and Engineers, John Wiley & Sons, 1975.

7. Ross, Sheldon: A First Course in Probability, 2nd ed., Macmillan Publishing Co., 1984.