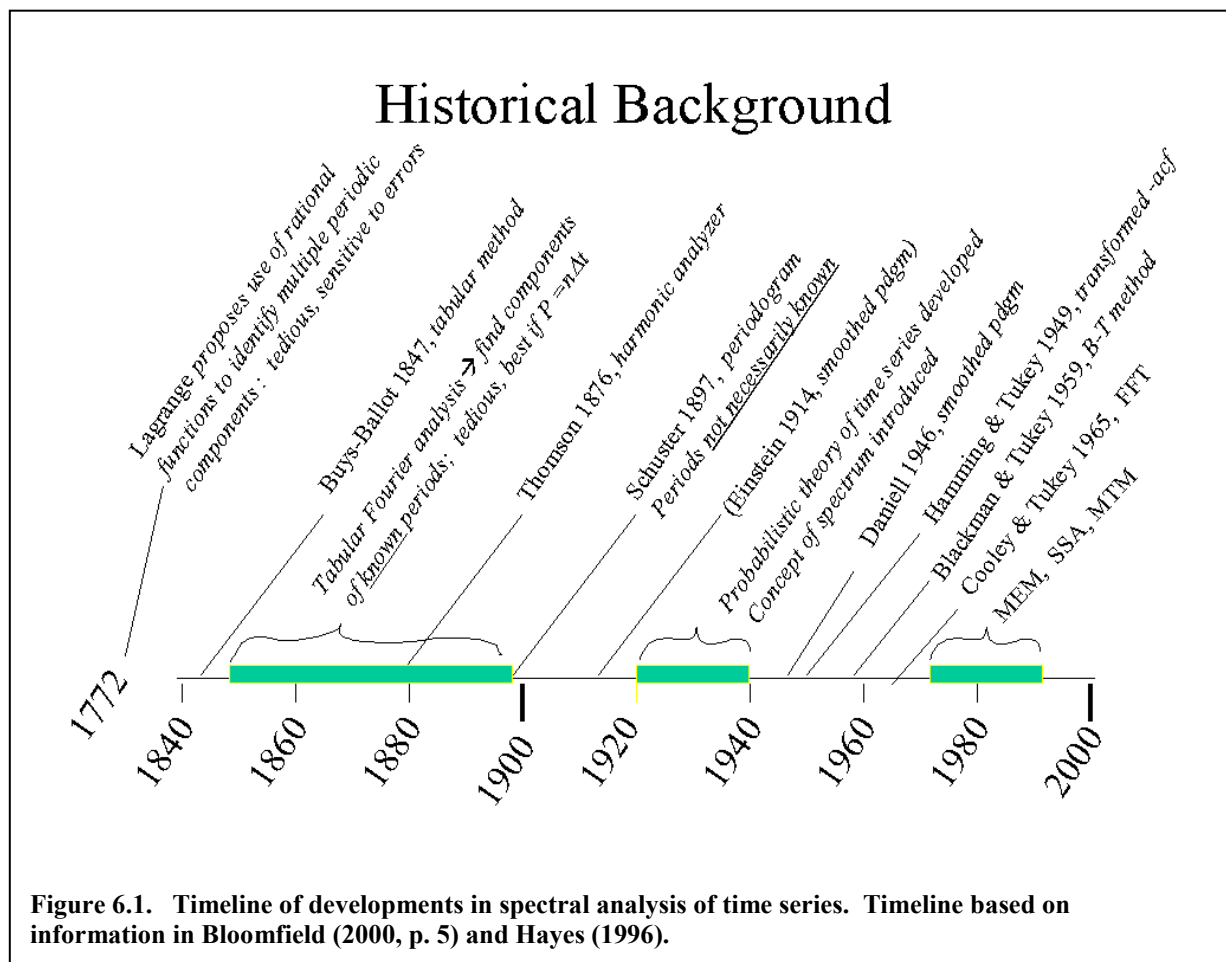


6 Spectral Analysis -- Smoothed Periodogram Method

6.1 Historical background

The spectrum of a time series can be estimated by a variety of methods. In lesson 4 we looked at the Blackman-Tukey method, which uses a Fourier transform of the smoothed, truncated autocovariance function. In contrast, the smoothed periodogram method uses a Fourier transform of the time series itself. The transform yields a quantity called the raw periodogram, which in essence is a highly resolved spectral estimate. The raw periodogram was introduced in the late 1800s for study of periodicity in time series. Unfortunately, the raw periodogram is a crude spectral estimate with high variance (great uncertainty). Smoothing the raw periodogram produces a better estimate of the spectrum – one with a smaller variance (tighter confidence interval). But smoothing reduces the ability of the spectrum to resolve periodic features that may be very close in wavelength. The smoothness, resolution and variance of spectral estimates is controlled by the choice of filters in smoothed periodogram analysis. Extremely broad smoothing produces an underlying smoothly varying spectrum, or null continuum, against which spectral peaks in less-severely smoothed periodogram can be tested for significance. This approach is an alternative to the specification of a functional form of the null continuum (e.g., red noise).

The periodogram was one of the earliest statistical tools for studying periodic tendencies in time series (Figure 6.1). Prior to development of the periodogram such analysis was tedious and generally feasible only when the periods of interest covered a whole number of observations.



Schuster (1897) showed that the periodogram could yield information on periodic components of a time series and could be applied even when the periods are not known beforehand. Following the development of statistical theory of the spectrum in the 1920s and 1930s, the smoothed periodogram was proposed as an estimator of the spectrum (Daniell 1946). Bloomfield (2000) notes that use of the smoothed periodogram in this sense had actually been anticipated much earlier – in 1941 -- by Albert Einstein.

The smoothed periodogram enjoyed a brief period of popularity as a spectral estimator. Another method, Fourier transformation of the truncated and smoothed autocorrelation function (e.g., the Blackman-Tukey method), gained prevalence over the smoothed periodogram in the 1950s because of computational advantages. The smoothed periodogram has become popular again recently. According to Chatfield (2004, p. 136), two factors have led to increasing use of the smoothed periodogram. First is the advent of high-speed computers. Second is the discovery of the fast Fourier Transform (FFT), which greatly speeded up computations (Cooley and Tukey 1965).

Today the smoothed periodogram is one of many alternative methods available for estimating the spectrum. Some of these methods are listed in Table 1. Pros and cons of the various methods – except MTM -- are discussed in Hayes (1996).

Table 1. Alternative methods for spectrum

Method	Summary	Reference
Blackman-Tukey	Fourier transformation of smoothed, truncated autocovariance function	Chatfield, 1975
Smoothed periodogram	Estimate periodogram by DFT of time series; Smooth periodogram with modified Daniell filter	Bloomfield, 2000
Welch's method	Averaged periodograms of overlapped, windowed segments of a time series	Welch, 1967
Multi-taper method (MTM)	Use orthogonal windows ("tapers") to get approximately independent estimates of spectrum; combine estimates	Percival and Walden, 1993
Singular spectrum analysis (SSA)	Eigenvector analysis of autocorrelation matrix to eliminate noise prior to transformation to spectral estimates	Vautard and Ghil, 1989
Maximum entropy (MEM)	Parametric method: estimate acf and solve for AR model parameters; AR model has theoretical spectrum	Kay, 1988

6.2 Steps in smoothed-periodogram method

The main steps in estimating the spectrum by the smoothed periodogram method are:

1. Subtract mean and detrend time series
2. Compute discrete Fourier transform (DFT)
3. Compute (raw) periodogram
4. Smooth the periodogram to get the estimated spectrum

The first step in estimation of the spectrum by the smoothed periodogram method is subtraction of the sample mean. This operation has no effect on the variance, and amounts to a shift of the series along the vertical axis (Figures 6.2, 6.3). The most obvious problem with not subtracting the mean is that an abrupt offset is introduced when the series is padded with zeros in a later step in the analysis.

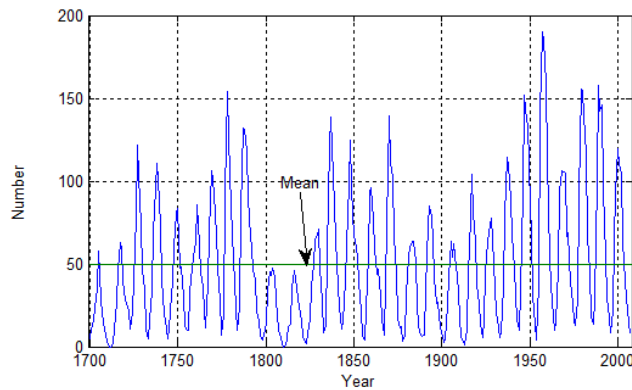


Figure 6.2. Time plot of Wolf Sunspot Number, 1700-2007. This time series is known to have an irregular cycle with period near 11 years. The long-term mean is 49.9. Data source: <http://sidc.oma.be/sunspot-data/>

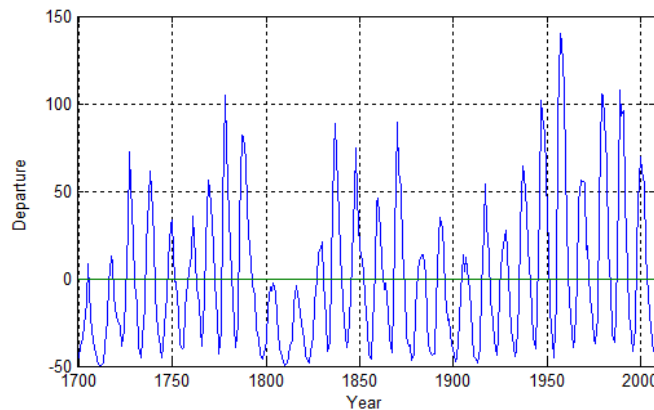


Figure 6.3. Sunspot series as departure from long-term mean.

Any obvious trend should also be removed prior to spectral estimation. Trend produces a spectral peak at zero frequency, and this peak can dominate the spectrum such that other important features are obscured. After detrending, the next steps are computation of the Fourier transform, computation of the raw periodogram, and smoothing of the periodogram.

Discrete Fourier transform. Say x_0, x_1, \dots, x_{n-1} is an arbitrary time series of length n . The time series can be expressed as the sum of sinusoids at the Fourier frequencies of the series:

$$x_t = A(0) + \left\{ 2 \sum_{0 < j < n/2} [A(f_j) \cos 2\pi f_j t + B(f_j) \sin 2\pi f_j t] \right\} + [\{A(f_{n/2}) \cos 2\pi f_{n/2} t\}, \quad t = 0, 1, \dots, n-1], \quad (1)$$

where the summation is over Fourier frequencies

$$f_j = \frac{j}{n}, \quad j = 1, 2, \dots, (n-1)/2,$$

and the last term in braces is included only if n is even (Bloomfield 2000, p. 38). Note that the total number of coefficients is n whether n is even or odd. The coefficients in (1) are given by

$$A(f) = \frac{2}{n} \sum_{t=0}^{n-1} x_t \cos 2\pi f t$$

$$B(f) = \frac{2}{n} \sum_{t=0}^{n-1} x_t \sin 2\pi f t. \quad (2)$$

Equations (2) are sine and cosine transforms that transform the time series x_t into two series of coefficients of sinusoids. The relationships in (2) can be more succinctly expressed in complex notation by making use of the Euler relation

$$e^{ix} = \cos x + i \sin x \quad (3)$$

and its inverse

$$\cos x = \frac{1}{2} \{e^{ix} + e^{-ix}\}, \quad \sin x = \frac{1}{2i} \{e^{ix} - e^{-ix}\} \quad (4)$$

In general, observed data are strictly real-valued, but they may be regarded as complex numbers with zero imaginary parts. Suppose x_0, x_1, \dots, x_{n-1} is such a real-valued time series expressed as complex numbers. The *discrete Fourier transform (DFT)* of x_t is given in complex notation by

$$d(f) = \frac{1}{n} \sum_{t=0}^{n-1} x_t e^{-2\pi i f t} \quad (5)$$

Periodogram. The relationships (2) transform the time series into a series of coefficients at its Fourier frequencies. The discrete Fourier transform is the complex expression of these coefficients

$$d(f) = \frac{A(f)}{2} - i \frac{B(f)}{2}, \quad (6)$$

where A and B are identical to the quantities defined in (2).

The original data can be recovered from the DFT using the *inverse transform*

$$x_t = \sum_j d(f_j) e^{2\pi i f_j t} \quad (7)$$

which is the complex equivalent of equation (1).

The discrete Fourier transform has two representations. The first is in terms of its real and imaginary parts, $A(f)/2$ and $-B(f)/2$. The second is in terms of its *magnitude* $R(f)$ and *phase* $\phi(f)$

$$d(f) = R(f)e^{i\phi(f)} \quad (8)$$

The magnitude, given by

$$R(f) = |d(f)| \quad (9)$$

measures how strongly the oscillation at frequency f is represented in the data. The strength of the periodic component is more often represented by the *periodogram* defined as

$$I(f) = n[R(f)]^2 = n|d(f)|^2 \quad (10)$$

The sine and cosine terms at the Fourier frequencies are orthogonal, and so the variance of the time series x_t can be decomposed into components at the individual frequencies. For the sine and cosine transforms, the sum of squares of the original data can be expressed as

$$\sum_{t=0}^{n-1} x_t^2 = nA(0)^2 + \left\{ 2n \sum_{0 < j < n/2} [A(f_j)^2 + B(f_j)^2] \right\} + nA(f_{n/2})^2 \quad (11)$$

where the last term is included only if n is even. The analog for the discrete Fourier transform in complex notation is

$$\sum_{t=0}^{n-1} |x_t|^2 = n \sum_j |d(f_j)|^2 = \sum_j I(f_j) \quad (12)$$

If x_t is a time series expressed as departures from its mean, the sums of squares in equations (11) and (12) are simply n times the variance. Equation (12) therefore says that (a) the sum of the periodogram ordinates equals the sum of squares of departures of the time series from its mean, (b) the sum of periodogram ordinates divided by the series length equals the series variance, and (c) the periodogram ordinate at Fourier frequency f_j is proportional to the variance accounted for by that frequency component. The periodogram at this stage is a “raw” periodogram, meaning it has not yet been smoothed.

The raw periodogram of the Wolf sunspot number is plotted in Figure 6.4. Each point represents the relative variance of the time series contributed by a frequency range centered at the point. Wavelengths near 11 years make relatively large contribution to the variance.

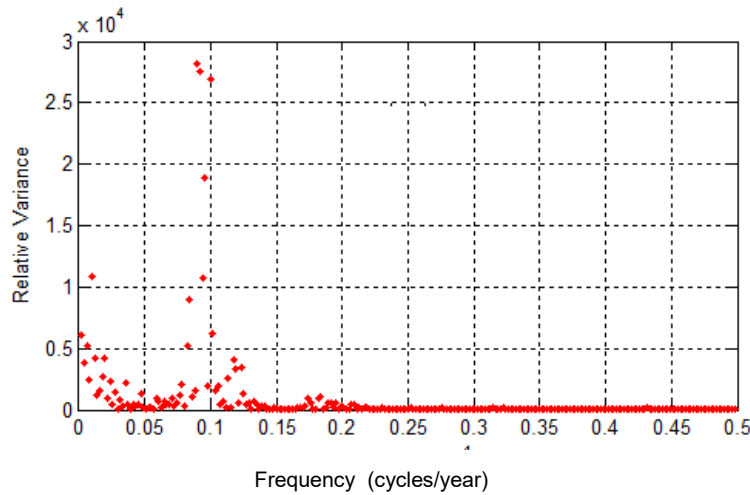


Figure 6.4. Raw periodogram of Wolf sunspot number, 1700-2007. Periodogram ordinates give the relative variance contributed at different frequency ranges centered on fundamental frequencies (after padding) of the series. The number of points in the plot is 256 because the series has been padded to length 512 before periodogram analysis (see Section 6.6).

Smoothing the periodogram. The periodogram is a wildly fluctuating estimate of the spectrum with high variance. For a stable estimate, the periodogram must be smoothed. Bloomfield (2000, p. 157) recommends the Daniell window as a smoothing filter for generating an estimated spectrum from the periodogram. The modified Daniell window of span, or length, m , is defined as

$$g_i = \frac{1}{2(m-1)}, \quad i = 1 \text{ or } i = m$$

$$= \frac{1}{m-1}, \quad i \text{ otherwise}$$
(13)

where m is the number of weights, or *span* of the filter, g_i is the i^{th} weight of the filter, and i is an index such that $i = 1, \dots, m$. The Daniell filter differs from an evenly weighted moving average (rectangular filter) only in that the first and last weights are half as large as the other weights. A plot of the filter weights therefore has the form of a trapezoid. Filter weights of 5-weight Daniell and rectangular filters are plotted in Figure 6.5. The advantage of the Daniell filter over the rectangular filter for smoothing the periodogram is that the Daniell filter has less *leakage*, which refers to the influence of variance at non-Fourier frequencies on the spectrum at the Fourier frequencies. The leakage is related to sidelobes in the frequency response of the filter. Successive smoothing by Daniell filters with different spans gives an increasingly smooth spectrum, and is equivalent to single application of a resultant filter produced by convoluting the individual spans of the Daniell filters (Bloomfield 2000, p. 157).

A smoothed periodogram of the Wolf sunspot number is plotted in Figure 6.6. The smoothing for this example was done with successive application of Daniell filters of length 7 and 11. Broader (longer) filters would give a smoother spectrum. Narrower filters would give a rougher spectrum. The proper amount of smoothing is somewhat subjective, and depends on the characteristics of the data. If the natural periodicity of a series is such that peaks in the spectrum are closely spaced in frequency, too broad a filter will merge the peaks. The tradeoffs in

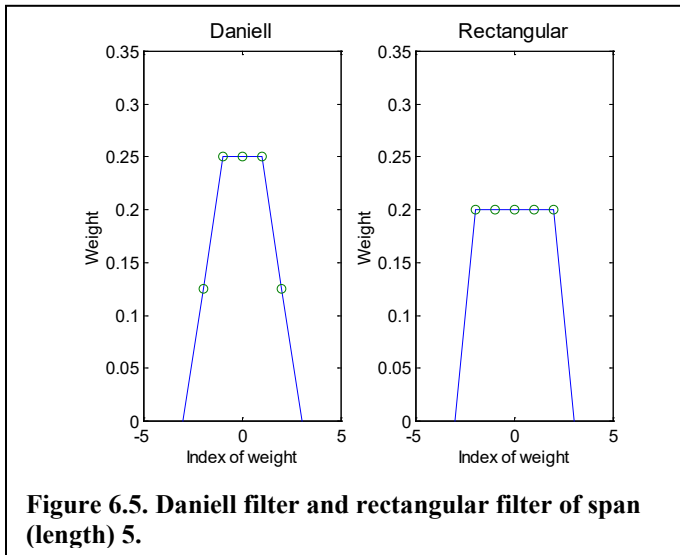


Figure 6.5. Daniell filter and rectangular filter of span (length) 5.

smoothness, stability and resolution should be considered in selecting widths of Daniell filters (see Section 6.4).

6.3 Null continuum

Although the spectrum of a time series is innately useful for describing the distribution of variance as a function of frequency, interest sometimes centers on how the sample spectrum for a given time series differs from that of some known generating process. Interest also sometimes centers on the

statistical significance of peaks in the spectrum. Significance can be evaluated only by reference to some standard of comparison. The question is “significantly different than what?” A standard for comparison is the null continuum. The null continuum is a general null hypothesis for the spectrum. The null continuum can be theoretically based, or data-based.

The simplest form of null continuum is white noise, which has an even distribution of variance over frequency. The white noise spectrum is accordingly a horizontal line. Variance is not preferentially concentrated in any particular frequency range. In testing for significance of spectral peaks, the white noise null continuum may be inappropriate if it is known that the series is persistent. Persistence, or positive autocorrelation, in a time series skews its variance toward the low-frequency side of the spectrum. One option for dealing with persistent processes is to use the theoretical spectrum of an autoregressive process as the null continuum.

Theoretical AR spectra. An autoregressive process has a characteristic spectrum whose shape depends on the model order and the values of the model parameters. An AR(1) process can have a spectrum that ranges from red-noise (emphasizing low frequencies) to blue noise (emphasizing high frequencies). Higher-order AR process can have very complicated spectral shapes. The AR(2) model, for example, can represent a pseudo-periodic process. A variety of theoretical spectra for various AR(1) and AR(2) processes are illustrated by Wilks (1995, p. 353). The equations for AR spectra in this section are taken from Wilks (1995), with one difference: signs on the AR coefficients have been changed for consistency with Matlab’s notation for AR models. Matlab’s System Identification Toolbox writes the AR(1) model as

$$y_t + a_1 y_{t-1} = e_t \quad (14)$$

and the AR(2) model by

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} = e_t \quad (15)$$

where y_t is the time series expressed as departures from its mean, a_1 and a_2 are autoregressive parameters, and e_t is the noise term. A positively autocorrelated AR(1) series has $a_1 < 0$ by this convention.

AR(1) model. For the simple case of the AR(1) process, negative values of the autoregressive parameter build a memory into the system that tends to smooth over short-term (high-frequency) variations and emphasize the slower variations. The effects are progressively stronger as the parameter gets closer to -1.

The theoretical spectrum of an AR(1) process can be written in terms of the AR parameter, the variance of the residuals and the sample size. For an AR(1) model,

$$S(f) = \frac{4\sigma_\varepsilon^2/N}{1 + a_1^2 + 2a_1 \cos(2\pi f)}, \quad 0 \leq f \leq 1/2 \quad (16)$$

where σ_ε^2 is the variance of the residuals, N is the number of observations, a_1 is the autoregressive parameter, f is frequency in cycles per year (or time unit), and $S(f)$ is the theoretical spectrum. The shape is entirely determined by the AR parameter; σ_ε^2 and N act to scale the spectrum higher or lower, but do not change the relative distribution of variance over frequency.

Equation (16) can be used to generate a theoretical spectrum for any observed time series. The series is first modeled as an AR(1) process. The autoregressive parameter and variance of residuals are estimated from the data. Plots of spectra for different values of a_1 show how the spectrum varies as a function of autoregressive parameter. The special case of $a_1 = 0$ corresponds to a white noise process. By analogy with visible light, white noise contains an equal mixture of variance at all frequencies. The theoretical spectrum of white noise is a horizontal line.

For $a_1 < 0$, the spectrum is enhanced at the low frequencies and depleted at the higher frequencies. By analogy with the light, the spectrum is called “red noise.”

For $a_1 > 0$, the process tends to create erratic short-term variations, with positive autocorrelation at even lags and negative autocorrelation at odd lags. The spectrum is enriched at the high frequencies, and depleted at the low frequencies. Such series are sometimes called “blue noise.”

AR (2) model. Theoretical spectra can also be expressed for other AR models. For the special case of the AR(2) model, the spectrum is :

$$S(f) = \frac{4\sigma_\varepsilon^2/N}{1 + a_1^2 + a_2^2 + 2a_1(1 + a_2)\cos(2\pi f) + 2a_2\cos(4\pi f)}, \quad 0 \leq f \leq 1/2 \quad (17)$$

The spectrum of an AR(2) process is particularly interesting because it can express pseudo-periodicity, with the spectral details depending on the size of the two parameters. For example, for $a_1 = -0.9$, $a_2 = +0.6$, the spectrum has a strong peak near a frequency of 0.15. For annual time series, this frequency corresponds to a wavelength of 6.7 yr. It is therefore possible that a time series that exhibits quasi-periodic behavior with a variance peak at 6.7 years results from a short memory process with dependence on the past restricted to the most recent two years.

Theoretical AR spectra can be useful in data interpretation. Similarity of the gross shape of the estimated spectrum with that of an AR process may suggest the process as a simple generating mechanism.

Allowable ranges of AR parameters. AR parameters must stay within a certain range for a process to be stationary. For the AR(1) model it is required that

$$-1 < a_1 < 1 \quad (18)$$

For an AR(2) process, several conditions must be satisfied:

$$\begin{aligned} -1 < a_2 < 1 \\ a_1 + a_2 &> -1 \\ a_1 - a_2 &< 1 \end{aligned} \quad (19)$$

The allowable ranges of AR(2) parameters and the portions of the bivariate range corresponding to pseudo-periodic behavior in a time series are discussed by Anderson (1976).

White noise, AR(1) and AR(2) theoretical spectra are sketched in figure 6.7. It should be noted that higher-order AR process can have very complicated spectra. The maximum entropy method of spectral estimation fits a high-order AR process to the series and uses the theoretical AR spectrum as the spectral estimate.

Another approach to a null continuum is empirically based and does not attempt to assign any particular theoretical generating model as a null hypothesis (Bloomfield 2000). This approach uses a greatly smoothed version of the raw periodogram as the null hypothesis. Figure 6.8 illustrates a smoothed-periodogram null continuum for the sunspot series using successive applications of Daniell filters of spans 43, 61 and 77. The estimation of a null continuum by smoothing the periodogram relies on subjective judgement and trial-and-error. In particular the null continuum should follow just the smooth underlying shape of the distribution of variance over frequency. If smoothed insufficiently, the null continuum will bulge at the important peaks in the spectrum. This would clearly be undesirable as the test of significance of the peak is that it is different from the null continuum.

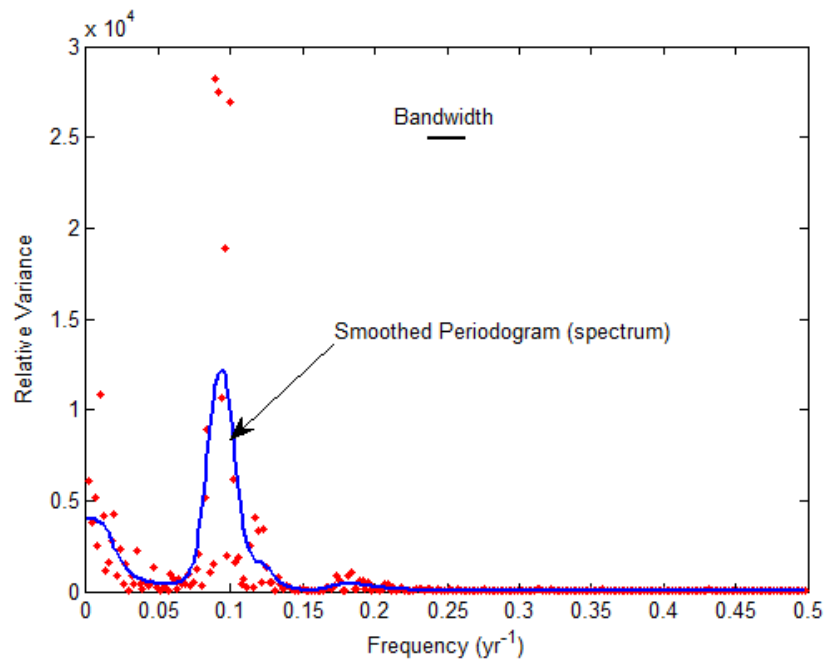


Figure 6.6. Smoothed periodogram estimate of spectrum of Wolf sunspot number. Raw periodogram (points) smoothed by Daniell filters of length 7 and 11. Bandwidth gives resolution of the spectral estimate. Spectral peak at 10.6 years.

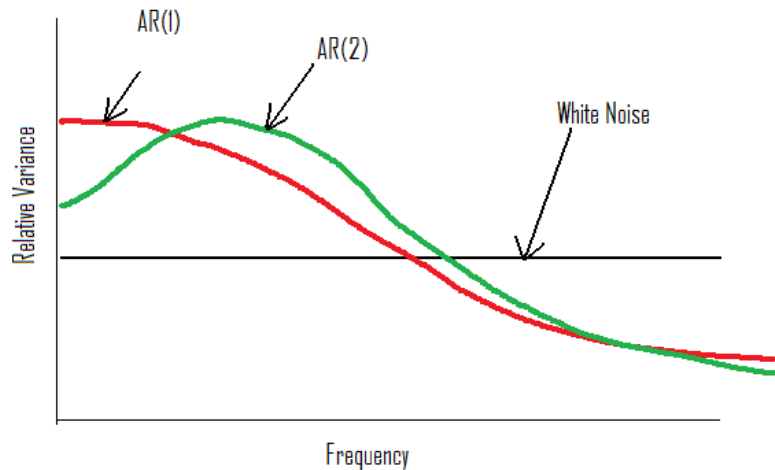


Figure 6.7. Sketch of general shapes of white noise, AR(1) and AR(2) theoretical spectra, or null continua. The AR(1) spectrum here is a red-noise model. The AR(2) spectrum can take various shapes depending on values of the two AR parameters.

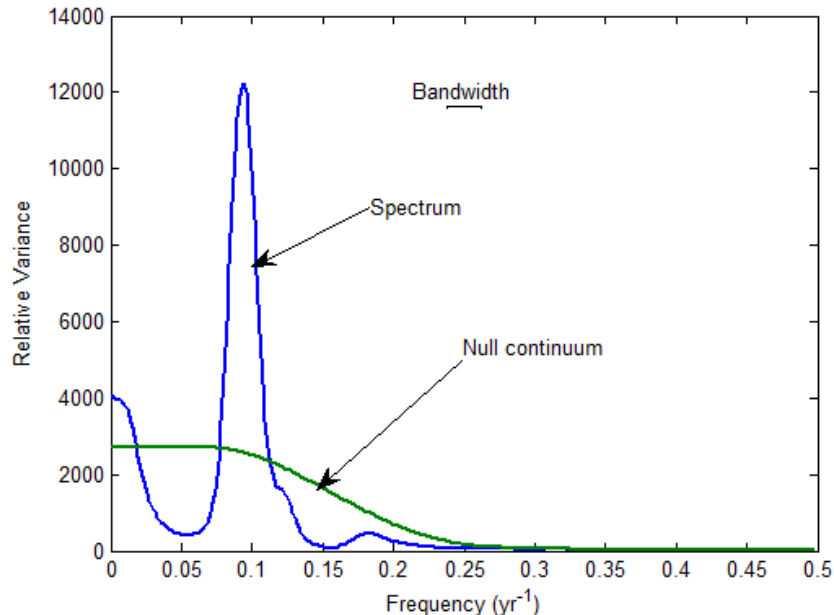


Figure 6.8. Spectrum and empirically-based null continuum of Wolf sunspot series. Daniell filters of spans 7 and 11 were applied to smooth the raw periodogram into the spectrum. Spans of length 43, 61 and 77 were used to smooth the raw periodogram into the null continuum. In evaluating the significance of the spectral peak, the null hypothesis H_0 is that the variance contribution at the frequency of the peak is no different than the corresponding contribution for the null continuum (green line) at the same frequency. H_0 would be tested by placing a confidence band (e.g., 95%) on the blue line and checking whether the lower confidence limit includes the green line.

6.4 Spectral properties

Smoothness. The raw (unsmoothed) periodogram is a rough estimate of the spectrum. The periodogram is proportional to variance contributed at the fundamental frequencies. Unfortunately, the raw periodogram is of little direct usefulness because of the high variance of the spectral estimates. Fluctuations in the raw periodogram can be driven by sampling variability, a problem that becomes more severe the shorter the series. Smoothing the raw periodogram reduces this problem. If the raw periodogram is smoothed with Daniell filters of various spans, the result is a spectrum smoother in appearance, and with a tighter confidence interval. But excessive smoothing obscures important spectral detail, and insufficient smoothing leaves erratic unimportant detail in the spectrum. Smoothness is closely related to *bias*, as discussed in the lecture on the Blackman-Tukey method of spectral estimation. As a spectrum is smoothed more and more, the estimated spectrum eventually approaches a featureless curve, “biased” towards the local mean.

Stability. The stability of the spectral estimate is “the extent to which estimates computed from different segments of a series agree, or, in other words, the extent to which irrelevant fine structure in the periodogram is eliminated” (Bloomfield 2000, p. 156). High stability corresponds to low variance of the estimate, and is attained by averaging over many periodogram ordinates. The number of periodogram ordinates averaged over in the smoothed periodogram method as

described by Bloomfield (2000) is defined by the *span* of the Daniell filter. Let $\hat{s}(f)$ be the spectral estimate computed as the average of periodogram ordinates at the Fourier frequencies. If the time series has not been padded or tapered, the variance of the spectral estimate is given by

$$\text{var}\{\hat{s}(f)\} = s(f)^2 \sum_u g_u^2 \quad (20)$$

where $\hat{s}(f)$ is the spectral estimate at frequency f , $s(f)$ is the true (unknown) value of the spectrum, assumed to be approximately constant over the interval of averaging, and the summation $\sum_u g_u^2$ is the sum of squared weights of the Daniell filter used to smooth the periodogram. The sum of periodogram weights must equal 1 for the spectral estimate to be an unbiased estimate of the true spectrum (Bloomfield 2000, p. 178). The broader the Daniell filter, the lower the sum of squares of weights and the lower the variance of the spectral estimate. For example, for the 3-weight Daniell filter $\{.25, .50, .25\}$ the sum of squares of weights is 0.375, while for the 5-weight filter $\{.125, .25, .25, .25, .125\}$ the sum of squares is 0.2188.

An approximate confidence interval for the spectral estimate can be derived by considering that the periodogram estimates are independent and exponentially distributed. The spectral estimate, as a sum of independent exponentially distributed quantities, is approximately χ^2 distributed. The distribution of $\hat{s}(f)$ can be shown to be approximately χ^2 with degrees of freedom

$$\nu = \frac{2}{g^2} \quad (21)$$

where $g^2 = \sum_u g_u^2$ is the sum of squared Daniell weights. The relationship in (21) can be used to place a confidence interval around the spectral estimates. As shown in Bloomfield (2000), $X = \nu \hat{s}(f) / s(f)$ is approximately chi-squared distributed with ν degrees of freedom (Notation slightly different here than in Bloomfield (2000).) Accordingly, 95% confidence interval for $\hat{s}(f)$ is given by

$$\frac{\nu \hat{s}(f)}{\chi_v^2(0.975)} \leq s(f) \leq \frac{\nu \hat{s}(f)}{\chi_v^2(0.025)} \quad (22)$$

where $\chi_v^2(0.025)$ and $\chi_v^2(0.975)$ are the 2.5% and 97.5% points of the χ^2 distribution with ν degrees of freedom

Resolution. Resolution is the ability of the spectrum to represent the fine structure of the frequency properties of the series. The fine structure is the variation in the spectrum between closely spaced frequencies. For example, narrow peaks are part of the fine structure of the spectrum. The raw periodogram measures the variance contributions at the Fourier frequencies, or the finest possible structure. Smoothing the periodogram, for example with a Daniell filter, averages over adjacent periodogram estimates, and consequently lowers the resolution of the spectrum. The wider the Daniell filter, the greater the smoothing and the greater the decrease in resolution.

If two periodic components in the series are close to the same frequency, the smoothed spectrum might be incapable of identifying, or *resolving*, the individual peaks. The width of the frequency interval applicable to a spectral estimate is called the *bandwidth* of the estimate. If a hypothetical periodogram were to have just a single peak at a particular Fourier frequency, the smoothed spectrum is roughly the image of the Daniell filter used to smooth the periodogram, and

the peak in the spectrum is spread out over several Fourier frequencies. How many Fourier frequencies the peak covers depends on the spans of the filter. A reasonable measure of the bandwidth of the spectral estimate is therefore the width of the resultant Daniell filter used to smooth the periodogram. Depending on how the resultant Daniell filter has been constructed, the shape of the filter also varies. Thus one filter may have only a few weights appreciably different from zero, while another filter of the same length may have fewer or more appreciably non-zero weights. Rather than the width of the Daniell filter, therefore, a more effective measure of bandwidth also takes into account the values of the Daniell filter weights. One such measure of bandwidth is *the width of the rectangular filter that has the same variance as the Daniell filter*.

The variance of the estimator is proportional to the sum of squares of the filter weights. The bandwidth for a given Daniell filter can therefore be computed as follows:

1. Compute the sum of squares of the Daniell filter weights
2. Compute the number of weights n_w of the evenly weighted moving average that has the same sum of squares as computed in (1)
3. Compute the bandwidth as $bw = n_w \Delta f$, where Δf is the spacing of the Fourier frequencies. (Note that if the series has been padded to length N' , the spacing is taken as $1/N'$.)

Differences in smoothness, stability and resolution are illustrated for spectra of the Wolf sunspot series in Figure 6.9. A lesser amount of smoothing of the raw periodogram yields the spectrum in Figure 6.9A. A greater amount of smoothing yields the spectrum in Figure 6.9B. The bandwidths indicate the differences in resolution of the two versions of the spectral. Both versions clearly show the main spectral peak near 11 years, but the peak is narrower and much higher for the spectrum with less smoothing. On the other hand, the confidence interval around the spectrum is much tighter for the spectrum with greater smoothing. Trial and error computing and plotting of spectral with different degrees of smoothing for the smoothed periodogram method is analogous to the “window smoothing” approach described in lesson 4 for the Blackman-Tukey method of spectral estimation.

Note that the sunspot series also exhibits a spectral peak at near frequency 0.01 (wavelength 100 years). This lower-frequency fluctuation is evident also in the time plot of the series (Figure 6.2). Too much smoothing (e.g., Figure 6.9B) makes it impossible to resolve this peak from trend (peak at zero frequency).

6.5 Testing for periodicity

A peak in the estimated spectrum can be tested for significance by comparing the spectral estimate at a given frequency with the confidence interval for the estimate. Two considerations for the testing are:

1. A significance test requires a null hypothesis. For the spectrum, the null hypothesis is that the spectrum at the specified frequency is not different from some “null” spectrum, or null continuum. An earlier section described a white noise null continuum, an autoregressive null continuum and a null continuum based on a greatly smoothed raw periodogram. The null hypothesis is then that the estimated spectrum is no different than this underlying spectrum.
2. The confidence bands developed above (equation 22) are not simultaneous. In other words, the bands should be used strictly to test for significance of a peak at a specified frequency, and that frequency should be specified before running the spectral analysis. This approach can be contrasted with a “fishing expedition”, in which the spectrum is

estimated and then browsed to identify “significant” peaks. Simultaneous confidence bands, which would be much wider than those given by equation 22, are needed if the spectrum is to be in such an exploratory mode to pick out significant peaks. A so-called “Bonferroni correction” (Cochran 1989, p 115 and 167) may be applied to adjust confidence intervals to account for multiple tests or multiple comparisons. This correction essentially amounts to using a $1-\alpha/m$ confidence interval in place of a $1-\alpha$ interval, where m is the number of independent tests evaluated. For the smoothed periodogram, spectral estimates separated by the bandwidth, bw , are essentially independent, such that however many bw ’s are required to span the frequency axis $[0, 0.5]$ is approximately the number of independent tests. In other words, $m=0.5/bw$. For example, if $bw=.05$, or $1/10$ the width of the frequency axis of the spectrum, a 99.5% confidence interval instead of the 95% confidence should be used in equation 22 as the “corrected” 95% confidence interval $[(1-\alpha/m) = (1-0.05/10) = 0.995]$. The effect of the correction is to widen the confidence band, making it less likely to attribute significance to spectral peaks.

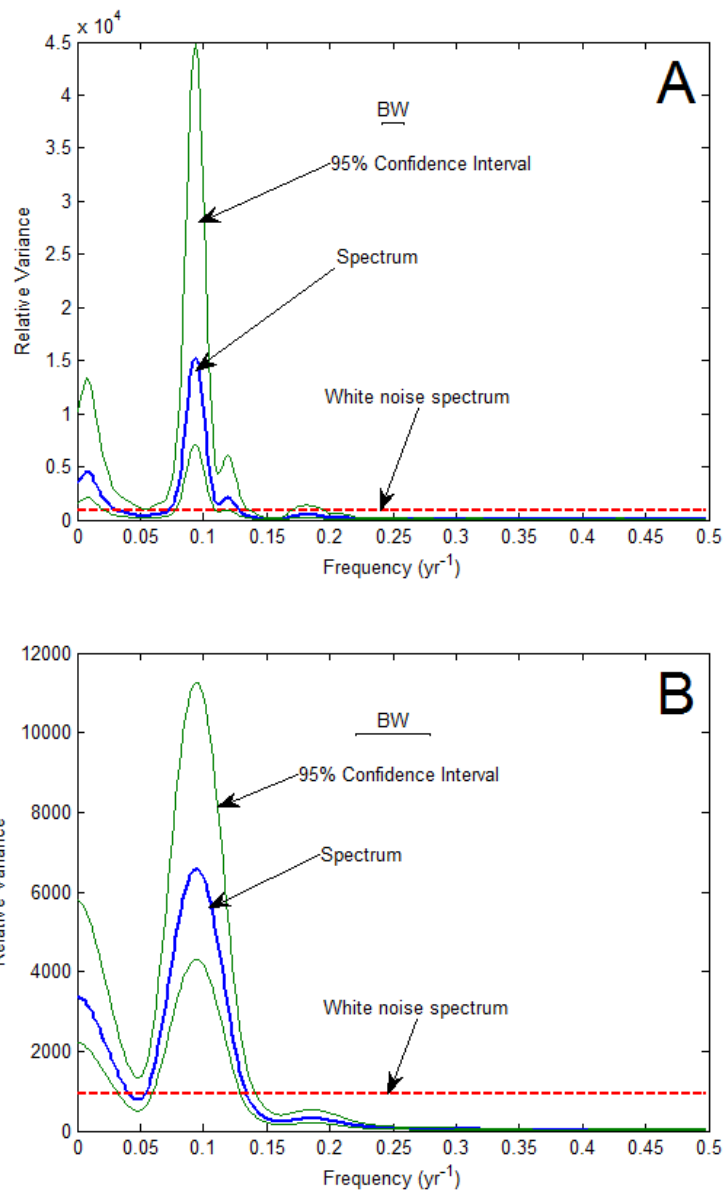


Figure 6.9. Spectra of Wolf sunspot number using two levels of smoothing of raw periodogram. (A) Smoothing with Daniell filter spans [3 5 7]. (B) Smoothing with Daniell filter spans [11 15 23]. Note the difference in range of y-axis.

To summarize, the test for periodicity begins with specification of a period or frequency of interest. Second, the spectrum and its confidence interval are estimated, possibly using a window-closing procedure. Third, a null continuum is drawn so that the peaks in the spectrum can be compared to a “null” spectrum without those peaks but with the same broad underlying spectral shape. Finally, the peak is judged significant at 95% if the lower CI does not include the null continuum.

6.6 Additional considerations: tapering, padding and leakage

Tapering and padding. Tapering and padding are mathematical manipulations sometimes performed on the time series before periodogram analysis to improve the statistical properties of the spectral estimates or to speed up the computations. In spectral analysis, a time series is regarded as a finite sample of an infinitely long series, and the objective is to infer the properties of the infinitely long series. If the observed time series is viewed as repeating itself an infinite number of times, the sample can be considered as resulting from applying a *data window* to the infinite series. The data window is a series of weights equal to 1 for the N observations of the time series and zero elsewhere. This data window is rectangular in appearance. The effect of the rectangular data window on spectral estimation is to distort the estimated spectrum of the unknown infinite-length series by introducing *leakage*. *Leakage* refers to the phenomenon by which variance at an important frequency (say a frequency of a strong periodicity) “leaks” into other frequencies in the estimated spectrum. The net effect is to produce misleading peaks in the estimated spectrum.

The objective of *tapering* is to reduce leakage. Tapering consists of altering the ends of the mean-adjusted time series so that they taper gradually down to zero. Before tapering, the mean is subtracted so that the series has mean zero. A mathematical taper is then applied. A frequently used taper function is the split cosine bell, given by

$$w_p(t) = \begin{cases} \frac{1}{2}\{1 - \cos 2\pi t / p\}, & 0 \leq t < p/2, \\ 1, & p/2 \leq t < 1 - p/2 \\ \frac{1}{2}\{1 - \cos 2\pi(1-t)/p\}, & 1 - p/2 \leq t \leq 1 \end{cases} \quad (23)$$

where p is the proportion of data desired to be tapered, t is the time index, and $w_p(t)$ are the taper weights. A suggested proportion is 10%, or $p = 0.10$, which means that 5% is tapered on each end (Bloomfield 2000, p. 69).

Padding. The Fast Fourier Transform (FFT), introduced by Cooley and Tukey (1965), is a computational algorithm that can greatly speed up computation of the Fourier transform and spectral analysis. The FFT is most effective if the length of time series, n , has small prime numbers. One way of achieving this is to *pad the time series with zeros* until the length of the series is a power of 2 before computing the Fourier transform. The padded data are defined as

$$x'_t = \begin{cases} x_t & 0 \leq t < n \\ 0 & n \leq t < n' \end{cases} \quad (24)$$

where x_t is the original time series, after subtracting the mean. It can be shown (Bloomfield 2000, p. 61) that the discrete Fourier transform of the padded series differs trivially from that of the original series

As a side effect of padding, the grid of frequencies on which the transform is calculated is changed to a finer spacing. This change suggests that padding with zeros can also be used to alter

the Fourier frequencies such that some period of a-priori interest falls near a Fourier frequency. This is an acceptable procedure (e.g., Mitchell et al. 1966). The resulting finer spacing of Fourier frequencies for a given span of Daniell filter gives a spectral estimate with a narrower bandwidth (see #3 under “Resolution” above), but the increase in resolution comes at the expense of a decrease in stability of the spectral estimate (see eqn (26) below).

Effect of padding and tapering on stability. Tapering and padding both have the effect of increasing the variance of the spectral estimate. If the time series is tapered by the split cosine bell taper and the total proportion of the series tapered is p , the variance of the spectral estimate (see eqn (20)) is increased by a factor of

$$c_T = \frac{128 - 93p}{2(8 - 5p)^2} \quad (25)$$

If the time series is padded from an initial length of N to a padded length of N' , the variance is increased by a factor of

$$c_P = \frac{N'}{N} \quad (26)$$

If a time series has been padded and tapered, an equation of form (22) can still be used for the confidence interval for the spectrum, except with an effective degrees of freedom defined as

$$\nu = \frac{2}{g_*^2} \quad (27)$$

where

$$g_*^2 = c_T c_P \sum_u g_u^2 \quad (28)$$

An example will serve to illustrate the computation of a confidence interval when the series has been padded and tapered before computation of the spectrum. Say the original time series has a length 300 years, a total of 20% of the series has been tapered, and that the tapered series has then been padded to length 512 by appending zeros. Equations (25) and (26) give variance inflation factors

$$c_T = \frac{128 - 93p}{2(8 - 5p)^2} = \frac{128 - 93(.2)}{2(8 - 5(.2))^2} = 1.1163 \quad (29)$$

and

$$c_P = \frac{512}{300} = 1.7067 \quad (30)$$

If the periodogram is smoothed by a 5-weight Daniell filter, $\{.125 \ .25 \ .25 \ .25 \ .125\}$, the quantity g_*^2 is given by

$$g_*^2 = c_T c_P \sum_u g_u^2 = 1.1163(1.7067)(.2188) = 0.4169, \quad (31)$$

equivalent degrees of freedom are

$$\nu = \frac{2}{g_*^2} = \frac{2}{0.4169} = 4.80 \approx 5, \quad (32)$$

and the 95% confidence interval is

$$\frac{5\hat{s}(f)}{12.83} \leq s(f) \leq \frac{5\hat{s}(f)}{.8312} \quad \text{or} \quad (33)$$

$$0.39\hat{s}(f) \leq s(f) \leq 6.01\hat{s}(f)$$

The estimated spectrum, $\hat{s}(f)$ is plotted with lower and upper confidence bands at $0.39\hat{s}(f)$ and $6.01\hat{s}(f)$, along with some null continuum (e.g., white noise). A peak is judged significant if the lower confidence interval does not include the null continuum. Note that the confidence interval is wider above the estimated spectrum than below the estimated spectrum (asymmetrical).

6.7 References

- Anderson, O.D.. 1976. Time series analysis and forecasting: the Box-Jenkins approach. Butterworths, London, 182 p.
- Blackman, R.B., and Tukey, J.W., 1959, The measurement of power spectra, from the point of view of communications engineering: New York, Dover.
- Bloomfield, P., 2000, Fourier analysis of time series: an introduction, second edition: New York, John Wiley & Sons, Inc., 261 p.
- Chatfield, C., 2004, The analysis of time series, an introduction, sixth edition: New York, Chapman & Hall/CRC.
- Cooley, J.W., and Tukey, J.W., 1965, An algorithm for the machine computation of complex Fourier series: Math. Comput., v. 19, p. 297-301.
- Daniell, P.J., 1946, Discussion on the symposium on autocorrelation in time series: J. Roy. Statist. Soc. (Suppl.), v. 8, p. 88-90.
- Einstein, A., 1914, Arch. Sci. Phys. Natur., v. 4.37, p. 254-256.
- Hamming, R.W., and Tukey, J.W., 1949, Measuring noise color, Bell Telephone Laboratories Memorandum.
- Hayes, M.H., 1996, Statistical digital signal processing and modeling: New York, John Wiley & Sons, Inc.
- Kay, S.M., 1988, Modern spectral estimation: Engelwood Cliffs, NJ, Prentice Hall.
- Lagrange, 1873, Recherches sur la manière de former des tables des planètes d'après les seules observations, in Oeuvres de Lagrange, v. VI, p. 507-627.
- Mitchell, J.M., Jr., Dzerdzevskii, B., Flohn, H., Hofmeyr, W.L., Lamb, H.H., Rao, K.N., and Wallén, C.C., 1966, Climatic change: Technical Note No. 79, report of a working group of the Commission for Climatology; WMO No. 195 TP 100: Geneva, Switzerland, World Meteorological Organization, 81 p.
- Percival, D.B., and Walden, A.T., 1993, Spectral analysis for physical applications: Cambridge University Press.
- Schuster, A., 1897, On lunar and solar periodicities of earthquakes: Proc. Roy. Soc., p. 455-465.
- Snedecor, G.W., and Cochran, William G., 1989, Statistical methods, eighth edition, Iowa State University Press, Ames, Iowa, 803 pp.
- The MathWorks, I., 1996, Matlab signal processing toolbox: Natick, MA, The MathWorks, Inc.
- Thomson, W., 1876, On an instrument for calculating the integral of the product of two given functions: Proc. Roy. Soc., v. 24, p. 266-68.
- Vautard, R., and Ghil, M., 1989, Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, Physica D 35, 395-424.

- Welch, P.D., 1967, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms: IEEE Trans. Audio Electroacoust., v. AU-15, p. 70-73.
- Wilks, Daniel S., 1995. Statistical methods in the atmospheric sciences. Academic Press, New York. 467p.