6 Spectral Analysis -- Smoothed Periodogram Method

6.1 Historical background

There are many available methods for estimating the spectrum of a time series. In lesson 4 we looked at the Blackman-Tukey method, which is based on Fourier transformation of the smoothed, truncated autocovariance function. The smoothed periodogram method circumvents the transformation of the acf by direct Fourier transformation of the time series and computation of the raw periodogram, a function first introduced for study of time series back in the 1800s. Smoothing of the periodogram into spectral estimates is accomplished by applying spans of filters. By choosing the spans of the filters, you control the smoothness, resolution and variance of the spectral estimates. The same smoothing procedure is extended to produce an underlying smoothly varying spectrum, or null continuum, against which spectral peaks can be tested for significance. This approach is an alternative to the specification of a functional form of the null continuum (e.g., AR spectrum).

The periodogram was one of the earliest statistical tools for studying periodic tendencies in time series (Figure 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.

Figure 1. Timeline of developments in spectral analysis of time series. Timeline based on information in Bloomfield (2000, p. 5) and Hayes (1996).
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). (This use of the smoothed periodogram in this sense had actually been described much earlier by Einstein (1914).)

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 (1975, p. 145), 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>
<thead>
<tr>
<th>Method</th>
<th>Summary</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackman-Tukey</td>
<td>Fourier transformation of smoothed, truncated autocovariance function</td>
<td>Chatfield, 1975</td>
</tr>
<tr>
<td><strong>Smoothed periodogram</strong></td>
<td>Estimate periodogram by DFT of time series; Smooth periodogram with modified Daniell filter</td>
<td>Bloomfield, 2000</td>
</tr>
<tr>
<td>Welch’s method</td>
<td>Averaged periodograms of overlapped, windowed segments of a time series</td>
<td>Welch, 1967</td>
</tr>
<tr>
<td><strong>Multi-taper method</strong></td>
<td>Use orthogonal windows (“tapers”) to get approximately independent estimates of spectrum; combine estimates</td>
<td>Percival and Walden, 1993</td>
</tr>
<tr>
<td>(MTM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Singular spectrum analysis (SSA)</strong></td>
<td>Eigenvector analysis of autocorrelation matrix to eliminate noise prior to transformation to spectral estimates</td>
<td>Vautard and Ghil, 1989</td>
</tr>
<tr>
<td><strong>Maximum entropy (MEM)</strong></td>
<td>Parametric method: estimate acf and solve for AR model parameters; AR model has theoretical spectrum</td>
<td>Kay, 1988</td>
</tr>
</tbody>
</table>

### 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
Subtraction of the mean is a vital preliminary step, especially if padding (described later) is used in the analysis. Detrending is also important, as spectral analysis is intended for the study of stationary series. Mathematical operations associated with steps 2-4 are described below.

**Discrete Fourier transform.** Say \( x_0, x_1, \ldots, x_{n-1} \) 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:

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

(1)

where the summation is over Fourier frequencies

\[
f_j = \frac{j}{n}, \quad j = 1,2,\ldots,(n-1)/2,
\]

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

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

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

(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

\[
e^{ix} = \cos x + i \sin x
\]

(3)

and its inverse

\[
\cos x = \frac{1}{2} \left\{ e^{ix} + e^{-ix} \right\}, \quad \sin x = \left\{ e^{ix} - e^{-ix} \right\}
\]

(4)

In general, observed data are strictly real-valued, but they may be regarded as complex numbers with zero imaginary parts. Suppose \( x_0, x_1, \ldots, x_{n-1} \) 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

\[
d(f) = \frac{1}{n} \sum_{t=0}^{n-1} x_t e^{-2\pi j f t}
\]

(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

\[
d(f) = \frac{A(f)}{2} - i \frac{B(f)}{2}
\]

(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

\[
x_t = \sum_j d(f_j) e^{2\pi j f t}
\]

(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 \(\phi(f)\)

\[
d(f) = R(f)e^{i\phi(f)}
\]  

(8)

The magnitude, given by

\[
R(f) = |d(f)|
\]  

(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

\[
I(f) = n\left[R(f)^2 + n|d(f)|^2\right]
\]  

(10)

The sine and cosine terms at the Fourier frequencies are orthogonal, and so the variance of the time series \(x\) 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

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

(11)

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

\[
\sum_{j=0}^{n-1} |x_j|^2 = n\sum_j |d(f_j)|^2 = \sum_j I(f_j)
\]  

(12)

If \(x\) 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 indicates 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 \(f\) is proportional to the variance accounted for by that frequency component.

**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

\[
g_i = \begin{cases} 
\frac{1}{2(m-1)}, & i = 1 \text{ or } i = m \\
\frac{1}{m-1}, & i \text{ otherwise}
\end{cases}
\]  

(13)

where \(m\) is the number of weights, or span of the filter, and \(g_i\) is the \(i^{th}\) weight of the filter. 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. For example, the following figures shows filter weights of 5-weight Daniell and rectangular filters. 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).

6.3 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. Smoothing the periodogram with Daniell filters of various spans results in a spectrum much smoother in appearance than the raw periodogram. Excessive smoothing obscures the important spectral detail; 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. If the time series has not been padded or tapered, the variance of the spectral estimate is given by

\[
\text{var}\{\hat{s}(f)\} = s(f)^2 \sum g_u^2
\]

where \(\hat{s}(f)\) is the spectral estimate at frequency \(f\), \(s(f)\) is the true, and unknown value of the spectrum, assumed to be approximately constant over the interval of averaging, and the summation \(\sum g_u^2\) 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

![Figure 2. Daniell filter and rectangular filter of span, or length, 5.](image-url)
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 \(\chi^2\) distributed. The distribution of \(\hat{S}(f)\) can be shown to be approximately \(\chi^2\) with degrees of freedom

\[
\nu = \frac{2}{g^2} \tag{15}
\]

where \(g^2 = \sum u g^2_u\) is the sum of squared Daniell weights. The relationship in (15) can be used to place a confidence interval around the spectral estimates. For example, a 95% confidence interval for \(\hat{s}(f)\) is given by

\[
\frac{\nu \hat{s}(f)}{\chi^2(0.975)} \leq s(f) \leq \frac{\nu \hat{s}(f)}{\chi^2(0.025)} \tag{16}
\]

where \(\chi^2(0.025)\) and \(\chi^2(0.975)\) are the 2.5% and 97.5% points of the \(\chi^2\) distribution with \(\nu\) 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 were 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 would depend 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 \( \Delta f \) is the spacing of the Fourier frequencies. (Note that if the series has been padded to length \( N' \), the spacing is taken as \( 1/N' \))

### 6.4 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. The confidence bands developed above (e.g., equation 1.16) are not simultaneous. In other words, the bands can be used strictly to test for significance of a peak at a specified frequency, which should be specified before running the spectral analysis.
2. 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. The null spectrum might be white noise, but more generally will not be. One approach is to specify a “null continuum” as a very smooth underlying spectrum (Mitchell et al. 1966). The null hypothesis is then that the estimated spectrum is no different than this underlying spectrum. Bloomfield (2000) suggests that a greatly smoothed version of the periodogram might serve as an empirically derived null continuum. This approach was used in a tree-ring evaluation of periodicity of tree growth in the corn belt of the U.S. (Meko et al. 1985).

The test for periodicity can then proceed as follows:
1. Specify beforehand your period or frequency of interest
2. Estimate the spectrum and its 95% CI using the smoothed periodogram method
3. Estimate the null continuum using much greater smoothing of the periodogram
4. Plot the spectrum, its CI, and the null continuum, and consider a peak significant at 95% if its lower CI does not include the null continuum

### 6.5 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

\[
\frac{1}{2} \left\{ \begin{array}{ll}
1 - \cos \frac{2\pi t}{p} & , \quad 0 \leq t < p/2, \\
1 & , \quad p/2 \leq t < 1 - p/2 \\
1 - \cos \frac{2\pi (1-t)}{p} & , \quad 1 - p/2 \leq t \leq 1
\end{array} \right.
\]

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

\[
x'_i = \begin{cases} 
x_i & , \quad 0 \leq t < n \\
0 & , \quad n \leq t < n'
\end{cases}
\]

where \( x_i \) 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 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 (20) 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 (14)) is increased by a factor of

\[
c_r = \frac{128 - 93p}{2(8 - 5p)^2}
\]

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

\[
c_p = \frac{N'}{N}
\]

If a time series has been padded and tapered, an equation of form (16) can still be used for the confidence interval for the spectrum, except with an effective degrees of freedom defined as
where

\[ g^2 = c_r c_p \sum g_u^2 \]  

A simple 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 (19) and (20) give variance inflation factors

\[ \frac{128-93p}{2(8-5p)^2} = \frac{128-93(0.2)}{2(8-5(0.2))^2} = 1.1163 \]  

and

\[ c_p = \frac{512}{300} = 1.7067 \]  

If the periodogram is smoothed by a 5-weight Daniell filter, \{0.125 0.25 0.25 0.25 0.125\}, the quantity \( g^2 \) is given by

\[ g^2 = c_r c_p \sum g_u^2 = 1.1163(1.7067)(0.2188) = 0.4169, \]  

equivalent degrees of freedom are

\[ v = \frac{2}{g^2} = \frac{2}{0.4169} = 4.80 \approx 5, \]  

and the 95% confidence interval is

\[ \frac{5\hat{s}(f)}{12.83} \leq s(f) \leq \frac{5\hat{s}(f)}{0.8312} \quad \text{or} \]  

\[ 0.39\hat{s}(f) \leq s(f) \leq 6.01\hat{s}(f) \]  

\[ \hat{v} = \frac{2}{g^2} \]
6.6 References


Lagrange, 1873, Recherches sur la manière de former des tables des planètes d'apres les seules observations, in Oeuvres de Lagrange, v. VI, p. 507-627.


