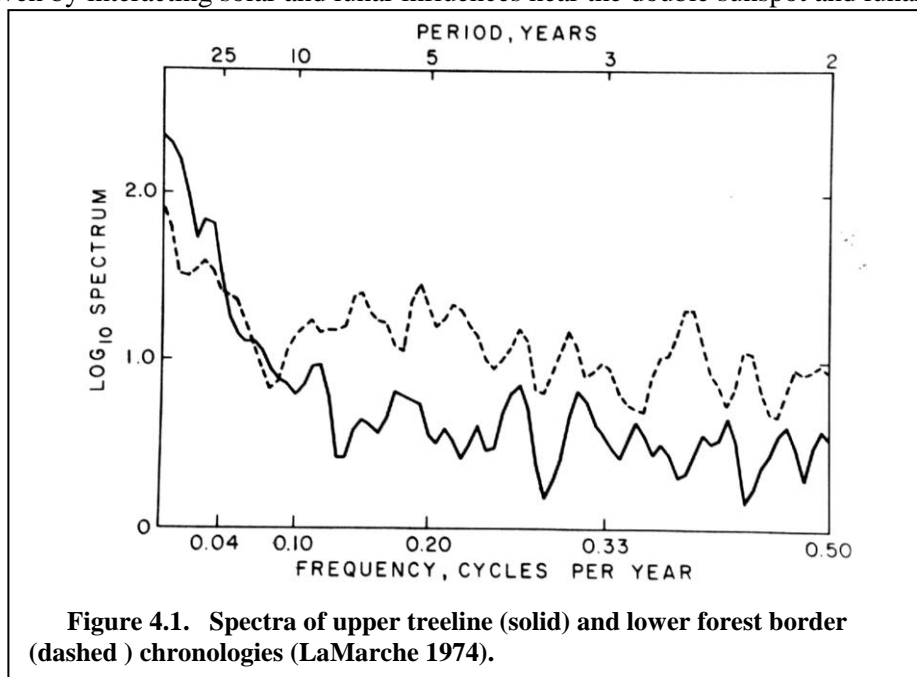


4 Spectrum

The *spectrum* of a time series is the distribution of variance of the series as a function of frequency. The object of *spectral analysis* is to estimate and study the spectrum. The spectrum contains no new information beyond that in the autocovariance function (acvf), and in fact the spectrum can be computed mathematically by transformation of the acvf. But the spectrum and acvf present the information on the variance of the time series from complementary viewpoints. The acf summarizes information in the *time domain* and the spectrum in the *frequency domain*.

Why analyze the spectrum?

The spectrum is of interest because many natural phenomena have variability that is frequency-dependent, and understanding the frequency dependence may yield information about the underlying physical mechanisms. Spectral analysis can help in this objective. A classic example from dendroclimatology is Val LaMarche's (1974) application of spectral analysis to study differences in frequency properties of tree growth at the upper treeline and lower forest border in eastern Nevada. LaMarche found that lower frequencies dominated the variance at the upper treeline, and that higher frequencies were more important at the lower treeline (Figure 4.1). From correlation analysis of frequency-stratified components of variability, he concluded that the low-frequency variations reflected temperature fluctuations and the high-frequency fluctuations precipitation variations. A couple of years earlier, LaMarche and Fritts (1972) had first applied modern spectral analysis methods to study a possible solar-variability signal in tree rings. This topic received much attention from dendroclimatologists in subsequent years. Ed Cook followed up on the solar-tree ring connection by applying spectral analysis methods to relate tree-ring variations at hundreds of sites in North America to an index of area covered by drought (Cook et al. 1997) and concluded that the data supports a bi-decadal rhythm in drought area possibly driven by interacting solar and lunar influences near the double sunspot and lunar nodal periods.



4.1 The frequency domain

In the time domain, variations are studied as a function of time. For example, the time series plot of an annual tree-ring index displays variations in tree-growth from year to year, and the acf summarizes the persistence of a time series in terms of correlation between lagged values for different numbers of years of lag. In the frequency domain, the variance of a time series is studied as a function of frequency or wavelength. The main building blocks of variation in the frequency domain are sinusoids, or sines and cosines. In discussing the frequency domain, it is helpful to start with definitions pertaining to waves. For simplicity, we will use a time increment of one year. Consider the simple example of an annual time series y_t generated by superimposing random normal noise on a cosine wave:

$$y_t = R \cos(\omega t + \phi) + z_t = w_t + z_t \quad (1)$$

where t is time (years, for this example), z_t is the random normal component in year t , w_t is the sinusoidal component; and R , ω and ϕ are the amplitude, angular frequency (radians per year), and phase of the sinusoidal component.

The plot in Figure 4.2 shows a time series generated by model (1) with the following settings:

- time series length of 201 years
- sinusoidal component with wavelength 100 years, amplitude 1.0, and phase 0 degrees relative to time 0
- noise component from normal distribution with mean 0 and variance 0.01

The *peaks* are the high points in the wave; the *troughs* are the low points. The wave varies around a mean of zero. The vertical distance from zero to the peak is called the *amplitude*. The variance of a sinusoid is proportional to the square of the amplitude: $\text{var}(w_t) = R^2/2$. The *phase* ϕ describes the offset in time of the peaks or troughs from some fixed point in time. From the relationship between variance and amplitude, the sinusoidal component in this example has a variance of 50 times that of the noise (0.5 is 50 times 0.01).

The angular frequency ω describes the number of radians of the wave in a unit of time, where 2π radians corresponds to a complete cycle of the wave (peak to peak). In practical applications, the frequency is often expressed by f , the number of cycles per time interval. The relationship between the two frequency measures is given by

$$f = \omega/(2\pi) \quad (2)$$

The *wavelength*, or *period*, of the cosine wave is the distance from peak to peak, and is the inverse of the frequency

$$\lambda = \frac{1}{f} \quad (3)$$

Thus a frequency of one cycle per year corresponds to an angular frequency of 2π radians per year and a wavelength of 1 year.

From the relationship (3), the frequency of the cosine wave in Figure 4.2 is

$$f = 1/100 = 0.01 \text{ cycles per year} \quad (4)$$

Since 0.01 cycle is covered in one year, 100 years is required for a full cycle. The the angular frequency is

$$\omega = 2\pi f = 0.0628 \text{ radians per year} \quad (5)$$

A frequency of one cycle every two years corresponds to an angular frequency of π radians per year, or a wavelength of 2 years. In the analysis of annual time series, this frequency of $f = 0.5$ cycles/yr or $\omega = \pi$ radians/yr corresponds to what is called the *Nyquist* frequency, which is the highest frequency for which information is given by the spectral analysis.

Another important frequency in spectral analysis is the *fundamental* frequency, also referred to as the first *harmonic*. If the length of a time series is N years, the fundamental frequency is $1/N$. The corresponding fundamental period is N years, or the length of the time series. For example, the fundamental period of a time series of length 500 years is 500 years – a wave that undergoes a complete cycle over the full length of the time series.

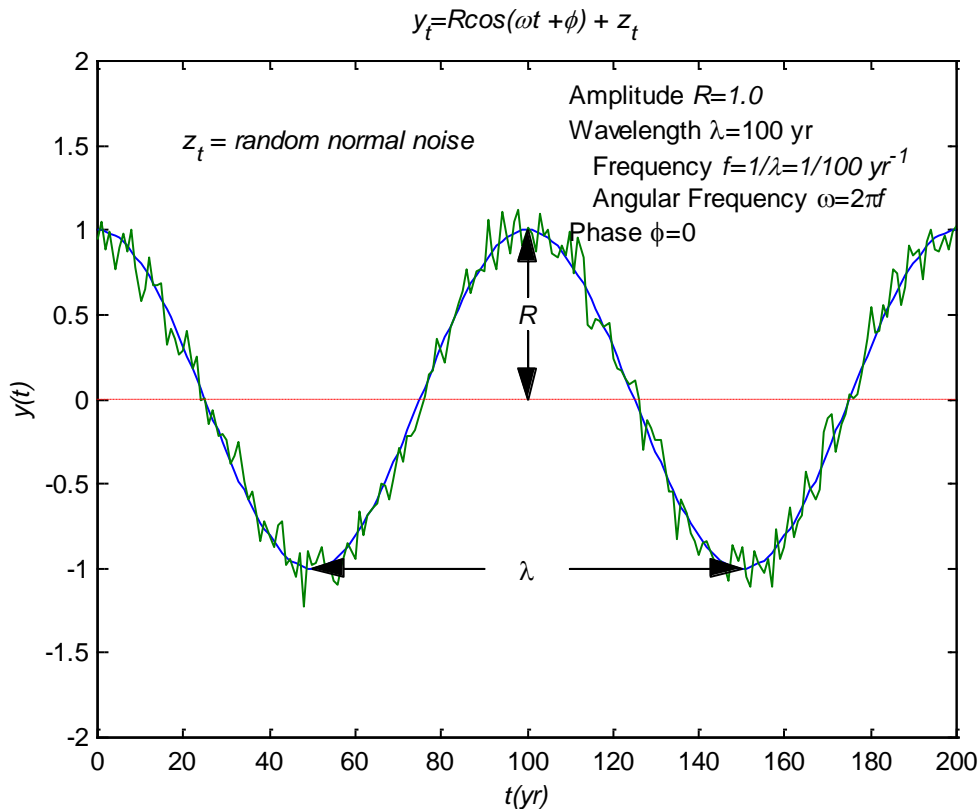


Figure 4.2. Simple example of periodic series with superimposed noise.

4.2 Sinusoidal model of a time series

The model given by equation (1) is extremely simple, consisting of just a single sinusoidal component with superposed noise. Following Percival and Walden (1993), the spectrum can be defined in terms of a more complicated model in which a time series X_t of length N consists of a linear combination of many sinusoids at fixed frequencies $\{f_j\}$ and random amplitudes $\{A_j\}$ and $\{B_j\}$:

$$X_t = \mu + \sum_{j=1}^{[N/2]} [A_j \cos(2\pi f_j t) + B_j \sin(2\pi f_j t)], \quad t=1,2,\dots,N \quad (6)$$

where μ is a constant term, the notation $[N/2]$ refers to the greatest integer less than or equal to $N/2$, and the frequencies f_j are related to the sample size N by

$$f_j \equiv j/N, \quad 1 \leq j \leq [N/2] \quad (7)$$

The frequencies of the sinusoids are at intervals of $1/N$ and are the *Fourier frequencies*, or *standard frequencies*. The frequency f_j is j th *standard frequency* (Figure 4.3). The standard frequencies clearly depend on the sample size. For example, for a 500-year tree-ring series, the standard frequencies are at $1/500, 2/500, \dots$ cycles per year. The highest standard frequency is $f = (N/2)/N = 1/2 = 0.5$, which corresponds to a wavelength of two years.

In developing the definition of the spectrum with this model, the additional assumptions must be made that the amplitudes $\{A_j\}$ and $\{B_j\}$ are random variables with expected values

$$E\{A_j\} = E\{B_j\} = 0 \quad (8)$$

and

$$E\{A_j^2\} = E\{B_j^2\} = \sigma_j^2 \quad (9)$$

From the relationship between variance and amplitude of sinusoid, equation (9) implies that the variance associated with the j th standard frequency is σ_j^2 . It must also be assumed that the amplitudes associated with various standard frequencies are uncorrelated:

$$E\{A_j A_k\} = E\{B_j B_k\} = 0 \quad \text{for } j \neq k \quad (10)$$

and

$$E\{A_j B_k\} = 0 \quad \text{for all } j, k. \quad (11)$$

With the assumptions above, it can then be shown that the expected value of X_t is

$$E\{X_t\} = \mu \quad (12)$$

the variance of X_t is

$$\sigma^2 = E\{(X_t - \mu)^2\} = \sum_{j=1}^{[N/2]} \sigma_j^2 \quad (13)$$

and the autocorrelation function of X_t is

$$\rho_k = \frac{\sum_{j=1}^{[N/2]} \sigma_j^2 \cos(2\pi f_j k)}{\sum_{j=1}^{[N/2]} \sigma_j^2} \quad (14)$$

Equation (13) says that the variance of the series X_t is the sum of the sum of the variances associated with the sinusoidal components at the different standard frequencies. Thus the variance of the series can be decomposed into components at the standard frequencies -- the variance can be expressed as a function of frequency.

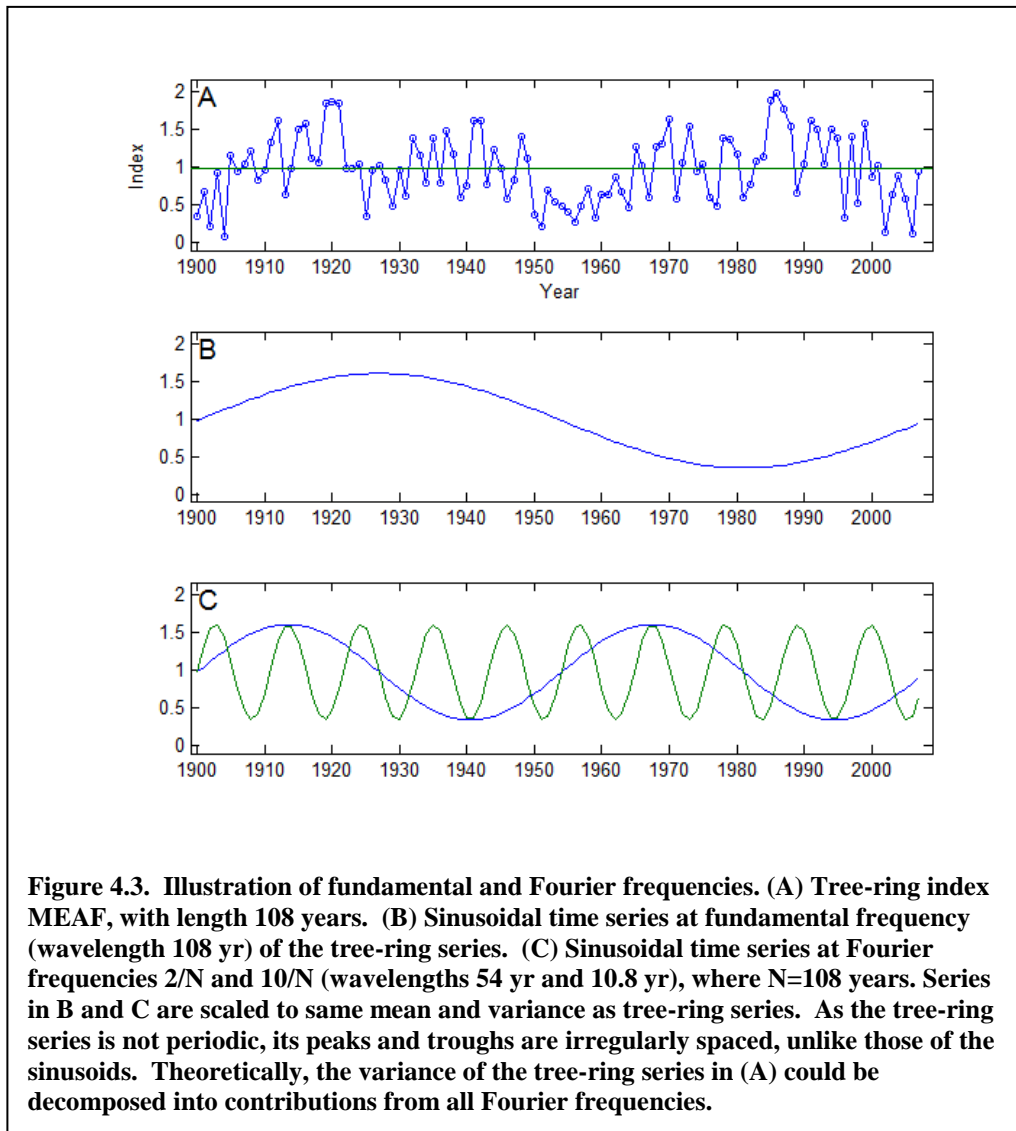
For the model given above, the spectrum can be defined as

$$S_j^2 \equiv \sigma_j^2, \quad 1 \leq j \leq [N/2] \quad (15)$$

A plot of S_j against frequencies f_j shows the variance contributed by the sinusoidal terms at each of the standard frequencies. From equation(13), the variance of X_t can then be expressed as the sum of the spectral components

$$\sigma^2 = \sum_{j=1}^{[N/2]} S_j \quad (16)$$

The variance contributed at frequency f_j is the spectrum S_j at that frequency. The shape of the spectral values S_j plotted against f_j indicates which frequencies are most important to the variability of the time series.



Considering that σ_j^2 are spectral values, equation (14) gives an important relationship between the spectrum and the autocorrelation function: the acf is expressed as a cosine transform of the spectrum. Similarly the spectrum can be shown to be the Fourier transform of the acf. The spectrum and acf are therefore different characterizations of the same time series information. The acf is a time-domain characterization and the spectrum is a frequency-domain characterization. From a practical point of view, the spectrum and acf are complementary to each other. Which is most useful depends on the data and the objective of analysis.

4.3 Harmonic analysis = periodogram analysis = Fourier analysis

For a given time series, it is possible to apply the above model to mathematically estimate the parameters of the sinusoidal terms at each Fourier frequency. Such an analysis is called *Fourier analysis*, or *harmonic analysis* Chatfield (2004; Panofsky and Brier 1958). Harmonic analysis is most appropriate for phenomena with known periodic components. For example, a single time series made up 12 values of the long-term mean of monthly temperature will likely have a well-defined annual (12-month) component at the fundamental frequency. Harmonic analysis can be used to quantify the importance of this annual wave relative to other components of the variability of the annual distribution of monthly means.

In harmonic analysis, the frequencies j/N , $j=1, \dots, N/2$ are referred to as the harmonics: $1/N$ is the first harmonic, $2/N$ the second harmonic, etc. Any series can be decomposed mathematically into its $N/2$ harmonics. The sinusoidal components at all the harmonics effectively describe all the variance in a series. A plot of the variance associated with each harmonic as a function of frequency has been referred to above as the “spectrum”, for the hypothesized model. Such a plot of variance (sometimes scaled in different ways) against frequency is also called the *periodogram* of the series, and the analysis is called *periodogram analysis* (Chatfield 2004). Spectral analysis, to be described next, departs from periodogram analysis in an important way: in spectral analysis, the time series is regarded as just one possible realization from a random process, and the objective is to estimate the spectrum of that process using just the observed time series.

4.4 Spectral analysis

In analyzing the spectrum, we want to acknowledge the uncertainty in trying to understand a process from a single sample. Spectral analysis is therefore concerned with estimating the unknown spectrum of the process from the data and with quantifying the relative importance of different frequency bands to the variance of the process. The spectrum that is being estimated in a sense then is not really the spectrum of the observed series, but the spectrum of the unknown infinitely long series from which the observed series is assumed to have come.

Various methods have been developed to estimate the spectrum from an observed time series. For an overview and comparisons of different methods, see Percival and Walden (1993), Chatfield (2004), and Bloomfield (1976). In this chapter, we use the Blackman-Tukey method, one of several available nonparametric methods (Percival and Walden 1993). (Later in the semester we will also use another spectral estimation method -- the smoothed periodogram.)

The presentation on the Blackman-Tukey method closely follows Chatfield (2004). Note the use of angular frequency in the equations.

Spectral distribution function

We previously defined the sample autocorrelation r_k and autocovariance function c_k . We will refer to these functions also as the “acf” and the “acvf.” In the notation of Chatfield (2004), the corresponding population statistics are the population acvf $\gamma(k)$ and population acf $\rho(k)$. An important time series theorem, called the Wiener-Khinchine theorem, says that for any stationary stochastic process with acvf $\gamma(k)$, there exists a monotonically increasing function, $F(\omega)$, such that

$$\gamma(k) = \int_0^\pi \cos \omega k dF(\omega) \quad (17)$$

Equation (17) is called the spectral representation of the autocovariance function, and $F(\omega)$ is called the *spectral distribution function*. $F(\omega)$ has a direct physical interpretation:

$$F(\omega) = \text{contribution to the variance of the series which is} \\ \text{accounted for by frequencies in the range } (0, \omega) \quad (18)$$

Note the similarity of equation (18) to the relationship between autocorrelation function and spectrum for the simple sinusoidal model discussed previously (14). Why is the range restricted to angular frequencies $(0, \pi)$? First, there is no variation at negative frequencies, so the lower limit on the range is 0. Second, if a process is measured at unit intervals, the highest possible frequency that can be studied corresponds to wave that undergoes a complete cycle in two intervals, or an angular frequency of $\omega = \pi$. Thus, all the variation is accounted for by frequencies less than π :

$$F(\pi) = \text{var}(X_t) = \sigma_x^2 \quad (19)$$

The function $F(\omega)$ increases monotonically between $\omega = 0$ and $\omega = \pi$, and in this way $F(\omega)$ is similar to a cumulative distribution function, or cdf. In fact, by scaling $F(\omega)$ by the variance, we get what is called the *normalized spectral distribution function*

$$F^*(\omega) = F(\omega) / \sigma_x^2 \quad (20)$$

which gives the *proportion* of variance accounted for by frequencies in the range $(0, \omega)$, and like a cdf, reaches a maximum of 1.0, since $F^*(\pi) = 1$.

Spectral density function, or spectrum

The spectral distribution function gives the variance of the process at frequencies less than some frequency. By differentiating the spectral distribution function we get the *spectral density function*, which gives the variance associated with each frequency

$$f(\omega) = \frac{dF(\omega)}{d\omega} \equiv (\text{power}) \text{ spectral density function} \quad (21)$$

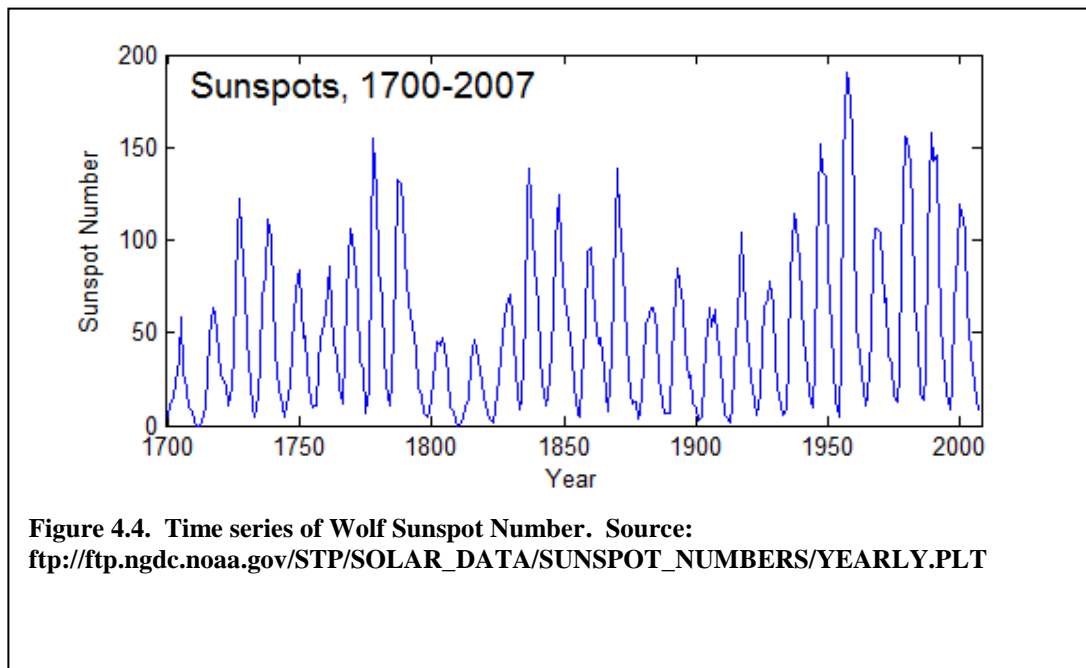
The term “spectral density function” is often shortened to *spectrum*. The adjective “power” is often omitted. “Power” comes from the application of spectral analysis in engineering, and is related to the passage of an electric current through a resistance. For a sinusoidal input, the power is directly proportional to the squared amplitude of the oscillation. We have seen that for a time series the variance of a sinusoid is proportional to its squared amplitude. Thus “power” is equivalent to “variance.”

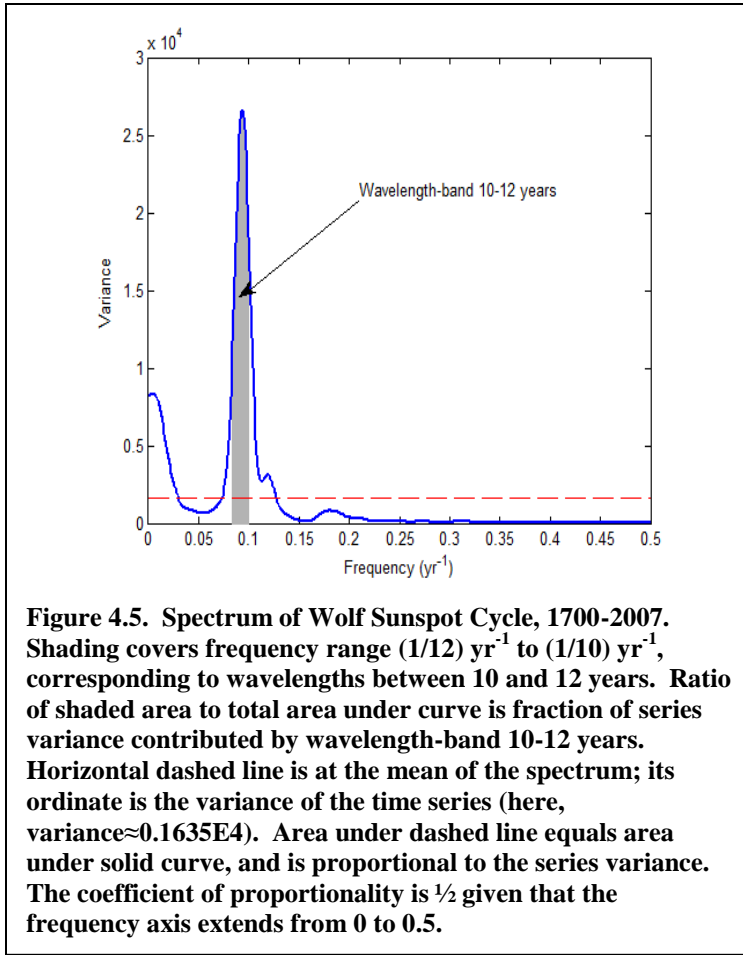
Relationship between variance and spectrum of a time series

A plot of the spectrum $f(\omega)$ against frequency ω is essentially a plot of the variance of a time series as a function of frequency. More precisely, a given value of $f(\omega)$ is a “variance per unit frequency”, such that by integrating over some increment of $f(\omega)$ we get the variance associated with that range of frequencies. In other words, if $d\omega$ is some increment of frequency, $f(\omega)d\omega$ is the contribution to the variance of components with frequencies in the range $(\omega, \omega + d\omega)$.

In a graph of the spectrum, therefore, the area under the curve bounded by two frequencies represents the variance in that frequency range, and the total area underneath the curve represents the variance of the series. A peak in the spectrum represents relatively high variance in a frequency band centered on the peak. The Wolf sunspot series appears to have about 10 peaks per century (Figure 4.4). The spectrum of this series peaks at a wavelength somewhat longer than 10 years, and indicates a large fraction of the variance is contributed by wavelengths between 10 and 12 years (Figure 4.5).

A flat spectrum indicates that variance is evenly distributed over frequencies. Theoretical “white noise” is characterized by such even distribution of variance, and is represented by a horizontal line in the spectral plot (see dashed line in Figure 4.5). A sampled time series referred to as white noise series therefore has no tendency for amplified variance at any particular frequency, and has no significant spectral peaks.





Relationship between acvf and spectrum

In equation (17) the acvf was expressed in terms of a cosine transform of increments of the spectral distribution function. That relationship also means that the acvf can be expressed as a cosine transform of the spectral density function, or spectrum. An inverse relationship can also be shown, such that the spectrum is the *Fourier transform* of the acfv

$$f(\omega) = \frac{1}{\pi} \left[\gamma(0) + 2 \sum_{k=1}^{\infty} \gamma(k) \cos \omega k \right]. \quad (22)$$

The Blackman-Tukey method of spectral estimation exploits this relationship to estimate the spectrum by way of the sample acf.

The normalized spectrum

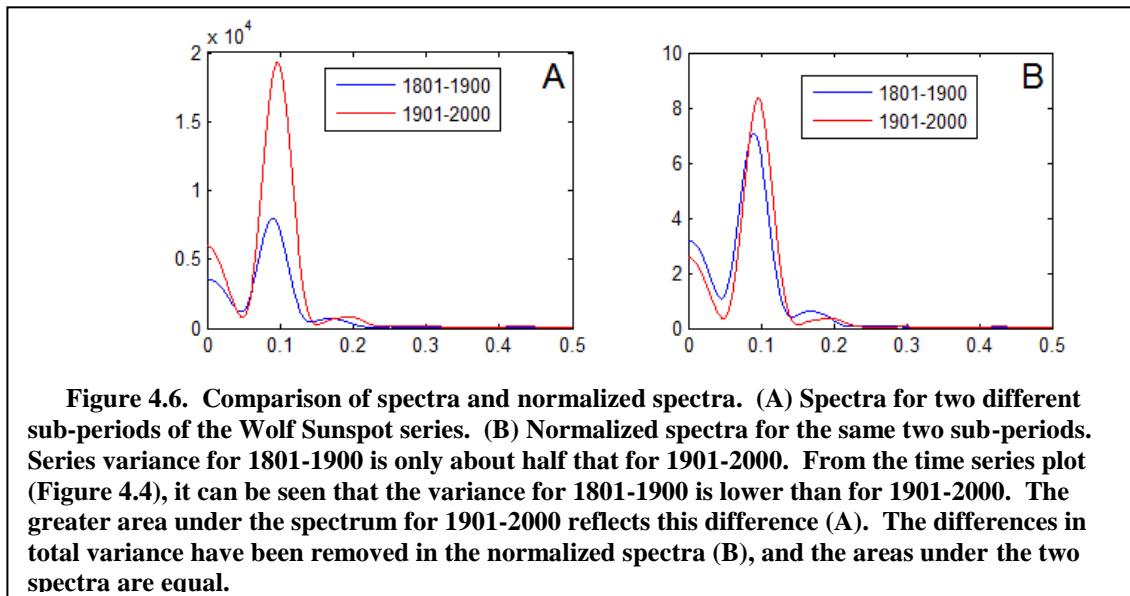
Because the area under the spectrum equals the variance, scaling the spectrum by dividing it by the variance σ_x^2 yields a plot for which the area under the curve is 1. The *normalized spectrum* is accordingly defined as

$$f^*(\omega) = f(\omega) / \sigma_x^2 \quad (23)$$

Recall that the acf is just the acvf divided by the variance. The normalized spectrum can therefore be written as the Fourier transform of the acf

$$f^*(\omega) = \frac{1}{\pi} \left[1 + 2 \sum_{k=1}^{\infty} \rho(k) \cos \omega k \right]. \quad (24)$$

The area beneath a portion of the normalized spectrum, $f^*(\omega)d\omega$, is the *proportion* of variance in the frequency range $(\omega, \omega + d\omega)$. Plots of the normalized spectrum are often preferable to plots of the spectrum for comparing spectral properties of time series with greatly differing variances, or different units of measurement. For example, a comparison of spectral properties of two different segments of the Wolf Sunspot series is distorted by gross differences in the total series variance for the sub-periods (Figure 4.6). In this case the normalized spectra show that the segments are really quite similar in the fraction of variance near wavelength 11 years.



4.5 Estimating the spectrum from data

The discussion above suggests that an “obvious” estimator for the spectrum is the Fourier transform of the complete sample acvf. Why not simply substitute the sample acvf for the population acf in equation (24) and have the summation run out to the maximum possible lag, which is lag $k = N - 1$? In fact this approach is one way of estimating the periodogram, which as we have seen is a spectrum with spectral estimates at the standard frequencies. The estimation of the periodogram in this way is described by Chatfield (2004). A problem with estimating the spectrum in this seemingly obvious way is that the estimator is not *consistent*, meaning that variance of the estimate does not decrease as the sample size N increases. One reason is that the method entails estimating N parameters from N observations, no matter how long the series is. Another problem is that the acvf at high lags is uncertain, and the method does not discount the higher lags. As a result, the spectrum (or periodogram) estimated in this way fluctuates wildly from one standard frequency to another and is extremely difficult to interpret. The Blackman-Tukey method circumvents this problem by applying the Fourier transform to a truncated, smoothed acvf rather than to the entire acvf.

The lag-window estimate of the spectrum

The Blackman-Tukey estimation method consists of taking a Fourier transform of the truncated sample acvf using a weighting procedure. Because the precision of the acvf estimates r_k decreases as lag k increases, it seems reasonable to give less weight to the values of the acvf at high lags. Such an estimator is given by

$$\hat{f}(\omega) = \frac{1}{\pi} \left\{ \lambda_0 c_0 + 2 \sum_{k=1}^M \lambda_k c_k \cos \omega k \right\} \quad (25)$$

where $\{\lambda_k\}$ are a set of weights called the *lag window*, and $M (< N)$ is called the *truncation point*. From equation (25) we see that the acvf at lags $M < k < N$ are no longer used, and that the acvf estimates at lower lags are weighted by a weighting function λ_k . Various weighting functions have been used to estimate the spectrum by equation (25); all have decreasing weight toward higher lags, such that the higher-lag acvf values are discounted. One popular form of lag window is the *Tukey window*.

Tukey window

The Tukey window, also called the Tukey-Hanning and Blackman-Tukey window, is given by

$$\lambda_k = 0.5 \left(1 + \cos \frac{\pi k}{M} \right) \quad k = 0, 1, \dots, M \quad (26)$$

where k is the lag, M is the width of the lag window – also called the *truncation point*, and λ_k is the weight at lag k . The Tukey window for a window-width of 30 lags is shown in Figure 4.7.

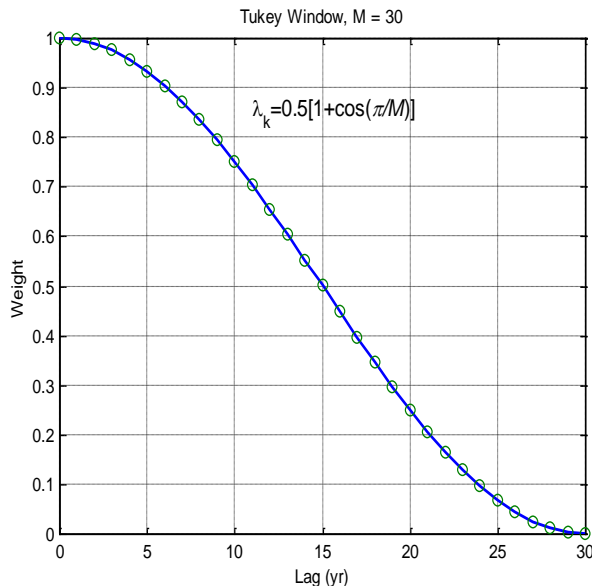


Figure 4.7. Weights in Tukey Window of width 30.

Notice that the weight decreases in the form of a bell-shaped curve from a maximum weight of 1 at lag 0 to a minimum weight of 0 at lag M . In estimating the spectrum by smoothing the acvf, you must choose the truncation point M . This is generally done by trial and error, with a subjective evaluation of which window best displays the important spectral features. The choice of M affects the *bias*, *variance*, and *bandwidth* of the spectral estimates.

Smaller $M \rightarrow$ increased bias

Bias refers to the tendency of spectral estimates to be less extreme (both highs and lows) than the true spectrum. Increased bias is manifested in a “flattening out” of the estimated spectrum, such that peaks are not as high as they should be, and troughs not as low. This bias is acknowledged, but not explicitly expressed as a function of M .

Smaller $M \rightarrow$ smaller variance of spectral estimates (narrower confidence bands)

Chatfield (2004) references Jenkins and Watts (1968, Section 6.4.2) for the general relationship between the truncation point, M , of the Tukey window and the variance of the spectral estimates. The relationship is that the quantity $\nu \hat{f}(\omega)/f(\omega)$ is approximately distributed as χ^2_ν , or chi squared with ν degrees of freedom, where for the Tukey window the degrees of freedom depends on the ratio of sample size to truncation point as follows

$$\nu = 2.67N/M \quad (27)$$

The relationship yields the following asymptotic $100(1 - \alpha)\%$ confidence interval for $f(\omega)$:

$$\frac{\nu \hat{f}(\omega)}{\chi^2_{\nu, \alpha/2}} \text{ to } \frac{\nu \hat{f}(\omega)}{\chi^2_{\nu, 1-\alpha/2}} \quad (28)$$

Let's look at a hypothetical example to illustrate the effect of changing M on the confidence interval. Say the time series has length 500 years, and you try two values of truncation point: $M_1 = N/5 = 100$ and $M_2 = N/10 = 50$. From equation (27), the corresponding values for degrees of freedom are

$$\nu_1 = 5 * 2.67 = 13.35 \quad (29)$$

and

$$\nu_2 = 10 * 2.67 = 26.7 \quad (30)$$

For $\nu_1 \cong 13$, associated with a choice of truncation point $M = 100$, the chi square values for the 95% confidence interval are

$$\chi^2_{13, 0.025} = 5.01 \quad , \quad \chi^2_{13, 0.975} = 24.73 \quad (31)$$

and the confidence interval is

$$\frac{13 \hat{f}(\omega)}{5.01} \text{ to } \frac{13 \hat{f}(\omega)}{24.73} \quad \text{or} \quad 2.59 \hat{f}(\omega) \text{ to } 0.52 \hat{f}(\omega) \quad (32)$$

For $\nu_2 \cong 27$, associated with a choice of truncation point $M = 50$, the chi square values for the 95% confidence interval are

$$\chi^2_{27, 0.025} = 14.57 \quad , \quad \chi^2_{27, 0.975} = 43.19 \quad (33)$$

and the confidence interval is

$$\frac{27 \hat{f}(\omega)}{14.57} \text{ to } \frac{27 \hat{f}(\omega)}{43.19} \quad \text{or} \quad 1.85 \hat{f}(\omega) \text{ to } 0.62 \hat{f}(\omega) \quad (34)$$

For a spectral estimate $\hat{f}(\omega)$, therefore, lowering the truncation point from $M = 100$ to $M = 50$ can be seen from (32) and (34) to result in a narrowing of the confidence interval around the spectral estimate.

Smaller $M \rightarrow$ increased bandwidth (decreased resolution of frequency of features)

Bandwidth roughly speaking is the width of the spectral window, which limits the resolution in frequency that spectral features can be identified. Spectral estimates apply to a *band* of frequency, and the width of that band is the bandwidth. If the bandwidth is wide, nearby peaks with similar frequencies might not be resolvable from one another. The bandwidth is a function of truncation point M , such that making M smaller broadens the bandwidth. For the Tukey window, Chatfield (2004) gives the relationship between bandwidth and truncation point is

$$bw = 8\pi/(3M) \quad \text{radians/yr} \quad (35)$$

or

$$bw = 4/(3M) \quad \text{cycles/yr} \quad (36)$$

By the relationship above, you can increase M to get a narrower bandwidth, and spectral features more finely resolved along the frequency axis. The trade-off, as shown in the previous section, is that increasing M also has the effect of widening the confidence interval around the spectral estimates.

General guidelines for choice of M

From the above, we see that the choice of truncation point M is important in getting the spectral estimates $\hat{f}(\omega)$. Chatfield (2004) offers some practical guidelines on the choice of M :

The choice of truncation point, M , is rather difficult and little clear-cut advice is available in the literature. It has to be chosen subjectively so as to balance ‘resolution’ against ‘variance’. The smaller the value of M , the smaller will be the variance of $\hat{f}(\omega)$ but the larger will be the bias (Neave, 1971). If M is too small, important features of $\hat{f}(\omega)$ may be smoothed out, while if M is too large the behaviour of $\hat{f}(\omega)$ becomes more like that of the periodogram with erratic variation. Thus a compromise value must be chosen, usually in the range $1/20 < M/N < 1/3$. For example if $100 < N < 200$, a value of M about $N/6$ may be appropriate, while if $1000 < N < 2000$, a value of M less than $N/10$ may be appropriate. The asymptotic situation we have in mind is that as $N \rightarrow \infty$, so does $M \rightarrow \infty$ but in such a way that $M/N \rightarrow 0$Jenkins and Watts (1968) suggest trying 3 different values of M . A low value will give an idea where the large peaks in $f(\omega)$ are, but the curve is likely to be too smooth. A high value is likely to produce a curve showing a large number of peaks, some of which may be spurious. A compromise can then be achieved with the third value of M . As Hannan (1970, p. 311) says, ‘experience is the real teacher and cannot be got from a book’.

A reasonable lag window can be achieved by viewing the spectrum while gradually increasing M . This approach results in a gradually narrowed bandwidth, and increasing detail in the estimated spectrum. The “window closing” approach is illustrated in spectral estimation for the sunspot series, 1700-2007 (Figure 4.8). Narrowing the bandwidth yields more detail in the spectrum, but with the tradeoff that the spectral estimate has greater variance, or uncertainty.

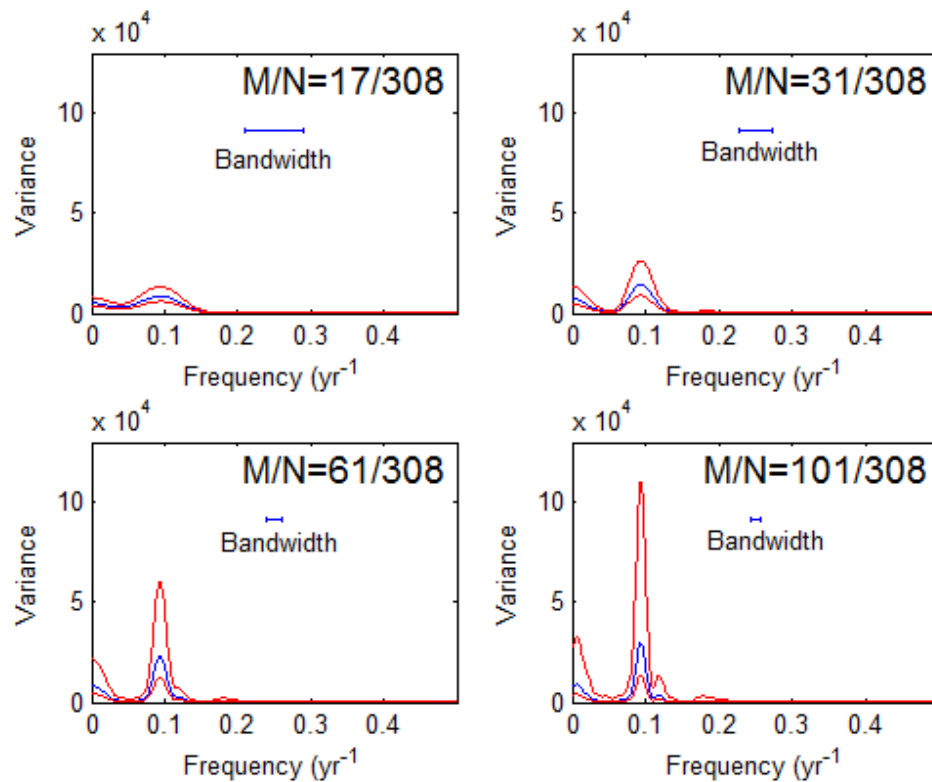


Figure 4.8. Illustration of “window closing” process in spectral estimation. Plots are spectra of 1700-2007 Wolf Sunspot Number using different setting of lag window M . With increasing M the following occurs: 1) bandwidth narrows, allowing greater frequency resolution of the main peak near 11 years, 2) spectral peak near 11 years becomes larger, and 3) confidence interval around spectrum widens.

Low-Frequency Spectrum

The broad shape of the estimated spectrum might indicate that variance tends to be higher at the low frequencies than at the high frequencies. This broad underlying shape is best seen in a much smoothed spectrum, which can be produced by selecting a small lag window, M , in the Blackman-Tukey method. For example, the upper left plot in Figure 4.8 clearly indicates that the sunspot series has relatively high variance at low frequencies. Such a spectrum is called a low-frequency spectrum. Conversely, if a spectrum tends to be higher at the high frequency than at the low frequencies the series has a high-frequency spectrum. Persistence, or positive autocorrelation, is characteristic of time series with a low-frequency spectrum.

Aliasing

Aliasing refers to the phenomenon in which spectral features on the frequency range $\{0, 0.5\}$ can be produced by variability at frequencies higher than the Nyquist frequency (i.e., higher than $f = 0.5$ cycles per year). Whether aliasing is a problem or not depends on the sampling interval and the frequencies of variability in the data, and is most easily illustrated for sampled rather than aggregated time series. For example, imagine a time series of air temperature

and a sampling interval of 18 hours. The Nyquist frequency corresponds to a wavelength of twice the sampling interval, or 36 hours. Air temperature has roughly a diurnal, or 24-hour, cycle – a cycle at a higher frequency than the Nyquist frequency. If the first sample happens to coincide with the time of daily peak temperature, the second sample – 18 hr later -- will be 6 hr before the peak on the second day, the third sample will be 12 hr before the peak on the third day, the fourth sample will be 18 hr before the peak on the fourth day, and the fifth sample will be 24 hr before the peak on the fifth day. This fifth sample is again at the daily peak. If the sampling is continued, the series would tend to peak at observations 1, 5, 9, 13, etc. The spacing between peaks is 4 sample points, or $4 \times 18 = 72$ hr. A spectrum of the series sampled in this way would have a spectral peak at wavelength 72 hrs. This is a false spectral peak, and is really the 24-hr cycle aliased to 72 hours.

An aliased cycle at a wavelength of 5 years (60 months) can be produced by sampling a time series of monthly mean-maximum temperature at intervals of 15 months (Figure 4.9). The monthly temperature series has a nature period of 12 months, and sampling every 12 months results in a series with no apparent cycle (Figure 4.9, top). But sampling every 15th month produces a strong cycle with a 5 year period (Figure 4.9, bottom). This bogus cycle could be predicted by considering that the lowest common denominator of the sampling interval (15 months) and the natural cycle (12 months) is 60 months.

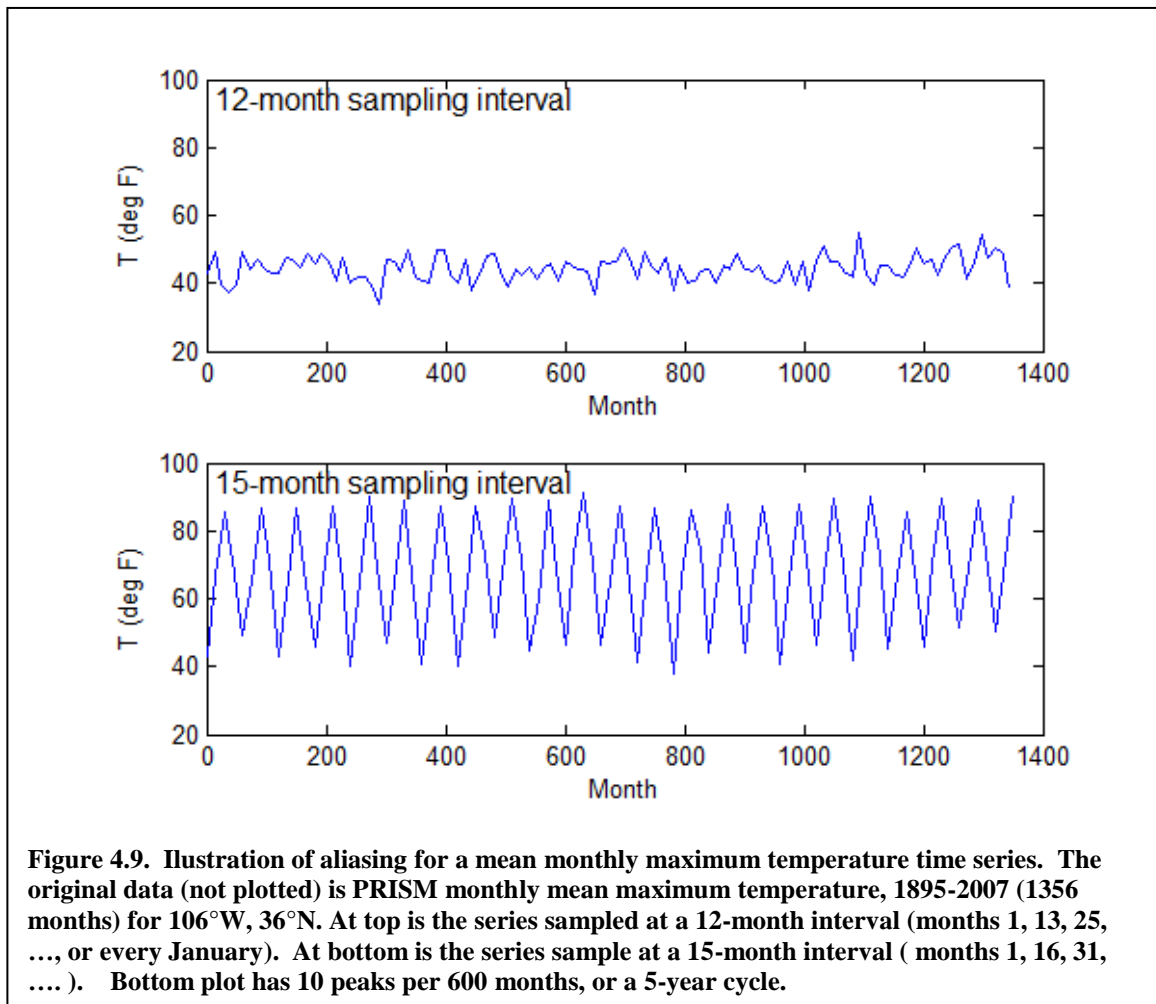


Figure 4.9. Illustration of aliasing for a mean monthly maximum temperature time series. The original data (not plotted) is PRISM monthly mean maximum temperature, 1895-2007 (1356 months) for 106°W, 36°N. At top is the series sampled at a 12-month interval (months 1, 13, 25, ..., or every January). At bottom is the series sample at a 15-month interval (months 1, 16, 31,). Bottom plot has 10 peaks per 600 months, or a 5-year cycle.

References

- Bloomfield, P., 1976, **Fourier analysis of time series: an introduction: John Wiley & Sons, p. 258 pp.**
[Smoothed periodogram method of spectral analysis]
- Chatfield, C., 2004, **The analysis of time series, an introduction, sixth edition: New York, Chapman & Hall/CRC. [A readable presentation of time series in frequency and time domains – a good introductory text]**
- Cook, E.R., Meko, D.M., and Stockton, C.W., 1997, **A new assessment of possible solar and lunar forcing of the bi-decadal drought rhythm in the western United States: Journal of Climate, v. 10, p. 1343-1356.**
- Granger, C.W.J., and Hatanaka, M., 1964, **Spectral analysis of economic time series: Princeton, New Jersey, Princeton University Press. [practical pointers for applied spectral analysis]**
- Hannan, E.J., 1970, **Multiple time series: New York, Wiley.**
- Jenkins, G.M., and Watts, D.G., 1968, **Spectral analysis and its applications: Holden-Day, 525 p.**
- LaMarche, V.C., 1974, **Frequency-dependent relationships between tree-ring series along an ecological gradient and some dendroclimatic implications, Tree-Ring Bulletin 34, 1-20.**
- LaMarche, V. C., and Fritts, H.C., 1972, **Tree-rings and sunspot numbers: Tree-Ring Bulletin, v. 32, p. 19-33.**
- Neave, H.R., 1971, **The exact error of spectral estimators: Ann. Math. Statist., v. 42, p. 961-975.**
- Panofsky, H.A., and Brier, G.W., 1958, **Some applications of statistics to meteorology: The Pennsylvania State University Press, 224 p. [Harmonic analysis; Climatology applications]**
- Percival, D.B., and Walden, A.T., 1993, **Spectral analysis for physical applications: Cambridge University Press. [Definition of spectrum in context of a sinusoidal model, eqns 1.6-1.16]**
- Salas, J.D., Delleur, J.W., Yevjevich, V.M., and Lane, W.L., 1980, **Applied modeling of hydrologic time series: Littleton, Colorado, Water Resources Publications, p. 484 pp. [Focuses on time domain, but a bit on spectral analysis of hydrologic series]**
- Wilks, D.S., 1995, **Statistical methods in the atmospheric sciences: Academic Press, 467 p. [Meteorology context]**