7 Detrending

Trend in a time series is a slow, gradual change in some property of the series over the whole interval under investigation. Trend is sometimes loosely defined as a long term change in the mean (Figure 7.1), but can also refer to change in other statistical properties. For example, tree-ring series of measured ring width frequently have a trend in variance as well as mean (Figure 7.2). In traditional time series analysis, a time series was decomposed into trend, seasonal or periodic components, and irregular fluctuations, and the various parts were studied separately. Modern analysis techniques frequently treat the series without such routine decomposition, but separate consideration of trend is still often required. *Detrending* is the statistical or mathematical operation of removing trend from the series. Detrending is often applied to remove a feature thought to distort or obscure the relationships of interest. In climatology, for example, a temperature trend due to urban warming might obscure a relationship between cloudiness and air temperature. Detrending is also sometimes used as a preprocessing step to prepare time series for analysis by methods that assume stationarity. Many alternative methods are available for detrending. Simple linear trend in mean can be removed by subtracting a least-squares-fit straight line. More complicated trends might require different procedures. For example, the cubic smoothing spline is commonly used in dendrochronology to fit and remove ring-width trend that might not be linear, or not even monotonically increasing or decreasing over time. In studying and removing trend, it is important to understand the effect of detrending on the spectral properties of the time series. This effect can be summarized by the *frequency response* of the detrending function.

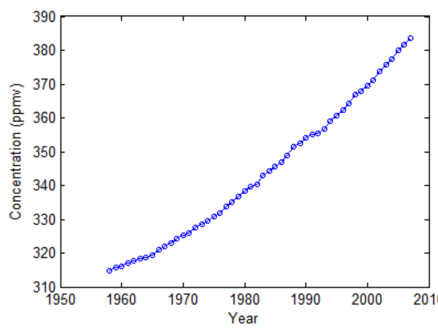

Figure 7.1. Trend in mean. A strong trend dominates the December atmospheric CO2 concentration at Mauna Loa, Hawaii, 1958-2007. Source: http://cdiac.ornl.gov/ftp/trends/co2/maunaloa.co2

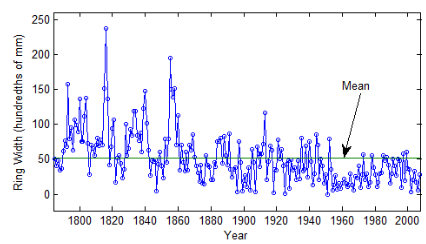

Figure 7.2. Trend in mean and variance. Ring widths from a Douglas-fir tree in Jemez Mountains, New Mexico, 1785-2007. Both mean ring width and variance of ring width decline with age of tree.

7.1 Identifying trend: a frequency-domain approach

Identification of trend in a time series is subjective because trend cannot be unequivocally distinguished from low frequency fluctuations. What looks like trend in a short time series segment often proves to be a low-frequency fluctuation – perhaps part of a cycle -- in the longer series. Sometimes knowledge of the physical system helps in identifying trend. For example, a decrease of ring width of a tree with time is expected partly on geometrical grounds: the annual increment of wood is being laid down on an ever-increasing circumference. If the volume of wood produced annually levels off as the tree ages, the ring width would be expected to decline. A hypothetical "age curve" in ring width can be computed assuming the cross-sectional area of wood added each year is constant (Figure 7.3). Such a conceptual model was used in dendrochronology as justification for "modified negative exponential" detrending (Fritts 1976).

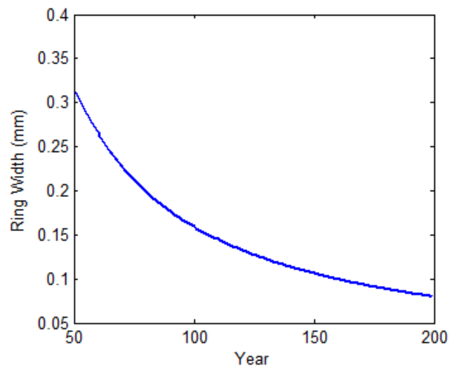

Figure 7.3. Mathematical form of expected ring width as function of time assuming constant increment of cross-sectional area.

If a physical basis is lacking, we need to rely on statistical methods to quantify trend. Statistical methods can help distinguish trend from other variations. The frequency domain is particularly useful here. Granger and Hatanaka (1964 p. 130) give some insight into spectral interpretation of trend. They conclude that we are unable to differentiate between a true trend and a very low frequency fluctuation, and give the following advice:

> It has been found useful by the author to consider as "trend" in a sample of size n all frequencies less than $1/(2n)$ as these will all be monotonic increasing if the phase is zero, but it must be emphasized that this is an arbitrary rule. It may also be noted that it is impossible to test whether a series is stationary or not, given only a finite sample as any apparent trend in mean *could* arise from an extremely low frequency.

If we apply the above reasoning to a 500-year tree-ring series, we would say that variation with period longer than twice the sample size, or 1000 years, should be regarded as trend. In another paper, Granger (1966) defines 'trend in mean' as comprising all frequency components whose wavelength exceeds the length of the observed time series. Cook et al. (1990) refer to Granger's (1966) "trend in mean" concept in giving suggestions for detrending tree-ring data:

> Given the above definition of trend in mean, another objective criterion for selecting the optimal frequency response of a digital filer is as follows. Select a 50% frequency-response cutoff in years for the filter that equals some large percentage of the series length, *n*. This is the %*n* criterion described in Cook (1985). The results of Cook (1985) suggest that the percentage is 67%*n* to 75%*n* based on using the cubic smoothing spline as a digital filter. The %*n* criterion ensures that little low-frequency variance, which is resolvable in the standardized tree rings, will be lost in estimating and removing the growth trend. This criterion also has a bias of sorts because of the stiff character of the low-pass filter estimates of the growth trend. It will not necessarily guarantee and, in fact, will rarely possess any kind of optimal goodness-of-fit.

7.2 Fitting the trend

Four alternative approaches to detrending are: 1) first differencing, 2) curve-fitting, 3) digital filtering and 4) piecewise polynomials. This section is weighted heavily toward the piecewise polynomials approach, which is widely used in dendrochronology.

First differencing. A time series that is non-stationary in mean (e.g., trend in mean) can be made stationary by taking the first difference. The first-difference is the difference of the value of the series at times t and $t-1$:

$$w_t = x_t - x_{t-1} \quad (1)$$

where x_t is the original time series and w_t is the first-differenced series. If the series is nonstationarity in not just the mean but in the rate of change of the mean (the slope), stationarity can be induced by taking the second difference, or the first difference of the first difference:

$$u_t = w_t - w_{t-1} \quad (2)$$

Higher orders of differencing can likewise be applied. First differencing has been applied in hydrology in the context of ARIMA (Autoregressive-Integrated-Moving-Average) modeling of streamflow series (Salas et al. 1980). As with any detrending method, first differencing can be expected to strongly attenuate the variance at the lowest frequencies in a time series. Salas et al. (1980) report that first differencing can be problematic in hydrology because it tends to introduce spurious high-frequency variation.

Anderson (1975) describes differencing as a way to remove nonstationarity from time series in general. According to Anderson (1975), each successive differencing will decrease the variance of the series, but at some point, higher-order differencing will lead to an increase in variance. When variance increases, the series has been *over-differenced*.

First-differencing of the Mauna Loa CO_2 time series does not effectively remove the trend (Figure 7.4). The first-differenced series is positive at all times, reflecting the accelerating rate of increase in the original CO2 curve (see Figure 7.1). Second-differencing appears to remove the trend. Interestingly, the variance increases with second-differencing, suggesting possible over-differencing. Standard deviations of the original, first-differenced and second-differenced series are 20.89, 0.60 and 0.68. The relative sizes of these standard deviations also attest to the overwhelming importance of trend to variance of the original time series.

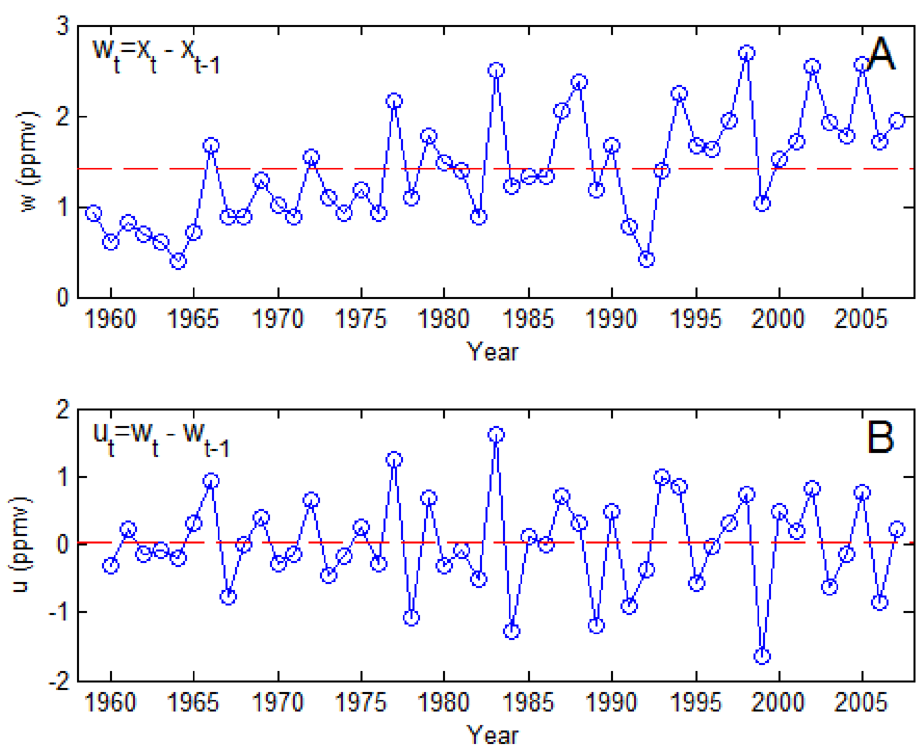

Figure 7.4. First-differencing and second-differencing to remove trend. (A) First-difference of December atmospheric CO2 concentration at Mauna Loa. (B) Second-difference of CO2 concentration. Original series, x_t, is plotted in Figure 7.1.

Conceptually, differencing can be problematic in that information on the level of the original data is missing from the differenced series. For example, a change from very wide tree-ring to moderately wide ring from one year to the next can produce the same first difference as a change from moderately narrow ring to very narrow ring. If precipitation is related directly to ring width, the difference series would be problematic for reconstruction of precipitation.

Curve-fitting. If a time series changes in level gradually over time, it makes sense to summarize trend by some simple function of time itself. A simple and widely used function of time is the least-squares-fit straight line, which assumes linear trend. Simple linear regression is used to fit the model

$$x_t = a + bt + e_t \tag{3}$$

where x_t is the original time series at time t, a is the regression constant, b is the regression coefficient, and e_t are the regression residuals. The trend is then described by

$$\hat{g}_t = \hat{a} + \hat{b}t \tag{4}$$

where $\hat{g}_t$ is the fitted trend, $\hat{a}$ is the estimated regression constant, and $\hat{b}$ is the estimated regression coefficient. The advantage of the straight-line method is simplicity.

The straight line may unrealistic, however, in restricting the functional form of the trend. Other functions of t (e.g., quadratic) might be better depending on the type of data. Sometimes the mathematical form of the trend function has physical basis. For example, a modified negative exponential curve with conceptual basis in the change of tree geometry with time has been used to remove the “age trend” from ring-width series (Fritts 1976). The modified exponential follows the equation

$$\hat{g}_t = \hat{a}e^{-\hat{b}t} + \hat{k} \tag{5}$$

where the coefficients, $\hat{a}, \hat{b},$ and $\hat{k}$ are estimated such that the sum of square of differences of the smooth curve $\hat{g}_t$ and the original time series is minimized.

Ring-width trend alternatively described by a straight line and a modified negative exponential is illustrated in Figure 7.5. The curvature in the time plot of ring width is so slight that the choice of curve makes little difference for this example. At the recent end of the time series, however, the expected ring width according to trend is about 30 percent higher for the modified exponential than for the straight line.

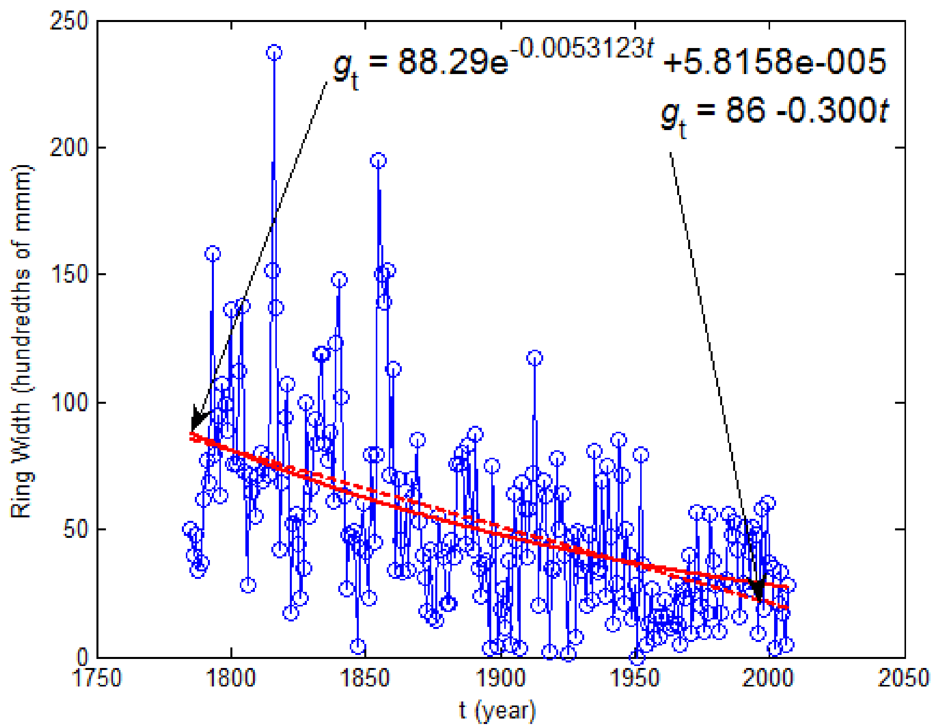

Figure 7.5. Trend in measured tree-ring width fit by straight line and modified negative exponential curve. Tree-ring series is same series described in caption to Figure 7.2. Time variable (year) was shifted to being with 1 before estimating parameters.

Digital filtering. Another procedure for dealing with trend is to describe the trend as a linearly filtered version of the original series. The original series is converted to a smooth "trend line" by weighting the individual observations

$$g_t = \sum_{r=-q}^{s} a_r x_{t+r} \tag{6}$$

where x_t is the original series, $\{a_r\}$ is a set of filter weights (summing to 1), and g_t is the smooth trend line. The weights are often symmetric, with $s = q$ and $a_j = a_{-j}$. If the weights are all equal, the filter is a simple moving average, which generally is not recommended for measuring trend (Chatfield 1975). Preferable is a symmetric filter with weights decreasing from the central weight.

Piecewise polynomials. An alternative to fitting a curve to the entire time series (curve fitting) is to fit polynomials of time to different parts of the time series. Polynomials used this way are called piecewise polynomials. The *cubic smoothing spline* is a piecewise polynomial of time, *t*, with the following properties:

- **The polynomial is cubic (t raised to third power)**
- **A separate polynomial is fit to every sequence of three points in the series**
- **The first and second derivatives are continuous at each point**

- **The “spline parameter” specifies the flexibility and depends on the relative importance given to “smoothness” of the fitted curve, and “closeness of fit”, or how close the fitted curve passes to the individual data points**

Given the approximate values $y_i = g(x_i) + \varepsilon_i$ of some supposedly smooth function g at data points $x_1, \ldots, x_N$ and an estimate δy_i of the standard deviation of y_i, the problem is to recover the smooth function from the data. Let $s(x_i)$ be the spline curve, or the approximation to the smooth function g. Following De Boor (1978, p. 235), the spline curve is derived by minimizing the quantity

$$p \sum_{i=1}^{N} \left[\frac{y_i - s(x_i)}{\delta y_i} \right]^2 + (1-p) \int_{x_1}^{x_n} \left[\mathrm{D}^2 s \right]^2 \tag{7}$$

over all functions s for a given *spline parameter*, p, where $\mathrm{D}^2 s$ refers to the second derivative of s with respect to time. The first term in (7) is similar to a sum-of-squares of deviations. The second term integrates curvature contributions (second derivative). Minimizing establishes a compromise between staying close to the given data (first term) and obtaining the smoothest possible curve (second term). The choice of p, where p can range from 0 to 1, depends on which of those two goals is given the greater importance. For $p = 0$, s is the least squares straight-line fit to the data. At the other extreme, $p = 1$, s is the cubic spline interpolant, and passes through each data point. As p ranges from 0 to 1, the smoothing spline changes from one extreme to the other (Figure 7.6). The term δy_i allows for differential weighting of data points. Following recommendations of Cook and Peters (1981) we use the default MATLAB© weighting (1 for all points).

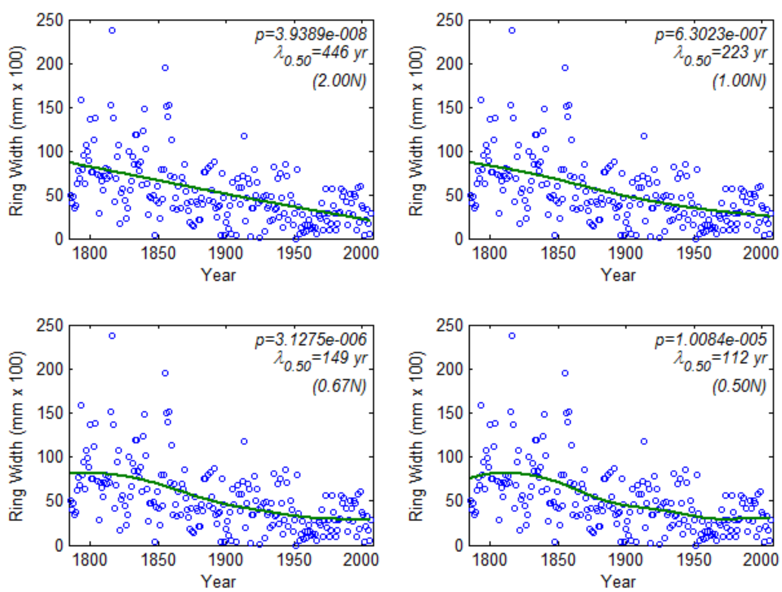

Figure 7.6. Cubic smoothing splines of differing stiffness fit to a time series of tree-ring ring width. Spline parameter is *p*. Spline has a frequency response of 0.50 at wavelength $\lambda_{0.50}$, which is expressed as a decimal fraction of the series length, *N*, in parentheses. Spline *p* for upper left plot is small enough that straight line is approximated.

7.3 Frequency response

The frequency response function describes how a linear system responds to sinusoidal inputs at different frequencies (Chatfield 2004, p. 198). The frequency response function has two components -- the *gain* and the *phase.* The gain at a given frequency describes how the amplitude of a sinusoid at that frequency is damped or amplified by the system. The phase describes how a wave at that frequency is shifted in absolute time.

In reference to a spline curve, the phase is zero, and the "frequency response" merely describes the gain, or the amplitude, of the response function. The input to the "system" in this case is the original time series; the output is the smoothed curve purported to represent the trend. The frequency response measures how strongly the spline curve responds to or tracks a periodic component of a given frequency, should the time series have such a component. The amplitude of frequency response at a given frequency is the ratio of the amplitude of the sinusoidal component in the smoothed series (the spline curve) to the amplitude in the original time series.

Relation of frequency response to spline parameter

The cubic smoothing spline has become increasingly popular as a detrending method in dendrochronology because the spline is adaptable and easily applied to a wide range of types of "age trend" or "growth trend" found in tree-ring data. Application of the spline to dendrochronology was first proposed by Cook and Peters (1981), who derived a mathematical relationship between a spline parameter and the frequency response of the spline. Jean-Luc Dupouey (INRA, Forest Ecology and Ecophysiology Unit, Champenoux, France, has pointed out (personal communication) that the spline parameter p is defined somewhat differently in MATLA B© Spline Toolbox than in Cook and Peters (1981), and has provided equations that give the correct relationship for use with MATLAB©.

In terms of the parameter p in equation (7), the frequency response of the spline is given by

$$u(f) = \frac{1}{\left[1 + 12\left(\frac{1-p}{p}\right)\frac{\left(\cos 2\pi f - 1\right)^2}{\left(\cos 2\pi f + 2\right)}\right]} \qquad (8)$$

where $u(f)$ is the amplitude of frequency response at frequency f, and p is the spline parameter as defined earlier. A plot of $u(f)$ against f shows the relative response of the spline to hypothetical input variations at different frequencies. For a smoothing spline, this response is higher toward the low-frequency end of the spectrum (Figure 7.7) .

Equation (8)can be rearranged with p on the left-hand side to get the parameter corresponding to a spline that has a desired amplitude of frequency response at a specified frequency:

$$p_{u_0(f_0)} = \frac{1}{\left\{\left[\left(\frac{1-u_0(f_0)}{u_0(f_0)}\right)\left(\frac{\left(\cos 2\pi f_0 + 2\right)}{12\left(\cos 2\pi f_0 - 1\right)^2}\right)\right] + 1\right\}} \qquad (9)$$

where f_0 is the target frequency, $u_0(f_0)$ is the desired amplitude of response at that frequency, and $p_{u(f_0)}$ is the corresponding spline parameter. Cook and Peters (1981) define an "*n*-year spline" as the spline whose frequency response is 50%, or 0.50, at a wavelength of n years.

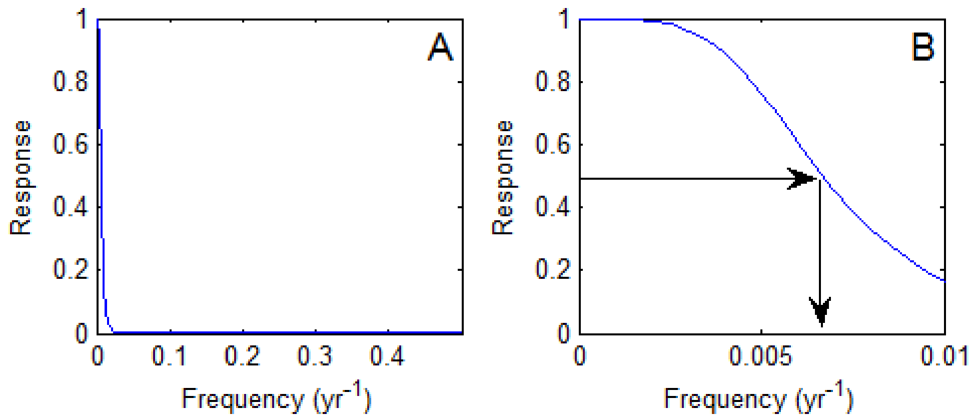

Figure 7.7. Frequency response function of a cubic smoothing spline with spline parameter *p*=3.1257E-6 as applicable to an annual time series. (A) Full frequency response (B) Plot zoomed to frequency range 0 to 0.01 (periods ∞ to 100 years). Response rises above 0.2 at a wavelength slightly longer than 100 years, and reaches 0.5 at a wavelength of about 149 years (frequency of 0.0067). Plot of full response emphasizes that this spline tracks only low frequencies.

Equation (9)can be used to compute the required spline parameter for the "n-year spline" by setting $u_0\left(f_{0)}\right)=0.50$. For example, substitution into equation (9)for the "100-year" spline yields:

$$p=\frac{1}{\left\{\left[\left(\frac{0.5}{0.5}\right)\left(\frac{(\cos 2\pi f_0+2)}{12(\cos 2\pi f_0-1)^2}\right)\right]+1\right\}}$$

$$=\frac{1}{\left\{\left[\left(\frac{(\cos(2\pi[1/100])+2)}{12(\cos(2\pi[1/100])-1)^2}\right)\right]+1\right\}}$$

$$=\frac{1}{\left\{\left[\left(\frac{(0.99802672842827+2)}{12(0.99802672842827-1)^2}\right)\right]+1\right\}}$$

$$=1.5585\text{e-}005$$

Function csaps in the Spline Toolbox of MATLAB© can be used generate the spline-smoothed curve for a given input time series and spline parameter (middle curve, Figure 7.8). It is important to keep in mind that the spline parameter in equation (9) will not give the desired spline smoothness if applied in tree-ring standardization program ARSTAN; for that application, the correct versions of the equations for the spline parameter and frequency response are those in Cook and Peters (1981).

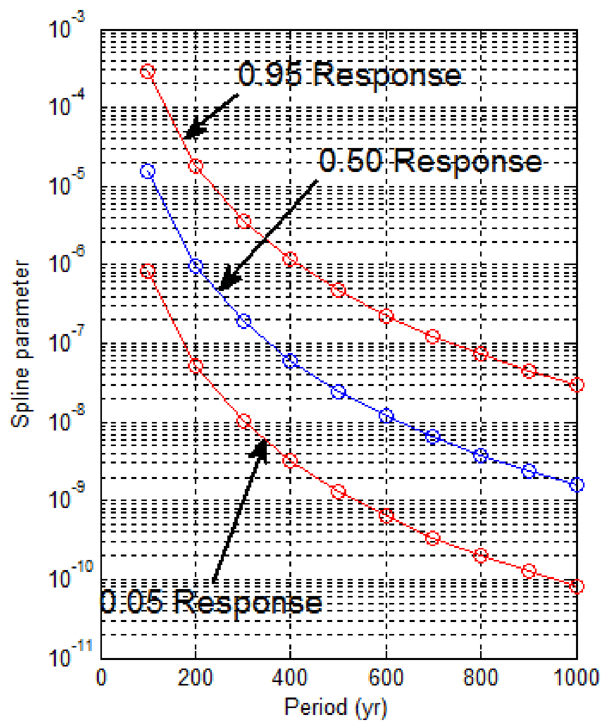

Figure 7.8. Spline parameter as a function of wavelength of frequency response. Curves shown correspond to amplitude of frequency response 0.05, 0.50 and 0.95. For example, a spline with parameter p=2E-07 has a frequency response of 0.50 at wavelength 300 years, a response of 0.95 at wavelength slightly longer than 600 years, and response of 0.05 at wavelength of about 150 years. Such a spline strongly tracks variations with wavelength longer than 600 yr and hardly tracks at all variations with wavelength shorter than 150 yr. If the original time series were a sine wave with wavelength 300 years, the smooth curve would be a damped sine wave with 25 percent (0.50 squared) the variance of the original series. The detrended series—the difference of the original series and smooth curve-- would have 75 percent of the original series' variance.

7.4 Removal of trend

Once a trend line has been fit to the data, we can regard that line as representing the "trend." The question remains, how to remove the trend? If the trend-identification method has identified a trend line, two options are available. First is to subtract the value of the trend line from the original data, giving a time series of residuals from the trend. This "difference" option is attractive for simplicity, and for giving a convenient breakdown of the variance: the residual series is in the same units as the original series, and the total sum of squares of the original data can be expressed as the trend sum-of-squares plus the residual sum-of-squares.

The "ratio" option is attractive for some kinds of data because the ratio is dimensionless, and the ratio operation tends to remove trend in variance that might accompany trend in mean. Tree-ring width is one such data type: variance of ring width tends to be high when mean ring width is high, and low when mean ring width is low. Ratio-detrending generally is feasible for non-negative time series only, and runs the risk of explosion of the detrended series to very high values if the fitted trend line approaches zero (Cook and Peters 1997).

7.5 Effect of detrending on spectrum

Detrending has the same effect on the spectrum as high-pass filtering. That is, the variance at low frequencies is diminished relative to variance at high frequencies. The frequency response of the spline is high for those frequencies *tracked* closely by the spline. In the subsequent removal of the trend line, these frequencies are mostly *removed.* Frequencies at which the frequency response of the spline is high are therefore those at which the spectrum of the detrended series is low. In general, at the lowest frequencies, the spectrum of the detrended series will be diminished relative to the spectrum of the original data. The more *flexible* the spline, the higher the frequency-range affected by the detrending.

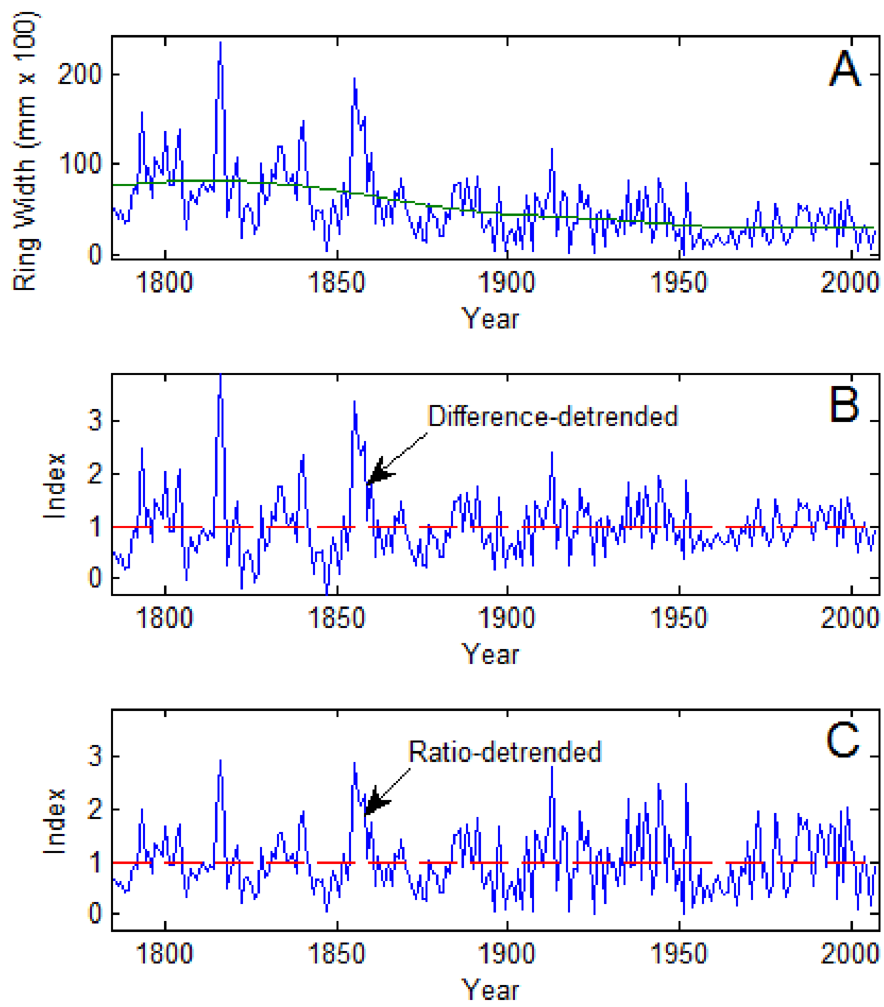

Figure 7.9. Contrast of time series detrended as difference and ratio. (A) Ring-width time series and fitted spline with 0.50 frequency response at 70 percent of series length. (B) Index computed as difference of original time series and fitted trend line. (C) Index computed as ratio of original time series to fitted trend line. Ratio-detrended series converted to same overall mean and variance as difference-detrended series before plotting. Time trend of variance evident in original ring width and difference-detrended series is not evident in ratio-detrended series. Data: ring-width from New Mexico, USA, described in caption to Figure 7.2.

Normalized spectrum. In comparing spectra of series detrended by different methods, it is convenient to plot the spectra on the same pair of axes. To eliminate spectrum differences due differences in total series variance, the spectra are best plotted as *normalized spectra*. The normalized spectrum is the spectrum standardized to have a unit area of 1.0[1], and can be computed by dividing each spectrum ordinate by the area under the spectrum (Figure 7.10). A

[1] Depending on plotting convention, the area under a plotted normalized spectrum may appear to differ from 1.0. For example, areas under the normalized spectra in Figure 7.10 are 1.0 only if the frequency axis is scaled such that the range {0 0.5} is scaled to {0 1}. The important point is that the total area under the normalized spectrum is the same regardless of total variance of the time series.

normalized spectrum can also be arrived at by first converting a time series to "Z-scores" (zero mean, unit standard deviation) before spectral estimation. (Recall that the area under the spectrum equals or is proportional to the variance.)

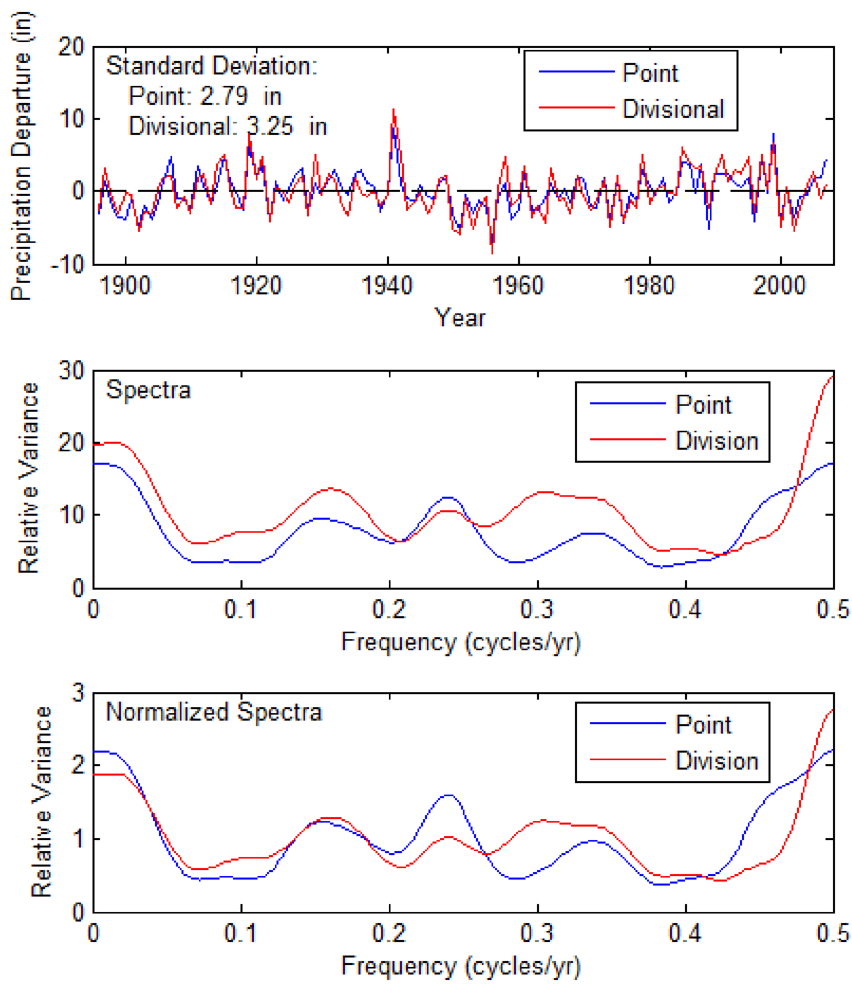

Figure 7.10. Normalized spectrum as tool for comparing spectral features in two time series. (A) Time plots of annual (Oct-Sept) point precipitation and divisional precipitation for northern new Mexico. (B) Spectra of the two precipitation series. (C) Normalized spectra. Without normalization, the spectra might be miss-interpreted as showing higher low-frequency variance in the divisional series. The vertical offset in spectra in the middle plot merely reflects the differing total variances of the two series. Total areas under the two normalized spectra are equal.

7.6 Quantifying the importance of trend

A simple measure of the practical importance of trend in a time series is the fraction of original variance of the series accounted for by the fitted trend line, which can be computed by

$$R^2 = 1 - \frac{\mathrm{var}(e_t)}{\mathrm{var}(x_t)} \qquad (10)$$

where $\text{var}(x_t)$ is the variance of the original time series, and $\text{var}(e_t)$ is the variance of the residuals from the trend line.

Equation (10)measures the importance of the tend component in a time series time series, and can range from 0 for no importance to 1 if the series is pure trend (Figure 7.10). Note that for ratio detrending, the total variance of the original series cannot be decomposed into variance due to trend and residual variance because the detrended series is NOT a residual.

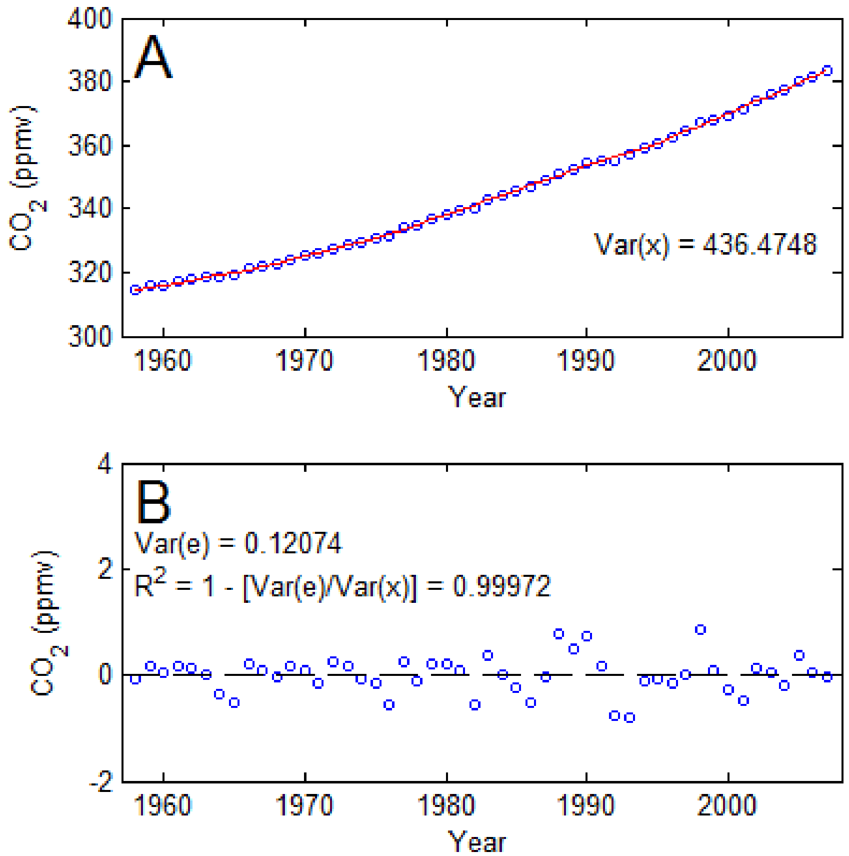

Figure 7.10. Variance due to trend in a trend-dominated series. (A) December free-air CO2 concentration at Mauna Loa, fit with cubic smoothing line to describe monotonic trend. (B) Residuals from the fitted trend line. While original time series varies over a range of more than 60 parts per million by volume (ppmv), the residuals have a range of less than 2 ppmv. The percentage of variance due to estimated trend according to equation (10) exceeds 99.9 percent.

7.7 References

Anderson, O., 1976, Time series analysis and forecasting: the Box-Jenkins approach: London, Butterworths, p. 182 pp.

Chatfield, C., 2004, The analysis of time series, an introduction, sixth edition: New York, Chapman & Hall/CRC.

Cook, E.R., 1985, A time series approach to tree-ring standardization, Ph. D. Diss., Tucson, University of Arizona.

-----, Briffa, K., Shiyatov, S., and Mazepa, V., 1990, Tree-ring standardization and growth-trend estimation, *in* Cook, E.R., and Kairiukstis, L.A., eds., Methods of Dendrochronology, Applications in the Environmental Sciences: Kluwer Academic Publishers, p. 104-123.

-----, and Peters, K., 1981, The smoothing spline: A new approach to standardizing forest interior tree-ring width series for dendroclimatic studies, Tree-Ring Bulletin 41, 45-53.

-----, and Peters, K., 1997, Calculating unbiased tree-ring indices for the study of climatic and environmental change: The Holocene, v. 7, no. 3, p. 361-370.

de Boor, C., 1978, A practical guide to splines: New York, Springer-Verlag, 392 p.

-----, 1999, Spline toolbox for use with MATLAB©, user's guide, version 2: Natick, MA, The MathWorks, Inc.

Fritts, H.C., 1976, Tree rings and climate: London, Academic Press, 567 p.

Granger, C.W.J., 1966, On the typical shape of an econometric variable: Econometrics, v. 34, p. 151-160.

Granger, C.W.J., and Hatanaka, M., 1964, Spectral analysis of economic time series: Princeton, New Jersey, Princeton University Press.

Panofsky, H.A., and Brier, G.W., 1958, Some applications of statistics to meteorology: The Pennsylvania State University Press, 224 p.

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.

World Meterorological Organization, 1966, Technical Note No. 79: Climatic Change, WMO-No, 195.TP.100, Geneva, 80 pp.