Methods of Dendrochronology

Applications in the Environmental Sciences

Edited by

E. R. Cook

Tree-Ring Laboratory, Lamont-Doherty Geological Observatory, Columbia University, New York, U.S.A.

and

L. A. Kairiukstis

IIASA, Laxenburg, Austria and Lithuanian Academy of Sciences, U.S.S.R.

Springer-Science+Business Media, B.V.

Library of Congress Cataloging in Publication Data

Methods of dendrochronology : applications in the environmental science / E. Cook and L. Kairiukstis, editors.
p. cm.
Includes bibliographical references.

1. Dendrochronology. 2. Ecology. I. Cook, E. (Edward) II. Kairikūkštis, Leonardas.
QK477.2.A6M48 1989
582.16'0372--dc20 89-48124

ISBN 978-90-481-4060-2 ISBN 978-94-015-7879-0 (eBook)
DOI 10.1007/978-94-015-7879-0

Reprinted 1992

Printed on acid-free paper

series and of the inadequacy of the negative exponential curve as a model for the observed growth trend. In such cases, the outer portion of a ring-width series may be systematically underfit or overfit for decades. Holmes *et al.* (1986) also showed that a stiff spline fit alone to the same series by the 67%n criterion was not sufficiently flexible to track the sharp curvature of the juvenile portion of the growth trend, even though it was quite adequate for the mature phase of the trend. Since each method of detrending was better for different portions of the growth trend, Holmes *et al.* (1986) reasoned that the sequential use of both techniques would correct the deficiencies of each method.

Cook (1985) examined the spectral properties of double-detrending and found that linear or negative exponential detrending followed by 67%n spline detrending worked well without removing too much low-frequency variance.

3.3.5. Concluding remarks on growth-trend estimation

The estimation and removal of growth trends from tree-ring series should be based, as much as possible, on the intended application of the tree-ring data. This means that there should be an *a priori* expectation of what the signal of interest is in the ring-width measurements. Given this expectation, the method of detrending should be chosen that will reduce the low-frequency noise not associated with that signal. It is difficult to accomplish this task within an objective framework because of the uncertainty distinguishing signal from noise in a given ensemble of ring-width data. Inevitably, some assumptions must be made that may have a great effect on the final standardized tree-ring chronology. It is imperative that these assumptions are carefully considered and justified in any application of standardized tree rings.

In general, stochastic methods are preferable to deterministic methods because of the freedom that the former possess in fitting the behavior of ring widths as they are observed, not as theory would have them behave. However, the consequence of this added flexibility are the problems of *ad hoc* model selection and overfitting, which are more likely to occur for stochastic models than for deterministic models. There also seems to be some utility in using a hybrid double-detrending approach, which can compensate for local lack-of-fit problems of single detrending methods.

3.4. Estimation of the Mean Chronology

E. Cook, S. Shiyatov, and V. Mazepa

3.4.1. Introduction

Once a collection of ring-width series has been detrended and indexed into a new ensemble of tree-ring indices, the estimation of the common signal, C_t, can proceed. As mentioned earlier, tree-ring indices can be treated as stationary, stochastic processes that allow them to be treated as a collective ensemble of realizations containing both a common signal in the form of C_t (and perhaps $D2_t$) and individual signals unique to the series ($D1_t$ and E_t).

3.4.2. Methods of computing the mean-value function

Three methods will be described that have been used in tree-ring studies: the arithmetic mean, the biweight robust mean that discounts outliers, and a mean based on testing for a mixture of normal distributions in the sample.

The Arithmetic Mean

The classical method of estimating C_t is by averaging the ensemble of detrended tree-ring indices across series for each year using the arithmetic mean (Fritts, 1976). This produces a time series mean-value function that concentrates the signal (C_t) and averages out the noise ($D1_t$ and E_t). The arithmetic mean of m indices available in year t is computed as

$$\mathbf{\bar{I}}_t = \sum_{j=1}^{m} I_t/m \quad . \tag{3.21}$$

This is an estimate of the signal in the tree-ring indices for year t. The variance or spread of the frequency distribution of m indices about the mean is computed as

$$\mathbf{S}_t^2 = \sum_{j=1}^{m} (I_t - \mathbf{\bar{I}}_t)^2/(m-1) \quad . \tag{3.22}$$

The square root of $\mathbf{S}_t^2$ is the standard deviation of m indices for year t. And, the variance of the mean, which is a measure of the noise or uncertainty in the estimation of the mean, is computed as

$$\mathbf{S}_{\mathbf{\bar{I}}}^2 = \mathbf{S}_t^2/m \quad . \tag{3.23}$$

The square root of equation (3.23) is the standard error of the mean. These statistics are described in virtually all basic statistics textbooks and in Fritts (1976). A measure of the strength of the resultant estimate of C_t is the signal-to-noise ratio (SNR) (Wigley *et al.*, 1984; Briffa *et al.*, 1987), an estimate of which is given in equation (3.15). The SNR can only provide information about the quality of the observed signal, but says nothing about its relationship to the expected signal, which is not known at this stage (Briffa *et al.*, 1987).

The Biweight Robust Mean

Other more involved methods are available for computing the mean-value function. If there are suspected outliers, or extreme values, in the tree-ring indices, than a robust mean such as the biweight mean (Mosteller and Tukey, 1977) can be used in place of the arithmetic mean. When outliers are present, the arithmetic mean is no longer a minimum variance estimate of the population mean, and it is not guaranteed to be unbiased. In contrast, robust means automatically discount the influence of outliers in the computation of the mean and, thus, reduce the variance and bias caused by the outliers. The use of a robust mean tacitly admits the likelihood of contamination by endogenous disturbance effects and other sources of noise having long-tailed, not normally distributed properties. Endogenous disturbance effects are likely to act as outliers because, as defined earlier, endogenous disturbances tend to behave as random events in space and time. The biweight mean for year t is computed by iteration as

$$\bar{\mathbf{I}}_t^* = \sum_{j=1}^{m} w_t I_t \quad , \tag{3.24}$$

where

$$w_t = \left[1 - \left(\frac{I_t - \bar{\mathbf{I}}_t^*}{c\mathbf{S}_t^*}\right)^2\right]^2 \quad ,$$

when

$$\left(\frac{I_t - \bar{\mathbf{I}}_t^*}{c\mathbf{S}_t^*}\right)^2 < 1 \quad ,$$

otherwise 0. The weight function, w_t, is symmetric, and, therefore, unbiased in its estimation of central tendency when the data are symmetrically distributed (Cook, 1985). $\mathbf{S}_t^*$ is a robust measure of the standard deviation of the frequency distribution, which will be the median absolute deviation (MAD)

$$\mathbf{S}_t^* = \text{median}\ \{|I_t - \bar{\mathbf{I}}_t^*|\} \quad , \tag{3.25}$$

and c is a constant, often taken as six or nine (Mosteller and Tukey, 1977). The constant c determines the point at which a discordant value is given a weight of zero. When this is the case, the outlier is totally discounted in computing the mean and, thus, has no influence on the estimation of the mean index. A constant c equal to 9 was used by Cook (1985) in developing a new tree-ring standardization procedure. From Mosteller and Tukey (1977), c equal to 9 is equivalent to totally rejecting any value exceeding ± 6 standard deviations from the mean, as estimated by equations (3.25) and (3.24), respectively. To start the iteration for computing the final $\bar{\mathbf{I}}_t^*$, the arithmetic mean or median can be used as an initial estimate. Ordinarily, only three to four iterations are needed to converge on an estimate of $\bar{\mathbf{I}}_t^*$ that does not change by more than 10^{-3}. A robust estimate of the variance, analogous to equation (3.22), is also available for the biweight mean as

$$mS_t{}^{*2} = \frac{m\Sigma'(I_t - \bar{\mathbf{I}}_t{}^*)^2 (1 - u_t{}^2)^4}{[\Sigma'(1 - u_t{}^2)(1 - 5u_t{}^2)]\,[-1 + \Sigma'(1 - u_t{}^2)(1 - 5u_t{}^2)]} \,, \qquad (3.26)$$

where $u_t = (I_t - \bar{\mathbf{I}}_t{}^*) \;/\; 9(\mathrm{MAD})$ and Σ' indicates summation for $u_t{}^2 \leq 1$, only (Mosteller and Tukey, 1977). Similarly, a robust estimate of the variance of the biweight mean, analogous to equation (3.23), is also available as

$$S_{\bar{\mathbf{I}}_t}{}^{*2} = \frac{\Sigma'(I_t - \bar{\mathbf{I}}_t{}^*)^2 (1 - u_t{}^2)^4}{[\Sigma'(1 - u_t{}^2)(1 - 5u_t{}^2)][-1 + \Sigma'(1 - u_t{}^2)(1 - 5u_t{}^2)]} \,, \qquad (3.27)$$

where u_t and Σ' are defined as above.

Aside from its added computational complexity, the biweight mean has a potential cost or premium associated with it. When the sample of indices is devoid of outliers and approximates a Gaussian distribution, the variance of the biweight mean will be greater than that of the arithmetic mean. That is,

$$S_{\bar{\mathbf{I}}_t}{}^{*2} > S_{\bar{\mathbf{I}}_t}{}^{2} \;.$$

This means that the biweight mean is less efficient in estimating the common signal when the assumed presence of outliers is false. However, when outliers are present in the sample, the variance of the biweight mean will be less than that of the arithmetic mean. That is,

$$S_{\bar{\mathbf{I}}_t}{}^{*2} < S_{\bar{\mathbf{I}}_t}{}^{2} \;.$$

In this case, the biweight mean is more efficient at estimating the common signal. A measure of statistical efficiency of the biweight mean relative to the arithmetic mean is (Mosteller and Tukey, 1977)

$$\mathrm{Efficiency} = S_{\bar{\mathbf{I}}_t}{}^{2} \,/\, S_{\bar{\mathbf{I}}_t}{}^{*2} \,, \qquad (3.28)$$

which is the ratio of the lowest variance feasible, under the Gaussian assumption, to the actual variance of the biweight mean. Under the Gaussian assumption and for moderate to large sample sizes (say $m > 10$), the efficiency of the biweight mean exceeds 90% of the arithmetic mean. This is a small premium to pay for protection from outliers and is very difficult to see in practice. When the sample size drops below 6, the simpler median can replace the biweight as the robust mean (Cook, 1985).

Extensive use of the biweight mean on closed-canopy forest tree-ring data (Cook, 1985) revealed that approximately 45% of the yearly means of 66 tree-ring chronologies showed some reduction in error variance using the biweight mean. This resulted in an average error variance reduction of about 20% in the robust mean-value functions compared with those based on the arithmetic mean. These results reveal the high level of outlier contamination in closed-canopy forest tree-ring data that can corrupt the estimated common signal if left unattended.

Mean-Values from a Mixture of Normal Distributions

Shiyatov and Mazepa (1987) and Mazepa (1982) describe another method of computing the mean-value function, which is based on examining the frequency distribution of the individual indices for each year. If the distribution is symmetrical and unimodal, the arithmetic mean is computed. However, if the distribution appears to be bimodal or multimodal, then the distribution is tested for a mixture of normal distributions. If mixed normal distributions are detected in the sample, then the mode of the grouping of largest indices is used as the best estimate of central tendency for that year. The selection of the grouping of largest indices is based on the notion that non-climatic effects, such as fruiting, will cause the ring widths of the affected trees to be narrower than the ring widths of the unaffected trees for the same years (Danilov, 1953; Kolischuk *et al.*, 1975). Therefore, the grouping of largest indices should more faithfully record the influence of climate.

In addition, the Law of Limiting Factors suggests that these anomalous effects are more likely to be seen when climate is less limiting to growth in a given year. That is, any expression of bimodality or multimodality in the sample distribution of tree-ring indices in a given year is more likely to be found when growth is not severely limited by climate. *Figure 3.9* shows six frequency histograms of tree-ring indices derived from Siberian larch (*Larix sibirica*) growing in the Ural Mountains, USSR. Four of the six examples show a clear indication of bimodality or multimodality, which tends to increase as the mean increases. The dispersion of the histograms increases as the mean level increases. This reflects the typical positive correlation between the means and standard deviations of tree-ring indices, which has been described by Fritts (1976).

Figure 3.10 provides an illustration of the proposed technique of Shiyatov and Mazepa (1987). Because of the limited sample size for each year (usually < 30), Shiyatov and Mazepa (1987) restricted the test for mixtures of normal distributions to two distributions. The authors than tested four different tree-ring chronologies for climatic signal enhancement by modeling the climatic signal in each chronology after estimation by their new method and by the arithmetic mean alone. The percentage of years in which skewness or multimodality was indicated was approximately 25%. Shiyatov and Mazepa (1987) found a statistically significant ($\alpha = .15$) increase in the strength of the modeled climatic signal in three of the four chronologies developed by their procedure. This encouraging

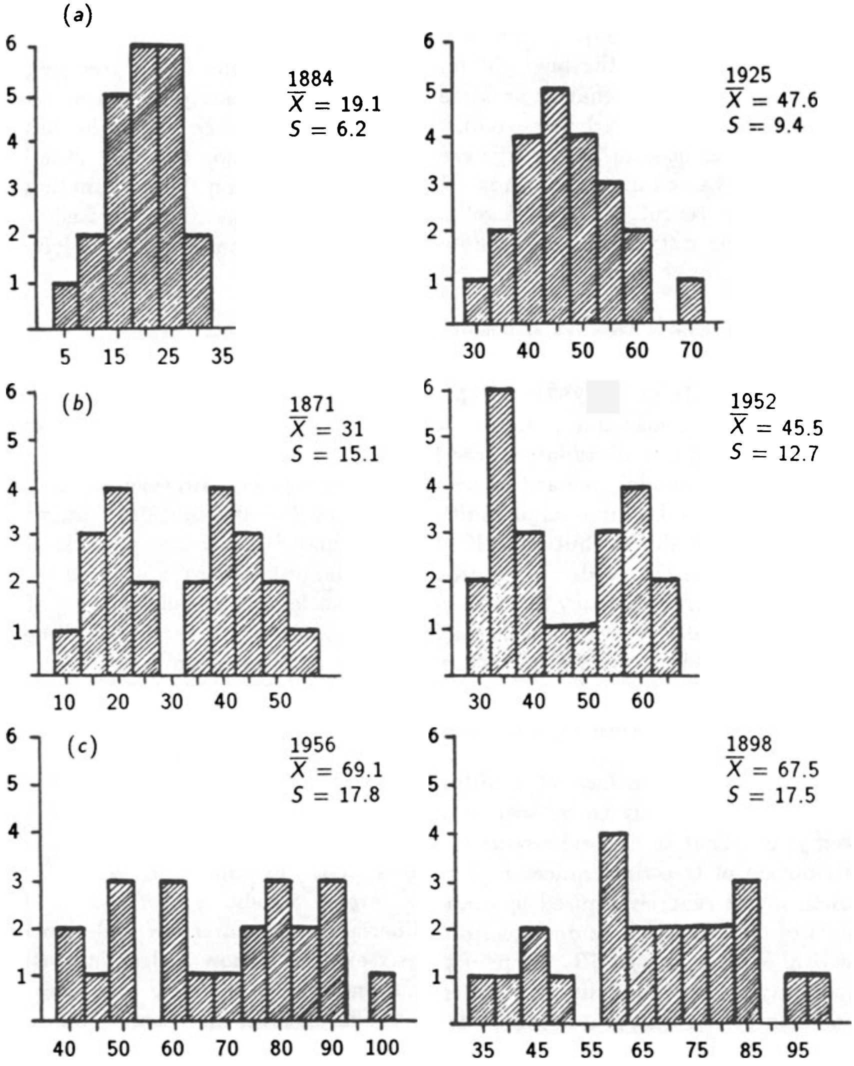

Figure 3.9. Some characteristic frequency distributions of *Larix sibirica* tree-ring indices for separate years. Note the clearly bimodal distributions in *3.9(b)* and the more multimodal forms in *3.9(c)*. $\bar{X}$ is the arithmetic mean, and S is the standard deviation for each distribution.

result is probably conservative because it is based on using all years, not just those years in which a mixture of normal distributions was detected.

As noted by Shiyatov and Mazepa (1987), their method has not yet dealt with the problem of serial dependence between the yearly sample distributions of indices and how it may affect any mixtures of normal distributions in each

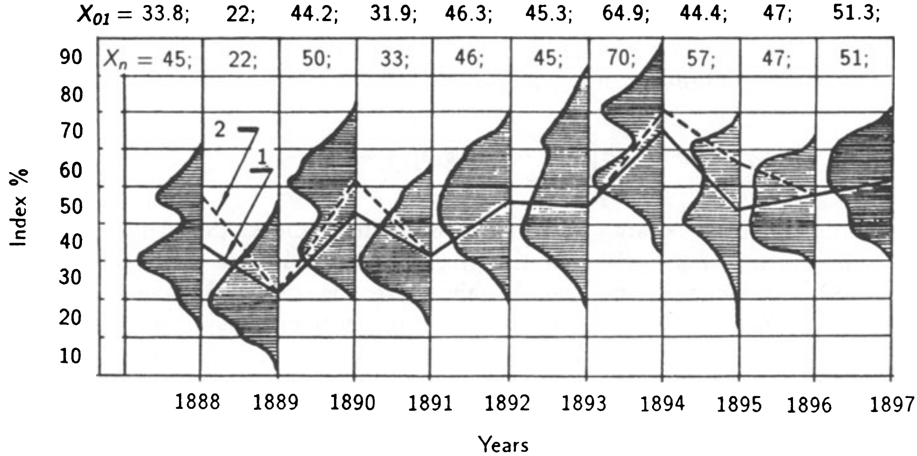

Figure 3.10. An illustration of the method of computing the mean-value function based on modeling the distributions of sample indices as mixtures of normal distributions. The mean-value function is computed using the arithmetic mean of all samples (1); the mean-value function is computed by testing for the presence of a mixture of normal distributions and adjusting the estimate of the mean, accordingly (2). (From Shiyatov and Mazepa, 1987.)

sample. Therefore, further research is needed before the method can be routinely used for estimating mean-value functions. There is almost certainly some overlap between the expected performance of the biweight robust mean and the method of Shiyatov and Mazepa (1987). This performance overlap is likely when the frequency distribution is principally skewed or long-tailed, rather than bimodal or multimodal. When the former condition is present, either method may provide outlier-resistant measures of central tendency. However, when the latter condition occurs, the biweight mean will iterate toward the mode of highest frequency or to a compromise position between balanced modes, without regard to biological considerations. In this case, the performance of the two techniques will diverge.

3.4.3. The use of autoregressive-moving average models in estimating the common signal

The computation of the mean-value function, by any of the above methods, is easily done with the tree-ring indices. However, if the autocorrelation within each series is high, then a more statistically efficient estimate of the mean-value function (i.e., a higher SNR) is possible, in many cases, through the use of time series modeling and prewhitening techniques. Tree-ring series have an autocorrelation structure that allows the estimation of C_t to be broken down into a two-stage procedure (Cook, 1985; Guiot, 1987a), based on ARMA time series modeling (Box and Jenkins, 1970).

Tree-ring indices can be expressed, in difference equation form, as an ARMA process of order p and q, *viz.*,

$$I_t = \phi_p I_{t-p} + \cdots + \phi_1 I_{t-1} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \quad , \tag{3.29}$$

where the values for e_t are serially random inputs or shocks that drive the tree-growth system as reflected in the tree rings, the ϕ_i values are the p autoregressive (AR) coefficients, and the θ_i values are the q moving average (MA) coefficients that produce the characteristic persistence or memory seen in the I_t. Equation (3.29) can be economically re-expressed in polynomial form using the backshift operator, B, as

$$I_t = [\theta(B)/\phi(B)]\, e_t \quad , \tag{3.30}$$

where

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

and

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$$

(Box and Jenkins, 1970). On an individual series basis, the values for e_t are assumed to be composed of inputs owing to climate (C_t) and those owing to disturbances and random variability ($D1_t$ and E_t). Component A_t is assumed to be nonexistent either in the raw ring-width series in total or for certain periods (Guiot, 1987a) (i.e., $I_t = R_t$), or to have been removed by detrending or differencing. Extensive ARMA modeling of tree-ring chronologies by Rose (1983) and Monserud (1986) indicates that western North American conifers are most commonly ARMA(1,1) processes, with the best competing models falling in the AR(1)–AR(3) classes. Cook (1985) restricted his analyses of eastern North American conifer and hardwood tree-ring chronologies to the AR process and found that AR(1)–AR(3) models were satisfactory, in most cases.

ARMA processes are examples of causal feedback-feedforward filters (Robinson and Treital, 1980) that are used extensively in geophysical signal analysis. The AR part of the process operates as a feedback filter, while the MA part operates as a feedforward filter. That is, the current I_t is a product of the current e_t plus past I_{t-i} inputs, which are fed back into the process, and past e_{t-i} inputs, which are fed forward upon the arrival of the current e_t. In this way, the potential for current growth is largely affected by previous radial growth (I_{t-i}) and by reflections of antecedent environmental inputs (e_{t-i}).

Thus, the ARMA process is an elegant mathematical expression of *physiological preconditioning* (Fritts, 1976).

An important concept of ARMA processes is the way in which they can operate as signal amplifiers. The amplifier mechanism can be seen in the variance formula of AR(p) processes, *viz.*,

$$\sigma_y^2 = \frac{\sigma_e^2}{1 - \rho_1\phi_1 - \rho_2\phi_2 - \cdots - \rho_p\phi_p} , \tag{3.31}$$

where σ_y^2 is the variance of the observed AR process, σ_e^2 is the variance of the unobserved random shocks, and ρ_i and ϕ_i are the theoretical autocorrelation and autoregression coefficients of the process. If both the ρ_i and ϕ_i are positive, which is usually the case for tree-ring indices, then σ_y^2 will always be greater than σ_e^2. A reflection of this amplifier mechanism is transience. That is, the effect of a given e_t in a tree-ring series, whether climatic or from disturbance, will last for several years or, in extreme cases, decades before it disappears (Cook, 1985). The consequence of transience, when endogenous disturbance shocks are present in the e_t, is a degradation of the SNR of the C_t in the mean-value function of tree-ring indices.

To remove the effects of unwanted, disturbance-related transience on the common signal among trees, the tree-ring indices can be modeled and prewhitened as AR(p) (Cook, 1985) or ARMA(p, q) (Guiot, 1987a) processes before the mean-value function is computed. The order of the process can be determined at the time of estimation using the Akaike Information Criterion (AIC) (Akaike, 1974). Once the ARMA(p, q) coefficients are estimated, the prewhitening is carried out as

$$e_t = I_t - \phi_1 I_{t-1} - \cdots - \phi_p I_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \tag{3.32}$$

in difference equation form, or

$$e_t = [\phi(B)/\theta(B)]\, I_t \tag{3.33}$$

in backshift operator form. The tree-ring series are now *white noise*.

The resulting e_t represent the contributions of C_t, $D1_t$, and E_t, with $D2_t$ assumed to be absent at this stage. The reduction of the transient effects of endogenous disturbance pulses results in an increase in fractional common variance (Wigley *et al.*, 1984) or $\%Y$ (Fritts, 1976) and in an improved SNR in the mean-value function of the e_t, denoted $\mathbf{e}_t$. This results in an improved estimate of C_t, especially if the biweight robust mean is also used in computing $\mathbf{e}_t$. Cook (1985) found that the average absolute increase in fractional common variance between sampled trees was about 7% for 66 tree-ring chronologies developed from closed-canopy forest stands. The average relative increase in fractional common variance, compared with that of a mean-value function developed without using AR modeling, was about 25%. There are no comparable figures for ARMA-based models.

The estimate of C_t, in the form of $\mathbf{e}_t$, is incomplete because it is missing the natural persistence owing to climate and tree physiology. To have a complete model of the common signal within the ensemble, an estimate of the common persistence structure among all detrended tree-ring series is necessary. For the pure AR model, a pooled estimate of autoregression, denoted $\Phi(\mathbf{B})$, can be computed directly from lag-product sum matrices of the ensemble that include information on persistence both within and among series (Cook, 1985). The method appears to be quite robust in the face of high levels of out-of-phase fluctuations among series that are caused by endogenous disturbances. Unfortunately, this pooling procedure is difficult to apply to the ARMA case because of the highly nonlinear MA coefficients.

Guiot (1987a) addressed the estimation of the common ARMA model by first creating a mean-value function of raw ring-width series from very old trees, which were selected by principal components and cluster analysis of the corresponding white-noise series. The common ARMA(p,q) model, denoted $\Theta(\mathbf{B})/\Phi(\mathbf{B})$, was then estimated for stationary subperiods in the mean ring-width series. However, this approach may be difficult to apply to the general closed-canopy forest where stationary subperiods rarely exist for any length of time (e.g., *Figures 3.1* and *3.2*). As an alternative, the ring-width series could be detrended first, a robust mean-value function created, and the mean series modeled as an ARMA process to produce $\Theta(\mathbf{B})/\Phi(\mathbf{B})$. This method would depend upon sufficient replication to diminish the effects of endogenous disturbance effects. Experiments in detrending ring-width series (Cook, 1985) indicate that the choice of the detrending method will have little effect on the order and coefficients of the mean ARMA process as long as the variance removed by the trend line is effectively all trend, as defined by Granger (1966).

The estimation of the common signal components is now complete. A final tree-ring chronology, $\mathbf{I}_t$, containing both common signal components, can be easily created by simply convolving the pooled AR, $\Phi(\mathbf{B})$, or the ARMA $[\Theta(\mathbf{B})/\Phi(\mathbf{B})]$ operators with the $\mathbf{e}_t$ (Cook, 1985; Guiot, 1987a), once suitable starting values are obtained. The q starting values for the MA component will ordinarily be set to zero, the unconditional expected value of $\mathbf{e}_t$. The p starting values for the AR component may be obtained from the I_t lost through prewhitening (Cook, 1985), by back-forecasting of the I_t past the beginning of the I_t (Box and Jenkins, 1970), or by using the unconditional expected value of the I_t. If the characteristic transience of the AR component decays rapidly, then the choice of starting values will have little effect on the final series.

For the full ARMA case, the final estimate of C_t is

$$\mathbf{I}_t = [\Theta(\mathbf{B})/\Phi(\mathbf{B})]\mathbf{e}_t \quad . \tag{3.34}$$

Aside from producing an efficient estimate of C_t, knowledge of $[\Theta(\mathbf{B})/\Phi(\mathbf{B})]$ and $\mathbf{e}_t$ are also useful for estimating $D1_t$ and E_t in the individual tree-ring series.