SO-70
GTR
C.2

United States
Department of
Agriculture

Forest Service

**Southern Forest
Experiment Station**

New Orleans,
Louisiana

General Technical Report
SO-70

# Theory and Programs for Dynamic Modeling of Tree Rings from Climate

Paul C. Van Deusen and Jennifer Koretz

## SUMMARY

Computer programs written in GAUSS™ for IBM compatible personal computers are described that perform dynamic tree ring modeling with climate data; the underlying theory is also described. The programs and a separate users manual are available from the authors, although users must have the GAUSS software package on their personal computer. An example application of the programs is given, and a number of relevant references should provide enough information for users to decide if they are interested in this system. The purpose of these programs is to make these techniques easily available to the community of tree ring analysts for application and further development.


August 1988

# Theory and Programs for Dynamic Modeling of Tree Rings from Climate

Paul C. Van Deusen and Jennifer Koretz

## INTRODUCTION

The field of dendrochronology is currently in a period of rapid development. This is due in part to emerging concern about air pollution effects on forests and the reality that tree rings represent a long-term record that may yield information on these effects. Linked with this is the fact that much of the research on air pollution effects on forests is being conducted in closed canopy forests. Traditionally, tree ring research was done in arid forests of the Southwestern United States where trees are less affected by stand dynamics.

One of the mainstays of dendrochronology that states that current conditions are analogues for past conditions, the uniformitarian assumption, may not be applicable in closed forest conditions. Trees being heavily affected by stand dynamics are likely to exhibit changing relationships with the environment over time, even though the environment itself has remained stable. The purpose of this report is to describe a procedure and computer programs for modeling climate relationships with tree ring data without making the uniformitarian assumption.

A simple method for dynamically modeling the relationship of climate and ring width was presented by Van Deusen (1987a) and expanded to include more tree level information such as diameter and age in Van Deusen (1988). The simpler method will form the basis of this paper, although some minor extensions are made to improve the model fit and to increase the level of statistical rigor. However, the prevailing purpose here is to make these procedures available to interested tree ring analysts for testing and evaluation without each having to write their own software.

## THE PROPOSED MODEL

The procedure presented herein has a number of distinct features:

(1) The uniformitarian assumption is not required. Climate models with dynamic parameters are employed, so changing climatic relationships can be modeled. The Kalman filter (Kalman 1960) becomes important in the estimation process, but it should be noted that any linear model could be formulated (as a state space model) in a manner suitable for estimation using the Kalman filter. Other applications of the Kalman filter have been made to tree ring analysis that are quite different from the one to be discussed below, e.g. Visser and Molenaar (1987) and Van Deusen (1987b). The estimation method is recursive and past data are given less weight than current data in determining the current parameter estimates. This seems biologically reasonable, because when the tree was growing a particular ring of wood it presumably was influenced more by recent circumstances than by the distant past.

(2) All tree ring series are maintained as individuals throughout the analysis and data reduction is achieved by estimating parameters. This is different from the traditional method of modeling climate with a single average ring width or index series, although the results could be quite similar in some cases.

(3) The method simultaneously estimates the mean value function and the climate model. The standardization procedure is also incorporated into the method proposed in Van Deusen (1988).

### The State Space Formulation

The formulation of the suggested model is simple when written in state space form as we demonstrate below. The first equation, known as the measurement or observation equation, is

$$Y_t = F_t \alpha_t + v_t, \tag{1a}$$

where $Y_t$ is an $n_t$x1 vector of standardized ring widths, $n_t$ is the number of observations for year t, $F_t$ is an $n_t$xp matrix that for our formulation has 1's in the first column and zeroes in the rest, $\alpha_t$ is a px1 parameter vector, and $v_t$ is an $n_t$x1 error vector with variance matrix $V_t$. The second equation in the formulation is known as the transition equation and is written generally as

$$\alpha_t = G_t \alpha_{t-1} + w_t, \tag{1b}$$

where $G_t$ is a pxp transition matrix, and $w_t$ is a pxl error vector with variance matrix $W_t$. Equations (1a) and

Paul C. Van Deusen is mathematical statistician and Jennifer Koretz is statistician at the Institute for Quantitative Studies, Southern Forest Experiment Station, Forest Service—USDA, New Orleans, Louisiana 70113. This research has been supported by the National Vegetation Survey as part of the USDA Forest Service/US EPA Forest Response program.

(1b) shed little light on our specific formulation, but the general notation is useful in describing the algorithms later.

More specifically, the parameter vector in our suggested formulation is

$$\alpha_t = \begin{bmatrix} m_t \\ b_0 \\ b_{1t} \\ \cdot \\ \cdot \\ \cdot \\ b_{p-2,t} \end{bmatrix},$$

where $m_t$ is roughly equivalent to the average chronology in classical dendrochronolgy, $b_0$ is a fixed intercept in the climate model, and the other b-parameters are multiplied by p-2 climate variables in the transition equation to partially explain the $m_t$ parameter movements. Notice that only the m-parameter appears in the observation equation, since the F-matrix consists of 0's except for the first column of 1's, which zeros out the b-parameters.

The transition equations for each parameter are

$$m_t = b_0 + b_{1,t-1}c_{1,t} + \ldots + b_{p-2,t-1}c_{p-2,t} + w_{mt}, \qquad (2)$$

$$b_0 = b_0 + 0,$$

$$b_{1,t} = b_{1,t-1} + w_{1t},$$

$$b_{2,t} = b_{2,t-1}\, w_{2t},$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$b_{p-2,t} = b_{p-2,t-1} + w_{p-2,t},$$

where the $c_{i,t}$, $i$=1,...,p–2 represent p–2 climate variables. Thus, equation (2) is analogous to the transfer function from classical methods, with the important difference being that the parameters are time varying. The following G-matrix will result in the above transition equations when substituted into equation (1b):

$$G_t = \begin{bmatrix} 0 & 1 & c_{1,t} & \cdots & c_{p-2,t} \\ 0 & 1 & 0 & \cdots & 0 \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ \cdot & & & & \cdot \\ 0 & & \cdots & & 1 \end{bmatrix}.$$

Basically, the G-matrix is created by altering the first row of an identity matrix of order p.

### Variance Specification

The observation variance matrix, $V_t$, and the transition variance matrix, $W_t$, must be specified before the Kalman filter equations can be solved. We specify $V_t$ to be $\sigma_t^2 I_t$, where $I_t$ is an identity matrix of order $n_t$ and

$$\hat{\sigma}_t^2 = \frac{\sum_{i=1}^{n_t} (y_{it} - \bar{y}_t)^2}{n_t - 1}. \qquad (3)$$

Thus, the error terms in the observation equation for year t are assumed to be independent, with variance estimated from the standardized ring widths for year t. However, the fact that $m_t$ is random and shared by all observations means that contemporaneous correlation is accounted for in this formulation.

The transition variance matrix, $W_t$, is assumed to be

$$W_t = \frac{\sigma_t^2}{n_t} \mathrm{DIAG}[\, q_m, 0, q_1, \ldots, q_{p-2}],$$

where DIAG[.] means to form a diagonal matrix from the arguments. Therefore, the parameters in the model are assumed a-priori to be independent, and the intercept term in equation (2) is fixed. The reason for multiplying the transition variance matrix by $\sigma_t^2/n_t$ is that the $m_t$ parameter should be roughly equal to the mean value in the vector $Y_t$, and the proper choice for $q_m$ allows this to happen even when the climate model is poor. In fact, it can be shown that when all of the b-parameters equal zero the mean value parameters are estimated as follows:

$$m_t = \frac{q_m}{q_m + 1}\, \bar{y}_t,$$

with variance

$$V(m_t) = \frac{q_m}{q_m + 1} \frac{\sigma_t^2}{n_t}.$$

Thus, the parameter $q_m$ allows the system to produce reasonable estimates of the mean value function even when the climate relationships are poor. When the climate model predicts well, less weight is given to the data and more to the climate model. The parameters $q_m,\ldots,q_{p-2}$ can be estimated via the method of maximum likelihood, which is discussed below. The $q_m$-parameter was assumed equal to 1 in Van Deusen (1987a), but allowing more flexibility can considerably improve the model performance as was shown above.

## KALMAN FILTER THEORY

To fully understand what the computer programs are doing, it is necessary to understand the basics of Kalman filter theory. The Kalman filter is a set of equations that provide estimates of the parameters in equations (1a) and (1b) that have the general property of being best linear unbiased estimates (BLUE) when the following assumptions hold:

$$E\begin{bmatrix} v_t \\ w_t \end{bmatrix} \sim \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} V_t & 0 \\ 0 & W_t \end{bmatrix} \right\},$$

i.e., the transition and observation error terms are distributed with mean zero and are not contemporaneously correlated. Furthermore, there should be no serial correlation either within or between the sequences $v_t$ and $w_t$. Estimation procedures have been developed when some of these conditions were violated (Jazwinski 1970), but these will not be required here. Also, the starting values used should be independent of all error terms. If the starting values and the transition and observation errors are normally distributed, the Kalman filter estimates become the best minimum variance estimators among all estimators.

The state parameters are estimated with equations that can be considered in three parts: prediction equations, updating equations, and smoothing equations. Let $a_t$ denote the optimal estimator of $\alpha_t$ as based on all information up to and including $Y_t$, and let the covariance matrix of $a_t - \alpha_t$ be $P_t$. The prediction equations for $\alpha_t$ and the associated covariance matrix conditional on $a_{t-1}$ and $P_{t-1}$ are

$$a_{t/t-1} = G_t a_{t-1}, \text{ and} \tag{4a}$$

$$P_{t/t-1} = G_t P_{t-1} G_t' + W_t. \tag{4b}$$

When $Y_t$ becomes available, the updating equations for the estimate of $\alpha_t$ and the associated covariance matrix are

$$a_t = a_{t/t-1} + P_{t/t-1} F_t' H_t^{-1} E_t, \text{ and} \tag{5a}$$

$$P_t = P_{t/t-1} - P_{t/t-1} F_t' H_t^{-1} F_t P_{t/t-1}, \text{ where} \tag{5b}$$

$E_t = Y_t - F_t a_{t/t-1}$, and

$H_t = F_t P_{t/t-1} F_t' + V_t$.

The estimate of $\alpha_t$ in (5a) is the sum of its estimate at time t–1 and a weighted average of the prediction errors, $E_t$. The covariance matrix of the prediction errors is $H_t$. A numerically superior equation for $P_t$ is used in our program:

$$P_t = (I_p - K_t F_t) P_{t/t-1} (I_p - K_t F_t)' + K_t V_t K_t',$$

where $K_t = P_{t/t-1} F_t' H_t^{-1}$ is the Kalman gain.

This form of the equation tends to produce a positive definite symmetric covariance matrix for less well conditioned problems than equation (5b).

At any time, $a_t$ is the optimal linear estimate given all previous information, but only the estimate at time T contains all available information. The optimal solution for any time t, given all available information, is referred to as smoothing or signal extraction. The state space solution to this problem (Harvey 1981) uses solutions at time T and recursively goes backward to time 1. This yields the optimal smoothed estimates of the state parameters with associated covariance matrices as follows:

$$a_{t/T} = a_t + P_t^* (a_{t+1/T} - G_{t+1} a_t), \text{ and} \tag{6a}$$

$$P_{t/T} = P_t + P_t^* (P_{t+1/T} - P_{t+1/t}) P_t^{*\prime}, \text{ where} \tag{6b}$$

$P^*_t = P_t G_{t+1}' P_{t+1/t}^{-1}$, and

$a_{T/T} = a_T$ and $P_{T/T} = P_T$ for the starting values.

Equations (4b), (5a), and (5b) require knowledge of variance matrices $W_t$ and $V_t$ that is not often available in statistical applications. There are a number of suggested methods for estimating the unknown parameters in these matrices, including a maximum likelihood approach. Schweppe (1965) has shown that the log-likelihood of a sample from a population described by equations (1a) and (1b) is

$$L = -\frac{1}{2} \sum_t \log |H_t| + E_t' H_t^{-1} E_t. \tag{7}$$

The likelihood function is specified completely in terms of the prediction errors and their covariance matrices, which are natural outputs of the Kalman filter solution equations. The only assumptions required for this approach are that $w_t$ and $v_t$ be normally and independently distributed. Harvey (1981) and Engle and Watson (1981) give good developments of this method. The likelihood function (7) is evaluated iteratively by solving for the state parameters with equations (5a) and (5b) conditional on the most recent values for the unknown parameters in the model. The log-likelihood is maximized by using a grid search for the variance parameters or the method of scoring (Engle and Watson 1981).

If the model is correct, the prediction errors form an innovations sequence (Anderson and Moore 1979) in that $E_t$ represents the new information in $Y_t$ in the sense that $E_t$ is orthogonal to all previous $E_t$'s. This is true with or without the normality assumption. A diagnostic check for a correctly specified model involves checking to see if the prediction errors form a white noise sequence with no lagged cross-correlations. Thus, the Kalman filter in combination with the method of maximum likelihood provides a feasible algorithm for fitting and testing the validity of equations (1a) and (1b) with tree ring data.

## COMPUTER PROGRAMS

### Data and Software Requirements

The authors offer a number of programs that will enable users to take their data through the entire analysis with a minimum of effort. The programs are written in the GAUSS™ programming language, which will run on IBM-compatible personal computers and requires 512K memory, the DOS 2.10+ operating system, and a math coprocessor. This software was chosen because it is inexpensive and efficient, but the programs could be rewritten in some other software language.

Two types of data are required: standardized tree ring data and climate data. Standardization in dendrochronology involves removing uninteresting long-term trends from each tree ring series so that most of the remaining variance is due to signals that all of the trees share in common. A portion of this common signal should be due to climatic factors. The data in the example to be presented here have been standardized by taking first differences of the natural logarithm of ring width. The rationale for this is discussed in Van Deusen (1987a,b). Users may want to standardize their data in other ways. The climate data used in the following example came from National Weather Service climatic division averages for total monthly rainfall and average temperature.

The main program requires the standardized tree ring data to be input as a matrix where each row corresponds to a tree and each column corresponds to a year, with the first column being the first year to be analyzed. The climate data are input in a matrix with each row corresponding to a year and each column corresponding to a climate variable. Thus, the number of rows in the climate matrix should equal the number of columns in the tree ring matrix. The users manual describes programs for preparing these matrices. The program is not currently designed to predict past values of climate, although this is certainly possible (see Van Deusen 1987a).

### Searching for Climate Variables

A program, called dynatry, is provided that searches for climatic variables by a brute force technique. The program asks the user to list the climate variables to try and then fits the model to each climate variable and outputs summary statistics before proceeding to the next variable. The program operates by allowing the user to force any number of variables into the model, then one variable from the 'try list' is added to the basic model, summary statistics are produced, the previous 'try list' variable is dropped, and the next 'try list' variable is added. When the 'try list' is exhausted, the program stops and the user can print out the results. The single most important summary statistic provided is the maximum likelihood value, and usually the variables with the largest likelihood values would be judged most important. The users manual contains more information on program dynatry.

### Main Program

The main program, called dynaclim, implements the Kalman Filter equations (4a)–(5b) and estimates the unknown variance parameters discussed above in the section on variance specification. This is done by maximizing the log-likelihood function (7) using the method of scoring described in Engle and Watson (1981). The method of scoring requires only first derivatives and provides asymptotically valid estimates of the variances of the estimated variance parameters as a by-product. The program computes exact derivatives recursively (see Ljung 1981) and exploits the special structure of this particular problem for increased efficiency. This results from the fact that the F-matrix in equation (1) is of a very simple form that leads to elementary equations for $H^{-1}$ and the determinant of H in equation (7) among other things.

The program begins by asking the user how many years are involved in the analysis. It is possible that the tree ring data go beyond the end of the climate data, and the user could analyze fewer years than there are columns in the tree ring matrix. Normally, years and columns would be equal. The program then requests the number of climate variables to be included, which effectively tells the program how many parameters to prepare to estimate. Finally, the program asks for the column number and lag to associate with each climate variable. Remember that the columns of the climate data matrix correspond to climate variables and the rows correspond to years. If lag 0 is entered, the program uses the climate variable from row t when analyzing year t. If lag 1 is entered the program uses the climate variable from row t-1 when analyzing year t, and lag 2 means use row t-2, etc.

The program then uses a grid search/bisection method to find starting values and begins the scoring algorithm with the best starting values it finds. After converging on a solution the user is asked if smoothing should be performed. If the answer is yes, the smoothed state variables and the prediction errors are saved to files for later analysis and plotting using other programs that are described below. In either case, the user can return to the point where the program asks for new climate variables and lag values or stop the program execution.

*Example Application of Main Program.*—Appendix A contains the output from our example application. The program output in Appendix A begins by asking the number of years, which is 82 (1902–1983) for this analysis. The data set being used here contains 28 cores from eastern hemlock trees (*Tsuga canadensis*) from the upper peninsula of Michigan and was provided by Dr. Edward R. Cook of the Lamont-Doherty Tree Ring Laboratory of Columbia University. The main program uses a data reading module, called dynadata, that was set up to read the hemlock data and climate data from the appropriate climatic division of Michigan. To read other data sets the user must edit this module as described in the users manual.

The program then asks for the number of climate variables, which was 1 for this example but could be any number in general. The 12 columns in the climate matrix correspond to the first difference of January through December temperature. Therefore, climate variable 7 is the first difference of July temperature. Lag 1 is chosen, which means that previous year's climate is used to predict current year standardized

ring width. First difference of July temperature was chosen after trying a number of variables with program dynatry.

After inputting the climate variables, the program beings outputting information on its initial grid search to find starting values for the unknown variance parameters. The grid search finds that $q_m$=31 and $q_1$=.0125 are the best starting values tried and begins actually using the scoring algorithm (Engle and Watson 1981) to produce the output after the message "Finished with Bisection" is printed (see Appendix A). The fact that the iteration number gets out of sequence is unimportant and results from the way the program uses the iteration number to control the grid search.

The program converges and produces the final results and prints the variance parameter estimates associated with the best likelihood value it could find, i.e. 2297.80. The variance parameter estimates are listed with $q_m$ first and $q_1$ second. The covariance matrix is the inverse of the information matrix, which had to be computed anyway as part of the scoring algorithm (see Theil 1971, p. 395). In this example, only $q_m$ appears to be significantly different from zero as judged by the following approximate 95-percent confidence interval: $29.29 \pm 2\sqrt{64.82}$.

The program then asks if the smoothed estimates are desired, which they are for this example. After smoothing is completed, the user has the option of either entering different climate variables for another run or stopping. In this case, we stop to look at the smoothed output.

## Plotting Program

A program called dynaplot is provided that enables the user to plot certain results from the main program output. The smoothed parameter estimates have been stored in a file called "ans," and the prediction errors are stored in "preder." Dynaplot will automatically access these files when executed by the user. Hardcopy from this program can be generated by GAUSS; see GAUSS enhanced graphics on pages 32 and 33 of the main manual. Module GRAPH2d of GAUSS must be installed before this program will run.

*Plot Program Output.*—Appendix B contains a series of plots produced by dynaplot. Plot B1 gives the average one step ahead prediction error defined for equation (5b) with approximate 95-percent confidence interval as dashed lines. If the climate model were able to predict well, these should not be significantly different from zero. The model, unfortunately, was not predicting well enough for that ideal situation to occur in this example. Plot B2 is the m-parameter, which is roughly equivalent to the average chronology from classical procedures. Plot B3 is the constant intercept from the climate model with 95-percent confidence intervals. Plot B4 is the time varying parameter that multiplies past July temperature differences. This shows a tendency to become less significant following about 1970. Plot B5 comes from another general purpose program called "overlay" for producing overlaid plots of climate and tree ring data. Plot B5 is previous July temperature differences overlaid on the average chronology for the standardized Michigan hemlock data. This shows the strong negative relationship before 1970 that was evident on plot B4. However, it indicates that the relationship did not disappear after 1970; it was in the process of becoming a positive rather than a negative relationship.

## Correlation Matrices Program

Program dynacorr generates lagged cross-correlation matrices from the prediction errors stored in file preder. This program computes the cross-correlations defined below:

$$\rho_\tau(i,j) = \frac{E[e_t(i)' e_{t-\tau}(j)]}{\sigma_i \sigma_j} ,$$

where $e_t(i)$ is the vector of prediction errors for tree i, $\sigma_i$ is the standard error of a prediction error for tree i, and $\tau$ is the lag value applied to tree j's prediction errors. These quantities are estimated from the data. Under the null hypothesis that the prediction errors are not causally related, the sample cross-correlations are independently and normally distributed with mean zero and variance 1/T (T is the number of years where the particular prediction errors exist) when the sample size is large. The program computes matrices of these cross-correlations and prints them out in schematic form. A "+" means the sample cross-correlation was greater than 2 standard errors (i.e., $2/\sqrt{T}$), a "-" means less than 2 standard errors, and a "0" means in between.

*Correlation Program Output.*—Appendix C contains the output from dynacorr for this example using the Michigan hemlock data with one climate variable. Figure C1 is the lag 0 matrix, which is necessarily symmetric so only the upper diagonal is printed. This shows that 83 percent of the values are significant and positive, which verifies what the average prediction error plot showed. The climate model is not performing well in that much of the variance remains unexplained. However, the fact that all significant correlations are positive suggests that the crossdating and standardization were effective in making all of the series quite similar. This matrix is interpreted by noting that the ith row and jth column are the cross correlation of tree i with tree j prediction errors. Therefore, the diagonals will always be significant at lag 0, because this is tree i's correlation with its own current prediction errors. If one core were completely uncorrelated with any other cores, it would suggest a crossdating error. Therefore, the lag 0 plot is a valuable check on crossdating.

The remaining schematics in Appendix C give cross-correlation matrices for lags 1, 2, and 3. There is no reason why these should be symmetric, so the proportion of significant values is based on all elements being independent and free to take on any value. There are a

large number of significant lag 1 negative correlations, which suggests that either the trees are competing or the climate model needs improvement. Frankly, the interpretation of the lagged values is difficult, and without knowing the physical location of these trees it is hard to make any biologically meaningful statements.

### Miscellaneous Programs

We have provided some other programs that are useful for manipulating and plotting either the tree ring or climate data associated with the analysis. These are described briefly below and more thoroughly in a users manual (available from the authors). GAUSS™ provides enough programming and plotting capability to allow users to develop other programs as needed within that system.

*Data Input Programs.*—A program called "makedata" has been provided to prepare an ascii tree ring data set in the standard "Tucson format." See Appendix D for examples of data in formats easily read by program makedata. This program requires padding missing values with dots in the raw tree ring data set, because GAUSS has problems with missing values on data input. The program outputs raw data or first differences of the natural log into a file suitable for dynaclim or dynatry.

Users can, of course, use any method to create GAUSS data sets suitable for input to dynaclim. The important thing is that the standardized tree ring data are organized so that rows represent trees and columns represent years. Furthermore, each year or column must have at least two non-missing values or dynaclim will abort. The climate data should be organized so that each row represents a year and each column a variable. Naturally, the order of the data sets is important, that is, column 1 of the tree ring data is the first year to be analyzed and corresponds to row 1 of the climate data. The ascii to GAUSS conversion section in the Gauss manual should be consulted on this.

*Plotting Program.*—A program called "plotraw" is provided for plotting tree ring or climate data. This program is easy to use and begins by requesting the name of the file where the data reside. The y-axis scales are chosen automatically, and the x-axis is assumed to take sequential integer values. Thus, you are asked for the first value of the x-axis that will correspond to the first column in the data set. The other columns will be numbered sequentially. This results in an x-axis that could go from 1 to 12, say, if you were plotting a climate data set with 12 columns, one for each month of the year. The program then asks for a row number, which with climate data corresponds to a year. A row number with tree ring data would correspond to a tree. This program is useful for visually inspecting the data. An enhanced version of plotraw, called overlay, was used to produce plot B5, which is discussed under "Plot Program Output."

*Data Editing.*—A program called dynaedit allows the user to do some basic editing of their GAUSS data sets. This asks for an input file and then allows you to delete whole rows or columns, or to change individual elements of the matrix. The results can be saved to the original or to a new file or you can abort at any time. This program is self explanatory after being executed.

## CONCLUSIONS

Theory and computer programs have been presented for predicting standardized ring widths dynamically from climate data. The method propounded is similar to the traditional approach used by dendrochronologists, in that a climate model is developed to predict a parameter that is analogous to the average chronology. Major differences between this system and the usual approach are that individual cores maintain their identity, and climate parameters are allowed to vary over time. Thus, the uniformitarian assumption of time invariant climate relationships is not required.

These programs are easy to use and will help dendrochronologists to familiarize themselves with the concept of dynamic tree ring modeling from climate. These procedures will no doubt be improved and enhanced over time, but it is unlikely that they will disappear.

## LITERATURE CITED


Anderson, B. D. O.; Moore, J. B. 1979. Optimal filtering. Englewood Cliffs, NJ: Prentice Hall, Inc.

Engle R.; Watson, M. 1981. A one-factor multivariate time series model of metropolitan wage rates. Journal of the American Statistical Association. 76(376): 774–781.

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

Harvey, A. C. 1981. Time series models. New York, NY: Wiley.

Jazwinski. A. H. 1970. Stochastic processes and filtering theory. New York, NY: Academic Press.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. Transactions ASME. Journal of Basic Engineering. 82: 34–45.

Ljung, L. 1981. Analysis of a general recursive prediction error identification algorithm. Automatica. 17(1): 89–99.

Schweppe, F. C. 1965. Evaluation of likelihood functions for Gaussian signals. IEEE Transactions on information theory. 11: 61–70.

Theil, H. 1971. Econometrics. New York, NY: Wiley.

Van Deusen, P. C. 1987a. Some applications of the Kalman filter to tree ring analysis. In: Proceedings, international symposium on ecological aspects of tree ring analysis. Tarrytown, NY: Marymount College; August 1986: 566–578.

Van Deusen, P. C. 1987b. Testing for stand dynamics effects on red spruce growth trends. Canadian Journal of Forest Research. 17: 1487–1495.

Van Deusen, P. C. 1988. A simultaneous approach to tree ring analysis. To appear as a chapter in the book "Forest growth: Process modeling of responses to environmental stress."

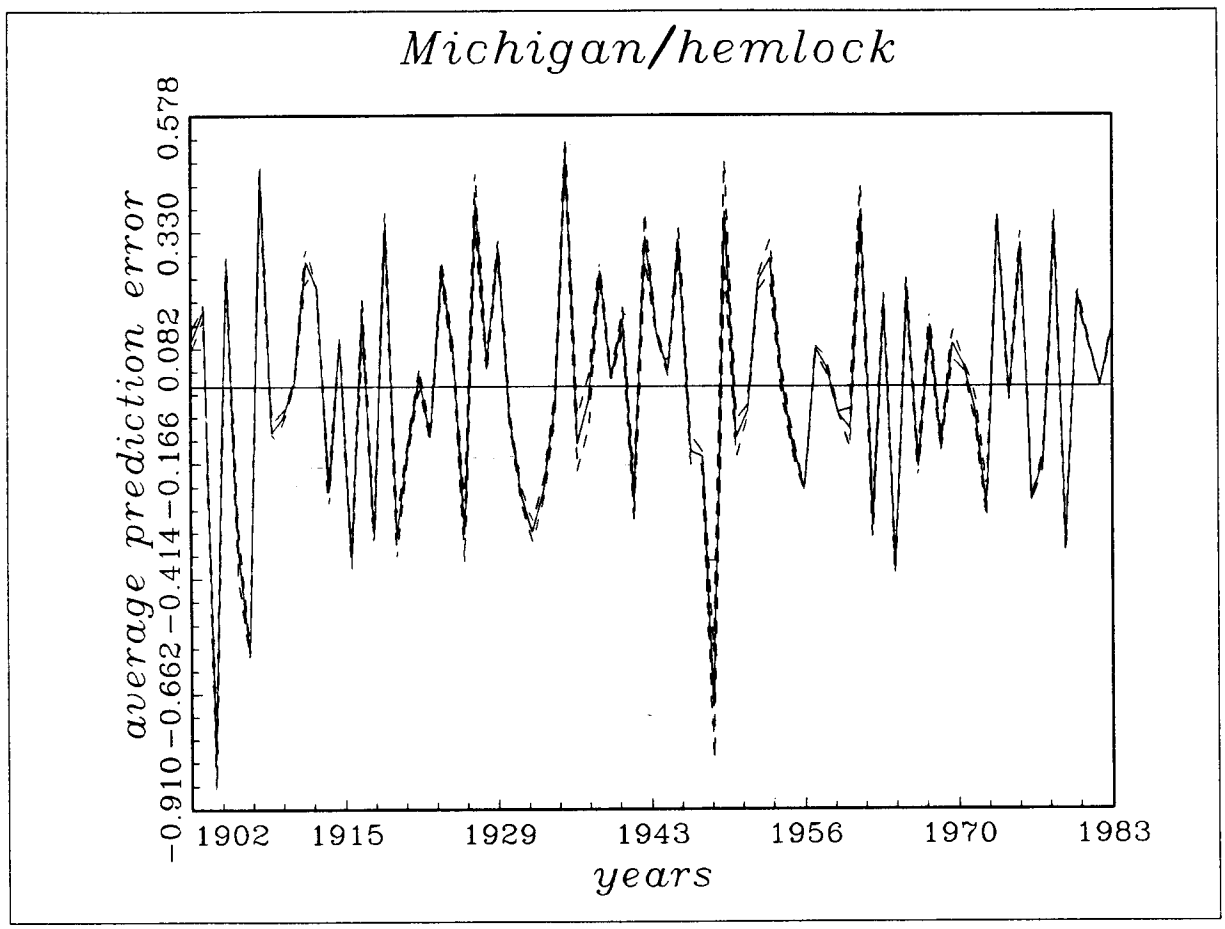
Visser, H; Molenaar, J. 1987. Time dependent responses of trees to weather variations: an application of the Kalman filter. In: Proceedings, international symposium on ecological aspects of tree ring analysis. Tarreytown, NY: Marymount College; August 1986: 579–590.

## APPENDIX A - OUTPUT FROM EXAMPLE APPLICATION

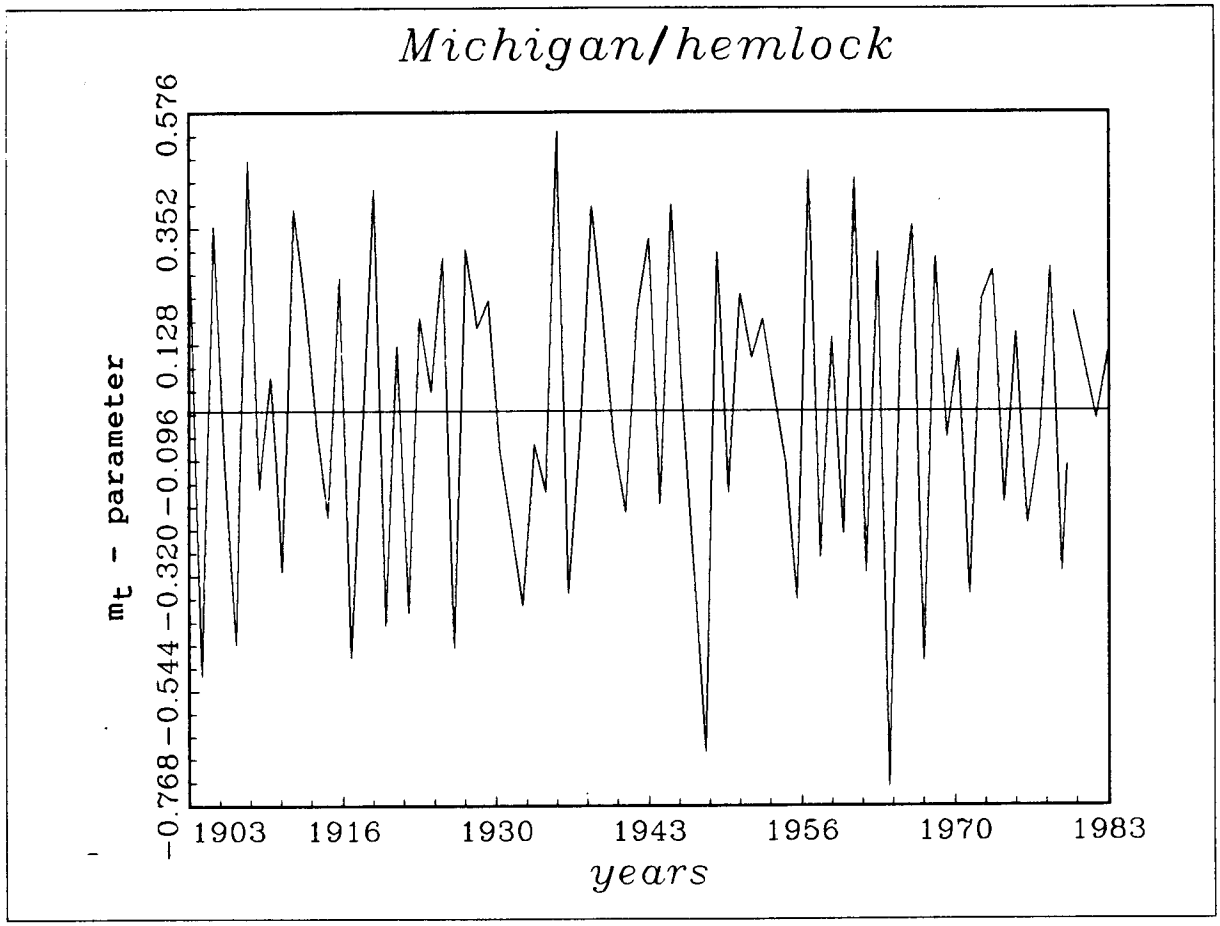
```
DYNACLIM VERSION 1.1 - BY PAUL VAN DEUSEN & JENNIFER KORETZ
USDA FOREST SERVICE, SOUTHERN FOREST EXPERIMENT STATION
*****************************************************************


ENTER NUMBER OF YEARS
     < 82 >

 You have      12.00000000  columns in climate matrix CLIM

     82.00000000 rows in climate data,     82.00000000 columns in treering data

ENTER NUMBER OF CLIMATE VARIABLES
     < 1 >


ENTER COLUMN AND LAG OF CLIMATE VAR     1  enter -1 to STOP
CLIMATE VARIABLE   < 7 >  LAG   < 1 >

------Begin bisection/grid search, ignore iteration numbers----
CURRENT LIKELIHOOD =   0.000000
BEGINNING ITERATION NUMBER   1 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
  1.000000
  0.000000
CURRENT LIKELIHOOD = 1687.771775
BEGINNING ITERATION NUMBER   2 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 11.000000
  0.002500
CURRENT LIKELIHOOD = 2270.400577
BEGINNING ITERATION NUMBER   3 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 16.000000
  0.005000
CURRENT LIKELIHOOD = 2288.091535
BEGINNING ITERATION NUMBER   4 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 21.000000
  0.007500
CURRENT LIKELIHOOD = 2294.123504
BEGINNING ITERATION NUMBER   4 ,PLEASE WAIT
```

```
CURRENT VARIANCE PARAMETERS
 26.000000
  0.010000
CURRENT LIKELIHOOD = 2295.971978
BEGINNING ITERATION NUMBER   4 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 31.000000
  0.012500
CURRENT LIKELIHOOD = 2295.979962
BEGINNING ITERATION NUMBER   4 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 36.000000
  0.015000
 ---------Finished with Bisection----------

CURRENT LIKELIHOOD = 2295.090842
BEGINNING ITERATION NUMBER   3 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 31.000000
  0.012500
CURRENT LIKELIHOOD = 2295.979962
BEGINNING ITERATION NUMBER   4 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 37.492555
  0.003756
CURRENT LIKELIHOOD = 2297.724042
BEGINNING ITERATION NUMBER   5 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 43.985110
  0.001570
CURRENT LIKELIHOOD = 2297.444017
BEGINNING ITERATION NUMBER   6 ,PLEASE WAIT
CURRENT VARIANCE PARAMETERS
 29.292810
  0.002924
--------------FINAL RESULTS-------------
FINAL LIKELIHOOD    2297.802358
BEST VARIANCE PARAMETER ESTIMATES
      29.292810
       0.002924
COVARIANCE OF VARIANCE PARAMETERS
      64.822559      -0.007322
      -0.007322       0.000006


ENTER 1 TO SKIP SMOOTHING
         < 0 >

ENTER COLUMN AND LAG OF CLIMATE VAR     1  enter -1 to STOP
        < -1 >-1
```

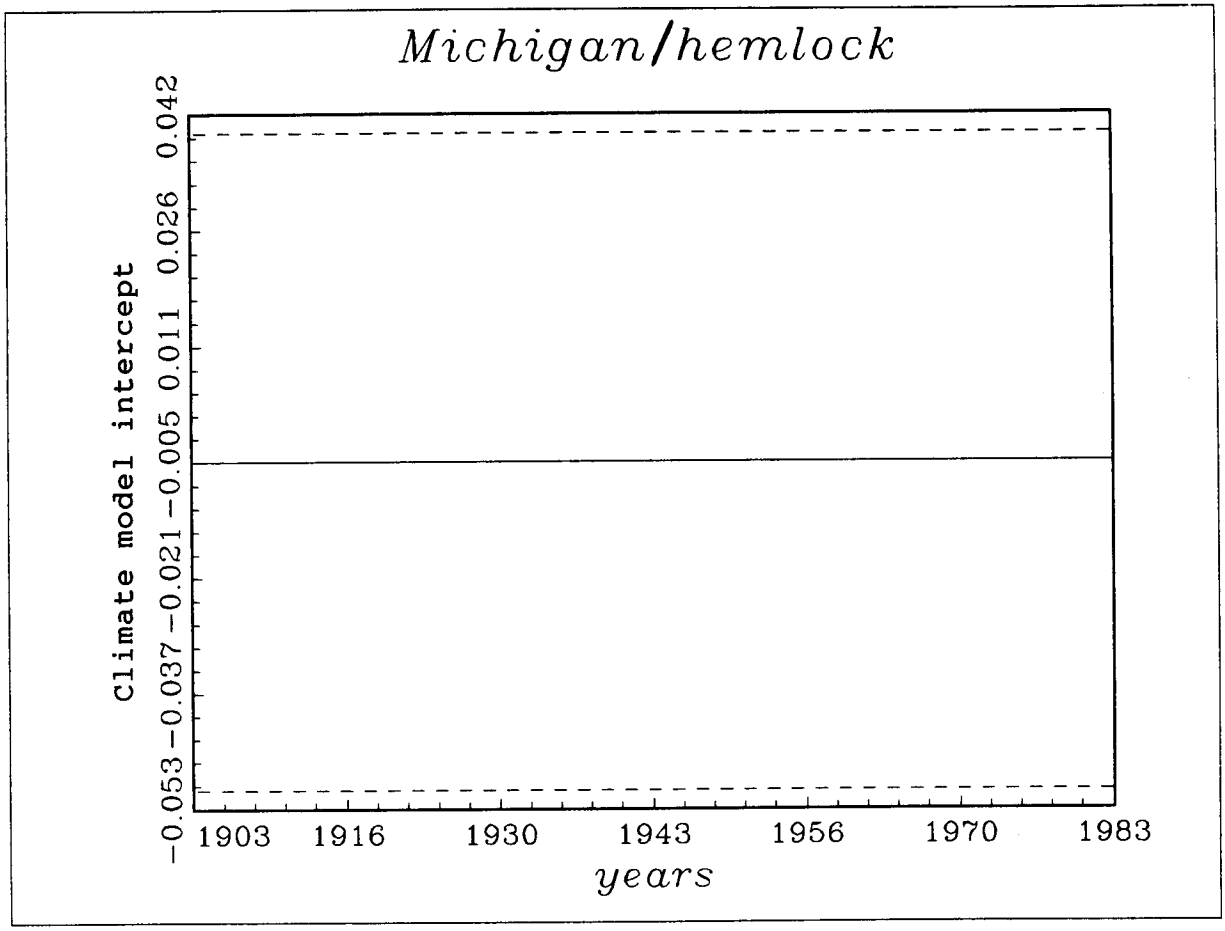
## APPENDIX B - DYNAPLOT PROGRAM OUTPUT



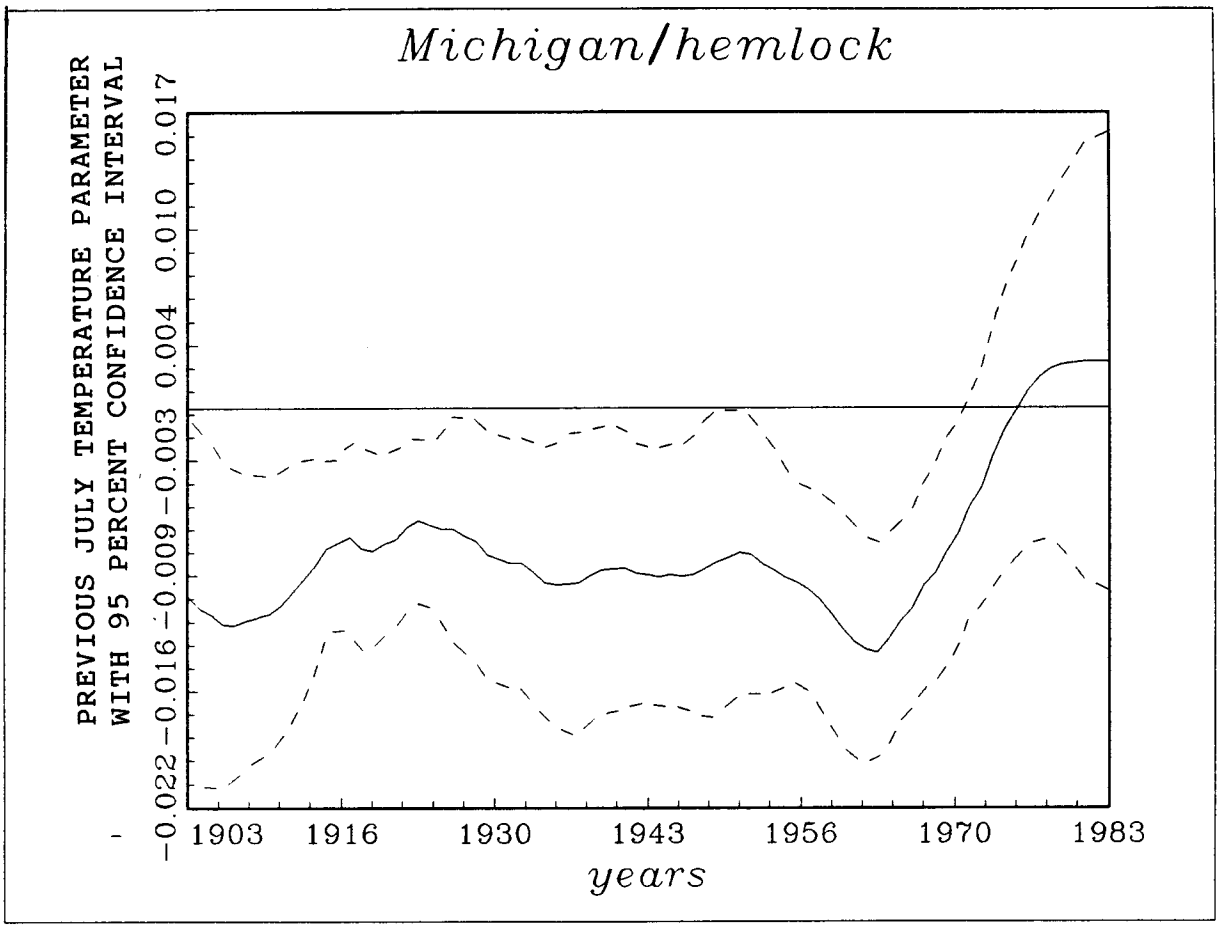

PLOT B-1. This plot was produced by program DYNAPLOT and shows the one step ahead prediction errors surrounded by 95 percent confidence intervals, which are very narrow for this example. The user can supply a title for each graph; Michigan/hemlock was used for this example to denote that the trees were hemlock from Michigan.



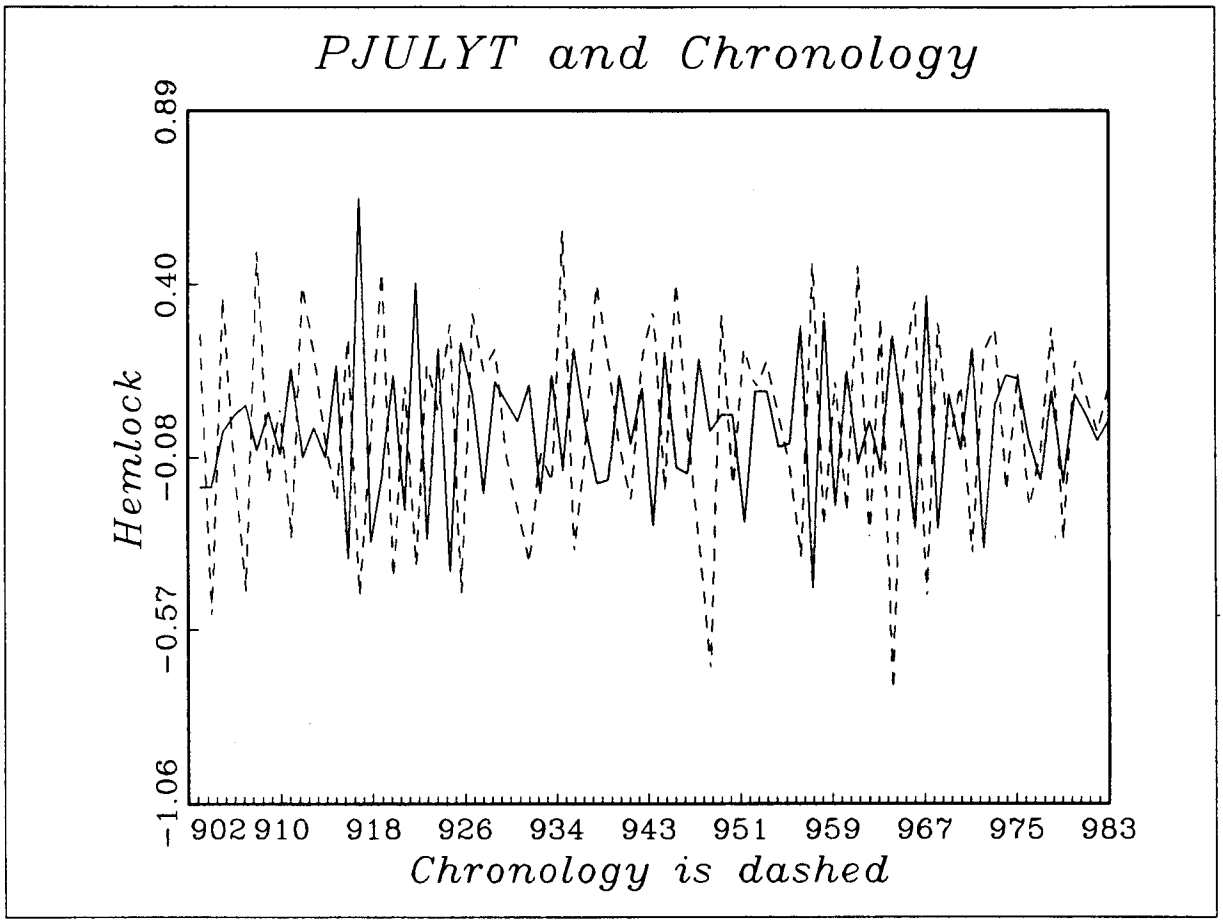

PLOT B-2. This plot was produced by program DYNAPLOT and shows the $m_t$ parameter (equation 2) plotted over time.




PLOT B-3. This plot was produced by program DYNAPLOT and shows the climate model intercept, $b_0$, (equation 2) plotted over time.




PLOT B-4. This plot was produced by program DYNAPLOT and shows the parameter multiplying previous first difference of July temperature plotted over time. This corresponds to $b_1$ in equation (2). The dashed lines are 95 percent confidence intervals.




PLOT B-5. This plot was produced by program OVERLAY and shows previous first difference of July temperature (solid line) overlaid on the average chronology (dashed line) plotted over time. The program allows the user to specify all labels.

## APPENDIX C - CORRELATION PROGRAM OUTPUT

```
 Prediction error cross-correlation matrix
++++++++++++++++++++++++++
 ++++++++++++++++++++++++++
  +++++++++++++++++++++++++++
   ++++++++++++++++++++++++++
    +++++++++++++++++++++++++
     ++++++++++++++++++++++++
      +++++++++++++++++++++++
       ++++++++++++++++++++++
        +++++++++++++++++++++
         ++++++++++++++++++++
          +++++++++++++++++++
           ++++++++++++++++++
            +++++++++++++++++
             ++++++++++++++++
              +++++++++++++++
               ++++++++++++++
                +++++++++++++
                 ++++++++++++
                  +++++++++++
                   ++++++++++
                    +++++++++
                     ++++++++
                      +++++++
                       ++++++
                        +++++
                         ++++
                          +++
                           ++
                            +

C1 - Correlation matrix for lag  0
 Out of  378  independent elements 378  are significant (100 %)
```

Prediction error cross-correlation matrix

```
---------------0------------
----00----------00----00--0--
----00--0-0-00---0-00-0------
-----------------------------
----------0-0---0--0-0------
----0------0---0---0--------
0---00----------0----00-----
0000000-0000000-000000000-00
0--000-----0-0--000-----0-0--0
--0-00------00---000---0000--
00000000000000000000000000000
0---000-00--000-00-00000---0
0--0000------0--00---00--00-
00-0000-0-000-000-0-000----0
----000-000-0--00-----0-0---
000000000000000-000000000000
00000000000000000000000000000
--00000--0000---0--0000---00
----000----00---0-----0-----
00000000000000000000000000000
-----0----------0-----0-----
000000000-000000000-0-00000-
--000-0------0-------------0
0-00000-00--00000000000--000
00000000000000000000000000000
--00000-000000000000000---00
--------0-0-----0--0--------
--0--00-00-0----000--0000---
```

C2 - Correlation matrix for lag 1
Out of 784 independent elements 388 are significant (49 %)

Prediction error cross-correlation matrix

```
0+0+000+0000+0+00000000+000+
000+000+0000+00000000000000+
0000000000000000000000000000
+0++000+0000+0+000+00000000+
++++000+000++++00++++0++0+++
000000000000+000000++0000000
000000000000+000000000000000
00+000000000+000000000000000
00+000000000+000000000000000
++++000+000++++0000++00000++
0000000000-00000000000000000
00000-000000000000000000000
0000000000000000000000000000+
00+000000000+0000000000+0000
00+000000000+0000000000+0000
0000000000-00000000000000000
0000000000000000000000000000
0000000000000000000000000000
0000000000000000000000000000
0000000000000000000000000000
00+000000000000000000000000
00000--00000000000000000000
000000000000+000000000000000
0000--00-0000000000000000000
0000000000-00000000000000000
0000000000000000000000000000
+0++00000000+0+000+0000+000+
+00+0000000000+000000000000+
```

C3 - Correlation matrix for lag 2
Out of 784 independent elements 88 are significant (11 %)

```
 Prediction error cross-correlation matrix
0000000-00000000000000000000
00000000000000000000000000000
00000000000000000000000000000
-0000000000000-000000-00000-
000000000000-0-0000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
-00-0000000-0000000000000000
00000000-00-00-0000000000000
0000000000000+000000000+0000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000000000
00000000000000000000000+00000
00000000000000000000000000000
00000000000000000000000000000
-00-000-000000-000000-00000-
-0000000000000-000000-000000

C4 - Correlation matrix for lag  3
 Out of  784  independent elements  25  are significant (3 %)
 These results are saved in the file corrmtx.out
```

## APPENDIX D - EXAMPLES OF TREE RING DATA FORMATS

Tucson Format - The essential features are: each line except the first begins a new decade; each line consists of an ID variable, then the year followed by up to 10 increments; variables are separated by spaces, and a tree ends with a 999. However, the program works without the 999 ending.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 108011 | 1856 | 190 | 192 | 144 | 161 | . | . | . | . | . | . |
| 108011 | 1860 | 147 | 89 | 96 | 126 | 81 | 86 | 106 | 145 | 103 | 84 |
| 108011 | 1870 | 63 | 58 | 87 | 88 | 72 | 94 | 101 | 60 | 87 | 81 |
| 108011 | 1880 | 93 | 93 | 117 | 127 | 96 | 84 | 83 | 82 | 83 | 107 |
| 108011 | 1890 | 89 | 61 | 61 | 74 | 72 | 64 | 56 | 78 | 57 | 62 |
| 108011 | 1900 | 69 | 69 | 68 | 86 | 82 | 67 | 80 | 80 | 92 | 69 |
| 108011 | 1910 | 92 | 59 | 85 | 97 | 66 | 68 | 50 | 67 | 56 | 77 |
| 108011 | 1920 | 99 | 72 | 61 | 84 | 80 | 69 | 39 | 46 | 45 | 41 |
| 108011 | 1930 | 61 | 47 | 58 | 42 | 24 | 35 | 59 | 42 | 30 | 45 |
| 108011 | 1940 | 44 | 51 | 44 | 20 | 25 | 27 | 19 | 17 | 17 | 8 |
| 108011 | 1950 | 21 | 24 | 31 | 16 | 9 | 23 | 17 | 15 | 18 | 24 |
| 108011 | 1960 | 18 | 18 | 21 | 20 | 16 | 19 | 30 | 14 | 21 | 13 |
| 108011 | 1970 | 26 | 19 | 18 | 14 | 17 | 17 | 27 | 14 | 17 | 11 |
| 108011 | 1980 | 25 | 4 | 24 | 20 | 999 | . | . | . | . | . |

Alternate Format - The same as Tucson format but all lines begin with an even decade. Both of these formats can be easily read by program makedata.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 108011 | 1850 | . | . | . | . | . | . | 190 | 192 | 144 | 161 |
| 108011 | 1860 | 147 | 89 | 96 | 126 | 81 | 86 | 106 | 145 | 103 | 84 |
| 108011 | 1870 | 63 | 58 | 87 | 88 | 72 | 94 | 101 | 60 | 87 | 81 |
| 108011 | 1880 | 93 | 93 | 117 | 127 | 96 | 84 | 83 | 82 | 83 | 107 |
| 108011 | 1890 | 89 | 61 | 61 | 74 | 72 | 64 | 56 | 78 | 57 | 62 |
| 108011 | 1900 | 69 | 69 | 68 | 86 | 82 | 67 | 80 | 80 | 92 | 69 |
| 108011 | 1910 | 92 | 59 | 85 | 97 | 66 | 68 | 50 | 67 | 56 | 77 |
| 108011 | 1920 | 99 | 72 | 61 | 84 | 80 | 69 | 39 | 46 | 45 | 41 |
| 108011 | 1930 | 61 | 47 | 58 | 42 | 24 | 35 | 59 | 42 | 30 | 45 |
| 108011 | 1940 | 44 | 51 | 44 | 20 | 25 | 27 | 19 | 17 | 17 | 8 |
| 108011 | 1950 | 21 | 24 | 31 | 16 | 9 | 23 | 17 | 15 | 18 | 24 |
| 108011 | 1960 | 18 | 18 | 21 | 20 | 16 | 19 | 30 | 14 | 21 | 13 |
| 108011 | 1970 | 26 | 19 | 18 | 14 | 17 | 17 | 27 | 14 | 17 | 11 |
| 108011 | 1980 | 25 | 4 | 24 | 20 | 999 | . | . | . | . | . |

Van Deusen, P. C.; Koretz, J. 1988. Theory and programs for dynamic modeling of tree rings from climate. Gen. Tech. Rep. SO-70. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 18 p.

The effect of climate on tree rings may change over time as a result of stand dynamics or environmental stress. These dynamic effects can be studied using theory and computer programs described herein. An example application is given, and the programs and further information on their use are available from the authors.

**Additional keywords:** Kalman filter, dendrochronology, dendroclimatology.

* U. S. GOVERNMENT PRINTING OFFICE: 1988 - 566019 / 85015