USDA

United States Department of Agriculture

Agricultural Research Service

Technical Bulletin Number 1857

August 1997

Processing of Climatic Data for Detection of Cycles and Trends

In cooperation with the University of Georgia Experiment Station

United States
Department
of Agriculture

Agricultural
Research
Service

Technical
Bulletin
Number 1857

August 1997

Processing of Climatic Data for Detection of Cycles and Trends

In cooperation with the University of Georgia Experiment Station

A.W. Thomas, A.L. Dillard,
W.M. Snyder, W.C. Mills

Thomas and Dillard are agricultural engineers, USDA–ARS, Southeast Watershed Research Laboratory, P.O. Box 946, Tifton, GA 31793. Snyder is hydrologic consultant, Athens, GA. Mills, who was a collaborator on this project, resides in Hendersonville, NC.

Abstract

Thomas, A.W., A.L. Dillard, W.M. Snyder, and W.C. Mills. 1997. Processing of Climatic Data for Detection of Cycles and Trends. U.S. Department of Agriculture, Technical Bulletin No. 1857, 46 pp.

Cycles and trends are considered structural information in historical data which, if detectable and quantifiable, can be used in simulating future climates. Three data-processing models were applied to climatic records to test detectability of such parameters of climatic change. One was the conventional ARIMA model. The other two were developed and tested for this study. The first model detected cycles of varying wavelength in data. It tested successfully with monthly data with known seasonal cycle. The second model detected persistence of high and low values during cycling, and included a time-trend component. This model tested successfully against purposefully designed synthetic data. The three models were applied to two rainfall and two temperature records in Georgia. No model detected any significant structure in these climatic records of 52–104 years. It is concluded that short-term climatic simulation for risk and uncertainty analysis in agricultural planning in this region need not presently include any time-shifting parameters. However, since slight climatic change may not be detected in single-site analysis, future research is suggested for multisite analyses. This publication will be useful to anyone concerned with examining climatic records for natural and human-induced cycles and trends.

Keywords: global climate change, least squares, modeling, sliding polynomials, time series

While supplies last, single copies of this publication may be obtained at no cost from USDA–ARS, Southeast Watershed Research Laboratory, P.O. Box 946, Tifton, GA 31793.

Copies of this publication may be purchased from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161; telephone 703–487–4650.

Issued August 1997

Contents

One of today's dominant environmental concerns is the changing climate. In particular, two publicized and popularized concerns are the greenhouse effect (which is due to the increase of carbon dioxide in the atmosphere) and increased ultraviolet radiation (due to a diminished ozone layer). The question of whether or not climatic changes are detectable in historical records is still somewhat controversial. This controversy diminishes the motivation for formulating and implementing programs for environmental protection.

Time-series analysis of historical records is a longstanding, vast, and complex topic. To some degree, interest in this topic dates from the time people first began making systematic observations and keeping records. In any new study, one cannot hope to treat such a subject exhaustively; one can only suggest methodology for specific and limited purposes. Briefly, historical series are analyzed in order to find some information structure in the record. If such structure—usually a cycle or a trend—is found, it is hoped that it can be used to improve the prediction of future items in the series.

Examples of current interest in cycles and predictability are easily found. The cover of *Eos* (1989) shows daily sunspot numbers from January 8, 1818, to December 31, 1988. *U.S. News & World Report* (Budiansky 1989) contains an article that reports the finding of linkage between solar energy and climate. A *Reader's Digest* article contains a popularized discussion of some controversial aspects (Bidinotto 1990). In *Eos Supplement* (1995), the American Geophysical Union lists the Committee on Global Environmental Change as one of several in its currently recognized research programs. Articles on methodology, case studies, and predictability for planning can be found in many journals in meteorology, water resources, and environmental sciences.

The responsibility and interest of the U.S. Department of Agriculture (USDA) were outlined in 1990 with the development of a strategic plan for the various affected agencies (USDA 1990). The objective of the plan was to establish a long-term strategy for research, education, technology transfer, extension activities, and policy analysis to assist in sustaining U.S. agriculture and forestry that are compatible with future global change. The Agricultural Research Service (ARS) followed USDA's lead by presenting its climate and hydrology research component of the global-change research program (USDA 1992). ARS targeted key areas of research within the hydrologic community where significant knowledge gaps still exist—gaps that must be filled before answers to specific global-change questions can be determined.

The end product of time-series inquiry—successful predictability—must depend on either finding a link to terrestrial or extraterrestrial events that are predictable, or finding predictable information in the series itself. In either case, successful prediction in the practical sense means that stochastic variability of the future series is kept significantly smaller than if the events in the series were purely random. The predominant methodology in time series is a composite of autoregressive and moving-average techniques, which are mostly dependent on the work of Box and Jenkins (1976). An overview of time-series analysis can be found in Hipel (1985), Yevjevich and Harmancioglu (1985), and Lewis (1985). An earlier work by Fiering and Jackson (1971) remains a readily understandable and practical introductory analysis and simulation approach. Snyder (1976b) modified this approach by the use of sliding polynomials to produce free-form seasonal patterns of the autoregression coefficients.

The first step in the development of a time-series predictive system is the identification of usable information in historical records of the series. As stated earlier, such information normally consists of cycles or trends in the data. It seems advantageous, then, to have a number of data-processing models, each of which can screen the record in varying ways, searching for varying forms of information. Present needs in prediction are not limited to simple simulation of possible future series of the events. Such simulation does provide the basis for planning for resource conservation and development.

But researchers may have other needs. For example, when field data were obtained on soil-plant-water relationships, was the climate representative or were the data biased by some climatic extreme? When climatic generators of input to crop yield-erosion models are calibrated, are the parameters affected by abnormalities in existing climate records? Can the results of time-series analysis be transported to other places and times? Can quantified parameters from the analysis be associated with regional physical processes?

Thomas and Snyder (1984) addressed questions like these. They developed a probability index through simulation that can be used to judge the climatic representativeness of short periods such as the duration of a field research project.

Objective

It is the objective of this publication to screen climatic records with three different data-processing models. The purpose of such screening is to test in various ways for cycles and trends in climatic data—in ways that would be useful in appraisal of and prediction with results from field research. Such an objective has some obvious overlap with climate-change modeling. Records of about 50–100 years are used because this length is considered representative of the bulk of climate records, and the screening objective is intended for testing with such representative record length.

Screening Models Used

One of the screening models used is the well-known ARIMA model of Box and Jenkins (1976). The other two are special-purpose models developed and programmed for this study. It is hoped that the demonstration of model concepts other than ARIMA will stimulate additional special-purpose modeling and provide useful alternatives in time-series analysis. Description of the two special-purpose models follows.

CYCLES Model

The CYCLES model was developed to screen a record for trends and cycles. Thus two separate traces are extracted: a trend line and a cycle pattern. Cycles of varying length are identified and extracted to describe the shape of the cycle as a free-form sliding polynomial function (Snyder 1976a). The time-trend line is extracted simultaneously from the data. The structure and operation of the model are shown schematically in figure 1.

Four nodes are used to define a segmented time trend (see figure 1, part a). The nodes are placed at the quarter points of the record length. The values of these nodes are derived by least squares. Myriad patterns of time trend are possible; two are shown for illustration in figure 1, part a. A true trend may be found, or long cycles may be found. The shortest detectable cycle would be one-half the record length.

A varying number of nodes are used to define a free-form cycle in the data. The number of nodes is increased during successive sweeps through the data. The first three sweeps are shown in figure 1, part b. The smallest number of cycle nodes is two. They are placed 2 years apart, beginning with year 1 of the record. As shown, the pattern of the two nodes repeats through the full record. The four time-trend nodes and the two cycle nodes are evaluated simultaneously by least squares. The optimization process always uses all years of the record, not just the years where the nodes are located. Sweep 1 defines a cycle with a length of 4 years.

On the second sweep, the number of cycle nodes is increased to three, and this pattern repeats through the entire record. The four time-trend nodes and the three cycle nodes are evaluated simultaneously. Sweep 2 defines a cycle with a length of 6 years. Note that the derived nodal values give the shape of the cycle.

On the third sweep, the number of cycle nodes is increased to four. Optimization by least squares yields simultaneous values of the four trend nodes and the four cycle nodes. Sweep 3 defines a cycle with a length of 8 years.

The sweeps through the record are continued through a cycle based on 12 nodes plus the time trend of 4 nodes. All cycle lengths from 4 years to 24 years in steps of 2 years are thus evaluated. Note that a 22-year cycle can also express two 11-year cycles. The computer program implementing the model performs the 11 sweeps without operator intervention. On each sweep the statistics of the fit are evaluated, and the best fit can be selected. Obviously, operating parameters other than those described here can be chosen. This first-generation program was designed primarily for detection of whole-year cycles. Cycles with fractional year lengths can be investigated with some modification because, for example, an 11-year cycle can actually be two 5 1/2-year cycles.

FFAUTO Model

FFAUTO is a model based on free-form autoregression. In conventional autoregression, the relationship of each item in a time series is fixed with respect to the preceding items. FFAUTO allows each individual item to have its own autoregression function, and this function is dependent on the average value of a number of antecedent items in the series. A time trend is included so that processing a record with FFAUTO evaluates simultaneously a time trend and a free-form autoregression function.

The FFAUTO model is expressed as follows:

$$x(t) = \sum_{i=1}^{4} c_i(t)T_i + b_1 * x(t-1) + \sum_{j=2}^{15} b(j,\bar{x}_j) * x(t-j), \qquad (1)$$

where

- x = an individual item of the series,
- t = time,
- T_i = a time-trend node,
- $c_i(t)$ = a linear interpolating coefficient between nodes adjacent to t,
- b_1 = first autoregression coefficient, and
- $b(j,\bar{x}_j)$ = free-form autoregression coefficient with dependence on lag time j and antecedent average of $x, \bar{x}_j$, shown in figure 2.

Basically, $b(j,\bar{x}_j)$ is given as a set of two-dimensional sliding polynomial contours in the j- $\bar{x}_j$ space. The contours are defined on eight nodes, of which nodes 1 and 5 are required to be zero to make the autoregression function go to zero at a 15-year lag time. The spacing of nodes in j scale is a simple exponential transform. Values of $\bar{x}_j$ are a backward-growing average. When j = 2, then $\bar{x}_j = x(t-1)$. When j = 3, then $\bar{x}_j = \{x(t-1) + x(t-2)\}/2$. When j = 4, then $\bar{x}_j = \{x(t-1) + x(t-2) + x(t-3)\}/3$, and so on.

The rationale behind equation 1 is as follows: Item x(t) in a series can have a serial dependence on x(t-1); hence we have the term $b_1\, x(t-1)$ in equation 1. Now item x(t) can also have serial dependence on x(t-2). However, this dependence might depend on the intervening value x(t-1). Perhaps persistence patterns develop in a series so that high values tend to beget a string of high values. Then $b_2(2,\bar{x}_2)$ would be large for large $\bar{x}_2$.

Such persistence can be expressed by the moving-average part of an ARIMA model, but in such a model it is a simple linear additive to the autoregression function. FFAUTO allows an interactively nonlinear expression of the possible persistence patterns. Evaluation of the FFAUTO parameters consists of obtaining optimized values of the four time-trend nodes T_i, the regression coefficient b_1, and the six non-zero nodes in figure 2. The structure is linear in all the parameters; therefore, least squares was attempted as the optimizing technique. But it was discovered that the intercorrelations of the coefficients of the six non-zero nodes of figure 2 were so high that the multivariate technique of components regression had to be used.

CYCLES Program Testing

Prior to its use in screening climatic records, the CYCLES program was tested on monthly rainfall and runoff records from two research watersheds. Thus the response of the model to the known seasonal cycle of 12 months could be observed. Such testing also helped to eliminate errors in the computer program. The two test watersheds were White Hollow in eastern Tennessee (TVA 1961) and Pine Tree Branch in western Tennessee (TVA 1962).

The records of monthly rainfall and runoff for White Hollow are shown in figure 3, and those for Pine Tree Branch are shown in figure 4. Superimposed on all the records are the monthly values of the fitted CYCLES model with a 12-month cycle. The monthly rainfall values are moderately cyclic whereas the runoff values are strongly cyclic. The residual error standard deviations for the 11 sweeps through the data are plotted in figure 5. It is clear from these plots that the sweep search does identify the natural 12-month cycle by minimum residual error. The 24-month cycle also identifies the natural 12-month cycle.

Some goodness-of-fit statistics for the four test data sets are given in table 1. The standard deviations of the errors residual to fitting the CYCLES model with a cycle length of 12 months are compared with the standard deviations of the data. It should be noted that the reduction in the standard deviation is small, amounting to only 7–10 percent for monthly rainfall and 22–27 percent for monthly runoff. Such small reduction indicates that simulation of future series with a CYCLES model plus randomization with the standard deviation of the residuals will yield a series with little less scatter than purely stochastic simulation. This in turn means that large uncertainties will remain when such simulated future series are used in planning. Stated another way, the precise detection of cyclicity in a series will not automatically yield a good predictor for this series.

Figure 6 shows the seasonal cycles of rainfall and runoff identified by sweeping the data with the CYCLES model. These cyclic patterns are similar to—but not identical with—simple arithmetic monthly averages since they are derived by smoothing. Also, they are influenced by the time trend simultaneously identified. Pine Tree Branch cyclic runoff shows some negative values, but again, these must be balanced with the identified time trend. The predicted runoff values in figure 4 do show a few negative values, and these must be judged a failure of the model. In practical work, such values are set to zero (Fiering and Jackson 1971).

The time trends of rainfall and runoff for the two test watersheds are shown in figure 7. The fitting of the model produced similar trends for rainfall and runoff for White Hollow and for Pine Tree. The trends are all measured relative to zero at the beginning of the record. The Pine Tree record starts during a period of low rainfall during the early 1940's. Therefore, its time trends are positive relative to these initial values. Both watersheds show a peak in the middle to late 1940's, with an essentially downward trend to the end of the record.

Both White Hollow and Pine Tree Branch were research watersheds operated for inquiry into forest-streamflow relationships. Both watersheds received conservation treatment consisting mainly of gully control and tree planting (TVA 1961, 1962). The time of the treatments is shown in figure 7. The CYCLES model does not reveal any shift in the relationship of the rainfall-runoff time trends with developing forest following treatment. The magnitudes of the time trends are generally less than the standard deviations of data shown in table 1.

In summary for this section, we conclude that the computer program of the CYCLES model is functioning properly. The model identifies the known seasonal cycle of monthly rainfall and runoff data. Time trends identify major shifts in rainfall. The four independently derived time trends are mutually supportive.

FFAUTO Program Testing

The FFAUTO program was tested on three synthetic data sets. These data sets are shown in figures 8–10. Also shown in these figures are the results of fitting the FFAUTO model. For comparison, the CYCLES model was fitted to the three synthetic sequences, and these results are also plotted.

Synthetic set no. 1 was generated with an 11-year cycle plus a time trend. Set no. 2 was a modification of set no. 1, consisting of broadening valleys and peaks in an arbitrary fashion to enforce a slight persistence on the data. Variability of the data was also reduced slightly. Set no. 3 was constructed with broad plateaus imposed on a major, irregular cycle to enforce a larger degree of persistence in the time sequence of the numbers.

Both the CYCLES and FFAUTO models fit the data reasonably well in figures 8 and 9. However, in the highly persistent sequence in figure 10, the FFAUTO model is a noticeably better fit.

The statistics of fitting the two models to the three synthetic data sets are given in table 2. The standard deviations of the original data are not greatly different. However, the performance of the two models is different. In data set no. 1, which has no pattern of persistence built in, CYCLES fits better than FFAUTO, with a reduction of standard deviation of 54 percent versus 41 percent. In set no. 3, with a significant amount of persistence, FFAUTO fits much better than CYCLES, with a 75 percent reduction in standard deviation versus 44 percent. For data set no. 2 with only slight persistence in the data, there is little difference in model performance.

In the FFAUTO model, each item in the series after the second item has its own individual autoregression function. As shown schematically in figure 2, the autoregression coefficients $b(j,\bar{x}_j)$ are a function of lag number j and the average of antecedent numbers in the series, $\bar{x}_j$. Selected autoregression functions for the three synthetic test data sets are shown in figure 11. The functions with the highest and the lowest regression coefficients for lag no. 2 are plotted. All other autoregression functions in the respective data sets will lie between these two. Remember that coefficient number one is the same for all autoregression functions in a set.

Set no. 1 with its essentially cyclic structure shows the largest amplitude of cyclic structure in the autoregression function in figure 11. The cycle length is 10 lags. Note in figure 8 that the lag distance between peaks or troughs is not always exactly 11 as built into the data set. Rather, it varies slightly due to the included stochastic component. The autoregression function in figure 11 predicts cyclic lengths of 10, 11, or 12, from peak to peak. The autoregression functions for set no. 2 in figure 11 show less cyclic amplitude than those for set no. 1, except that the fixed coefficient for lag no. 1 is much larger. Also the difference between the highest and lowest coefficient for lag no. 2 is much greater for set no. 2. These results are consistent with the structure of the synthetic data set no. 2. This set has some persistence structure in its broader valleys and peaks compared to set no. 1. The purely cyclic content of the data is thus reduced. The autoregression functions vary through the data set. In synthetic data set no. 3 with the large amount of imposed persistence, the difference between the regression coefficients for lag no. 2 is large. The autoregression functions will vary greatly through data as the antecedent average, $\bar{x}_j$, in figure 2 varies greatly due to the long sequences of high or low values of the sequential items x.

The positions of the autoregression functions with highest and lowest regression coefficients number two are marked in figures 8–10. Generally, they occur with the highest and lowest data values except as they are modified by the time-trend component of the model. In figure 10, the highest and lowest $b(2,\bar{x}_2)$ are the same for successive times with plateaus of equal values of x.

Figure 12 displays the time trends derived from the three synthetic data sets by application of the two models. The time trend incorporated in data set no. 1 is shown for comparison. This time trend is also present in data set no. 2, since set no. 2 was a persistence modification of set no. 1. However, no attempt was made to ensure that these modifications did not corrupt the trend incorporated in set no. 1. No attempt was made to incorporate or to exclude a trend in set no. 3.

Caution is necessary in how time trends are visualized. If data are not highly variable, then any time trend present should be readily found during data processing. If data are highly variable, then a trend may be difficult to quantify. The true trend may be a small shift in mean value. However, the derived analytical trend may be conditioned by the accidental occurrence of a few extreme events with large errors residual to model fitting. In table 2, the greatest reduction in standard deviation occurred with FFAUTO applied to set no. 3. The time trend in figure 12 for this model is small. The time trend for CYCLES applied to this set is large, and considering the fit in figure 10 is probably not as reliable. The second greatest reduction in standard deviation in table 2 resulted from processing data set no. 1 with CYCLES. The time trend shown for this fitting in figure 12 agrees reasonably well with the time trend incorporated in the data. The third greatest reduction in table 2 is for FFAUTO applied to set no. 2. This produced the very small time trend in figure 12. CYCLES applied to set no. 2 achieved nearly as great a reduction in standard deviation, yet produced a different time trend shown in figure 12. This difference illustrates the caution needed in interpretation of a weak time trend.

The time trends plotted in figure 12 have variations that are about the same magnitude as the standard deviations of the residuals given in table 2. The exception is the time trend for CYCLES applied to data set no. 3. This exception needs explanation to achieve confidence in the FFAUTO model and program and confidence in its comparison with the CYCLES model. Figures 13 and 14 were prepared to provide this explanation.

The cycles derived from the CYCLES program applied to the three data sets are shown in figure 13 for the "best fit" of the successive sweeps through the data. For sets 1 and 2, the cycles are pronounced. Although the cycle width is 22 periods, these cycles are actually composed of two 11-period half-cycles, in agreement with the construction of the synthetic data sets. The derived cycle in figure 13 for data set no. 3 is much less pronounced. The cycle amplitude is only about one and one-half units. Therefore, the reduction in standard deviation of 44 percent in table 2 must be due to the time trend rather than the cyclic component of the model, which is confirmed by the plots of residual error in figure 14.

The residual errors for synthetic data set no. 1 in figure 14 have uniform variability throughout the length of the series of data. No residual cycling appears. No relationship exists between the errors and the location of the nodes of the time-trend function. The residual errors from fitting FFAUTO to set no. 3 are relatively small and also have no relationship to the location of the nodes of the time-trend function. However, the errors residual to fitting CYCLES to set no. 3 are large and are highly related to the time-trend nodes. With this perception it is possible to see in figure 10 that the fitted CYCLES is actually a very small cyclic function superimposed on the segmented time trend.

In summary, the free-form autoregression model, FFAUTO, tested satisfactorily when applied to synthetic data sets designed with varying amounts of tendency toward persistence of high or low values in a time series. Comparison with the previously tested CYCLES program produced rational and explainable differences. The derived time trends of both models must be carefully examined. The piecewise time trends may actually represent long-wave cycling.

Processing of Historical Records

Four data sets were processed in the search for cycles and trends in natural or environmental data. These four data sets were examined using the CYCLES, FFAUTO, and ARIMA models.

The data sets are described briefly in table 3. Annual totals of rainfall and annual average temperatures were studied at Watkinsville in the Georgia Piedmont and at Tifton in the Georgia Coastal Plain. The record lengths varied from 52 to 104 years. The last year of record was 1988 for all four stations. Figures 15–18 display the four data sets. Superimposed on the data are the results of fitting the CYCLES and FFAUTO models.

Figures 15 and 16 both show a major pattern of temperature variation. Temperatures were high from 1940 to 1950, were low from 1960 to 1970, and rose again following 1970. The rise from the 1970 low was more pronounced at Watkinsville than at Tifton. Because the temperature records contain only the one incomplete wave, no conclusions can be drawn about cycles or wavelengths.

Figure 17 shows no major shifts in rainfall at Watkinsville. The only outstanding feature is the four successive years of low rainfall ending in 1988, which is the only such occurrence in the 104-year record. Therefore, little reliance can be placed on the predictability of such events in the future. Figure 18 shows that little shift has occurred in annual rainfall at Tifton, but a pronounced change in variability of rain may be noted. From the beginning of record until about 1965, wide swings of rainfall occurred. Exceptional high and low values were recorded, superficially—much the same as at Watkinsville. From 1965 to 1988, however, no exceptionally high or low values occurred.

A brief statistical summary of results of fitting the three models to the four data sets is given in table 4. Generally, all three models follow the same pattern of results. Significant reduction of standard deviation resulted for all models for the Tifton temperature record, but little reduction for the Watkinsville rainfall record. The ARIMA model had the least reduction in standard deviation except for the Tifton rainfall.

FFAUTO Model

Figures 15–18 show that the FFAUTO model tends to follow a smoothed average value of the data. It does not predict well the high and low swings of the data. This means that the adjustments of variability shown in table 4 are due more to the time-trend portion of the model, rather than to any autoregressive structure. The derived time trends, which are straight-line segments by quarters of the record length, are shown as solid lines for the temperature data in figure 19 and for the rainfall data in figure 20.

The time trends shown for FFAUTO analysis of temperature in figure 19 do not show definitive structure. A slight tendency is shown for low values in the middle of the records and for higher values toward the end of the record, but the overall correspondence with the clearly visible partial wave in figures 15 and 16 is not good. The time trends for rainfall in figure 20 are more understandable. No significant time trend is visible in figure 17, and the trend line in figure 20 correspondingly shows little change. The high values of time trend for Tifton rainfall at 1972 in figure 20 is not inconsistent with the absence of years with very deficient rainfall beginning around 1955. Also, the years following 1970 tend toward lower amounts except for the year 1983.

Examples of the free-form autoregression functions are shown in figure 21 for annual temperatures and in figure 22 for annual rainfall. The regression functions shown are those with the highest and the lowest regression coefficient at lag 2. The separation of these functions indicates that some persistence is present in the record. Also, for Tifton, the regression functions show a tendency to cycle at a 9-year lag. However, the coefficients are small, and no visible cycles are generated in the FFAUTO curves superimposed on the data in figures 15–18.

CYCLES Model

The results of smoothing with the CYCLES program (figs. 15–18) show that the CYCLES model predicts with wider swings than does the FFAUTO model. The exception is the Tifton annual rainfall in figure 18, where the FFAUTO model also has wide swings of predicted values. It should be noted again that for this data set the FFAUTO model produced a significant reduction in standard deviation, whereas the CYCLES model did not (table 4).

The residual standard deviations for all cycle searches for the four data sets are shown in table 5. Noted in this table are the cycles with minimum residual standard deviation as shown in table 4. In table 5, there are no clear and distinct "best" cycle lengths such as shown for the test of monthly data shown in figure 5. For example, the range from high to low residual standard deviations for Tifton rainfall (10.188–9.576 = 0.612 inch) is very small compared to the 10.267 inches for the original data.

The nodal values for increasing cycle length are shown in table 6 for Tifton annual temperature. Very little change of amplitude, or wave form, can be found here. The position of the largest and smallest nodes in the cycle move in a haphazard manner with changing cycle length.

The minimum residual error for Tifton annual temperature was found with the 24-year cycle length (table 5). The nodes for the solution are plotted in figure 23. This figure clearly illustrates the lack of any definitive cycling of temperatures.

The time-trend portions of the CYCLES model are shown with the FFAUTO trends in figures 19 and 20. In figure 19 the temperature trends for the CYCLES model for Watkinsville and Tifton are in agreement, as the FFAUTO runs for the two locations were in agreement. In general, all four temperature analyses tend to outline the long cycles of temperature shown in figures 15 and 16. Differences are due to the different record lengths and to the happenstance way in which quarters of the record length can correspond to the long-cycle shape.

In figure 20 the CYCLES time trends tend to match the FFAUTO time trends for the two locations. Little trend is noted for the long Watkinsville rainfall record. A high-rainfall period is noted for the early 1970's for Tifton annual rainfall.

None of the derived time trends indicate a clear single-directional shift in either annual temperature or rainfall.

ARIMA Model

The integrated, autoregression, moving-average analyses were performed using commercial software that follows Box and Jenkins (1976). It is difficult to make direct comparisons between the ARIMA results and the FFAUTO and CYCLE results. The ARIMA approach appears to be more concerned with forecasting future occurrences over some lead time and to be less concerned with abstracting structural information from a data set to test a physical hypothesis. For example, ARIMA does not incorporate a specific time-trend structure to parallel a direct relationship between succeeding elements in a series. Rather, one must construct differences between the elements in the series and then correlate the differences. Specifically then, one is studying the series of differences, not the series of events. The integration of the difference structure yields a single linear time trend.

ARIMA analysis is performed in three steps: identification, estimation, and forecasting. Step 1 is identification. In this step, autocorrelation of the original series is used to decide on the model structure. This means deciding how many autoregression coefficients to use, and whether to use the original series of elements or difference of the elements. One also decides whether "seasonal" coefficients should be used. Rules of model selection are given by Box and Jenkins (1976).

Step 2 in analysis is estimation of the parameters of the model selected in step 1. Step 3 is forecasting future elements of the series for some specified lead time.

In the identification step, it was determined that differencing should be used on the temperature data for both Watkinsville and Tifton in order to express a time trend. The need for differencing was not indicated for the annual rainfall data. Consequently, ARIMA was used for temperature data, but the simpler ARMA was used for rainfall. The autocorrelation structure indicated that two autoregression coefficients and two moving-average coefficients should be used for the Watkinsville temperature data. One autoregression coefficient and one moving-average coefficient were selected as the model for the other three data sets.

The results of analysis of the four data sets are given in table 7. No clear and definitive pattern is evident in the results. For example, in the identification step it appeared that differencing should be used for the temperature data, yet the constants of the model are far from significant. Autoregression coefficients are significant only for Watkinsville temperature, while moving-average coefficients are significant for temperature but not for rainfall. It was mentioned earlier that the residual standard deviations in table 4 are highest for the ARIMA model except for Tifton rainfall. In contrast to this, the autocorrelation of the residual errors are lowest for the ARIMA model (shown in figs. 24–27). These figures show that the ARIMA model is the best of the three at reducing the residuals to pure noise. The coefficients for the ARIMA model are well within the error limits of 2 standard deviations. The FFAUTO and CYCLES coefficients lie mainly within the error limits. An occasional value does lie outside the limits, but a consistent pattern to these larger values is not evident.

Figure 27 requires separate consideration. Most of the autocorrelation coefficients for the CYCLES model lie outside the error limits. However, these larger values simply oscillate with each unit change of lag. It is noted in table 4 that the Tifton rainfall data cannot be represented by the CYCLES model; standard deviations were reduced by only 7 percent. The annual oscillation of the coefficients in figure 27 does not have a rational basis, and because of the nonconformance of data and model, it is judged to be statistical nonsense.

Time trends for the ARIMA models are not explicitly expressed. Therefore, forecasts for a lead time of 20 periods were prepared. These forecasts contain little information. For the two rainfall records, the forecasts are essentially 20 repetitions of the average rainfall. Watkinsville temperatures are forecast to decline from 65.8 °F to 65.1° in the 20 years following 1988. Tifton temperatures are forecast to decline from 61.2° to 60.4° in 20 years. The slightly shifting temperatures are, of course, the result of performing autocorrelations on the first differences of the temperature data. Therefore, ARIMA yields a decline of temperature of 0.07° per year through the record for Watkinsville and 0.08° per year through the record for Tifton. The acceptability for these simple linear trends must be judged against the cyclic appearance of the data in figures 15 and 16.

Summary

Four data sets were screened for the detection of possible cycles and trends. Three different models were used. Two models were developed for this study. The third was the conventional ARIMA model of Box and Jenkins (1976).

The two models developed for this study were first tested against natural and synthetic data to assure that they would perform properly in known conditions. This testing was successful in deriving known cycles and trends in the data.

Processing of the four data sets by the three models did not produce sufficient consistency of results to support any conclusion that predictable cycles and trends had been detected. It is evident that some long-term shifts in mean value or a change of variability does exist. Such can be seen in simple plots of the data. But it has not been possible to demonstrate that any regular cyclicity exists, and obviously, irregular cyclicity cannot be used to improve predictability. It was also not possible to demonstrate that consistent time trends are present, since the noted shifts cannot be separated into true trend or portions of long irregular cycles.

The inability to demonstrate conclusively the presence of cycles or trends in the data sets analyzed cannot be used as an argument that such do not exist. Rather, this result should be considered a demonstration of the complexity of the problem. Natural data are highly variable, and the derived cycles and trends contribute only a small amount to this variability and hence contribute little to improved prediction. It is hoped that the demonstration of model concepts other than ARIMA will stimulate additional special-purpose modeling and in the future will provide useful alternatives in time-trend analysis.

An announcement in the *New York Times*, December 20, 1994, stated that new satellite data showed that the sea level had risen about 6 millimeters over the previous 2 years. This was based on measurements at 500,000 places per day. Such data, compared to the results in this study, likely indicate that climate change may not be readily detectable in single-site analysis. Additional research should focus on multisite analyses and further improvement in model structure.

References

American Geophysical Union. 1989. Eos, Transactions, American Geophysical Union 70:32.

American Geophysical Union. 1995. Eos supplement, AGU handbook. American Geophysical Union 76:5.

Bidinotto, R.J. 1990. What is the truth about global warming? Reader's Digest, February 1990, p. 814.

Box, G.E.P., and G.M. Jenkins. 1976. Time series analysis: Forecasting and control. Holden-Day, Oakland, CA.

Budiansky, S. 1989. Sunny weather (literally). A long-sought link between the sunspot cycle and climate. U.S. News & World Report, March 6, 1989, pp. 52-54.

Fiering, M.B., and B.B. Jackson. 1971. Synthetic streamflows. American Geophysical Union, Water Resources Monograph 1, Washington, DC.

Hipel, K.W. 1985. Time series analysis in perspective. *In* Time Series Analysis in Water Resources, pp. 609–624. American Water Resources Association Monograph Series No. 4, Bethesda, MD.

Lewis, P.A.W. 1985. Some simple models for continuous variate time series. *In* Time Series Analysis in Water Resources, pp. 635–644. American Water Resources Association Monograph Series No. 4, Bethesda, MD.

Snyder, W.M. 1976a. Interpolation and smoothing of experimental data with sliding polynomials. Agricultural Research Service, U.S. Department of Agriculture, ARS–S–83.

Snyder, W.M. 1976b. Time series data analysis and synthesis for research watersheds. Agricultural Research Service, U.S. Department of Agriculture, ARS–S–76.

Tennessee Valley Authority. 1961. Forest cover improvement influences upon hydrologic characteristics of White Hollow watershed 1935–1958. Tennessee Valley Authority, Knoxville.

Tennessee Valley Authority. 1962. Reforestation and erosion control influences upon the hydrology for the Pine Tree Branch watershed 1941–1960. Tennessee Valley Authority, Knoxville.

Thomas, A.W., and W.M. Snyder. 1984. Testing the climatic representativeness of short-period records through simulation. Transactions of ASAE 27:1027–1033.

U.S. Department of Agriculture. 1990. USDA global change strategic plan.

U.S. Department of Agriculture, Agricultural Research Service. 1992. Global change, water resources and agriculture. NWRC 92–3.

Yevjevich, V., and N.B. Harmancioglu. 1985. Past and future of analysis of water resources time series. *In* Time Series Analysis in Water Resources, pp. 625–634. American Water Resources Association Monograph Series No. 4, Bethesda, MD.

Table 1. Statistics of fit for monthly rainfall and runoff

Data set	Standard deviation of data (inches)	Standard deviation of residuals (inches)	Reduction in standard deviation (inches)	Percent reduction
White Hollow rainfall	2.304	2.150	0.154	7
Pine Tree rainfall	2.563	2.298	.265	10
White Hollow runoff	1.553	1.127	.426	27
Pine Tree runoff	0.963	0.749	.214	22

Table 2. Statistics of fit for synthetic data sets

Data set	Data	Residuals	Reduction	Percent reduction
	(Standard deviation units)			
FFAUTO				
1	2.113	1.257	0.856	41
2	2.034	1.152	.882	43
3	2.149	0.531	1.618	75
CYCLES				
1	2.113	.980	1.133	54
2	2.034	1.200	.834	41
3	2.149	1.198	.951	44

Table 3. Data sets analyzed for cycles and trends

Set number	Location	Data	Record length (years)
1	Watkinsville, GA	Average annual temperature	52
2	Tifton, GA	Average annual temperature	64
3	Watkinsville, GA	Average annual rainfall	104
4	Tifton, GA	Annual total rainfall	64

Note: 1988 is last year for all records.

Table 4. Statistical summary of fittings

Data set	Standard deviation of data	Standard deviation of residuals	Reduction	Percent reduction
FFAUTO				
1	1.138 °F	0.940 °F	0.198 °F	17
2	1.494 °F	.891 °F	.603 °F	40
3	8.814 inches	8.353 inches	.461 inch	5
4	10.267 inches	7.919 inches	2.348 inches	23
CYCLES				
1	1.138 °F	.917 °F	.221 °F	19
2	1.494 °F	.898 °F	.596 °F	40
3	8.814 inches	8.559 inches	.255 inch	3
4	10.267 inches	9.576 inches	.691 inch	7
ARIMA				
1	1.138 °F	.957 °F	.181 °F	16
2	1.494 °F	.999 °F	.495 °F	33
3	8.814 inches	8.700 inches	.114 inch	1
4	10.267 inches	8.373 inches	1.894 inches	18

Table 5. Variation of residual standard deviation with cycle length

	Residual standard deviation			
	Temperature (°F)		Rainfall (inches)	
Cycle	Watkinsville	Tifton	Watkinsville	Tifton
4	0.976	0.949	8.890	9.582
6	.975	.928	8.871	9.651
8	.978	.964	8.969	9.576*
10	.992	.909	9.007	9.735
12	.981	.911	8.896	9.790
14	.992	.955	8.594	9.929
16	.928	.938	9.034	9.651
18	.926	.920	8.559*	10.037
20	.920	.902	8.939	10.140
22	.917*	.980	8.921	10.143
24	.946	.898*	8.908	10.188

* Minimum residual standard deviations are shown in table 4.

Table 6. Variation of nodes with cycle length (Tifton annual temperature)

Position of node in cycle	Nodal values for wave form in increasing cycle length — Cycle length (years)										
	4	6	8	10	12	14	16	18	20	22	24
1	67.78	67.87*	67.86*	68.00*	68.28*	67.48	68.00*	67.77	68.09	67.23	68.38*
3	67.45	67.63	67.41†	67.59	67.47	67.49	67.89	68.04	67.45	68.31*	67.78
5		67.16†	67.69	66.92†	67.03†	67.79	67.52	66.78†	66.68†	67.11	66.87
7			67.49	67.36	67.46	67.90*	67.08	67.73	67.55	67.41	67.31
9				67.73	67.88	67.60	67.76	67.38	68.12*	67.40	67.68
11					67.37	67.05†	66.96†	67.84	67.39	66.88	66.67†
13						67.14	67.90	68.36*	67.24	67.13	67.96
15							67.94	67.51	66.71	66.83†	66.84
17								67.11	67.46	67.20	67.01
19									66.88	67.14	67.45
21										67.46	67.93

* Maximum node
† Minimum node

Table 7. Results of ARIMA analysis

Data set	Coefficient	Value	Standard error	T-ratio
Watkinsville	Constant	–0.074	0.110	–0.677
temperature	AR1*	–1.051	0.176	–5.980
	AR2*	–0.260	0.157	–1.654
	MA1†	–0.410	0.139	–2.947
	MA2†	0.600	0.150	3.966
Tifton	Constant	–0.037	0.039	–0.968
temperature	AR1	0.020	0.179	0.111
	MA1	0.700	0.127	5.497
Watkinsville	Constant	59.083	48.808	1.210
rainfall	AR1	–0.200	0.991	–0.202
	MA1	–0.291	0.968	–0.301
Tifton	Constant	89.804	38.594	2.327
rainfall	AR1	–0.890	0.811	–1.097
	MA1	–0.876	0.856	–1.024

*AR1 and AR2 are lag 1 and lag 2 autoregression coefficients.

† MA1 and MA2 are lag 1 and lag 2 moving average coefficients.

Figure 1. Schematic of CYCLES model

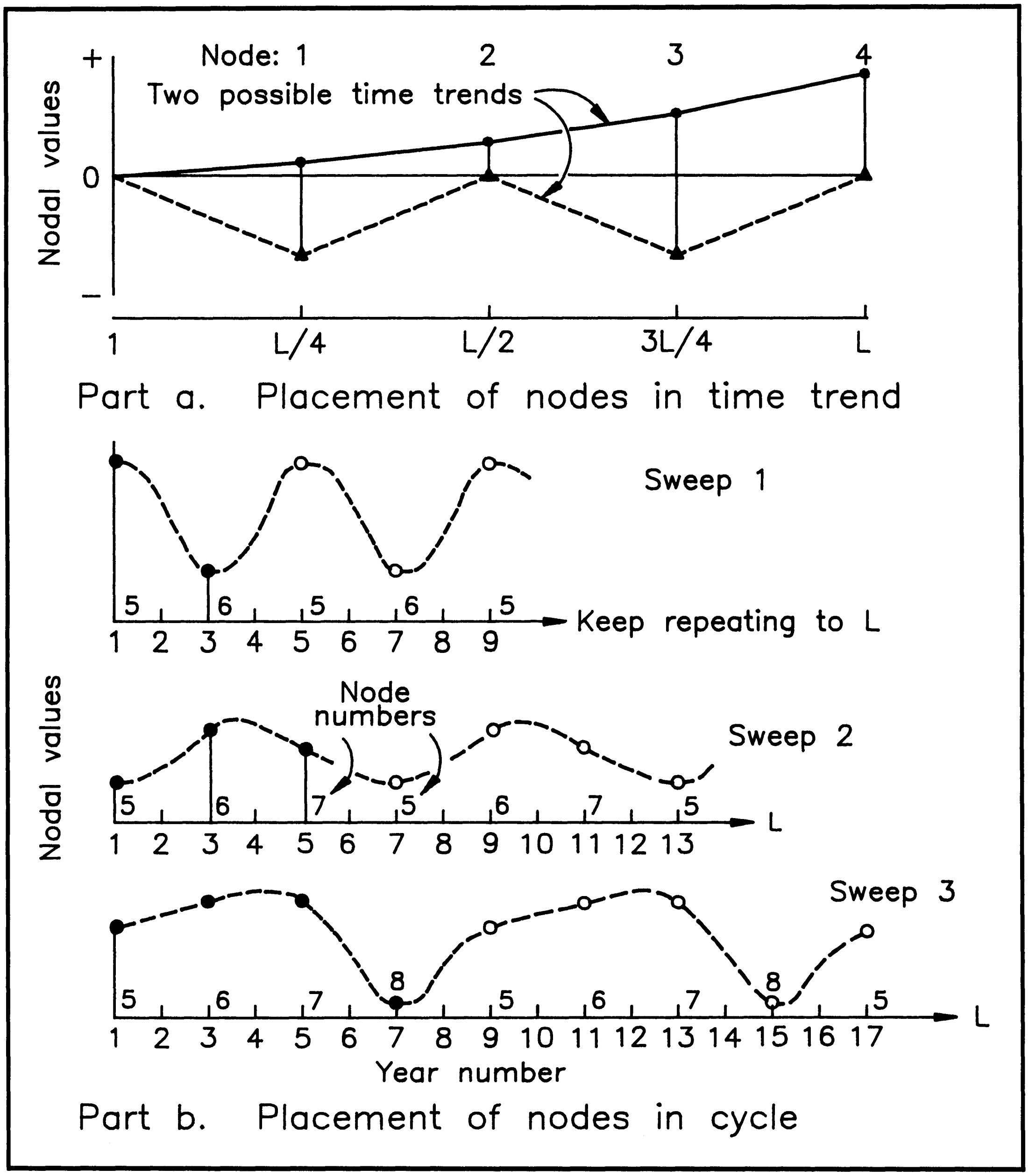


Figure 2. Schematic for evaluation of $b(j, \bar{x}_j)$

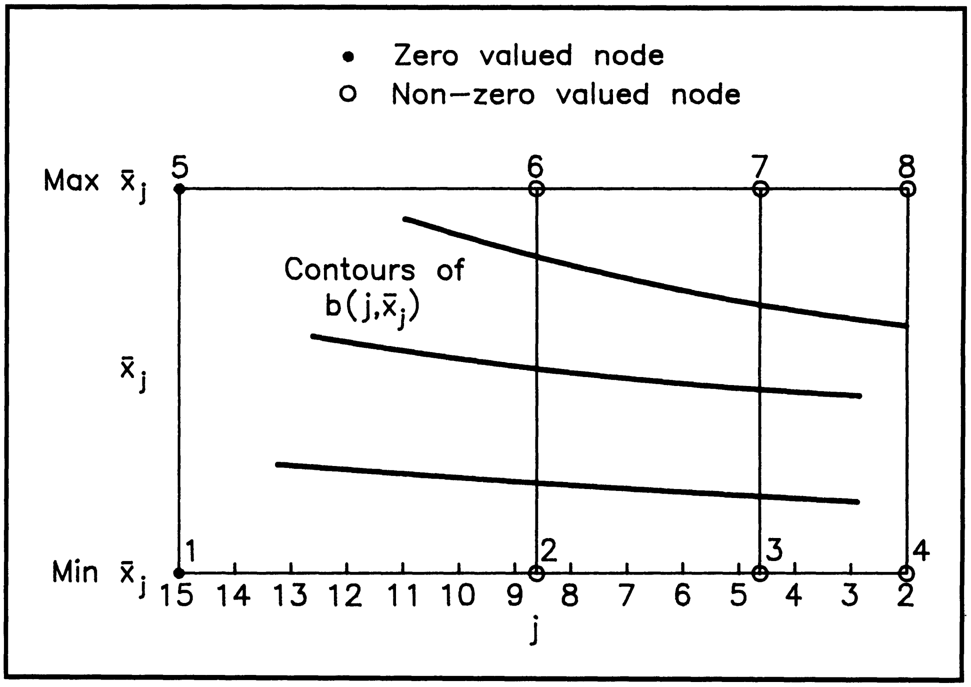


Figure 3. White Hollow monthly rainfall and runoff

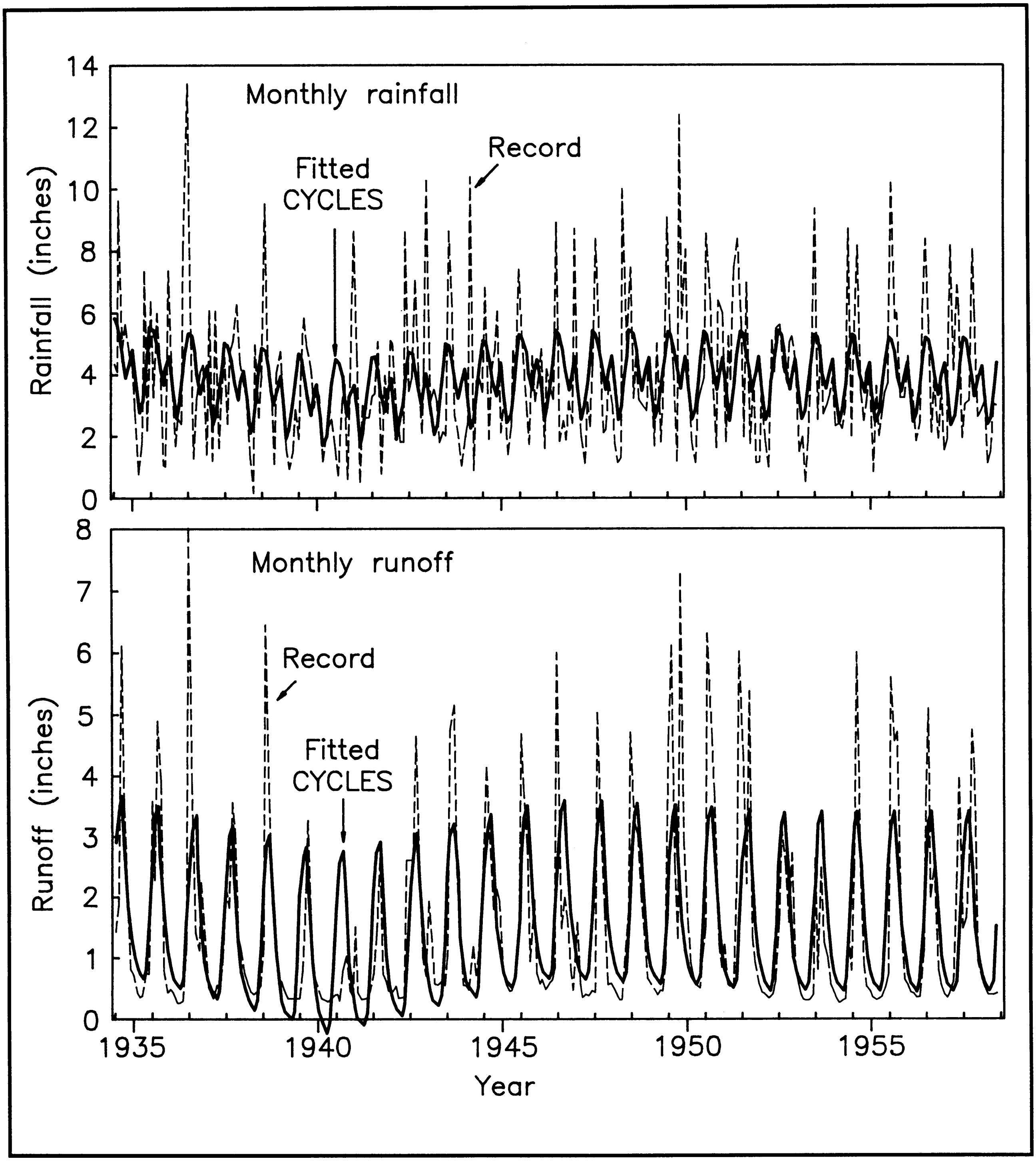


Figure 4. Pine Tree Branch monthly rainfall and runoff

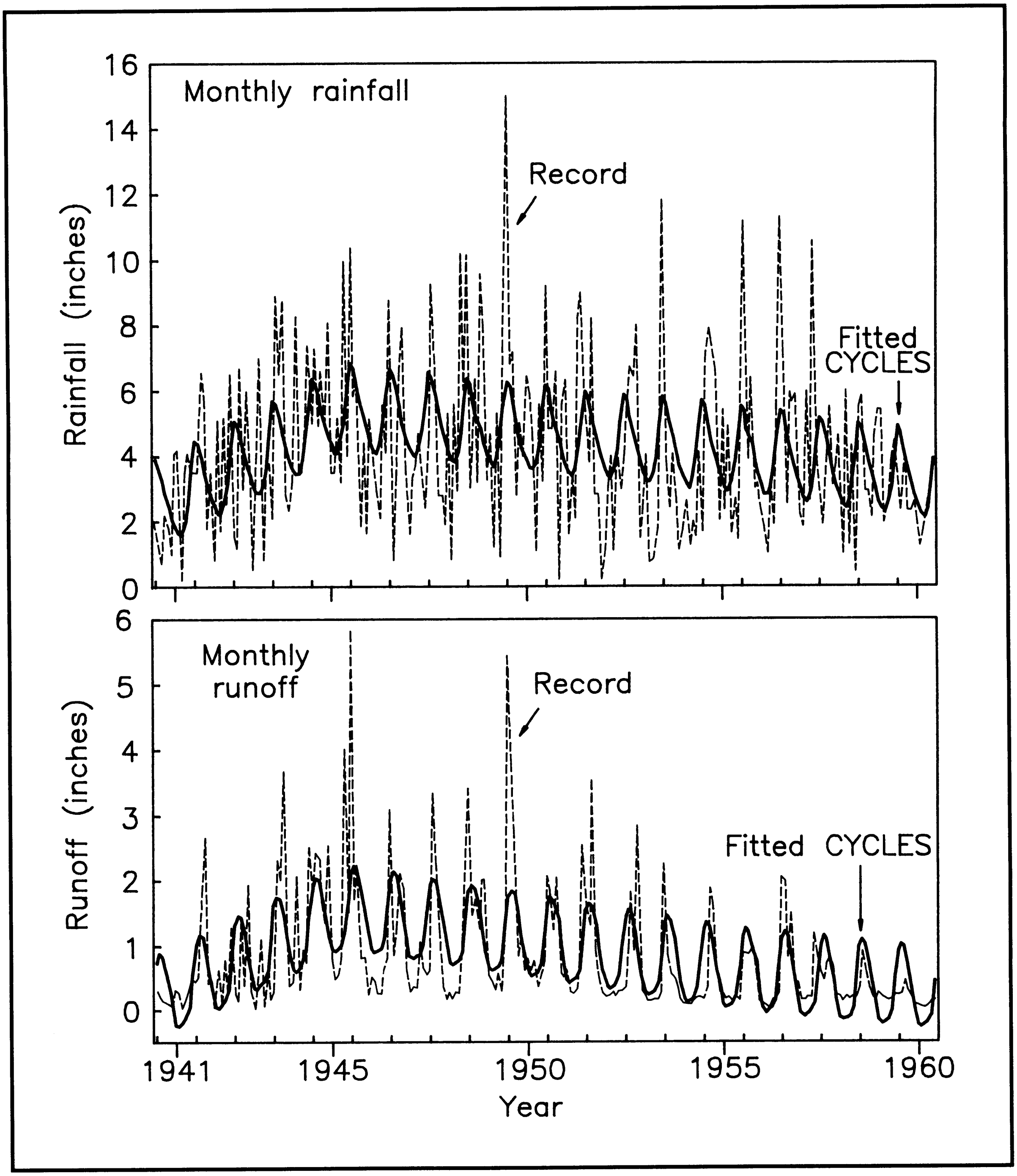


Figure 5. Residual errors versus length of cycle

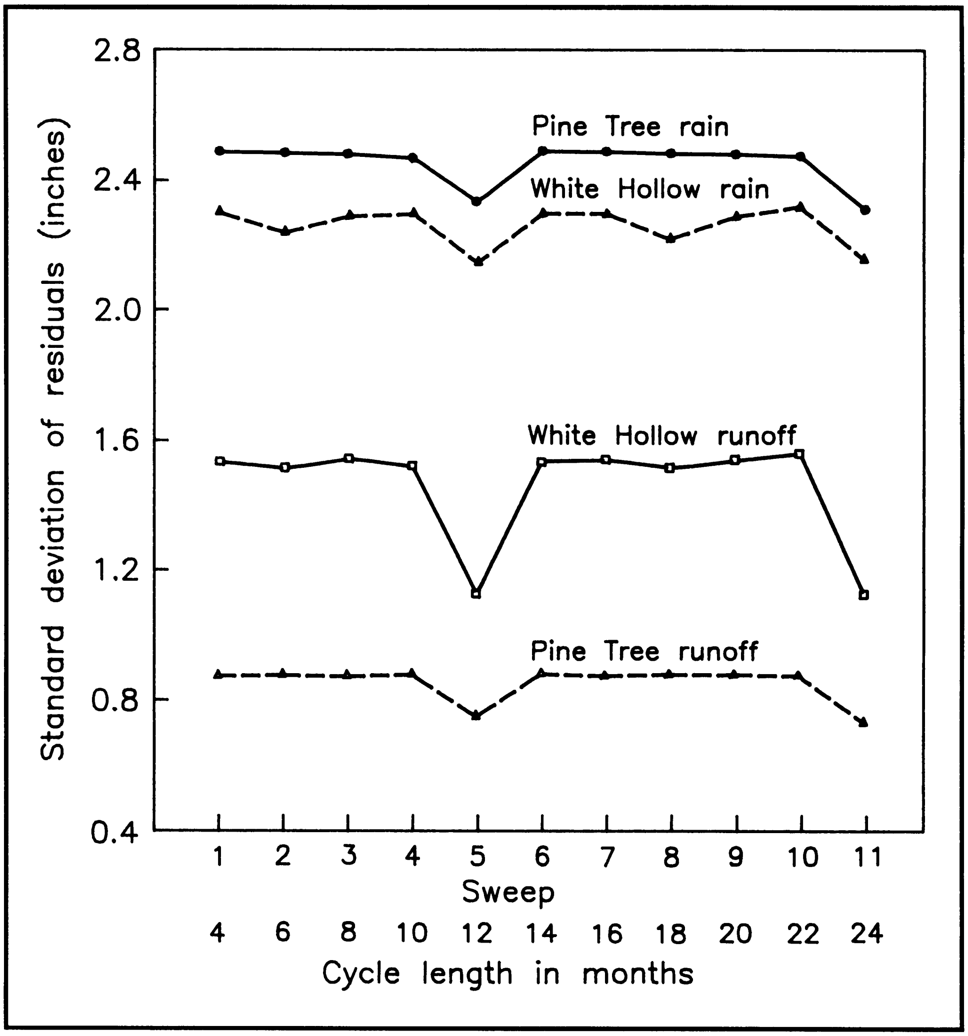


Figure 6. Seasonal cycles of rainfall and runoff

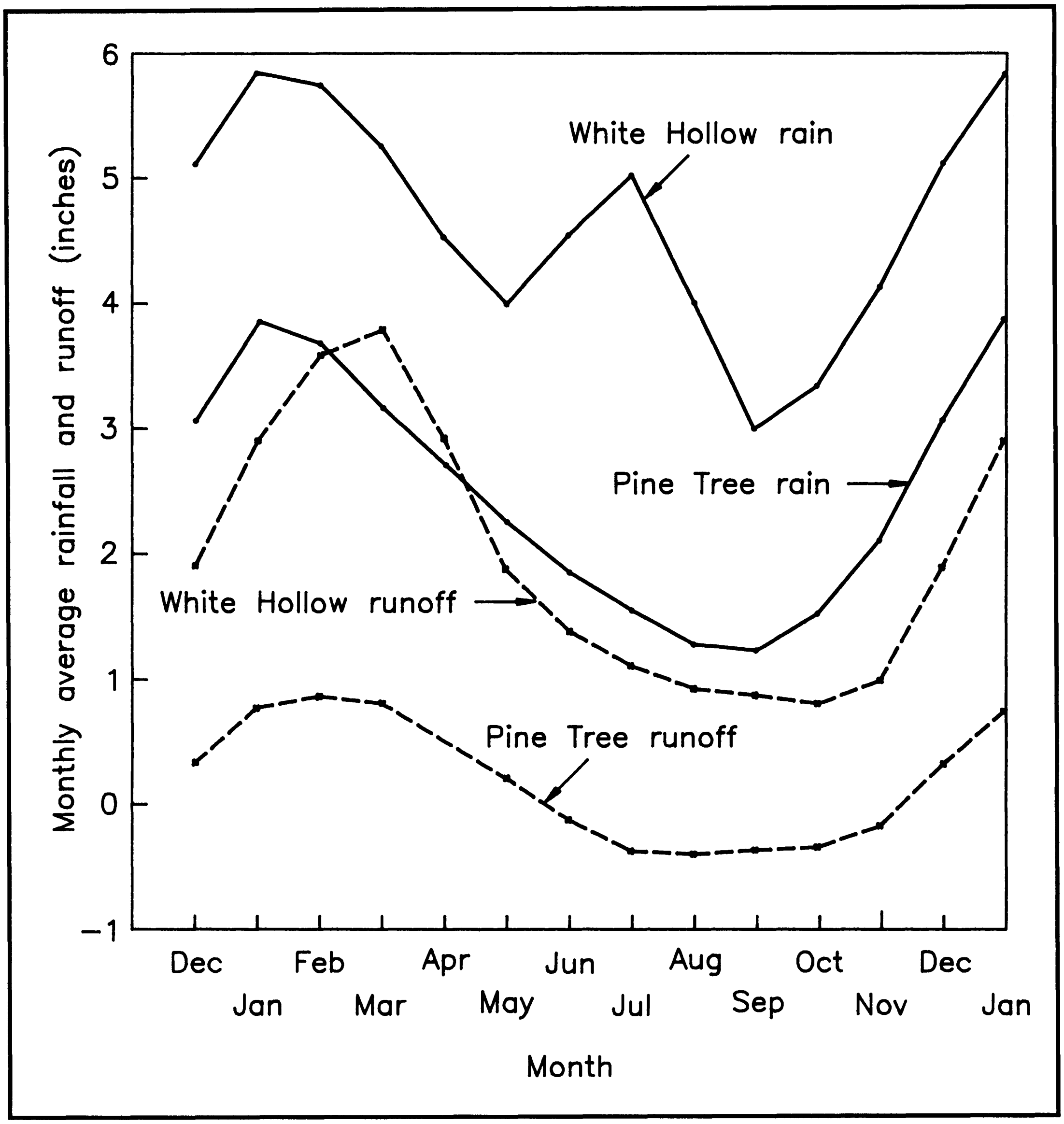
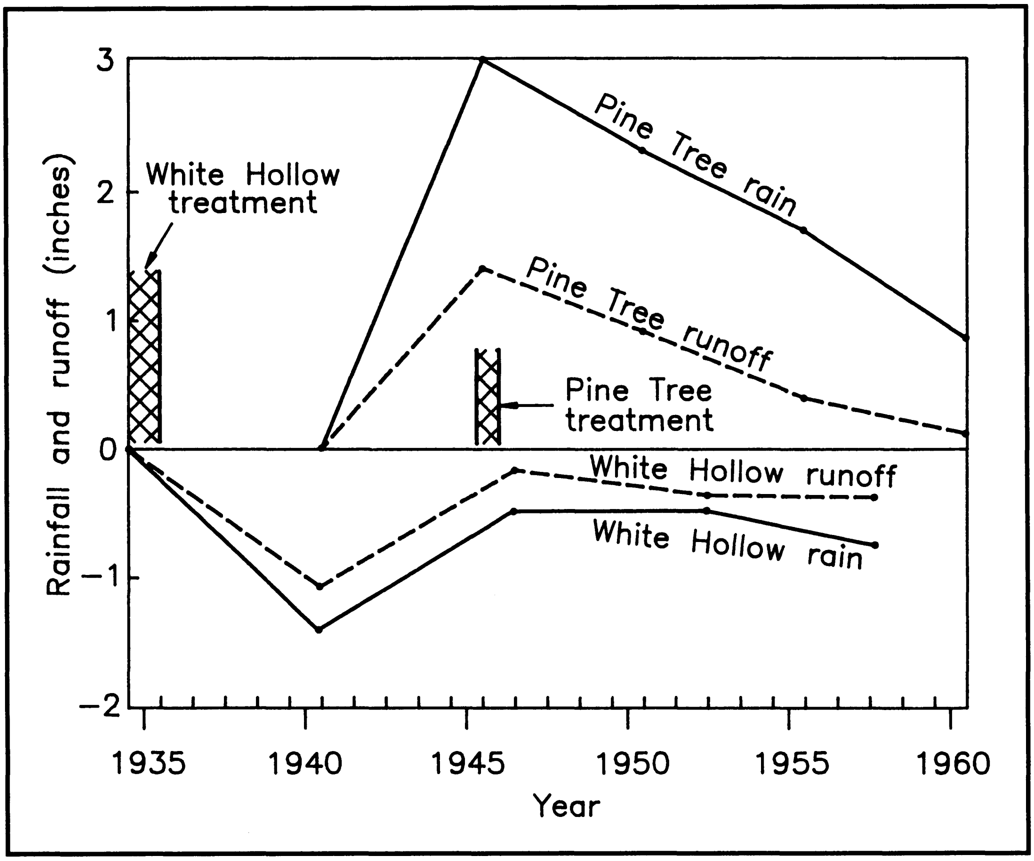


Figure 7. Time trends of rainfall and runoff



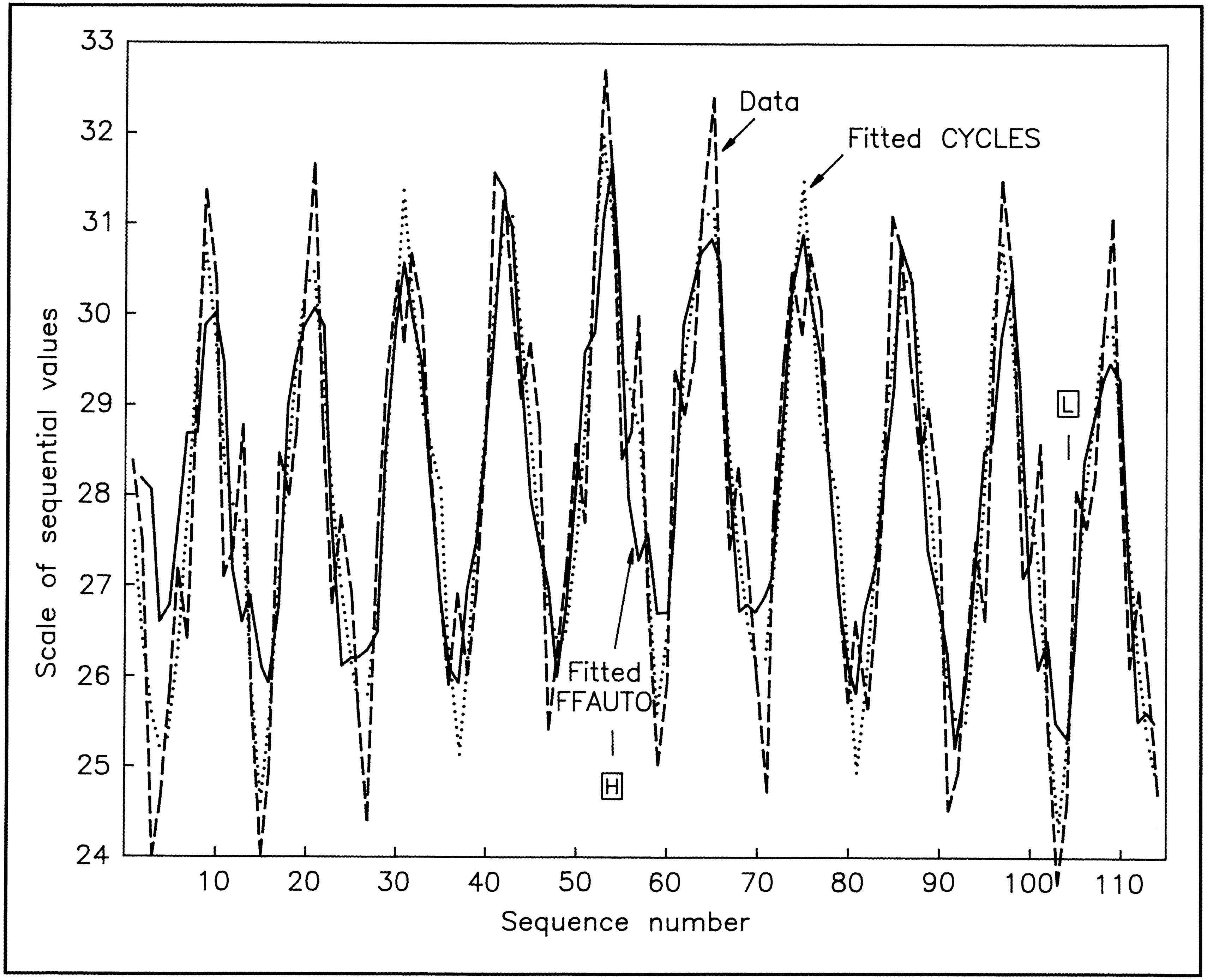


Figure 8. Synthetic data set number 1

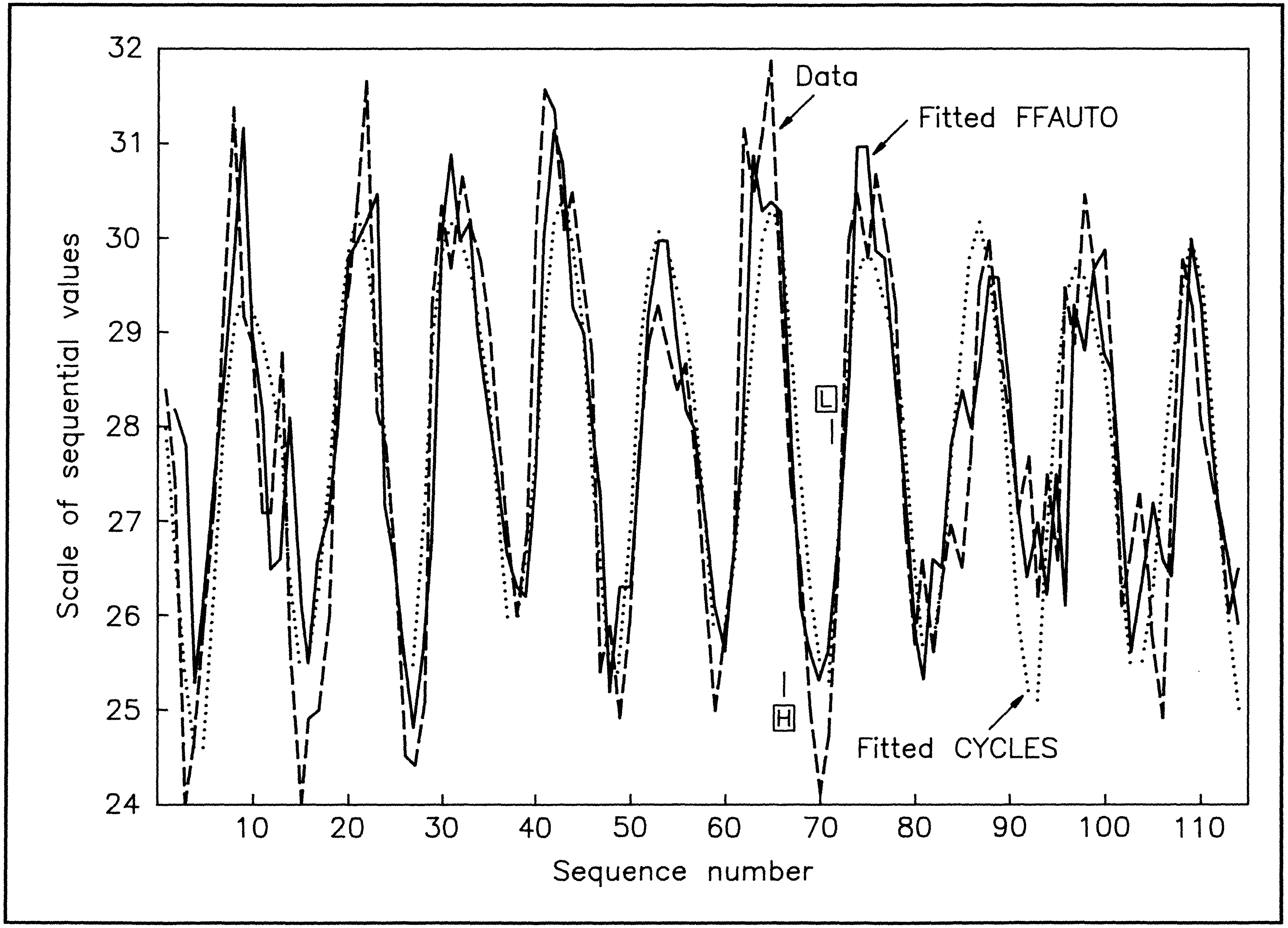


Figure 9. Synthetic data set number 2

Figure 10. Synthetic data set number 3

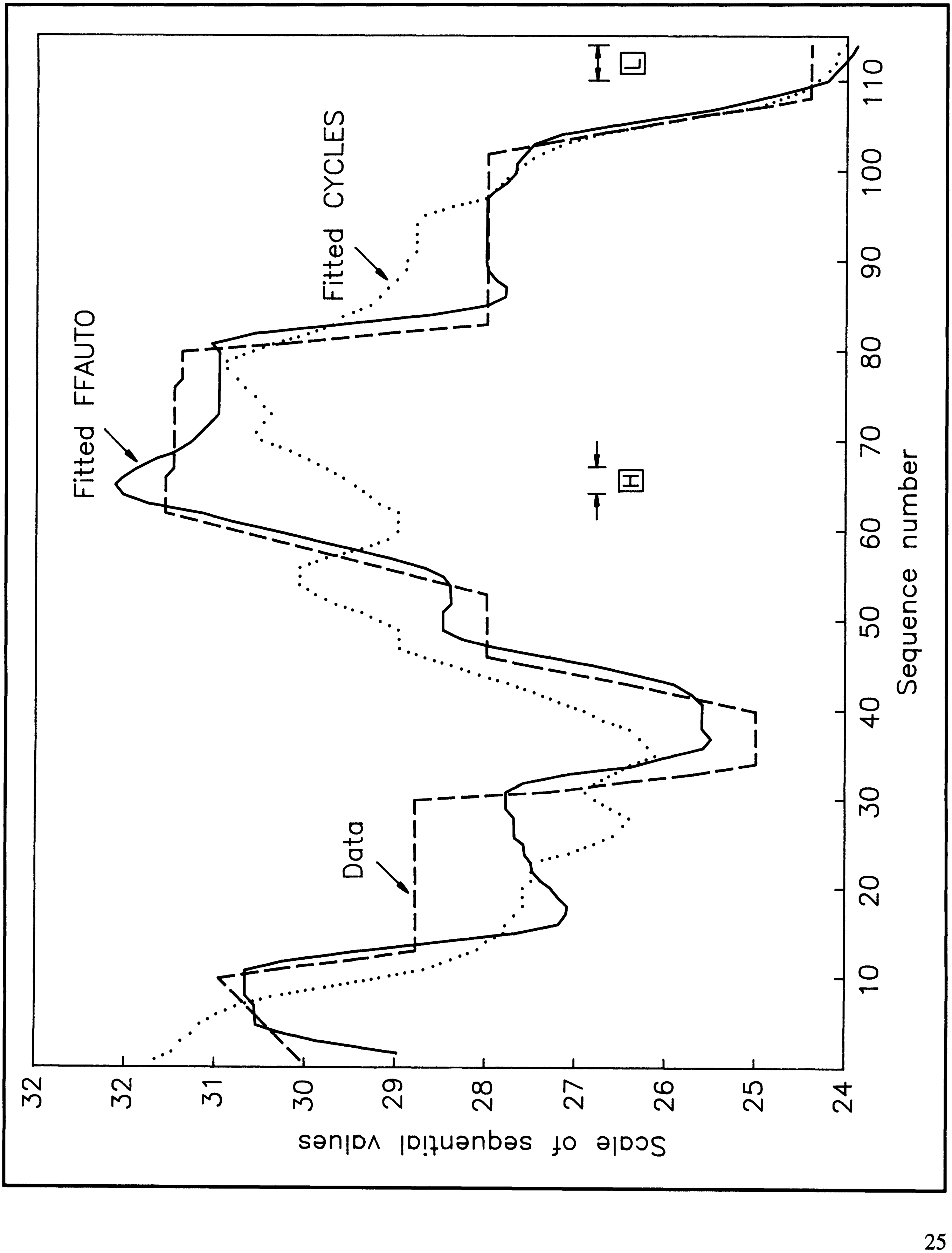


Figure 11. Autoregression functions

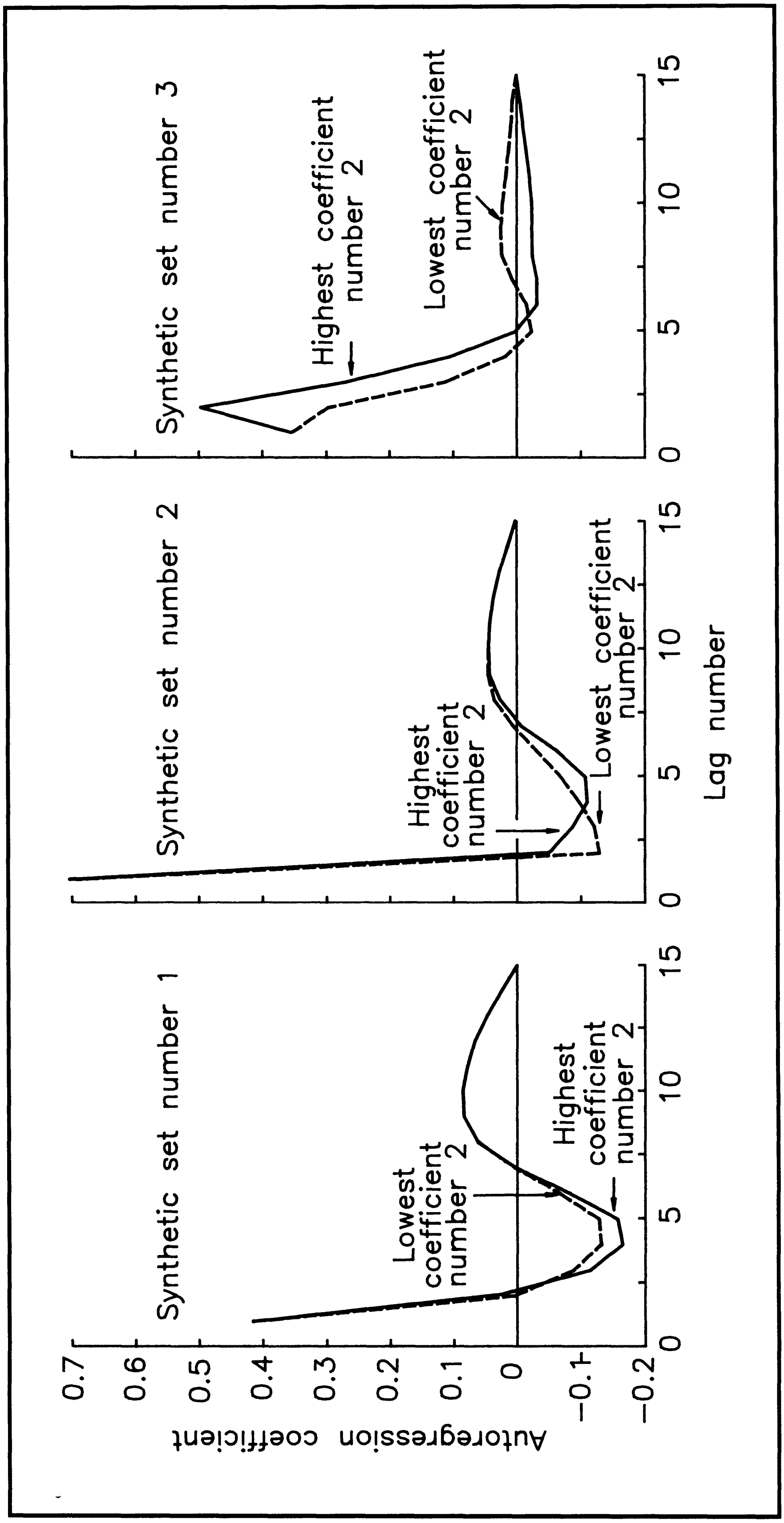


Figure 12. Time trends in synthetic data

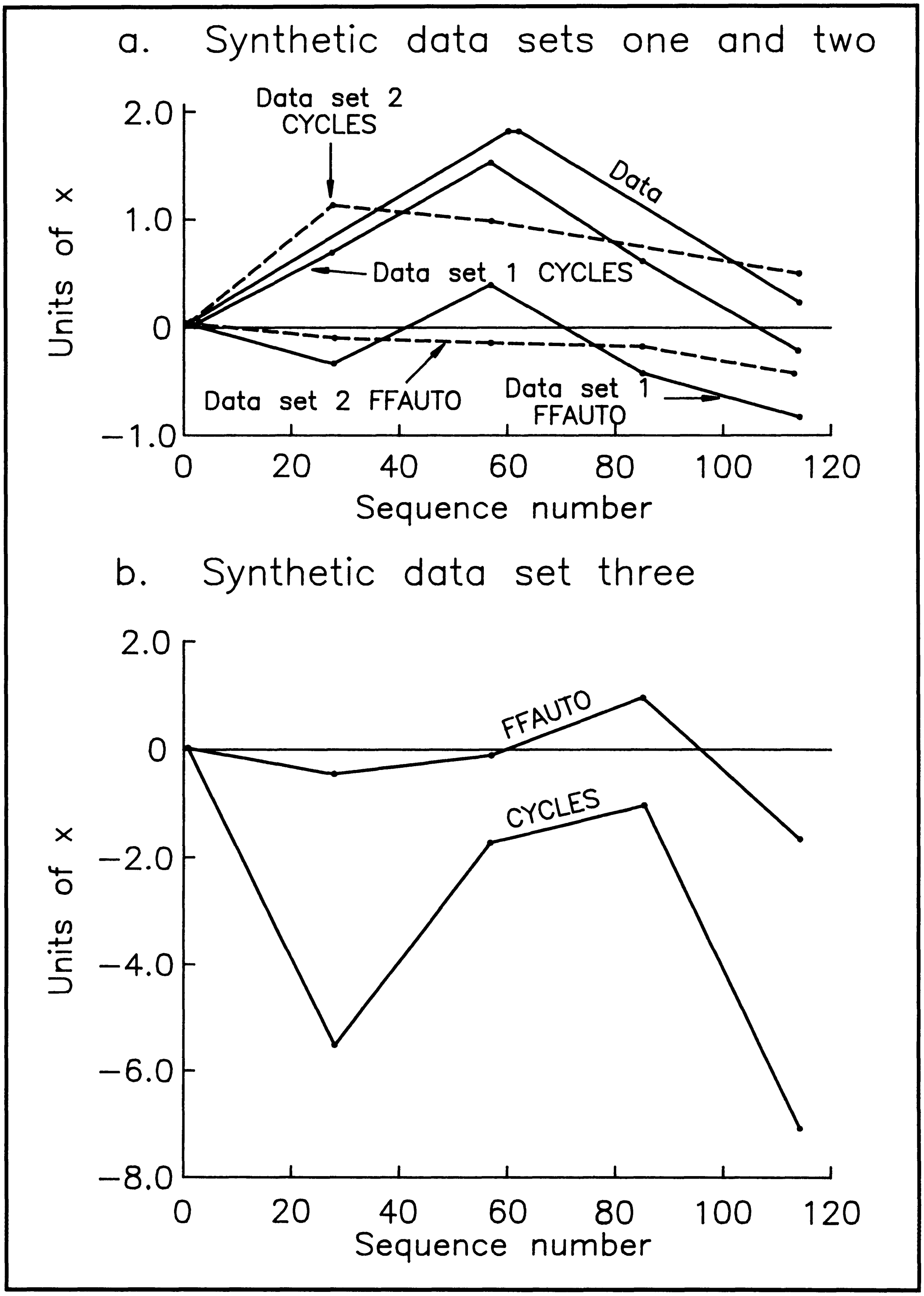


Figure 13. Cycles derived from synthetic data sets

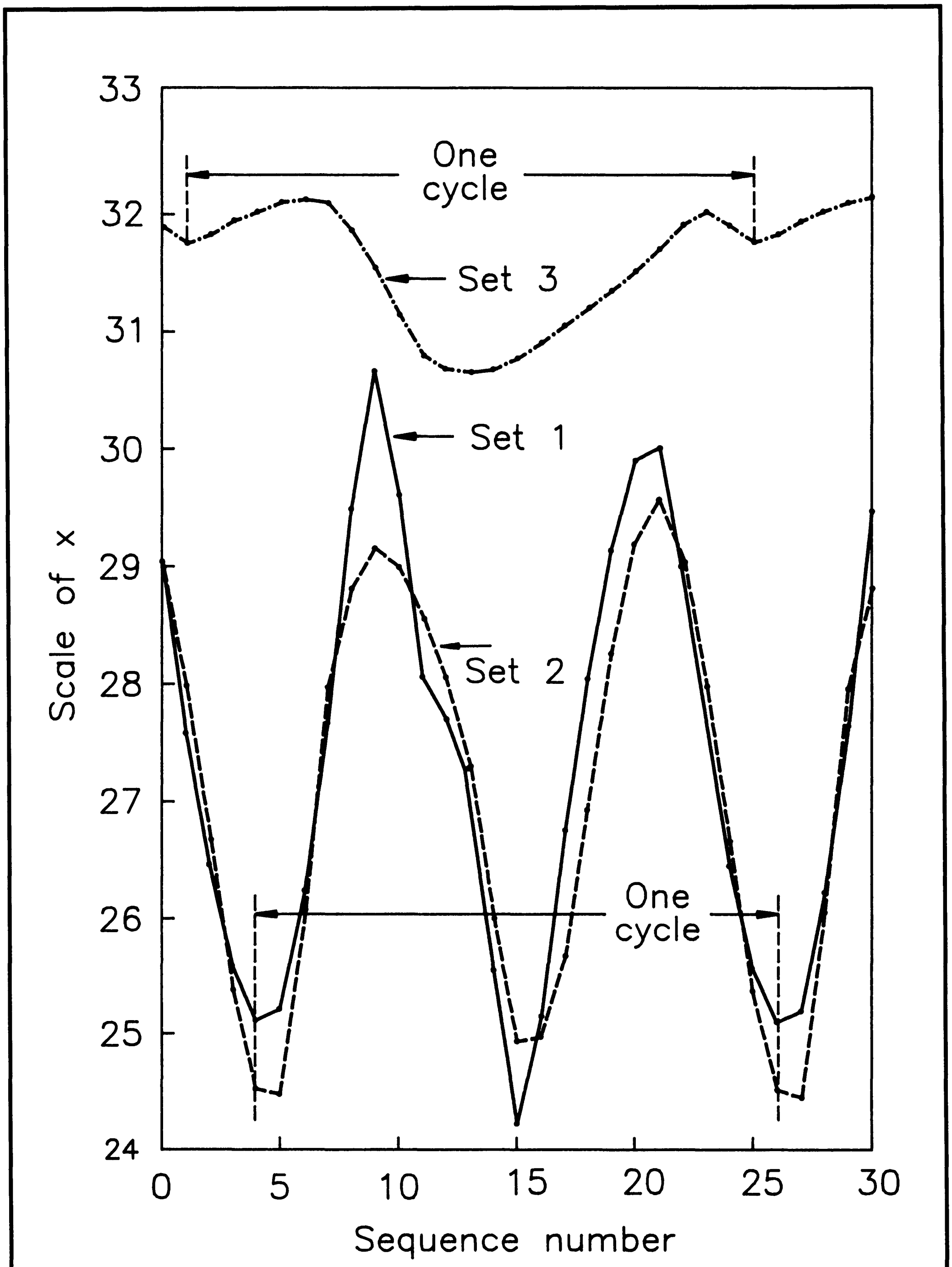


Figure 14. Residual errors for synthetic data sets 1 and 3

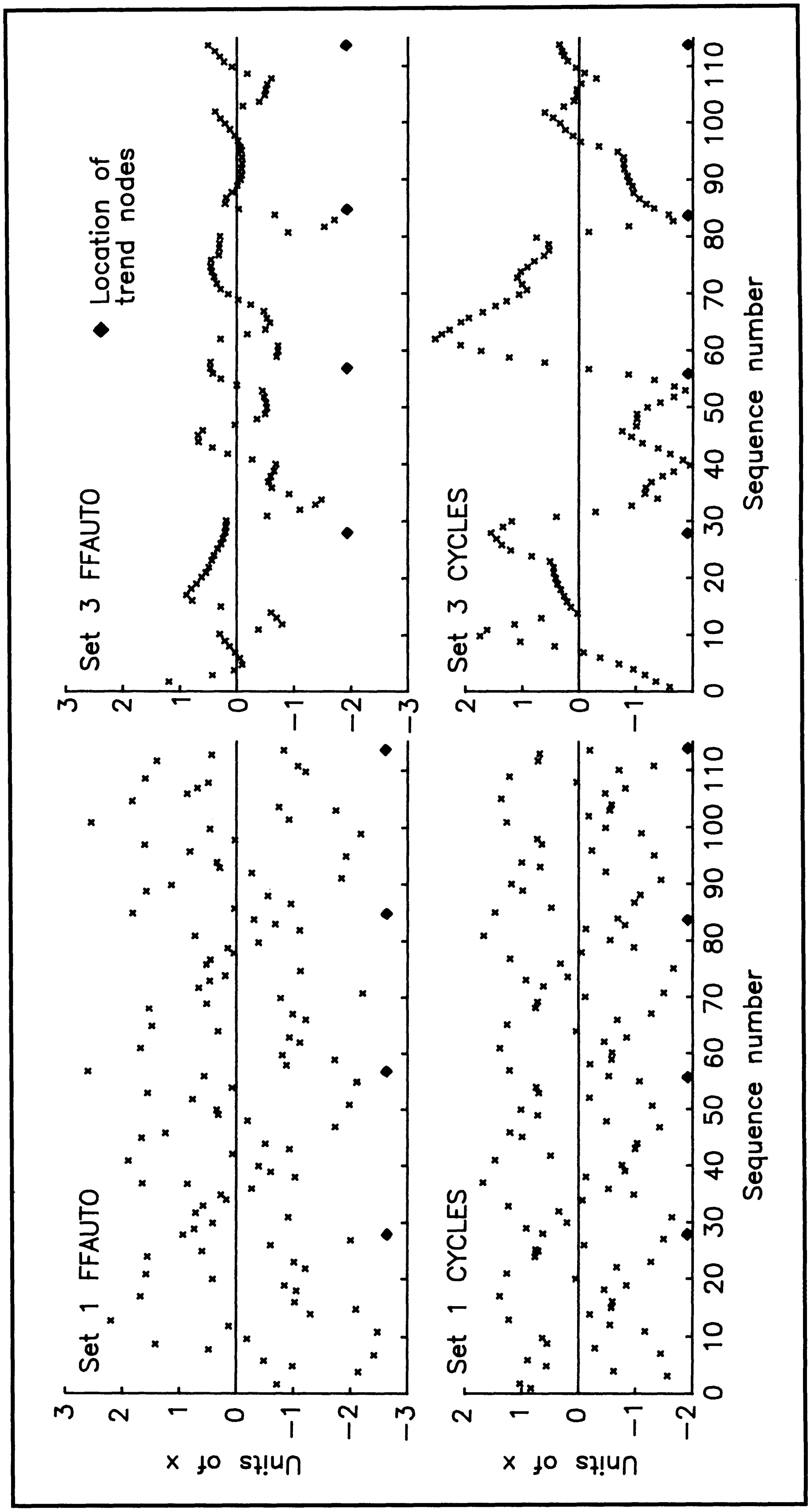


Figure 15. Watkinsville average annual temperature

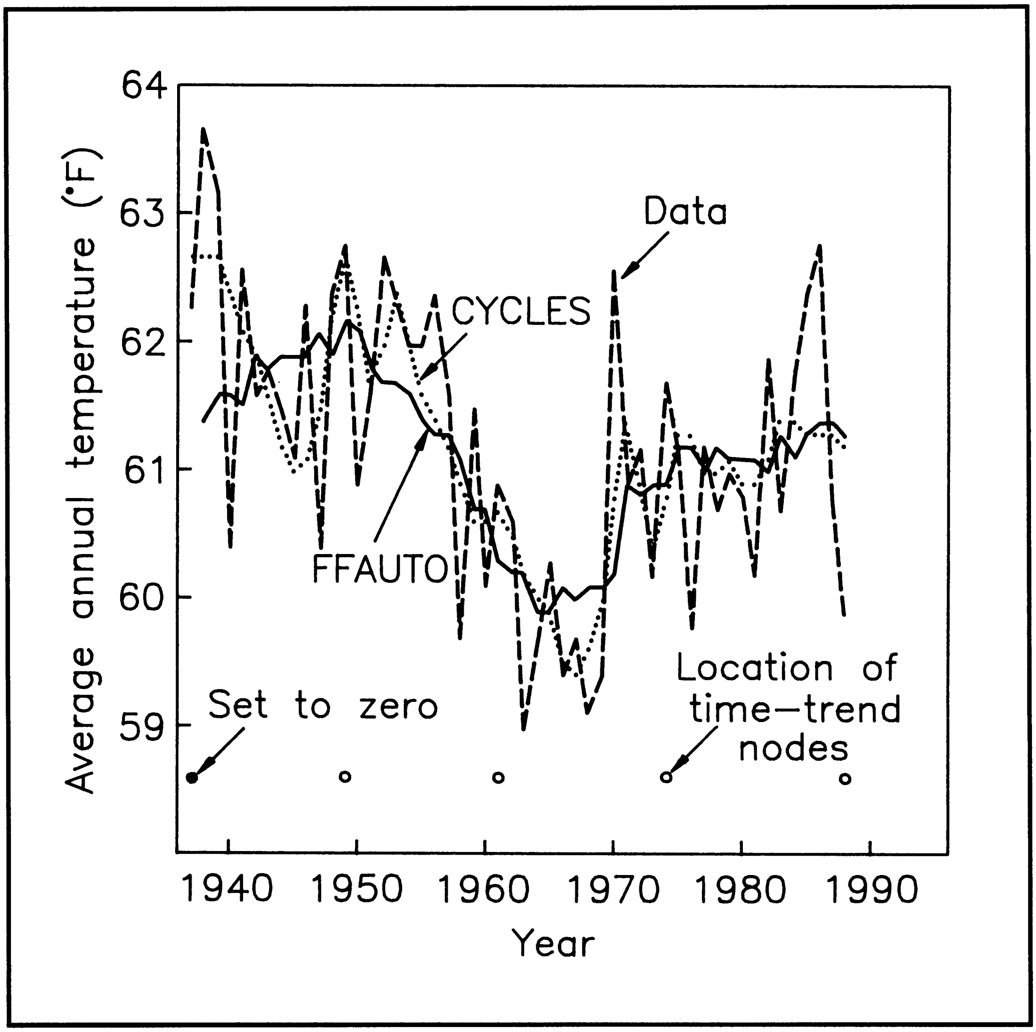


Figure 16. Tifton average annual temperature

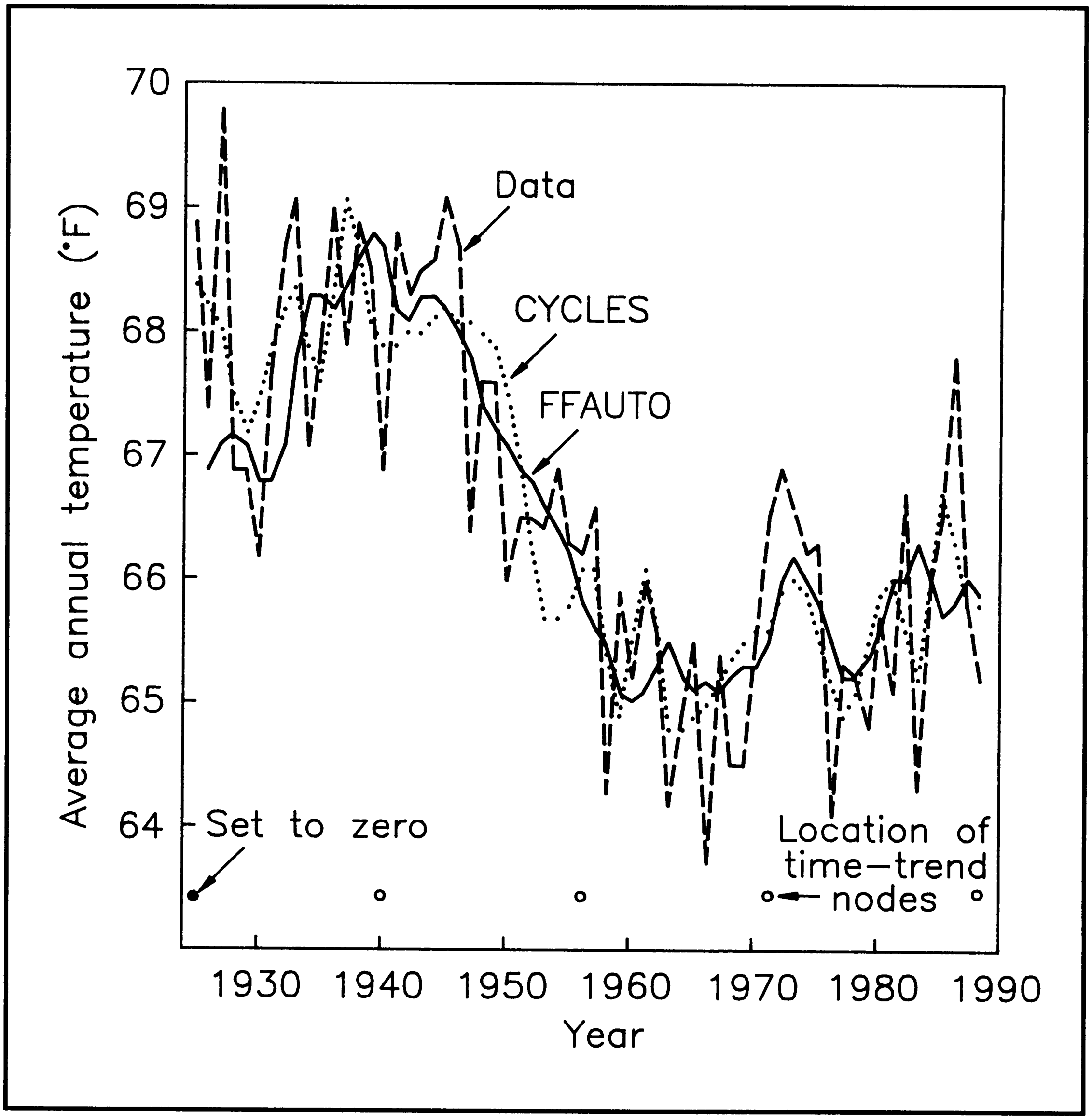


Figure 17. Watkinsville annual rainfall

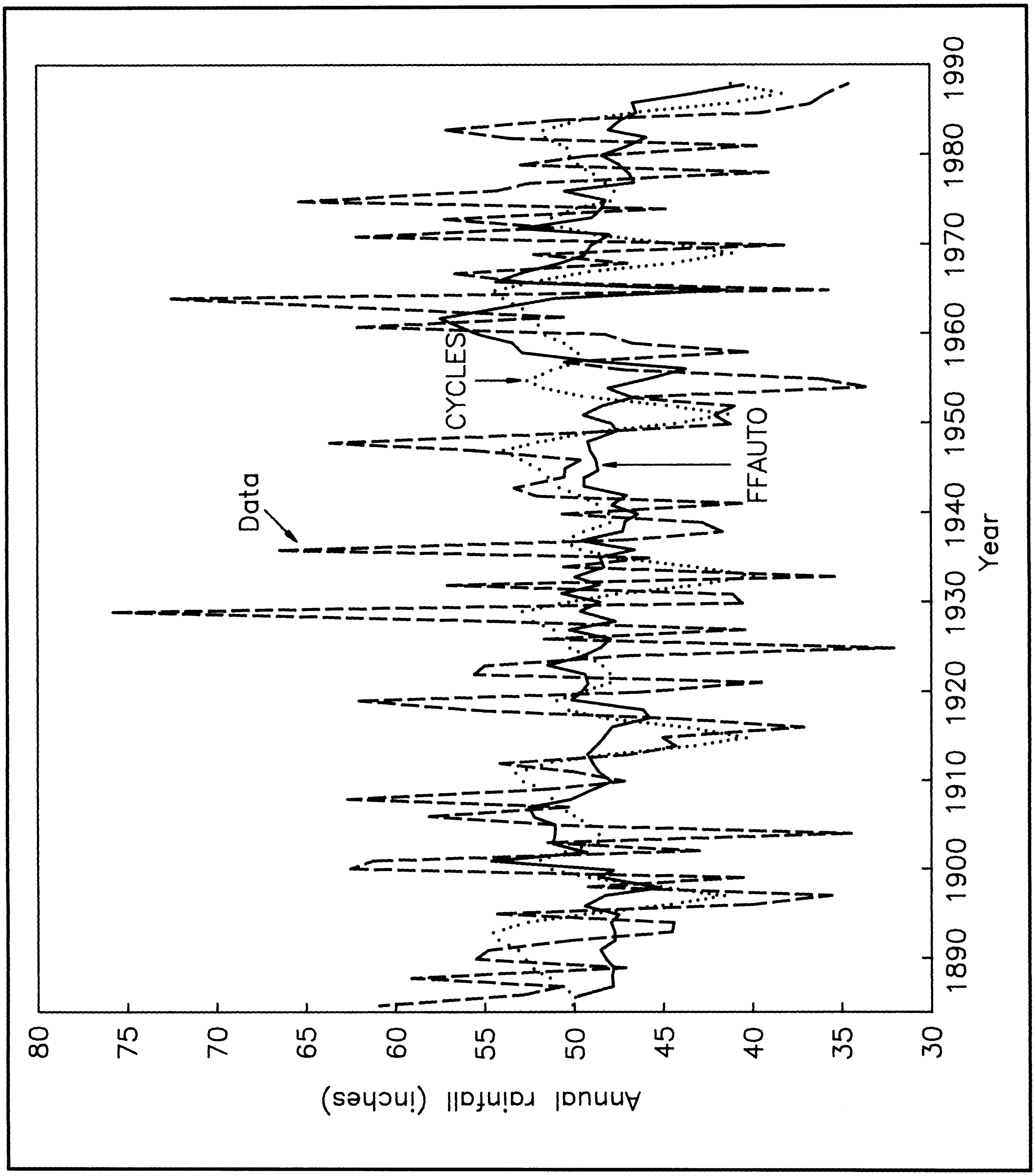


Figure 18. Tifton annual rainfall

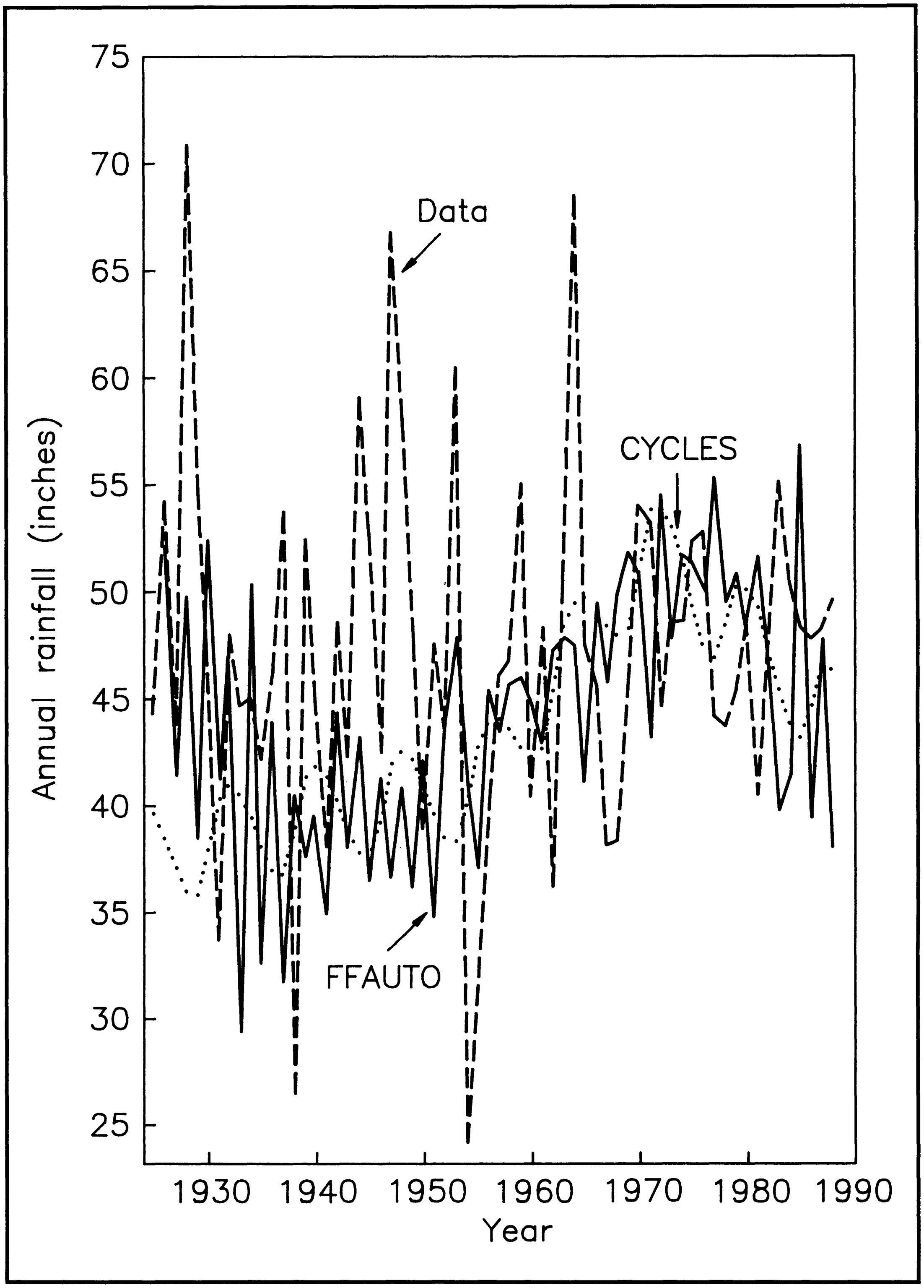


Figure 19. Time trends for temperature from fitted models

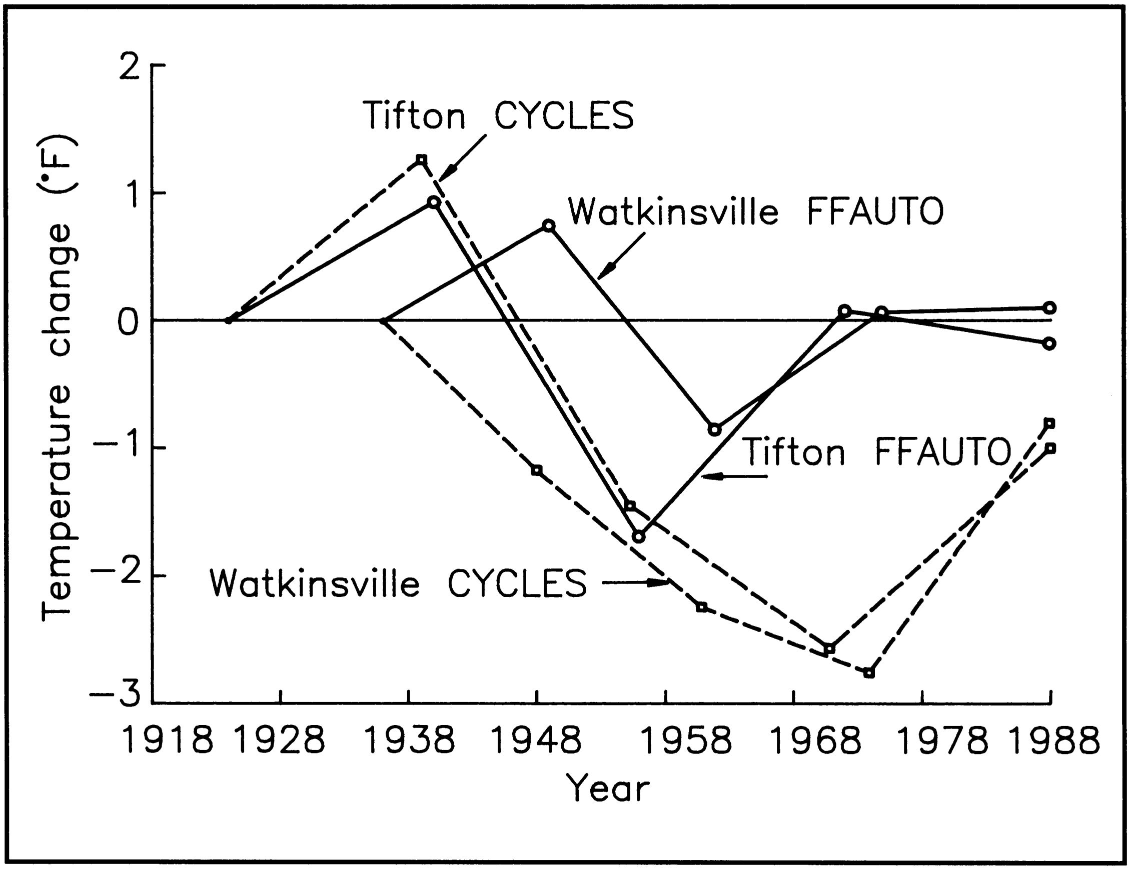


Figure 20. Time trends for rainfall from fitted models

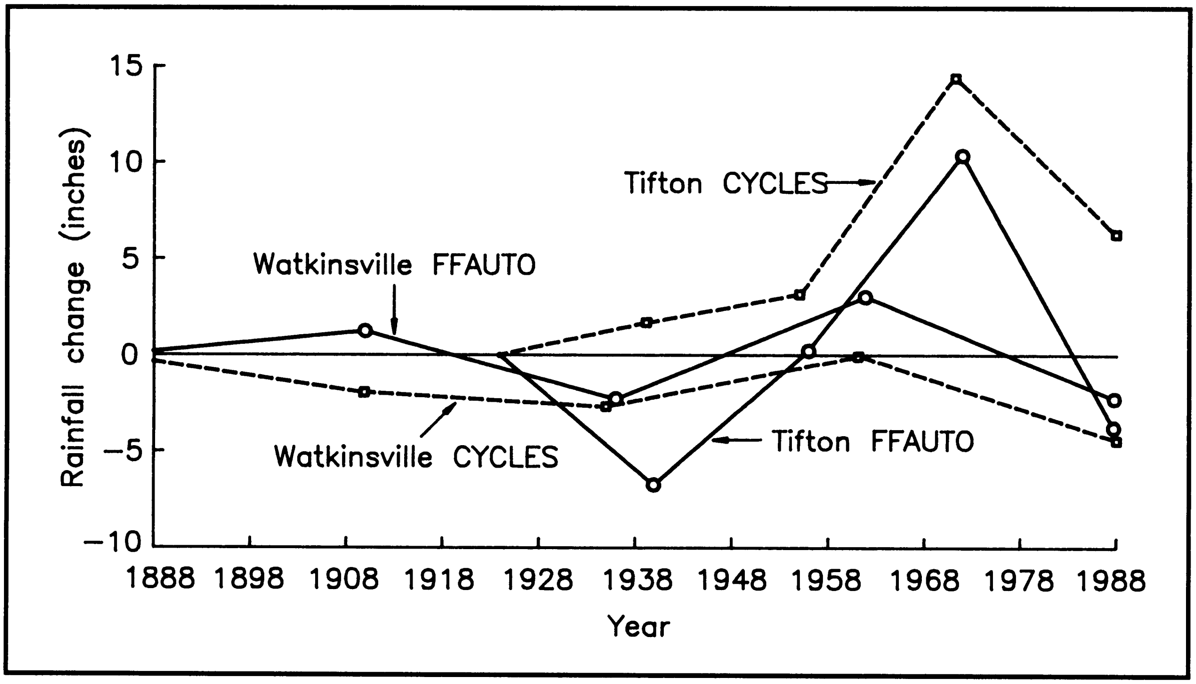


Figure 21. Free-form autoregression functions for temperature

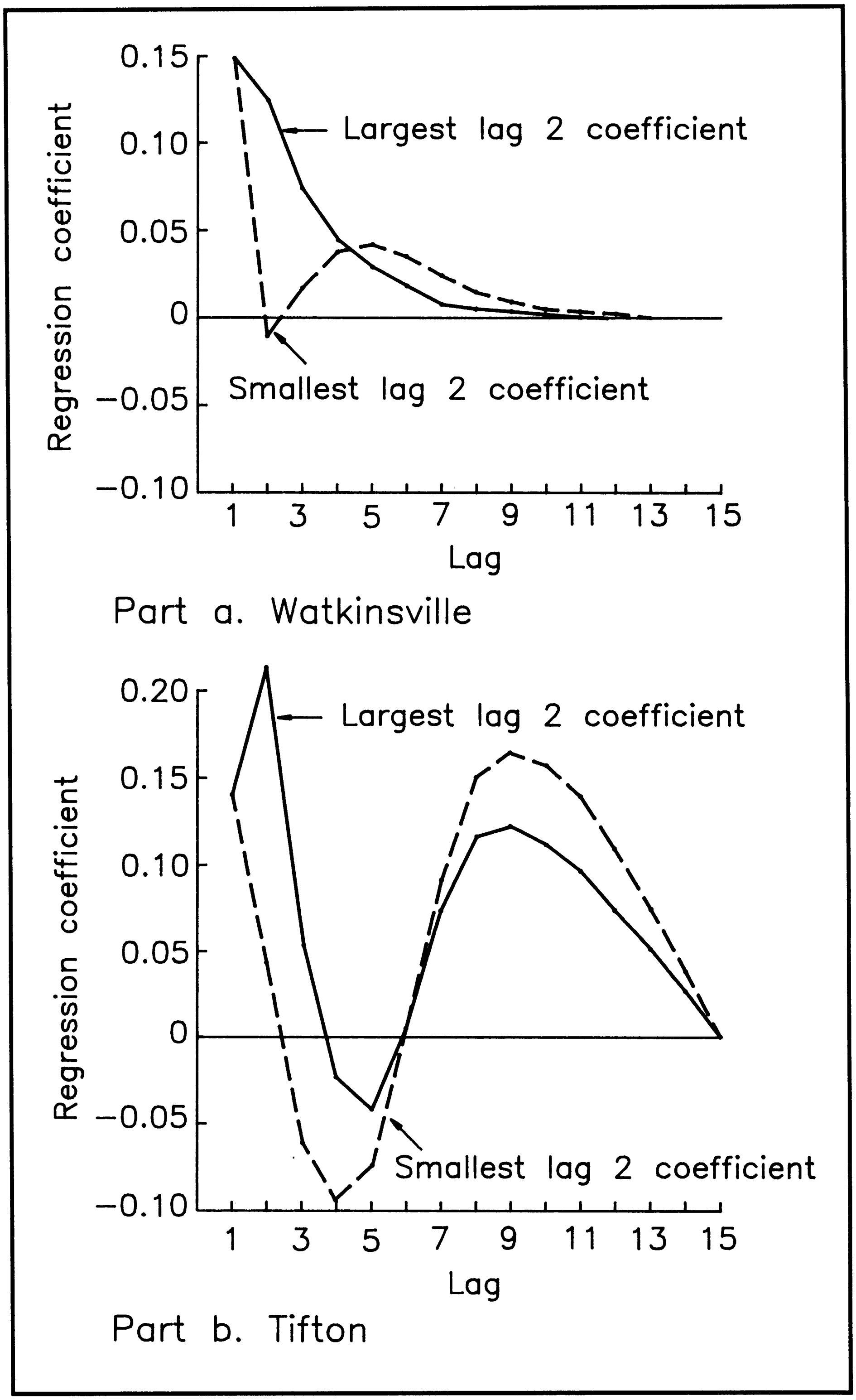


Figure 22. Free-form autoregression functions for rainfall

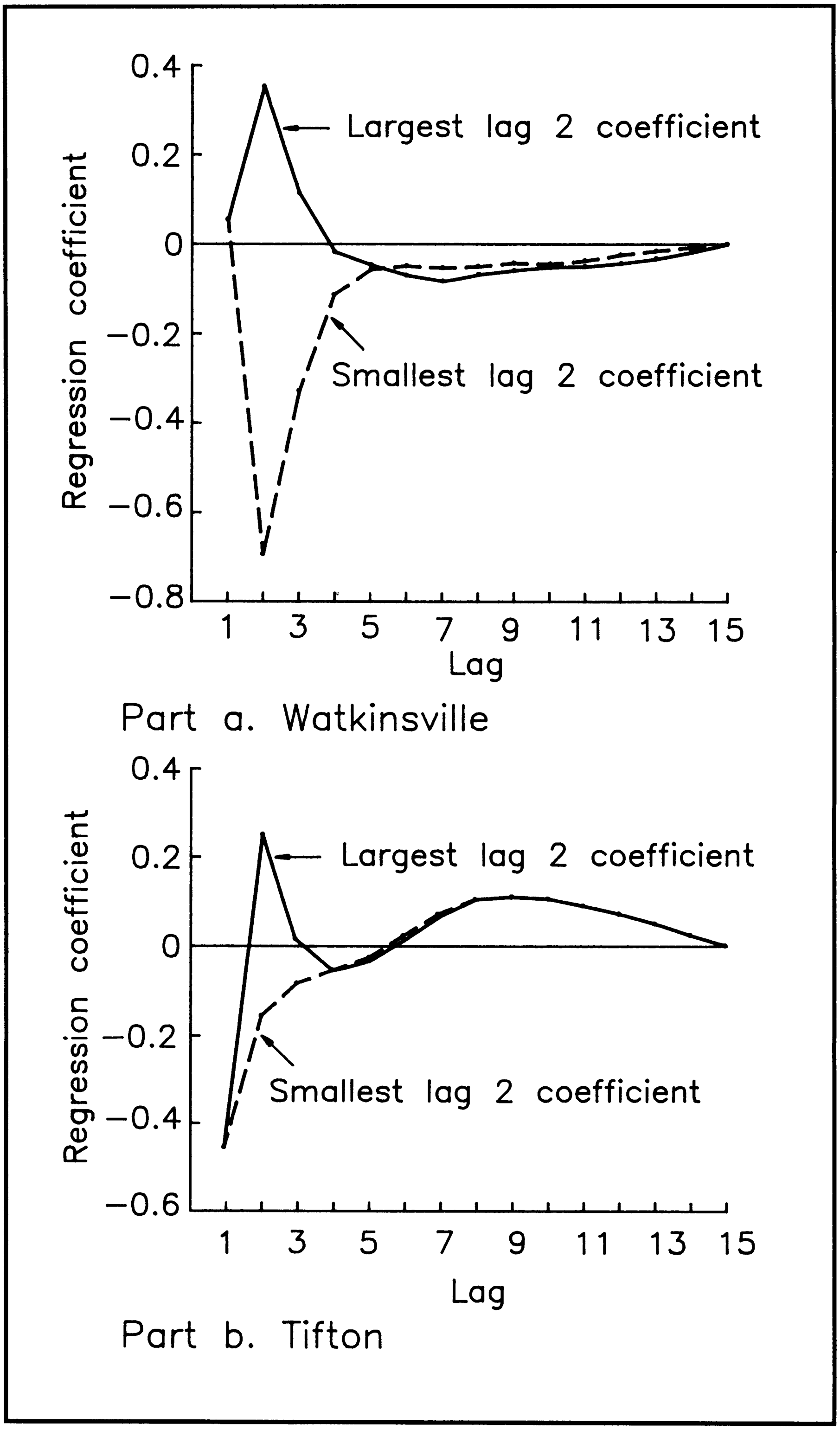


Figure 23. Best-fit free-form cycle of Tifton temperature

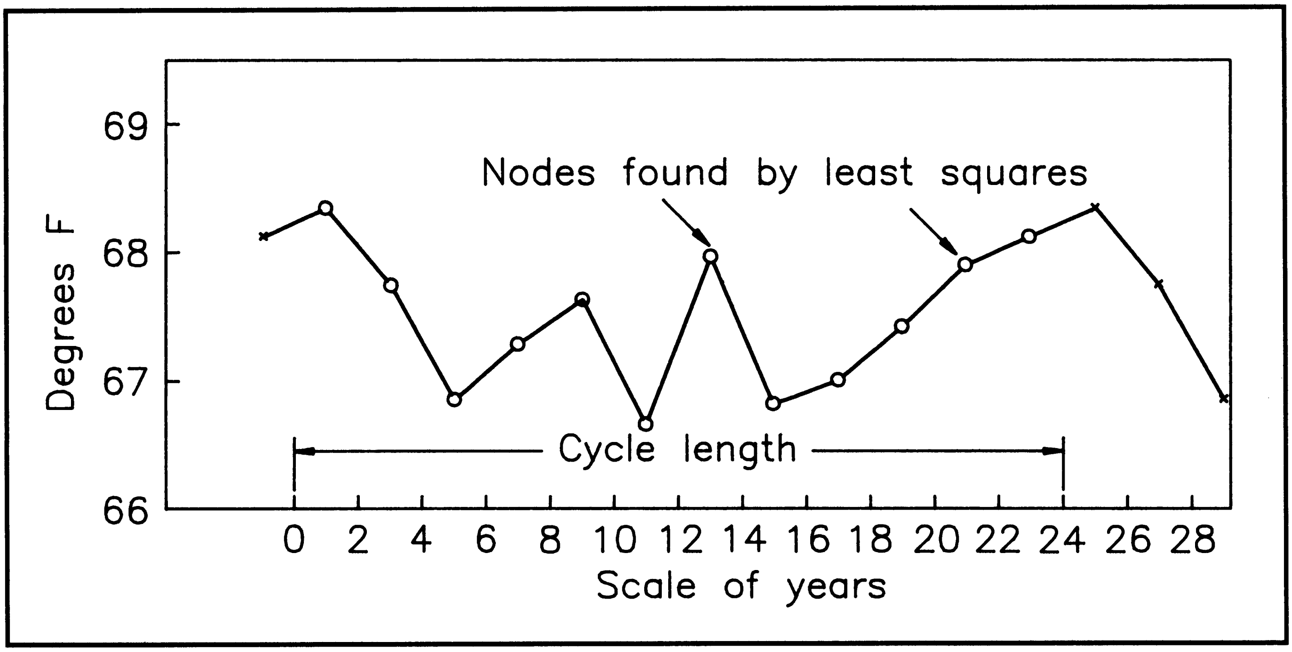


Figure 24. Autocorrelation of residuals for Watkinsville temperature

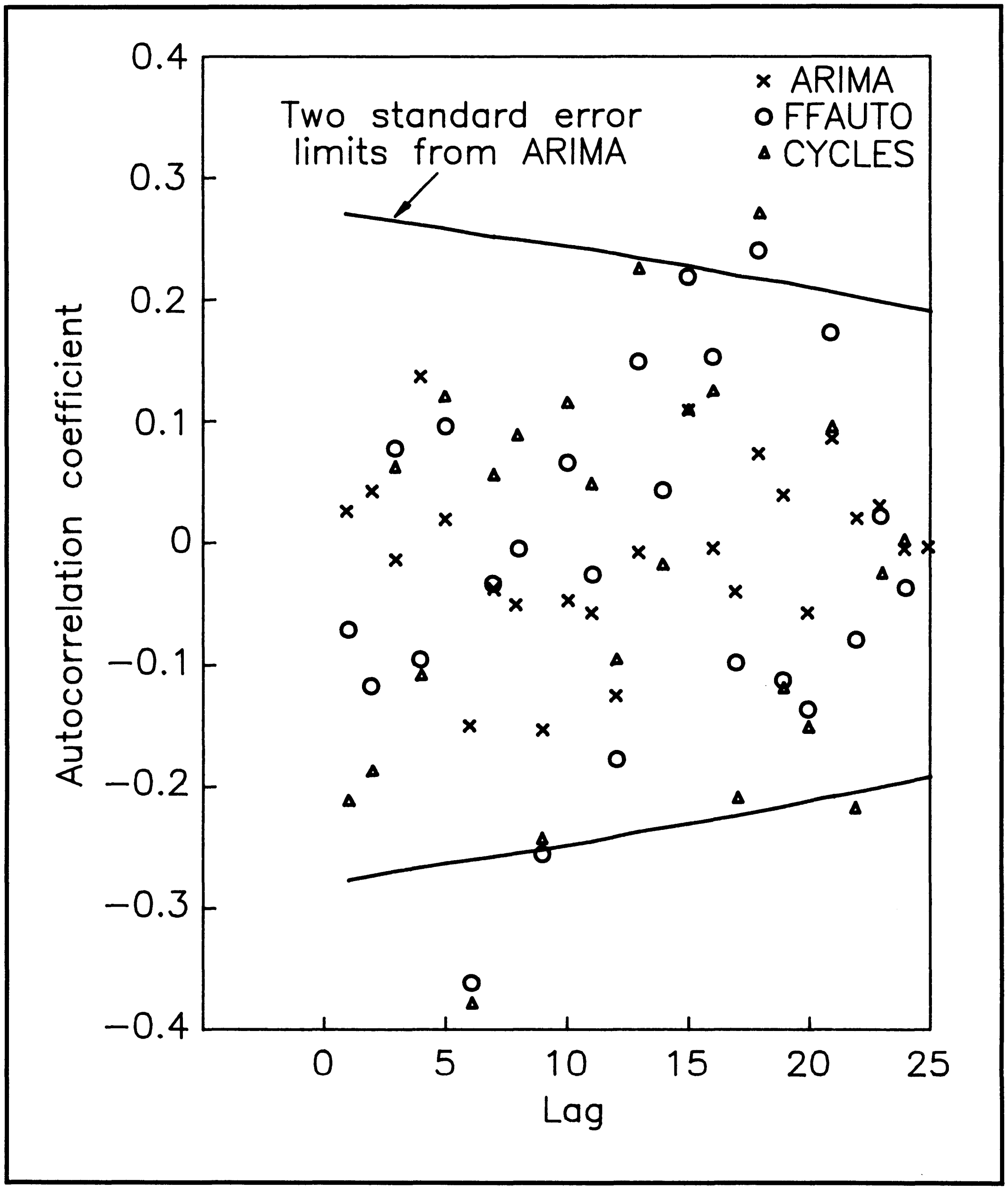


Figure 25. Autocorrelation of residuals for Tifton temperature

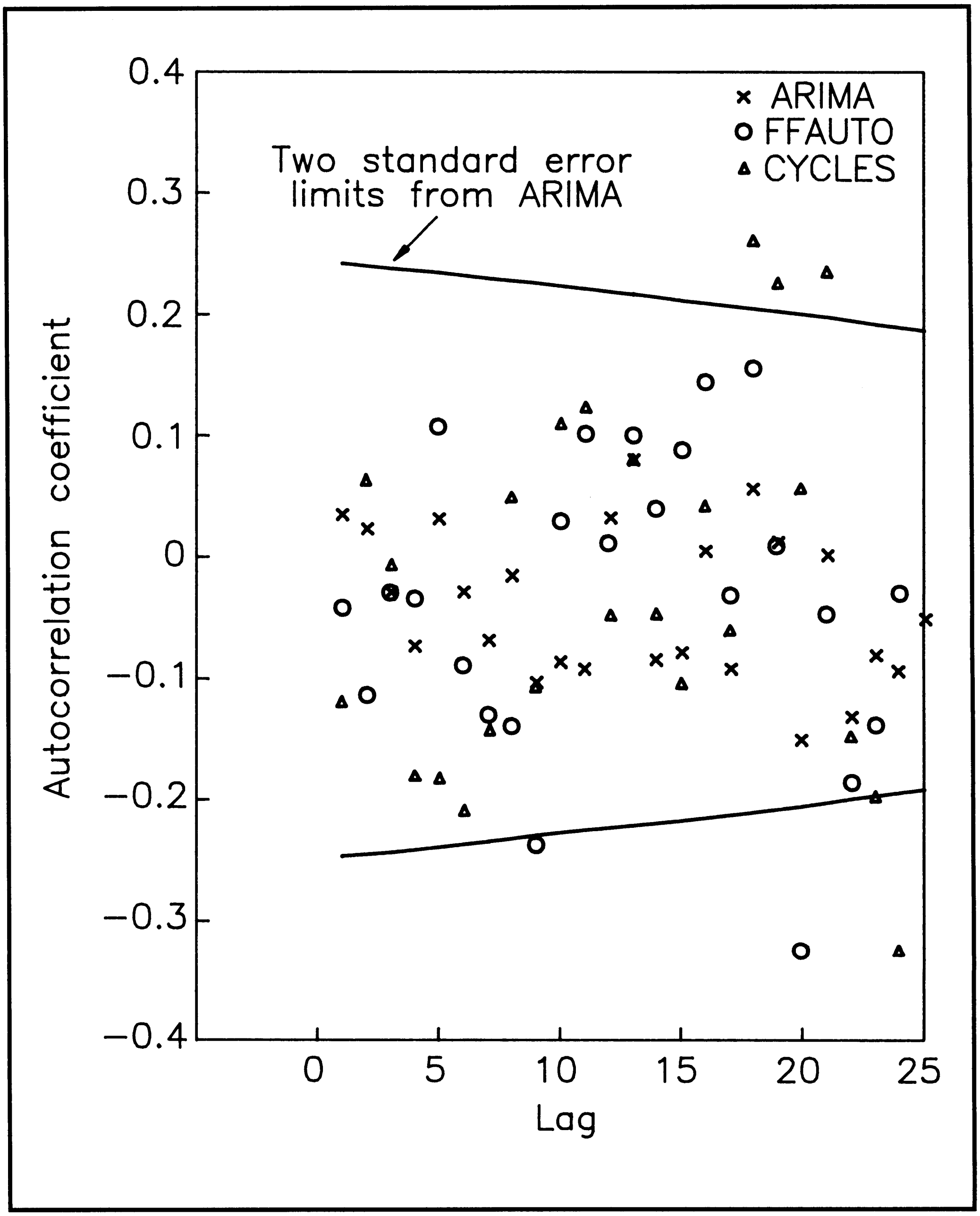


Figure 26. Autocorrelation of residuals for Watkinsville rainfall

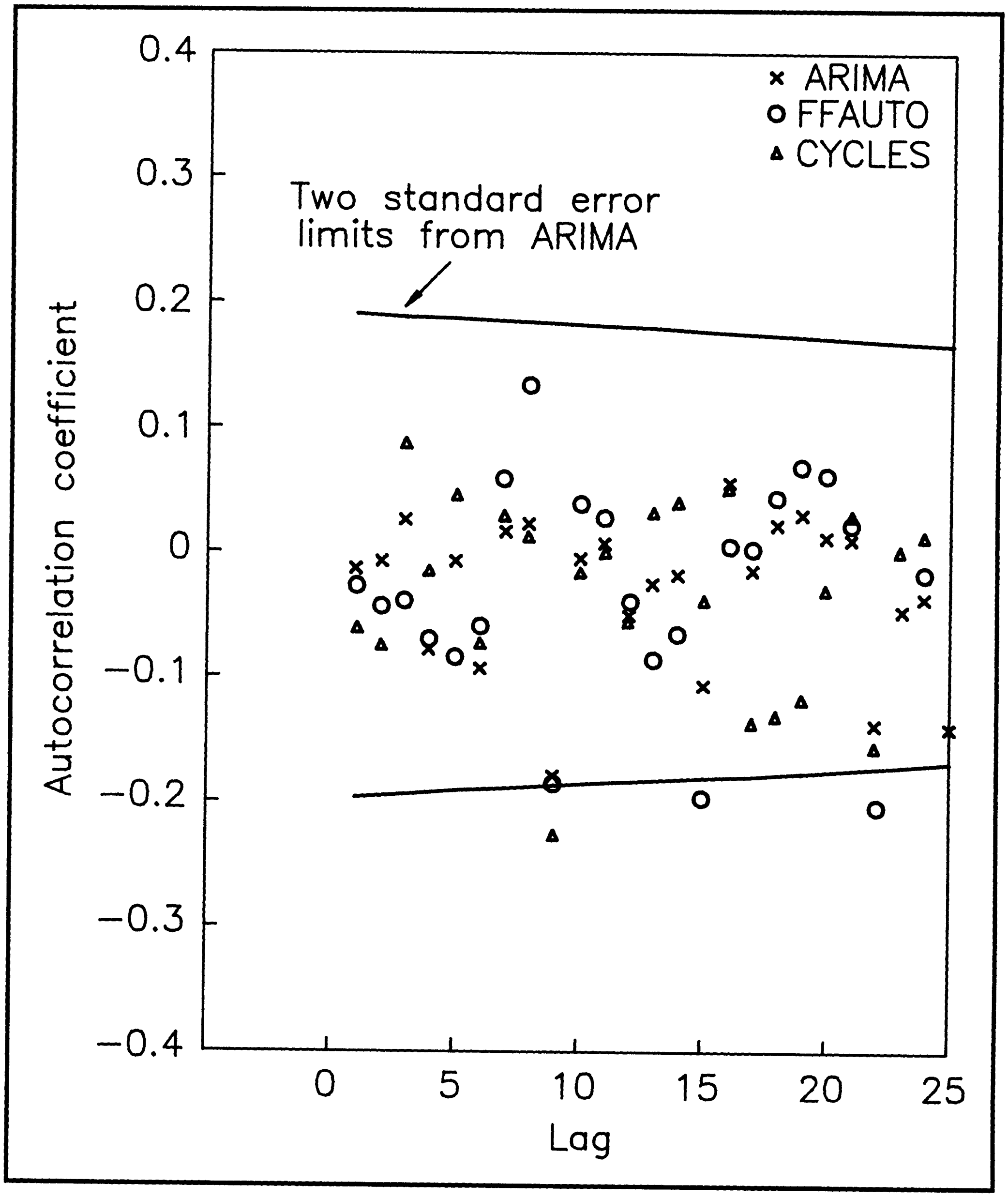


Figure 27. Autocorrelation of residuals for Tifton rainfall