Ecology, 71(6), 1990, pp. 2044–2052

DETECTING CHANGES IN ECOLOGICAL TIME SERIES[1]

ALAN D. JASSBY AND THOMAS M. POWELL[2]
Division of Environmental Studies, University of California, Davis, California 95616 USA

Abstract. Some practical techniques are discussed for analyzing time series whose statistical properties are changing with time. We first consider how principal component analysis can reduce the multidimensional nature of certain series and, in particular, apply this technique to the analysis of changing seasonal patterns. Discussions of trend, changes in oscillatory behavior, and "unusual" events follow. The problem of making inferences regarding causation is briefly considered. We conclude with a call for flexibility in approach.

INTRODUCTION

Time series are ubiquitous in all branches of ecology. Perhaps most of the publications in this discipline present a plot of one or more quantities as a function of time. The analysis of time series is a well-developed branch of statistics, and very readable introductions to the subject have been prepared by Kendall (1976), Chatfield (1989), who includes a brief review of other books on time series methods, and Wei (1990). Shumway (1988) presents an excellent practical treatment of time series analysis applications and there are numerous examples of such applications in the various ecological sciences (Platt and Denman 1975, 1980, Poole 1978, Shugart 1978, Steele 1978). Many of the techniques are also applicable to transect data, i.e., to spatial series (Ripley 1981).

The focus of this article is quite specific: how can one objectively determine whether a time series has changed after some perturbation has occurred? In time series terminology, this question is similar to asking if the series is *stationary,* although here we examine only certain types of nonstationarity. The level of our presentation is limited. There is no attempt to give an introductory overview of time series analysis in ecology, nor to delve into formal mathematical and statistical questions; the references cited above can be used as a guide to these topics. Rather, we offer a few practical comments on selected techniques for analyzing series that are, indeed, changing. Many readers will, no doubt, have some favorite technique that has been omitted.

We first address the important issue of reducing multivariate sets of data to manageable proportions and the related issue of variability in seasonal pattern. These techniques receive the most emphasis, because of their proven applicability to geophysical series and their as yet unrealized promise for analyzing ecological data. We then consider techniques for assessing changes in trend, changes in oscillatory behavior, and the impact of "unusual" events. The treatment of these latter three topics is cursory, as they have already received much attention in the ecological literature. Finally, we examine inference about causation.

SEASONALITY AND REDUCING DIMENSIONALITY

The study of ecosystem response to perturbation often involves an examination of several different variables simultaneously. In the case of an aquatic environment, for example, we might wish to investigate the time course of dissolved nutrients, primary productivity, and the biomasses of various trophic groupings. In general, there may be p univariate time series of interest, where $p > 1$, and together these series can be thought of as forming a $p \times 1$ *vector* time series.

A different but related problem concerns the analysis of a single variable that is indexed on p coordinates in addition to the basic time step t. For example, a single variable measured at p different stations on a lake or ocean surface, at p different depths at a single station, or for p subdivisions of the basic time step (e.g., monthly values when $t = 1$ yr, in which case $p = 12$) also gives rise to a collection of p related univariate series. In this case, however, the p univariate series are records of the same property and are measured in the same units. They can be thought of as forming a *multidimensional* time series, as opposed to a *vector* time series in which the measured properties and perhaps the units are fundamentally different.

Univariate time series methods have various extensions to both the vector and multidimensional cases (e.g., Ripley 1981, Tiao and Box 1981). Applications of these methods are rare, however, for biological data

[1] For reprints of this Special Feature, see footnote 1, page 2037.

[2] The order of authorship was determined by simulated coin tosses.

in the environmental sciences, probably because of the large number of samples required. Here, we examine a different approach, namely, reduction in the dimension of multivariate time series to the univariate case, enabling use of univariate methods that require fewer data.

Various forms of eigenvector analysis, such as principal component analysis (PCA) or common factor analysis, have been applied to reduce the dimension of problems involving multiple variables (Pielou 1984, Jolliffe 1986). The goal in PCA is to replace the original variables by a smaller number of new variables, linear combinations of the original variables, that capture most of the total original variance but are uncorrelated with each other. The new variables are called principal components (PC's) and are arranged in descending order according to the amount of the original variance they reproduce.

Although the number of variables is reduced from, say, p to m ($1 \leq m < p$) with the use of PCA, the ability to interpret the m new uncorrelated variables is not guaranteed. The difficulty arises because principal components are linear combinations of variables that may have an essentially different nature and be measured in completely different units. This difficulty applies to vector time series. No such problem, however, plagues the use of PCA for multidimensional time series and it is this application that we focus on here.

Meteorologists, oceanographers, and hydrologists have successfully applied PCA to time-varying spatial distributions such as rainfall (Richman and Lamb 1985), sea surface temperature (Hsuing and Newell 1983) and streamflow (Lins 1985). Similar analyses, but for coordinates other than those in the horizontal plane, have been published (Denman and Platt 1978). The individual PC's in this context represent the dominant spatial modes of variation for the variable under study (Preisendorfer 1988). PCA also results in a scalar series for each mode, known as the *amplitude time series* or *score,* that expresses the relative importance of each mode over time. Thus, if the first m PC's or modes can account for most of the variability in the original p variables, we have succeeded in reducing the original p-dimensional time series to m univariate series. Often, the first few modes do represent most of the variability in the original data, and m is typically an order of magnitude smaller than p.

Craddock (1965) described a related application of PCA to multidimensional data in which monthly mean temperature is indexed on each month of the year, and the 12 monthly means are followed from year to year. The characteristic patterns in this case can be interpreted as the major modes in which the annual pattern deviates from the long-term mean pattern. Here, we illustrate a similar use of PCA with data for monthly births in the United States from 1948 through 1978 (Shumway 1988: Appendix I, Table 10). The data form a complex long-term pattern, dominated by the "baby boom" peaking around 1960. This peak is best removed to prevent obscuring of the seasonal patterns (Fig. 1). First, a centered 12-term moving average is used to filter the series, resulting in a smooth series representing the long-term behavior of the data. Then the seasonal patterns are estimated by subtracting the smoothed from the original series (e.g., Chatfield 1989). A PCA is applied to the covariance matrix of the 12 extracted monthly variables for the years 1949 through 1977, as filtering results in the loss of 6 mo of data for both 1948 and 1978. As each monthly series is first adjusted by its long-term mean, we are actually examining monthly anomalies. In classical time series terminology, these anomalies are referred to as the *disturbance* component of the series. The first three principal components, which account for 65% of the total variability, are judged to be significant and retained for an orthogonal (varimax) rotation. The first rotated component accounts for 34% of the variability and is characterized by a contrast between negative coefficients in spring, around the time of minimum birth rate, and positive coefficients in summer, around the time of maximum birth rate (Fig. 2A). Thus, the most common form of deviation from the mean seasonal pattern can be thought of as a change in amplitude of the seasonal cycle. The intensity of this first mode, i.e., the amplitude time series, exhibits a decreasing trend between approximately 1954 and 1973, implying that births became distributed more evenly throughout the year (Fig. 2B). An unusually low value for 1951 is also clear. Although these conclusions can be deduced from the original series (Fig. 1), often such structure will not be apparent in the raw data; furthermore, the amplitude time series (Fig. 2B) provides a quantitative description of the structure for further analysis. The remaining rotated components can be investigated in a similar manner; the main point to note here is that a third of the variability for the original multidimensional series has been captured in a single series, which amplifies trends and anomalies that can then be investigated with conventional univariate methods.

The completion of an informative PCA cannot be entrusted solely to the numerous computer programs available. Careful judgment is needed at each step of the process. Certain key technical points are summarized as follows:

1) Sampling plans inevitably change over the years, yet PCA assumes that the coordinates for a multidimensional series are fixed. Interpolation of the data may be required to place it on an equal footing from year to year, in which case care must be taken to avoid the introduction of spurious correlation among inter-

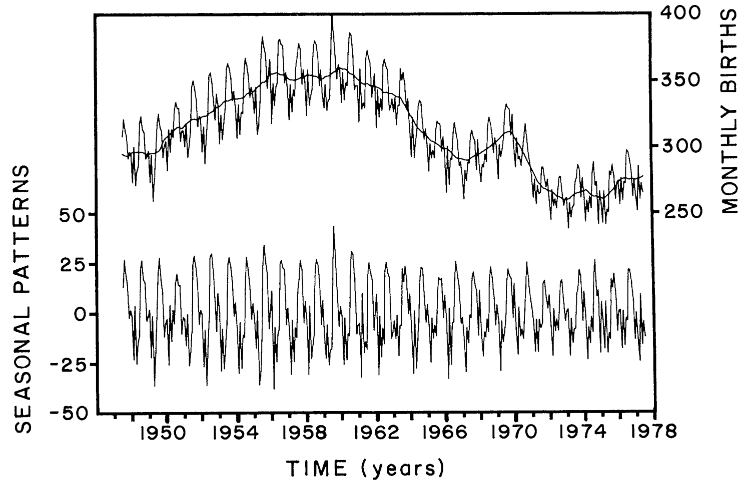

FIG. 1. The upper graph shows monthly births in the United States for the years 1948 through 1978 (data from Shumway 1988: Appendix I, Table 10). The smooth line in the upper graph is the centered 12-term moving average. The bottom graph is the seasonal pattern of monthly births, as determined by the residuals from the moving average.

polated sampling points. Approximation with piecewise polynomials, also known as *splines,* is often appropriate (Wegman and Wright 1983).

2) Coordinates, whether spatial or temporal, for a multidimensional time series should be equally spaced to avoid certain biases that can enter PCA (Karl et al. 1982).

3) The data may require filtering to prevent unwanted frequencies from dominating the analysis, as in the example of monthly births.

4) The form of the dispersion matrix must be chosen, such as the covariance matrix, the correlation matrix, or various robust alternatives that guard against sensitivity to outliers (Devlin et al. 1981, Jolliffe 1986).

5) Various diagnostic tests may suggest that rotation of the principal components is necessary (Richman 1986).

6) If rotation is necessary, an objective determination of *m,* the number of "significant" components to be retained for rotation, is required. Jolliffe (1986) devotes a chapter to this difficult issue, and Preisendorfer (1988) and Stauffer et al. (1985) discuss some Monte Carlo alternatives in an environmental science context. If asymptotic theory is used to make inferences about the number of significant components, then transformation of the data may be required to satisfy multivariate normality. In the case of time series, however, the presence of autocorrelation usually violates the assumption of independent observations and complicates the inference procedures developed for PCA.

DETECTING TREND

Inferences about trend can often be made from the original time series, but sometimes obfuscating fea-

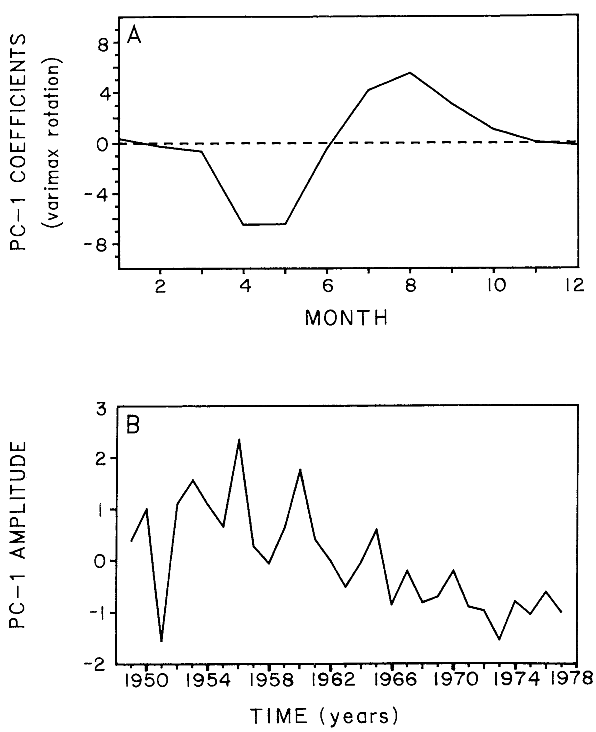

FIG. 2. (A) Coefficients for the first principal component of seasonal birth patterns shown in Fig. 1. The principal component analysis was performed on the monthly anomalies (deviations from the long-term mean for each month) and the results were subjected to a varimax rotation. (B) Amplitude time series for the first principal component illustrated in (A).

tures of the series, such as seasonality or interannual fluctuations, must first be removed. A time series can be thought of as a mixture of several constituents. The birth series of Fig. 1, for example, has been decomposed into a sum of seasonal patterns and a smoothed birth series. A classic approach is to consider the series as being composed of a long-term movement or trend, regular oscillations about the trend, fixed seasonal factors, and residual irregular movements or "disturbances." The trend and other components can be separated by a variety of techniques, e.g., with moving average filters, by taking successive first differences, or by polynomial regression (Chatfield 1989), and the trend then examined in isolation without the complications of the other features.

The decomposition of a series into trend and other components, however, has no unique solution and, to a large extent, personal judgment must be employed in choosing the decomposition technique (Kendall 1976). Usually, components other than the trend are most dependent on the decomposition method, particularly the spectral properties of residuals after trend removal. Linear detrending of a random walk, for example, creates spurious autocorrelation in the residuals at small lags (Chan et al. 1977). A fixed algorithm for approaching time series decomposition will inevitably introduce distortions in certain cases. Also, no decomposition can resolve the difference between a monotonic trend and a natural oscillation with a period much longer than the series length. In some cases, natural populations, e.g., certain marine fish, may undergo cycles with periods measured by centuries (Schindler 1987).

Decomposition of a time series is not always possible, even when desirable. Typical problems with ecological time series include lack of sufficient data, the presence of censored data (i.e., data below the detection limit), missing data, and irregularly spaced sampling times. These same difficulties can prevent application of Box-Tiao intervention analysis (Box and Tiao 1975, Carpenter 1990) to the detection of step trends. In general, then, we require trend detection methods that can overcome these obstacles as well as the various forms of autocorrelation present in time series. Classical statistical tests are inadequate, and typically have additional requirements for constancy of variance and normality that are rarely satisfied by ecological time series. Fortunately, distribution-free tests have been devised recently that permit trend analysis for a wide variety of series, obviating any need for a time series decomposition (Gibbons 1985, Neave and Worthington 1988). A recent issue of *Water Resources Bulletin* (June 1988, Volume 24, Number 3) features the application of these techniques to water quality trends.

The exact character of the time series will determine which of the many distribution-free tests is appropriate. Berryman et al. (1988) list nine tests for monotonic trend, seven tests for step trend, and three tests for multistep trend, as well as an algorithm for deciding among them. Choosing a suitable test depends not only on the type of trend under investigation, but also on the autocorrelation structure, the homogeneity of the separate monthly trends, and the series length. It is also important to understand that distribution-free tests are not a panacea for all problems with data; they require various assumptions, which may be violated for particular data sets.

Detecting Changes in Oscillatory Behavior

Cyclic behavior is common in ecosystems. Strong seasonal variation, for example, is the rule rather than the exception for plant and animal communities (witness the birth rates of Fig. 1). Two statistical quantities are crucial to the study of cyclic phenomena: the autocorrelation function (ACF) and the spectrum (S). The ACF is the basic quantity used to analyze time series in the "time domain," and S performs the same role in the "frequency domain." Many references carefully define and extensively discuss these quantities (e.g., Kendall 1976, Chatfield 1989, Wei 1990). Although the results obtained from consideration of the ACF must be mathematically equivalent to those obtained from consideration of the spectrum, and vice versa, one or the other of the quantities will be much easier to calculate or interpret in practice.

The spectrum allows a straightforward answer to the question: Have the properties of the cyclic variation changed? The spectrum of the time series is compared before and after the presumed perturbation. Is the peak in the same place (i.e., in the same frequency band or bands) both before and after? If yes, then the frequency of the cyclic behavior is unchanged. Is the area under the curve of S "near the peak" the same before and after the perturbation? If yes, then the variance associated with cyclic behavior is unchanged, i.e., the amplitude of the cyclic oscillation is unchanged. In some cases, it may be possible to make these before–after comparisons simply by inspection. Emanuel et al. (1978) illustrate this approach in their study of the effects of three environmental perturbations on a model of forest succession in Appalachian deciduous forests. But most often the changes are slight, and a satisfactory before-after comparison reduces to a question about confidence levels for spectral estimates (Shumway 1989). The Kolmogorov-Smirnov test (Neave and Worthington 1988) can be used for an overall comparison of two spectra.

Two cautions concerning the comparisons of variance before and after a perturbation are in order. Often one is most interested in the relative contribution of a cyclic component to the total variance, not its absolute contribution. One should then normalize the spectra to the same total variance before any comparison is made (e.g., Denman 1976). Second, it is common for the cyclic peak to lie atop some noise background in the curve for *S*. This background should be subtracted before any comparisons are attempted.

One of the gravest difficulties an analyst of cyclic behavior faces is the large amount of data needed to calculate spectra. The (conservative) rule of thumb, that to unambiguously resolve a cyclic component one must record 10 oscillations, places great restrictions on what one can say about low-frequency phenomena, often the phenomena of greatest interest. For series that do not have many entries, as is usually the case with ecological time series, one can often obtain more reliable results from the ACF. To obtain statistically significant *spectral* estimates, one must average over many frequency bands. Hence, few estimates remain and they apply to very wide spectral windows. For the ACF, on the other hand, confidence intervals are relatively narrow. For example, in an annual time series of duration, say, 32 yr (a short time series, but a *very* long ecological time series), the $\approx$95% confidence limits on the ACF are $\approx$0.35. Thus any lagged correlations that are greater than 0.35 (not a large correlation) are significant. Again, one should be confident about only the ACF values for small lags, about one-fourth or less of the series length.

Work on so-called autoregressive spectral estimators suggests that cyclic components can be resolved with much shorter records than suggested above. The best known approach is Maximum Entropy Spectral Analysis (MESA), pioneered by Burg (1975). In an influential note, Ulrych (1972) demonstrated how the MESA technique could resolve spectral components from much shorter records than the common Fourier approach. The method appears to resolve the frequency of a component that contributes less than a full cycle to a time series record! If the spectrum of a process is known to be dominated by one or a few distinct cyclic components, the MESA technique gives excellent frequency resolution. Unfortunately, many spectra from geophysical or ecological records are continuous (Denman 1976, Abbott et al. 1982), and contributions to the time series arise from *all* frequencies, not just a few. In these cases, the spectra calculated from Maximum Entropy techniques will have a very large number of spikes; every rise and fall in the record will look approximately like a half-cycle of some sine or cosine component. One is then forced to average over many spikes (negating the method's high-frequency resolution) to obtain a stable spectral estimate. The MESA technique is superior only when one or a few harmonics dominate, for example, in a tidally dominated environment.

Unusual Events

The term "unusual events" refers to short-term, yet substantial, discontinuities in the underlying behavior of a time series. Unusual events are among the most difficult, but most crucial, phenomena that ecologists (and time series analysts, in general) must handle effectively. Carpenter (1988) illustrates how single unusual events can vastly improve our understanding of an ecosystem and our ability to forecast future behavior. Indeed, this illuminating quality of single events is the rationale behind experimental large-scale perturbations of ecosystems.

Often we may have no a priori knowledge of cause, but our attention is struck by some anomalous short-term behavior of the series, an outlier. The time series of monthly births again presents an excellent example (Fig. 1). The large negative value of the amplitude time series for 1951 is the lowest for the entire 30-yr record (Fig. 2B). As the amplitude represents the contribution of a 12-mo oscillation to the seasonally varying birth rate, the 1951 value corresponds to a year with much less seasonal variation than normal. Inasmuch as the Korean War began in August 1950, the anomalous birth rate in 1951 may reflect the effects of the conflict, but the exact cause of this "unusual" event is beyond our scope here.

How do we decide whether an atypical value merits special consideration; i.e., how can we quantify the concept of "unusual"? If the effects of the outlier on the series are moderate, the outlier can be detected by postulating a probability model for the "underlying" series (Fox 1972, Wei 1990). The probability models most commonly used are ARIMA models (Auto-Regressive-Integrated-Moving-Average models; Box and Jenkins 1976), in which the value of a variable is expressed as a linear combination of previous values, a white noise process, and previous values of the white noise. Extreme outliers make it very difficult to characterize the "underlying" series, however, as they seriously distort the autocorrelation and spectral pattern. On the other hand, they will usually show up in a time plot as indisputably anomalous.

A related problem is posed when an ecosystem is subjected to an intentional perturbation. In this case, the "cause" is known and we are interested in the size and nature of any concomitant changes occurring in the series of interest. As discussed elsewhere (Carpenter 1990), intervention analysis offers a fruitful approach when the timing and qualitative nature of the response is clear from the raw data. Probability models are fit to the "underlying" series and various dynamic models used for the effects of intervention. In addition to the pioneering study of Box and Tiao (1975) on Los Angeles air pollution, intervention analysis has been used in ecology to examine such disparate topics as the effect of the 1965 New York City blackout on birth rates (Izenman and Zabell 1981); power plant impacts on fish populations (Madenjian et al. 1986); and an abrupt decline in British Columbia Dungeness crab landings (Noakes 1986).

Intervention analysis requires subjective judgments regarding the form of the dynamic model used to model the intervention effects. For example, is a step change, an exponential decay back to the long-term mean, or perhaps some other description most accurate? Using Bayesian techniques, a variety of models can be entertained simultaneously, one of which is to be chosen as the best representation of the process (Bolstad 1986; see also Reckhow 1990).

Inference Regarding Causation

Often we are led to investigate the relation between a given ecological time series and a second series that is hypothesized to constitute a perturbing influence: Does one series "cause" the other? Usually the question involves a statistical evaluation of the extent to which the two series covary. Determining how two time series might be associated involves singular issues that are not encountered when comparing *unordered* sets of data. These issues arise from the autocorrelation present in almost all real-world processes when sampled closely enough in time. A number of separate concerns regarding autocorrelation can be delineated.

First, note that certain data transformations of a completely random process can give rise to an autocorrelated series. A well-known example is the "Slutzky-Yule effect," which refers to the way in which moving averages generate apparently systematic oscillations in a random series. Kendall (1976) describes this effect and a related example (due to H. Working) in which aggregation of even a random series, followed by taking the first differences of successive aggregates, will lead to serial correlation. This situation can arise easily with ecological data that are averaged for the year, and then differenced to remove a trend in the annual averages and accentuate interannual variation.

Whether serial correlation is introduced through moving averages or is present in the raw data, it must be accounted for when examining the relationship between two series. The variance of the cross-correlation function is dependent on the autocorrelation present in the individual series, and a large cross-correlation coefficient is not necessarily statistically significant; the effective number of degrees of freedom may have been reduced. The converse problem may also occur, i.e., when a true relationship between two series is obscured by some strong feature of one of the series, such as a trend. One solution is to filter each series to remove the obfuscating features and then compute the cross-correlations between the two residual series (Chatfield 1989). Specifically, a filter is developed to convert the "causal" or input variable to white noise (a process called *prewhitening*), and the filter is applied to both variables (Box and Jenkins 1976). Goldman et al. (1989) give some practical examples of filtering to remove trend and other forms of autocorrelation in an ecological time series (annual lake primary production) before cross-correlation analysis. Although prewhitening filters greatly reduce the problem of arriving at spurious relationships, especially for nonstationary series, the method is conservative and may underplay true causal connections. For example, if two series do, in fact, have a causal relationship that manifests as a trend in each series, this evidence will be masked by the prewhitening process.

Once the series have been filtered to remove obscuring features, the usual tests of association become applicable. If the series have been converted to true white noise, then the cross-correlations greater than $\approx \pm 2/\sqrt{N}$ are significant at the .05 level, where N is the length of the series. If filtering has removed dependence within each series, but the filtered series does not appear to arise from a normal distribution, then Spearman's rank correlation coefficient or the Kendall tau statistic are powerful alternatives (Gibbons 1985).

An alternative to prewhitening is to use theoretical results for the variance of estimated cross-correlations between autocorrelated series. Bartlett (1966) developed an expression for the variance at lag k, $k = 0, 1, 2, \ldots$, in terms of the actual cross-correlations and autocorrelations for the two series at all lags. In practice, the variance must be approximated from this expression by using estimated values for the correlations and truncating the infinite summation to a finite number of lags. Different methods of truncation have resulted in a variety of expressions for the variance or, equivalently, for the effective number of degrees of freedom. Several of these expressions have been compared in Monte Carolo simulations and an interesting conclusion has emerged (Botsford 1987, Kope and Botsford 1988). In cases with weak intra-series correlation, the usual variance estimate for series with no autocorrelation, namely $1/(N - k)$, leads to more accurate rejection rates than all the other expressions. Although the "corrected" expressions make a better estimate of the *average* variance, the *variance* of these estimates is high. Thus, unless the intraseries correlations are strong, it is better to ignore them. Note also that, if either series has zero autocorrelation, Bartlett's expression reduces to the uncorrected expression for the variance under the null hypothesis of no cross-correlation. On the other hand, it is essential to correct the degrees of freedom when both series do have strong internal structure. Drinkwater and Myers (1987) provide an example of this kind of calculation in regard to the relation between fish catch and environmental variables, showing how a failure to account for the loss of degrees of freedom can lead to erroneous conclusions.

In many cases, a series may be measured on a *nominal* scale; for example, the perturbation series may only record the presence or absence of a perturbation at each point in time. These problems are best reduced to an analysis of contingency tables using the chi-squared test of independence (Neave and Worthington 1988). Alexander and Smith (1988) give some examples of this test in assessing the relation between riverine lead concentrations and gasoline consumption. Agresti et al. (1979) describe an alternative procedure for those cases where entries in the contingency table are too small to employ the chi-squared approximation for the test statistic. Unfortunately, the alternative tests are conditional on the observed marginal frequencies of the contingency table, whereas in practice these frequencies often cannot be treated as fixed quantities. (Neave and Worthington [1988] discuss this issue in relation to Fisher's exact test, the alternative for 2 × 2 contingency tables.)

Goldman et al. (1989) encountered this difficulty while investigating the effects of the El Niño/Southern Oscillation (ENSO) phenomenon on annual primary production in Castle Lake, California. The issue was whether or not ENSO years resulted in anomalous production, where "anomalous" was defined as greater than one standard deviation away from the long-term mean. The chi-squared test was inapplicable, and Fisher's exact test required unwarranted assumptions. As an alternative, they chose to examine whether the values for ENSO years had a larger spread or scale than for non-ENSO years, i.e., whether an ENSO year had a greater probability of being extreme. The Siegel-Tukey test (Gibbons 1985) turned out to be appropriate, and indeed the null hypothesis was rejected. Strub et al. (1985) used an identical approach to show that heat storage at Castle Lake tended to be extreme during ENSO years. Perhaps similar analyses will be useful when applied to still higher moments, like the skewness and kurtosis, of ecologically important quantities. One can, of course, test for the equivalence of entire distributions, disregarding considerations of the individual moments, using an appropriate goodness-of-fit test (Neave and Worthington 1988). In summary, hypotheses regarding association can often be rephrased to find a suitable test when the "obvious" choice turns out to be inapplicable.

Statistical evaluations of the extent to which two series covary, regardless of what test is chosen, are not sufficient for establishing significant causal relationships. One only has to point to the numerous correlative studies that have failed the test of time (Walters and Collie 1988). Often an investigator scans numerous series for cross-correlations, in which case the probability of identifying a spurious relationship is greater than the individual probabilities. Statistical evaluations should be completely described, and must be accompanied by a variety of independent measurements supporting a plausible causal mechanism.

Concluding Remarks

Flexibility is an indispensable feature of any attempt to investigate changing time series. A "classical" analysis has been emphasized here in which the series is decomposed into a trend, seasonality, long-term cycles, and residual fluctuations or unusual events. But, as noted earlier, even within the boundaries of a classical decomposition, several lines of attack must be considered. Flexibility also entails the use of approaches other than decomposition of the series. In ARIMA models, for example, there is no recourse to a complete decomposition, although trends and seasonality are removed implicitly through differencing operations. Such models with surprisingly few terms often provide a compact way of expressing lengthy time series. Goldman et al. (1989) present a limnological example.

Occasionally, less formal approaches to modelling the "unperturbed" series are very instructive. Rust and Kirk (1978) provide three examples in which time series for striped bass, atmospheric CO_2, and Canadian lynx are modelled by an iterative procedure using regression and spectral analysis. Regression is used for detrending and spectral analysis for identifying harmonic terms. As exemplified particularly by the study of the CO_2 series, quite complex combinations of effects can be uncovered by this inductive approach. These informal "time series" techniques may not always be productive, but the analyst should not eschew their use simply because they are somehow less elegant.

Finally, recent software developments for microcomputers are proving to be part of the key to flexibility. As is true in so many fields, the advent of powerful desktop personal computing has radically changed the accessibility of time series analysis methods. Virtually every scientist who deals with field or laboratory data in ecology has statistics procedures readily available on a microcomputer. Most of these have quite sophisticated time series components that are easy to use. It is commonly argued that this accessibility increases the misuse of sophisticated statistical techniques by those who are inadequately versed in the underlying theory. Although inappropriate use may sometimes occur, this same accessibility is also a tremendous motivation for learning statistical methods and, in particular, time series analysis. The ease of use and insight gained may surprise even the most skeptical of readers.

ACKNOWLEDGMENTS

We profited from discussions with K. Denman and J. Armengoal. T. M. Powell is grateful to W. T. Edmondson and S. R. Emerson of the University of Washington for their hospitality in Seattle during a portion of the writing of this manuscript. A. D. Jassby acknowledges the support of the National Science Foundation (BSR 87-05170, C. R. Goldman, principal investigator). T. M. Powell thanks the National Science Foundation for its support through grants (OCE 86-13749 and OCE 87-17678).

LITERATURE CITED

Abbot, M. R., T. M. Powell, and P. J. Richerson. 1982. The relationship of environmental variability to the spatial patterns of phytoplankton biomass in Lake Tahoe. Journal of Plankton Research **4**:927–941.

Agresti, A., D. Wackerly, and J. M. Boyett. 1979. Exact conditional tests for cross classification: approximation of attained significance levels. Psychometrika **44**:75–83.

Alexander, R. B., and R. A. Smith. 1988. Trends in lead concentration in major U.S. rivers and their relation to historical changes in gasoline-lead consumption. Water Resources Bulletin **24**:557–569.

Bartlett, M. S. 1966. An introduction to stochastic processes. Second edition. Cambridge University Press, Cambridge, England.

Berryman, D., B. Bobée, D. Cluis, and J. Haemmerli. 1988. Nonparametric tests for trend detection in water quality time series. Water Resources Bulletin **24**:545–556.

Bolstad, W. M. 1986. Harrison-Stevens forecasting and the multiprocess dynamic linear model. American Statistician **40**:129–135.

Botsford, L. 1987. Analysis of environmental influences on population dynamics. *In* T. L. Vincent, V. Cohen, W. J. Grantham, G. P. Kirkwood, and J. M. Skowronski, editors. Modeling and management of resources under uncertainty. Lecture Notes in Biomathematics **72**:54–64.

Box, G. E. P., and G. M. Jenkins. 1976. Time series analysis, forecasting and control. Revised edition. Holden-Day, San Francisco, California, USA.

Box, G. E. P., and G. C. Tiao. 1975. Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association **70**: 70–79.

Burg, J. P. 1975. Maximum entropy spectral analysis. Dissertation. Stanford University, Stanford, California, USA.

Carpenter, S. R. 1988. Transmission of variance through lake food webs. Pages 119–135 *in* S. R. Carpenter, editor. Complex interactions in lake communities. Springer-Verlag, New York, New York, USA.

———. 1990. Large-scale perturbations: opportunities for innovation. Ecology **71**:2038–2043.

Chan, K. H., J. C. Hayya, and J. K. Ord. 1977. A note on trend removal methods: the case of polynomial regression versus variate differencing. Econometrica **45**:737–744.

Chatfield, C. 1989. The analysis of time series: an introduction. Chapman and Hall, London, England.

Craddock, J. M. 1965. A meteorological application of principal component analysis. Statistician **15**:143–156.

Denman, K. L. 1976. Covariability of chlorophyll and temperature in the sea. Deep-Sea Research **23**:539–550.

Denman, K., and T. Platt. 1978. Time series analysis in marine ecosystems. Pages 227–242 *in* H. H. Shugart, Jr., editor. Time series and ecological processes. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.

Devlin, S. J., R. Gnanadesikan, and J. R. Kettenring. 1981. Robust estimation of dispersion matrices and principal components. Journal of the American Statistical Association **76**:354–362.

Drinkwater, K. F., and R. A. Myers. 1987. Testing predictions of marine fish and shellfish landings from environmental variables. Canadian Journal of Fisheries and Aquatic Sciences **44**:1568–1573.

Emanuel, W. R., H. H. Shugart, Jr., and D. C. West. 1978. Spectral analysis and forest dynamics: the effects of perturbations on long-term dynamics. Pages 193–207 *in* H. H. Shugart, Jr., editor. Time series and ecological processes. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.

Fox, A. J. 1972. Outliers in time series. Journal of the Royal Statistical Society, Series B **43**:350–363.

Gibbons, J. D. 1985. Nonparametric methods for quantitative analysis. American Sciences Press, Columbus, Ohio, USA.

Goldman, C. R., A. Jassby, and T. Powell. 1989. Interannual fluctuations in primary production: meteorological forcing at two subalpine lakes. Limnology and Oceanography **34**: 310–323.

Hsuing, J., and R. E. Newell. 1983. The principal nonseasonal modes of variation of global sea surface temperature. Journal of Physical Oceanography **13**:1957–1967.

Izenman, A. J., and S. L. Zabell. 1981. Babies and the blackout: the genesis of a misconception. Social Science Research **10**:282–299.

Jolliffe, I. T. 1986. Principal component analysis. Springer-Verlag, New York, New York, USA.

Karl, T. R., A. J. Koscielny, and H. F. Diaz. 1982. Potential errors in the application of principal component (eigenvector) analysis to geophysical data. Journal of Applied Meteorology **21**:1183–1186.

Kendall, M. 1976. Time-series. Charles Griffin, London, England.

Kope, R. G., and L. W. Botsford. 1988. Detection of environmental influence on recruitment using abundance data. Canadian Journal of Fisheries and Aquatic Sciences **45**: 1448–1458.

Lins, H. F. 1985. Streamflow variability in the United States: 1931–1978. Journal of Climatology and Applied Meteorology **24**:463–471.

Madenjian, C. P., D. J. Jude, and F. J. Tesar. 1986. Intervention analysis of power plant impact on fish populations. Canadian Journal of Fisheries and Aquatic Sciences **43**: 819–829.

Neave, H. R., and P. L. Worthington. 1988. Distribution-free tests. Unwin Hyman, London, England.

Noakes, D. 1986. Quantifying changes in British Columbia Dungeness crab landings using intervention analysis. Canadian Journal of Fisheries and Aquatic Sciences **43**:634–639.

Pielou, E. C. 1984. The interpretation of ecological data: a primer on classification and ordination. John Wiley, New York, New York, USA.

Platt, T., and K. L. Denman. 1975. Spectral analysis in ecology. Annual Review of Ecology and Systematics **6**:189–210.

Platt, T., and K. L. Denman. 1980. Patchiness in phytoplankton distribution. Pages 413–431 *in* I. Morris, editor.

The physiological ecology of phytoplankton. University of California Press, Berkeley, California, USA.
Poole, R. W. 1978. The statistical prediction of population fluctuations. Annual Review of Ecology and Systematics **9**: 427–448.
Preisendorfer, R. W. 1988. Principal component analysis in meteorology and oceanography. Developments in Atmospheric Science 17. Elsevier, New York, New York, USA.
Reckhow, K. H. 1990. Bayesian inference in non-replicated ecological studies. Ecology **71**:2053–2059.
Richman, M. B. 1986. Rotation of principal components. Journal of Climatology **6**:293–335.
Richman, M. B., and P. J. Lamb. 1985. Climatic pattern analysis of three- and seven-day summer rainfall in the central United States: some methodological considerations and a regionalization. Journal of Climate and Applied Meteorology **24**:1325–1343.
Ripley, B. D. 1981. Spatial statistics. Wiley, New York, New York, USA.
Rust, B. W., and B. L. Kirk. 1978. Inductive modelling of population time series. Pages 154–192 *in* H. H. Shugart, Jr., editor. Time series and ecological processes. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.
Schindler, D. W. 1987. Detecting ecosystem responses to anthropogenic stress. Canadian Journal of Fisheries and Aquatic Sciences **44**(Supplement):6–25.
Shugart, H. H., Jr., editor. 1978. Time series and ecological processes. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.
Shumway, R. H. 1988. Applied statistical time series analysis. Prentice-Hall, Englewood Cliffs, New Jersey, USA.
Stauffer, D. F., E. O. Garton, and R. K. Steinhorst. 1985. A comparison of principal components from real and random data. Ecology **66**:1693–1698.
Steele, J. H., editor. 1978. Spatial pattern in plankton communities. Plenum, New York, New York, USA.
Strub, P. T., T. Powell, and C. R. Goldman. 1985. Climate forcing: effects of El Niño on a small, temperate lake. Science **227**:55–57.
Tiao, G. C., and G. E. P. Box. 1981. Modelling multivariate time series with applications. Journal of the American Statistical Association **76**:802–816.
Ulrych, T. J. 1972. Maximum entropy power spectrum of truncated sinusoids. Journal of Geophysical Research **77**: 1396–1400.
Walters, C. J., and J. S. Collie. 1988. Is research on environmental factors useful to fisheries management? Canadian Journal of Fisheries and Aquatic Sciences **45**:1848–1854.
Wegman, E. J., and I. W. Wright. 1983. Splines in statistics. Journal of the American Statistical Association **78**:351–365.
Wei, W. W. S. 1990. Time series analysis: univariate and multivariate methods. Addison-Wesley, New York, New York, USA.