A method to determine rates and patterns of variability in ecological communities

Scott L. Collins, Fiorenza Micheli and Laura Hartt


It is well known that ecological communities are spatially and temporally dynamic. Quantifying temporal variability in ecological communities is challenging, however, especially for time-series data sets of less than 40 measurement intervals. In this paper, we describe a method to quantify temporal variability in multispecies communities over time frames of 10–40 measurement intervals. Our approach is a community-level extension of autocorrelation analysis, but we use Euclidean distance to measure similarity of community samples at increasing time lags rather than the correlation coefficient. Regressing Euclidean distances versus increasing time lags yields a measure of the rate and nature of community change over time. We demonstrate the method with empirical data sets from shortgrass steppe, old-field succession and zooplankton dynamics in lakes, and we investigate properties of the analysis using simulation models. Results indicate that time-lag analysis provides a useful quantitative measurement of the rate and pattern of temporal dynamics in communities over time frames that are too short for more traditional autocorrelation approaches.

It is well documented that populations and communities are spatially and temporally variable (Dutilleul and Legendre 1993, Hastings et al. 1993). Variation, however, comes in diverse forms, such as cyclic (microtine rodents, Bjørnstad et al. [1999]), directional (plant succession, Myster and Pickett [1994]), stochastic (patch dynamics, Glenn and Collins [1990]), or chaotic (predator-prey cycles, Schaffer [1985], Hastings et al. [1993]). Moreover, mathematical models have shown that different temporal dynamics are generated by a number of underlying mechanisms (May 1973, DeAngelis and Waterhouse 1987, Tilman and Wedin 1991, Hanski and Gyllenberg 1997). Although laboratory experiments have yielded patterns that support some theoretical predictions derived from these models (e.g., Constantino et al. 1995, Dennis et al. 1995), our ability to test theoretical predictions in more complex natural systems has been limited by a lack of long-term data sets (Tilman 1989).

Investigations of temporal variability in complex natural systems have also been limited by the lack of adequate statistics to quantify and compare variability across communities each comprised of many species (Micheli et al. 1999). The metrics and analytical tools for describing and comparing temporal variation in aggregate community variates, such as total density or biomass, are relatively well developed. Commonly used variability metrics include the standard deviation or the coefficient of variation of the density or biomass time-series data (Gaston and McArdle 1994, Doak et al. 1998). However, it is often critical to quantify temporal variation in the relative abundance of species in a
community. For example, anthropogenic perturbations, such as acidification of lake waters, can cause no change in the total biomass of the plankton, but can cause dramatic changes in the relative abundance of different species (Frost et al. 1995). Community composition, and possibly ecological processes, may be altered profoundly without detectable changes in aggregate community variates (Frost et al. 1995, Micheli et al. 1999).

In this paper, we present a novel analytical method that can be used to measure the degree of cyclical, stochastic, and/or directional patterns in relatively short time-series data sets. The analytical approach is related to temporal autocorrelation. Our method measures similarity in community composition at increasing time lags to determine the degree of temporal variability and the potential for clear patterns of change over time. In this paper, we describe our analytical approach and demonstrate the utility of the method using simulation models and empirical data from vegetation in short-grass prairie, an old-field successional sere, and zooplankton community dynamics in an experimental lake.

**Methods**

**Analytical approach**

The analytical approach is a straightforward analysis of differences in species composition between samples at increasing time lags (Fig. 1). The first step in the analysis is calculation of a triangular dissimilarity matrix from the species-by-time rectangular data matrix (i.e., species being the variables and samples in time being the observations). For community data we recommend using Euclidean distance because of the clear geometric properties of this metric, although other metrics may be appropriate, as well. Euclidean distance is calculated as:

\[
ED = \left( \sum_{i=1}^{S} (x_{ij} - x_{ik})^2 \right)^{1/2}
\]

where \( x_{ij} \) is the abundance of the \( i \)th species in the \( j \)th sample, and \( x_{ik} \) is the abundance of the \( i \)th species in the \( k \)th sample, and \( S \) is the total number of species. As used here, samples correspond to measurements of community composition over time. Because differences are squared, ED emphasizes change in abundant species from one sample to the next (Ludwig and Reynolds 1988). Other measures of similarity (quantitative, rank order, or presence/absence) could be used for time-lag analyses depending on the research questions being asked and the abundance measurements of the constituent species.

Next, the Euclidean distance values are plotted against time lag for all lags below the diagonal in the triangular resemblance matrix. For example, a data set of 10 species sampled over 6 time intervals would have five one-year time lags (year 1 vs year 2, year 2 vs year 3, . . . year 5 vs year 6), four two-year time lags (year 1 vs year 3, year 2 vs year 4, etc.) through one six-year time lag (year 1 vs year 6), for a total of 15 time lags over the entire data set (Fig. 1). Linear and non-linear regressions can then be calculated for Euclidean distance as a function of the square root of the time lag. The square root transformation reduces the probability that the smaller number of points at larger time lags will bias the analysis. It is important to note that we use regression analysis as a means to summarize trends in temporal data sets, not as a statistical method to forecast or model temporal change. Our approach is conceptually similar to that of Venrick (1990) who used Kendall’s correlation coefficients (\( t \)) between the rank order of abundance in samples separated by increasing time intervals to describe long-term trends and community persistence of phytoplankton in the central North Pacific. Our goal, however, is to demonstrate the ability of time-lag analysis to discriminate among different patterns of community change (stochastic variation, directional change, and cyclical dynamics).

The time-lag analytical approach can produce a number of general theoretical patterns with time-series data (Fig. 2). If the regression line is significant, positive, and linear, then it implies that the assemblage in question is undergoing directional change. If the regression line is not significant or the slope is not significantly different from zero, then it implies fluctuation or stochastic variation over time. If the slope of the line is negative, then it implies species composition is converging on a community-type characteristic of one of the early sample periods. This result might occur when temporal change is non-linear as in cyclical dynamics (van der Maarel 1988, Turchin 1993, Bjørnstad et al. 1999) or community resilience following perturbation (e.g., Loucks 1970). In general, the slope of the regression line indicates the rate and direction of change, and the
Fig. 2. Some theoretical possibilities with time-series data. If the distance between samples does not change as time-lags increase, then the community is considered to be stable. If sample distance increases over time, the community is unstable and undergoing directional change. If sample distance decreases over time, then the community is unstable and undergoing convergence.

regression coefficient is a measure of signal versus noise. For example, a significant positive relationship ($P < 0.01$) with a small slope and a small $r^2$ value suggests that directional change is occurring, but change is slow and stochastic variation between sample intervals is high. A steeper slope and larger $r^2$ value would indicate a stronger signal of directional change and less noise.

**Empirical data and simulations**

To evaluate the utility and some properties of this approach, we conducted time-lag analyses on three empirical data sets and constructed simulation models that reflect some of the general properties of community change shown within those data sets. The first data set from shortgrass prairie was used to demonstrate the analytic response of vegetation presumed to exhibit stochastic variation over time. Species cover values were estimated in 25 0.1-m$^2$ permanent quadrats in an ungrazed shortgrass prairie in Morton County, KS (Glenn and Collins 1992). Vegetation was sampled annually from 1962 until 1980. Annual cover values were determined by averaging species cover in the 25 quadrats. Given the lack of disturbance in this grassland, the time-lag analysis is predicted to show high variation, and little or no slope for the regression.

A second data set on old-field succession from the Hutchinson Memorial Forest (HMF) long-term succession experiment (Small et al. 1971, Myster and Pickett 1988, 1994) was used to test whether a community undergoing directional succession (Myster and Pickett 1988, 1994) shows the expected positive, linear relationship in the time-lag analysis. Field 1 at HMF was left fallow in the autumn of 1957. Forty-eight permanent 1-m$^2$ quadrats were established in the first year following abandonment. Plant species composition has been sampled annually in this old field from 1958 to 1980, and every other year thereafter. For this analysis, we used the first 20 years of annual data for our time-lag analysis.

The third data set is derived from zooplankton dynamics in Little Rock Lake, Wisconsin (Frost et al. 1995). In 1984 the two basins of this lake were divided by a curtain. After a baseline period, one basin was acidified incrementally to a pH of 4.7 over a six-year period, and the control basin remained at pH 6.1. After the six years, acidification treatments were stopped, and the acidified basin was allowed to recover for five years. Zooplankton abundances (measured as biomass) were sampled over this time period in the treatment and control basins (Frost et al. 1995). For the respective time-lag analyses, the control basin should show stochastic variation (no slope) because it was never disturbed, while the treatment basin should show divergence (positive slope) from the initial community composition during acidification, followed by convergence (negative slope) during the recovery phase.

In addition to the time-series analyses, all three empirical data sets were subjected to a DCA ordination to help visualize general patterns in community change over time. Ordinations were performed independently on the three species-by-time matrices. In all cases, species were deleted if they occurred in less than 10% of samples in each time-series data set, and the remaining rare species were downweighted to minimize distortion in the ordination (Gauch 1982).

We used simulation models to test some of the basic properties of the time-lag analyses. We used a random-walk approach to simulate community change in multi-dimensional (multispecies) space. This null approach assumed that autocorrelated movement along any single axis was independent of movement along any of the other axes (Berg 1993). Thus, changes in species abundances may be autocorrelated over time but are independent of one another. Two versions of the random-walk model were constructed, one for stochastic dynamics and the other for directional change. In both versions, the initial community position in Euclidean space was found by selecting values at random from a uniform distribution of starting values along each species axis. In the stochastic version, the position along each axis was randomly chosen at every time step. In the directional version, an attractor was set at a random positive distance away from the initial value along each axis. Movement along each axis at every time step was random but bounded by a percentage of the distance between the position and the attractor at the previous time step. Both versions assumed a maximum time-series length of 20 units, a total of 20 species (axes), and 25 replicate communities. The simulations
were performed using Matlab (Mathworks 1997). As for the empirical data sets, we analyzed the simulation outputs by calculating linear regressions of the square root of the time lag (independent variable) versus Euclidean distance (dependent variable).

**Results**

**Shortgrass prairie**

The ordination of the shortgrass prairie plant community suggests that this vegetation changes stochastically over time. Based on the ordination, community composition in 1962 is similar to that in 1980. The time-lag analysis supports this conclusion. The slope of the regression line is not significantly different from zero and the overall regression is not significant, both factors indicating stochastic variation over time (Fig. 3A).

The stochastic model simulations yield comparable results in that the regressions are not significant (F-statistic ranges from 0.6 to 0.86), with relatively poor fits to the predicted lines ($r^2 < 0.002$) and slopes close to zero ($-0.6 < m < 1.8$) (Fig. 4A–C). Increasing the bounds for positions at each time step (from 100 to 200 to 500), analogous to increasing the amount of stochastic variation possible at each time step, results in higher Euclidean distance values but does not change the linear relationships appreciably.

**Hutchinson Memorial Forest succession data**

The ordination of HMF data from field 1 exhibits a clear pattern of directional change in composition over time. This pattern reflects compositional differences as the vegetation changes from annuals early in succession to dominance by a series of perennial herbaceous species, through the invasion of woody vegetation. Temporal variation from year 1 to 2 and from year 2 to 3 is very high as the community changes dramatically during the early years of succession (Myster and Pickett 1988). After this compositional reorganization occurs, the trajectory settles into a clear linear trend in the ordination. This conspicuous directional change in the ordination is reflected in the strong linear and positive regression of the time-lag analysis (Fig. 5). The low $r^2$ (0.05) reflects the uniqueness of vegetation in years one and two. When years one and two are removed from the regression, the $r^2$ increases dramatically to 0.61.

The directional change model simulations, likewise, produced significant regressions with positive slopes and good fits to the predicted lines ($r^2 > 0.7$). In this case, the assumption is that the larger the distance between the starting point and the attractor the longer the successional sequence. The simulations showed that increasing the distance to the attractor resulted in larger slopes ($m$ increasing from 55.1 to 105.9 to 200.7 for 10% movement, Fig. 6A, B, D), suggesting that community dissimilarity increased as a function of the distance to the attractor. For any given distance to the attractor, increasing the possible amount of change in abundance at each time step also increased the slope (from 105.9 to 124.5 for attractor distances up to 200, Fig. 6B, C) and the variance, thereby reducing the $r^2$ from 0.92 to 0.71. Increasing attractor distance also resulted in larger Euclidean distances after the first time lag (Fig. 6A, B, D).

**Little Rock Lake**

The ordination of compositional trends for the treatment and control basins of Little Rock Lake clearly showed the effects of the experimental acidification. The treated half exhibits directional change away from the control followed by a recovery phase after the cessation of acid additions (Fig. 7A). Zooplankton...
Time-lag analysis of simulated data showing stochastic community dynamics. Random-walk models were used to generate dynamics of communities with 20 species over 20 time steps. Each point is an average of 25 replicates. Species abundances change independently from one time step to the next. Maximum variability was constrained to (A) 100, (B) 200 and (C) 500 abundance units per species at each time step.

Discussion

Analyses of empirical datasets indicate that the time-lag method effectively discriminated between stochastic, directional and cyclical patterns of community variation in complex multispecies assemblages. In addition, simulation models demonstrated that this method was robust over a range of values of the degree and rate of community change. We do not intend to provide a detailed description of the dynamics of each empirical data set because community dynamics in these systems have been described elsewhere (e.g., Myster and Pickett 1988, 1994, Glenn and Collins 1992, Frost et al. 1995). Rather, our purpose was to determine if the time-lag analytical approach could capture the essence of the dynamics in these thoroughly described and somewhat predictable systems. If so, we can then use this analysis with some confidence when analyzing data sets where the patterns of temporal change are less obvious. In addition, the method allows a quantitative comparison among sets of data (comparison of slopes), along with a means to extract general long-term trends in systems with high inter-sample variability.

The time-lag analysis elicited the expected pattern of change in each empirical data set. As predicted, stochastic variability occurred in the undisturbed shortgrass steppe vegetation (Fig. 3). The high degree of interannual variation in this plant community resulted from stochastic immigration and extinction dynamics (Glenn and Collins 1992) and interannual variation in climate. Indeed, other evidence suggests that shortgrass steppe vegetation is highly stable even under heavy grazing by domestic cattle (Milchunas et al. 1990, 1998). More recently, changes in community structure in shortgrass steppe vegetation have been reported in response to increasing minimum annual temperatures (Alward et al. 1999), but such directional changes were not evident in our data set.

Directional change was evident in the old-field succession data set, even though initial interannual variation was quite high. This reflects the unique starting conditions of early successional communities dominated by many short-lived annual species (Pickett 1982) and leads to the caveat that results from time-lag communities in the control half of the lake fluctuate stochastically over time (Fig. 7B). The time-lag analysis captures these dynamics as predicted. Time-lag regression analysis of the control half has high variability at all time lags, and no significant slope (Fig. 7C). The treatment basin, on the other hand, exhibited directional change for the early time lags, followed by convergence as this lake basin recovered from acidification. The regression is significant and non-linear as predicted.

Fig. 4. Time-lag analysis of simulated data showing stochastic community dynamics. Random-walk models were used to generate dynamics of communities with 20 species over 20 time steps. Each point is an average of 25 replicates. Species abundances change independently from one time step to the next. Maximum variability was constrained to (A) 100, (B) 200 and (C) 500 abundance units per species at each time step.
analysis, like other time series analyses, are somewhat dependent on initial sample composition. However, by year three, stochastic dynamics on this abandoned field decreased and a clear pattern of linear, directional change in community composition was evident. Overall, results from this successional data set demonstrated that starting conditions can create non-linear patterns and potentially increase temporal heterogeneity in some dynamic systems.

The time-lag analysis of zooplankton composition in the treatment and control lake basins documented the stochastic annual variation in the control half of the lake and the divergence-convergence dynamics in the treatment half. This result clearly illustrated that our analytical approach can also quantify non-linear trends in community dynamics.

With respect to the model simulations, because the random-walk represents a null model where changes in species abundances are independent of one another, it represents a good first approximation of variation in community composition over time. The model output clearly shows that the time-lag analysis method effectively captures community dynamics when those dynamics are believed to be due to stochastic or directional processes. Still, the time-lag approach is designed to detect patterns in time series data, it does not provide a mechanistic understanding of the cause of temporal change in a community.

In all of our empirical cases, the time-lag analysis provided a quantitative assessment of change that matched the general temporal variation patterns previously documented in these communities, and also shown less clearly using DCA ordination. This begs the question, “Why bother?” However, time-lag analysis has several advantages that cannot be derived from a perceived understanding of general patterns or using standard statistical techniques. First, the analysis can be applied to data sets in which the general temporal trends are not obvious, and thus it can be used as a
diagnostic tool when the particular type of temporal variation of a community is not known. Second, this method also reduces complex, multivariate changes in a suite of species to a simple univariate metric describing community change. Third, by comparing slopes, the method provides a quantitative measure of the rate of change over time that can be used to compare differences among data sets. Fourth, by comparing $r^2$ values, the analysis yields a measure of stochasticity along with a measure of rate of change. In combination, these measurements provide strong inference on the nature and pattern of temporal variability in multispecies communities.

Results of time-lag analysis of long-term data on plant and animal communities from tallgrass prairie (Konza Prairie Research Natural Area, KS, USA) illustrate this point. Time-lag analysis was used to quantify temporal dynamics of plant, grasshopper, small mammal, and breeding bird communities in mature tallgrass prairie subjected to 1-yr, 4-yr or 20-yr fire frequencies (Collins 2000). The time-lag analysis rendered several surprising results. First, plant communities subjected to annual spring burning exhibited little interannual variation and strong directional change over a 15-yr time frame. In contrast, unburned sites were highly variable from one year to the next and directional change was weak. The opposite was expected given the fundamental role of fire in maintaining tallgrass ecosystems (Daubenmire 1968). Grasshopper, small mammal, and breeding bird communities all showed high interannual variation and little directional change regardless of burning interval. That is, animal community dynamics were decoupled from plant community dynamics in this tallgrass ecosystem. None of these patterns was evident prior to this time-lag analysis (Collins 2000).

Our analyses were based on relatively straightforward patterns of temporal change in different empirical data sets. More complex cyclical and non-linear patterns of change may occur in some communities, and
Fig. 6. Time-lag analysis of simulated data showing directional change. Random-walk models were used to generate dynamics of communities with 20 species over 20 time steps. Each point is an average of 25 replicates. Species abundances change independently from one time step to the next, but in this case directional change is generated by constraining species to move towards an attractor located at increasing distances from a random starting point.

such complexity deserves further exploration with simulated and empirical data sets. Time-lag analysis is useful for determining the rate and pattern of variation, but it does not provide a mechanistic understanding of what causes change over time in a community. For example, random drift rather than successional dynamics could cause directional change in a community. Random drift may occur in forest communities, for example, if species gained or lost some fraction of abundance at each time step via gap dynamics and these changes at each time step are uncorrelated. In this random walk process, the position at time $t+1$ is highly correlated with current position (time $t$), but changes are uncorrelated. Over time, this could lead to random drift away from the starting point that would look like directional change.

To test whether or not random drift could lead to directional change, we simulated drift with three levels of proportional change (10%, 25% and 50%) and three ranges in species abundance units (0–100, 0–200 and 0–500 abundance units). For moderate levels of proportional change regardless of range in abundance values we found significant linear regressions with positive slopes ($r^2 > 0.9$ in all cases). Thus, time-lag analysis can provide a useful measure of change over time and serve as a means to derive hypotheses regarding the mechanisms that produce change in communities.

In summary, although patterns of directional, cyclical and stochastic variation may be predictable in some systems prior to quantitative analysis, time-lag analysis provides a statistical evaluation of trends in time-series data that are not long enough to be subjected to more traditional forms of time-series analysis. This is especially true given the high levels of inter-annual variability that are inherent in many time-series data sets. The overall utility of this approach can only be assessed by applying the technique to different data sets collected from different systems, at different spatial scales and over variable lengths of time. In addition, simulation models exploring the trends generated by different mechanisms of community change will help establish a link between patterns of community variability and underlying processes. For example, models could examine the patterns generated when species interactions are added to our model assuming that species are indepen-
dent of each other. Also, simulations may be used to compare the effects of different disturbance regimes, by varying the magnitude and frequency of perturbations imposed on the simulated system. Our preliminary analyses suggest, however, that time-lag analysis provides a valuable quantitative measurement of the rate and pattern of temporal dynamics in communities over time frames that are too short for more traditional autocorrelation approaches.

Acknowledgements – We are extremely grateful to Peter Kareiva for suggesting this analytical approach for multispecies time-series data sets and for encouraging us to pursue this analysis, and to Bill Murdoch for helpful comments on the analysis. We thank Tom Frost, Tim Kratz and Steward Pettet for allowing us to test our method on their long-term data sets. Bruce Kendall, Don Jackson and Brian Shuter provided many helpful comments on earlier versions of the manuscript. Portions of this research were completed at the National Center for Ecological Analysis and Synthesis, a center funded jointly by the National Science Foundation (DEB-9421535), the University of California – Santa Barbara, the California Resources Agency, and the California Environmental Protection Agency.

References