

University of Massachusetts Amherst

From the Selected Works of Joseph Elkinton

June, 2005

Measuring and testing for spacial synchrony

Joseph Elkinton, *University of Massachusetts - Amherst*

W. A Ruscoe

D. Choquenot

R. B Allen



Available at: https://works.bepress.com/joseph_elkinton/4/

MEASURING AND TESTING FOR SPATIAL SYNCHRONY

JOHN P. BUONACCORSI,¹ JOSEPH S. ELKINTON,² SCOTT R. EVANS,^{3,5} AND ANDREW M. LIEBHOLD⁴

¹Department of Mathematics and Statistics, University of Massachusetts, Amherst, Massachusetts 01003 USA

²Department of Entomology and Graduate Program in Organismic and Evolutionary Biology,
University of Massachusetts, Amherst, Massachusetts 01003 USA

³Department of Biostatistics, University of Massachusetts, Amherst, Massachusetts 01003 USA

⁴Northeast Research Station, USDA Forest Service, Morgantown, West Virginia 26905 USA

Abstract. Spatial synchrony in abundance among populations at different locations has been studied for many species. Different statistics have been used as measures of synchrony, and various techniques have been employed to test the hypothesis that there is no synchrony. In this paper we first describe and contrast various measures of synchrony and then discuss testing for no synchrony. Tests that ignore the serial correlation are commonly employed but are incorrect if there is serial correlation present, as is often the case with populations followed over time. Alternative approaches and their limitations are presented including tests based on residuals, adjusted degrees of freedom tests, and bootstrap procedures. We recommend tests based on residuals in a model-based setting. We also discuss some of the difficulties of finding model-free approaches and suggest some methods based on confidence intervals for future study.

Key words: autoregressive models; bootstrapping; concordance; correlation; hypothesis testing; Moran effect; spatial synchrony; time series.

INTRODUCTION

The densities of populations in different locations are frequently observed to fluctuate synchronously. Such behavior, known as spatial synchrony, has been described, for example, in small mammals (Bonrup-Nielson and Ims 1988, Krohne et al. 1988, Hansson 1990, Machin-Rogalska and Nabaglo 1990, Heikkilä et al. 1994), butterflies (Pollard 1991, Sutcliffe et al. 1996, 1997), grouse (Ranta et al. 1995b), various species of moths and aphids (Hanski and Woiwod 1993), sheep (Grenfell et al. 1998), forest defoliators (Mason 1978, Williams and Liebhold 1995, Myers 1998), and acorn crops (Crawley and Long 1995).

Two basic causes of synchrony have been proposed: dispersal and the effects of exogenous factors, frequently weather, that influence each of the populations in the same way. Populations that are linked by dispersal may fluctuate synchronously because increases in density at one location produce emigrants that move into nearby populations (Barbour 1990, Holyoak and Lawler 1996). Exogenous factors such as weather typically vary in the same direction over a large region and thus can have the same influence on many populations (Haydon and Steen 1997). Moran (1953) demonstrated that populations with common linear autoregressive models have the same correlation as shared exogenous factors, such as the weather. Since most populations are influenced by weather to some degree,

the widespread occurrence of spatial synchrony is not surprising. Other exogenous factors include natural enemies that disperse between sites or are synchronized themselves for reasons independent of prey density. For example, Myers (1988) proposed that virus infections might spread between populations and serve to synchronize densities of forest tent caterpillars. Heikkilä et al. (1994) proposed that mustelid predators might synchronize densities of microtine rodents in Scandinavia.

Most observations of synchronous fluctuations in density have concerned different populations of the same species. A number of studies, however, have documented synchrony between species. The most famous example is the snowshoe hare–lynx oscillation in boreal forests of North America, which is also correlated with several other species of mammalian predators and their prey (Elton and Nicholson 1942, Royama 1992). The principal focus of this paper is on the spatial synchrony of densities of the same species at different locations, but the same statistical issues apply to the wider context.

Despite a surge of interest in spatial synchrony, limited attention has been given to the issue of how synchrony of populations should be measured and tested. Synchrony has been measured in a variety of ways, and various methods have been employed to determine whether the observed pattern of synchrony differs statistically from that expected under “no synchrony.” These topics are the central focus of our paper. We will first examine various ways to measure synchrony in observed series of data and then view measures in terms of models for random processes underlying the data.

Manuscript received 14 September 1998; revised 28 April 2000; accepted 9 May 2000; final version received 11 July 2000.

⁵ Present address: Statistical and Data Analysis Center and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115 USA.

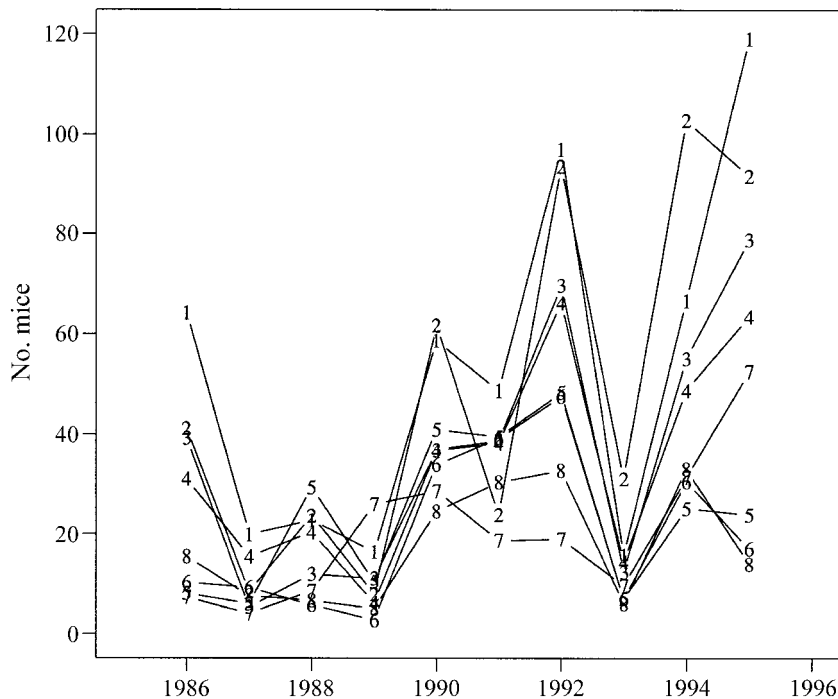


FIG. 1. Midsummer densities of white-footed mice for eight populations in western Massachusetts, USA (Elkinton et al. 1996).

These models provide the basis for our investigation into tests for no synchrony. After arguing that many of the commonly used procedures can have serious problems, we outline some alternative methods and outline the challenges in conducting statistical inference in this setting.

The spatial synchrony of population densities that we examine here is but one of various forms of synchrony that abound in the biological world. Many organisms exhibit synchronized behavior or physiology, such as synchronized life cycles, mating behaviors, estral cycles, etc. Many of the statistical issues raised here may also be pertinent in assessing these other types of synchrony.

MEASURES OF SYNCHRONY

Suppose the data consist of n different series measured at the same T points in time, where x_{it} is the population density (or some function of it such as $\log(\text{density})$) for series i at time t . For location i , the series is represented by $x_{i1}, \dots, x_{iT}$.

To illustrate, we use a 10-yr time series of densities of the white-footed mouse, *Peromyscus leucopus*, that we and our colleagues (Elkinton et al. 1996) collected at eight sites in western Massachusetts, USA. At each site we estimated densities with standard mark-recapture methods (Otis et al. 1978) in August of each year. For this problem, there are $n = 8$ locations and $T = 10$ time points. Figs. 1 and 2 display the data using density and $\log(\text{density})$, respectively.

Most studies that document synchrony begin with plots like Figs. 1 and 2, and synchrony is often said to exist when the population densities rise and fall together. Figs. 1 and 2 show that the visual impression of just "how synchronized" the series are can depend on whether densities or transformed densities are used. In many cases, from visual inspection, the trajectories are so similar that the investigators feel no need for further statistical demonstration of synchrony. Examples include Hornfeldt (1978) and Mason (1978). In other cases a variety of measures of synchrony are calculated.

Measures for two series

We first consider measures of synchrony between two series, say series i and j , followed by consideration of measures for more than two series.

1. Correlation in series values.—Many studies (e.g., Machin-Rogalska and Nabaglo 1990, Pollard 1991, Adler 1994, Garber and Burger 1995) have measured synchrony between two series using either Pearson or Spearman correlations. The Pearson correlation is a well-known measure of the strength of the linear relationship between two series, but one that changes under nonlinear transformations of the data such as taking logarithms. The Spearman correlation is invariant to monotonic transformations of the data and so remains the same for either density or $\log(\text{density})$.

Synchrony between populations can also be quantified using cross-correlograms, which are graphs of

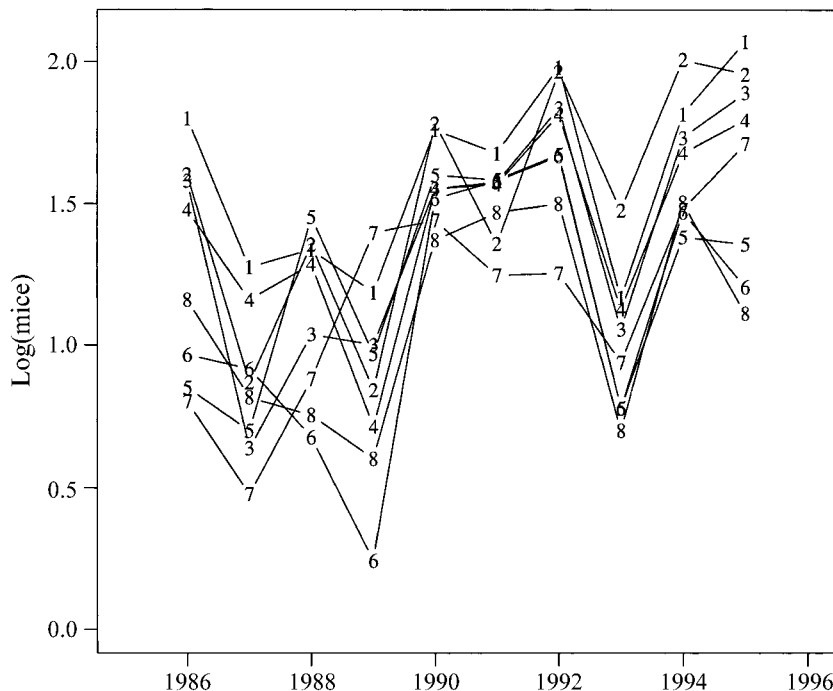


FIG. 2. Log(mice densities) of white-footed mice for eight populations (Elkington et al. 1996).

lag correlations between series vs. lag intervals (varying from negative to positive values). This is usually more appropriate when measuring synchrony among different species, as in Miller and Epstein (1986) and Ranta et al. (1995a), but can also be applied to study spatial synchrony among disjunct populations when population peaks are not coincident (Liebhold et al. 1996). This type of lag in synchrony may arise via dispersal effects or as a result of geographically varying responses to synchronous exogenous influences.

2. *Correlation in residuals.*—Another approach is to fit a model for each series and then calculate the correlation between the residuals. A common strategy has been to first fit an autoregressive model (usually an AR(1) or AR(2)) in which the current value depends on a specified number of previous values in a linear fashion. For example, Hanski and Woiwod (1993) and Sutcliffe et al. (1996) fit AR(1) models while Williams and Liebhold (1995) used AR(2) fits, all in terms of log(density). The idea behind this approach is to try and measure the correlation in the noise, possibly arising from an exogenous factor such as weather, after accounting for local dynamics. This is discussed further from a modeling perspective in *Population measures and models*.

Some (e.g., Ranta et al. 1995a, Koenig 1998) use correlation among residuals after “detrending” the series to remove a linear trend. Others (e.g., Ranta et al. 1995b) used residuals after removing periodic behavior modeled as a function of time with trigonometric functions. While detrending makes sense when all of the

series are either increasing or decreasing, removing periodic behavior may be removing the synchrony of interest.

3. *Measures based on change.*—To many investigators, synchrony means simply the tendency of the series to move in the same direction, especially for large fluctuations. Letting $d_{it} = x_{it} - x_{it-1}$, another measure of synchrony is the correlation in these changes (e.g., Steen et al. 1996, Bjørnstad et al. 1999). The differences are also sometimes used to remove nonstationarity in the series.

We propose a simpler related measure of how the two series change together, namely

$$A_{ij} = (\text{number of times series } i \text{ and } j \text{ move in same direction}) / (T - 1). \quad (1)$$

Notice that A is the same whether we use density or any monotonic transformation of it. The A_{ij} , which ranges from 0 to 1, can be converted to a measure similar to a correlation, ranging from -1 to 1 , by defining $\tau_{ij} = 2A_{ij} - 1$. We can also view τ as a modified version of Kendall's tau (Kendall 1970). Kendall's tau uses all possible pairs of points, but given the temporal nature of the data, we use only adjacent time points.

4. *Coincidence of peaks.*—Other investigators have measured the coincidence of peaks in the density data. For example, Crawley and Long (1995) studied the synchrony in acorn production among individual oak trees in England. Myers (1998) compared outbreak vs. nonoutbreak years of various species of forest lepi-

TABLE 1. Measures of synchrony for selected pairs of locations and mean value over all 28 pairs for the mouse data (Elkinton et al. 1996).

Series	r	$r(\log)$	SP	$\tau = 2A - 1$	N	C	$r(\text{res})$	SP(res)
1, 2	0.850	0.836	0.818	0.778	3	0.750	0.833	0.817
1, 5	0.497	0.613	0.503	0.778	3	0.750	0.730	0.583
1, 6	0.574	0.773	0.685	0.333	1	0.333	0.795	0.767
1, 7	0.682	0.559	0.527	0.778	2	0.667	0.387	0.317
2, 5	0.499	0.588	0.455	1.000	4	1.000	0.670	0.467
2, 6	0.607	0.719	0.697	0.556	2	0.500	0.715	0.667
2, 7	0.640	0.558	0.576	0.556	2	0.500	0.419	0.367
5, 6	0.861	0.687	0.685	0.556	2	0.500	0.718	0.750
5, 7	0.304	0.610	0.455	0.556	2	0.500	0.449	0.200
6, 7	0.271	0.348	0.345	0.111	1	0.500	0.062	0.133
Mean, all pairs	0.680	0.718	0.672	0.595		0.577	0.692	0.6310

Notes: Abbreviations are: r = Pearson correlation for density; $r(\log)$ = Pearson correlation for log(density); SP = Spearman correlation; A = proportion of time the series change in the same direction; N = number of coincident peaks; C = proportion of coincident peaks expressed as a fraction of the maximum number of peaks in the two series; $r(\text{res})$ = Pearson correlation on residuals; SP(res) = Spearman correlation on residuals. The residuals resulted from fitting a lag-1 model on the log scale.

doptera at different locations around the world. Both papers used a chi-square test for no synchrony.

Measures of synchrony can also be constructed from the coincidences of peaks as has been done in the literature examining synchrony in physiological pulse series (see Yang and Rao 1993 and references therein). One approach is to use N = number of times that series i and j both have peaks. This measure fails to account for how regularly either of the series peaks, so a more useful measure is $C = N/M$, where if M_i and M_j are the number of peaks in series i and j , respectively, then $M = \text{maximum}[M_i, M_j]$. Note that if the series always move up and down together, then $C = 1$. Other measures might use troughs or both peaks and troughs.

5. *Coefficient of variation.*—Several investigators have used the coefficient of variation in density as indicators of synchrony. Ims and Steen (1990), for example, form the coefficient of variation (sample standard deviation among locations/mean over locations) at each time and use the average coefficient of variation over years as a measure of synchrony. We believe however, that the average coefficient of variation is a poor choice as a measure of synchrony as it takes no account of possibly different levels of the populations. One can readily construct different collections of series, each collection in perfect synchrony in that they all move up and down in exactly the same way, and yet the average coefficient of variation can be made arbitrarily small or large by manipulating the overall means of the different series. While standardizing values in each series may help, it is difficult to interpret the resulting measure.

Measures for multiple series

With three or more series, a single measure of synchrony might be desired, although it should typically be used in conjunction with some measure of pairwise synchrony which conveys more information. One possibility is to average any of the pairwise measures, such as the correlation, over the $n(n - 1)/2$ different pairs

of series. The average measure of going up and down together,

$$\bar{A} = \frac{\sum_i \sum_{j < i} A_{ij}}{(n - 1)n/2} \quad (2)$$

has the particularly nice property of being the same as the overall proportion of time (over all pairs) that pairs of series agree in their change in direction. The average Spearman correlation, say S , is related to a measure of concordance among series based on ranks. Rank the data in each series and let R_{it} denote the rank of the density at time t in series i , $\bar{R}_{.t}$ denote the average rank for the n series in year t , and $\bar{R}_{..} = \sum_i \sum_t R_{it}/nT$ denote the overall mean. The “coefficient of concordance” (Kendall 1970:95, Herrera 1998), which measures synchrony in terms of the ranks, is defined by

$$W = \frac{12 \sum_{t=1}^T (\bar{R}_{.t} - \bar{R}_{..})^2}{T^3 - T}. \quad (3)$$

The maximum value of W is 1, which is achieved when the rankings in each series agree perfectly over the years. The relationship to S is that $S = (nW - 1)/(n - 1)$.

Numerical illustrations

We first illustrate using the mouse data, where the various measures of synchrony are computed for each of the 28 pairs of series. Table 1 provides values on selected pairs and Fig. 3 displays the relationships between some of the different measures over all pairs. The measures include r = correlation using density, $r(\log)$ = correlation using log(density), and $r(\text{res})$ = correlations using residuals after fitting an AR(1) model on the log scale. The different measures can provide conflicting information. For example, with series 5 and 7, $r = 0.304$, but $r(\log) = 0.610$, while for series 2 and 5 the correlation measures are low to moderate (ranging from 0.455 to 0.670) yet A and C are both 1. Fig. 3

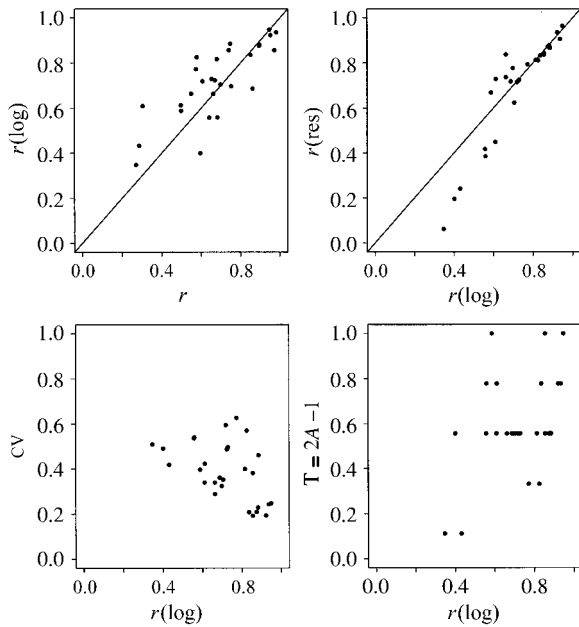


FIG. 3. Relationship among various measures of synchrony for each of the 28 pairs of stands in the mouse data (Elkington et al. 1996); r = Pearson correlation on density; $r(\log)$ = Pearson correlation on log(density); A = proportion of time the series change in the same direction; $r(\text{res})$ = Pearson correlation on residuals after lag-1 fit on log scale; CV = average coefficient of variation over time. The diagonal lines in the first two plots indicate where the two measures are equal.

also demonstrates the weak relationship of $r(\log)$ with A and the average coefficient of variation.

Further comparisons can be demonstrated with artificial data, as displayed in Fig. 4. In Fig. 4a both correlations are 1, and the two series always move in the same direction but the visual impression of synchrony is weak due to the low temporal variability in one series relative to the other. (Although one of the series looks constant over time, it does vary somewhat.) In Fig. 4b there is low to moderate correlation but perfect agreement in moving up and down together and in coincidence of peaks. The correlation between changes in the two series were also moderate, being 0.5700 and 0.5345 for Pearson and Spearman, respectively. In Fig. 4c there is very low correlation between the series and $A = 0.55$ but very high correlation in the residuals after fitting an autoregressive model of order 2. The data in the last case were generated by simulating from an AR(2) model with high correlation in the errors but different dynamic coefficients in the two series and show that high correlation in the errors, arising from an exogenous variable, may not manifest itself in terms of high correlation between the series values.

Population measures and models

Population measures, analogous to the data-based ones, can be defined in terms of stochastic models gov-

erning the series. Models are essential for describing the mechanisms associated with synchrony and also provide the framework for a discussion of statistical inference. Let X_{it} denote the random quantity for the i th location at time point t and $\mathbf{X}_t = (X_{1t}, \dots, X_{nt})$ the collection of values, over locations, at time t . We will be concerned primarily with *jointly* stationary models for which our measures of synchrony will not depend on time. Joint second order stationarity assumes that the mean, variance, and correlation structure of $(\mathbf{X}_t, \mathbf{X}_{t+1}, \dots, \mathbf{X}_{t+k})$ is the same for any t and arbitrary k ; see Brockwell and Davis (1996:15) for a formal definition. The stronger notion of strict joint stationarity assumes that the joint distribution of $(\mathbf{X}_t, \mathbf{X}_{t+1}, \dots, \mathbf{X}_{t+k})$ is the same for any t and arbitrary k . Stationarity allows many models, including multivariate ARMA models, but the model cannot involve any nonconstant deterministic functions of time.

Under second order stationarity we can define the population correlation

$$\rho_{ij} = \text{corr}(X_{it}, X_{jt}) \quad (4)$$

between series i and j , which will not depend on t . This is the population version of r . Under strict stationarity we define

$$\pi_{ij} = \text{probability that series } i \text{ and } j \text{ change in same direction} \quad (5)$$

which is the parametric version of the measure of agreement A .

For nonstationary series, one can define similar quantities by taking expected values over time, but we will not pursue this here.

There are three immediate ways to define no synchrony among series; (1) $\rho = 0$, for each pair of series, (2) the series are stochastically independent, and (3) $\pi = 0.5$ for each pair. These are related but not equivalent. For example, independence implies zero correlation, and independence along with the distribution of either $X_{i,t+1} - X_{it}$ or $X_{j,t+1} - X_{jt}$ being symmetric around 0 (as happens for Gaussian models) implies $\pi_{ij} = 0.5$.

Some investigators might prefer to focus on changes $D_t = X_t - X_{t-1}$, where typically $X = \log(\text{density})$. While it is always possible to move back and forth between a model for X and one for D , there are no general results relating the correlations among series for the two models. The serial correlation in D_t over time is also important, because of its potential impact on statistical inference. If X follows a stationary process with $\omega_1 = \text{lag 1 correlation}$ and $\omega_2 = \text{lag 2 correlation}$, then $\text{corr}(X_t - X_{t-1}, X_{t-1} - X_{t-2}) = [(\omega_2 - 1)/2(1 - \omega_1)] - 0.5$. Even if the original series has no serial correlation, there is a correlation of -0.5 between successive differences.

Further discussion requires specification of a multivariate time series model. We will focus on autoregressive models $X_{it} = f(\mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \dots, \mathbf{X}_{t-q}, \boldsymbol{\beta}_i) + \varepsilon_{it}$,

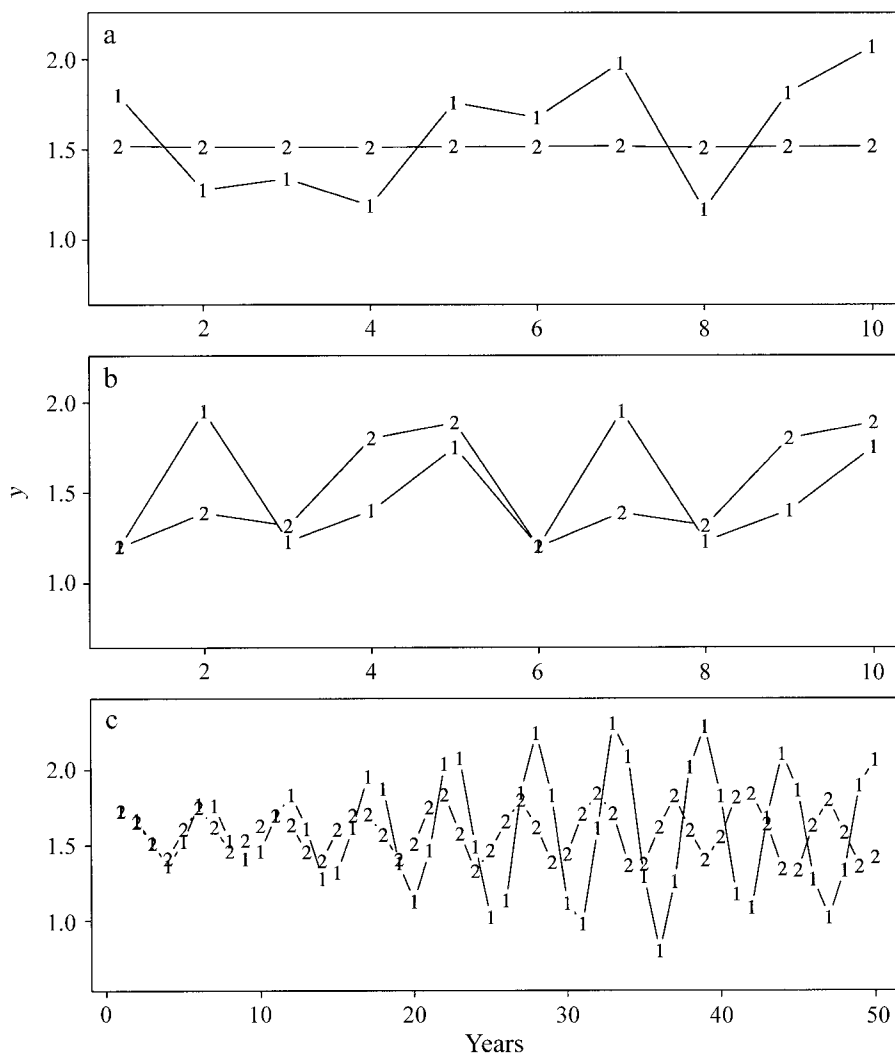


FIG. 4. Plots of three artificially generated data sets wherein (a) $r = 1$, $P = 1$, and $A = 1$; (b) $r = 0.3838$, $\rho = 0.700$, $P = 1$, $A = 1$; and (c) $r = -0.1447$, $A = 0.55$, $r(\text{res}) = 0.9612$ after a lag-2 fit. For an explanation of the abbreviations, see *Measures of synchrony*.

where f is a known function, β_i is a vector of parameters, and ε_{it} has conditional mean 0, where the conditioning is on past values of the series. The correlation among errors in series i and j ,

$$\tilde{\rho}_{ij} = \text{corr}(\varepsilon_{it}, \varepsilon_{jt}) \quad (6)$$

can arise in part from exogenous variables influencing all of the series. This is the population version of $r(\text{res})$, the correlation in residuals.

The two correlations $\tilde{\rho}$ and ρ are related, but often not the same. For models where each series depends only on its past values, $\tilde{\rho} = 0$ and $\rho = 0$ are usually equivalent. It is possible however, that ρ can be nonzero even if $\tilde{\rho}$ is zero. This can happen with models that allow past values of one series to influence future values of other series, as with dispersal, or when there are lagged exogenous factors.

We illustrate some of the above points using the linear multivariate AR(q) model:

$$\mathbf{X}_t = \boldsymbol{\beta}_0 + \mathbf{B}_1 \mathbf{X}_{t-1} + \mathbf{B}_2 \mathbf{X}_{t-2} + \dots + \mathbf{B}_q \mathbf{X}_{t-q} + \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\varepsilon}_t$ is the $n \times 1$ vector of errors at time t , $\boldsymbol{\beta}_0$ is an $n \times 1$ vector of intercepts, and each $\mathbf{B}_k$ is an $n \times n$ matrix of coefficients. If each of the $\mathbf{B}_k$ is diagonal then this corresponds to separate AR(q) models for each series, and zero correlation between the series is equivalent to zero correlation in the errors. However, nondiagonal $\mathbf{B}_k$ allows past values of one series to influence future values of other series, and the ρ 's can be nonzero even if the $\tilde{\rho}$'s are zero. One also has to be cautious about interpreting $\tilde{\rho}$ as capturing all of the correlation due to exogenous variables such as weather. Suppose for example $\mathbf{X}_t = \boldsymbol{\beta}_0 + \mathbf{B}_1 \mathbf{X}_{t-1} + \boldsymbol{\delta}_t$, where the $\boldsymbol{\delta}$ itself follows an autoregressive model $\boldsymbol{\delta}_t = \Phi \boldsymbol{\delta}_{t-1}$

+ ε_t , where ε_t has conditional mean 0. (Note that δ_t does not have conditional mean 0.) This describes a model where the exogenous variables have a lag structure and can operate on a global level. This leads to an AR(2) model $X_t = \beta_0^* + \mathbf{B}_1^* X_{t-1} + \mathbf{B}_2^* X_{t-2} + \varepsilon_t$, where $\mathbf{B}_1^* = \mathbf{B}_1 + \Phi$ and $\mathbf{B}_2^* = -\Phi \mathbf{B}_2$. If Φ is not diagonal, then the correlation in the ε 's captures only part of the correlation due to the exogenous factors. It is possible for the ε 's to be uncorrelated yet there are nonzero correlations among the series.

For separate AR(1) models for each series

$$X_{it} = \beta_{i0} + \beta_{i1} X_{i,t-1} + \varepsilon_{it} \quad (7)$$

where the ε_{it} are random "errors," independent over t within a series, with a conditional mean 0 and variance σ_i^2 , and β_{i1} measures the serial correlation in series i . In this case

$$\rho_{ij} = \tilde{\rho}_{ij} \left[\frac{(1 - \beta_{i1}^2)^{1/2} (1 - \beta_{j1}^2)^{1/2}}{1 - \beta_{i1} \beta_{j1}} \right] \quad (8)$$

for which $\rho_{ij} \leq \tilde{\rho}_{ij}$ and equality occurs only when $\beta_{i1} = \beta_{j1}$ or when $\tilde{\rho}_{ij} = \rho_{ij} = 0$. Similar results hold for separate AR(2) models (a popular choice in modeling population dynamics) with the equality of the two correlations under equal coefficients being related to the so-called "Moran effect" (Moran 1953, Royama 1992). We emphasize that when the coefficients in the dynamic models differ between two series, high correlation in the errors may not translate into high correlation in the series as is illustrated in Fig. 4c.

There are other models that could be considered, but a full discussion of these is outside the objectives of this paper. See for example Ranta et al. (1997), Grenfell et al. (1998), and Hudson and Cattadori (1999), for discussion of other models.

TESTING FOR NO SYNCHRONY

In this section we discuss testing for "no synchrony," which most investigators equate with independence or zero correlation among the series. Construction of confidence intervals will be discussed later. Until the summary section, most of our discussion is limited to the context with separate autoregressive models for each of the series.

Naive tests

The most common naive test for comparing two series is the standard test for zero correlation (e.g., Sokal and Rohlf 1981:581, SAS Institute 1990:224), which compares $r(T - 2)^{1/2}/(1 - r^2)^{1/2}$ to a t distribution with $T - 2$ degrees of freedom. Despite its continued widespread use in the context of assessing synchrony (e.g., Mackin-Rogalska and Nabaglo 1990, Pollard 1991, Adler 1994, Garber and Burger 1995, Hawkins and Holyoak 1998), this test is well known to be invalid when there is serial correlation in both series (Brockwell and Davis 1996:230). It is valid if one of the series is white noise.

Root and Cappuccino (1992) tested the hypothesis that $\pi = 0.5$ by comparing $Z = (A - 0.5)/(0.25/(T - 1))^{1/2}$ to a standard normal distribution or equivalently $C = Z^2$ to a chi-square distribution with one degree of freedom. Even with no serial correlation over time within the series, differences from adjacent time points are correlated over time, and consequently this test is invalid. The approaches suggested by Goodman and Grunfeld (1961) to address this do not handle the dynamic models of interest here. In a similar vein, the methods applied by Krohne et al. (1988) and Crawley and Long (1995) do not account for serial correlation and cannot be recommended.

Residual and adjusted degrees of freedom tests

There are a few strategies to deal with the problems caused by serial correlation. One option is to use a standard test for zero correlation on the residuals after fitting time series models to each series. This is often called "pre-whitening" (see Haugh [1976], Brockwell and Davis [1996:230], and Pyper and Peterman [1998] for some general discussion and Williams and Liebhold [1995] for an application after fitting AR(2) models). For an AR(q) model, this test compares $\tilde{r}(T - q - 2)^{1/2}/(1 - \tilde{r}^2)^{1/2}$ to a t -distribution with $T - q - 2$ degrees of freedom, where $\tilde{r}$ is the correlation in the residuals. For T large this is approximately the same as comparing $Z = \tilde{r}T^{1/2}$ to a standard normal distribution (see Brockwell and Davis 1996:230).

Another approach is to create a standardized test statistic $Z = M/\text{SE}(M)$, where M is a measure of synchrony and $\text{SE}(M)$ is an estimated standard error of M which takes account of the serial correlation. A large sample test based on Z can be carried out using the standard normal distribution. This is related to "adjusted degrees of freedom" tests based on the correlation. If the estimated standard error of r is of the form $(1/T)^{1/2}$, then the suggestion is to compare $r(T' - 2)^{1/2}/(1 - r^2)^{1/2}$ to a t distribution with $T' - 2$ degrees of freedom. For example, in Fromentin et al. (1998) $T' = T(1 - b_1 b_2)/(1 + b_1 b_2)$ where b_1 and b_2 are the estimated autoregressive coefficients from fitting AR(1) models to two series. While they did not provide any details, this results from the fact that for two independent AR(1) processes, the variance of r is $\sim 1/J$, where $J = T(1 - \beta_1 \beta_2)/(1 + \beta_1 \beta_2)$; this follows from Bartlett's formula (Brockwell and Davis 1996:232). Hence J is like a relative sample size with T' being an estimate of it. In a closely related spatial problem, Clifford et al. (1989) used an adjusted degrees of freedom test, based on a nonparametric estimate of the standard error of a correlation.

The residual test is based on large sample arguments, so it is only approximately correct, but it often performs well with moderate sample sizes. This approach also has the advantage of being easily implemented after fitting any type of time series models. The use of standardized or adjusted degrees of freedom tests depends

on estimating the standard error of the measure involved, which in turn depends on the dynamic models used. Determining the standard error can be difficult for measures other than for correlation, and even for correlation, it is only easy for the low order autoregressive models. In addition, the use of $T' - 2$ degrees of freedom in the adjusted degrees of freedom test is based on a heuristic argument. For these reasons, we feel the tests based on residuals is preferred.

Bootstrapping

Some authors have tried simulation or resampling methods to assess synchrony. For example, Miller and Epstein (1986) examined synchrony in multiple series by first getting standard correlations and P values for each pair of series. They then assessed significance by comparing summary measures over pairs of series, based on either the correlations or P values, to simulated behavior under a null model with independence among series and use of what they call a “matched control” model within each series. One difficulty is that the null distribution that is generated depends on the use of the matched control model, and other null models (i.e., with independence among series but a different model than the matched control one for the series) can lead to different results. Also, for some of the measures this would only make sense with enough series so that a reasonable number of pairs enter into the analysis. Ranta et al. (1995a, b) used resampling when assessing the significance of correlations among 11 Finnish provinces with 20 yr of data. However, it is unclear how their resampling incorporates the null hypothesis of zero correlation. The bootstrap samples to which the observed correlation is compared should be generated under the null hypothesis. Also, the 11 provinces in the study are fixed, yet they bootstrap by resampling pairs of locations. Since bootstrap resampling must be carried out in a manner that reflects the original randomness in the data (see Efron and Tibshirani 1993:91), the locations should not be resampled. Yang and Rao (1993) investigated a model-free method for testing independence among series based on the coincidence of peaks, by resampling from the waiting times between peaks. This approach is attractive, since it does away with the need for a specific dynamic model, although it must be assumed that the waiting times are independently and identically distributed. While this method is potentially useful, it requires longer series than are usually available in ecological applications, in order to have enough peaks to get a distribution of waiting times.

We describe a model-based bootstrap approach to testing for no synchrony using the AR(1) model in Eq. 7 for illustration. For each series estimates of the coefficients, say b_{i0} and b_{i1} , are obtained from which residuals $e_{it} = x_{it} - (b_{i0} + b_{i1}x_{i,t-1})$, $t = 2, \dots, T$ are formed. The k th bootstrap sample, $k = 1$ to K (large) consists of $\{x_{it(k)}, i = 1, \dots, n \text{ and } t = 1, \dots, T\}$, where $x_{i1(k)}$

$= x_{i1}$ (the same initial value in the respective series is used) and

$$x_{it(k)} = b_{i0} + b_{i1}x_{i,t-1(k)} + \varepsilon_{it(k)}, t = 2, \dots, T \quad (9)$$

where the $\varepsilon_{it(k)}$ are simulated errors. Since we want to sample under the null hypothesis of independence among the series, the $\varepsilon_{it(k)}$ are generated separately for each series. Within series i , $\varepsilon_{it(k)}$ arises by sampling with equal probability from a set of modified residuals $e_{i2}^*, \dots, e_{iT}^*$. With least squares estimates (for which the residuals have mean 0), $e_{it}^* = ((T-1)/(T-3))^{1/2}e_{it}$, where the rescaling is such that the variance of a randomly generated error agrees with $\hat{\sigma}_i^2 = \sum_{t=2}^T e_{it}^2/(T-3)$. This modification is often omitted for moderate to large T . If Q is the observed value of the test statistic and the test is set up in a form that rejects H_0 if Q is large, then the bootstrap P value is

$$P_{\text{boot}} = (\text{number of times } Q_k \geq Q)/K$$

where Q_k is the test statistic computed from the k th bootstrap sample.

This procedure can be readily extended to handle dynamic models other than the AR(1) and methods of estimation besides least squares. (See Davison and Hinkley [1997: section 8.2] for a general discussion and Dennis and Taper [1994] for some ecological applications). One thing to note is that the residuals should be centered if they do not have mean 0 (Efron and Tibshirani 1993:95, Davison and Hinkley 1997: 390). As emphasized in *Simulations*, while the bootstrap test in this form is attractive for its simplicity and this is the manner in which it is often used, it is not a panacea and can encounter some problems.

Example

Table 2 displays two-sided P values for various tests of no synchrony for the mouse data (Elkinton et al. 1996). The bootstrap samples are generated under independence among series and using AR(1) models, which were found to provide an adequate fit to the data. The $B_{r(\log)}$ P value is obtained by finding the bootstrap P value using the simple correlation r between the $\log(\text{densities})$, while the $B_{r(\text{res})}$ finds the bootstrap P value using the correlation between the residuals after fitting an AR(1) model to the $\log(\text{densities})$. For these data the bootstrap P values for $r(\log)$ are in general agreement with the naive P values based on $r(\log)$ which ignore the serial correlation in the series. This is due to the fact that the serial correlations are modest, all below 0.2 except for 0.6 in series 7. As noted earlier, the naive test that ignores serial correlation is essentially correct as long as there is no serial correlation in at least one of the series. Notice however that there can be disagreements between the P value based on $r(\log)$ and that based on $r(\text{res})$. Theoretical results, along with simulations, suggest that we should use the tests based on $r(\text{res})$. For these data, the bootstrap P

TABLE 2. Two-sided P values for testing “no synchrony” on selected pairs of stands from the mouse data (Elkington et al. 1996).

Series	$r(\log)$	$r(\text{res})$	A	$B_{r(\log)}$	$B_{r(\text{res})}$	B_A
1, 2	0.003	0.005	0.020	0.004	0.006	0.047
1, 5	0.059	0.026	0.020	0.051	0.027	0.047
1, 6	0.009	0.010	0.317	0.013	0.013	0.370
1, 7	0.093	0.304	0.020	0.087	0.311	0.026
2, 5	0.074	0.048	0.003	0.075	0.063	0.000
2, 6	0.019	0.030	0.096	0.039	0.043	0.043
2, 7	0.094	0.261	0.096	0.081	0.269	0.041
5, 6	0.028	0.029	0.096	0.031	0.036	0.056
5, 7	0.061	0.225	0.096	0.077	0.229	0.056
6, 7	0.324	0.875	0.739	0.327	0.867	0.510

Notes: Here, $r(\log)$ and $r(\text{res})$ use a standard test for zero correlation based on log(density) and residuals after an AR(1) fit on log(density), respectively. A uses a naive chi-square test for probability 0.5 of moving in the same direction. The last three columns give the bootstrap P values associated with $r(\log)$, $r(\text{res})$, and A , respectively.

values essentially confirm the nonbootstrap P values for either $r(\log)$ or $r(\text{res})$.

Simulations

The performance of some of the tests was simulated for two series using the AR(1) model in Eq. 7, with normal errors having correlation $\tilde{\rho} = 0$. This also implies $\rho = 0$; see Eq. 8. Throughout, based in part on the mouse data, we use $\beta_{10} = \beta_{20} = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.14$. The initial values (X_{11} , X_{21}) are generated from a bivariate normal distribution corresponding to the appropriate stationary distribution. Table 3 provides simulated rejection rates for different values of $\beta = \beta_{11} = \beta_{21}$ and T and for certain of the naive, adjusted, and bootstrap tests. Since $\rho = 0$, we are simulating the probability of a Type I error, that is, of rejecting the null hypothesis of no synchrony when there is no synchrony. When $\beta = 0$, there is no serial correlation within the series.

The commonly employed naive test based on r performs as expected, being adequate when each series is white noise ($\beta = 0$) but having rejection rates that are too high with serial correlation ($\beta \neq 0$). The poor performance of this test is exacerbated for longer series. The adjusted degrees of freedom test suffers from the opposite problem of having rejection rates that can be too low and hence can lead to a loss of power. The naive test using A has rejection rates that are almost twice the desired size throughout. The t test based on the residuals does quite well in all cases, even for short series.

Turning to the bootstrap tests, the test based on r , while improving on the naive test, still has difficulty with high serial correlation ($\beta = 0.8$). The test based on A overcompensates in three of the four cases, and now has rejection rates that are too small, while the test based on W has generally inflated error rates. The bootstrap test based on residuals is the only one that performs fairly well throughout (the one low estimate

of 0.04 has a standard error of 0.006 associated with it). For these settings not much is gained by using the bootstrap test based on residuals rather than the t test based on residuals. But nothing is lost either and the bootstrap would provide better protection under non-normal errors.

There are two important points reinforced by the simulation results. The first is the obvious deficiency of naive methods that fail to account for serial correlation in the series, when it is present. The second, concerning the bootstrap methods, is more subtle. Notice that the bootstrap tests based on r , A , and W improve on the naive methods since they account for the model, but they still encounter some problems. The difficulty is that r , A , and W are not “pivotal quantities,” that is, their null distribution depends on the parameters in the models. Different dynamic models lead to different null distributions. By removing the dynamic part of the model, the bootstrap test based on residuals does not suffer from this problem to the same extent. One approach to correcting the problems with tests based on measures such as r , A , or W is to bootstrap a standardized test statistic, but this requires an estimated standard error for the statistic and offers no advantage over the residual test in a model-based context. In the model-based setting, we repeat our earlier recommendation that tests for no synchrony be based on the residuals.

Confidence intervals

Rather than testing for no synchrony, interest may lie in estimating the amount of synchrony (as defined by a suitable population measure, such as ρ , $\tilde{\rho}$, or π) via estimated standard errors and/or confidence intervals. Confidence intervals also offer a direct way to test for no synchrony in a way that will avoid some of the problems with tests described in *Simulations*.

In the model-based context, approximate procedures for ρ are available for certain models based on the large

TABLE 3. Simulated rejection rates of tests for zero correlation.

Statistic	$\beta = 0$		$\beta = 0.4$		$\beta = 0.8$	
	$T = 10$	$T = 30$	$T = 10$	$T = 30$	$T = 10$	$T = 30$
A	0.1114	0.092	0.098	0.084	0.104	0.084
r	0.051	0.049	0.072	0.086	0.152	0.186
$r(\text{adj})$	0.024	0.05	0.042	0.053	0.016	0.03
$r(\text{res})$	0.055	0.048	0.058	0.050	0.054	0.052
B_W	...	...	0.048	0.064	0.062	0.086
B_A	...	...	0.032	0.028	0.030	0.066
B_r	...	...	0.056	0.056	0.080	0.090
$B_{r(\text{res})}$	...	...	0.054	0.040	0.048	0.052

Notes: β = amount of serial correlation; T = number of time points; A = naive chi-square test; r = naive correlation test; $r(\text{adj})$ = adjusted degrees of freedom test; $r(\text{res})$ = test based on correlation in residuals. B indicates bootstrap test based on the designated statistic. One thousand simulations per setting were performed with a desired rejection rate of 0.05.

sample distribution of r (e.g., Brockwell and Davis 1996:232). Rather than detail these procedures, we describe the calculation of bootstrap standard errors and confidence intervals. Once again we illustrate with the AR(1) models in Eq. 7, where the collection of errors, $\varepsilon_{1t}, \dots, \varepsilon_{nt}$, at time t are allowed to have a general correlation structure (see Eq. 6). The bootstrap model is as given in Eq. 9 with the exception that at time t we do not generate the errors independently for each series. Instead, a random time, say t^* , is sampled from 2 to T , and the generated errors are set equal to the modified residuals corresponding to time t^* . Notice that this resampling, from the empirical joint distribution of the residuals, retains the correlation structure among series. (One *cannot* simply resample vectors of the actual densities over time as this will lose the serial correlation within series.) For each of the K bootstrap samples, the statistic of interest is computed and bootstrap standard errors and confidence intervals are obtained (see Efron and Tibshirani [1993] and Davison and Hinkley [1997] for details). This could be done for r , $r(\text{res})$, and A leading to standard errors and intervals for ρ , $\hat{\rho}$, and π respectively. The confidence interval can also be used to test at level α by rejecting the hypothesis of zero correlation between two series if the $100(1 - \alpha)\%$ confidence interval for the correlation does not contain 0. Similarly a test of $H_0: \pi = 0.5$ would reject H_0 if 0.5 is not in the interval for π . The advantage of this approach is that we get both a test and a range of plausible values for the parameter of interest. This is a promising direction based on past successes of bootstrap confidence intervals, but we leave numerical evaluations of these techniques for further investigation.

Model-free considerations

The above methods are model-based and would be wrong if an incorrect model is fit, say an AR(1) instead of an AR(2) or a linear model when a nonlinear one is needed. This issue is certainly not unique to this problem, as all statistical inferences depend on some model assumptions. It is important to utilize standard model-building and assessment techniques for choosing a model.

Is there a model-free way to address the problem? One suggestion is to try and develop a model-free estimate of the standard error of the statistic involved, as done by Clifford et al. (1989) for r in a closely related spatial context, and then bootstrap a standardized version of the test statistic. This needs further development and assessment in the time-series context. Another technique is the use of block resampling (Li and Maddala 1996, Davison and Hinkley 1997:389) but the value of this for small to moderate length series is questionable, and the same issues as earlier remain regarding constructing tests of no synchrony.

Model-free resampling appears easier when there is a random sample of locations and we are estimating

overall synchrony (such as an average population correlation) over the larger area from which they are sampled. An example would be in assessing synchrony in acorn production over a large region based on a sample of trees. Each bootstrap sample uses a sample of n series *with replacement* from the observed series. The resulting bootstrap samples can be used to construct a confidence interval, as suggested by Bjørnstad et al. (1999) for obtaining a bootstrap confidence interval for the average correlation between differences in log(abundance). Suppose we condition on the values of the series at all of the locations in the population and define M to be the average sample value of the measure of interest over all pairs of locations in the area of interest (e.g., the average ρ over all pairs in the area). Then the bootstrap interval can be interpreted conditionally as a confidence interval for M , but it remains to be seen how this interval performs for the average population measure (e.g., the average ρ over all pairs of locations in the area of interest). This issue needs further study. Resampling in this way should not be used when the locations in the sample are the only ones of interest, so the locations are fixed. Finally, even with random locations, these bootstrap samples should also not be used directly to construct the bootstrap P value, P_{boot} , for testing no synchrony since the resampling is not done under the null hypothesis.

SUMMARY

Our first objective was a discussion of various measures of synchrony that fall into four major categories: correlation among series values, association in changes (e.g., correlation in change, agreement in change in direction), correlation in residuals after fitting dynamic models, and coincidence of peaks or other extreme behavior of the series. As seen with our examples, these measures may or may not coincide in their assessment of synchrony. The measure of agreement and the correlation of changes in the series most closely capture the dictionary definition of synchrony, as they directly relate to how the two series move together. The measure of agreement A has the nice features of being easy to interpret and invariant to monotonic transformations, and the combined measure over many series is the simple average of the pairwise measures. Measures that capture how often the peaks (and possibly troughs) of the series coincide are also appealing for their ease of interpretation. These would be of most use to investigators interested primarily in whether outbreaks (or crashes) occur in common over the region of interest, but are only useful for longer series. Although Pearson correlation between series values is the most commonly used measure of synchrony, it may not capture the concept of synchrony between populations in an ecologically meaningful way. It is difficult to understand just how synchronized two series with a given correlation are and, as our examples demonstrated, this correlation can be low while measures of movement together are

high. High correlation in the errors or residuals can be interpreted as indicating that there are some exogenous variables operating in common over the different locations. This does not always translate into strong associations between either series values or changes in the series. It does however offer the *potential* for synchronized behavior if the population dynamics are similar in the different locations (e.g., the Moran effect).

Our second objective was to heighten awareness of the challenges in attempting statistical inferences, especially the popular approach of testing the hypothesis of no synchrony. Since serial correlation is often present in populations over time, tests that ignore it will have problems and should generally be avoided unless preliminary investigations clearly establish a lack of serial correlation. When there are separate autoregressive models for each series, our recommendation is to use a test based on residuals, with the bootstrap version offering some protection for small sample sizes. This does not mean that the correlation in residuals is always the appropriate way to measure synchrony, but simply that tests based on residuals perform better than the existing alternatives. The residual tests of course require a model for each series, and it is worth emphasizing that to this point there is no satisfactory model-free approach for testing for independence among series.

We also recall that if there is feedback among the series or there are global exogenous variables with a lag structure, then zero correlation in the errors is not equivalent to zero correlation in the series. In this case, tests for no synchrony should not be carried out based on residuals. The bootstrap can be extended if a model is specified (e.g., a linear multivariate autoregressive model), but as in our simpler setting, there will be difficulty in generating bootstrap tests directly, and the recommendation is to approach the problem via confidence intervals for suitable quantities.

Additional complications arise when there is measurement error, which is typically present since population abundances must be estimated. One ramification is that the correlations among the estimated abundances are different than those among true abundances. The model will also change. For example, if the true abundances follow an AR(q) model and there is additive measurement error with constant variance, then the estimated abundances follow an ARMA(q, q) model (see Box and Jenkins 1976:121). In practice, there are typically standard errors to accompany the estimated abundances and future work is needed to investigate how to use this information to get at synchrony in true abundance values. This is an important area that has received no attention in the context of assessing synchrony.

With many series, there are also approaches to assessing synchrony that relate a measure of synchrony (usually correlation) between pairs of series to the distance between the locations. (See Koenig [1999],

Bjørnstad et al. [1999], and references therein for discussion of this approach.)

The problem of assessing synchrony, whether through estimation or testing, is essentially a problem in multivariate time series and should be examined in this context. We have outlined the fundamental issues and strategies in a relatively simple context, but much work remains before comprehensive recommendations can be made that will cover the many complex models encountered in practice.

ACKNOWLEDGMENTS

This research was funded in part by NSF grant 9427792 to J. Elkinton and J. Buonaccorsi. We are very grateful to Associate Editor Allan Stewart-Oaten and two referees, Bruce Kendall and Ottar Bjørnstad, for their extremely helpful comments.

LITERATURE CITED

- Adler, G. H. 1994. Tropical forest fragmentation and isolation promote asynchrony among populations of a frugivorous rodent. *Journal of Animal Ecology* **63**:903–911.
- Barbour, D. A. 1990. Synchronous fluctuations in spatially separated populations of cyclic forest insects. Pages 339–346 in A. D. Watt, S. R. Leather, M. D. Hunter, and N. A. Kidd, editors. *Population dynamics of forest insects*. Intercept, Andover, UK.
- Bjørnstad, O. N., R. A. Ims, and X. Lambdin. 1999. Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends in Ecology and Evolution* **14**:427–431.
- Bonrup-Nielsen, S., and R. A. Ims. 1988. Predicting stable and cyclic populations of *Clethrionomys*. *Oikos* **52**:178–185.
- Box, G. E. P., and G. M. Jenkins. 1976. *Time series analysis: forecasting and control*. Second edition. Holden Day, San Francisco, California, USA.
- Brockwell, P. J., and R. A. Davis. 1996. *Introduction to time series and forecasting*. Springer, New York, New York, USA.
- Clifford, P., S. Richardson, and D. Hémon. 1989. Assessing the significance of the correlation between two spatial processes. *Biometrics* **45**:123–134.
- Crawley, M. J., and C. R. Long. 1995. Alternate bearing, predator satiation and seedling recruitment in *Quercus robur* L. *Journal of Ecology* **83**:683–696.
- Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap methods and their application*. Cambridge University Press, Cambridge, UK.
- Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. *Ecology* **64**:205–224.
- Efron, B., and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall, New York, New York, USA.
- Elkinton, J. S., W. M. Healy, J. P. Buonaccorsi, G. H. Boettner, A. M. Hazzard, H. R. Smith, and A. M. Liebhold. 1996. Interactions among gypsy moths, white-footed mice, and acorns. *Ecology* **77**:2332–2342.
- Elton, C. S., and M. Nicholson. 1942. The ten year cycle in numbers of lynx in Canada. *Journal of Animal Ecology* **11**:215–244.
- Fromentin, J. M., N. C. Stenseth, J. Gjøsæter, T. Johannessen, and B. Planque. 1998. Long-term fluctuations in cod and pollack along the Norwegian Skagerrak coast. *Marine Ecology Progress Series* **162**:265–278.
- Garber, S. D., and J. Burger. 1995. A 20-yr study documenting the relationship between turtle decline and human recreation. *Ecological Applications* **5**:1151–1162.
- Grenfell, B. T., K. Wilson, B. F. Finkenstadt, T. N. Coulson,

- S. Murray, S. D. Albon, J. M. Pemberton, T. H. Clutton-Brock, and M. J. Crawley. 1998. Noise and determinism in synchronized sheep dynamics. *Nature* **394**:674–677.
- Goodman, L. A., and Y. Grunfeld. 1961. Some nonparametric tests for comovements between time series. *Journal of the American Statistical Association* **56**:11–26.
- Hanski, I., and I. P. Woiwod. 1993. Spatial synchrony in the dynamics of moth and aphid populations. *Journal of Animal Ecology* **62**:656–668.
- Hansson, L. 1990. Spatial dynamics in fluctuating vole populations. *Oecologia* **85**:213–217.
- Haydon, D., and H. Steen. 1997. The effects of large- and small-scale random events on the synchrony of metapopulation dynamics: a theoretical analysis. *Proceedings of the Royal Society of London, B* **264**:1375–1381.
- Haugh, L. D. 1976. Checking the independence of two time series. *Journal of the American Statistical Association* **71**: 378–385.
- Hawkins, B. A., and M. Holyoak. 1998. Transcontinental crashes in insect populations. *American Naturalist* **152**: 480–483.
- Heikkilä, J., A. Below, and I. Hanski. 1994. Synchronous dynamics of microtine rodent populations on islands in Lake Inari in northern Fennoscandia: evidence for regulation by mustelid predators. *Oikos* **70**:245–252.
- Herrera, C. M. 1998. Population-level estimates of interannual variability in seed production: what do they really tell us? *Oikos* **82**:612–616.
- Holyoak, M., and S. P. Lawler. 1996. The role of dispersal in predator–prey metapopulation dynamics. *Journal of Animal Ecology* **65**:640–652.
- Hörnfeldt, B. 1978. Synchronous population fluctuations in voles, small game, owls and tularemia in northern Sweden. *Oecologia* **32**:141–152.
- Hudson, P. J., and I. M. Cattadori. 1999. The Moran effect: a cause of population synchrony. *Trends in Ecology and Evolution* **14**:1–2.
- Ims, R. A., and H. Steen. 1990. Geographical synchrony in microtine population cycles: a theoretical evaluation of the role of nomadic avian predators. *Oikos* **57**:381–387.
- Kendall, M. G. 1970. Rank correlation methods. Griffin, London, UK.
- Koenig, W. D. 1998. Spatial autocorrelation in California land birds. *Conservation Biology* **12**:612–620.
- Koenig, W. D. 1999. Spatial autocorrelation of ecological phenomena. *Trends in Ecology and Evolution* **14**:22–26.
- Krohne, D. T., J. F. Merritt, S. H. Vessey, and J. O. Wolff. 1988. Comparative demography of forest *Peromyscus*. *Canadian Journal of Zoology* **66**:2170–2176.
- Li, H., and G. S. Maddala. 1996. Bootstrapping time series models (with discussion). *Econometric Reviews* **15**:115–195.
- Liebholt, A., N. Kamata, and T. Jacob. 1996. Cyclicity and synchrony of historical outbreaks of the beech caterpillar, *Quadricalcarifera punctatella* (Motschulsky) in Japan. *Researches on Population Ecology* **38**:87–94.
- Mackin-Rogalska, R., and L. Nabaglo. 1990. Geographic variation in cyclic periodicity and synchrony in the common vole, *Microtus arvalis*. *Oikos* **59**:343–348.
- Mason, R. R. 1978. Synchronous patterns in an outbreak of the Douglas-fir tussock moth. *Environmental Entomology* **7**:672–675.
- Miller, W. E., and M. E. Epstein. 1986. Synchronous population fluctuations among moth species (Lepidoptera). *Environmental Entomology* **15**:443–447.
- Moran, P. A. P. 1953. The statistical analysis of the Canadian lynx cycle. II. Synchronization and meteorology. *Australian Journal of Zoology* **1**:291–298.
- Myers, J. H. 1988. Can a general hypothesis explain population cycles of forest Lepidoptera? *Advances in Ecological Research* **18**:179–242.
- Myers, J. H. 1998. Synchrony of outbreaks of forest Lepidoptera: a possible example of the Moran effect. *Ecology* **79**:1111–1117.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monograph* **62**:1–135.
- Pollard, E. 1991. Synchrony of population fluctuations: the dominant influence of widespread factors on local butterfly populations. *Oikos* **60**:7–10.
- Pyper, B. J., and R. M. Peterman. 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. *Canadian Journal of Fisheries and Aquatic Sciences* **55**:2127–2140.
- Ranta, E., V. Kaitala, J. Lindström, and E. Helle. 1997. The Moran effect and synchrony in population dynamics. *Oikos* **78**:136–142.
- Ranta, E., V. Kaitala, J. Lindström, and H. Lindén. 1995b. Synchrony in population dynamics. *Proceedings of the Royal Society of London* **262**:113–118.
- Ranta, E., J. Lindström, and H. Lindén. 1995a. Synchrony in tetraonid population dynamics. *Journal of Animal Ecology* **64**:767–776.
- Root, R. B., and N. Cappuccino. 1992. Patterns in population change and the organization of the insect community associated with goldenrod. *Ecological Monographs* **62**:393–420.
- Royama, T. 1992. Analytical population dynamics. Chapman & Hall, New York, New York, USA.
- SAS Institute. 1990. SAS procedures guide. Version 6. Third edition. SAS Institute, Cary, North Carolina, USA.
- Sokal, R. R., and F. J. Rohlf. 1981. Biometry: the principles and practices of statistics in biological research. W. H. Freeman, New York, New York, USA.
- Steen, H., R. A. Ims, and G. A. Sonnerud. 1996. Spatial and temporal patterns of small-rodent population dynamics at a regional scale. *Ecology* **77**:2365–2372.
- Sutcliffe, O. L., C. D. Thomas, and D. Moss. 1996. Spatial synchrony and asynchrony in butterfly population dynamics. *Journal of Animal Ecology* **65**:85–95.
- Sutcliffe, O. L., C. D. Thomas, T. J. Yates, and J. N. Grea-torex-Davis. 1997. Correlated extinctions, colonizations and population fluctuations in a highly connected ringlet butterfly metapopulation. *Oecologia* **109**:235–241.
- Williams, D. W., and A. M. Liebhold. 1995. Influence of weather on the synchrony of gypsy moth (Lepidoptera: Lymantriidae) outbreaks in New England. *Environmental Entomology* **24**:987–995.
- Yang, M. C. K., and P. V. Rao. 1993. Testing concomitancy between two physiological pulse series. *Statistics in Medicine* **12**:2043–2055.