Testing for spatial autocorrelation in ecological studies

Walter D. Koenig and Johannes M. H. Knops

Koenig, W. D. and Knops, J. M. H. 1998. Testing for spatial autocorrelation in ecological studies. – Ecography 21: 423–429.

We describe a statistical method appropriate for the analysis of spatial autocorrelation in data varying in time as well as space. In particular, the technique was developed to address the issue of geographic synchrony in ecological variables that may change markedly from year to year such as population density of animals or seed production of trees. The method yields "modified correlograms" that test for significant autocorrelation between sites located within any given range of distances apart. This technique facilitates detecting and understanding spatial processes in a variety of ecological phenomena, including testing the plausibility of causational hypotheses using cross-correlational analyses. Several examples are discussed, including population densities of squirrels in Finland, winter densities of two hawk species in California, and acorn production and radial growth by individual blue oak *Quercus douglasii* trees in central coastal California.

W. D. Koenig (wicker@uclink.berkeley.edu), Hastings Reservation, Univ. of California, 38601 E. Carmel Valley Rd, Carmel Valley, CA 93924, USA. – J. M. H. Knops, Dept of Ecology, Evolution and Behavior, Univ. of Minnesota, 1987 Upper Buford Dr., St. Paul, MN 55108, USA.

Spatial autocorrelation examines the degree of synchrony between variables observed across geographic space and is important to a wide variety of ecological phenomena (Legendre 1993). Unfortunately, the techniques used to measure and, in particular, statistically test for spatial autocorrelation are poorly developed (Thompson et al. 1996). In particular, most methods for dealing with spatial variation are concerned with the situation in which data are available at a single point in time. What if the variable of interest varies temporally as well as spatially? Typical examples concern synchrony in population size or performance from year to year over a given geographic area. In these cases, both the absolute degree of autocorrelation as well as relative values are potentially of interest.

Few attempts have been made at dealing with such data, and those that exist have used varying methods to describe and statistically analyze the data. Here we describe a modification of standard statistical techniques appropriate for measuring and testing for spatial autocorrelation in cases where data vary temporally and illustrate its use.

Description of the technique

The typical data set with which we are concerned involves a series of measurements overlapping in time taken at multiple sites over some geographic area. Examples include the radial growth of oaks (genus *Quercus*) as measured by tree-ring chronologies at 39 sites in Britain and western Europe over a 120-yr period (Kelly et al. 1989), the number of moths captured each year at a network of light traps scattered throughout Britain (Hanski and Woiwod 1993), synchrony between microtine rodents and their specialist predators (Ims and Steen 1990, Hanski et al. 1991, Heikkila et al. 1994), and population sizes of red squirrels *Sciurus vulgaris* in Finland (Ranta et al. 1997).

The primary question of interest in these examples is not whether the absolute number of animals (or the radial growth of the trees) is similar across sites in a particular year, but whether the relative population sizes from year to year are similar at different sites. Specifically, whether populations are at identical densities in a particular year is irrelevant; what is of interest is whether or not population sizes tend to be relatively large or small during the same years in different areas.

As with many statistical tests, raw data used in analysis of spatial autocorrelation must often be modified in an appropriate fashion. For example, it may be desirable to transform the raw values to reduce the correlation between the mean and the variance or to calculate residuals from a linear regression in order to avoid spurious cross-correlations due to either local dynamics or large-scale global trends rather than the regional processes of direct interest (Hanski and Woiwod 1993).

The next step is to calculate the geographic distance between pairs of sites along with the correlation (r-value) between the ecological phenomenon of interest using all years for which data from both sites are available. This yields two matrices: one of the distance between sites and the other of the correlation between the shared ecological variable.

A standard statistical test for such data is the Mantel test (Fortin and Gurevitch 1993), which examines the overall relationship between distance and the correlation coefficient between pairs of sites. The Mantel test consists of comparing the correlation coefficients (normalized Mantel statistic) obtained from the observed data with correlation coefficients derived following randomization trials in which the values of one of the matrices are shuffled at random. A modification of this technique is the Mantel correlogram (Oden 1984, Oden and Sokal 1986, Legendre and Fortin 1989), in which the correlation coefficients are divided into distance categories and the values for each distance category are tested against the overall average degree of autocorrelation among sites present in the complete data set. Consequently, positive and negative values of the Mantel product z indicate autocorrelations greater than and less than the overall mean autocorrelation among sites, respectively, not an average degree of autocorrelation between sites greater than or less than zero. Thus, a complete Mantel correlogram usually yields a roughly equivalent number of positive and negative z-values even if autocorrelation among sites is always positive.

In most ecological phenomena measured over an appropriate geographic area, spatial autocorrelation declines with distance whether it varies temporally or not. Examples from the literature include numbers of British moths, butterflies, and aphids (Hanski and Woiwod 1993, Sutcliffe et al. 1996) and population dynamics of seven species of Finnish wildlife (Ranta et al. 1995). One obvious reason why this will usually be true is that

spatial autocorrelation in most environmental factors also declines with distance; that is, conditions are typically more similar between sites located close together than between sites located long distances apart. As a result, standard Mantel tests and Mantel correlograms, which are designed to test for the existence of such patterns, will often reveal little of interest to ecologists beyond the existence of the pattern of declining spatial autocorrelation with distance expected a priori from knowledge of spatial patterns in environmental variation.

The question typically of interest in many ecological phenomena is not whether any spatial structure exists per se, but rather how far spatial autocorrelation, if any, extends geographically and whether there is statistically significant cross-correlation between sites located a given distance apart. This question is not addressed by Mantel tests. For example, consider the question of whether or not there is significant positive spatial autocorrelation among sites that are very far apart. Because of the expected decline in environmental synchrony, such sites are likely to exhibit mean values below the global average and will thus yield negative z-values in a Mantel correlogram in which they are compared to the overall mean autocorrelation present in the complete data set.

To address the issue of how far spatial autocorrelation extends, we propose the following modification yielding what we refer to as "modified correlograms". To calculate modified correlograms, the correlations (r-values) between each pair of sites are divided into appropriate categories depending on the distance between the sites. Within each category, r-values are then tested by performing trials in which sets of correlation coefficients are chosen at random from the entire pool such that individual sites are used only once. For example, if the correlation between sites A and B is chosen, all other pairwise combinations involving either site A or site B (i.e., not only the correlation between A and B but also that between sites A and C, A and D, B and C, etc.) are eliminated from the remaining pool of available values. This procedure is continued until no unused sites remain. To avoid pseudoreplication, sets of r-values are tested to ensure that a particular set of correlation coefficients is used only once.

Once a complete set of correlations is chosen, the mean r-value is calculated and the number of positive and negative correlation coefficients present in the set determined. After a sufficiently large number of trials are conducted, statistical significance can be determined either based on 1) the number of trials in which positive correlations outnumbered negative correlations or 2) based on the z-value obtained by dividing the mean by the standard deviation (SD) of the mean r-values for the trials. For example, if positive mean r-values outnumbered negative mean r-values in 990 of 1 000 trials, the overall significance of the test would be considered

significant at the $p = 0.01$ level by method 1. If the mean $\pm$ SD of mean r-values for 1 000 trials were 0.068 ± 0.029, method 2 would yield a z-value of 2.35 and a one-tailed p-value of 0.009. P-values based on these alternative methods are likely to be approximately similar under most, but not all, circumstances.

In cases where all sites are being compared against all other sites, only positive mean r-values can be significant by this procedure for the intuitive reason that correlations between pairs of multiple sites cannot all be negative. Consequently, one-tailed tests are appropriate. However, this is not true if analysis involves cross-correlations between two separate data sets (see below), in which case mean values can be positive or negative and two-tailed tests should be used. When tests are being performed on more than one distance category, as will usually be the case, corrections for multiple comparisons using the sequential Bonferroni method (Rice 1989) will generally be desirable.

The fact that in many cases correlations can all be positive but not all negative raises the question of whether expected values have an equal probability of being greater than or less than zero. To check this, we performed tests on randomly-generated data to ensure that the expected mean correlation coefficients are unbiased. In order to ensure that results are unaffected by the number of data sets available for analysis, we generated between 5 and 200 sets of 10 random numbers yielding between 10 and 19 900 pairwise correlation coefficients. Ten thousand trials were performed using the above procedure for choosing sets of correlation coefficients in which each data set was represented once. The absolute values of resulting mean correlation coefficients were all <0.05 and their confidence intervals (± 1.96 SD) all broadly overlapped zero (Fig. 1). We conclude that there is no significant degree of bias associated with the test that yields biased r-values, even with relatively small numbers of data sets.

The modified correlogram procedure provides a statistical test that measures whether changes through time (typically, but not necessarily, from year to year) at sites a given distance apart tend to vary synchronously, defined as having mean r-values greater than zero. However, this definition of synchrony is much less strict than often envisioned; with large amounts of data, sites may be statistically synchronous according to the test even though mean r-values are small and close to zero.

Examples

Population dynamics of red squirrels

Ranta et al. (1997) present data on population sizes of red squirrels over a 20-yr period from 11 sites in Finland. Data within sites were standardized to zero
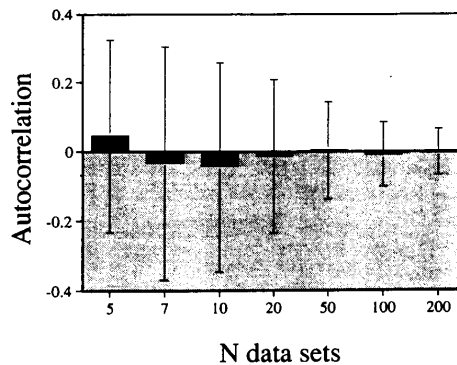
Fig. 1. Mean (± 1.96 SD) autocorrelation coefficients resulting from applying the modified correlogram technique to differing numbers of random data sets. Values plotted are for 5–200 sets of 10 random numbers. For each test, 10 000 randomizations were performed (fewer when the number of data sets ≤ 10; these were limited by the number of possible pairwise combinations). Mean r-values are all $-0.05 < r < 0.05$ and none differs significantly from zero.

mean and unit variance. The mean r-value for all 55 correlations between sites was 0.716; Ranta et al. (1997) tested the significance of this value using a randomization procedure in which they chose random sites and calculated the level of synchrony between that site and the remaining sites. This allowed them to conclude that the population dynamics of Finnish red squirrels were statistically synchronous over a large area.

A plot of the 55 correlation coefficients derived from Ranta et al.'s (1997) data is shown in Fig. 2. Using a Mantel test, the data exhibit a highly significant decline with distance, indicating that synchrony is less between
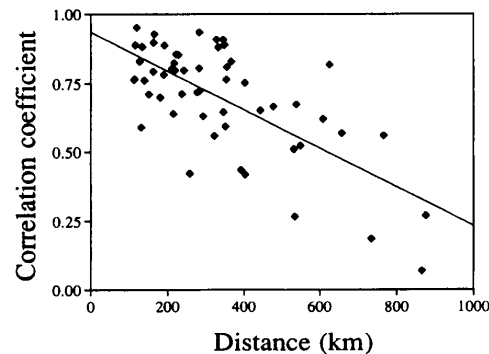
Fig. 2. Scattergram of all 55 pairwise correlations of red squirrel population densities measured between 1964 and 1983 at 11 sites in Finland by Ranta et al. (1997). Also plotted is the linear regression line based on these values, which decline significantly with distance based on a Mantel test (standardized z-value = -0.80, $p < 0.001$).
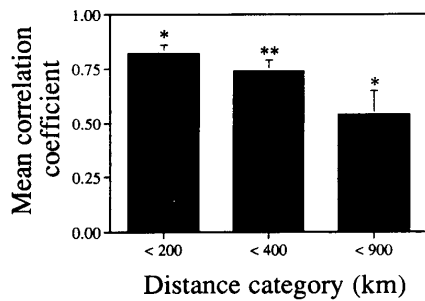
Fig. 3. Modified correlogram of red squirrel population densities measured between 1964 and 1983 at 11 sites in Finland. based on data from Ranta et al. (1997). Pairwise correlations (Fig. 2) were divided into three distance categories (0 200 km. 200–400 km. and >400 km); the most distant sites were ca 875 km apart. Mean (±SD) r-values are plotted; significance values are limited by the number of permutations available given the small sample sizes. * = p < 0.05; ** = p < 0.01.

geographically more distant sites. Although the populations taken together are statistically synchronous, is this because of the high similarity between sites that are close together, or are sites more distant from each other also statistically synchronized?

This latter question is addressed by modified correlograms (Fig. 3). Although the actual distance categories used are arbitrary, the relatively small sample sizes in this example preclude more than a small number of such categories. All three categories tested yield significantly positive mean r-values. This indicates that geographic synchrony, although declining with distance, is significantly positive even between sites 400 to 900 km apart.

Winter abundance of California land birds

Koenig (1998) analyzed spatial autocorrelation in relative abundance of California land birds using 30 yr of Christmas Bird Count data from 100 sites throughout California (Bock and Root 1981); thus, in contrast to the first example. a potentially large number of correlations were generated for the analysis. For each census. the number of birds counted was log-transformed after standardizing by dividing by the total number of party hours. Residuals from a linear regression of year on log-transformed values were used in order to ensure that results were due to regional processes rather than long-term trends in population numbers. Here we illustrate results from two species, the black-shouldered kite *Elanus caeruleus* and the red-shouldered hawk *Buteo lineatus*.

Modified correlograms (Fig. 4) indicate that black-shouldered kites exhibit significant spatial autocorrelation between sites separated by up to 500 km, while

there was no significant spatial autocorrelation for red-shouldered hawks. This result matches that intuitively expected based on the ecology of these two species; black-shouldered kites are microtine specialists (Warner and Rudd 1975), which themselves are often apparently spatial autocorrelated over considerable areas (Heikkila et al. 1994), while red-shouldered hawks are generalists that feed on a variety of vertebrates (Ehrlich et al. 1988).

In this case it was important to standardize the data by using residuals from linear regressions at each site; significantly positive spatial autocorrelation at all distance categories otherwise results not because of local synchrony between sites but because red-shouldered hawk populations increased at virtually all sites during the time period covered by the analysis. In other words, over the 30 yr of the study there were positive correlations in the numbers of red-shouldered hawks counted due to a shared population increase on a statewide scale. However, after statistically removing this trend, there was no significant correlation between sites in the relative numbers of birds counted from one year to the next.

Acorn production and growth in oaks

A third example from our own work involves spatial autocorrelation in the annual acorn production patterns of individual oaks measured using visual surveys at Hastings Reservation in central coastal California (Koenig and Knops unpubl.). Thus, in contrast to the prior examples, this case involves comparing data for individual trees rather than for means across sites. Here we focus on one of the species, the blue oak *Quercus douglasii*.
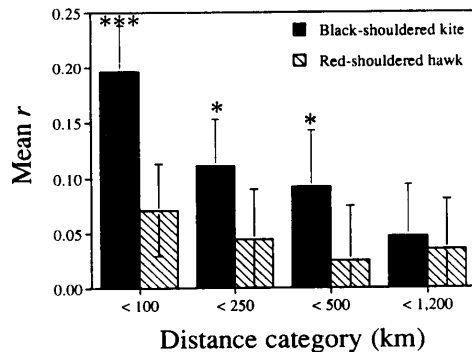

Fig. 4. Modified correlograms of relative numbers of black-shouldered kites and red-shouldered hawks based on 30 years of Christmas Bird Counts from 100 localities in California (Koenig 1998). Correlations were divided into four distance categories as shown. * = p < 0.05; *** = p < 0.001.

426

Legend (top of figure):
■ Mean r (acorn production)
■ Standardized Z (acorn production)
☑ Cross-correlation r (acorn production vs tree-rings)

Y-axis: Mean value (0.75, 0.50, 0.25, 0.00, -0.25)
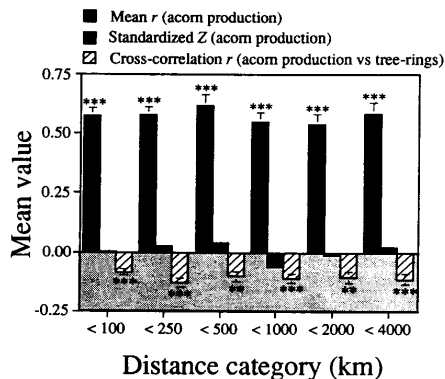X-axis (Distance category (km)): < 100, < 250, < 500, < 1000, < 2000, < 4000

Fig. 5. Analysis of 55 blue oaks over a 17-yr period at Hastings Reservation in central coastal California; values are divided into six distance categories depending on the distance between individual trees. Included are: 1) mean r-values of acorn production (modified correlogram); 2) standardized z-value of acorn production (Mantel correlogram); and 3) mean cross-correlation r-values between acorn production and radial growth (modified correlogram). ** = p < 0.01; *** = p < 0.001; standardized z-values all p > 0.05.

For each of 55 trees we collected acorn production data for each of 17 yr between 1980 and 1996. Trees were between 10 m and 3.5 km apart. As before, all pairwise correlations (of which there were 1 485) were calculated and divided into six distance categories depending on the distance between individual trees.

Annual patterns of acorn production were highly correlated between all trees at all distance categories (Fig. 5). Examination of the mean r-values from the modified correlogram fails to suggest any decline with distance, a conclusion supported by a non-significant Mantel test on the overall data set (standardized z-value = 0.02, p > 0.20). Thus, there is no overall spatial pattern in the data; trees growing 3.5 km apart have acorn production patterns just as similar as trees growing next to each other. Also graphed in Fig. 5 are the standardized z-values derived from a Mantel correlogram using the same distance categories; none of these values is significant after sequential Bonferroni correction.

As an example of a cross-correlation analysis, Fig. 5 also plots the modified correlogram obtained by cross-correlating acorn production with radial (tree-ring) growth by the same set of trees (Knops and Koenig unpubl.). This analysis involves correlating acorn production values for each tree over the 17-yr period with the standardized radial growth of each tree (including itself) during the same set of years, along with the geographic distance between the two trees involved in the analysis. As mentioned earlier, such cross-correlations involve examining the relationship between two

different data sets and thus can be either positive or negative, in contrast to the examples discussed previously. In this case, the resulting values are all significantly negative across all distance categories (Fig. 5). That is, in years when trees in our study site have relatively large acorn crops, other blue oaks in the population (including the same tree) tend to grow relatively little. As with the acorn production values alone, there is no indication of a decline with distance; this is confirmed by a Mantel test on the overall cross-correlational data set (standardized z = 0.0, p > 0.5).

Discussion

Most natural ecological phenomena display spatial structure (Legendre 1993, Ranta et al. 1995). Consequently, it is becoming increasingly clear that ecologists need to incorporate spatial autocorrelation into their models and analyses (Thompson et al. 1996). Beyond this general significance, however, there are a series of important ecological issues that hinge directly or indirectly on the existence and extent of spatial autocorrelation in various environmental and ecological variables, including questions related to population persistence, the relative importance of dispersal, and other basic issues of metapopulation dynamics and conservation biology (Hanski and Woiwod 1993). Improved methods to measure and detect spatial autocorrelation are thus of more than academic interest to ecologists.

Testing for spatial autocorrelation when the variable of interest varies temporally involves more than a normal complement of statistical problems. In particular, because there are invariably many more pairwise correlation coefficients than sites (data from n sites yields $[(n - 1) \times n]/2$ pairwise correlation coefficients), the potential for pseudoreplication biasing the results of statistical tests cannot be ignored. Authors have dealt with these problems in various ways. Hanski and Woiwod (1993) used the y-intercept of the linear regression of r-values against distance for all pairwise correlations to estimate differences in spatial synchrony among taxa, while Ranta et al. (1995, 1997) used a resampling technique in which correlation coefficients were calculated using the r-values from all pairwise correlations involving a single randomly chosen focal site. In the genetics literature, spatial data are common in the form of differing allele frequencies or heterozygosity exhibited by populations of the same species sampled at a series of sites. Standard analyses for comparing distance matrices such as these are Mantel tests (Sokal 1978, Oden 1984, Oden and Sokal 1986, Legendre and Fortin 1989).

Each of these methods focuses on the shape of the regression line between the variable of interest and distance rather than the significance of spatial autocorrelation between sites a given distance apart. Thus,

although these techniques may detect important patterns, they will often reveal little of interest to ecologists beyond the existence (or nonexistence) of a relationship between spatial autocorrelation and distance. A common alternative approach, particularly with phenetic data, is to use a clustering technique to generate phenograms. However, besides the shortcoming of not testing the significance of autocorrelations between sites a given distance apart, such techniques suffer from additional pitfalls that may obscure general spatial patterns (de Queiroz and Good 1997). Clearly techniques are needed to compare and statistically test the degree of spatial autocorrelation in different ecological or environmental phenomena varying both in time and space.

We propose such a procedure here in the form of modified correlograms that can be statistically tested using randomization procedures. The procedure can be performed on any data set involving measures of similarity between individuals or sites collected over multiple, overlapping time periods and can test for significant spatial autocorrelation within any number of distance categories, limited only by sample size.

Once spatial autocorrelation is detected, the next question is often that of causation. For example, can the significant spatial autocorrelation detected in black-shouldered kites in California (Fig. 4) be attributed to spatial autocorrelation in their microtine prey? Can the geographic synchrony in red squirrel populations (Figs 2 and 3) be attributed to synchrony in rainfall patterns? There are at least two possibilities for pursuing such questions. First, one can compare values from modified correlograms for both variables of interest calculated from data taken over the same geographic area or, ideally, from the same set of sites. For example, if analysis of data on microtine abundances taken annually throughout California revealed a pattern of spatial autocorrelation similar in magnitude and shape to that exhibited by black-shouldered kites, we could reasonably conclude that the spatial pattern in abundance of the birds was at least correlated with a comparable pattern in their prey. In contrast, if microtines exhibited no significant spatial autocorrelation, they would be unlikely to be the cause of the pattern observed in the kites.

An alternative, stronger approach is to examine the multiple data sets for existence of spatial cross-correlation between them. An example of such a cross-correlation analysis is that between acorn production and radial growth of blue oaks already discussed. In that particular case, the cross-correlation analysis allows us to conclude that acorn production by individual trees in the population is inversely correlated with radial growth by other trees in the population, even those more than several kilometers away. This indicates that a tradeoff exists between these two life-history parameters on a population level.

The essence of population ecology is to understand fluctuations in numbers over time and space (Krebs 1972, May 1976). Data sets appropriate for examining such fluctuations are becoming more commonly available, especially with the ability to post and subsequently download large data sets over computer networks. Equally importantly, the statistical power needed to conduct such tests is already well within the capabilities of most desktop computers, although the time required may be considerable for larger data sets involving many hundreds of sites. Increased awareness of the importance of spatial processes in ecological and environmental phenomena should yield significant advances in the understanding of population dynamics, both at local and larger scales.

References

Bock, C. E. and Root, T. L. 1981. The Christmas Bird Count and avian ecology. – Stud. Avian Biol. 6: 17–23.
de Queiroz, K. and Good, D. A. 1997. Phenetic clustering in biology: a critique. – Q. Rev. Biol. 72: 3–30.
Ehrlich, P. R., Dobkin, D. S. and Wheye, D. 1988. The birder's handbook. – Simon and Shuster, New York.
Fortin, M.-J. and Gurevitch, J. 1993. Mantel tests: spatial structure in field experiments. – In: Sheiner, S. M. and Gurevitch, J. (eds), Design and analysis of ecological experiments. Chapman and Hall, pp. 342–359.
Hanski, I. and Woiwod, I. P. 1993. Spatial synchrony in the dynamics of moth and aphid populations. – J. Anim. Ecol. 62: 656–668.
– , Hansson, L. and Henttonen, H. 1991. Specialist predators, generalist predators, and the microtine rodent cycle. – J. Anim. Ecol. 60: 353–367.
Heikkilä, J., Below, A. and Hanski, I. 1994. Synchronous population dynamics of microtine rodent populations on islands in Lake Inari in northern Fennoscandia: evidence for regulation by mustelid predators. – Oikos 70: 245–252.
Ims, R. A. and Steen, H. 1990. Geographical synchrony in microtine population cycles: A theoretical evaluation of the role of nomadic avian predators. – Oikos 57: 381–387.
Kelly, P. M. et al. 1989. Climate and signature years in west European oaks. – Nature 340: 57–60.
Koenig, W. D. 1998. Spatial autocorrelation in California land birds. – Conserv. Biol. in press.
Krebs, C. J. 1972. Ecology. The experimental analysis of distribution and abundance. – Harper and Row.
Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? – Ecology 74: 1659–1673.
– and Fortin, M.-J. 1989. Spatial pattern and ecological analysis. – Vegetatio 80: 107–138.
May, R. M. 1976. Models for single populations. – In: May, R. M. (ed.), Theoretical ecology. Principles and application. Blackwell, pp. 5–29.
Oden, N. L. 1984. Assessing the significance of a spatial correlogram. – Geograph. Anal. 16: 1–16.
– and Sokal, R. R. 1986. Directional autocorrelation: an extension of spatial correlograms to two dimensions. – Syst. Zool. 35: 608–617.
Ranta, E. et al. 1995. Synchrony in population dynamics. – Proc. R. Soc. Lond. B. 262: 113–118.

428

– et al. 1997. The Moran effect and synchrony in population dynamics. – Oikos 78: 136–142.

Rice, W. R. 1989. Analyzing tables of statistical tests. – Evolution 43: 223–225.

Sokal, R. R. 1978. Spatial autocorrelation in biology. 1. Methodology. – Biol. J. Linn. Soc. 10: 199–228.

Sutcliffe, O. L., Thomas, C. D. and Moss, D. 1996. Spatial synchrony and asynchrony in butterfly population dynamics. – J. Anim. Ecol. 65: 85–95.

Thompson, J. D. et al. 1996. Untangling multiple factors in spatial distributions: lilies, gophers, and rocks. – Ecology 77: 1698–1715.

Warner, J. S. and Rudd, R. L. 1975. Hunting by the white-tailed kite (Elanus leucurus). – Condor 77: 226–230.