Untangling Multiple Factors in Spatial Distributions: Lilies, Gophers, and Rocks
Author(s): James D. Thomson, George Weiblen, Barbara A. Thomson, Satie Alfaro, Pierre Legendre
Source: *Ecology*, Vol. 77, No. 6 (Sep., 1996), pp. 1698-1715
Published by: Ecological Society of America
Stable URL: http://www.jstor.org/stable/2265776
Accessed: 21/02/2011 15:30

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=esa.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Ecological Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Ecology*.

http://www.jstor.org

Concepts

Emphasizing
new ideas
to
stimulate research
in ecology

Editor's Note

This paper raises some general points about the interpretation of ecological data and focuses on the ways in which standard statistical correlations fail to serve ecological ends. Although it may not be obvious why "ecological data" differ from data in general, field-gathered data on the abundance, phenotype, and performance of organisms will in fact frequently pose two problems for correlation analysis. First, the habitats and landscapes from which data are gleaned tend to be spatially patchy, so that nearby data points cannot be considered as independent samples. New techniques are available to handle this problem. The second problem is more basic and more recalcitrant. The theory of limiting factors leads us to expect that ecological factors may often impose upper ceilings on response variables, without strongly influencing the distribution of the response variable below the ceiling. Bivariate scattergrams convey ecological information not only when they show linear or curvilinear relationships, but also when they fall into diffuse clouds, as long as the clouds have informative edges. Although ecologists care about such clouds, correlation analyses do not, and many ecologically informative data sets may be written off as "insignificant" because the wrong question is being asked. Asking the right question—is this cloud informative?—requires a new approach to data, new statistics, and new terminology. Thomson et al. bring some new statistical approaches to the attention of ecologists. Another goal of their paper is simply to encourage the objective examination of data for ecological meaning before, after, and beyond the application of "standard" analyses.

Carol Augspurger

Ecology, 77(6), 1996, pp. 1698–1715

UNTANGLING MULTIPLE FACTORS IN SPATIAL DISTRIBUTIONS: LILIES, GOPHERS, AND ROCKS[1]

James D. Thomson,[2,3] George Weiblen,[3] Barbara A. Thomson,[2]
Satie Alfaro,[3] and Pierre Legendre[4]

Abstract

Despite broad consensus on the power of experiments, correlational studies are still important in ecology, and may become more so as spatial studies proliferate. Conventional correlation analysis, however, (1) fundamentally conflicts with the basic ecological concept of limiting factors, and (2) ignores spatial structure in data, which can produce spuriously high correlations. Especially for field data, bivariate scattergrams often show "factor-ceiling" distributions wherein data points are widely scattered beneath an upper limit, due to the action of other factors. Although most ecological information in such a graph resides in the upper limit, standard correlation/regression does not characterize such limits. If other factors have been measured, path analysis may be useful, but otherwise, direct description of ecological ceilings is desirable. Objective methods for doing so are barely known to ecologists; we review recent proposals for statistical testing and data display. For correcting correlations for spatial patchiness of the variables,

---

[1] Manuscript received 7 March 1995; revised 8 October 1995; accepted 21 October 1995; final version received 18 December 1995.

[2] Department of Ecology and Evolution, State University of New York at Stony Brook, New York 11794 USA.

[3] Rocky Mountain Biological Laboratory, Post Office Box 519, Crested Butte, Colorado 81224 USA.

[4] Département de Sciences Biologiques, Université de Montréal, C.P. 6128, Succursale A, Montréal, Québec, Canada H3C 3J7.

another new technique has been proposed by Clifford, Richardson, and Hémon: by reducing the effective sample size to account for the autocorrelation it allows significance tests.

We discuss these issues with reference to counts of glacier lily (*Erythronium grandiflorum*) seedlings, vegetative plants, and flowering plants in a square grid of 256 contiguous 2 × 2 m quadrats in subalpine meadow in western Colorado. We also measured soil moisture, pocket gopher activity, and soil rockiness. All six variables showed significant patchiness (spatial autocorrelation) at similar scales. The abundance of flowering plants was positively correlated with rockiness and negatively correlated with moisture and gopher activity. Although limited seed dispersal suggests that seedlings should be spatially associated with flowering plants, no such correlation existed: indeed, examination of the bivariate scatterplot suggests a negative association, in the particular and restricted sense that seedlings are abundant only in quadrats where flowering is low. We hypothesize that seed germination is higher in less rocky areas of deeper, moister soil than in the rocky areas where most seeds land, but that seedlings seldom reach maturity unless they are in a rocky refuge from predation. Results from path analysis are consistent with this hypothesis. Such an ecological situation should weaken natural selection on characters enhancing seed dispersal.

Key words: demography; enemy-free space; Erythronium grandiflorum; *exploratory data analysis; factor-ceiling distribution; path analysis; refugia; seed dispersal and predation; source–sink distribution; spatial autocorrelation; subalpine meadow;* Thomomys talpoides.

Introduction

A more explicit treatment of spatial patterns and scaling can often enrich ecological field research. Achieving this goal may, however, require new techniques and conceptual stances. Here, we are concerned with a cluster of conceptual and statistical issues that confronted us in a descriptive study that linked plant demography, physical factors of the environment, and plant–predator interactions. Although most of these issues have been raised before, their interactions have been little discussed, and some of them remain unsolved problems. We attempt to show how an explicit consideration of these neglected issues can, first, permit us to form better hypotheses from spatial data and, second, point toward solutions for persistent statistical problems that crop up frequently in distributional studies.

Distributional restriction by species interactions.—The microdistribution of a plant population within a generally suitable habitat will reflect the process of seed dispersal, plus those factors that influence the success of dispersed seeds in germinating, establishing, and growing to a size at which they can produce seed themselves. Chance plays a role, of course, but it is commonplace to find that a population's density maps closely onto an easily identifiable edaphic variable or other abiotic factor. Plant ecologists often call such correlations habitat preferences, and describe plants as preferring shade, or well-drained soils, or chalk downs. We know that these are not preferences in the volitive sense, nor can they be analogized to habitat selection as shown by mobile animals; rather, we tend to interpret such preferences as indicating physiologically optimal habitats. In some cases, such interpretations may be warranted, but we know that the realized niche is frequently more restricted than the fundamental niche (Hutchinson 1957), and that such restriction may be indicated by reduced occupancy of some microsite types. When diffuse interspecific competition is the agent that restricts a population's ecological range, we might expect general agreement between microhabitat characteristics and physiological optima, because ceteris paribus an organism should compete best against a diverse array of competitors when it is under the least physiological stress. This general prediction loses force as competitive communities become simpler, and it may not hold at all when predators, parasites, or pathogens are the restricting agents, rather than competition. When natural enemies are involved, it is quite plausible that organisms may be concentrated in physiologically stressful but enemy-free space, as seen in both classic (e.g., the textbook example of Klamath weed [Huffaker 1957, Krebs 1978:359–360]) and recent work (Hacker and Bertness 1995).

Source–sink phenomena.—A different way in which an extant distribution may conceal the true determinants of the well-being of organisms is through source–sink effects (Shmida and Ellner 1984, Pulliam 1988, Pulliam and Danielson 1991). A sink habitat may contain a large and apparently stable standing crop of émigrés or propagules from source habitats, even though these organisms may be unable to survive and reproduce in the sink habitat. In this case, the realized habitat niche appears to exceed the fundamental niche, although this is an illusion caused by the continuous, unidirectional migration of doomed organisms. Source–sink effects and distributional restriction therefore will tend to have opposing effects on the match between observed spatial distributions and physiologically favorable areas of habitat: they could in principle cancel each other out. In general, sorting out their combined effects will be difficult and will require experimentation. However, one particular combination may leave a characteristic signature in the spatial distribution, if we examine demography as well. Suppose that predation falls disproportionately on those individuals in the most physiologically suitable sites, but a continual rain of propagules keeps a sink population present in those physiologically favorable areas. In this case, one could see spatially structured demography: younger individuals would be concentrated in the less stressful sites, where per-propagule establishment rates are higher, but these would be thinned by predation as they grew, to the point where older individuals would be mostly found in the physiologically poor sites.

Spatially autocorrelated variables.—Searching for such patterns in the field, however, is likely to entail some serious statistical impediments, which form one focus of this paper. The patterns postulated above explicitly involve both the spatial patterning and the correlation of variables. It is not universally appreciated in the ecological literature that traditional methods of assessing the strength of correlations are inappropriate for data that are spatially autocorrelated, because spatially constrained samples lack independence. If neighboring quadrats tend to have the same value for a variable, those quadrats cannot be considered truly independent. Suppose that two variables show a checkerboardlike patch structure where each patch encompasses about 100 contiguous quadrats. With a sample area only a few hundred quadrats in extent, the large patches of the two variables could easily fall into alignment by pure chance, but a correlation coefficient calculated across all quadrats would appear to be highly significant. Recent techniques (Clifford et al. 1989; see also Dutilleul 1993, Legendre 1993) allow a more conservative evaluation of such relationships. In essence, patch structure reduces the degrees of freedom, and this loss is estimated by the Clifford–Richardson–Hémon (CRH) procedure; it calculates the "effective sample size" (less than the actual number of quadrats), allowing an autocorrelation-corrected significance test. Although some aspects of the CRH techniques are controversial (Dutilleul 1993), this general approach seems capable of providing practical solutions to the lack-of-independence problem.

Factor-ceiling distributions.—A deeper problem arises from a seldom-resolved tension between ecological theory and correlation analysis. Ecological theory embodies the concept of limiting factors; correlation looks for controlling factors. In consequence, correlation analysis may be shortsighted or even blind to informative aspects in ecological data sets. Imagine a physiological response variable (e.g., growth rate) that is linearly determined by a single independent variable (e.g., local food density). If all other potential influences are at optimal or nonlimiting levels, the scattergram of the growth rates of many individuals vs. their local food densities would yield a line of points; simple regression/correlation methods would easily detect the relationship (Fig. 1a). Suppose, however, that in some cases, additional uncontrolled factors other than food limit the growth rates of some individuals. Then, some data points will be dragged below the ceiling line, and if these additional factors act strongly and frequently enough (i.e., if growth rate

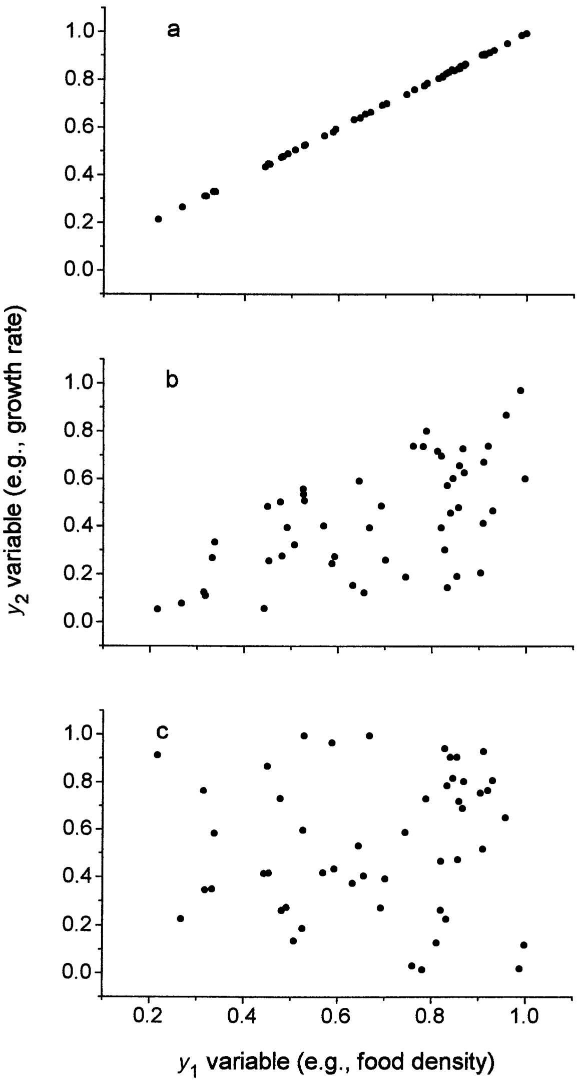

FIG. 1. Examples of possible bivariate distributions. (a) A tight, simple correlation of the form $y_2 = y_1 \pm$ a small amount of random variation. This relationship is well fit by standard techniques. (b) A ceiling function in which both x and y are drawn from random uniform distributions but y_2 is constrained to be $<y_1$. Standard regression would not locate the ceiling. (c) Both y_1 and y_2 are drawn from uniform distributions.

is frequently driven to zero), the correlation between food and growth may approach zero (Fig. 1b). Even with no correlation, however, the upper edge of the point cloud may still provide evidence of an ecologically important connection between the variables, in the sense that this edge may still delineate a ceiling that food density places on growth. A lopsided point cloud of this sort conveys a different meaning than one in which there is no relationship whatsoever (Fig. 1c), but conventional correlation analyses are poorly suited to make the distinction. Kaiser et al. (1994) develop this argument at length, with respect to Liebig's law of the minimum and nutrient limitations of productivity in lakes.

Terminology.—Bivariate patterns such as Fig. 1b have been termed "triangular distributions" by Maller (1990, Maller et al. 1983). Kaiser et al. also restrict their treatment to triangular cases, but potentially limiting relationships need not be triangular. Following our usual reasoning about ecological ranges of tolerance, we would expect that such ceilings could often be bell shaped, or at least humped. Indeed, an observed triangular distribution may frequently represent one side of a bell-shaped one, where the driving variable has not been sampled over a wide enough range to reveal the whole relationship. As there seems to be no general ecological term for this class of relationships, we propose one: "factor-ceiling distribution." There may be some correspondence between ecological ceilings and "production frontiers" as studied in economics (e.g., Färe et al. 1994). Production frontiers represent upper boundaries of production efficiency toward which organizations theoretically strive but generally fall short. However, the goals and methods of this branch of economics seem so remote from ecological concerns that we see potential confusion in adopting the term "frontier" for ecology.

Although the law of limiting factors provides the clearest theoretical basis for factor-ceiling relations, ecological ceilings need not be imposed by limiting physical factors. For example, two species of plants or sessile invertebrates may compete for space over a gradient in such a way that, in the absence of disturbance, they crowd each other so that their densities are negatively correlated across a set of sample plots. If plots are frequently swept clean by disturbance, however, there may be many plots where both species are absent. Enough data points at or near the origin can make the correlation coefficient positive, even though the competitively generated ceiling is negative. Different parts of a point cloud illuminate different ecological processes.

Despite the importance of ceilings in ecology, there is no generally accepted, objective statistical method for detecting them, i.e., for distinguishing Fig. 1b from Fig. 1c. Many ecologists, on finding no significant correlation in a pattern like that in Fig. 1b, would simply stop at that point. Others might offer a verbal argument for why the data are ecologically indicative despite the lack of correlation, but it is hard to make convincing arguments without a compact terminology and some objective mathematical procedures (witness the length and the labored wording of the previous paragraphs). Maller (1990, Maller et al. 1983) has proposed an objective approach that simultaneously fits a boundary function and a mixing function that describes how much the points have been dragged below the boundary by unaccountable factors. It has been used at least once in ecology (Rabinowitz et al. 1985). Its dependence on a particular model, however, robs it of general applicability and introduces a serious problem of statistical nonidentifiability (see *Discussion*). The approach of Kaiser et al. (1994) is similar in that it also involves simultaneously fitting the upper boundary, the distribution of points falling below the boundary, and the degree of random error. These methods have great promise, but their assumptions and performance have not yet been subjected to much scrutiny by ecologists. We do not apply these tests here.

In some cases, the additional uncontrolled factors may have been measured. If so, traditional techniques of multiple regression and (especially) path analysis may be sufficient to reveal the action of the multiple factors that produce ambiguous factor-ceiling scattergrams. In field data, such cases are probably fortunate exceptions. Individual data points are subject to too many influences, including idiosyncratic historical events. We cannot generally depend on path analysis to clean up these data. We do not offer any true statistical tests here, but we present some new display techniques for exploratory data analysis (EDA, sensu Tukey 1977). We hope that these techniques will make it easier to find, describe, and discuss subtle relationships that might otherwise be overlooked in ecological data.

We develop these issues in reference to a study of the distribution of *Erythronium grandiflorum* Pursh (Liliaceae) in subalpine meadows of the Colorado Rocky Mountains, where microsite-dependent predation and source–sink effects both appear to play roles as outlined above. Our attention was first drawn by the obvious patchiness of flowering plants, then by the patchiness of seedlings, then by the discordance of these distributions. By measuring the patch structure of various biotic and abiotic variables, we developed a hypothesis for the observed distribution: it is similar to the rock-refuge hypothesis advanced by Cantor and Whitham (1989) for aspen. In addition to discussing the statistical intricacies, we outline some implications for population dynamics, and show how the ecological situation provides opportunities for microhabitat-dependent gender asymmetry, as well as setting up a spatially structured demography. *Erythronium grandiflorum,* in contrast to ant-dispersed congeners, may be under selection to minimize long-distance dispersal (cf. Zohary 1937, Ellner and Shmida 1981).

Methods

Study species

The glacier lily *Erythronium grandiflorum* is a long-lived perennial geophyte that emerges and flowers immediately after snow melt. Most plants are vegetative and produce a single leaf. Flowering individuals produce a single scape, usually with a single flower, although two- and more-flowered plants may comprise up to 50% of the flowering plants in some localities. Fruits typically contain 30–50 large (3 mm) seeds. Seed dispersal is very restricted. Primary dispersal is aided only by the post-flowering elongation of the scape, which increases the distance to which seeds can be tossed out of the apically dehiscing capsule when it sways in the wind. However, this mechanism seldom throws seeds >1 m (median = 28 cm; Weiblen and Thomson 1995).

Other species of *Erythronium* (e.g., *E. americanum*: Handel et al. 1981, Ruhren and Dudash 1996) produce seeds with elaiosomes that are secondarily dispersed by ants, but *E. grandiflorum* seeds have no elaiosome and are not taken by ants or rodents in short-term cafeteria trials (Weiblen and Thomson 1995). They are too heavy to be much affected by wind. There may be some movement by the runoff of spring meltwater, but this effect seems likely to be important only along streams or erosion channels, which are not primary habitats for *E. grandiflorum.*

In garden plots, broadcast seeds of *E. grandiflorum* are usually unsuccessful, but seeds that are lightly buried by raking produce seedlings prolifically (J. D. Thomson, *personal observation*). We do not know why exposed seeds fail; desiccation is likely, but predation may also become important on longer time scales than examined by Weiblen and Thomson (1995). Seedlings comprise a single, grasslike, cotyledonary seed leaf, and they are thus easily distinguished from older vegetative plants, which have broader bladed leaves. After the seedling season, the plant dies back to a buried corm. In succeeding years, the corms not only increase in size (up to 15 g fresh mass), but also become more deeply planted through the production of vertical droppers (Robertson 1906). Based on growth rates in garden studies (J. D. Thomson, *personal observation*), it takes >5 yr before a plant can flower. By this time, the corm is ≈30 cm below the surface, a depth at which dropper production seems to cease, causing the corms to stay at this depth. When plants are grown in deep pots, the corms take on the straight conical form that lends *Erythronium* one of its common names, the dog-tooth violet. However, when excavating wild plants, it is unusual to find corms in such condition. Instead, most are distorted and asymmetrical because they are growing in crevices between rocks. In contrast to the situation in some congeners (e.g., *E. americanum*), vegetative reproduction is rare, and plants do not expand into clonal patches.

Habitats

Our study sites were in subalpine meadows on glacial till in wide valleys. There are frequent outcrops that may represent bedrock, but because the till contains vast quantities of rocks of all sizes, the outcrops may also be the tops of large boulders. Much of the rock, especially Mancos shale, is highly broken up and fissured. The meadows, which are surrounded by spruce–fir forest, support a diverse herbaceous community, with forbs dominant over grasses. The principal aboveground grazers are cervids, sciurids, and mice. Belowground, pocket gophers (*Thomomys talpoides fossor* J. A. Allen: Geomyidae; Armstrong 1972) are very active. Soil from their burrowing activities often covers a substantial portion of the ground.

Sampling grid

In May 1992, we laid out a 16 × 16 grid of contiguous 2 × 2 m quadrats (Fig. 2) in a moderately dense population of *E. grandiflorum* in Washington Gulch, Gunnison County, Colorado. This population is continuous for several square kilometres in this valley. The study site was slightly sloping, with a southern exposure, and included some areas of outcropping rock, as is typical for these meadows. In each 2 × 2 m square, we measured the following variables.

1. *Flowers.*—We counted flowering plants of *E. grandiflorum.* These were categorized by flower number, which ranged from 1 to 4. For the analyses, we use the total number of flowers in a quadrat. This number is highly correlated with the number of flowering plants; we use it because we view a site with a higher proportion of multiple-flowered plants as more favorable for flower production than a site with the same number of single-flowered plants. We did not count a quadrat until all flowers had presented themselves.

2. *Rockiness.*—To sample surface and subsurface rockiness in the growing zone of *E. grandiflorum,* we used a set of four tempered steel rods, 9 mm diameter,

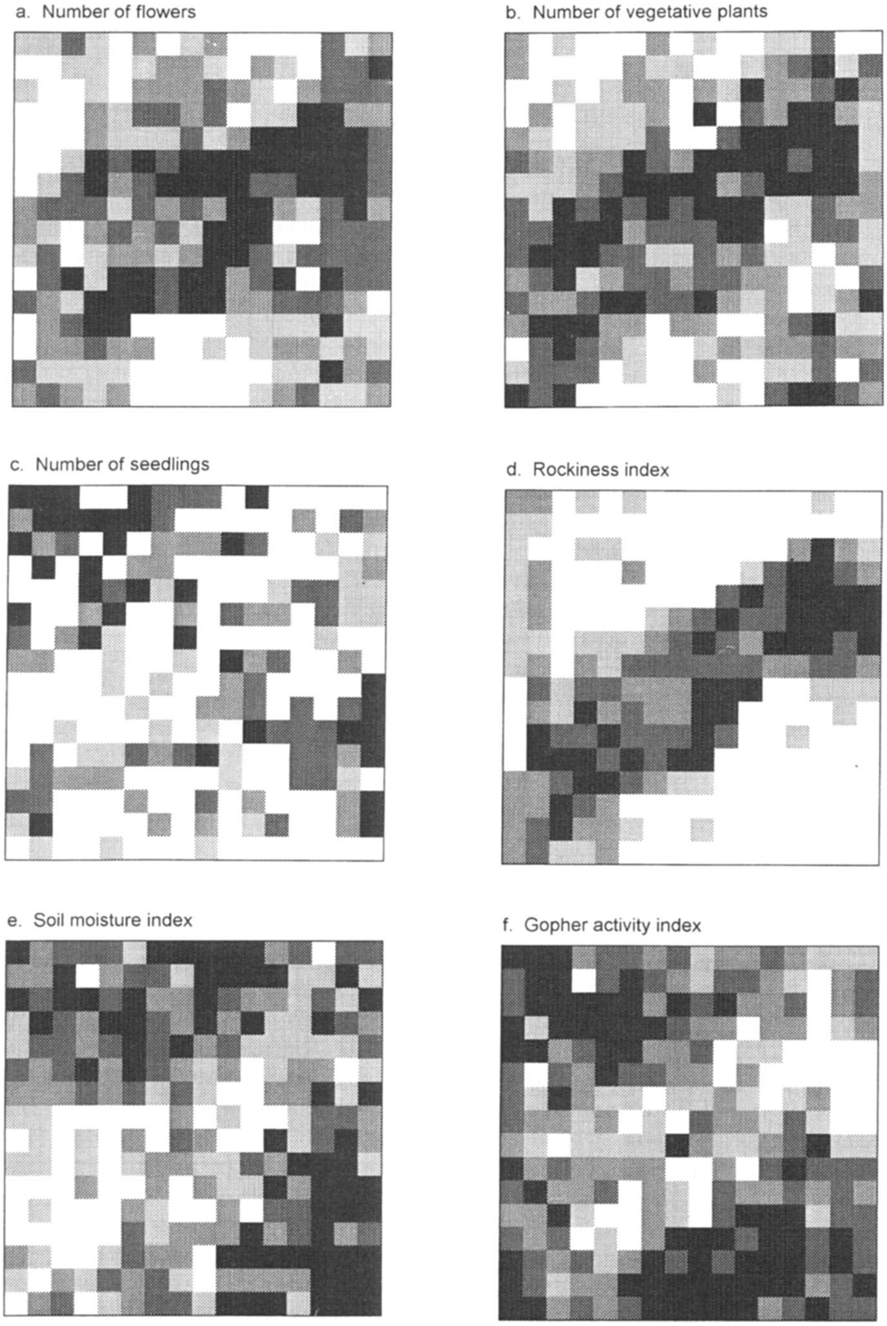

FIG. 2. Maps of the Washington Gulch grid of 256 2 × 2 m quadrats for the six variables described in *Methods*. Darker shading corresponds to a higher value of the variable. Cutoff values for shading levels were chosen to produce approximately equal frequency quintiles, insofar as the distributions permitted.

held vertically in a frame bolted together from perforated steel angle. The rods were spaced 18 cm apart. The rods could slide up and down in their frame, with a travel of 31.5 cm. We assessed rockiness by hammering the rods into the ground as far as possible, then measuring the height of rod remaining above the frame. Eight rods were driven per quadrat, oriented along a diagonal. Our index of rockiness was the sum of the protruding rod heights, with a potential range of 0–252 cm.

3. *Gopher activity.*—On 30 May, we roughly assessed winter activity of pocket gophers by counting how many of their winter soil cores crossed the two diagonals of the quadrat. These cores, also called "gopher garlands," represent casts of tunnels that the rodents make to store the soil brought up from their subterranean burrowing. Their quantity therefore provides an index of gopher burrowing activity over a period of months (Ellison and Aldous 1952). Easily seen shortly after snow melt, they are soon obliterated by rain. The counts ranged from 0 to 13, with the highest values corresponding to over 50% coverage of the quadrat by soil casts.

4. *Soil moisture.*—On 3 July, after soils had begun to dry from their drenching by snowmelt, we obtained a relative measure of soil moisture with a conductivity meter (Aqua/Light Meter, J and M Instruments Corporation). Although this is only an inexpensive device designed for houseplant care, it provided stable, reproducible readings on a scale from 0 (dry) to 8 (wet), readable to the nearest 0.5 unit. We placed the probe 15 cm into the soil at the center of the quadrat; if rock prevented sampling there, we moved away from the center until a 15-cm reading could be obtained. Readings ranged from 1.0 to 8.0.

Juvenile plants were too numerous to count over the entire 1024-m^2 plot. Instead, we laid out a strip of 10 subplots, each 20 × 20 cm, across the center of each 2 × 2 m quadrat for enumerating the last two variables:

5. *Seedlings.*—After their emergence was complete, first-year plants (cotyledonary leaves) were counted within each 20 × 20 cm subplot ($n = 2560$).

6. *Vegetative plants.*—Nonflowering plants (with a single true leaf) varied greatly in size. Some were young plants that had never flowered; others may have flowered in previous years but had declined to the vegetative state. We counted them in the 20 cm × 2 m strips, not breaking the strips down into the 10 subplots.

Scattered quadrats

In similar habitats at nearby Irwin, Colorado, we sought more spatially extensive data on the relationship between the density of flowering plants, gopher activity, and soil rockiness. An observer walked long, straight transects and placed a 1 × 1 m quadrat frame at 10-pace intervals. In each quadrat, the number of *E. grandiflorum* flowers and buds was counted. Rock and gopher activity were scored subjectively as follows. Rock scores were: 0 = open soil; 1 = 1–39% quadrat coverage by gravel; 2 = 40–100% gravel; 3 = 1–20% rock; 4 = 21–40% rock; 5 = 41–100% rock. Gopher scores were based on the estimated proportion of the quadrat covered by soil casts: 0 = none; 1 = 1–5%; 2 = 5–20%; 3 = 20–100%. One hundred quadrats were sampled along two parallel transects, ≈50 m apart.

Analyses

To consider various scales of spatial patchiness, for the variables measured in the 16 × 16 grid, we constructed all-directional spatial autocorrelograms (Sokal and Oden 1978*a*, *b*), using the SAAP package (Wartenberg 1994) with Moran's I as our autocorrelation statistic. For seedling data, we additionally compared the frequency distribution of the 2560 subplots to a Poisson distribution, using a G test following Sokal and Rohlf (1995).

To examine relationships among the variables we first plotted frequency distributions and bivariate scattergrams to examine the variables visually and to suggest hypotheses. We then took two very different approaches to interpreting the interrelationships among the variables. One uses standard, parametric, least squares regression procedures to construct a path analysis (Sokal and Rohlf 1995) that embodied a particular hypothesis about the interaction of the variables. This should be considered our "optimistic" approach, in that some of the data do not strictly fit the assumptions of path analysis. First, although most of the variables can be transformed to conform to the assumption of normality, rockiness cannot. Second, standard path analysis assumes that the data are represented by 256 independent data points, but in fact the variables all show spatial autocorrelation, which means that points are not truly independent. Third, some of the bivariate scattergrams (particularly, seedlings vs. flowers) may indicate factor-ceiling relationships that are worth describing and discussing directly, in addition to applying path analysis.

Instead of downplaying these features that make our data resistant to traditional methods, our alternative "pessimistic" analysis tries to confront them directly. First, we use nonparametric Spearman coefficients to gauge the strength and direction of the bivariate relationships. Second, we correct the significance test for those correlations by using the CRH procedure (Clifford et al. 1989) to take the underlying spatial autocorrelation into account. Third, we present new suggestions for examining factor-ceiling distributions. Two are data display techniques that we term "partitioned regression" and "logistic slices"; the third is the "range-restriction test," a test for bivariate skewness proposed by P. Wilson (*unpublished manuscript*). All are more easily explained with reference to a particular data set, so we defer their description to the

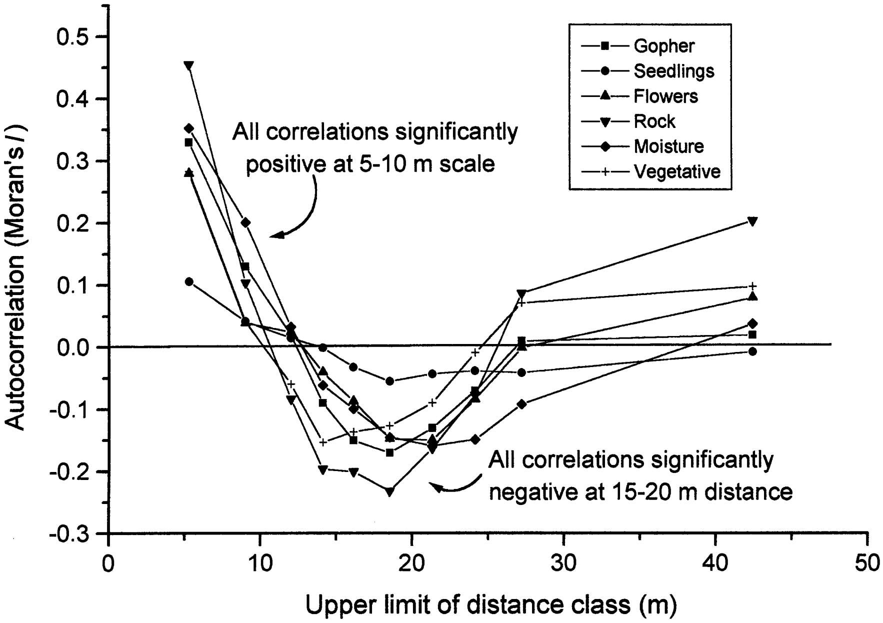

FIG. 3. Correlograms from the Washington Gulch grid of 256 2 × 2 m quadrats for the six variables described in *Methods*. Moran's autocorrelation statistic I indicates how similar are values of the variable in quadrats that are separated by the distance class indicated. For all variables, there is significant positive autocorrelation in nearby quadrats and significant negative autocorrelation at distances of 15–20 m, indicating that all the variables are highly patchy and show similar scales of patch structure.

Results section, where we apply all three to the seedlings vs. flowers relationship.

Results

Spatial pattern of variables considered separately

Maps of surfaces for the six variables on the 256-quadrat grid (Fig. 2a–f) suggest that all of them show spatial patch structure. The clearest expression of this is in rockiness: a strong outcrop bisects the study plot from southwest to northeast, with deeper soils in the southeastern corner and the central northern edge. The other variables appear to show patch structure at similar scales, and the patterns suggest correlations with rockiness. The visual impression of similar scales of patch structure is confirmed by the spatial correlograms for these variables (Fig. 3), which all start out positive at the smallest distance class tested (0.0–6.3 m), then all cross over to become negative in either the 9.0–12.0 m or the 12.0–14.1 m distance class. The variable with the least marked patchiness (at the scales examined here) is the total number of seedlings, but even this relatively flat correlogram shows significant autocorrelation: positive at 0.0–6.3 m (two-tailed $P < 0.00002$), and negative at 16.1–18.5 m ($P < 0.002$). All the other variables showed even more significant autocorrelation at these same scales.

Because seedlings were scored in the small 20 × 20 cm subplots, we can also analyze them for patchiness at this smaller scale. Although the spatially constrained sampling plan renders a significance test somewhat inappropriate, the very poor Poisson fit seems sufficient to reject the null hypothesis of randomness. The seedlings were highly aggregated (Fig. 4; goodness-of-fit $G = 1953$, df = 5); in fact, it was common to see tight clusters of many seedlings (right tail of Fig. 4). We believe that these represent cases where a ripe fruit capsule has fallen without dispersing its seeds. Rodent caching is another possibility, but is inconsistent with the unattractiveness of seeds to rodents (Weiblen and Thomson 1995).

In summary, flowering plants, vegetative plants, seedlings, gopher activity, soil moisture, and rockiness are all patchy in the study plot, with a patch dimension of roughly 10 m. This patchiness is least marked for seedlings. Seedlings also show strong patchiness at a much smaller scale, ≈20 cm. Although the other variables were not sampled at this scale, it is our strong impression that established plants are much less clustered than seedlings at small scales.

Inter-relationships of variables

For brevity, we have dropped vegetative plants from these analyses; basically, they show the same patterns as flowering plants. Fig. 5 gives a scatterplot matrix

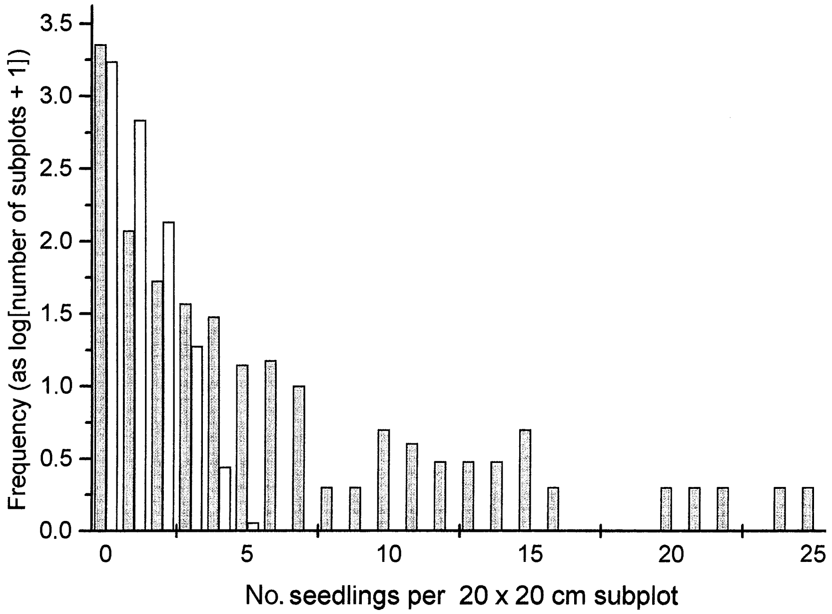

FIG. 4. Observed frequency distribution of seedlings per subplot for the 2560 20 × 20 cm subplots (shaded bars) compared to a fitted Poisson distribution (open bars). The log transformation simply improves the visibility of the right-hand portion of the distribution. There are far more quadrats with high numbers of seedlings and with no seedlings than expected by chance.

(from SYSTAT; Wilkinson 1990) of the raw data for the remaining variables. In addition to showing the frequency distribution of each variable, this display gives all of the bivariate scatter diagrams, with the major and minor axes indicated by 95% ellipses, following Sokal and Rohlf (1995). The data are untransformed, and several of the frequency distributions are skewed; therefore the ellipses should be considered only approximate guides to the eye. The Spearman correlations (Table 1) reinforce this picture in a more reliable fashion.

All of the variables except total seedlings are strongly intercorrelated, as can be seen from the scatter diagrams, the Spearman coefficients, the orientation of the ellipses, or even the surface maps directly (Fig. 2): flowering is concentrated in the rockier areas, while gopher activity and soil moisture are lower where rock outcrops exist. Concomitantly, gophers and moisture are positively associated, while flowering is negatively associated with moisture and gopher activity. Considering the data as 256 independent points, all of these correlations are highly significant, although none of them are very tight (Table 1). However, the patchiness of each of the underlying variables casts doubt on these conventional significance levels (Legendre 1993). Applying the CRH procedure to the six correlations discussed above causes all of the significance levels to decline (Table 1), but only the total flower-moisture and the gopher activity-moisture correlations lose significance ($P > 0.05$).

For seedlings, the uncorrected Spearman coefficients indicated significant positive associations with gophers and soil moisture, a negative association with rock, and no relationship with flowering plants. When CRH-corrected, however, all of these correlations lose their significance. This lack of significance does not necessarily imply a lack of ecological interest. Indeed, the one pairwise correlation with absolutely no trace of significance, that between flowers and seedlings, is the most noteworthy, because the extremely limited seed dispersal documented by Weiblen and Thomson (1995) leads to an a priori expectation of a strong positive spatial association between flowering and seedlings. Therefore, the lack of a correlation is unexpected and demands an explanation.

One possible explanation is that seeds usually desiccate in the rocky areas where most of them fall; that seed germination is higher in less rocky areas of deeper, moister soil; that gophers kill most seedlings in the deep-soil patches; that gophers are excluded from the rocky patches; and that therefore the rocky patches comprise source areas from which some seeds disperse outward into surrounding sink areas, where the seedlings eventually perish. If the seed mortality differential is strong enough between rocky areas and deep soil, we might even expect a negative relationship be-

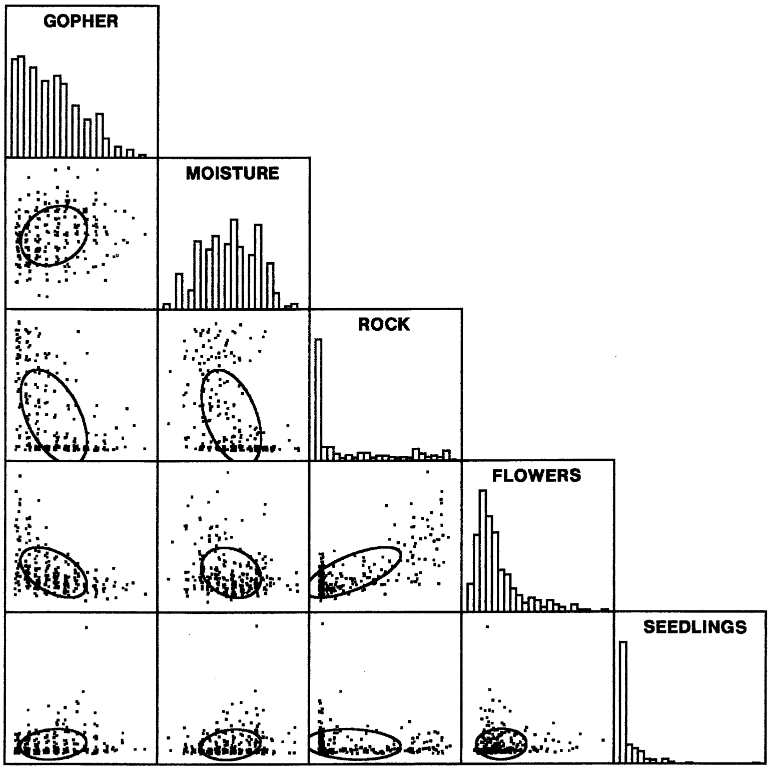

FIG. 5. Scatterplot matrix showing frequency distributions and all pairwise scatter diagrams, with the major and minor axes indicated by 95% ellipses, for the five principal variables from the Washington Gulch grid of 256 2 × 2 m quadrats. Vegetative plants (not shown) are distributed similarly to flowers. Variables are counts (flowers, vegetative plants, seedlings) or indices in arbitrary units (others). Duplicate points with the same location have been "jittered" or offset slightly for visibility. Means, medians, and standard deviations are: flowers, 29.9, 23, 23.3; vegetative plants, 11.8, 8, 11.2; seedlings, 3.9, 1, 7.4; moisture, 3.8, 4, 1.6; rockiness, 53.0, 14.5, 68.5.

tween flower and seedling abundances, i.e., a spatially structured demography of the sort described in the *Introduction.*

Indeed, if one considers the scattergram for seedlings vs. flowers (Fig. 5) as a factor-ceiling problem rather than a correlation problem, there is a suggestion of a negatively sloped ceiling. The data points generally fill in a triangle: one can find quadrats with zero or few seedlings at any number of flowers, but seedlings are plentiful only where flowers are sparse. That

TABLE 1. Correlation analyses for the 256 2 × 2 m quadrats at Washington Gulch. Below the diagonal, we report Spearman rank correlation coefficients (r_s). Above the diagonal, the uncorrected significance (*P*) value for the correlation coefficient is followed by the value when corrected for spatial autocorrelation by the Clifford–Richardson–Hémon (CRH) procedure. These are based on testing r_s as if it were a Pearson coefficient, which should be accurate because sample sizes exceed 29 even after their reduction to "effective" sample sizes by CRH.

	Gopher activity	Moisture	Rockiness	Total flowers	Total seedlings	Vegetative plants
Gopher activity		**0.001 → 0.208**	0.000 → 0.005	0.000 → 0.001	**0.032 → 0.197**	0.000 → 0.005
Moisture	0.241		0.000 → 0.014	**0.000 → 0.136**	**0.008 → 0.128**	0.000 → 0.013
Rockiness	−0.544	−0.514		0.000 → 0.000	**0.027 → 0.216**	0.000 → 0.001
Total flowers	−0.503	−0.197	0.515		0.995 → 0.997	0.000 → 0.000
Total seedlings	0.071	0.099	−0.054	0.163		0.085 → 0.250
Total veg. plants	−0.450	−0.364	0.546	0.582	−0.057	

Note: Boldface type indicates loss of significance. Underlined values reflect correlations that retained their significance after correction.

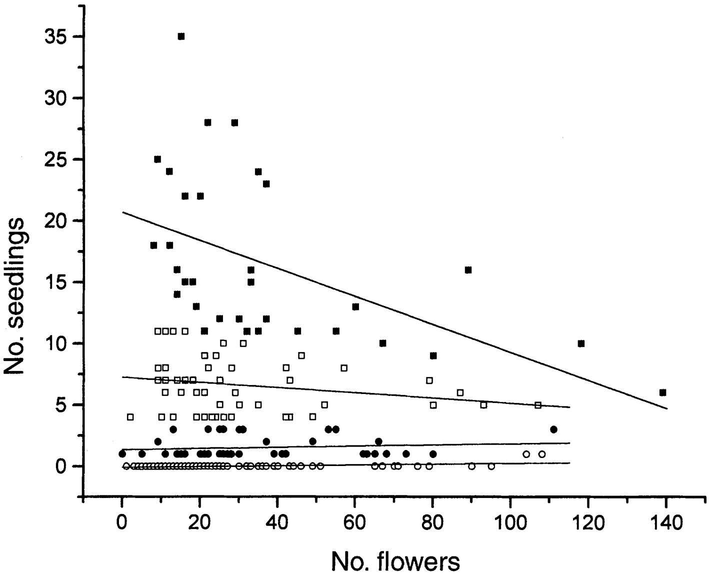

FIG. 6. Two-cycle partitioned regression of seedlings vs. flowers for the Washington Gulch grid of 256 2 × 2 m quadrats. First, a major-axis regression line through all the points was used to partition the data into upper and lower subsets. Then each subset was similarly partitioned, yielding four subsubsets denoted here by different symbols. Major axes are shown for each of the subsubsets. One outlier, evident in Fig. 5, has been removed. Points are not jittered, so some near the origin are obscured.

is, whatever the ecological relationship between flowering and seedling densities, it concerns upper limits, not central tendency, and is best revealed in the upper edge of the cloud of points. The interior points indicate quadrats where seedling number is depressed by factors independent of flower density, and we are therefore less concerned with it. A negatively sloped ceiling is even more unexpected than the neutral relationship suggested by correlation methods; we would therefore like an objective statistical method to measure and confirm this aspect. Because the possibility of spatial disjunction of seedlings and mature plants is important to the source–sink hypothesis discussed, we will consider this particular relationship at length.

Analysis of the 100 noncontiguous quadrats from Irwin gives similar correlations among gopher activity, flowering, and soil rockiness. Triangular distributions prevail here as well, but simple correlations are all highly significant and adequate to indicate the directions of the relationships: gopher activity and total flowers, $r_s = -0.52$; gopher activity and rockiness, $r_s = -0.54$; rockiness and total flowers, $r_s = +0.51$, all with $P < 0.0001$. Therefore, the patterns found in the 16 × 16 grid at Washington Gulch are not unique to that site or that scale, but occur elsewhere in transects spanning 0.5 km. Our further analyses, however, concern only the Washington Gulch grid.

EDA for factor-ceiling distributions

Partitioned regression.—Standard, familiar regression techniques may be used to produce a visual display of the internal structure within a point cloud, by successively partitioning the data by major-axis regression. First, a line is fit through the entire cloud. Second, the residuals are examined, and the data set split into two subsets, those above and below the line. Third, separate regression lines are fit to the two subsets. If the data are plentiful, the process can be continued through more cycles of subdivision, producing a spray of regression lines through different zones of the original point cloud. Two cycles (i.e., four lines) are shown for the seedlings vs. flowers relationship in Fig. 6. The uppermost line, which describes the bivariate relationship at the top of the point cloud, has a definite negative slope. In the data set as a whole, however, this trend is canceled by the lack of structure lower in the point cloud and the large number of quadrats with no seedlings at all. The result is an overall correlation of zero, but the spray of lines indicates the internal heterogeneity of the point cloud. If one uses the same technique on truly random data, all of the lines will be approximately horizontal, denoting lack of internal structure.

Edge detection; logistic slices.—A second strategy is to try to locate the upper edge of the point cloud. There are probably many ways to do this, and ongoing advances in geographical methods will doubtless provide more. Raster-based geographical information systems, for example, provide edge-enhancement filters that sharpen the boundary between blurrily separated areas in which pixels have different values. These work by changing the values of pixels based on their previous values and those of nearby pixels (Eastman 1993:78). Although we rejected the idea of changing data values, we did adopt the general strategy of ren-

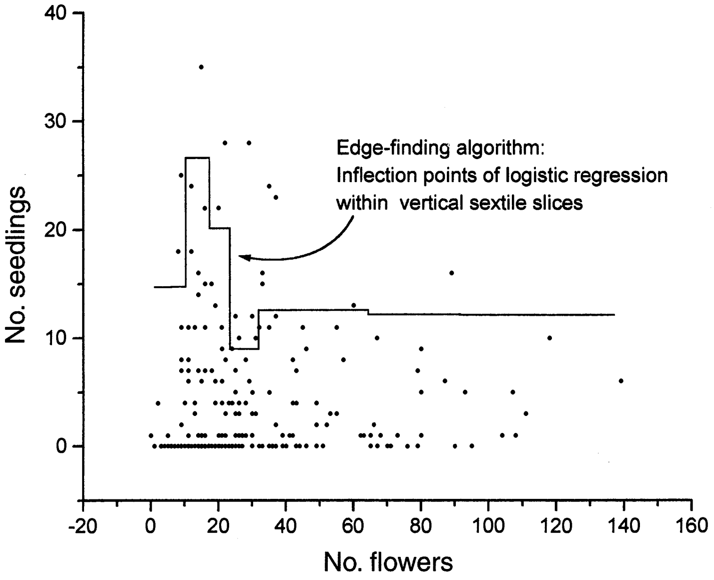

FIG. 7. Logistic slicing of the seedlings vs. flowers relationship. The lengths of the horizontal line segments indicate the widths of the six vertical slices, and their vertical position indicates the height of the inflection points that approximate the edge of the point cloud.

dering a scattergram into pixel-like areas. In this method, we first divide the scattergram cloud into a number of quantile slices, vertical strips each containing equal numbers of data. These slices vary in width depending on the local density of data points. We then make horizontal slices (of equal width), dividing the scattergram plane into rectangular cells that can be analogized to pixels. We used the presence or absence of a data point to classify each cell as having a value of 1 or 0. The problem then is to locate the point within each vertical slice where mostly data-filled cells give way to mostly data-empty cells. We used logistic regression to do this by (in effect) turning each vertical strip on its side and fitting a sigmoidal curve through the 0's and 1's. The inflection point of the logistic curve provides an objective location for the transition from primarily 0's to primarily 1's, i.e., of the edge of the point cloud within each strip.

Logistic slicing of the seedlings vs. flowers relationship (Fig. 7) does suggest the negative relation suggested above, but it is apparent that the method suffers from the excessive width of the right-most slice. This method requires a high density of points so that slices can be narrow enough to be informative. Thus it is not likely to perform well where the bivariate distribution tapers out into a sparsely filled wedge. On the other hand, Fig. 7 suggests the interesting point that the actual edge of the scattergram might be humped, with the highest numbers of seedlings occurring in quadrats with intermediate numbers of flowers. This is biologically reasonable (see *Discussion*), but of course was not detected by linear partitioned regression.

We discovered a similar approach by Blackburn et al. (1992) while our paper was in press. These authors also slice the point cloud vertically, but into strips of equal width rather than equal point density. Where we use logistic regression to locate the edge of the cloud, they simply choose the highest datum in each slice and fit a regression through those points to estimate the edge. Their method successfully recovers known upper edges in artificial data sets if there are many points and not too many slices. Although we have not applied it to the distribution in Fig. 7, it would presumably detect the declining slope of the edge on the right-hand portion of the cloud, but not the humped curve suggested by logistic slicing.

Tests for lopsidedness.—Although the EDA techniques provide objective ways to dissect scatter diagrams and reveal their internal structure, they cannot be regarded as significance tests. One step in this direction is the range-restriction test (P. Wilson, *unpublished manuscript*). Essentially, one examines the point cloud; if the points appear to cluster in one corner, extreme points of the distribution are used to draw the diagonal line that separates the more crowded corner from the less crowded. Then the numbers of points above and below this diagonal are compared. (Subextreme points may also be examined for additional robustness.) In a triangular point cloud, these counts will differ sharply; in the absence of any "shape" to the point cloud, the counts will be similar. Wilson

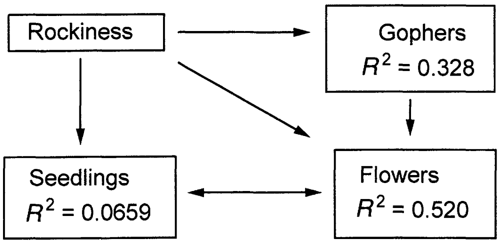

FIG. 8. Diagram of the influences modeled in the path analysis. The R^2 values indicate the proportion of variance in each variable explained by the model (see Table 2).

derives a statistic to measure this asymmetry, then uses a randomization technique to assess the probability of finding the observed statistic under a randomization of one of the variates with respect to the other. Wilson applied this procedure to the seedlings vs. flowers data considered here, and concluded that there was a marginally significant negative relationship (P. Wilson, *personal communication*; the relationship is significant at $P < 0.05$ for some endpoints but not for others). Although Wilson regards this test more as an EDA technique than a true statistical test, his result does support the paradoxical conclusion that, at the scale of 2 × 2 m quadrats, high densities of *E. grandiflorum* seedlings are significantly associated with low densities of flowers. Although we do not present analyses here, inspection of Fig. 5 suggests similar triangularity in several of the other scatter diagrams.

Path analysis

Before using path analysis (Sokal and Rohlf 1995) to address our hypothesized relationships, we sought appropriate transformations for the variables whose distributions are shown in Fig. 5. Soil moisture needed no transformation. For flowers and gopher activity, we used square roots, and for seedlings, ln(seedlings + 1). Rockiness showed a nonnormal but untransformable distribution, nearly flat except for a large number of zeros, and was left raw. We examined the residuals after fitting path models; they were reasonably well distributed. In classical path analysis, there is no allowance for spatial autocorrelation of variates; to make our interpretation more conservative, we report as significant only those relationships for which the conventionally reported P values were <0.001.

Exploration of a series of models indicated that soil moisture had no measurable effect of its own if the model also contained rockiness, and that rockiness explained the spatial variability of the plants better than moisture, making moisture a redundant source of variation. We eliminated moisture from the final model, which was chosen to explore the hypothesized relationship described (Fig. 8, Table 2). Examination of the standardized partial regression coefficients in Table 2 shows that, as expected, gopher activity is significantly and negatively affected by rockiness; flowers are significantly affected negatively by gopher activity, positively by seedlings, and positively by rockiness. The positive effect of rockiness on flowers has both direct and indirect components. There is also an indirect negative effect of moisture on flowers (coefficient = −0.275, not shown in Table 2), acting through rockiness, but this is highly confounded by the protective effect of rockiness on flowers. Seedlings are significantly and positively affected by flowers. In the model with moisture removed, rockiness has a significantly negative effect, although rockiness also has a positive indirect effect on seedlings acting through gopher activity and flowers.

In short, the path analysis generally accords with the hypothesis derived from examining pairwise in-

TABLE 2. Standardized partial regression coefficients, plus associated probabilities, from a path analysis (Fig. 8) on data from the 16 × 16 grid at Washington Gulch.

Origin of path	Total covariance (ordinary correlation)	P (total covariance = 0)	Direct causal covariance	P (direct causal covariance = 0)	Indirect causal covariance	Total direct and indirect causal covariance	Noncausal covariance
a) Direct path toward gopher activity from rockiness ($R^2 = 0.328$, $P > 0.00001$)							
Rockiness	−0.569	0.00000	−0.569	0.00000	0.000	−0.569	0.000
b) Direct paths toward seedlings from gopher activity, flowers, and rockiness ($R^2 = 0.066$, $P > 0.00018$)							
Gophers	0.093	0.135	0.000		−0.082	−0.082	0.175
Flowers	0.133	0.034	0.329	0.00007	0.000	0.329	−0.196
Rockiness	−0.075	0.23107	−0.295	0.00036	0.224	−0.070	−0.005
c) Direct paths toward flowers from gopher activity, seedlings, and rockiness ($R^2 = 0.520$, $P > 0.00000$)							
Gophers	−0.538	0.00000	−0.249	0.00000	0.000	−0.249	−0.289
Seedlings	0.133	0.034	0.197	0.00001	0.000	0.197	−0.064
Rockiness	0.667	0.00000	0.540	0.00000	0.084	0.624	0.043

teractions. Gophers negatively affect adult plants but not seedlings, in agreement with the sink hypothesis. The relationship between seedlings and adults, which appears neutral by correlation analysis and negative by ceiling analysis, becomes positive when important intervening variables are accounted for. Unfortunately, it is hard to say much about the independent role of moisture.

Discussion

The rock-refuge hypothesis

First, we elaborate on the hypothesis that prompted our close examination of the flower and seedling distributions and that structured our path analysis. The subalpine meadow habitats of *Erythronium grandiflorum* comprise a mosaic of dry rock outcrops and moist deep-soil pockets. Fossorial pocket gophers are excluded from the former areas but are very active in the latter. Corms of *E. grandiflorum* are quite palatable to gophers in feeding trials (M. Soulé and R. Smith, *personal communication*) and were a principal item found in gopher caches in Utah (Aldous 1951). Consequently, seedlings that become established in the soil pockets are likely to be discovered and killed before they reach flowering size. Flowering plants are thus more common in the enemy-free space of rocky areas. Therefore, most of the seeds fall in such areas, and as far as we can tell, stay where they fall (Weiblen and Thomson 1995). However, something makes those seeds unlikely to germinate in situ. Therefore, most of the seedlings come from the relatively small subset of seeds that are either (1) dispersed off the outcrops, or (2) produced by the relatively few plants that do manage to flower in deep soil areas. We do not know why seeds germinate better in the deep-soil pockets. The greater moisture may be directly responsible, or the greater soil perturbation caused by gopher activity, or other factors. Of course, seeds on the outcrops must succeed occasionally, perhaps in wetter years, but it may take very little recruitment to maintain populations of this long-lived perennial in protected sites where it is invulnerable to a voracious predator.

This hypothesis describes a source–sink relationship, but one that has the special feature of involving different life history stages. The paradoxical spatial disjunction between seed production and seed success, which at first seems puzzling and necessarily unstable, may actually reflect an equilibrial situation: the outcrops are sources, the pockets sinks.

The argument above presupposes a negative relationship or at least a negative ceiling between seedling abundance and flowering. Recall that the logistic-slice display actually suggested a humped distribution. This makes sense, and it does not fundamentally conflict with our argument. There are, rather, two ways in which low numbers of seedlings can come about. First, seeds may be present but unlikely to germinate, which is what we think happens on rock outcrops with high densities of flowers. Second, seeds may not be present, which is what must happen when there are no flowers in the vicinity. This could happen in a pocket where gopher pressure is extremely high, or on an outcrop where flowers are rare for any other reason, including chance. Our primary argument concerns the first situation only. The path analysis confirms that seedling and flower abundances correlate positively when the other factors are held constant.

Cantor and Whitham (1989) proposed that rock outcrops similarly act as gopher refuges for aspen clones (*Populus tremuloides*) in northern Arizona. In field experiments, they found that aspen survival and growth were increased by removing *Thomomys bottae.* These findings lend plausibility to our hypothesis, which is of course founded on correlational patterns only. Experiments (see Ellison and Aldous 1952) will be necessary to confirm a causative role for gopher predation in any particular case. However, we can adduce some additional evidence to address our contention that the restriction of *E. grandiflorum* to drier, rocky soils does not reflect its physiological optimum. First, plants grown at Irwin in a gopher-free transplant garden, and watered during dry spells, grow very vigorously (J. D. Thomson, *unpublished data*). Second, some of the largest and most vigorous plants in our study plots are actually in the deepest soil areas. When seeking corms to produce potted study plants, we have pursued this class of plants extensively because their large size gives a greater probability of flowering and the deep soil promises easier digging. As mentioned earlier, however, their corms almost always turn out to be wedged between rocks not visible from the soil surface. Apparently, deep soil habitats are good for *E. grandiflorum* growth if the plant finds itself in a microsite that protects it from gophers.

There are many studies of the effects of gophers on vegetation (Huntly and Inouye 1988). Many of these concern annual grasslands and emphasize the role of soil disturbance in creating new sites for seedling recruitment (e.g., Hobbs and Mooney 1985). Such effects may also be important in communities dominated by perennials, but direct attacks on perennials, especially tap-rooted species, also become important (Davis et al. 1991*a, b*). Burrows themselves change the physical characteristics of soil in the rooting zone (e.g., Reichman and Smith 1985, Reichman et al. 1993). In general, it is hard to separate observed gopher effects into direct effects of plant damage and indirect effects of mound making and burrowing. Our interpretation of direct predation on corms is reasonable, based on the known palatability of *Erythronium* to gophers, but we cannot exclude indirect effects.

Demography and reproductive success of lilies

If our hypothesis is correct, and if the patch structure of rockiness that we found in our study area is typical for *E. grandiflorum* habitats in general, unusual selection pressures should be operating. Most of the flowering, and therefore the seed production, takes place in rocky patches surrounded by inimical habitat. Seeds that are dispersed more than a few metres will, on average, land in habitat less suitable than those that stay closer to home. Natural selection might therefore lead to the loss of characters that increase dispersal (e.g., elaiosomes), and the reinforcement of characters that increase survival in rocky, desiccating microsites (e.g., seed size, seed coat impermeability). There is a substantial literature on the absence of dispersal structures (atelechory) and the development of characteristics that prevent seed dispersal (antitelechory) in desert plants (Zohary 1937, review by Ellner and Shmida 1981). Zohary's mother-site hypothesis suggested that "continued occupation of a few favorable sites in an otherwise hostile territory is often the best evolutionary strategy" (Stebbins 1971). Ellner and Shmida (1981) reject this hypothesis, in part because they find many open microsites and rapid species turnover in the deserts studied by Zohary. However, it may apply to *Erythronium grandiflorum* in subalpine meadows, where neither rocks nor lilies show much turnover in ecological time. If mother sites are selectively advantageous in our lily populations, it is not because of a global lack of gopher-free habitat, but because of the spatial scale (grain size) at which such habitat is distributed. Suppose that some individuals of *E. grandiflorum* bore elaiosomes and consequently dispersed their seeds 10 m instead of <1. Those seeds would be more likely to land in deep-soil pockets, where their expected fitness would be low.

The ecological patchiness of seed success should also affect functional gender (Lloyd and Bawa 1984) in *E. grandiflorum.* Pollen dispersal has been quantified in this species, and it is considerably more extensive than seed dispersal (Thomson and Thomson 1989, Weiblen and Thomson 1995). If our hypothesis is correct, and successful establishment is very rare in the deep-soil areas, the plants growing there may be effectively female sterile despite their substantial seed production. However, they should occasionally export some pollen to recipients in rockier areas, thus achieving some male reproductive success and therefore being male biased in functional gender.

Exploratory data analysis for ecological data

Standard regression approaches for summarizing and testing bivariate scattergrams are based on viewing one variable as a mathematical function of another: a given value of x leads us to look for values of y clustered around a particular predicted value. In ecology, we more often are interested in the less restrictive question of whether one variable might be influencing another. Our concept of limiting factors suggests that a common form of such influence might be that x may impose on y not a particular value but an upper bound (see Kaiser et al. 1994). Then, a value of x leads us to look for values of y that occupy a whole range of locations below some predicted ceiling. The upper edge of the scattergram is delineated by points from localities where x might be acting as a limiting factor for y; the interior of the point cloud shows where other factors, including history and chance, intervene.

Communicating the results of descriptive ecological studies would be simpler if we had a commonly accepted terminology and statistical machinery for treating such distributions. We have neither. Maller's term *triangular distribution* is probably known to few ecologists, and is not really general enough to cover humped distributions. Kaiser et al. (1994) avoid coining a new term, and their discussion of the problem is narrowly focused on limiting resources and triangular distributions. Blackburn et al. (1992) also focus on one specific type of relationship with a clear theoretical expectation. We offer the term factor-ceiling distribution not from a conviction that it is the best possible term, but from a strong feeling that *some* general term is needed.

Our proposals for partitioned regression and logistic slicing as ways to dissect point clouds are similarly tentative. We present them not as a complete solution, but to stimulate research into the development of better approaches (and to provide alternatives to the techniques of Maller (1990) and Kaiser et al. (1994), which also deserve much wider attention and testing). We hope that they exemplify one of the tenets of EDA, ". . . the value of graphs in forcing the unexpected on the reader" (Mosteller and Tukey 1977). Tentative though they are, they did help us to see and describe patterns in the data. In particular, before doing the logistic slice analysis, we had been thinking of the seedlings vs. flowers relationship as simply a triangular, factor-ceiling distribution in which the maximum number of seedlings per quadrat declined with the number of flowering plants. The sliced display suggests that the distribution is humped instead, which is ecologically more reasonable and interesting. That is, we have so far emphasized the right-hand portion of Fig. 7, where we hypothesize that abundant flowering characterizes rocky microsites that are inimical to seed germination, even though those sites have a plentiful supply of seeds. However, the extreme left-hand portion points out an alternative, "supply-side" reason why seedlings might be rare in particular quadrats: when flowers are extremely rare, the number of seedlings may be limited by the supply of seeds. Therefore, seedling density is highest at an intermediate position along a gradient of flowering intensity. Here, looking for the edge of a scatterplot produced a new hypothesis. There is no reason to suppose that logistic slicing is an optimal method, however; indeed, it clearly loses its ability to pinpoint an edge where data become sparse and the slices become wide. The results also depend heavily on the number of pixels created, which is arbitrary. Sliding-window algorithms may be worth investigating as an alternative to fixed slices.

In addition to descriptive techniques, we desire significance tests to assess the reliability of apparent patterns in the data. P. Wilson's (*unpublished manuscript*) approach provides one route to a test, although it too contains an arbitrary element: results depend on which extreme or subextreme data points are used to anchor the line. An alternative to consider would be the two-dimensional Kolmogorov–Smirnov statistic proposed by Fasano and Franceschini (1987; see also Press et al. 1992). The Kolmogorov–Smirnov statistic can test whether two bivariate point clouds have the same location, or whether one cloud differs from a theoretical expectation. It should be a more robust indicator of overall differences between observed and randomized distributions than Wilson's statistic, in that it does not hinge on a particular arrangement of the data the way Wilson's does. It can also measure the probability that an observed lopsided point cloud (e.g., Fig. 1b) could have been produced by chance from a symmetrical generating process. There is a great opportunity here for developing new techniques or discovering established but unfamiliar ones.

Factor-ceiling EDA vs. path analysis

Why engage in esoteric manipulations and displays of data if path analysis is capable of reaching similar conclusions? First, because the success of path analysis is a special case that depends on the critical variables having been measured. In ecological situations, numerous factors can intervene, and it will often be impossible to account for them (Kaiser et al. 1994). Second, although path analysis can, in a sense, uncover ceiling relationships in favorable cases, its use may draw one's attention away from the direct examination of bivariate scattergrams, with a consequent loss of insight. Both approaches are worth exploring.

Spatial scale of patchiness and correlation

This study raised some other methodological issues that deserve comment. Essentially, we documented spatial patchiness in several variables, then sought to use common ecological principles and some special knowledge of the system to erect a causative hypothesis to account for the correlations among them. The basic decision to be made is whether a correlation (say, the negative association of flowering and soil moisture) is to be taken at face value as potentially causative, or whether it is a spurious correlation driven by some other spatial process. We hypothesize that this correlation is indeed spurious and that the apparent "preference" of *Erythronium grandiflorum* for drier sites has nothing to do with dryness per se but is driven by patchiness in gopher predation, which is itself driven by the rock pattern. Are there any general rules for deciding about spurious correlation?

We used two sampling approaches, the contiguous grid of quadrats at Washington Gulch and the scattered quadrats at Irwin. We argue that the former approach is more useful for refining causative hypotheses because the spatial scale of the patchiness can also be assessed. Our data are poorly suited to illustrate this contention because all of our variables showed patchiness at similar scales. Imagine, however, a large study area in which gopher activity is uniformly high in the southern half and totally absent in the northern half. Throughout the whole area, patches of dry soil are scattered through a moist-soil matrix; these patches are more common in the northern half. Suppose that one throws quadrats randomly throughout the area and finds that flowering is negatively correlated with both moisture and gophers. Two completely different patterns could produce this result: (1) flowering might be confined to the northern half of the site; or (2) flowering might be confined to dry patches throughout the site. In the former case, we should entertain the possibility that the apparent moisture effect might actually be a side effect of gopher activity. In the latter case, this would be far less likely because the scales are discordant. We thus propose that examining the spatial scale of correlated processes can help identify variables that might be driving spurious correlations. For one variable C to be the driving force between two other variables A and B, all of them should act at the same spatial scale.

Explicitly considering spatial pattern as a modulator of ecological interactions poses difficult challenges, both in data collection, which becomes more arduous, and in analysis, which becomes more arcane. We risk bogging down in unfamiliar statistics. In compensation, however, we are able to address familiar issues with more precision, and our attention may be drawn to less familiar possibilities. For example, we hypothesize that the lack of an elaiosome in *Erythronium grandiflorum* may be adaptive, not to any particular microhabitat in the environment, nor to global characteristics of the environment, but to the spatial *grain* of the environment. Such hypotheses, which we hope are of broad interest, arise only from a spatial perspective.

Acknowledgments

P. Dixon, P. Dutilleul, M.-J. Fortin, C. Janson, M. Kaiser, G. LeBuhn, R. A. Maller, P. Neal, J. Patton, O. J. Reichman, F. J. Rohlf, R. Smith, M. Soulé, D. E. Wartenberg, T. Whitham, and P. Wilson provided useful comments or information. Field work was supported by NSF grant BSR 9006380 to J. D. Thomson and by an NSF REU grant to RMBL that supported S. Alfaro. Analysis and writing were supported by an NSF Mid-Career Fellowship to J. D. Thomson. This is Contribution 955 in Ecology and Evolution from SUNY–Stony Brook.

Literature Cited

Aldous, C. M. 1951. The feeding habits of pocket gophers (*Thomomys talpoides moorei*) in the high mountain ranges of central Utah. Journal of Mammalogy **32**:84–87.

Armstrong, D. M. 1972. Distribution of mammals in Colorado. Monograph of Museum of Natural History, University of Kansas, Number **3**.

Blackburn, T. M., J. H. Lawton, and J. N. Perry. 1992. A method of estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. Oikos **65**:107–112.

Cantor, L. F., and T. G. Whitham. 1989. Importance of belowground herbivory: pocket gophers may limit aspen to rock outcrop refugia. Ecology **70**:962–970.

Clifford, P., S. Richardson, and D. Hémon. 1989. Assessing the significance of the correlation between two spatial processes. Biometrics **45**:123–145.

Davis, M. A., J. Villinski, K. Banks, J. Buckman-Fifield, J. Dicus, and S. Hofmann. 1991*a*. Combined effects of fire, mound-building by pocket gophers, root loss and plant size on growth and reproduction in *Penstemon grandiflorus*. American Midland Naturalist **125**:150–161.

Davis, M. A., J. Villinski, S. McAndrew, H. Scholtz, and E. Young. 1991*b*. Survivorship of *Penstemon grandiflorus* in an oak woodland: combined effects of fire and pocket gophers. Oecologia **86**:113–118.

Dutilleul, P. 1993. Modifying the *t* test for assessing the correlation between two spatial processes. Biometrics **49**: 305–314.

Eastman, J. R. 1993. IDRISI technical reference, version 4.0. Clark University School of Geography, Worcester, Massachusetts, USA.

Ellison, L., and C. M. Aldous. 1952. Influence of pocket gophers on vegetation of subalpine grassland in central Utah. Ecology **33**:177–186.

Ellner, S., and A. Shmida. 1981. Why are adaptations for long-range dispersal rare in desert plants? Oecologia **51**: 133–144.

Farë, R., S. Grosskopf, and C. A. K. Lovell. 1994. Production frontiers. Cambridge University Press, New York, New York, USA.

Fasano, G., and A. Franceschini. 1987. A multidimensional version of the Kolmogorov–Smirnov test. Monthly Notices of the Royal Astronomical Society **225**:155–170.

Hacker, S. D., and M. D. Bertness. 1995. A herbivore paradox: why salt marsh aphids live on poor-quality plants. American Naturalist **145**:192–210.

Handel, S. N., S. B. Fisch, and G. E. Schatz. 1981. Ants disperse a majority of herbs in the mesic forest community in New York State. Bulletin of the Torrey Botanical Club **108**:430–437.

Hobbs, R. J., and H. A. Mooney. 1985. Community and population dynamics of serpentine grassland annuals in relation to gopher disturbance. Oecologia **67**:342–351.

Huffaker, C. B. 1957. Fundamentals of biological control of weeds. Hilgardia **27**:101–167.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology **22**:415–427.

Kaiser, M. S., P. L. Speckman, and J. R. Jones. 1994. Statistical models for limiting nutrient relationships in inland waters. Journal of the American Statistical Association **89**: 410–423.

Krebs, C. J. 1978. Ecology. Harper and Row, New York, New York, USA.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology **74**:1659–1673.

Lloyd, D. G., and K. S. Bawa. 1984. Modification of the gender of seed plants in varying conditions. Evolutionary Biology **17**:225–338.

Maller, R. A. 1990. Some aspects of a mixture model for estimating the boundary of a set of data. Journal du Conseil International pour l'Exploration de la Mer **46**:140–147.

Maller, R. A., E. S. de Boer, L. M. Joll, D. A. Anderson, and P. J. Hinde. 1983. Determination of the maximum foregut volume of western rock lobsters (*Panulirus cygnus*) from field data. Biometrics **39**:543–551.

Mosteller, F., and J. M. Tukey. 1977. Data analysis and regression. Addison-Wesley, Reading, Massachusetts, USA.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical recipes in FORTRAN. Cambridge University Press, New York, New York, USA.

Pulliam, H. R. 1988. Sources, sinks, and population regulation. American Naturalist **132**:652–661.

Pulliam, H. R., and B. J. Danielson. 1991. Sources, sinks, and habitat selection: a landscape perspective on population dynamics. American Naturalist **137**:S50–S66.

Rabinowitz, D., J. K. Rapp, P. M. Dixon, and A. T. Khieu. 1985. Separating structural and developmental variability in growth rate estimates for *Andropogon scoparius* Michx. Bulletin of the Torrey Botanical Club **112**:403–408.

Reichman, O. J., J. H. Benedix, Jr., and T. R. Seastedt. 1993. Distinct animal-generated edge effects in a tallgrass prairie community. Ecology **74**:1281–1285.

Reichman, O. J., and S. C. Smith. 1985. Impact of pocket gophers burrows on overlying vegetation. Journal of Mammalogy **66**:720–725.

Robertson, A. 1906. The "droppers" of *Tulipa* and *Erythronium*. Annals of Botany **20**:429–440.

Ruhren, S., and M. R. Dudash. 1996. Manipulation of seed release of *Erythronium americanum* (Liliaceae), a temporally restricted eastern deciduous forest myrmecochore. American Journal of Botany, *in press*.

Shmida, A., and S. Ellner. 1984. Coexistence of plant species with similar niches. Vegetatio **58**:29–55.

Sokal, R. R., and N. L. Oden. 1978*a*. Spatial autocorrelation in biology. 1. Methodology. Biological Journal of the Linnean Society **10**:199–228.

Sokal, R. R., and N. L. Oden. 1978*b*. Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. Biological Journal of the Linnean Society **10**:229–249.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W. H. Freeman, San Francisco, California, USA.

Stebbins, G. L. 1971. Processes of organic evolution. Prentice Hall, Englewood Cliffs, New Jersey, USA.

Thomson, J. D., and B. A. Thomson. 1989. Dispersal of *Erythronium grandiflorum* pollen by bumble bees: implications for gene flow and reproductive success. Evolution **43**:657–661.

Tukey, J. W. 1977. Exploratory data analysis. Addison-Wesley, Reading, Massachusetts, USA.

Wartenberg, D. E. 1994. SAAP spatial autocorrelation program, version 4.3. Exeter Software, Setauket, New York, USA.

Weiblen, G. D., and J. D. Thomson. 1995. Seed dispersal in *Erythronium grandiflorum* (Liliaceae). Oecologia **102**: 211–219.

Wilkinson, L. 1990. SYGRAPH: the system for graphics. Systat, Evanston, Illinois, USA.

Zohary, M. 1937. Die verbreitungsokologischen Verhaltnisse der Pflanzen Palaestinas. Beihefte zum Botanische Centralblatt **A56**:1–155.