



Concepts of Scale

Instructor: K. McGarigal

Assigned Reading: Turner et al. 2001 (Chapter 2); Wiens (1989)

Objective: Provide a basic understanding of concepts related to scale to serve as a foundation for understanding landscape ecology topics. Clarify commonly misused terms and concepts. Highlight importance of considering scale in resource management planning and analyses.

Topics covered:

1. Scale in ecology
2. Components of scale
3. Why is scale important?
4. Characteristic scale
5. Scaling techniques
6. Dealing with scale

Comments: Some material taken from Wiens (1989), Turner et al. (2001), Rosenberg (2001) and Dean Urban's Landscape Ecology course notes, Duke University.

Scale in Ecology

Acts in what Hutchinson (1965) has called the “ecological theatre” are played out on various scales of space and time.

“The problem of pattern and scale is the central problem in ecology, unifying population biology and ecosystems science, and marrying basic and applied ecology. Applied challenges...require the interfacing of phenomena that occur on very different scales of space and time, and ecological organization. Furthermore, there is no single natural scale at which ecological phenomena should be studied; systems generally show characteristic variability on a range of spatial, temporal, and organizational scales” (Levin 1992).

1. Scale in Ecology

Scientists nearly universally recognize the central role that scale plays in determining the outcome of observations (Levin 1992, Schneider 1994, Peterson and Parker 1998). Acts in what Hutchinson (1965) has called the “ecological theatre” are played out on various scales of space and time. To understand these dramas, we must view them on the appropriate scale.

"The problem of pattern and scale is the central problem in ecology, unifying population biology and ecosystems science, and marrying basic and applied ecology. Applied challenges ... require the interfacing of phenomena that occur on very different scales of space, time, and ecological organization. Furthermore, there is no single natural scale at which ecological phenomena should be studied; systems generally show characteristic variability on a range of spatial, temporal, and organizational scales." (Levin 1992)

Scale in Ecology

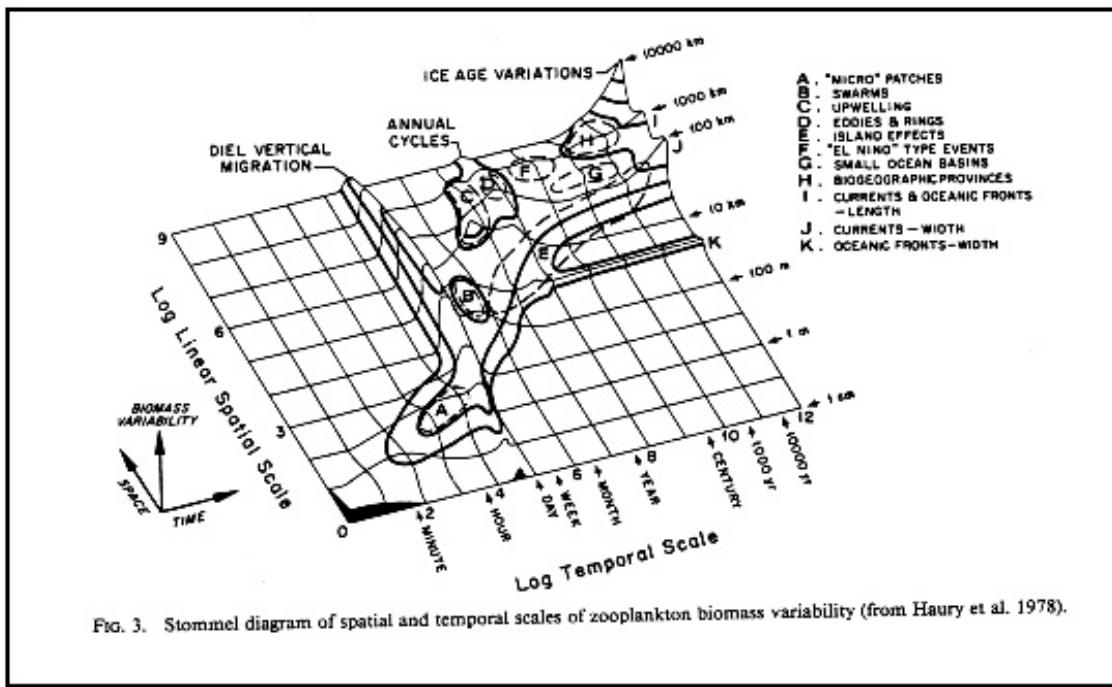
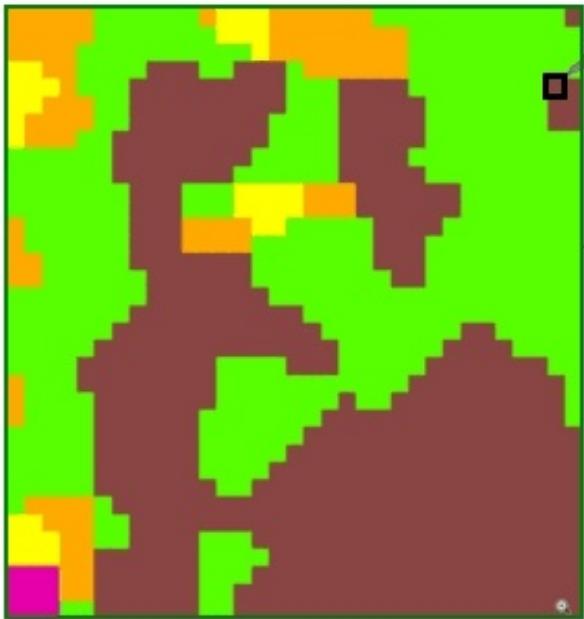


FIG. 3. Stommel diagram of spatial and temporal scales of zooplankton biomass variability (from Haury et al. 1978).

These ideas are perhaps best illustrated by “space-time” diagrams, which portray the hierarchical structure in nature and the positive correlation in spatial and temporal scales of varying processes. The most widely cited example is given by Delcourt et al. (1983) to illustrate environmental disturbance regimes, biotic responses, and vegetation patterns in the context of space-time domains, in which the scale for each process or pattern reflects the sampling intervals required to observe it (see textbook for this example). Shown here is a different example involving the spatial and temporal scaling of variation in zooplankton biomass. Note the positive correlation between spatial and temporal scales of variability. For example, “micro” patches of zooplankton exist at the spatial scale of 1's to 10's of meters, and these patches fluctuate over a period of hours. Upwellings, on the other hand, extend over 100's of kilometers and fluctuate seasonally over the course of a year. Thus, as the spatial extent of the phenomena increases, so too does the temporal extent over which the dynamics can be observed.

Components of Scale



GRAIN: The minimum resolution of the data, defined by the cell or minimum polygon size.

EXTENT: The scope or domain of the data, defined as the size of the landscape or study area under consideration.

Minimum Patch Size?

2. Components of Scale

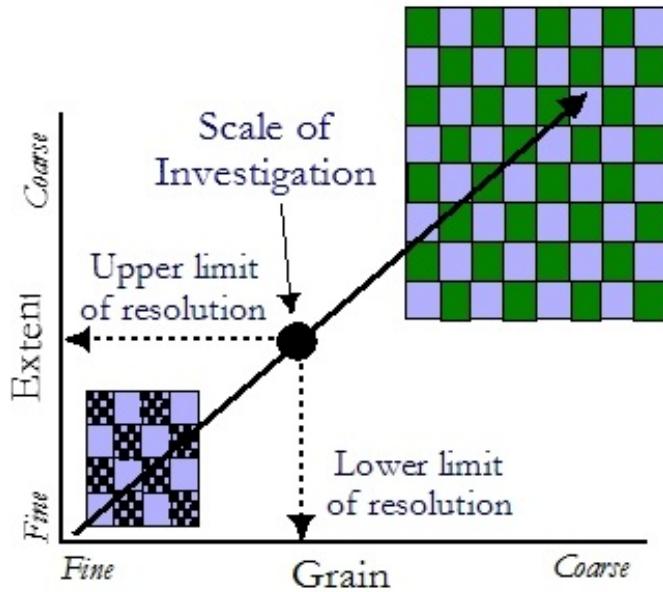
There are several components of scale, but the two most important for our purposes are “grain” and “extent”:

Grain The minimum spatial resolution of the data—in raster lattice data, the cell size; in field sample data, the quadrat size; in imagery, the pixel size; in vector GIS data, the minimum mapping unit (polygon). Grain is the size of the individual units of observation; i.e., the smallest entities that can be distinguished. For example, a fine-grained map might structure information into 1-ha units, whereas a map with an order of magnitude coarser resolution would have information structured into 10-ha units (Turner et al. 1989).

Extent The scope or domain of the data--defined as the size of the study area or landscape, typically. Extent is simply the spatial domain over which the system is studied and for which data are available; i.e., the overall area encompassed by an investigation or the area included within the landscape boundary. From a statistical perspective, the spatial extent of an investigation is the area defining the population we wish to sample.

Components of Scale

- Grain and extent are correlated.
- Information content is often correlated with grain.
- Grain and extent set lower and upper limits of resolution in the data.

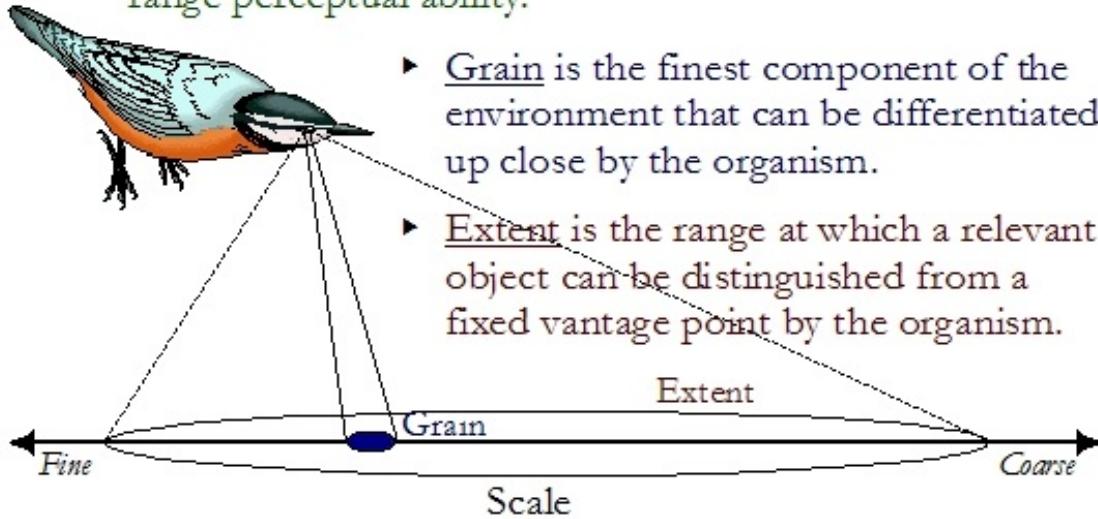


There are some general rules governing the relationship between grain and extent:

- Grain and extent are correlated, a result of logistical constraints in measurement; i.e., as the grain increases the logical extent increases as well.
- Information content is often correlated with grain; we tend to measure more variables in fine-grained studies.
- Grain and extent set the lower and upper limits of resolution in the data; i.e., we cannot detect patterns at finer or coarser scales than the grain and extent of the data, respectively, and any inferences about scale-dependency in a system are constrained by the extent and grain of investigation.

Components of Scale

- From an “organism-centered” perspective, grain and extent may be defined as the degree of acuity of a stationary organism with respect to short- and long-range perceptual ability.



There are some important considerations in selecting the grain and extent for any application:

- Grain and extent can be defined in a functionally relevant manner (i.e., relative scale). For example, from an organism-centered perspective, grain and extent may be defined as the degree of acuity of a stationary organism with respect to short- and long-range perceptual ability.
 - Grain is the finest component of the environment that can be differentiated up close by the organism.
 - Extent is the range at which a relevant object can be distinguished from a fixed vantage point by the organism.

Components of Scale



- From an “anthropogenic perspective”, grain and extent may be defined on the basis of management objectives.
 - ▶ Grain is the finest unit of management (e.g., stand).
 - ▶ Extent is the total area under management consideration (e.g., forest).



- Alternatively, from an anthropogenic perspective, grain and extent may be defined on the basis of management objectives.
 - ▶ Grain is the finest unit of management (e.g., stand).
 - ▶ Extent is the total area under management consideration (e.g., forest).

Components of Scale

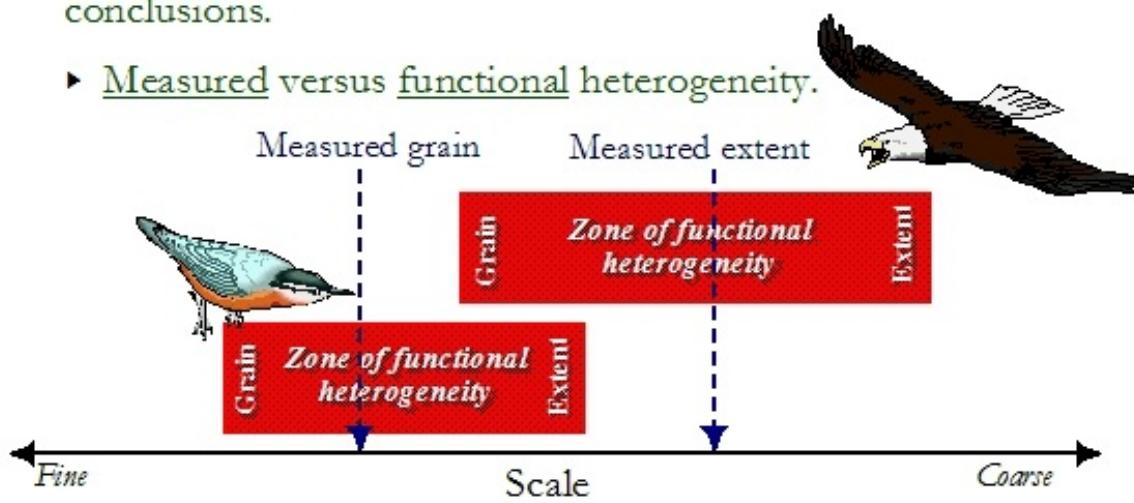
- In practice, grain and extent are often dictated by the scale of the imagery (e.g., photo scale) being used or the technical capabilities of the computing environment.



- In practice, grain and extent are often dictated by the scale of the imagery (e.g., aerial photo scale, Landsat resolution) being used or the technical capabilities of the computing environment. Or, in many applications, the grain and extent have already been established, for example when using the national forest GIS database where the minimum vegetation polygon size and the extent of the forest are set, and the investigator or manager must work within these constraints regardless of the application.

Components of Scale

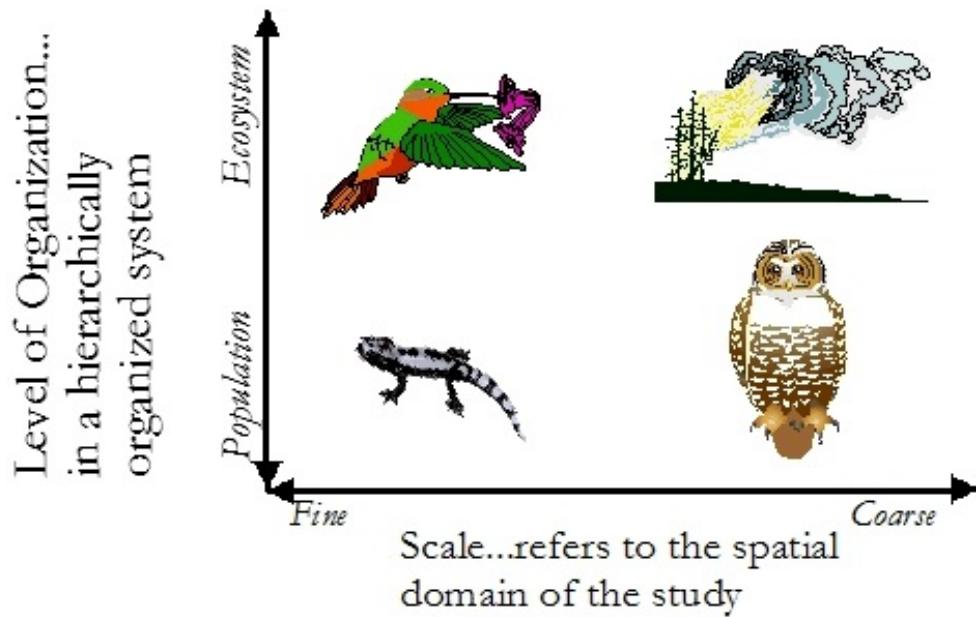
- It is critical that grain and extent be defined for a particular study and represent the ecological phenomenon or organism under study, otherwise the patterns detected will have little meaning and there is a good chance of reaching erroneous conclusions.
 - ▶ Measured versus functional heterogeneity.



- It is critical that grain and extent be defined for a particular study and represent, to the greatest possible degree, the ecological phenomenon or organism under study, otherwise the patterns detected will have little meaning and there is a good chance of reaching erroneous conclusions. For example, it would be meaningless to define grain as 1-ha units when the organism under consideration perceives and responds to habitat patches at a resolution of 1-m². A strong landscape pattern at the 1-ha resolution may have no significance to the organism under study. Similarly, it would be meaningless to define the landscape extent as 1-km² when the organism under consideration has a home range size several times that size. Typically, however, we do not know what the appropriate resolution should be. In this case, it is much safer to choose a finer grain than is believed to be important because the grain sets the minimum resolution of investigation. Once set, we can always resample to a coarser grain. In addition, we can always specify a minimum mapping unit that is coarser than the grain. That is, we can specify the minimum patch size to be represented in a landscape, and this can easily be manipulated above the grain size. Indeed, it may be useful to reanalyze the same landscape using progressively coarser minimum patch sizes to better assess landscape heterogeneity across a range of potentially relevant scales. Thompson and McGarigal (2002) used this approach successfully to define the “best” scale (grain and extent) for representing bald eagle habitat along the Hudson River, New York (discussed below).

Clarification of Scale Terms

- ‘Scale’ is not the same as ‘level of organization’.

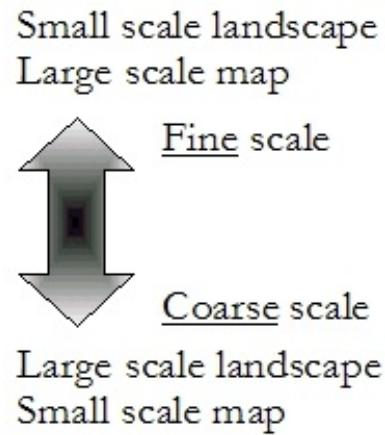


There are some important scale terminology issues to clarify:

- ‘Scale’ is not the same as ‘level of organization.’ Scale refers to the spatial domain of the study, while level of organization depends on the criteria used to define the system. For example, population-level studies are concerned with interactions among conspecific individuals, while ecosystem-level studies are concerned with interactions among biotic and abiotic components of some process such as nutrient cycling. One could conduct either a small- or large-scale study of either population- or ecosystem-level phenomena.

Clarification of Scale Terms

- ‘Ecological scale’ and ‘map scale’ are exactly opposite!
 - ▶ Ecological Scale = size (extent) of the landscape under consideration.
 - ▶ Map Scale = ratio of map to real distance.



- ‘Ecological scale’ and ‘map scale’ are exact opposites; in cartography, ‘scale’ is a ratio of map to real distance, so a large-scale map is fine-grained and of small extent.

Clarification of Scale Terms

- ‘Absolute scale’ versus ‘relative scale’

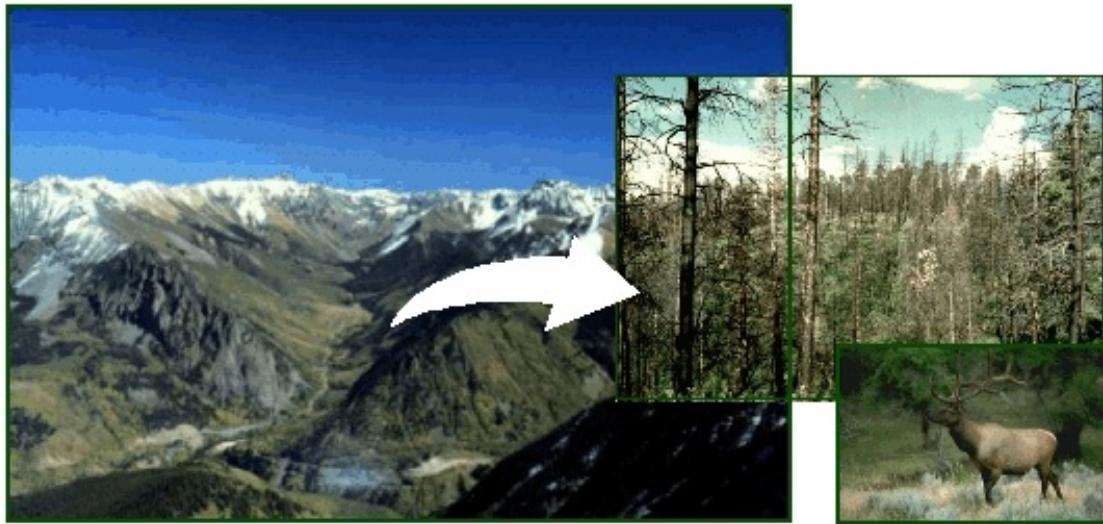
- ▶ Absolute Scale = actual distance, direction, shape, and geometry.
- ▶ Relative Scale = relative distance, direction, etc., based on some functional relationship to organism or process under consideration.



- ‘Scale’ can be defined in both absolute and relative terms. ‘Absolute scale’ represents the actual distance, direction, shape, and geometry of the pattern under investigation; ‘relative scale’ represents a transformation of absolute scale to a scale that describes the relative distance, direction, shape, or geometry based on some functional relationship to the organism or process under investigation.

Why is Scale Important?

- As one changes scale, controls on pattern and process change – e.g., local biological interactions can decouple systems from direct physical determination of patterns.



3. Why is scale important?

Scale is a critical consideration in all landscape ecological studies for several reasons:

- As one changes scale, controls on pattern and process change.—Local biological interactions can decouple systems from direct physical determination of patterns. Local biological interactions (e.g., competition and other processes) have the effect of decoupling systems from direct physical determination of patterns by introducing temporal and spatial lags in system dynamics or creating webs of indirect effects. However, at broader scales, physical processes may dominate or dissipate these biological effects.

For example, the relationship between climate and vegetation that are evident at broad scales, may disappear at finer scales, overridden by the effects of competition and other biological processes.

Krummel et al. (1987). . . .

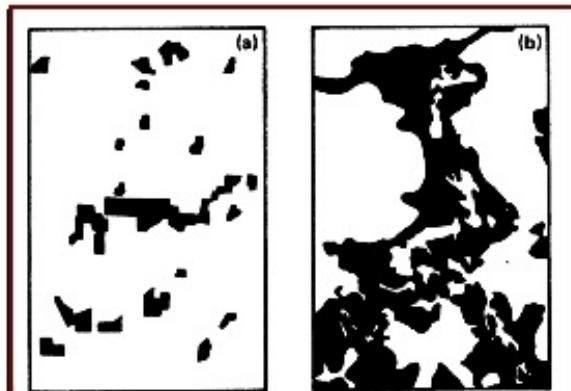


Fig. 2. Computer plots (both at same scale) of deciduous forest in two different areas of the Natchez Quadrangle that illustrate the differences in shape complexity between small forest patches in the Mississippi flood-plain (a) and a larger forested area on the eastern edge of the floodplain (b).

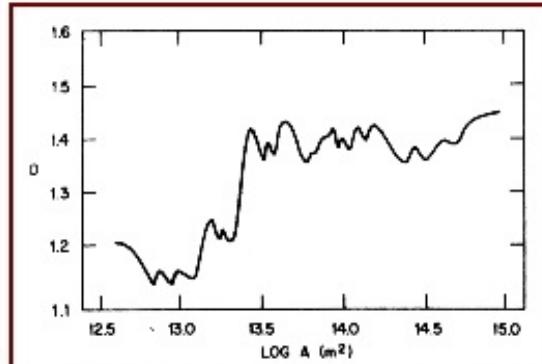
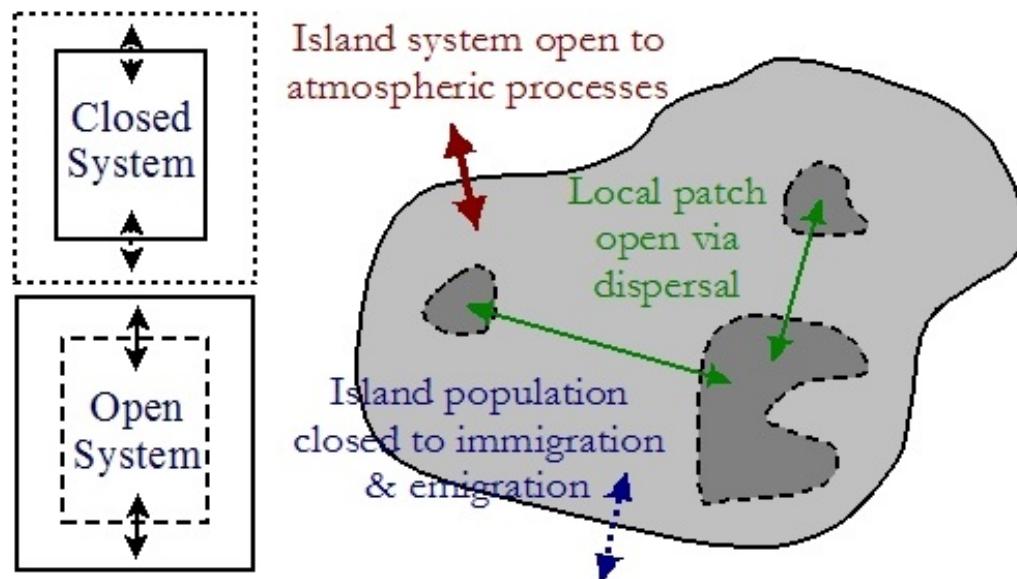


Fig. 1. Changes in fractal dimension (D) values as the log of area (A) increases; as determined by successive regressions of the log of perimeter (P) on log of A.

In this example, the major agent of landscape pattern formation in Mississippi flood-plain forests differs between the coarse-grained and fine-grained elements (Krummel et al. 1987). Specifically, the geometric complexity of the deciduous forest patches, as measured by the fractal dimension index, increases dramatically as forest patch size increases. Moreover, the shift in geometric complexity exhibits threshold-like behavior. The authors interpret this as a shift between natural flood-plain disturbances that maintain a complex geometric mosaic of forest in parts of the study area and anthropogenic disturbance processes associated with farmland development that leave geometrically simple remnant forest stands in other parts.

Why is Scale Important?

- As one changes scale, systems may switch between “closed” and “open.”

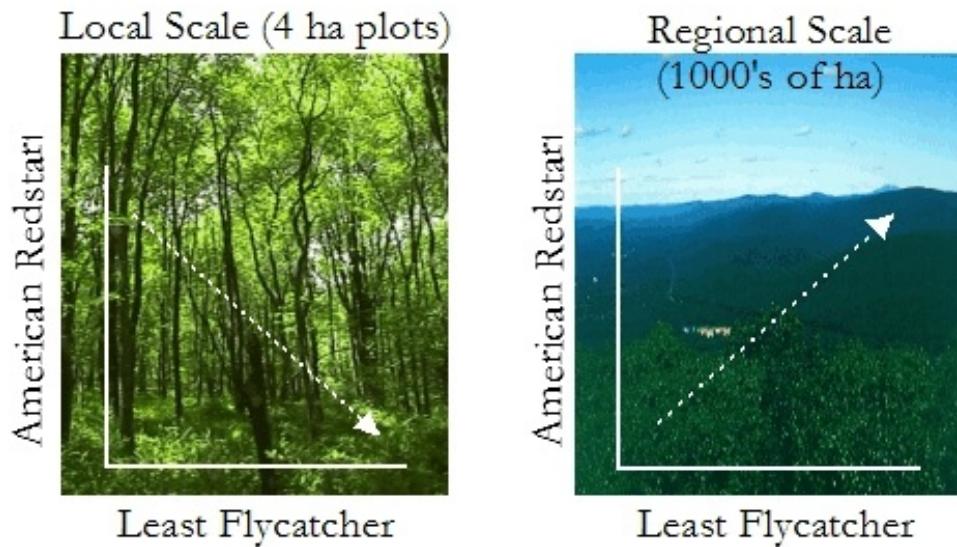


- As one changes scale, systems may change between “closed” and “open”.—Ecosystems do not exist in isolation; they are typically “open” with respect to the movement of energy, materials, and organisms into and out of the system. In open systems, transfer rates among elements are relatively high, and the dynamics of patterns at a given scale are influenced by factors at broader scales. In open systems, ecosystem context matters a lot. However, ecological systems can become “closed” when transfer rates among adjacent systems approach zero or when the differences in process rates between adjacent elements are so large that the dynamics of the elements are effectively decoupled from one another. In closed systems, ecosystem context doesn’t matter as much. Importantly, as one changes the scale, the system may switch between “closed” and “open”.

For example, a local habitat patch on an island may be an open system to dispersal, immigration and emigration, wherein the local population is demographically linked to the other habitat patches on the island. At the scale of the island, however, the population may be relatively closed to immigration and emigration, wherein the population can be treated as an autonomous (i.e., closed) demographic unit. On the other, the island is open to atmospheric processes. Thus, the issue of system openness depends on the process under consideration.

Why is Scale Important?

- As one changes scale, statistical relationship may change.

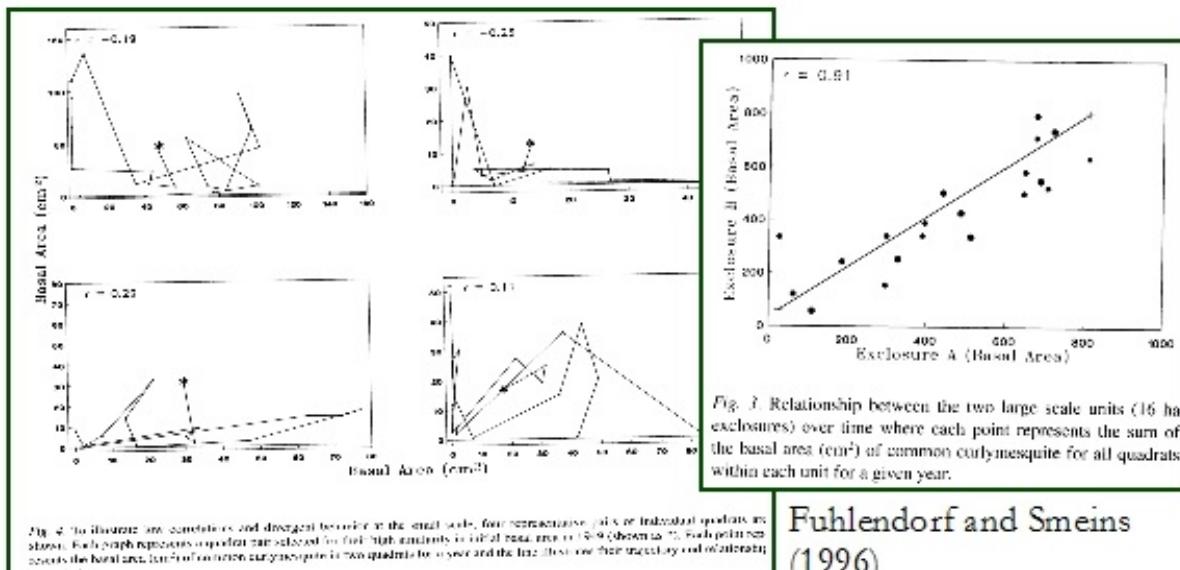


- As one changes scale, statistical relationships may change.—Different patterns emerge at different scales of investigation of virtually any aspect of any ecological system. The magnitude or sign of correlations may change, important variables may change, and variance relationships may change.

For example, in hardwood forests of the northeastern United States, least flycatchers negatively influence the distribution of American redstart territories at the scale of 4-ha plots. Regionally, however, these species are positively associated. Apparently, the broad-scale influences of habitat selection override the local effects of interspecific competition. Similar scale-dependency has been found in the habitat relationships of many other species.

Why is Scale Important?

- As one changes scale, statistical relationship may change.

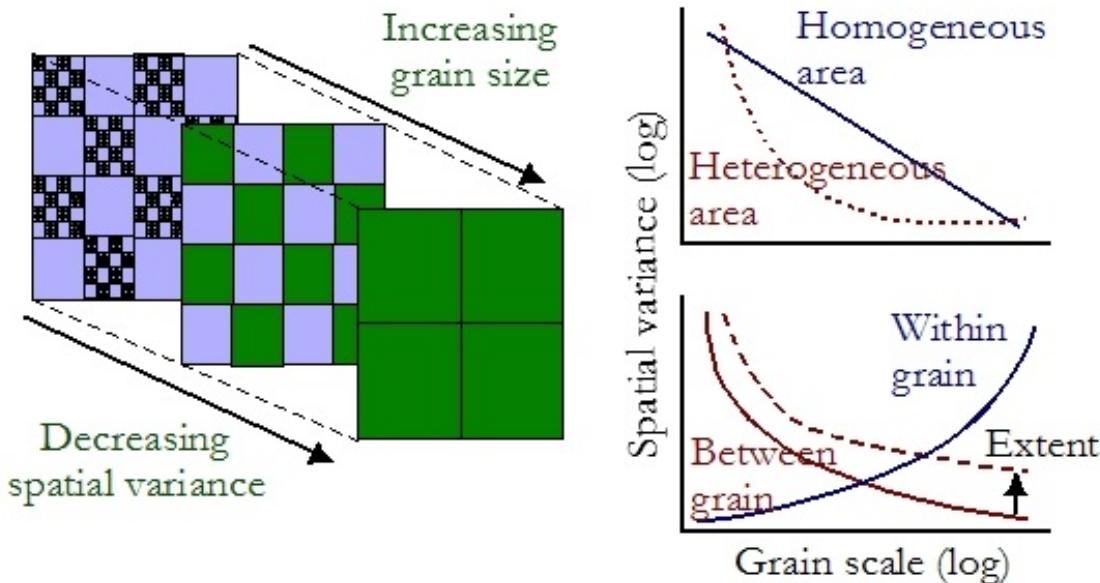


Fuhlendorf and Smeins
(1996)

In this example, the relationship between sample plots in the basal area of common curly mesquite varies with spatial scale (Fuhlendorf and Smeins 1996). At the meter-quadrat scale, any two selected quadrats show no relationship at all in the change in basal area of mesquite over time; i.e., they fluctuate randomly with respect to each other. However, at the scale of 16-ha plots, there is a strong and positive relationship over time, indicating that while overall stand development is occurring similarly among the larger plots, there is considerable fine-scale variation over time in the distribution of mesquite stems – perhaps resulting from fine-scale disturbance processes subsumed at the level of the larger plot.

Why is Scale Important?

- As one changes scale, statistical relationships may change.



Similarly, when the scale of measurement of a variable is changed, the variance of that variable changes. Holding *extent* constant, an increase in the *grain* of measurement generally decreases spatial variance. As grain increases, a greater proportion of the spatial heterogeneity of the system is contained within a sample or grain and is lost to our resolution, while between-grain heterogeneity (= variance) decreases. In addition, holding *grain* constant, an increase in *extent* will typically incorporate greater spatial heterogeneity, as a greater variety of patch types or landscape elements is included within the area being studied. Similarly, an increase in temporal extent will typically incorporate greater temporal variability, as a greater range of conditions are encountered. These observations have important implications since we rely heavily on the concept of variance to characterize patterns and infer pattern-process relationships.

Why is Scale Important?

- As one changes scale, statistical relationships may change.

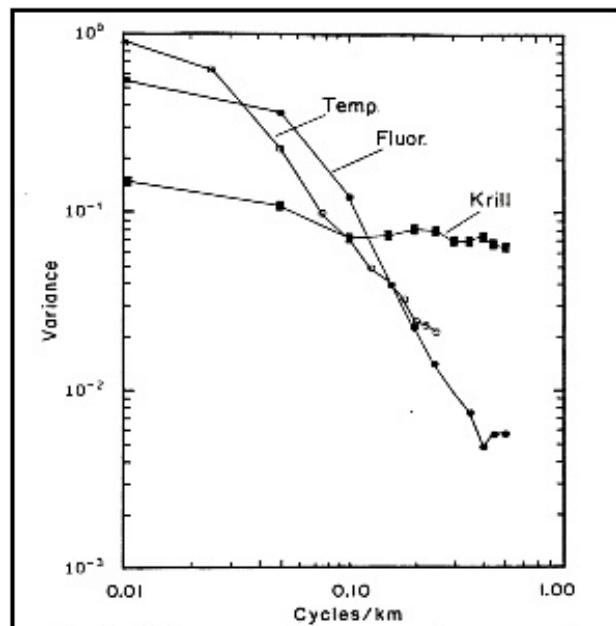
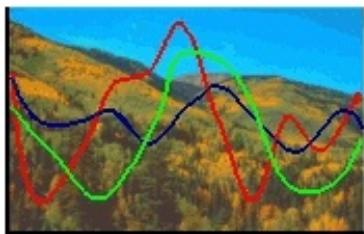
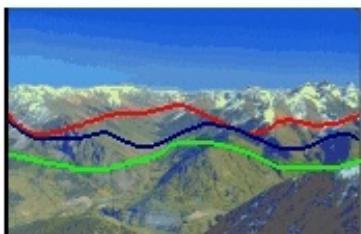


FIG. 7. Variance spectra for temperature, fluorescence, and krill in the Scotia Sea (from Weber et al. 1986).

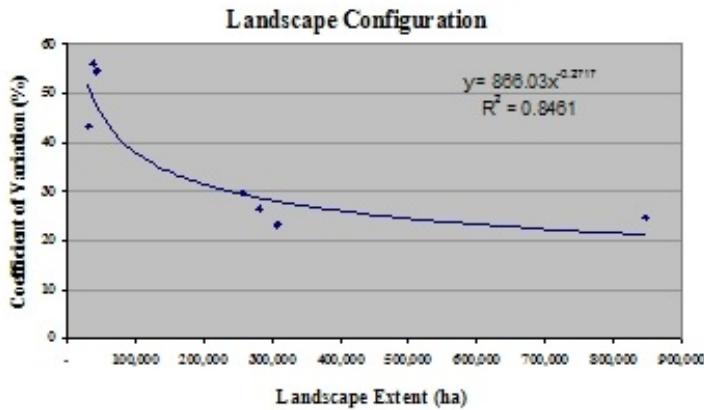
For example, variance in temperature, fluorescence, and krill along a transect in the Scotia Sea decreases as *grain size* increases (Weber et al. 1986). As Levin (1992) explains, the spatial distributions of krill populations of the Southern Ocean have been shown to be patchy on almost every scale of description. This has fundamental importance both for the dynamics of krill, and for their predator species. Various studies have characterized the Fourier spectrum of variability of krill, and shown that variance decreases with scale. However, substantial differences exist between the spectra for krill and those for temperature and fluorescence. On broad scales, temperature (a passive marker of phytoplankton activity), fluorescence (a measure of phytoplankton activity), and krill all have spectra that are consistent with predictions based on water movement. On fine scales, however, the krill spectrum is noticeably flatter, suggesting that krill are much more patchily distributed than their resource, or than can be explained by water movements alone. The interpretation of these data is that large patches of krill and phytoplankton are being moved about by water movements, but that on fine scales some other mechanism must be invoked to explain pattern. Thus, a two-scale model is needed, and a general lesson learned: no single mechanism explains pattern on all scales. In this case, pattern seems extrinsically driven on broad scales (i.e., water movement), and autonomously generated on fine scales.

Why is Scale Important?

- As one changes scale, statistical relationship may change.



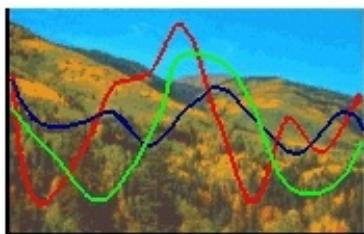
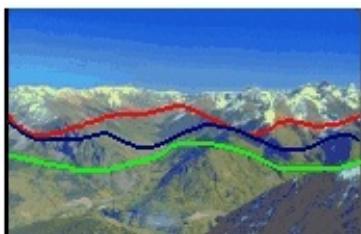
Range of variability decreases with increasing spatial extent.



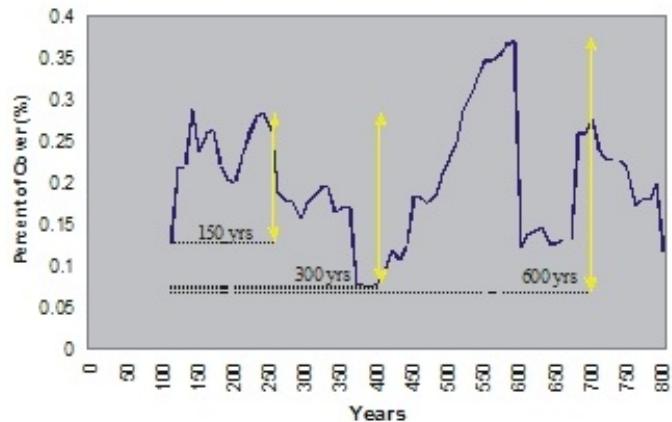
This example involves quantifying the historic range of variation (HRV) in landscape structure on the San Juan National Forest in southwest Colorado. First, with regards to *spatial extent*, when we examined progressively smaller spatial units of the entire simulated landscape, temporal variability in landscape structure increased – as would be expected, and there was an apparent threshold in the relationship. Specifically, the magnitude of variability in landscape structure increased only modestly as the landscape extent decreased from the forest scale (847,638 ha) to the district scale (average = 282,546 ha), but increased dramatically as the landscape extent decreased to the watershed scale (average = 38,469 ha). We interpreted this to mean that at the district extent (and larger), the landscape is large enough to fully incorporate the disturbance regime and exhibit stable dynamical behavior. An important conclusion of this study is that the statistical summary of HRV depends on the spatial extent of the landscape; i.e., that scale matters.

Why is Scale Important?

- As one changes scale, statistical relationship may change.



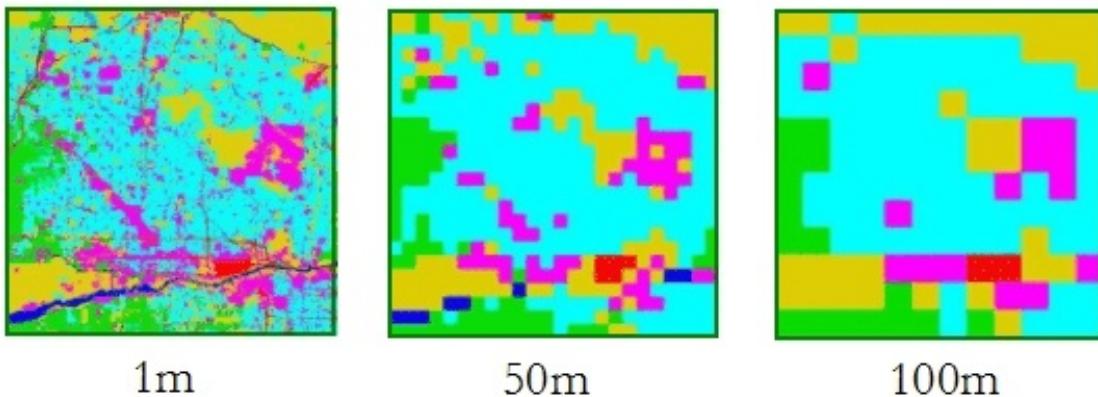
Range of variability increases with increasing temporal extent.



Second, with regards to *temporal extent*, not surprisingly the range of variability in landscape structure increased with increasing temporal extent. Note, we established the temporal extent of the simulations based on our desire to capture and describe a stable range of variation in landscape structure. In general, a minimum of 100-300 years was needed to capture the full range of variation in most metrics, and twice that long to confirm that the range was stable. Thus, a management strategy designed to emulate the natural disturbance regime would take 100-300 years to see the landscape fluctuate through its full range of conditions. This is a humbling thought given that most professional careers last no more than 30 years - a blip on the scale of these landscape dynamics - and that most policies are geared toward 10- to 20-year planning horizons. An important conclusion of this study is that the statistical summary of HRV depends on the temporal extent considered; i.e., that scale matters.

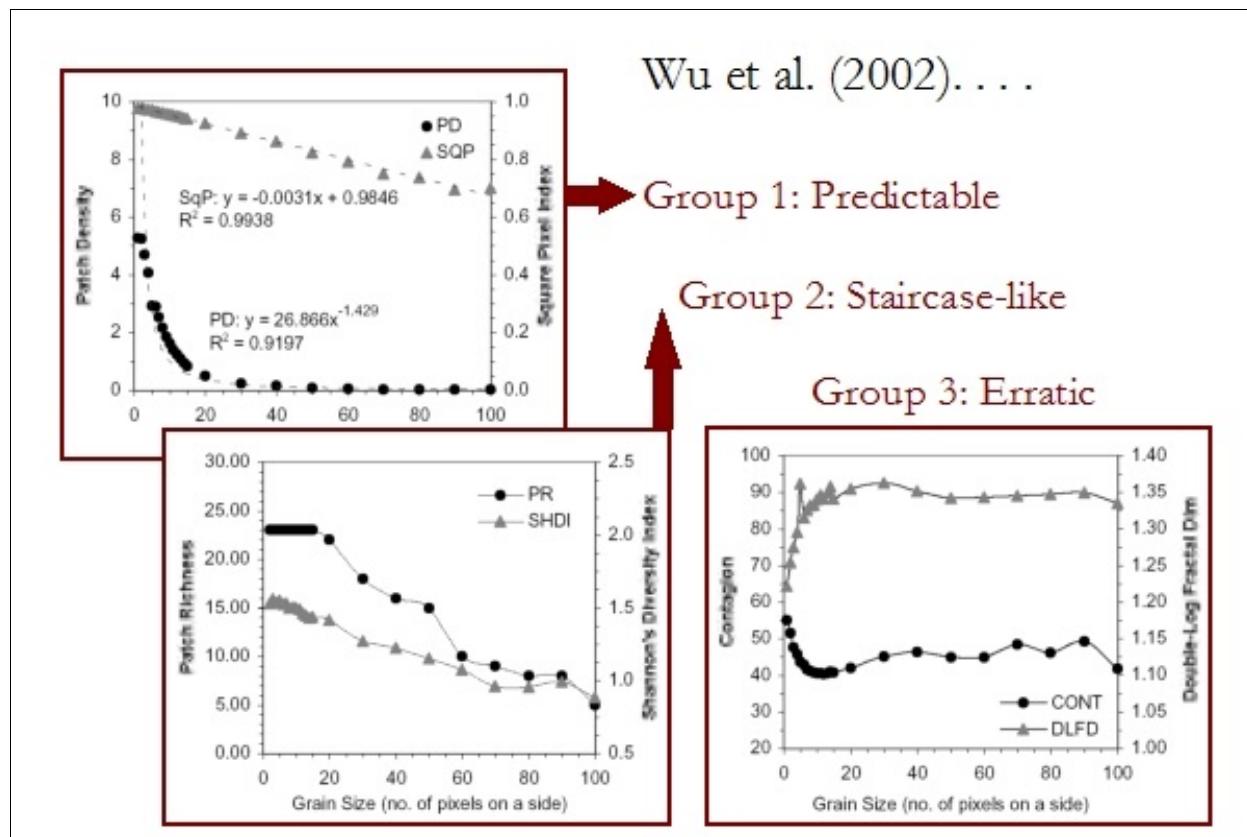
Why is Scale Important?

- As one changes scale, statistical relationship may change.



Wu et al. 2002. Empirical patterns of the effects of changing scale on landscape metrics. *Landscape Ecology* 17:761-782.

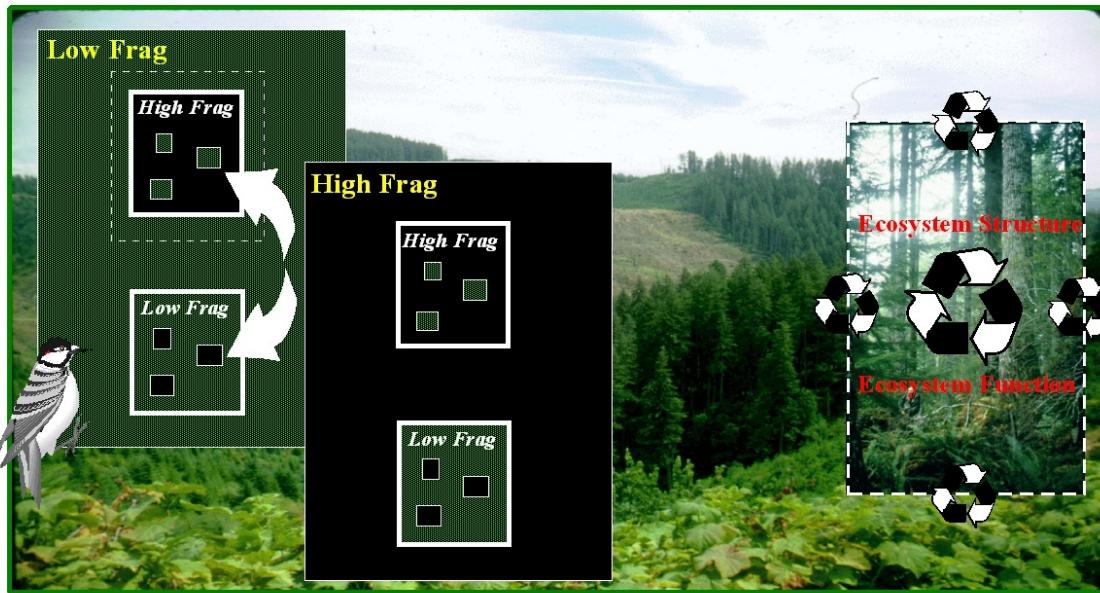
Let's consider one last example: the sensitivity of landscape metrics (indices of landscape pattern derived from categorical patch mosaics) to measurement scale. Consider the consequences of changing the grain (shown here) and extent (not shown) of a landscape mosaic on the measurement of landscape pattern. The most comprehensive assessment of this relationship to date is the study published in 2002 in *Landscape Ecology* by Jingle Wu and his colleagues. They explored in a systematic manner the effects of grain and extent on a suite of 19 common landscape metrics in real landscapes from four different geographic regions of North America.



In a nutshell, they found that metrics fell into three groups based on their scaling behavior. The first group consisted of metrics that had a predictable relationship with changing scale; changing either linearly or according to a power-law or logarithmic function, such as shown here for the patch density index and square pixel index. The second group exhibited staircase-like behavior, changing in a step-like fashion as the scale changed. The third group exhibited erratic and unpredictable behavior. They concluded that only metrics in the first group, those with predictable behavior, could be extrapolated or interpolated across scales. While this study and others have improved our understanding of landscape metric behavior in relation to scale, these findings highlight the need to pay particular attention to scale when quantifying landscape pattern-process relationships.

Why is Scale Important?

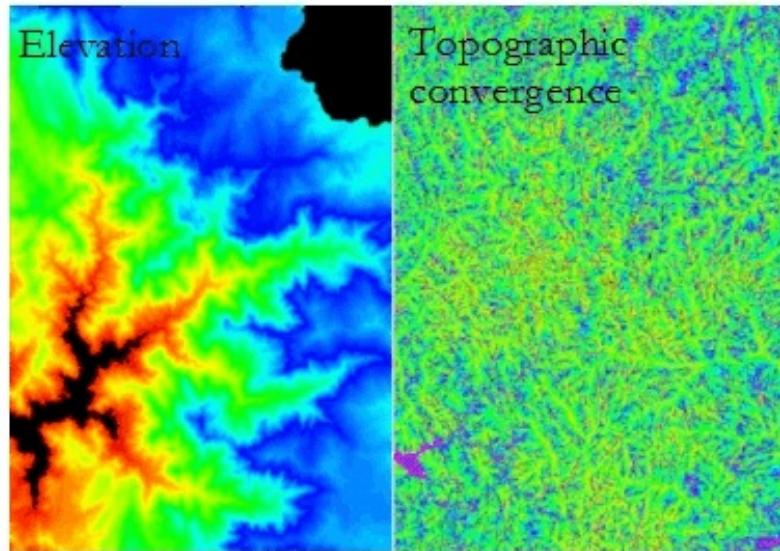
- As one changes scale (and context), pattern-process relationship may change.



- As one changes scale (and context), the pattern-process relationship may change.--Landscape context is especially important when measuring how pattern effects process. Landscape metrics quantify pattern within the designated landscape boundary only. Consequently, the interpretation of these metrics and their ecological significance requires an acute awareness of the landscape context and the openness of the landscape relative to the phenomenon under consideration. These concerns are particularly important for certain metrics. For example, nearest-neighbor distances are computed solely from patches contained within the landscape boundary. If the landscape extent is small relative to the scale of the organism or processes under consideration and the landscape is "open" relative to that organism or process, then nearest-neighbor results can be misleading. Consider a small subpopulation of a species occupying a patch near the boundary of a somewhat arbitrarily defined landscape. The nearest neighbor within the landscape boundary might be quite far away, yet the closest patch in reality might be very close, but just outside the landscape boundary. The magnitude of this problem is a function of scale. In general, the larger the ratio of extent to grain (i.e., the larger the landscape relative to the average patch size), the less likely these and other metrics will be dominated by boundary effects.

Characteristic Scale

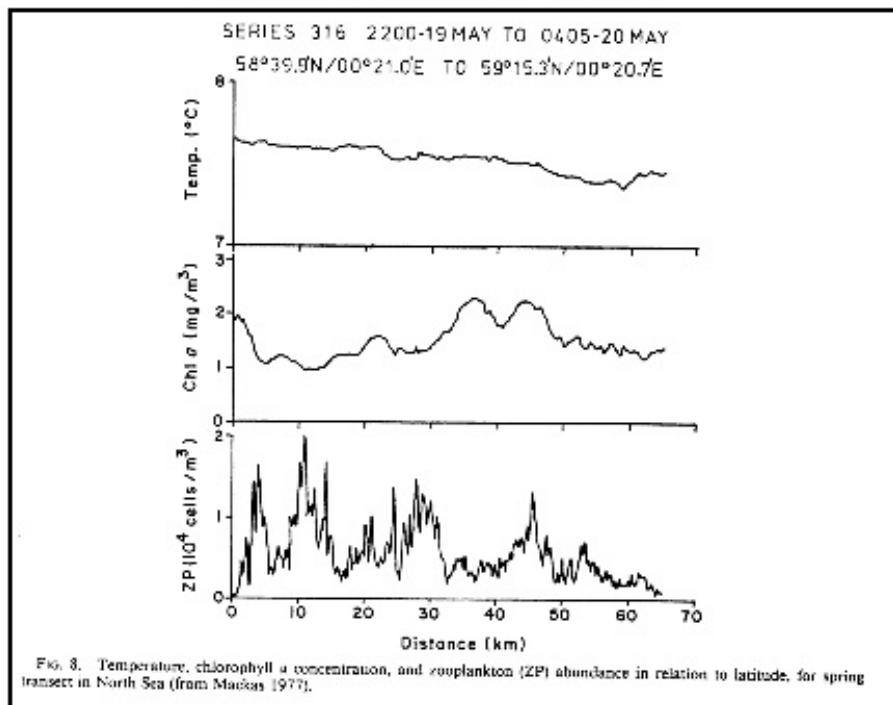
- Ecological phenomena have characteristic spatial and temporal scales, or spatiotemporal domains, and should be addressed at their characteristic scales.



4. Characteristic scale

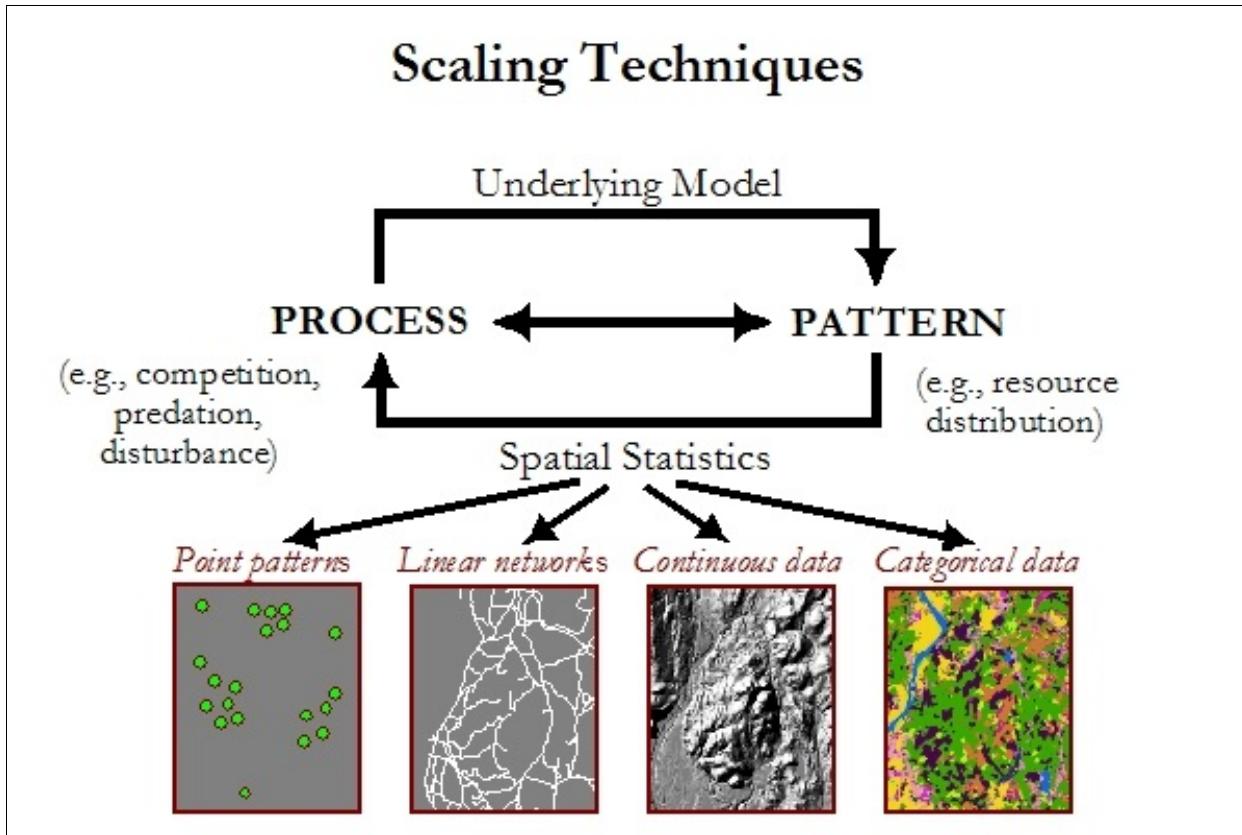
Many have argued that ecological phenomena tend to have characteristic spatial and temporal scales, or spatiotemporal domains (e.g., Delcourt et al. 1983, Urban et al. 1987). The “characteristic scale” is the scale at which the dominant pattern emerges; i.e., the scale with the maximum spatial variance among locations. Particular phenomena should be addressed at their characteristic scales. In the example shown here (from Urban), the characteristic scale of elevation appears to be much coarser than the characteristic scale of the topographic convergence index (TCI) for the same landscape. Elevation reflects the coarse-scale elevation gradient of this montane landscape, whereas TCI reflects fine-scale variation in landform associated with changes in slope and topography.

Characteristic Scale



In this example, the characteristic scale of temperature, chlorophyll *a* concentration, and zooplankton abundance are shown in relation to latitude for a one-dimensional spring transect in the North Sea (from Mackas 1977, as discussed in Levin 1992). Clearly, temperature shows little variation along the transect other than perhaps a gradient, reflecting the course-scale movement of water. Conversely, zooplankton show much more fine-scale variation than either phytoplankton (their food resource) or water temperature (presumably a driver of phytoplankton distribution and abundance). The different characteristic scales of variation of zooplankton, phytoplankton and water temperature indicate that different mechanisms are involved in determining their pattern.

Scaling Techniques



5. Scaling techniques

We have considered the utility of knowing the characteristic spatial scale of a system of interest, so that we could use this to isolate particular phenomena of interest. This is well and good for familiar systems where the characteristic scale is known. But what if you don't know? This overview is intended as a introduction and map of techniques that can be used to detect or define the scale of variation in a system (i.e., for a measured variable of interest).

There are several approaches to defining pattern reflecting the different classes or types of spatial data. There are four basic kinds of spatial data (points, linear networks, continuous surfaces, and categorical mosaics) and these can be defined in a largely intuitive way. These look rather different numerically, but they share a concern with the relative concentration of spatial variability. The examples we'll look at are designed for different kinds of data commonly encountered in landscape ecology. The goal here is not to delve too deeply into these, but rather, to look at the conceptual basis for these and moreover, to look at some ways of finding scaled pattern in field data. The goal: you'll appreciate what the techniques aim to do, and (intuitively) how they do it. And you'll be able to assess new ones as you encounter them in your readings.

5.2. Why Use Scaling Techniques?

Scaling techniques, or more generally, spatial statistics, can be useful to ecologists in a variety of ways:

- To describe spatial correlation and pattern in the data (i.e., to identify the scale of pattern). This is perhaps the primary use of spatial statistics in landscape ecology, where a major focus is to find the characteristic scale of pattern in the system in an effort to better understand the structure of the system, often as a precursor to experimental studies of pattern-process relationships.
- To choose an optimal experimental design. One of the key principles of experiment design is the independence of experimental units (i.e., the absence of any unaccounted for confounding sources of variation). This is necessary so that we can be certain (or reasonably confident) that the measured responses to the treatments is due exclusively to the treatments themselves. Understanding the scale of spatial dependencies in a system will allow investigators to choose experimental units or sampling units that are statistically independent.
- To model correlated (measurement) error. One of the key assumptions of most statistical tests is that the error (i.e., the unexplained variation in each sampling unit) is random; that is, that the error (or residuals) associated with each unit is uncorrelated. If there is spatial dependencies among the sampling units that is unaccounted for, then the error will in fact be correlated with the distance between units. Spatial statistics allow us to estimate the degree of correlation in our measurements due to spacial dependencies and then remove (or account for) this source of correlated error in the model.
- To interpolate data and construct contour maps. Often times we have sparsely sampled data and wish to predict the values for other unsampled locations. This is often the case with continuous surface data, where it is logically and practically difficult or perhaps impossible to exhaustively sample every location. This is the purview of geostatistics, where a major goal is to model the spatial dependencies in sampled continuous data and use this model to interpolate data for all unsampled locations. Often, the ultimate objective is to produce a contour map of the variable of interest.

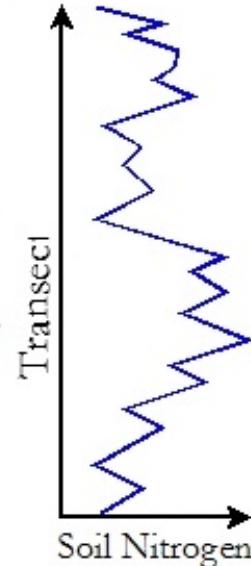
Why Use Scaling Techniques



- To describe spatial correlation and pattern in the data (i.e., identify the scale of pattern).
- To choose an optimal experimental design (e.g., spacing for independent sampling units).
- To model correlated (measurement) error.
- To interpolate data and construct contour maps.

Why Use Scaling Techniques

- What questions can scaling techniques address?
 - ▶ What is the dominant scale of Nitrogen pattern in the transect?
 - ▶ At what distance does soil Nitrogen become effectively uncorrelated?
 - ▶ How heterogeneous is the soil Nitrogen along the transect (i.e., patchiness)?
 - ▶ How large is spatial heterogeneity relative to error?
 - ▶ What does the soil Nitrogen distribution look like in space?
 - ▶ How steep is the soil Nitrogen gradient as you move across the study area?

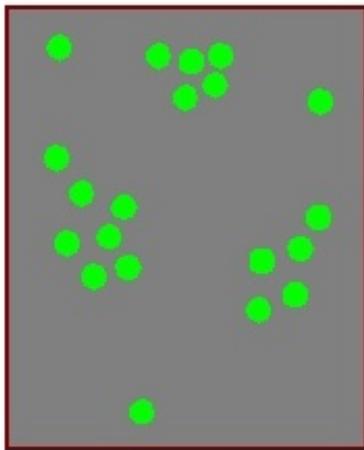


Given the many uses of spatial statistics, it should be of no surprise that there are myriad questions in ecology that require the use of spatial statistics. Consider, for example, how the spatial pattern of soil nitrogen influences and is in turn influenced by plants in a forested landscape. If we sampled soil nitrogen along a 1 km transect, we would expect to find substantial variation in nitrogen levels corresponding perhaps to variation in terrain features and vegetation. If we sampled nitrogen along a 100 m transect, would we expect to find less variation. Further still, if we sampled a 10 m transect, would we expect perhaps to find very little variation. These expectations motivate several important questions of scale:

- ▶ What is the dominant scale of nitrogen pattern (or variation) along this transect?
- ▶ Alternatively, at what distance does soil nitrogen become effectively uncorrelated?
- ▶ How heterogeneous is the soil nitrogen along the transect (i.e., is it patchily distributed or does it vary continuously??)
- ▶ How large is the spatial heterogeneity relative to the error (i.e., variation from the mean)?
- ▶ What does the soil nitrogen distribution look like in space if we were to produce a map of soil nitrogen?
- ▶ How steep is the soil nitrogen gradient as you move across the study area?

These and other questions can be addressed with spatial statistics, and answers to questions like these are an important step in understanding the structure and function of an ecological system.

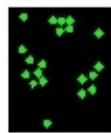
Scaling Techniques for Point Patterns



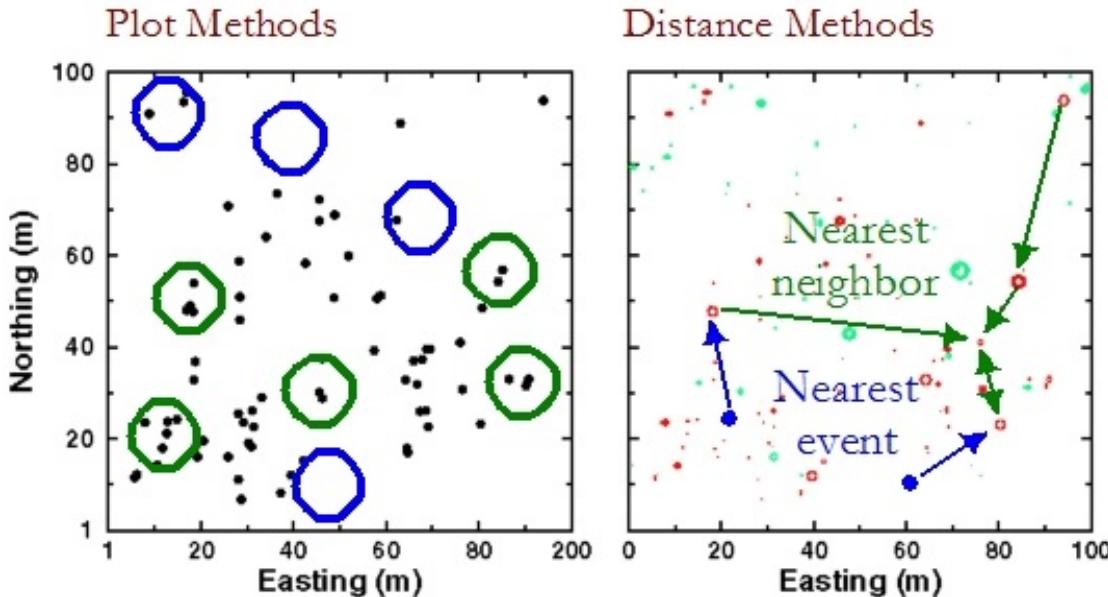
- Data Characteristics
 - ▶ Data comprise collections of the locations of entities of interest
- Examples
 - ▶ Map of trees in a forest stand
 - ▶ Map of vernal pools in a watershed
- Goal of Analysis
 - ▶ To find the spatial scale(s) at which the points tend to be more or less clustered than expected by chance

5.3. Scaling Techniques for Point Patterns

- Characteristics of spatial point patterns:
 - ▶ Point pattern data comprise collections of the locations of entities of interest.
 - ▶ A familiar example is a map of all trees in a forest stand, wherein the data consists of a list of trees referenced by their (x,y) locations. Typically, the points would be labeled by species, and perhaps further specified by their sizes (a marked point pattern).
 - ▶ The goal of pattern analysis with such data is to find the spatial scale(s) at which the points tend to be more or less clustered than expected by chance.



Scaling Techniques for Point Patterns



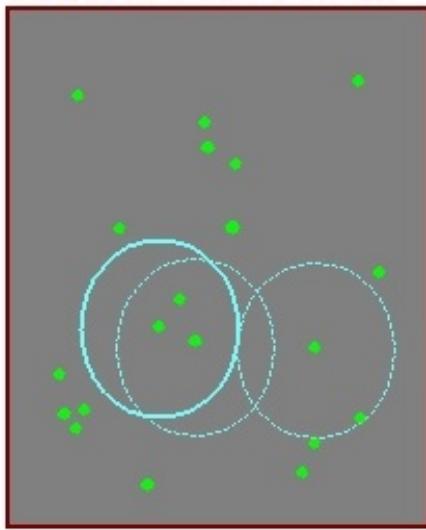
The primary data for point pattern analysis consist of n points tallied by location within an area of size A (e.g., hundreds of individual trees in a 1-ha stand). A plethora of techniques have been developed for analyzing spatial point patterns, some based on sample quadrats or plots and others based on nearest-neighbor distances. A typical distance-based approach is to use the mean point-to-point distance to derive a mean area per point, and then to invert this to get a mean point density (points per unit area), Λ , from which test statistics about expected point density are derived. There are nearly uncountable variations on this theme, ranging from quite simple to less so (e.g., Clark and Evans 1954). Most of these techniques provide a single global measure of point pattern aimed at distinguishing clumped and uniform distributions from random distributions, but do not help to distinguish the characteristic scale or scales of the point pattern.



Scaling Techniques for Point Patterns

■ Ripley's K-Distribution

Random Point Pattern



Expected number of points within circle of radius d_s from an arbitrary point:

$$E(d_s) = \frac{N}{A} K(d_s)$$

Where $K(d_s)$ is the area of a circle defined by radius d_s

Thus, under complete spatial randomness (csr):

$$E(d_s[\text{csr}]) = \frac{N}{A} \pi \cdot d_s^2$$

Ripley's K-distribution

The most popular means of analyzing (i.e., scaling) point patterns is the use of second-order statistics (statistics based on the co-occurrences of pairs of points). The most common technique is *Ripley's K-distribution* or K-function (Ripley 1976, 1977). The K-distribution is the cumulative frequency distribution of observations at a given point-to-point distance (or within a distance class); that is, it is based on the number of points tallied within a given distance or distance class. Because it preserves distances at multiple scales, Ripley's K can quantify the intensity of pattern at multiple scales.

Consider a *spatially random* distribution of N points. If circles of radius d_s , are drawn around each point, where s is the order of radii from the smallest to the largest, and the number of other points that are found within the circle are counted, and then summed over all points (allowing for duplication), then the expected number points within that radius $E(d_s)$ from an arbitrary point are:

$$E(d_s) = \frac{N}{A} K(d_s)$$

where N is the sample size, A is the total study area, and $K(d_s)$ is the area of circle defined by

radius d_s . For example, if the area defined by a particular radius is one-fourth the total study area and if there is a spatially random distribution, on average approximately one-fourth of the cases will fall within any one circle (plus or minus a sampling error). More formally, with *complete spatial randomness* (csr), the expected number of points within distance, d_s , is:

$$E(d_s[csr]) = \frac{N}{A} \pi \cdot d_s^2$$

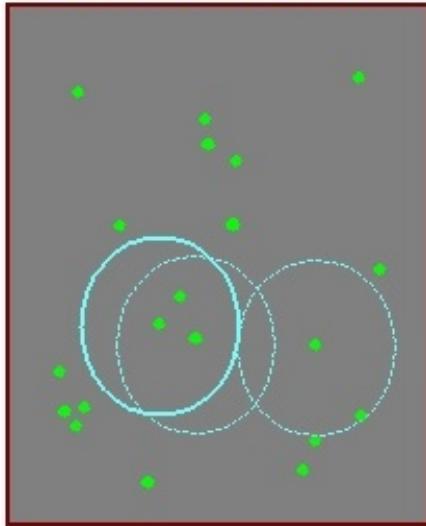
On the other hand, if the average number of points found within a circle for a particular radius placed over each point, in turn, is greater than that found by the equation above, this points to clustering for that radius. Conversely, if the average number of points found within a circle for a particular radius placed over each point, in turn, is less than that found by the equation above, this points to dispersion; that is, points are, on average, farther apart than would be expected on the basis of chance for that radius. By counting the total number of events within a particular radius and comparing it to the number expected on the basis of complete spatial randomness, the K-statistic is an indicator of non-randomness. The K-statistic applies to all distances up to the limit of the study area because the count is conducted over successively increasing radii. Indeed, this is the great utility of Ripley's K for investing point intensity over a range of spatial scales.



Scaling Techniques for Point Patterns

- Ripley's K-Distribution

Random Point Pattern



If $E(d_s) > E(d_s[\text{csr}])$ ---Clumped

If $E(d_s) < E(d_s[\text{csr}])$ ---Dispersed

$$E(d_s) = \frac{N}{A} K(d_s) \rightarrow K(d_s) = \frac{E(d_s)}{\lambda}$$

$$K(d_s) = \frac{1}{\lambda} \sum_{i=1}^n \sum_{j=1}^n \frac{\delta_{ij}(\|x_i - x_j\| \leq d_s)}{N}$$

Given the above, Ripley's K is defined as:

$$K(d_s) = \frac{E(d_s)}{\lambda}$$

where $E(d_s)$ is the expected number of points within a distance d_s from an arbitrary point. Again, the mean intensity Lambda is estimated simply as N/A , where N is total number of points and A is total area sampled. The cumulative distribution $E(d)$ is estimated empirically, and so in practice this requires completely surveyed data (not sparse samples of points). The typical estimate of $K(d)$ is tallied as:

$$K(d_s) = \frac{1}{\lambda} \sum_{i=1}^n \sum_{j=1}^n \frac{\delta_{ij}(\|x_i - x_j\| \leq d_s)}{N}$$

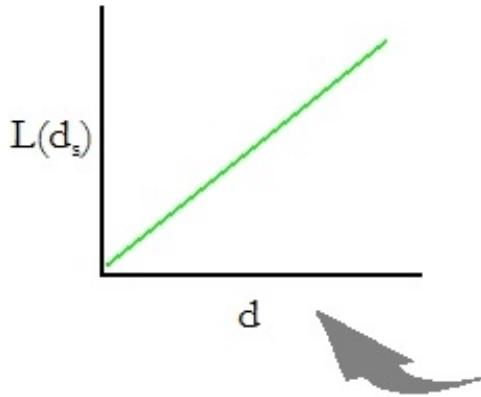
for $i \neq j$, where delta is an indicator function that takes on a value of 1 if the condition is true, else 0. Here, the condition is that the Euclidean distance between points is less than or equal to d . That is, $K(d)$ is a tally of the average number of points that fall within each distance class from an arbitrary point. The inverse of Lambda in the formula converts K to units of area.



Scaling Techniques for Point Patterns

- Ripley's K-Distribution

Random Point Pattern



Typical transformation

$$K(d_s) = \frac{1}{\lambda} \sum_{i=1}^n \sum_{j=1}^n \frac{\delta_{ij}(\|x_i - x_j\| \leq d_s)}{N}$$



$$L(d_s) = \sqrt{\frac{K(d_s)}{\pi}}$$

Under complete spatial randomness:

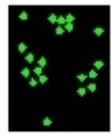
$$L(d_s) = \sqrt{\frac{\pi \cdot d_s^2}{\pi}} = d$$

Other variants of the equation can be used to correct for edge effects, in which cases delta takes on values less than 1.0 in the summation. Edge effects can seriously degrade distance-based statistics, and there are at least two ways to deal with these. One way is to invoke a buffer area around the study area, and to analyze only a smaller area nested within the buffer. By common convention, the analysis is restricted to distances of half the smallest dimension of the study area. This, of course, is expensive in terms of the data not used in the analysis. A second approach is to apply an edge correction to the indicator function for those points that fall near the edges of the study area; Ripley and others have suggested a variety of geometric corrections.

As noted above, for completely random data, the expected value of $K(d_s)$ is $\pi \cdot d_s^2$. For clarity of presentation, the K distribution is often transformed in one of two ways. The transformed distribution is defined:

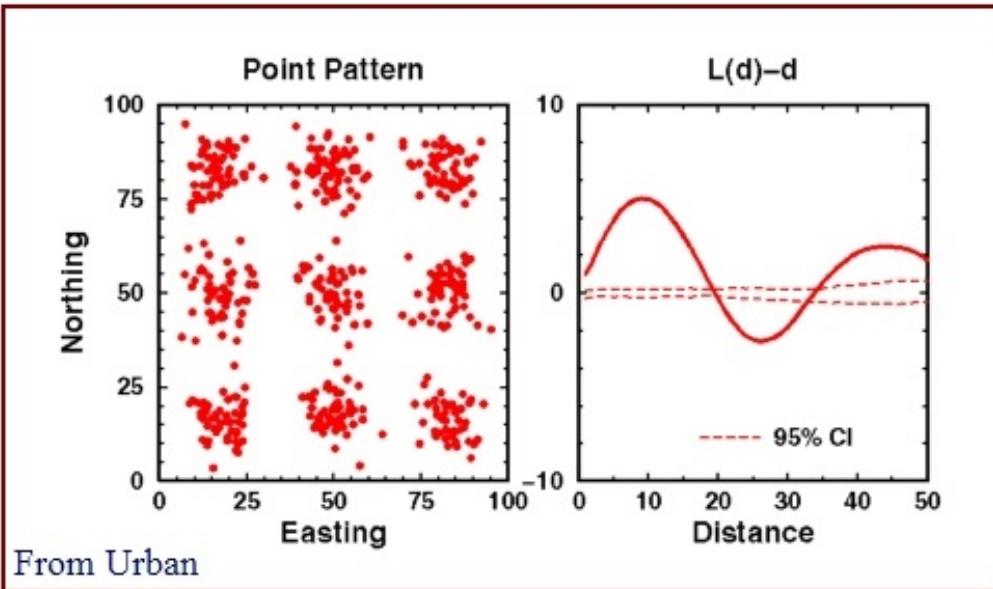
$$L(d_s) = \sqrt{\frac{K(d_s)}{\pi}}$$

which produces a plot of $L(d)$ against d where the expectation under randomness is a line with slope 1.0 (i.e., $L(d)=d$). Subtracting d from $L(d)$ transforms this expectation to a horizontal line under randomness. Which transformation to use is purely cosmetic.



Scaling Techniques for Point Patterns

- Ripley's K-Distribution



Ripley derived approximations of the test of significance for normal data. But data are often not normal, and assumptions about normality are particularly suspect under edge effects. So in practice, the K function is generated from the test data, and then these data are randomized to generate the test of significance as confidence limits. For example, if one permuted the data 99 times and saved the smallest and largest values of $L(d)$ for each d , these extrema would indicate the confidence limits at $\alpha=0.01$; that is, an observed value outside these limits would be a 1-in-a-hundred chance. Likewise, 19 randomizations would yield the 95% confidence limits [Note that these estimates are actually rather imprecise; simulations suggest that it might require 1,000-5,000 randomizations to yield precise estimates of the 95% confidence limits, and >10,000 randomizations to yield precise 99% limits.]

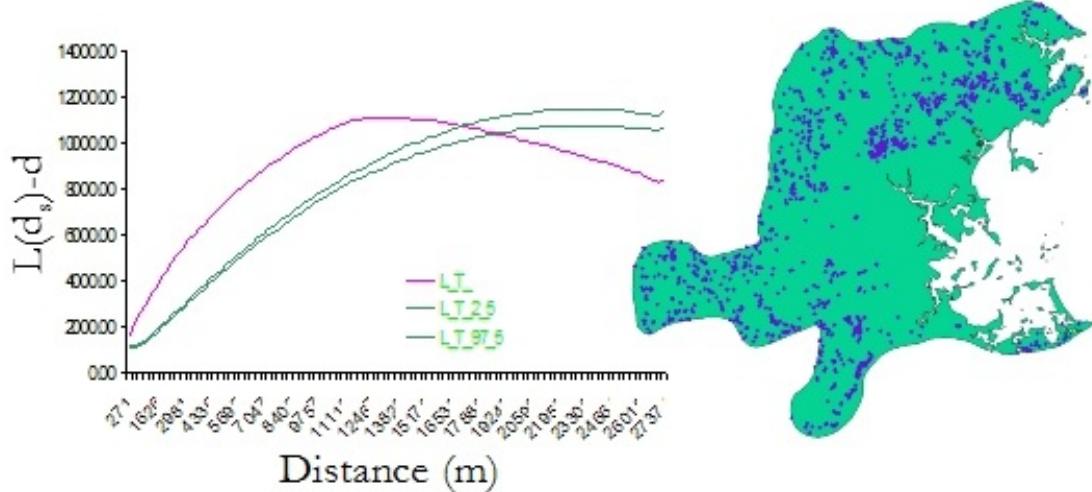
In the artificial data shown here, the K function reveals the pattern quite nicely. The first peak in the K distribution corresponds to the size of the point clusters (~ 10 m); the trough corresponds to the uniform spacing of the clusters (~ 25 m); and the second peak corresponds roughly to the size of two groups of point clusters (~ 50 m).



Scaling Techniques for Point Patterns

- Ripley's K-Distribution

Vernal pool patterns in a Massachusetts ecoregion



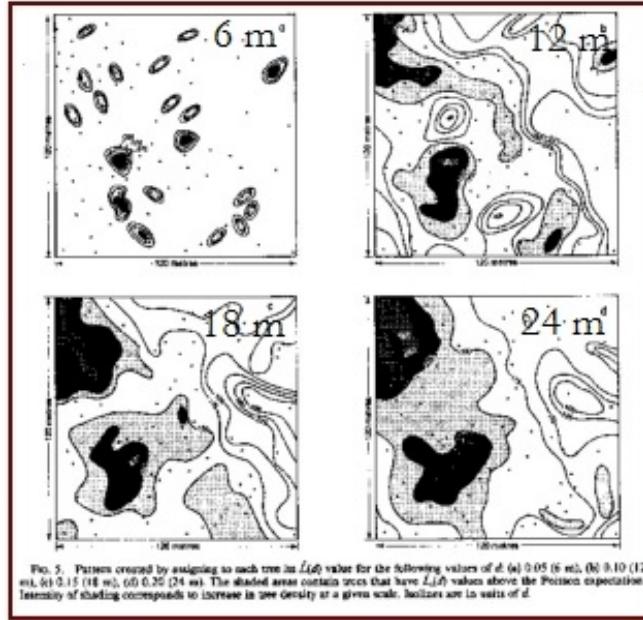
In the real-world example shown here for the distribution of vernal pools in an ecoregion in Massachusetts, the K function reveals that pools are more clumped than expected under a spatially random distribution out to a distance of roughly 1.7 km, but at distances greater than roughly 2 km, the pools are more uniformly distributed than random. The scale of clumping of pools is interesting given the dependence of many vernal pool amphibians on metapopulation processes. For many of these species the maximum dispersal distance roughly corresponds to the maximum scale of pool clumping.



Scaling Techniques for Point Patterns

▪ Extensions to Ripley's K-distribution

- Pattern intensity maps
 - ▶ Modified Ripley's K for each point in the data set (Getis and Franklin 1987)
 - ▶ Contours of point cluster intensity at each scale (neighbor distance)
 - ▶ Shaded areas have $L(d)$ values above the Poisson expectation



Extensions to Ripley's K

Ripley's K lends itself to two logical modifications that extend its applicability:

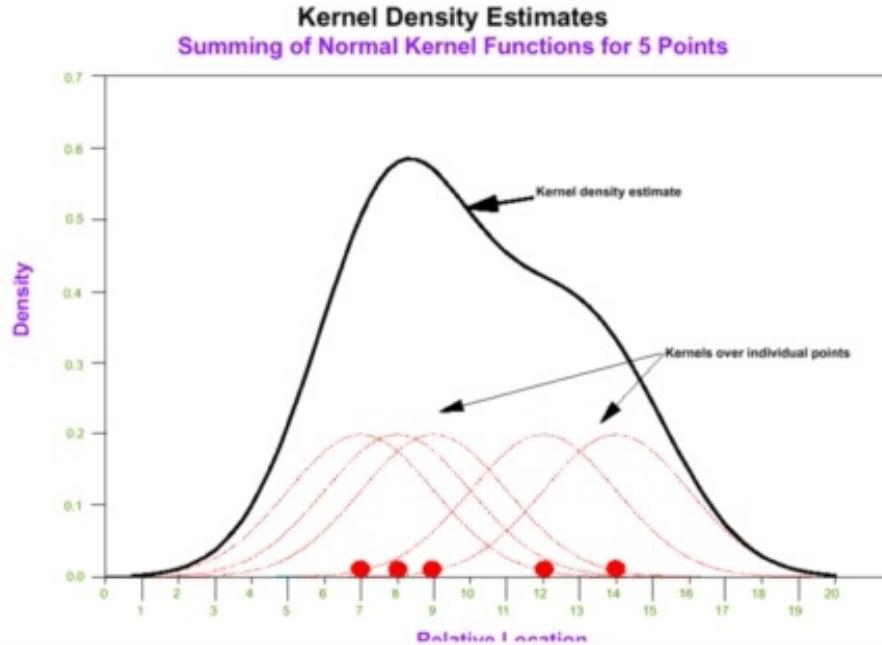
(1) Bivariate Ripley's K.—One extension to Ripley's K is to consider the distances among points or individuals of differing types, for example, different tree species, or trees of different size classes. Such applications are often of interest to ecologists, and can be performed using a rather straightforward extension to Ripley's K.

(2) Pattern Intensity Maps.—Getis and Franklin (1987) illustrated a variation on Ripley's K in which they consider only half of the summation in the previously given equation for $K(d)$; that is, they sum only over each point i . With the further adjustment of dividing by $(n-1)$ to correct for the one-pass summation, their transformed index is equivalent to $L(d)$, but it yields an index that can be computed for each point in the data set (each tree, in their case). This extension allows them to map the K function for each tree, and then to contour the density of the overlaid K functions from all points. The result is a contour plot that emphasizes regions of varying tree cluster density, with this cluster density (intensity) quantitatively defined. This provides a means of using the information about pattern at each point in the data set, in contrast to techniques that average the pattern over all points. While their approach hasn't exactly caught on among ecologists, it does have great potential as an aid to examining complex point patterns.



Scaling Techniques for Point Patterns

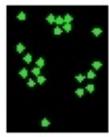
■ Kernel Density Approach



Kernel density maps as an alternative to Ripley's K pattern intensity maps

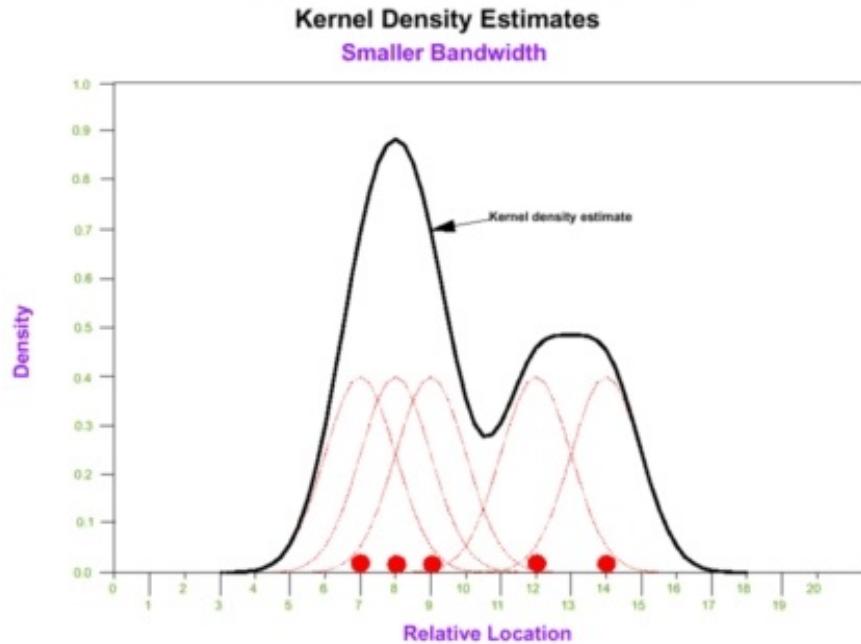
The kernel estimator (Silverman 1986; Worton 1989) is a density estimator commonly used for home-range analysis in radiotelemetry studies. However, kernels have a wide variety of applications in ecological studies. Here, we show how kernels can be used to create pattern intensity maps and, much like the Ripley's K extension, how they can be used to represent pattern intensity across a range of spatial scales.

It is easiest to illustrate the kernel estimator as applied to a one-dimensional set of points, e.g., representing trees or vernal pool locations along a transect. A kernel of any specified shape and width is placed over each point. In this example, a normal (or Gaussian) curve is placed over each point to reflect the nonlinear strength of ecological interaction between a focal point and its ecological neighborhood. Note, a major strength of the kernel approach is the ability to reflect non-linear and/or non-parametric ecological distance relationships. By adding up all the individual kernels, the cumulative kernel represents a distance-weighted point density estimate. As seen here, points in the center of clusters end up with higher densities.

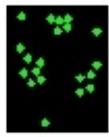


Scaling Techniques for Point Patterns

■ Kernel Density Approach

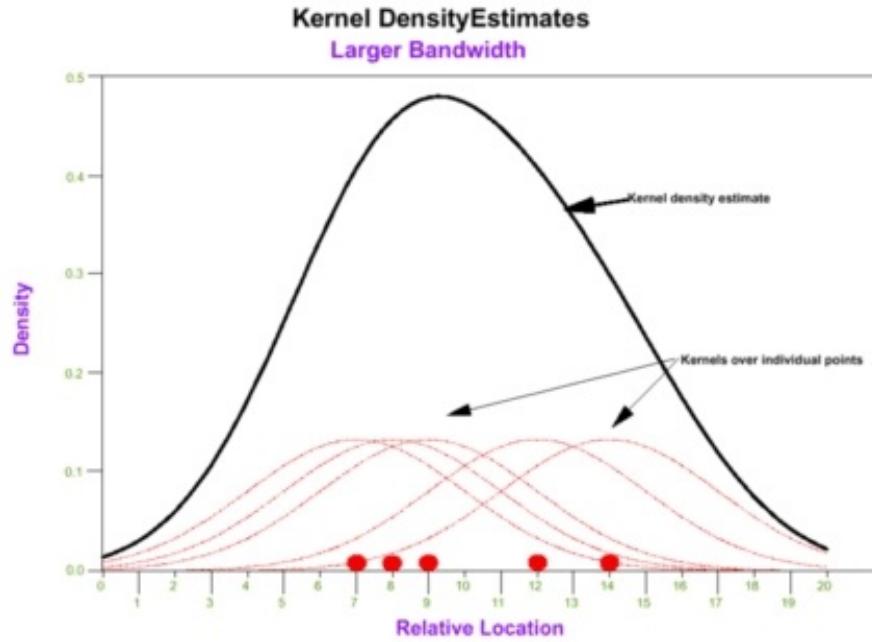


The great utility of kernels lies in the ability to adjust the width of each kernel to capture various scales of pattern intensity. For example, if we set the width of the normal curve, usually referred to as the *bandwidth*, h (which in the case of a normal curve is usually set equal to 1 standard deviation) to a relatively small value, then the cumulative kernel surface will reflect relatively fine-scale heterogeneity in point densities – the kernel surface will be relatively “rough”.

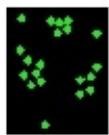


Scaling Techniques for Point Patterns

■ Kernel Density Approach

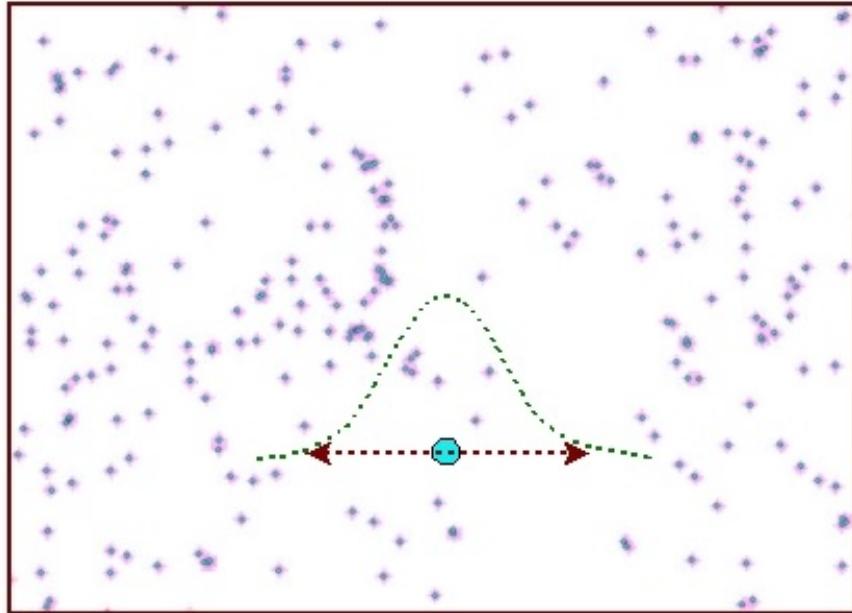


On the other hand, if we set the width of the normal curve to a relatively large value, then the cumulative kernel surface will reflect relatively coarse-scale heterogeneity in point densities – the kernel surface will be relatively “smooth”. Consequently, the single parameter h allows us to capture pattern intensity at any spatial scale. By varying h we can smooth or roughen the map.



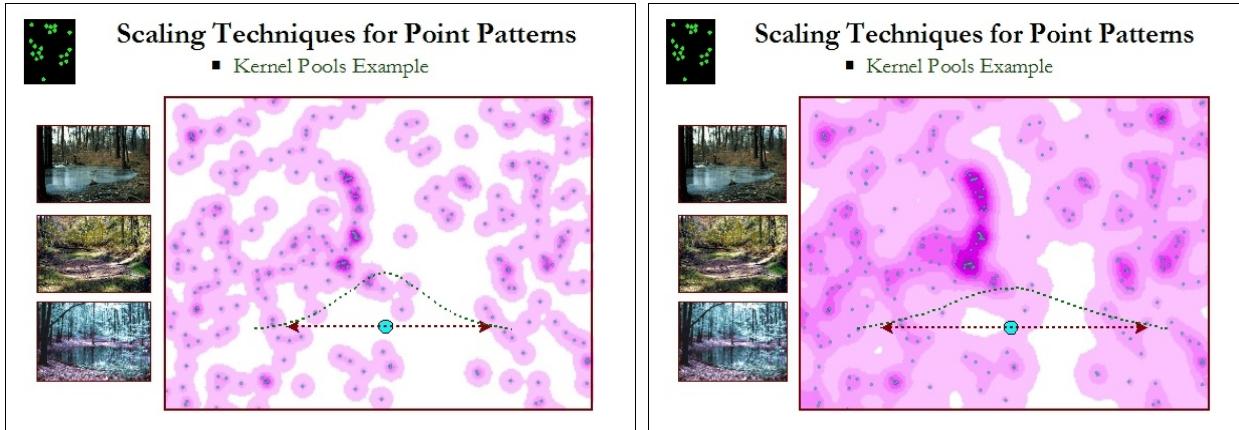
Scaling Techniques for Point Patterns

■ Kernel Pools Example

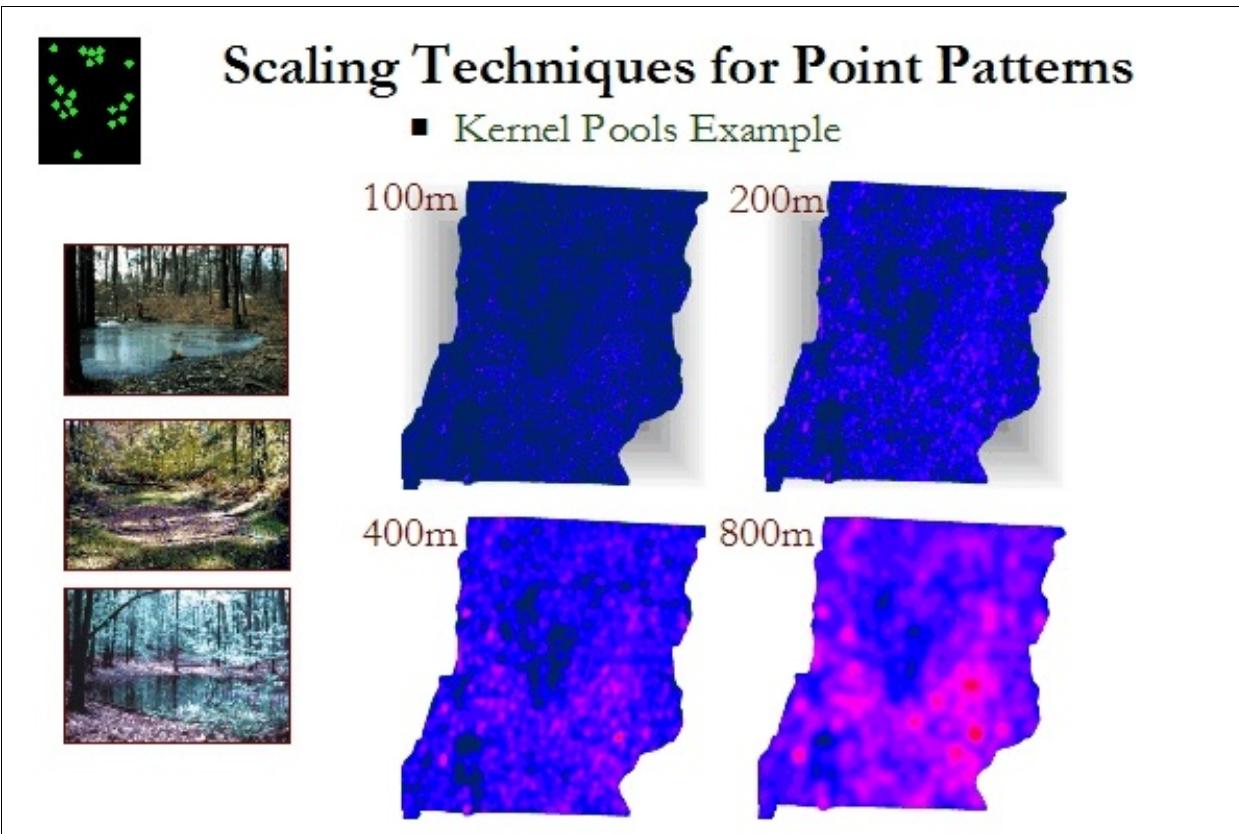


If we have a two-dimensional set of points, as is usually the case in landscape studies, we can place a bivariate kernel over each point, where the kernel has width along both the x-axis and y-axis and the kernel height represents density. A bivariate normal kernel is like placing a bell of some specified size over each point. As with the univariate kernel, adding up all the bivariate kernels produces a cumulative kernel surface of point density. Peaks in the surface represent regions of high point density, whereas troughs represent regions of low point density.

In the example shown here, a bivariate normal kernel is placed over each vernal pool in small study area from western Massachusetts. In this case, the bandwidth is specified to be relatively small compared to the spacing of vernal pools, so there is very little overlap among the individual kernels except where pools are closely spaced together. The cumulative kernel surface is extremely rough, reflecting the fine-scale heterogeneity in pool clusters.



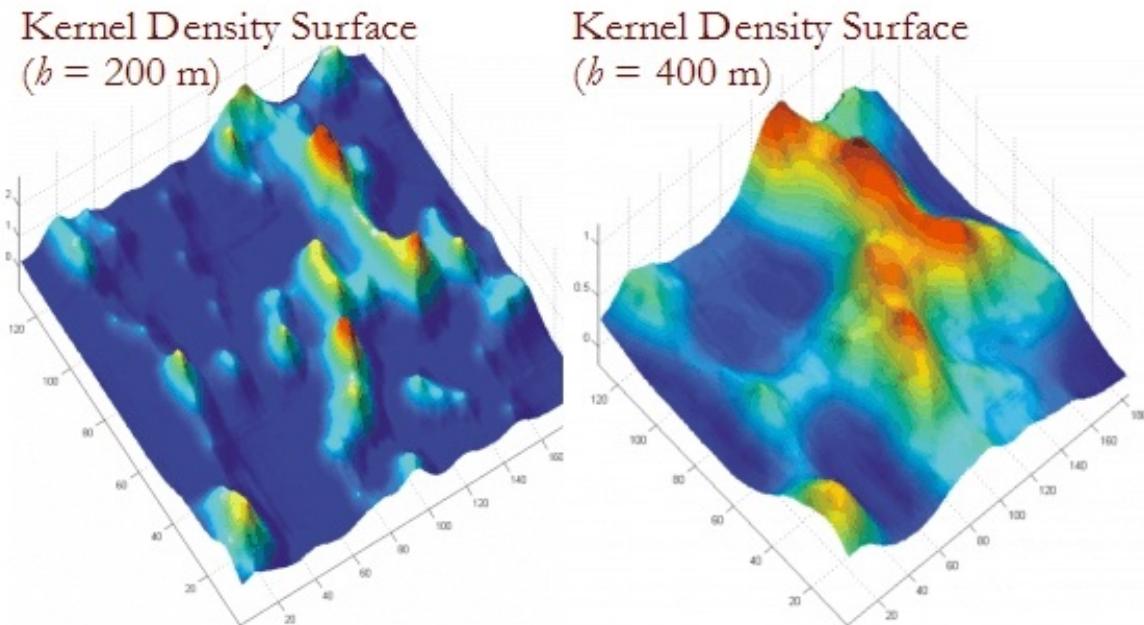
As we increase the bandwidth, the individual kernels reach out to greater distances and begin to overlap each other. As we increase the bandwidth even further, the individual kernels greatly overlap and the cumulative kernel surface smoothens, reflecting relatively coarse-scale heterogeneity in pool clusters. Varying the bandwidth allows us to represent the spatial heterogeneity in the point patterns at different scales. These pattern intensity maps are much like those based on Ripley's K; however, the kernel approach allows one to incorporate ecological information into the kernel to reflect non-linear ecological distance relationships.





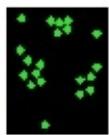
Scaling Techniques for Point Patterns

■ Kernel Pools Example



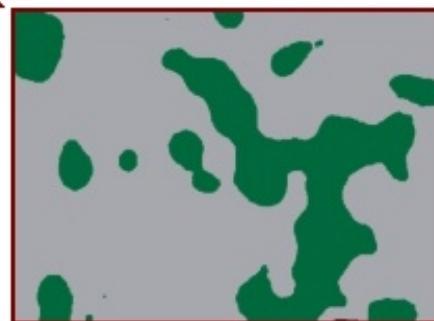
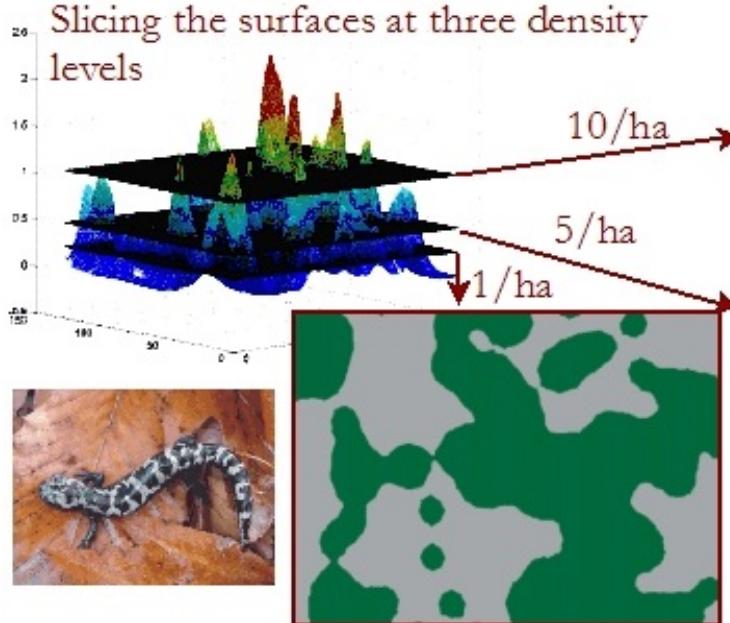
These cumulative kernel surfaces are perhaps best illustrated in three dimensions, as shown here. The kernel surface on the left results from a bivariate normal kernel with a bandwidth (standard deviation) of 200m. The one on the right has a bandwidth of 400m. Clearly, the smaller the bandwidth, the rougher the surface. The peaks represent regions of high vernal pool density, the troughs represent regions of low vernal pool density.

The challenge, of course, is in choosing the right bandwidth. In many cases, we don't know what the bandwidth should be; that is, we have no a priori reason for assuming one bandwidth is better than another. In such cases, there are cross-validation procedures for selecting the "best" bandwidth, where the best bandwidth is the one that best represents the intrinsic scale of the point pattern; i.e., its characteristic scale. However, one of the real strengths of the kernel method lies in the ability to represent pattern intensity across a range of scales, and then relate these scaled patterns to ecological processes to better understand and characterize the scaling relationships.



Scaling Techniques for Point Patterns

■ Kernel Pools Example

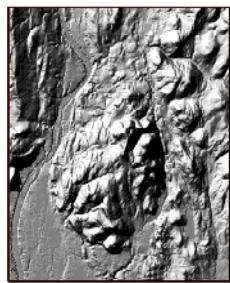


For example, based on a long-term field study of marbled salamanders in western Massachusetts we have empirical data on true dispersal distances – based on recorded juvenile salamanders emerging from their natal ponds and dispersing to new ponds to breed as adults. By specifying the shape and width of the kernel to reflect these empirical observations, we can effectively map the intensity of vernal pools across the landscape from the perspective of dispersing animals. The cumulative kernel surface is scaled to represent the probability of a dispersing individual from neighboring pools arriving at a pool. By “slicing” this surface at different levels, we can map the functional connectivity of vernal pools for dispersing organisms at different levels of connectivity.

5.4. Scaling Techniques for Continuous Surface Data

- Characteristics of continuous surface patterns:
 - ▶ Continuous surface patterns represent quantitative measurements that vary continuously across the landscape; there are no explicit boundaries (i.e., patches are not delineated). Here, the data can be conceptualized as representing a three-dimensional surface, where the measured value at each geographic location is represented by the height of the surface. A variable that takes on values based on its spatial location is called a “regionalized” variable.
 - ▶ A familiar example is a digital elevation model, but any quantitative measurement can be treated this way (e.g., plant biomass, leaf area index, soil nitrogen, animal density).
 - ▶ In many cases the data is collected at discrete sample locations separated by some distance. Analysis of the spatial dependencies (or autocorrelation – the ability to predict values of a variable from known values at other locations, as described by a ‘structure function’) in the measured characteristic is the purview of **geostatistics**, and a variety of techniques exist for measuring the intensity and scale of this spatial autocorrelation. Techniques also exist that permit the kriging or modeling of these spatial patterns; that is, to interpolate values for unsampled locations using the empirically estimated spatial autocorrelation. These particular surface pattern techniques were developed to quantify spatial patterns from sampled data (n) collected either along a one-dimensional transect or within a two-dimensional study area. When the data is exhaustive (i.e., the whole population, N) over the transect or study landscape, like it often is with remotely sensed data, other techniques are available as well (e.g., quadrat variance analysis, lacunarity analysis, fractal analysis, spectral analysis, and wavelet analysis) and may be more appropriate.
 - ▶ All scaling techniques for continuous surface patterns share a goal of describing the intensity and scale of pattern in the quantitative variable of interest. In all cases, while the location of the data points (or quadrats) is known and of interest, it is the values of the measurement taken at each point that are of primary concern. Here, the basic question is, “Are samples that are close together also similar with respect to the measured variable?” Alternatively, “What is the distance(s) over which values tend to be similar?”

Scaling Techniques for Continuous Data

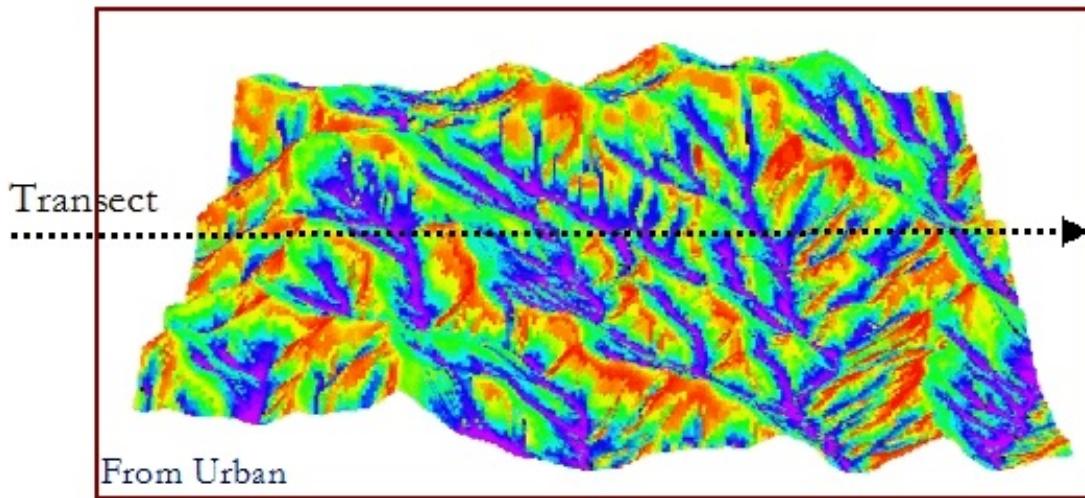


- Data Characteristics
 - ▶ Data comprise quantitative measurements at discrete sample locations separated by some distance.
- Examples
 - ▶ Map of elevation
 - ▶ Map of leaf area index
- Goal of Analysis
 - ▶ To determine whether samples that are close together are also similar with respect to the measured variable, and to identify the distance (scale) over which values tend to be similar.



Scaling Techniques for Continuous Data

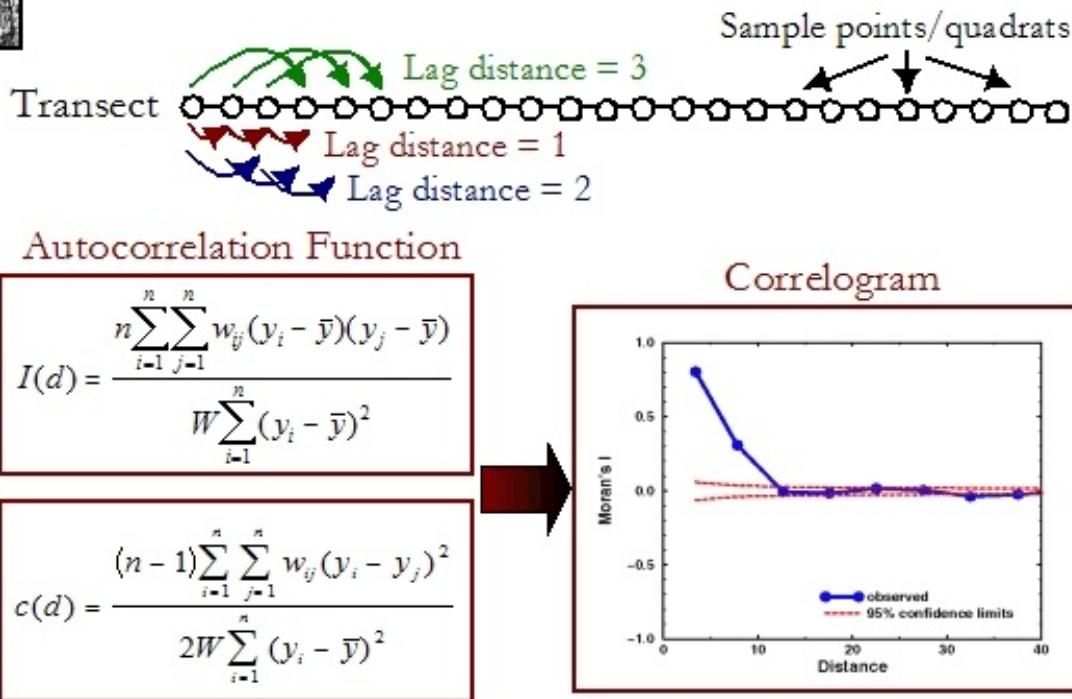
Topographic moisture index for
Coweeta Experiment Station



Consider an example in which you would like to discover the characteristic spatial scale (grain) of topography, because you know that many ecological phenomena are influenced by this. First, you must choose a variable to measure to index topography (you might choose elevation, or slope, or aspect, or some combination index). Measure this value at a given point on the side of a hillslope. Now walk a few steps away, and measure this variable again. The values would likely be similar. Walk farther, and repeat. Continue moving farther way, remeasuring the variable and comparing it to that measured at the starting point. At some point, you would find yourself on a different hillslope facet, because you crossed a divide or passed through a cove. Suddenly, the measured values are different. If you could do this for a very large number of starting points, you would discover the average size of a hillslope facet, i.e., the characteristic scale of topography. So how can we do this more quantitatively? Two methods, *autocorrelation* and *semivariance* analysis, are popular tools to discover the scale of pattern in continuous surface data (Legendre and Fortin 1989).



Scaling Techniques for Continuous Data



Autocorrelation

First, let's consider a way to index similarity in the variable as measured for two samples. A simple correlation coefficient will suffice:

$$r_{ij} = \frac{n \sum_{i=1}^n \sum_{j=1}^n (y_i - \bar{y})(y_j - \bar{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

where the values y for two samples denoted i and j are deviations $y_i - y_{mean}$ and there are a total of n of these. Now we need a way to qualify this index so that it is scale-specific, that is, so it reflects the correlation between points separated by some particular distance. An easy way to do this is to introduce an indicator or switching function (a weight) such that the weight takes on a value of 1 only if the two points are separated by a given distance d ; otherwise the weight is set to 0.

The new index, computed as **Moran's I** , is:

$$I(d) = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{W \sum_{i=1}^n (y_i - \bar{y})^2}$$

where w_{ij} is the indicator (weight) for samples i and j , and W is the sum of all the weights. Like a familiar Pearson's correlation coefficient, $I(d)$ is bounded on $[-1, 1]$, with 1 being perfect positive autocorrelation. The expected value under randomness is a value near 0, signifying a lack of autocorrelation.

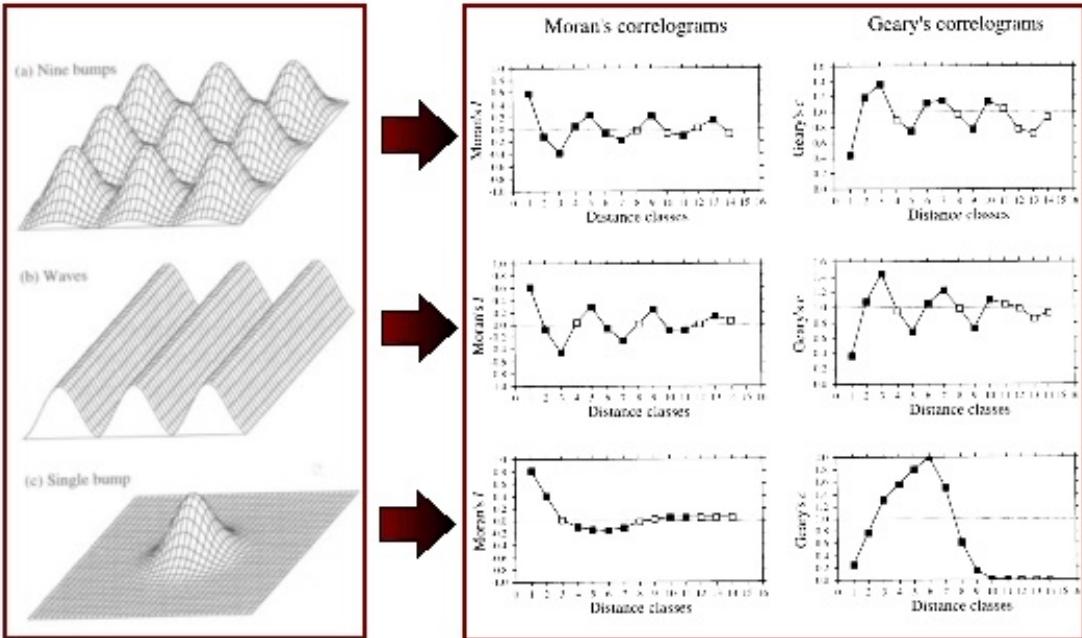
Geary's c is a closely related alternative index based on the ecological distance between points at each lag distance:

$$c(d) = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - y_j)^2}{2W \sum_{i=1}^n (y_i - \bar{y})^2}$$

Geary's c ranges from zero to infinity, with an expected value of 1 under no autocorrelation. Values from zero to one indicate positive spatial autocorrelation, values above 1 negative spatial autocorrelation. Geary's c is closely related to semivariance (see below); Geary's c is essentially a standardized semivariance. Geary's c and Moran's I are usually strongly, but inversely, correlated and thus usually yield similar interpretations. Moran's I appears to be the more popular statistic.



Scaling Techniques for Continuous Data

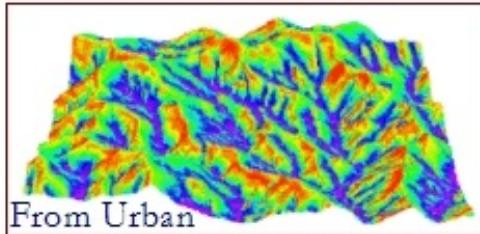


A *correlogram* is a graph of autocorrelation versus distance. As a scaling technique, we are interested in distances at which a variable is strongly autocorrelated. With ecological data, this is usually positive autocorrelation, since things closer together tend to be more alike; negative cases are not common and at short distances can reflect either avoidance (e.g., allelopathy) or that the sampling step is too large compared to patch size. In gradients, short distance positive autocorrelation is coupled with negative autocorrelation at longer lag distances. For a repeating pattern of clumps separated by some distance, the clump size appears as a peak in autocorrelation for distances up to average clump size; the distance between clumps appears as a second peak. For very regular patterns such as a sine wave, the size of the pattern is indicated by the distance between peak and trough (maximum dissimilarity of measurements).



Scaling Techniques for Continuous Data

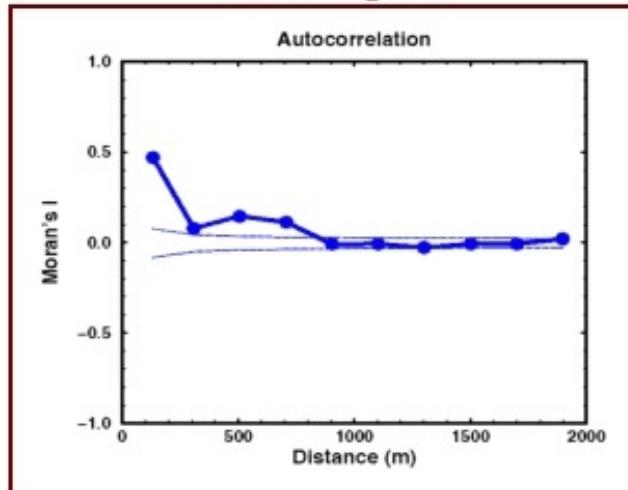
Topographic moisture index



1600 ha watershed
400 random samples



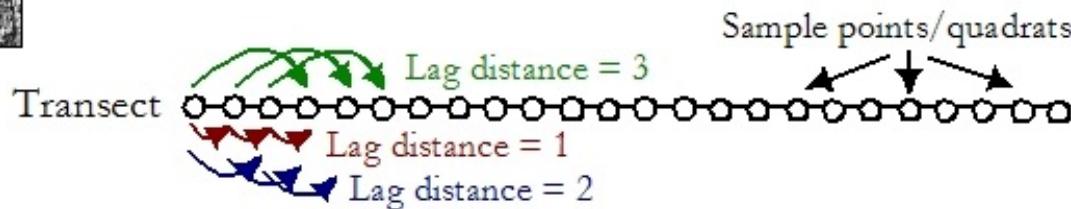
Correlogram



In the case of topography, we'd expect positive autocorrelation within the scale of slope facets, with $I_{(d)}$ approaching 0 at scales beyond the grain of topography. In this example, 400 random samples of a topographic moisture index (derived from local slope and drainage area) were taken from a 1600 ha watershed at the Coweeta Experimental Station. The correlogram indicates strong positive autocorrelation at distances of 100-300 m, indicating the scale of the dominant slope facets affecting topographic moisture (e.g., riparian areas), but there is also some positive autocorrelation at 500-700 m, representing a second scale of variation reflective of the scale of the major drainage patterns.

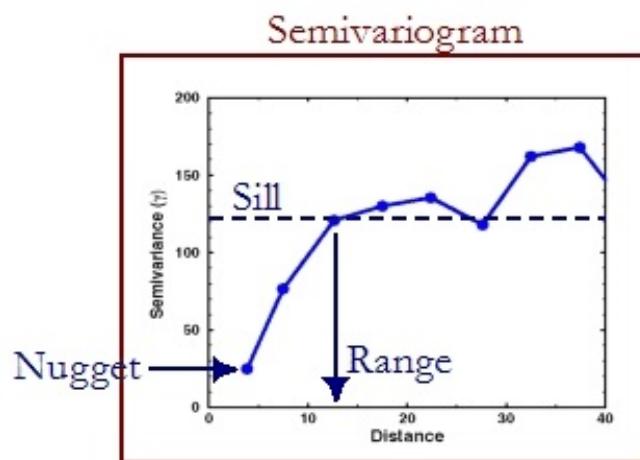


Scaling Techniques for Continuous Data



Semivariance

$$\gamma(d) = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - y_j)^2}{2n_d}$$



Semivariance

An alternative way to look at spatial structure in geostatistical data is to consider the dissimilarity in values measured at two points some distance apart. Following the notation used for autocorrelation, where y_i is an attribute value measured at location i and n is the total number of points, semivariance g (gamma) at distance d is defined:

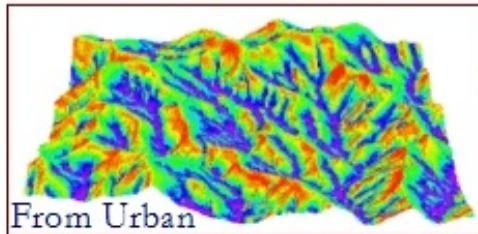
$$g(d) = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - y_j)^2}{2n_d}$$

where j is a point at distance d from i , and n_d is the number of points at that distance (or in that distance class). The summation term is a distance measure or variance term; by convention this is divided by 2 to rescale it. Note that if two points close together are similar, the difference term in the numerator is small, and they have low semivariance at that distance. At some distance farther removed, the association between points is no different from that of a random pair of points sampled from the grid (as rescaled, semivariance asymptotes to simple variance as estimated over all points).

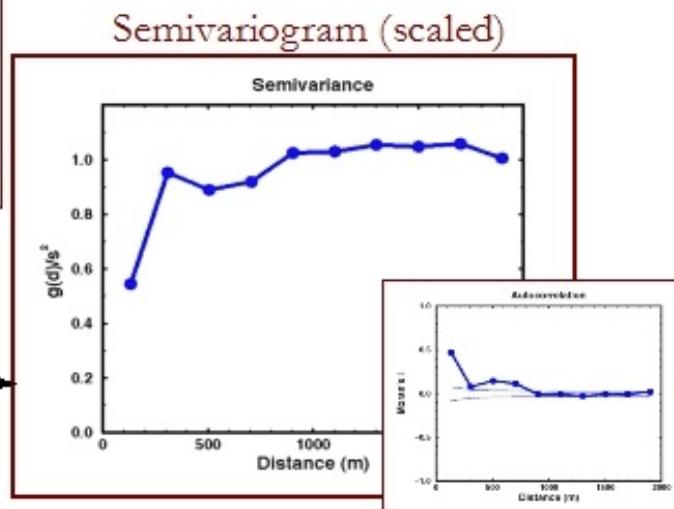


Scaling Techniques for Continuous Data

Topographic moisture index



From Urban
1600 ha watershed
400 random samples



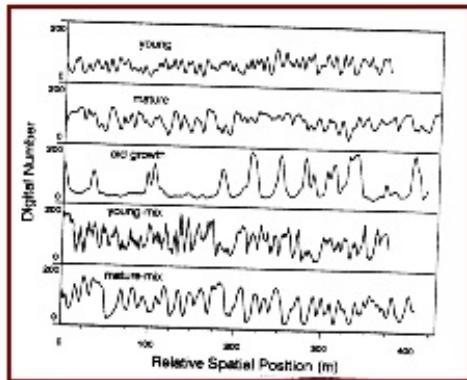
A *semivariogram* is a plot of semivariance against distance (or distance class). A semivariogram is defined by its **sill** (asymptotic value), its **range** (the distance at which the sill is reached), and its **nugget** semivariance (the intercept). Semivariance may be divided by the simple variance to rescale it so that it should approach a value of 1 beyond the range of the variogram.

As a scaling technique, the range is of primary interest as it indicates the characteristic scale of the variable. The nugget is often of interpretative value, in that a large (nonzero) nugget might indicate spatial structure at scales smaller than the minimum resolution of the sample points (i.e., at scales finer than the minimum between-sample distance); this would indicate that some finer-scale pattern was missed in sampling.

In the case of topography, we'd expect the range in the semivariogram to reflect the scale of slope facets, with semivariance approaching the sill at scales beyond the grain of topography. In the example of the topographic moisture index from the Coweeta Experimental Station, the range is roughly 300 m, indicating that beyond this distance the topographic moisture index varies randomly. As you can see from the paired diagrams, the semivariogram and correlogram are inversely related and often convey the same information regarding the scale of pattern. The choice of methods is partly a matter of personal preference, but there are some other subtle differences as well that may be important in choosing one over the other in a particular application.



Scaling Techniques for Continuous Data



Cohen et al. (1990)

Digital aerial video
400 m transect
1 m resolution

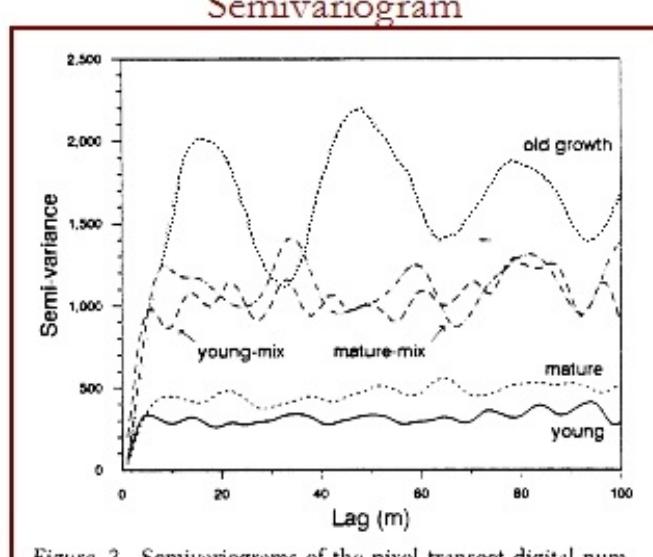


Figure 3. Semivariograms of the pixel transect digital numbers selected from the 1 m spatial resolution images of the five forest stands.

In this example, digital number from remotely sensed images was recorded at 1-m intervals along transects through five forest stands representing a gradient of seral stages of stand development in the Pacific Northwest (Cohen et al. 1990). The raw transect profiles are shown in the left panel. It is quite obvious to the naked eye that the characteristic scale varies among stands. The right panel shows the empirical semivariograms derived from these data. The patterns in the semivariograms reveal several important differences among stands in vegetation structure. For example, the range increases progressively from the young to old stands, reflecting the increasing size of the dominant tree crowns as stands age. Variability in the sill increases with stand age as well, reflecting the increasing heterogeneity in crown structure in older stands (i.e., increasing prevalence of canopy gaps). The average distance between tree crowns is revealed by the repeating peaks in the semivariogram. Clearly, a great deal of the differences in the scale of crown heterogeneity is revealed by these semivariograms.

General Considerations- Caveats on the use of Structure Functions

There are a number of general considerations or caveats on the use of autocorrelation and semivariance structure functions, and while it is beyond the scope of this presentation to describe each of these in detail and provide examples, a short description of each is warranted.



Scaling Techniques for Continuous Data

■ General Considerations and Caveats

- Statistical independence...
- Global significance...
- Sample size and statistical power...
- Non-Euclidean distances...
- Stationarity...
- Isotropy...



- **Statistical independence.**—Note that even if spatial scaling is not the primary goal of the analysis, autocorrelation and semivariance analysis can be an important part of screening geostatistical data. Samples that are autocorrelated are not independent; indeed one definition of autocorrelation is that it implies that the value of a variable at one location can be predicted from values measured at other (nearby) locations. This lack of independence violates a fundamental assumption in parametric statistical tests, and so autocorrelation analysis should be conducted prior to parametric statistical analysis.
- **Global significance.**—Most programs that compute autocorrelation also provide the test of significance. One caveat to bear in mind, however, is that autocorrelation tests performed over a large number of distance classes is actually a set of simultaneous tests. Global significance of the autocorrelation function for m distance classes is adjusted by alpha global = α/m (the Bonferroni correction). The autocorrelation is globally significant if any distance class is significant at alpha global; likewise, only those distance classes meeting this adjusted significance criterion should be interpreted. By convention, correlograms are displayed with open symbols for the data points, and significant correlations at lags are highlighted with filled symbols.
- **Sample size and statistical power.**—In general, the number of points (pairs) per distance class decreases as distance increases, and consequently the power of the test decreases as well. One typically truncates the lag distance to a threshold that retains a sufficient number of observations at the largest distance (say, 30 pairs of points); alternatively, the maximum distance is set to, at most, half the minimum dimension of the data grid. Another way around the problem of decreasing power is to use distance classes of differing intervals. For example, one might define distance classes according to the frequency distribution of inter-point distances, so that the number of points per class is equal.

- Non-Euclidean distance.—Of course, there is no reason why the distances used in the analysis need be integer values nor strictly Euclidean. For example, one could compute a single autocorrelation index for data in which the weights w_{ij} are computed $w_{ij} = 1/d_{ij}^2$, where d_{ij} is the distance between points i and j . This invokes an inverse relationship between weight and distance; other functions of distance could be used as well. Various 'weighted distance' or 'functional distance' terms might be used, depending on the process of interest. For example, in modeling the spatial structure of seed dispersal or air pollution, one might want to weight downwind distances differently than upwind distances; similarly, many gravity-driven processes might call for higher downhill weights. In these cases the distance matrix would be asymmetric but the calculations proceed in exactly the same way. One could also use distances that are weighted by a resistance term, for example, to indicate the relative ease of travel along the distance (this would make sense for animal dispersal studies where different habitats have varying navigability). Similarly, one might choose to specify an adjacency matrix exactly for an analysis, assigning each of the weights w_{ij} in a matrix of 1's and 0's for 'connected' and 'unconnected' respectively. This flexibility to assign weights (connections) explicitly provides for quite specific hypothesis tests. For example, one might explore the spatial structure of a metapopulation linked by dispersal, by specifying presumed dispersal corridors among habitat patches and then testing for autocorrelation in species abundance for this spatial structure. Of course, you would explore alternative structures as well, perhaps using this analysis iteratively to discover a most-likely spatial structure for the metapopulation.
- Stationarity.—The most difficult assumption of autocorrelation is that there is a single dominant pattern that is invariant across the entire study area. In other words, patterns do not change over irregular gradients, or patches, or domains of structure across the study area, or if they do they follow a repeated and stable overall pattern. This assumption is needed for any structure function. Does the world behave this way? Are we interested specifically in how they diverge from this? Irregular patterns are at the heart of many issues in spatial ecology – so can spatial autocorrelation analysis help us for these questions? Given a stationary pattern though, autocorrelation does a great job of describing the scales of positive or negative correlation - which is very useful.

Formally, the stationarity of interest is second-order (weak) stationarity, which requires that (a) the mean is constant over the region; (b) variance is constant and finite; and (c) covariance depends only on between-sample spacing. In many cases this is not true because of larger trends in the data; in these cases, the data are detrended before analysis. One way to detrend data is to fit a regression to the trend, and save only the residuals for autocorrelation analysis; many other ways are used commonly. Newer (computationally intense) approaches allow one to adjust the structure function locally to correct 'on the fly' for nonstationarity (Rossi et al. 1992).



Scaling Techniques for Continuous Data

■ Isotropy and Anisotropy

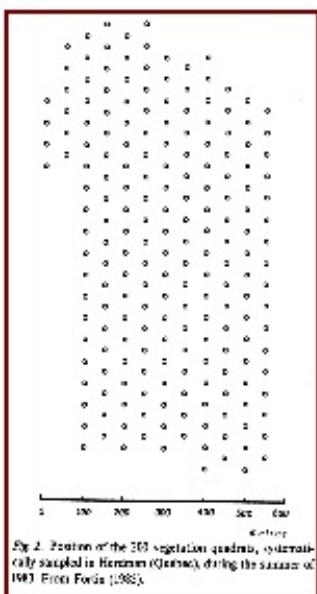


Fig. 2. Position of the 203 vegetation quadrats systematically sampled in Rennan (Quebec) during the summer of 1987. From Fortin (1988).

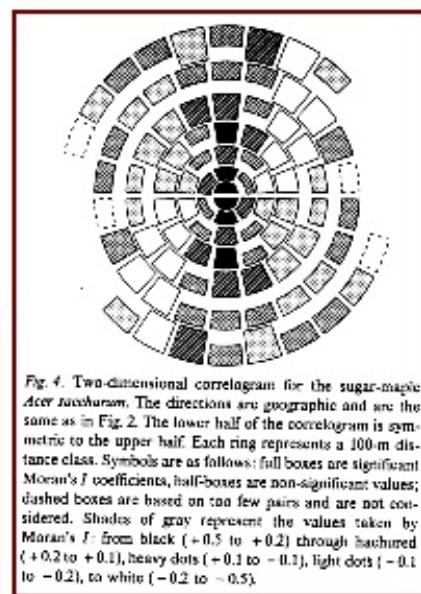
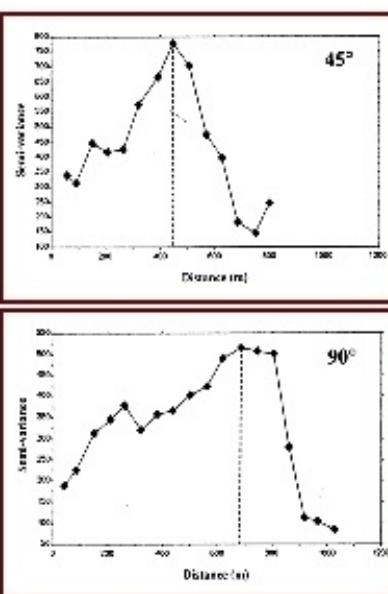


Fig. 4. Two-dimensional correlogram for the sugar-maple *Acer saccharum*. The directions are geographic and are the same as in Fig. 2. The lower half of the correlogram is symmetric to the upper half. Each ring represents a 100-m distance class. Symbols are as follows: full boxes are significant Moran's I ; half-boxes are non-significant values; dashed boxes are based on too few pairs and are not considered. Shades of gray represent the values taken by Moran's I : from black (+0.5 to +0.2) through hatched (+0.2 to +0.1), heavy dots (+0.1 to -0.1), light dots (-0.1 to -0.2), to white (-0.2 to -0.5).

- **Isotropy and Anisotropy.**--A simple autocorrelation is all-directional in that the index is averaged over all points regardless of their orientation in space. This is tantamount to assuming isotropy in the data, that is, that the pattern is similar in all directions. (In fact, autocorrelation does require the formal assumption of isotropy.) But ecological patterns are often anisotropic (e.g., when conditioned by topography). To make an autocorrelation directional, one adjusts each weight w_{ij} to reflect the angle point j forms with point i relative to a specified test direction (say, due north) (Oden and Sokal 1986). A separate autocorrelation could be computed for any given direction or range of directions, as shown in this example of sugar maple (*Acer saccharum*) distribution in a forest. Note in this example that the range is noticeably greater along the East-West gradient. Programs now exist that compute omni-directional correlograms such as the one shown in the right panel of this figure, by slicing space into a finite number of 'wedges' and computing the autocorrelation within each wedge separately (e.g., for 8 45° wedges) (see Legendre and Fortin 1989 for a review of available software).

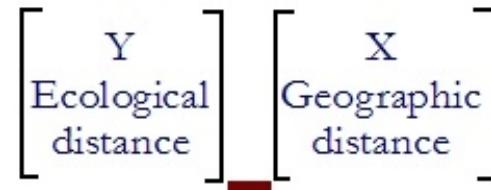


Scaling Techniques for Continuous Data

■ Extensions to Autocorrelation Methods

- Nominal and ordinal data...
- Bivariate (cross-) correlograms...
- Mantel's test – differences between two distance matrices, representing ecological and geographic distances between sample points.

Normalized
Mantel statistic



Mantel statistic

$$z = \sum_i \sum_j x_{ij}y_{ij}$$

$$r = \frac{\sum_i \sum_j \left[\frac{(x_{ij} - \bar{x})}{s_x} \right] \left[\frac{(y_{ij} - \bar{y})}{s_y} \right]}{n - 1}$$

Extensions to these Methods

Nominal and ordinal data.--While the previous discussion has focused on interval-scale data, autocorrelation can also be computed for ranked data. The method for ranked data is similar to those for interval-scale data; the difference is in the calculation of expected values (Sokal and Oden 1978). For categorical data, autocorrelation focuses on 'join counts' of like or unlike 'edges' in the data (e.g., AA, AB, or BB adjacencies). Sokal and Oden (1978) provide formulae for the modified indices and tests of significance.

Bivariate data.--One can compute the spatial correlation between two variables, yielding a spatial cross-correlation and a cross-correlogram. There are some interpretative differences between auto- and cross-correlograms. A simple (auto-) correlogram is symmetric about a lag distance of 0, and so by convention the correlogram is illustrated for lags greater than 0 only. By contrast, cross-correlation need not be symmetric because which point 'comes first' can make a difference. For example, if the two variables of concern are canopy leaf area and ground-level light, then this relationship is asymmetric because of solar inclination angles: shadows fall longer to the north than to the south. Thus, cross-correlograms are illustrated for positive as well as negative lag distances.



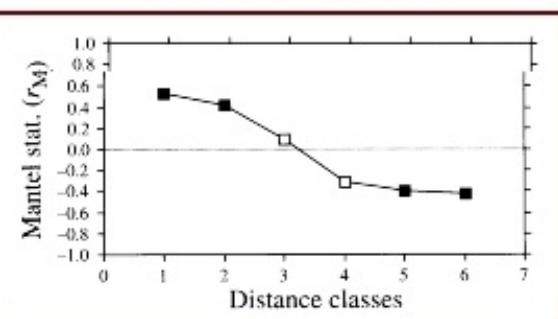
Scaling Techniques for Continuous Data

■ Mantel's Correlogram

$$\begin{bmatrix} Y \\ \text{Ecological} \\ \text{distance} \end{bmatrix} \quad \begin{bmatrix} X \\ \text{Geographic} \\ \text{distance} \end{bmatrix}$$
$$\begin{bmatrix} 0 & & r_{m2} \\ 2 & 0 & \\ 2 & 0 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & & r_{m1} \\ 0 & 1 & \\ 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Etc.

Mantel correlogram

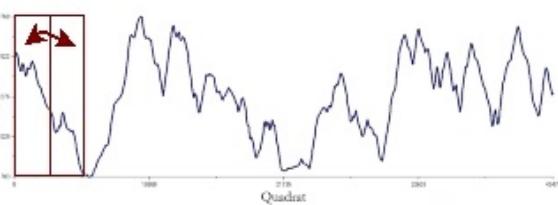
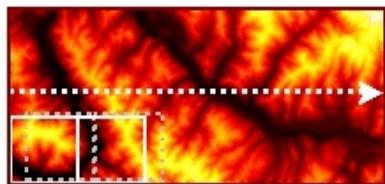


Mantel's Test.—Mantel's test evaluates the correlation between two (or more) distance matrices (Mantel 1967). As commonly used in ecology, a matrix of spatial distances is used to partial out the spatial autocorrelation in environmental variables, where the autocorrelation is averaged over all distances. However, the analysis can also be repeated at specific distances to yield a Mantel correlogram (Legendre and Fortin 1989). The key distinction of the Mantel test is the use of two (or more) distance matrices. One matrix represents the geographic distances (e.g., Euclidean or functional) between all pairs of sample points; the other matrix represents the ecological distances between the same set of points. The Mantel test is analogous to a multivariate regression, where the independent variable is geographic distance and the dependent variable is ecological distance. The great utility of this method stems from the flexibility in defining ecological distance among sample points in either univariate or multivariate terms (e.g., dissimilarity among communities based on species presences or abundances, dissimilarity among individuals or populations based on genetic information, etc.), and there are numerous ecological distance metrics developed for this purpose.



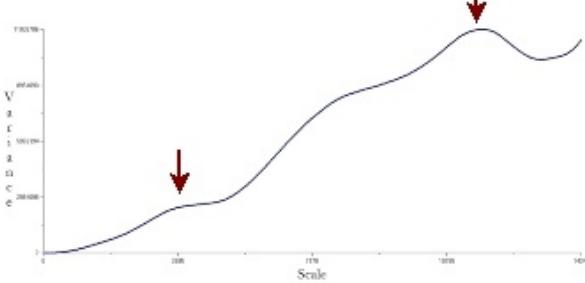
Scaling Techniques for Continuous Data

■ Quadrat Variance Analysis



Two-term local quadrat variance:

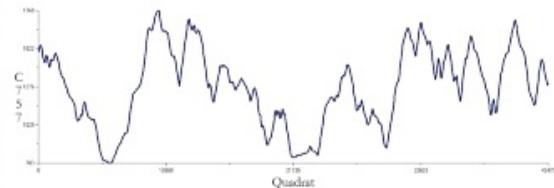
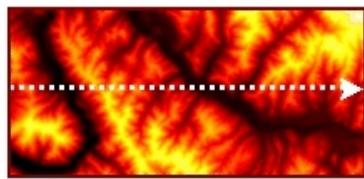
$$TTLQV = \frac{\sum_{i=1}^{n+1-2b} \left(\sum_{j=i}^{i+b-1} y_j - \sum_{j=i+b}^{i+2b-1} y_j \right)^2}{2b(n+1-2b)}$$



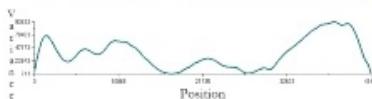
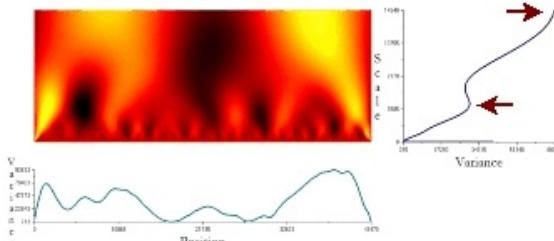
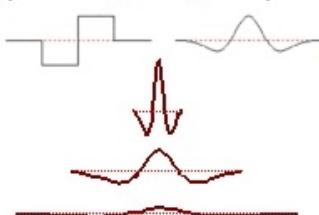


Scaling Techniques for Continuous Data

■ Wavelet Analysis



Wavelet coefficient =
 $f(\text{wavelet, scale, data})$



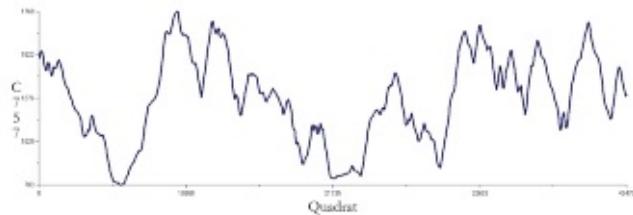
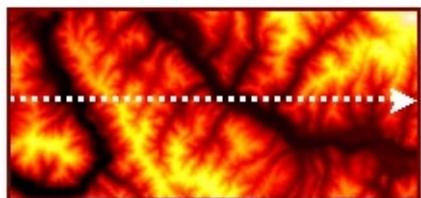
Wavelet Analysis.—Wavelet analysis is similar to many of the quadrat variance methods, although it is much more flexible (Bradshaw and Spies 1992, Dale and Mah 1998). For our purposes, a wavelet function, $g(x)$, is a scalable windowing function whose integral over all x equals zero. One way to think of this is that the wavelet function describes a shape template that can be scaled to a desired size, then slid along the transect or across the surface. When the template fits the observed data well, the value of the wavelet transform at that position is high; when it does not, the value is low. The overall variance for a given scale is equal to the average of the squared wavelet transforms at all positions. The overall variance for a given position is equal to the average of the squared wavelet transform over all scales. There are numerous wavelet functions, $g(x)$, that may be used for data analysis. The simplest is known as the Haar wavelet. It is a square wavelet function that is essentially identical to TTLQV; it compares the values of neighboring blocks of the same size. The wavelet identical to 3TLQV is known as the French Top Hat wavelet; it compare the values of a central block to those of its neighbors. A key result of wavelet analysis is a plot of the variance against scale, identical to the output from the quadrat variance methods. This plot shows the average variance (across all positions) for a given scale. Peaks in this plot indicate the scale of pattern being investigated. A plot of variance against position in the data series reveals locations where something is happening, and a surface plot of scale by position by variance shows the interaction between position in scale.

For the same elevation transect, the wavelet analysis reveals local peaks in the scalogram at approximately 3.5 and 14.5 km, and in this case the interpretation is similar to that of the local quadrat variance, except additional information is provided. In particular, the surface plot shows where along the transect the wavelet at each scale best fits the data; the alternating light and dark areas reflect the distribution of elevation peaks and valleys, respectively, at each scale.



Scaling Techniques for Continuous Data

■ Spectral Analysis

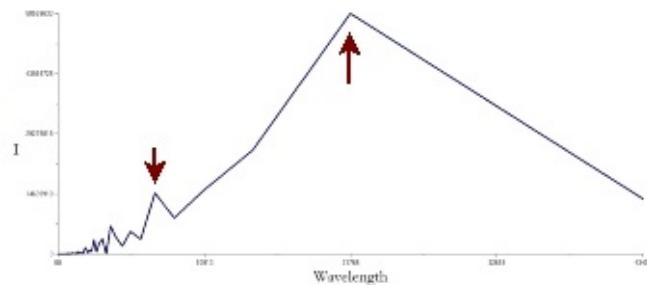


Spectral Fourier transformation:

$$y_i = \bar{y} + \sum_{p=1}^{(n/2)-1} c_p \cos\left(\frac{2\pi i p}{n}\right) + s_p \sin\left(\frac{2\pi i p}{n}\right)$$

Periodogram:

$$I_p = n(c_p^2 + s_p^2)$$



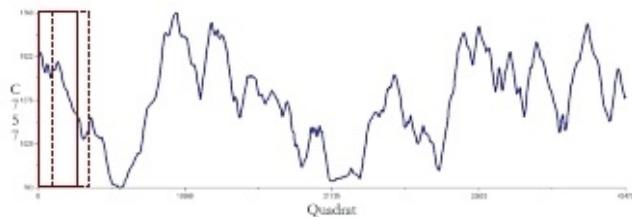
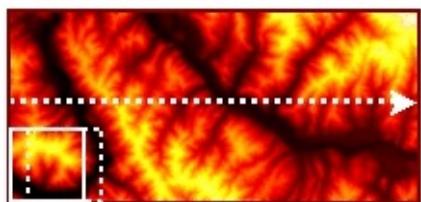
Spectral Analysis.—Spectral analysis describes spatial pattern by fitting sine and cosine waves to the data; the relative magnitudes of the coefficients of the fitted waves with different periods allows one to infer spatial scale. This is conceptually similar to TTLQV if you think of TTLQV as the fitting of a square wave to the data. In spectral analysis, the observed values are estimated by a linear combination of sine and cosine waves at each periodicity. The periodogram, I_p , is an estimate of the importance of the fitted waves of period p . Peaks in I_p reveal scales of pattern. The period is a measure of the frequency of the pattern across the data series. A measure of the scale of the pattern that is more equivalent to the block size measures from quadrat variance analysis is *wavelength*, measured as n/p . Because wavelength is the distance from peak to peak (or trough to trough), it is twice the value you would obtain from a quadrat variance analysis. A major limitation of spectral analysis is that it assumes evenly repeating patterns. Given that most ecological data is not regular, in practice the results can be very difficult to interpret compared to other methods.

For the same elevation transect, the periodogram (unsmoothed) reveals local peaks at approximately 7 and 22 km. Because the x-axis is given in wavelengths, these scales represent the distances between elevation peaks or between valleys and are roughly twice the distances given by the local quadrat variance and wavelet methods. In this case, because the topographic feature under consideration is somewhat regular in its pattern of variation, the periodogram is effective in revealing the dominant scales of pattern.



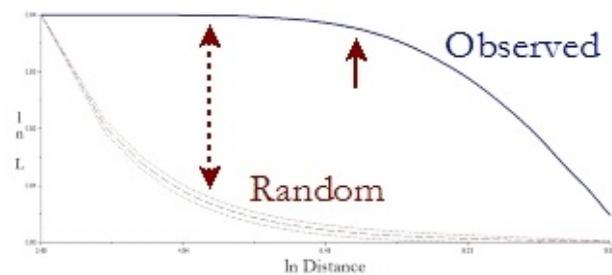
Scaling Techniques for Continuous Data

■ Lacunarity Analysis



Lacunarity:

$$\Lambda_r = \frac{s_{y(r)}^2}{\bar{y}_r^2} + 1$$



Lacunarity Analysis.—Lacunarity analysis derives from fractal mathematics and is a scale-dependent measure of heterogeneity or texture determined by the ratio of the variance to the square of the mean, determined within a gliding box of varying sizes (Plotnick et al.. 1993). Lacunarity was originally developed to examine binary data (discussed later), but was later extended for use on continuously distributed variables (Plotnick et al. 1996). Lacunarity for a specific range (scale) is estimated as the ratio of the first and second moments of counts within all possible boxes of that range width. The first moment is the average sum of values in all possible blocks of size r . The second moment is the average squared sum of values in all possible blocks of size r . The lacunarity for scale r is then measured as the ratio of the second moment to the square of the first moment, which is essentially equal to the variance to mean square ratio. Lacunarity is usually plotted on a log-log plot of lacunarity versus scale. The lacunarity plot can suggest scales of analysis at which the data appear random or scales over which the patterns exhibit self-similarity, and it can reveal hierarchical structure within data (Plotnick et al., 1996; but see Dale, 2000).

For the same elevation transect, the lacunarity plot reveals a strongly non-random pattern of variation at virtually all scales, which is not too surprising. Changes in the slope of the lacunarity curve suggest dominant scales of pattern. In this case, the curve begins to decline noticeably at a scale of roughly 7, which represents a distance of approximately 13 km because it is on a log scale (i.e., the x-axis represents the $\ln(\text{box size})$, and $\ln(13)=7.1$).



Scaling Techniques for Continuous Data

■ Complementary Analyses

Table 1

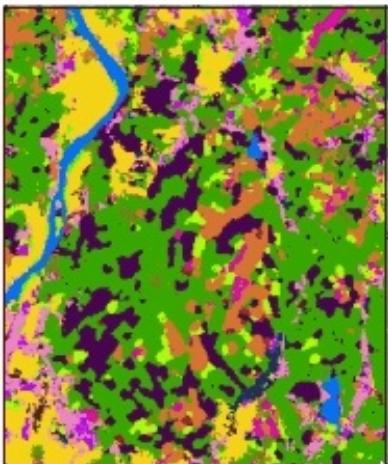
Comparison of characteristics of lacunarity, spectral and wavelet analysis for examining scales of pattern in ecological data

Characteristic	Lacunarity	Spectral	Wavelet
Scale or type of pattern best detected	Inconsistent (imprecise estimates ^a)	Repeating, cyclical, fine scale	Specific to analyzing wavelet
Type of data	Binary or continuous	Continuous	Continuous
Data dimensions	1 or 2 (also 3 ^b)	1 or 2	1 or 2
Missing data OK	Yes	No	No
Robust to nonstationarity	Yes	No	Yes
Sensitive to edges	No	Yes	Yes
Robust to nonnormality	Yes	No (may not be a concern)	Yes
Global summary	Yes	Yes	Yes (variance)
Retention of locational info	No	No	Yes (transform, position variance)
Detection of multiple scales	Yes?	Yes	Yes
Detection of hierarchical structure	Yes?	No	Yes
Hypothesis testing possible	Yes ^b	Yes (use of traditional inference methods, e.g., ANOVA ^c ; comparison to hypothesized distributions ^d)	Yes (use of bootstrapping, Monte Carlo techniques ^e)

Complementarity of Scaling Techniques for Continuous Surface Patterns

No one scaling technique has been found to be superior to all others for all applications. Each technique is based on a different approach and thus works better in some situations than others. While there are many potential criteria for comparing techniques, the table shown here from Saunders et al. (2005) lists a number of criteria used to differentiate among lacunarity analysis, spectral analysis and wavelet analysis for data collected continuously along transects or two-dimensional surfaces. As these authors concluded, the appropriate technique for assessing scales of pattern depends on the type of data available, the question being asked, and the detail of information desired. Moreover, they highlight the importance of: (1) using multiple techniques to examine scales of pattern in ecological data; (2) interpreting analysis results in concert with examination and ecological knowledge of the original data; and (3) utilizing results to direct subsequent descriptive and experimental examination of features or processes inducing scales of pattern.

Scaling Techniques for Categorical Data



- Data Characteristics
 - ▶ Data represent a region of a given area and can be intuitively defined by the notion of “edge” (a region is an area with edges adjoining other regions)
- Examples
 - ▶ Map of land cover types
 - ▶ Map of ownership parcels
- Goal of Analysis
 - ▶ To determine the scale(s) at which adjoining regions tend to have similar values

5.5. Scaling Techniques for Categorical Data

- Characteristics of categorical (or thematic or choropleth) map patterns:
 - ▶ Categorical map patterns represent data in which the system property of interest is represented as a mosaic of discrete patches, and can be intuitively defined by the notion of "edge" (a patch is an area with edges adjoining other patches). From an ecological perspective, patches represent relatively discrete areas of relatively homogeneous environmental conditions at a particular scale. The patch boundaries are distinguished by abrupt discontinuities (boundaries) in environmental character states from their surroundings of magnitudes that are relevant to the ecological phenomenon under.
 - ▶ A familiar example is a map of land cover types, where the data consists of polygons (vector) or grid cells (raster) classified into discrete classes. There are many methods for deriving a categorical map of patches. Patches may be classified and delineated qualitatively through visual interpretation of the data (e.g., delineating vegetation polygons through interpretation of aerial photographs), as is typically the case with vector maps constructed from digitized lines. Alternatively, with raster grids, quantitative information at each location may be used to classify cells into discrete classes and to delineate patches by outlining them, and there are a variety of methods for doing this.

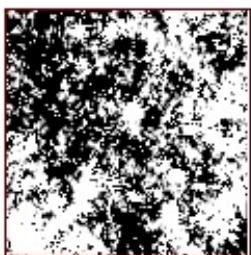
- ▶ Regardless of data format (raster or vector) and method of classifying and delineating patches, the goal of categorical map pattern analysis with such data is to characterize the composition and spatial configuration of the patch mosaic, and a plethora of metrics has been developed for this purpose.
- ▶ The scaling question with categorical map data tends to focus on the question, "At what scale(s) do adjoining regions tend to have similar values?"

While the field of geostatistics for continuous surface data is well developed, owing to its use in the geosciences, and quite familiar to landscape ecologists, scaling techniques for categorical map data are less commonly employed in landscape ecology. This may seem somewhat surprising given the predominant use of categorical data in landscape ecological investigations – after all, the predominant “landscape mosaic” model of landscape structure is based on a categorical data format. However, there has been a plethora of landscape metrics developed for quantifying various aspects of categorical map patterns and these have largely taken the place of the more conventional scaling techniques. In addition, in applications involving categorical map patterns, the relevant scale of the mosaic is often defined *a priori* based on the phenomenon under consideration. In such cases, it is usually assumed that it would be meaningless to determine the so-called characteristic scale of the mosaic after its construction. However, there are many situations when the categorical map is created through a purely objective classification procedure and the scaling properties of the patch mosaic is of great interest. We will explore one useful technique, lacunarity analysis.



Scaling Techniques for Categorical Data

■ Lacunarity Analysis



- Lacunarity represents the distribution of gap sizes.
- Geometric objects are more “lacunar” if gap sizes are distributed over a greater range.
- Lacunarity is a measure of the “gappiness” or “hole-iness” of a geometric structure.

High Lacunarity = heterogeneous objects with a wide range of gap sizes; not translationally invariant

Low Lacunarity = homogeneous objects with a narrow range of gap sizes; translationally invariant

Lacunarity Analysis

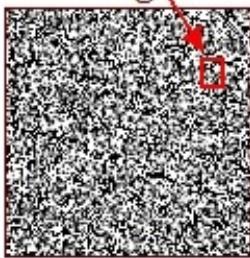
Lacunarity analysis holds a great deal of promise for detecting multiple scales of pattern in gridded (lattice) data and it is applicable to both categorical map patterns (Plotnick et al. 1993) and continuous surface patterns (Plotnick et al. 1996). Here we will consider its application to categorical maps. In this context, *lacunarity* represents the distribution of gap sizes. Geometric objects are more “lacunar” if gap sizes are distributed over a greater range. Lacunarity is a measure of the “gappiness” or “hole-iness” of a geometric structure. Heterogeneous objects with a wide range of gap sizes have high lacunarity. Such objects are not translationally invariant; that is, the pattern in terms of gap sizes is not uniform across the object. Thus, a window placed over one part of the object will reveal a different set of gap sizes than a window placed over another part of the object. Conversely, homogeneous objects with a narrow range of gap sizes have low lacunarity. Such objects are translationally invariant because a window placed over one part of the object will reveal about the size set of gap sizes as a window placed over any other part of the object.



Scaling Techniques for Categorical Data

■ Lacunarity Analysis

Gliding Box



Box Mass

	$S = 0$
	$S = 1$
	$S = 2$
	$S = 3$
	$S = 4$

Box Size

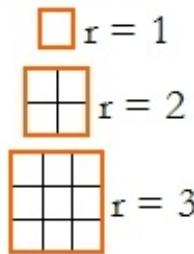


Table 1. Lacunarity calculation for 12×12 random map (Fig. 2a). The size of the gliding box is $r \times r$, with $r = 2$. S is the number of occupied sites or mass of the gliding box; $n(S,r)$ is the frequency of boxes of size r with mass S ; $Q(S,r)$ are the corresponding probabilities; $Z^{(1)} = \sum S Q(S,r)$ and $Z^{(2)} = \sum S^2 Q(S,r)$ are the first and second moments, respectively. The lacunarity, $\Lambda(r) = Z^{(2)}/(Z^{(1)})^2$.

	S	$n(S,r)$	$Q(S,r)$	$SQ(S,r)$	$S^2 Q(S,r)$
$r = 2$	0	3	0.024	0	0
	1	35	0.289	0.289	0.289
	2	46	0.380	0.760	1.520
	3	29	0.239	0.719	2.157
	4	4	0.066	0.264	1.057

$Z^{(1)} = 2.033$ $Z^{(2)} = 5.024$
 $\Lambda(2) = 1.215$

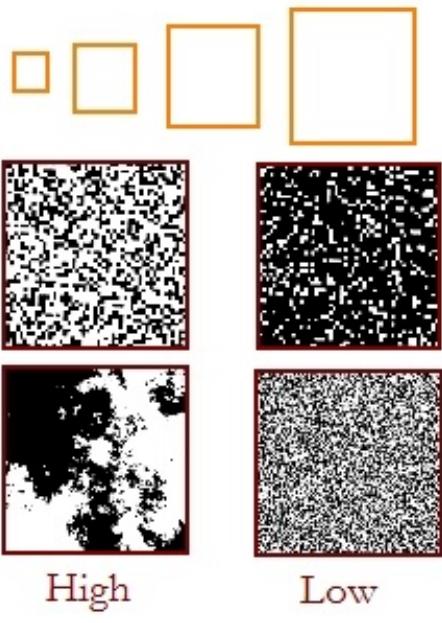
$$\Lambda(r) = \frac{S_s^2(r)}{\bar{S}^2(r)} + 1$$

Consider a binary (0 or 1) raster map of dimension m , where the 1's represent, say, habitat cells. Lacunarity analysis is concerned with the frequency with which one encounters habitat cells in a window of size r . To do this, a gliding box is constructed of dimension r , and a frequency distribution of tallies of $S=1,2, \dots, r^2$ habitat cells is constructed. In this gliding, the box is moved over one cell at a time, so that the boxes overlap during the sliding. Define $n(S,r)$ as the number of boxes that tally S habitat cells for a box of size r . There are $N(r)=(m-r+1)^2$ boxes of size r , and the probability associated with $n(S,r)$ is $Q(S,r)=n(S,r)/N(r)$. The 1st and 2nd moments of Q are used to define lacunarity, L (lambda), for this box size as a variance to mean-square ratio. This is repeated for a range of box sizes r .



Scaling Techniques for Categorical Data

■ Lacunarity Properties



■ Box size

- Proportion (p) of map occupied by the habitat of interest

- Geometry of the map (spatial contagion)

Lacunarity varies as a function of three things: box size, proportion of the map occupied by the habitat of interest, and the geometry of the map. Specifically, as box size increases lacunarity decreases monotonically – because the variation in habitat among sliding windows always decreases as the box increases in size relative to the landscape. As the proportion of the landscape occupied by habitat increases lacunarity decreases monotonically – because the variation in habitat among sliding windows of any size always decreases as landscape increasingly infills with habitat. Lastly, as the spatial contagion of the habitat decreases lacunarity decreases monotonically – because the variation in habitat among sliding windows always decreases as the habitat is more uniformly dispersed across the landscape. It is this latter property of lacunarity that is of most interest to us, because it provides a measure of scale-dependent contagion or aggregation of habitat. Essentially, highly fragmented habitat distributions at a particular scale (window size) have low lacunarity, whereas highly aggregated distributions have high lacunarity. By analyzing lacunarity across a range of scales we can summarize the degree of habitat fragmentation in relation to scale.



Scaling Techniques for Categorical Data

■ Lacunarity Properties

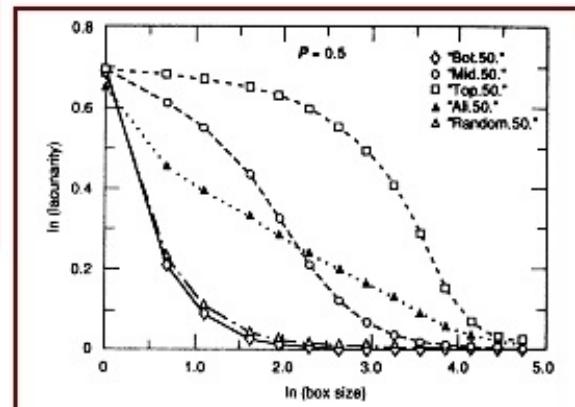
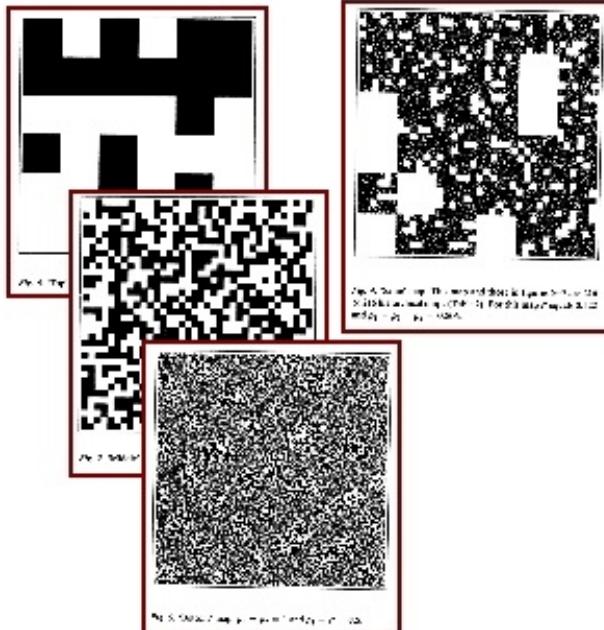


Fig. 8. Log-log plot of lacunarity versus gliding box size for maps with $P = 0.5$ (Figs. 3–7; Table 2).

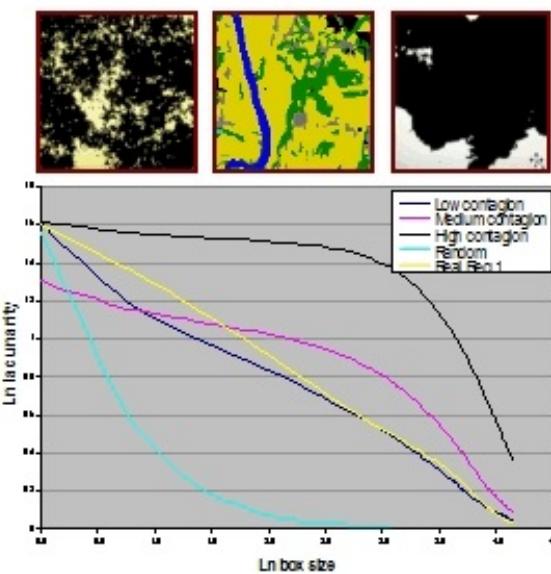
Lacunarity is summarized as a log-log plot of $L(r)$ versus r . Note that as r increases, the tally $n(S,r)$ converges on the mean and its variance converges on 0. Thus, for very large boxes $L=1$ and $\ln(L)=0$. At the smallest box size ($r=1$), $L(1) = 1/p$, where p is the proportion of the map occupied. Thus, for the finest resolution of a map, L depends solely on p . At intermediate scales, lacunarity expresses the contagion of the map, or its tendency to clump into discrete patches. So the lacunarity plot summarizes the texture of the map across all scales.

Lacunarity curves for patterned landscapes are particularly useful when plotted against the lacunarity curve for random distributions. As seen here, the lacunarity curve for a random distribution declines rapidly as box size increases and then asymptotically approaches zero. The curves for the hierarchically-structured artificial landscapes are generally much higher because the habitat is more clumped than random at most scales. The “top 50%” map (upper left), has the most aggregated or clumped distribution of habitat and its lacunarity curve remains higher than the others at all scales. Conversely, the “bottom 50%” map (lower left) is not much different than a random distribution and its lacunarity curve is almost indistinguishable from the random curve. Only at the largest box sizes do the lacunarity curves for the structured landscapes approach that of the random landscape. At these large box sizes, there is very little variation in habitat among the sliding windows regardless of the habitat distribution.



Scaling Techniques for Categorical Data

■ Lacunarity Properties



■ Lacunarity Graphs reveal:

- ▶ Overall fraction of map (P) occupied by the habitat type of interest.
- ▶ Level of contagion between occupied sites at a particular scale.
- ▶ Scale at which the map approximates a random pattern.
- ▶ Range of scales over which the map exhibits self-similarity.

Lacunarity graphs reveal several things about habitat pattern:

- ▶ Overall fraction of map occupied by the habitat type of interest. The Y-intercept is solely a function of P , with larger P 's resulting in lower lacunarity values.
- ▶ Level of contagion (clumping or aggregation) between occupied sites (habitat) at a particular scale. The greater the lacunarity the greater the contagion in habitat.
- ▶ Scale at which the map approximates a random pattern. When the lacunarity curve falls close to the random curve, the pattern is indistinguishable from random. Note, at the largest box size all curves will have lacunarity of zero. Thus, even highly structured landscapes will not be distinguishable from random at scales approaching the size of the landscape.
- ▶ Range of scales over which the map exhibits self-similarity. If the pattern exhibits perfect self-similarity (i.e., fractal properties), then the curve will be a straight diagonal line.

Lacunarity analysis has several advantages: 1) gliding box algorithm exhaustively samples the map to quantify changes in contagion and self-similarity with scale, 2) results are not sensitive to the boundary of the map, and 3) technique can be reliably used for the analysis of very sparsely occupied maps.

Scaling Techniques

■ General Rules and Reflections

- All techniques work best at finding regular patterns (i.e., repetitive or truly periodic variation); none of them work well at characterizing irregular patterns.
- All techniques require that you have measured the variable with sufficient resolution and extent to capture any patterns of interest ($\frac{1}{2}$ the finest-grain pattern you want, and an extent of 3-5 repetitions of a recurring pattern).
- None of the techniques is "best" and in fact they have complementary strengths and weaknesses (e.g., in finding fine- versus coarse-scale patterns).

5.6. General Rules and Reflections

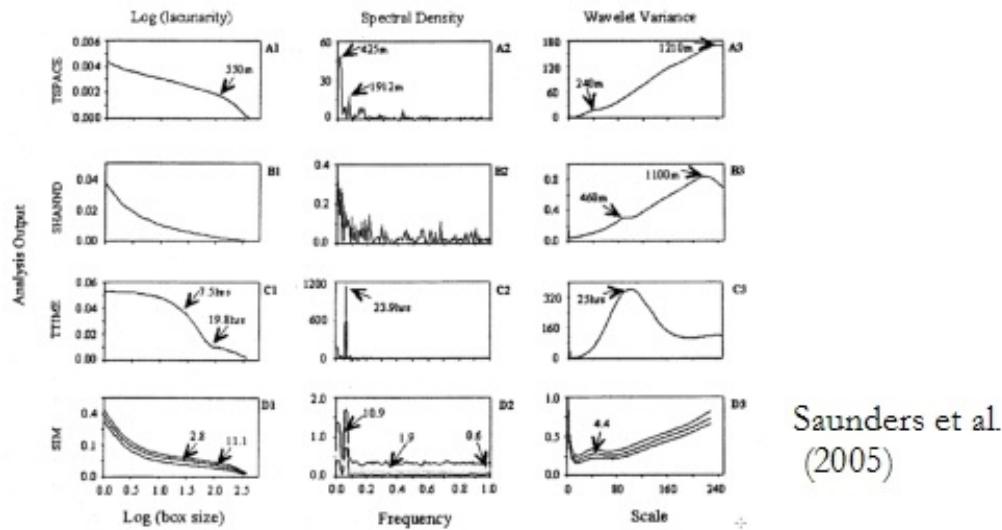
As even our cursory review indicates, there are myriad scaling techniques available depending on the type of data. They vary considerably in their numerical details but are remarkably similar conceptually. Generally:

- They all work best at finding regular patterns, i.e., repetitive or truly periodic variation; none of the them is very good at characterizing irregular pattern.
- Each technique requires that you've measured the variable with sufficient resolution and extent to capture any patterns of interest. Operationally, this means a resolution of $\frac{1}{2}$ of the finest-grain pattern you want, and an extent of 3-5 repetitions of a recurring pattern. This requires a lot of data and lends itself well to GIS and remote sensing technologies.
- None of the techniques is "best" and in fact they have what seem to be complementary strengths and weaknesses (e.g., in finding fine- versus coarse-scale patterns; none of them seem to be good at both). So many investigators recommend using more than one technique and interpreting only those patterns (scales) that appear consistently important.

Discovering scaled pattern in data, of course, is the first step in a process of understanding landscape pattern, and understanding of landscape pattern is considered prerequisite to understanding pattern-process relationships.

Dealing with Scale

- Employ scaling techniques to identify the characteristic scale(s) of the ecological phenomena or to elucidate the nature of the scaling relationship.

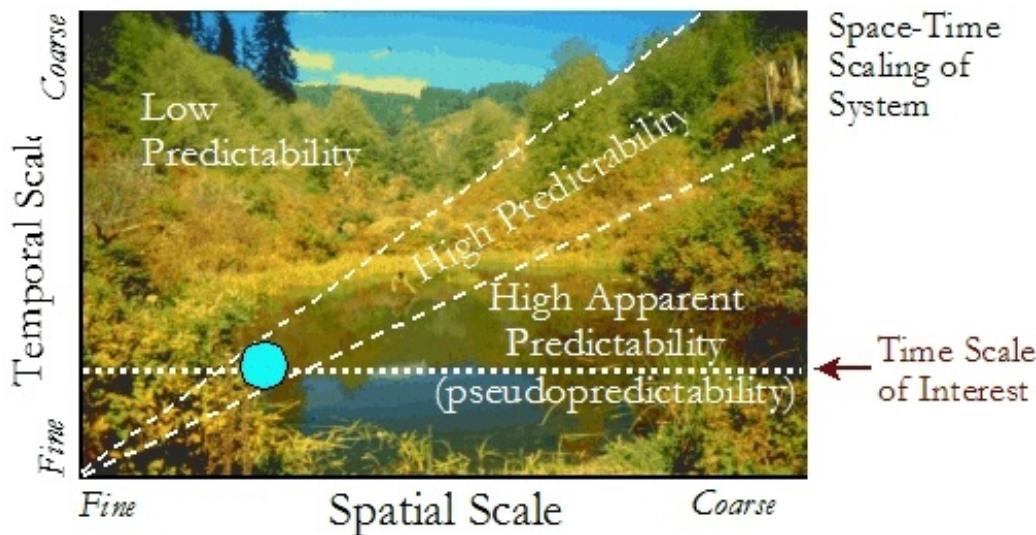


6. Dealing with Scale

- Employ scaling techniques (as discussed above) to identify the characteristic scale or scales of the ecological phenomena or to elucidate the nature of the scaling relationship.—The ability to characterize pattern is generally considered prerequisite to understanding pattern-process relationships. And one thing that is true of pattern, as we have already seen, is that its expression varies with scale. Thus, a careful evaluation of the relationship between scale and pattern is an important first step in the characterization of pattern. In addition, quantifying the scaling relationship of the ecological phenomenon of interest can provide unique insights into potential pattern-process relationships. For example, multi-scaled patterns (i.e., those with two or more distinctive characteristic scales) revealed through scaling techniques can indicate the existence of multiple agents of pattern formation operating at different scales. Conversely, scaling techniques can be used to assess a priori hypotheses concerning agents of pattern formation; e.g., is the expected scale or scales of pattern expressed in the data?

Dealing with Scale

- Recognize the relationship between space and time scaling of ecological phenomena and scale investigations accordingly.

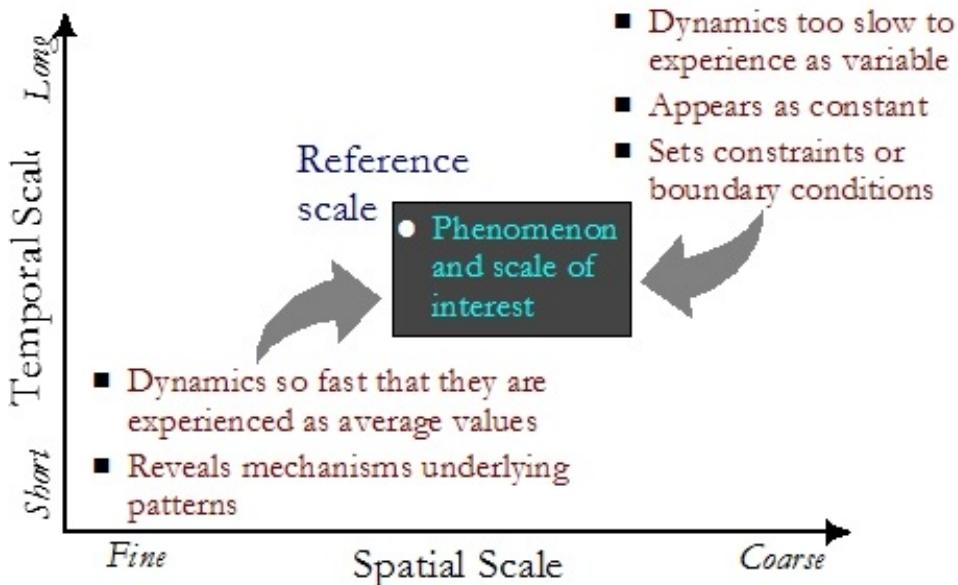


- Recognize the relationship between space and time scaling of ecological phenomena and scale investigations accordingly.—Our ability to predict ecological phenomena depends on the relationships between spatial and temporal scales of variation. With increased spatial scale, the time scale of important processes also increases because processes operate at slower rates, time lags increase, and indirect effects become increasingly important. Consequently, any predictions of the dynamics of spatially broad-scale systems that do not expand the temporal scale are “pseudopredictions”. The predictions may seem to be quite robust because they are made on a fine time scale relative to the actual dynamics of the system, but the mechanistic linkages will not be seen because the temporal extent of the study is too short.

For example, if we seek to predict the occurrence of active beaver ponds in a small drainage based on their presence during the past few years, our predictions may appear to be quite robust, i.e, our prediction that there will be active beaver ponds in the drainage during the next year will be right most of the time. However, these predictions would eventually prove to be wrong since the underlying dynamic of beaver occupancy in small drainages occurs over decadal (or longer) time scales. Thus, if we study a system at an inappropriate scale, we may not detect its actual dynamics and patterns, but may instead identify patterns that are artifacts of scale. Because we are clever at devising explanations of what we observe, we may think we understand the system when we have not even observed it correctly. This is sometimes referred to as a “type III error” – asking the right question at the wrong scale.

Dealing with Scale

- Adopt a hierarchical framework to help focus on scale.



- Adopt a hierarchical framework to provide focus on scale.—Hierarchy theory is based on a simple premise: that ecosystems do not exist in isolation; they are nested within larger ecosystems that influence processes occurring within the system. These larger ecosystems are nested within even larger ecosystems, that are nested within even larger ecosystems, and so on. In other words, each ecosystem has a context, and this context may constrain processes operating within the ecosystem.

Hierarchically organized systems can be divided, or decomposed, into relatively discrete functional components operating at different scales where, in general, lower-level units interact to generate higher-level behaviors and higher-level units control those at lower levels. Hierarchical organization typically results from differences in process rates.

Ecosystems operate over a wide spectrum of rates. Behaviors can be grouped into classes with similar rates that constitute levels in the hierarchy. Phenomena on a given level will be relatively isolated from lower and higher levels due to differences in process rates. An important outcome of this hierarchical organization is that patterns and processes at one scale provide a context that constrain the behaviors of the lower levels of the hierarchy.

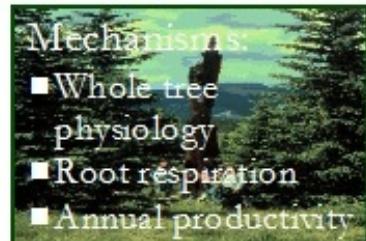
Of course, natural phenomena often are not perfectly decomposable: spatial boundaries may be difficult to define precisely and components may interact. Yet many complex, natural phenomena are nearly decomposable and thus can be conceptualized usefully as hierarchical systems.

Dealing with Scale

- Adopt a hierarchical framework to provide focus on scale.

Objective: To predict the increase in biomass of a forest stand at a time scale of approximately 100 yr.

Tree/gap



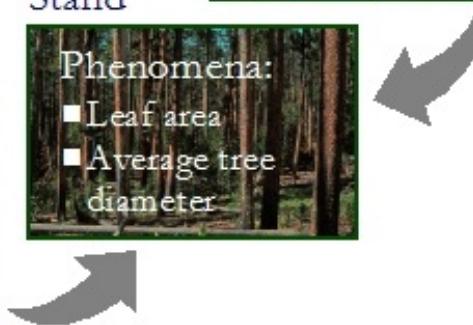
Stand



Landscape

Constraints:

- Pedogenesis of the soil
- Evolutionary responses of the trees



Casting the relationships in the context of hierarchy theory sharpens our focus on scaling by emphasizing logical and functional linkages among scales, specifically:

- ▶ The scale of an event defines the observational level (reference or focal level) through which the rest of the hierarchy is accessed.
- ▶ The next lower level (finer scale) provides the components of the event and mechanistic explanation of the dynamics of the focal level. That is, the next lower level becomes the “part” whose functioning and interactions provide “mechanisms” that explain the phenomenon of interest.
- ▶ Signals from still lower levels are so attenuated that they appear as background noise that can be ignored.
- ▶ Higher levels (coarser scale) provide the context that gives the event greater significance. Higher levels impose constraints or boundary conditions; that is, the level of interest is constrained to operate within the bounds set by the higher-level system of which it is a part. The next higher level becomes the “environment.” These higher-level factors can be treated as constants when viewed from the reference level, though they may be quite variable at larger scales.

- ▶ Levels of the hierarchy further removed from the reference level are not of immediate concern, in that they contribute very little toward understanding the event of interest.

The usefulness of a hierarchical framework to provide focus on scale is best demonstrated by example.

Example 1.—Consider the following objective: Predict the increase in biomass of a forest stand at a time scale of approximately 100 yr. Hierarchy theory provides a useful framework for dealing with this problem. The reference phenomena is biomass increase and the focal scale is the forest stand (space) and its development over 100 years (time). Appropriate state variables include leaf area (LAI) and average tree diameter. The hierarchical framework suggests that we consider patterns and processes operating at higher levels (coarser scales) and lower levels (finer scales) to gain a more complete understanding of the phenomena – and therefore better predictions.

At a higher level, stand development is influenced by patterns and processes operating at much coarser scales in space and time. For example, pedogenesis of the soil and the evolutionary response of the trees operate at much coarser scales (in space and time) and provide constraints on stand productivity – we must understand these constraints to accurately predict stand biomass. Soil productivity, for example, is influenced by processes occurring on geologic time scales, and for all practical purposes can be seen as a constant at the scale of the individual forest stand over a period of 100 years. Stand development is limited by the existing and relatively constant (from the perspective of a 100 yr time frame) productivity of the soils. Indeed, this is the basis for the concept of “site index”, is it not?

At a lower level, stand development is ultimately determined by patterns and processes affecting individual trees, since trees are the units that actually accumulate biomass. For example, individual tree energy budgets as determined by whole tree physiology, root respiration and annual productivity determine how fast a tree grows. The collective behavior of the individual trees determines the behavior of the stand, and we cannot fully understand stand development without understanding the factors affecting individual tree growth.

Ultimately, the hierarchical framework forces us to explicitly consider scale and identify the relevant patterns and processes at each scale affecting the reference phenomena.

Dealing with Scale

- Useful hierarchical frameworks?

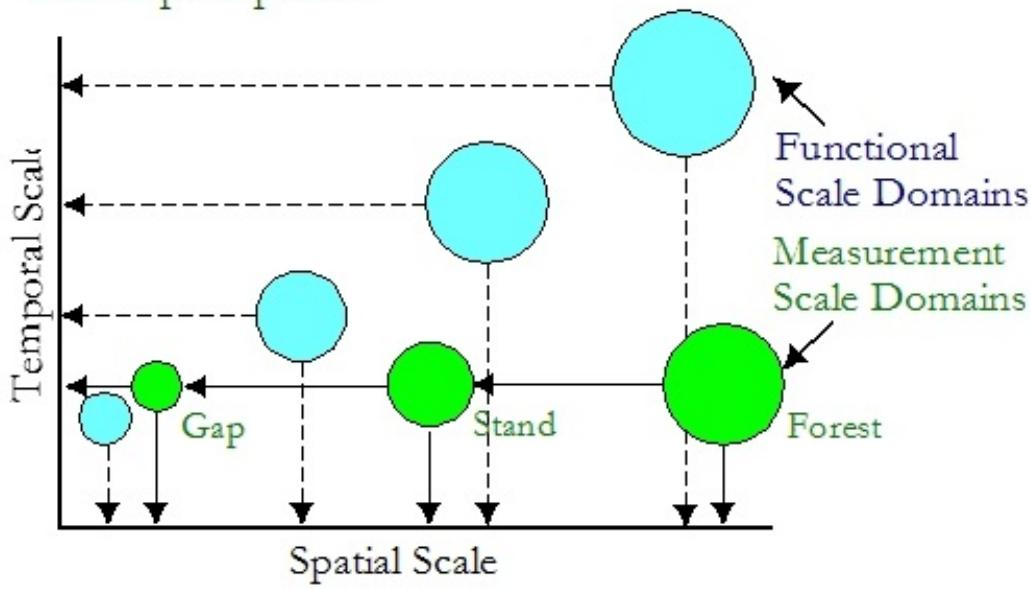
- River hydrology...
- Metapopulation dynamics...
- Disturbances...
- Ecological classification...
- Others?



There are numerous other examples of useful hierarchies in resource management – some are listed here. Can you think of others? Note, often we construct an intuitive hierarchy to explain the patterns and processes of interest without formally considering hierarchy theory per se. In practice, this may be quite sufficient. However, greater attention to the principles of hierarchy theory may stimulate us to observe important factors not revealed initially. And at the very least, adoption of a hierarchical framework forces us to make scaling issues explicit.

Dealing with Scale

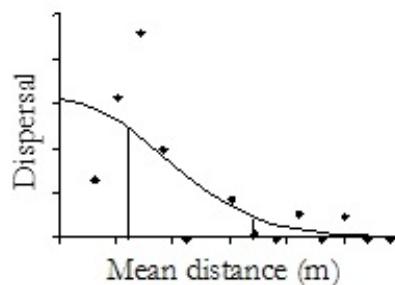
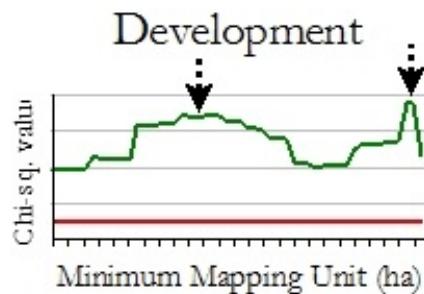
- Use ecological data to determine the functionally relevant scale(s) instead of arbitrary scales based on human perceptions.



- Use ecological data to determine the functionally relevant scale(s).—The scales chosen for analysis are often arbitrary; they tend to reflect hierarchies of spatial scales that are based on our own perceptions of nature. Just because these particular scales seem right to us is no assurance that they are appropriate to the organisms and/or processes under consideration. We need nonarbitrary, operational ways of defining and detecting scales. Our ability to predict ecological phenomena depends on our observing them at the functionally relevant scale(s).

Dealing with Scale

- Analyze pattern-process relationships across a range of scales and empirically determine the functionally relevant scale or scales.
- Use biological data such as frequency distributions for the space and time components of the behavior of organisms to determine the correct scale of analysis.



A variety of approaches have been used to identify the “correct” scale from the perspective of the organism or phenomenon under consideration.

- The most common approach is to analyze pattern-process relationships across a range of scales and empirically determine the functionally relevant scale or scales.
- Use biological data such as frequency distributions for the space and time components of the behavior of organisms to determine the correct scale of analysis; in other words, employ the concept of ecological neighborhoods (Addicott et al. 1987).

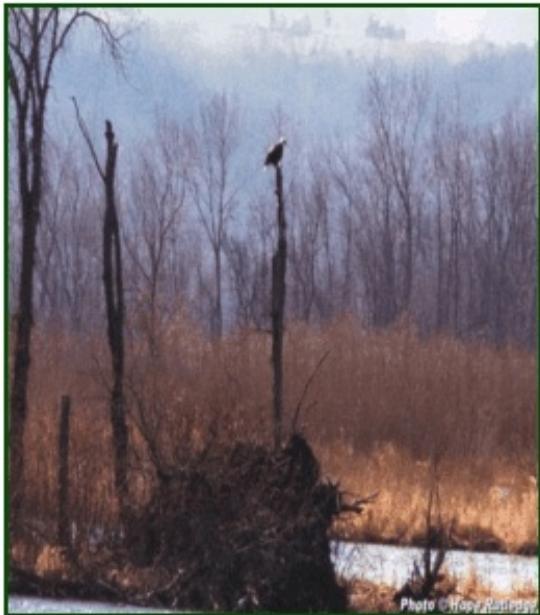
These approaches are best illustrated by examples:

Thompson and McGarigal. 2002. The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). *Landscape Ecology* 17:569-586.



This example involves habitat selection by bald eagles on the lower Hudson River, New York (Thompson and McGarigal 2002). We studied the influence of research scale (grain and extent) on bald eagle habitat selection and sought to define landscape patterns from an eagle's perspective. Here we measured habitat patterns at a range of spatial scales and let the eagles determine which scale or scales were most relevant; i.e., which scale or scales they perceived and responded to the most.

Thompson and McGarigal (2002) . . .

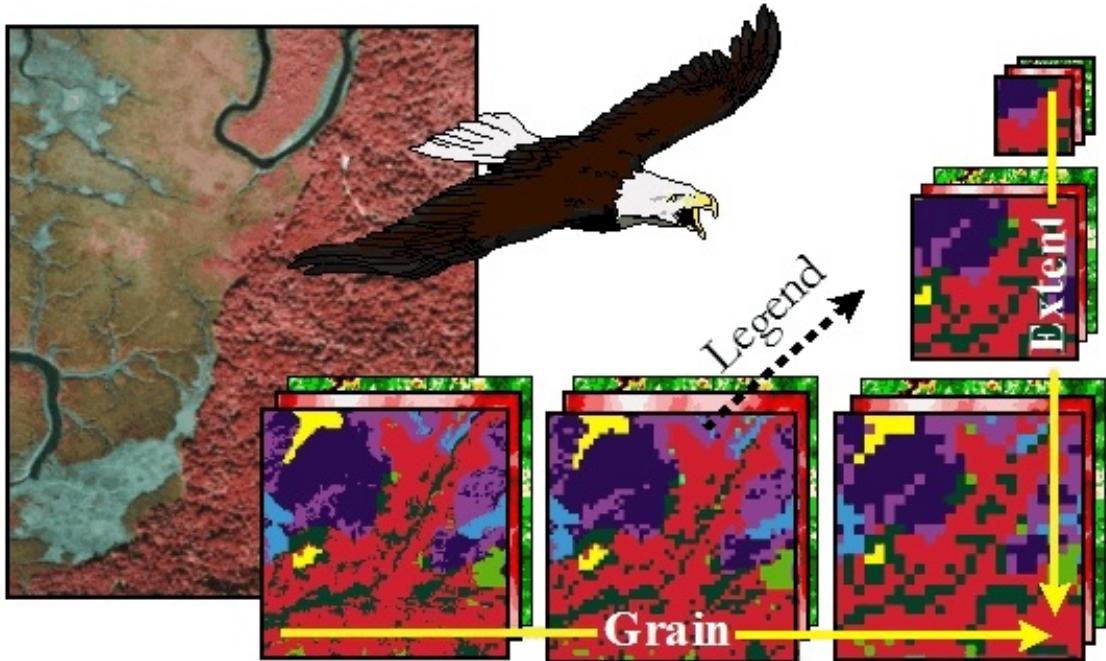


Bald eagle habitat requirements:

- Access to foraging areas
 - Water depth
- Available perch trees
 - Canopy structure
- Isolation
 - Human activity
 - Shoreline development

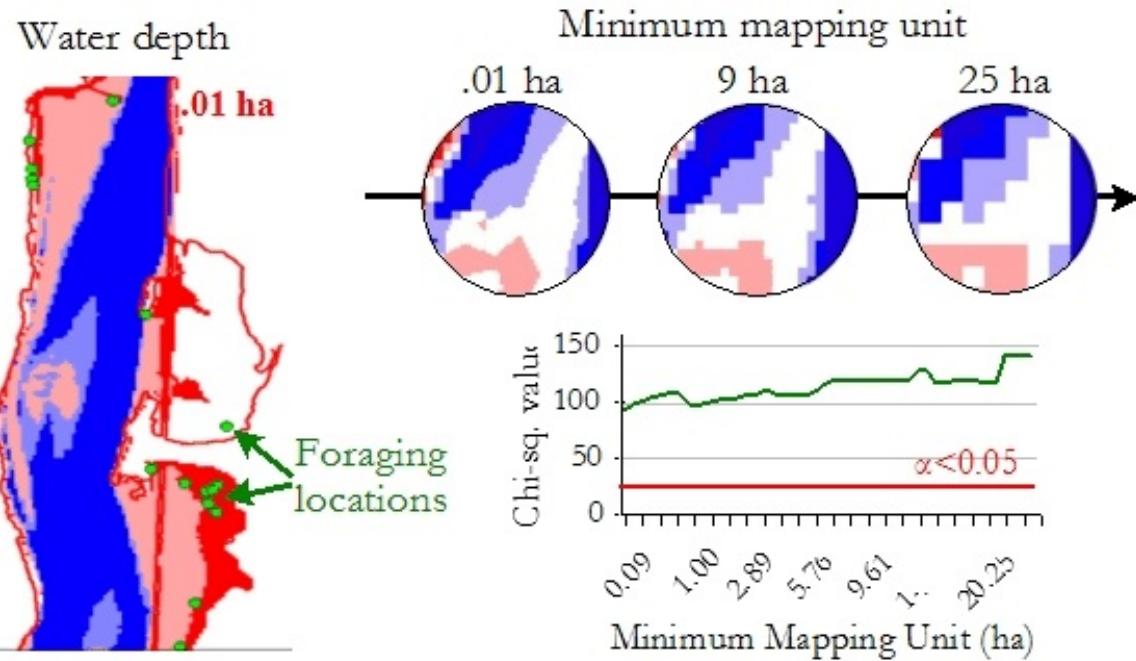
First, it is important to recognize that the bald eagle is perhaps the most studied bird in North America and much is known about their life history and habitat requirements. Indeed, there is consensus among eagle experts that the principal habitat components are: (1) access to foraging areas - represented by water depth in tidal systems, (2) available perch and nest trees - represented by canopy structure, and (3) isolation from human disturbance - represented by freedom from human activities and shoreline development.

Thompson and McGarigal (2002)....



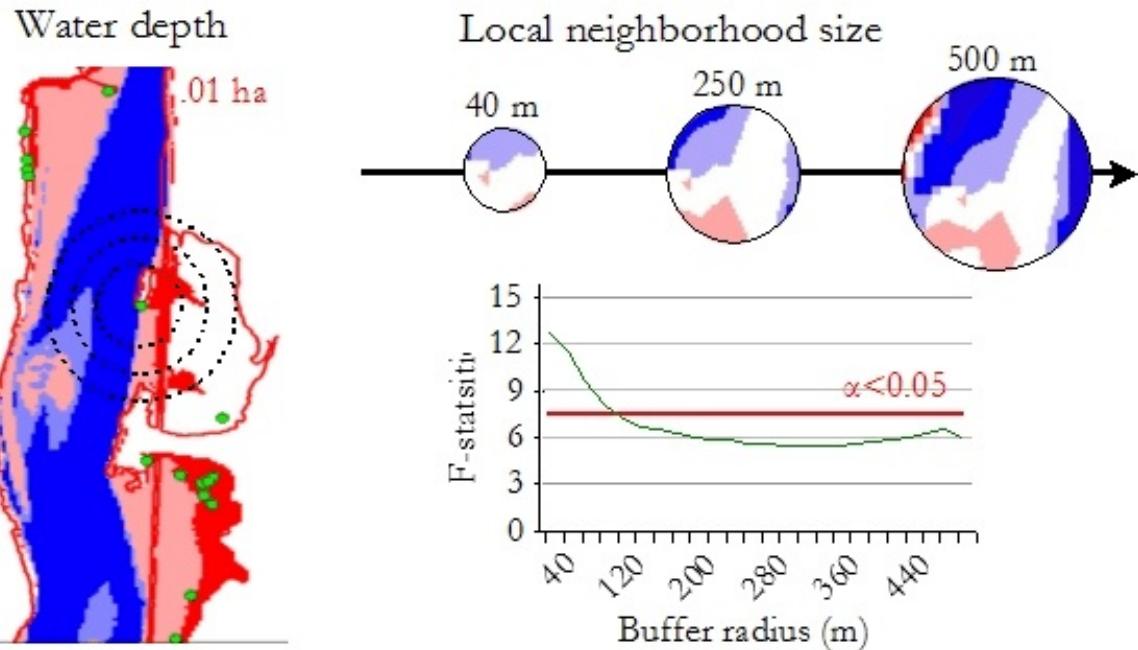
Consequently, we were able to create a meaningful map legend for each of these habitat components. Unfortunately, we had no idea what the appropriate spatial scale should be for representing each of these components. For example, we knew that eagle's preferred to forage on tidal mudflats, but we had no idea whether it was important to represent every tiny little patch of mudflat or only large mudflats. Similarly, we had no idea over what local spatial extent eagle's perceive and respond to habitat patterns.

Thompson and McGarigal (2002) . . .



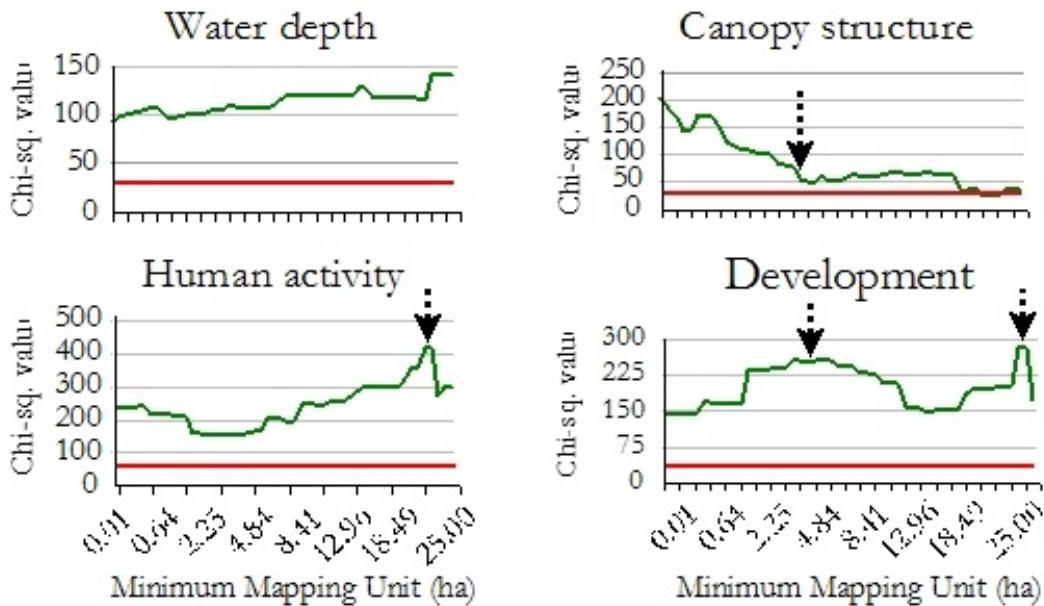
So, for each habitat component we systematically varied the minimum mapping unit (i.e., grain) to create a gradient from very fine- to very coarse-grained maps, and then used statistical procedures to assess habitat use versus availability at each scale. We then plotted the test statistic in relation to scale and looked for peaks which would be suggestive of the scale eagle's were most responsive to, and hence the "best" scale for mapping that habitat component.

Thompson and McGarigal (2002) . . .



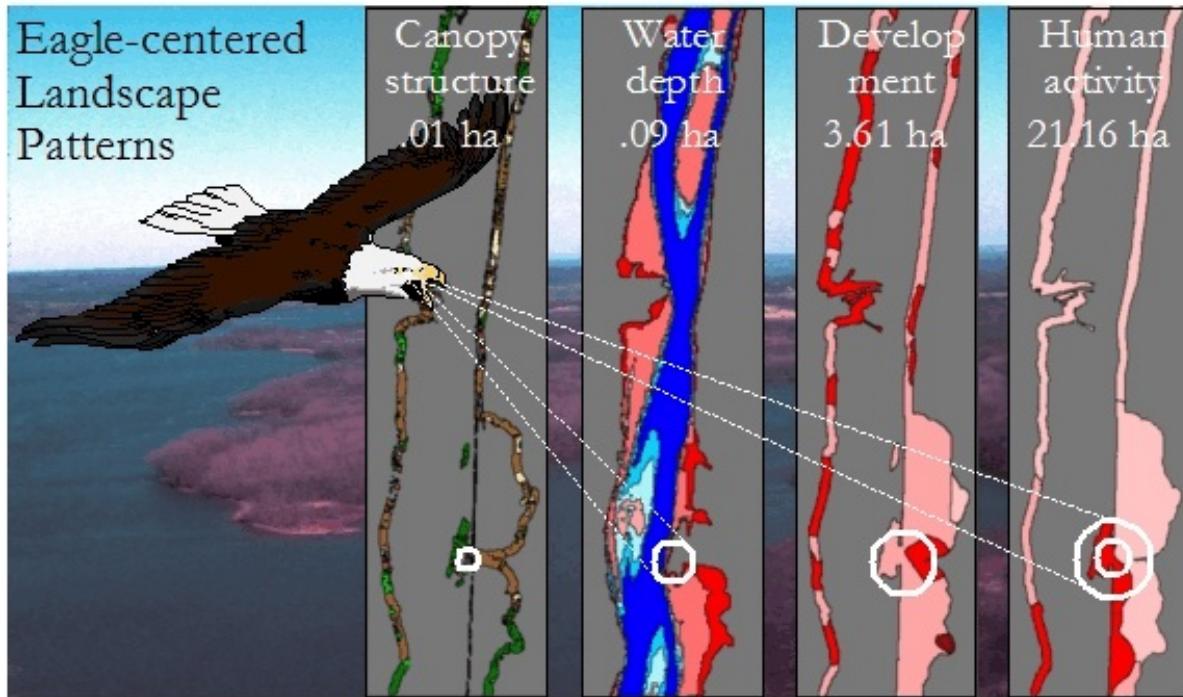
Using the "best" minimum mapping unit scale (i.e., grain) for each habitat component, we then systematically varied the local extent (or window size) and used statistical procedures to identify the local ecological neighborhood size in which to assess habitat use versus availability on the basis of local habitat patterns. The combination of best grain size and best local extent was deemed the scale or scales at which eagle's most strongly respond to habitat patterns, and thus the appropriate scale or scales for assessing habitat use.

Thompson and McGarigal (2002) . . .



In a nutshell, we found a variety of scaling relationships, including (1) strong selection at all scales associated with water depth, (2) threshold-like behavior - in this case, strong selection below a certain scale, associated with canopy structure (3) a single dominant scale associated with human activity, and (4) multi-scale selection associate with shoreline development. Among other things, these results indicated that a single scale was insufficient for representing eagle habitat.

Thompson and McGarigal (2002) . . .

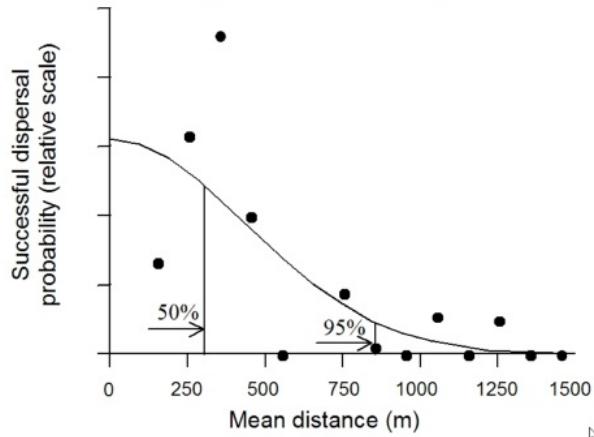


Ultimately, using this empirical approach we were able to define and map the landscape and measure landscape patterns in several dimensions corresponding to those presumably most strongly related to how eagle's perceive and respond to landscape structure. These eagle-defined maps then provide the basis for habitat capability modeling, among other things. For our purposes here, an important conclusion of this study is that statistical relationships can change with scale; i.e., that scale matters.

Compton, B.W., K. McGarigal, S.A. Cushman, and L.R. Gamble. 2007. A resistant-kernel model of connectivity for amphibians that breed in vernal pools. *Conservation Biology* 21:788-799.



- Direct observation of ecological neighborhood size for focal process of dispersal.



This example involves multi-scale modeling of landscape connectivity for vernal pool Ambystomatid salamanders in Massachusetts (Compton et al. 2007). We constructed a model to assess connectivity at three ecological scales: local breeding pool and surrounding upland (migration), clusters of neighboring pools and intervening uplands (dispersal), and broader regional clusters of pools (gene flow), and applied to nearly 30,000 potential vernal pools across Massachusetts to help identify those that have intact upland habitat and are highly connected across the landscape for ground-truthing and conservation action. To determine the appropriate spatial scale for modeling migration and dispersal processes, we directly observed individual salamanders and quantified the frequency distributions of movements. For example, to quantify the scale of dispersal, we monitored breeding populations of marbled salamanders (*Ambystoma opacum*) among 14 vernal pools in western Massachusetts over a 7 year period. The distribution of observed dispersal distances allowed us to characterize the size of the dispersal landscape around a vernal pool. Thus, in this study, rather than choosing scales arbitrarily, we used direct observations of the focal organism to determine the correct scale for analysis and management. A useful concept embodied in this approach is that of “ecological neighborhoods” introduced by Addicott et al. (1987). An ecological neighborhood is an area scaled to a particular ecological process, a time period, and an organism’s mobility or activity.

Why is Scale Important?

- THUS, explanatory models are scale dependent and what is an appropriate scale depends on the question one asks and the phenomenon under consideration.



- Ultimately, what is an appropriate scale(s) depends on the questions one asks and the phenomenon under consideration. What is a resource patch to one organism is not necessarily to another. A beetle does not perceive and respond to the environment at the same scale as a hawk; thus, the relevant landscape pattern-process relationships operate at fundamentally different scales. This is the most important take-home message.