Available online at www.sciencedirect.com
SCIENCE DIRECT®

Ecological Complexity 2 (2005) 339–356

ECOLOGICAL COMPLEXITY

http://www.elsevier.com/locate/ecocom

Heterogeneity versus homogeneity: A conceptual and mathematical theory in terms of scale-invariant and scale-covariant distributions

Rusty A. Feagin *

Spatial Sciences Laboratory, Texas Agricultural Experiment Station and Department of Forest Science, Texas A&M University, 2135 TAMU, College Station, TX 77843-2135, USA

Received 2 December 2004; received in revised form 5 April 2005; accepted 8 April 2005
Available online 23 June 2005

Abstract

Ecologists often use the words heterogeneity and homogeneity when they are describing the distribution of a variable, yet there has been no formal exploration into the relation of these two states to one another. This work formalizes their relationship within statistical theory with a conceptual framework that includes five parameters: L, K, S, D, and R. These parameters provide an increasing degree of specificity about a distribution as they are enumerated and accord with measurements of system extent, richness, evenness, variance, and scale-covariance. Moreover, a mathematical and physical basis for understanding heterogeneity and homogeneity is outlined in terms of Brownian motion, where the Hurst exponent H and the notion of variance are utilized to delimit these two states. Spatial and temporal patterns are quantified and classified according to the mathematical and physical basis at multiple scales, as well as within the conceptual framework that can be applied to other metrics. Concepts of scale-invariance and scale-covariance are discussed in terms of hierarchy theory. It is argued that in hierarchical distributions, we should expect more than simple scale-covariance; we should expect division by a transitional state that divides heterogeneity from homogeneity at some scale. By revisiting the underlying statistical theory behind these concepts, a more efficient approach to quantifying ecological distributions can result.

Keywords: Ecological complexity; Scale-covariance; Scale-invariance; Lacunarity; Hierarchy theory; Scaling

1. Introduction

The nature of variation has been crucial to methodological procedures in ecology, as evidenced by the synergism between early statistical and ecological work such as that between the origins of ANOVA methods and those of genetic analysis (Fisher, 1918). Statistics has provided a framework for the understanding of variability and set categorization, the importance of which has not been ignored by those who have investigated the theory of ecotones (Gosz, 1993; Kolasa and Zalewski, 1995; Cadenasso et al., 2003a,b), discrete versus continuous patch

* Tel.: +1 979 862 2612; fax: +1 979 862 2607.
E-mail address: feaginr@tamu.edu.

1476-945X/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecocom.2005.04.009

boundaries (Ludwig and Cornelius, 1987; Cornelius and Reynolds, 1991; Fortin, 1994; Fortin et al., 2000), the delineation of ecosystems (van der Maarel, 1976), and the practice of parsing variance components for community composition analyses (Braun-Blanquet, 1932; Orloci and Beshir, 1976; Barkman, 1989). Yet, there has been no systematic exploration at the most basic level into how statistical measures of variation can be incorporated into the same conceptual framework as measures of heterogeneity, homogeneity, richness, evenness, Shannon entropy, and scaling.

Complexity theory and its influence upon ecological research (Loehle, 2004), has only further illuminated the importance of variance in the study of multi-scaled structural (Sole and Manrubia, 1995; Keitt and Marquet, 1996; Ricotta et al., 1999) and functional (With and Crist, 1995; Milne et al., 1996) distributions, particularly where scale-invariance (Stanley et al., 2000; Allen et al., 2001; Schneider, 2001) and scale-covariance (Allen and Holling, 2002; Li, 2002) exist. Already, much work has been done in quantifying the covariance at multiple scales (Moellering and Tobler, 1972; Hulbert, 1990; Anand and Li, 2001) of a variable, such as vegetative cover (Li, 2000) or genetic type (Kareiva, 1990; Epperson, 2003; Rauch et al., 2003; Maurer, 2005) in ecological data sets, yet these efforts have not yielded a singular mathematical basis for heterogeneity and homogeneity. Heterogeneity and homogeneity are simply words that are often used in this relative sense, but undoubtedly convey a quantitative meaning about the distribution of a variable.

In hierarchy theory (Allen and Starr, 1982), it has been a tradition to quantify the variance of a distribution across multiple scales, and then define levels as discontinuous breaks in the power-law structure of the variance (O'Neill et al., 1991; Gardner, 1998). However, there has been no conceptual or mathematical justification as to why we follow this tradition. Nor has there been an investigation into how our quantification of those objects that are relatively the ''same'' and those that are relatively ''different'' reconcile with notions of heterogeneity and homogeneity.

The objective of this work is to return to the basics for discerning ''same'' from ''different'' and to codify the relevant notions of heterogeneity and homogeneity. We will pursue this objective in two different veins, while using examples from ecological distributions. First, we will compartmentalize the various components that contribute to the perception of heterogeneity into discrete functions: L, K, S, D, and R. Second, we will present the mathematical and physical logic for dividing heterogeneity from homogeneity. We will then illustrate these both of these notions visually and discuss the importance of heterogeneity and homogeneity in defining hierarchical structure.

1.1. Heterogeneity and homogeneity: intuitive or quantitative?

Heterogeneity and homogeneity are concepts that have been a part of much ecological research in recent years. The common understandings behind their usage imply something about the distribution of a variable. For a variable distributed across space, it might be qualitatively patchy (Forman, 1997; Turner et al., 2001) or quantitatively defined (Li and Reynolds, 1995) by an index such as contagion (Li and Reynolds, 1994), lacunarity (Plotnick et al., 1993), or wavelet analysis (Bradshaw and Spies, 1992; Li and Loehle, 1995; Saunders et al., 2005) among many others (O'Neill et al., 1988). For a variable distributed across time, or even any type of variable that does not cross a vector, one could also describe it as heterogeneous or homogeneous based upon its distribution.

The variability of a distribution, and thus the shape of a distribution, has a lot to do with our interpretation of these concepts. For example, the frequency distribution in Fig. 1a intuits heterogeneity and Fig. 1b intuits homogeneity.

''Hetero'' is ''different'' or ''other'' in Greek, implying that a difference or stark change occurs across the distribution of a studied variable. For example, a close-up digital photo shows two land cover types along a shoreline with a sharp division between the two: water and land. Or, as an ecological researcher samples for a particular genetic type while walking along a transect that straddles two sides of a mountain range, the frequency of the possible alleles abruptly changes between one population and another population. Fig. 1a could represent either of these examples, with the ordinate axis representing the two categories of classified pixel types, or the two different categories of alleles found.

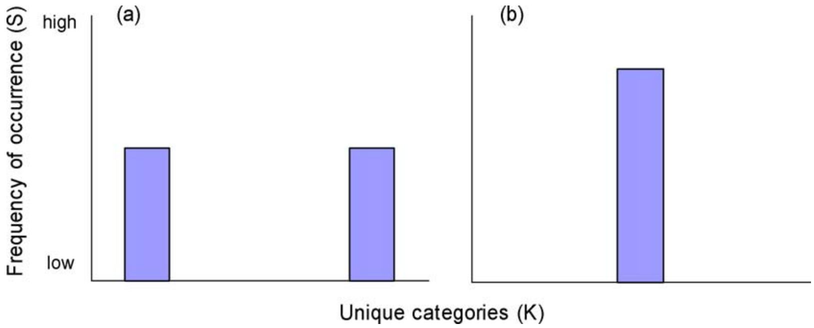

Fig. 1. Heterogeneous vs. homogeneous distributions at the simplest conceptual level. If the total number of samples, L, is 256 and there are only two possible categories, K, then the limit-heterogeneous distribution (a) would be represented by 128 samples in each of the two categories. The limit-homogeneous distribution (b) would be represented by all 256 samples being similar. The possible difference in the set, D, can be calculated with the quantities L, K, and S (see text). Abscissa is the number of unique categories, K, and ordinate is the frequency of each category, S.

''Homo'' is ''same'' in Greek, with the implication that a studied variable is in some way stationary across its distribution. For example, an aerial photo of the same shoreline area that has been zoomed-out shows one cover type within its extent: saltmarsh. Or, the frequency of the genetic type does not markedly change as the researcher walks across the precipice of the mountain, signaling that there is only one population. Fig. 1b would be representative.

An important theoretical question then presents itself: how do we divide these two states, heterogeneity and homogeneity? For example, the photos of the shoreline may appear to be heterogeneously distributed water and land from close-up, homogeneously distributed saltmarsh from far away, but what about in between? At what point does the crossover occur and what does the distribution of pixels look like at that scale? What happens to our notions of categorization in such a situation? Similarly, when does the mountain-climbing researcher choose to define when she/he has found one population or two populations? What would the distribution look like at the point where neither option takes precedence?

2. Conceptual theory

2.1. Vector: L

A conceptual basis for categorizing the variability of a distribution into either heterogeneity or homogeneity may require several dimensions. We may further clarify this idea with notation. Given the distribution of a variable along a vector L, each sample becomes l. This vector may be spatial, temporal, or of another type.

2.2. Richness: K

Along L, there may be K different forms of the variable: $k_1, k_2, k_3, \ldots, k_{max}$. With this information, we can calculate the richness of the distribution as K. We can also compute a measure of relative richness, where the measure allows comparison between distributions with different L as:

$$\text{relative richness} = \frac{K}{L} \tag{1}$$

Many other standardized measures of richness exist in ecology (Gotelli and Colwell, 2001), but all are based upon K in the numerator.

2.3. Evenness: S

Moreover, each K form may occur S times along the distribution: $s_1, s_2, s_2, \ldots, s_{max}$. With this information, we can calculate Shannon entropy by deriving the probability that any one K will be encountered when sampling the length L, as p_i = S/L for each K, and then plugging each p_i into the common formula:

$$\text{entropy} = -\sum p_i \log(p_i). \tag{2}$$

Fig. 2a shows an example, where L, K, and S have been calculated.

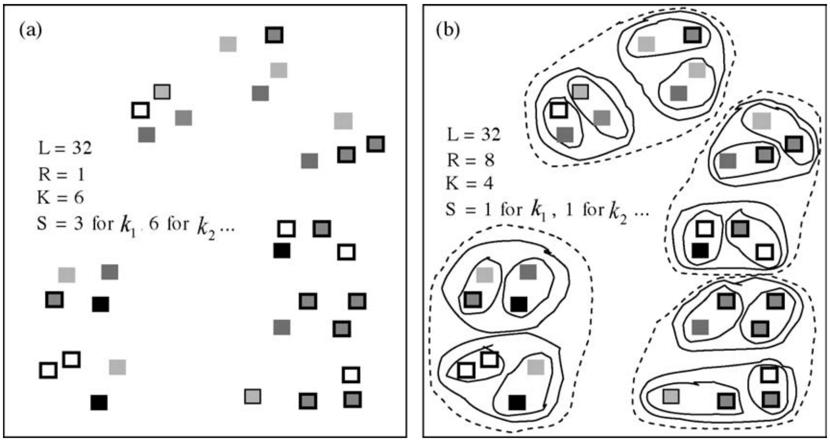

Fig. 2. A hypothetical distribution of a variable in space. At the finest scale of resolution (a), $R = 1$, there are $K = 6$ different categories of color-coded samples with S as the number of samples that are contained in each category. Values that have been calculated for L, K, S, and R are displayed. At the $R = 8$ scale (b), the samples have been re-aggregated based upon their proximity to one another, generating $K = 4$ unique categories of re-aggregations. The solid rings around the samples in (b) represent re-aggregations at the $R = 2$ and 4 scales with the dotted rings as $R = 8$.

2.4. Difference: D

We can also construct a factor D, which measures how different each K is from the other K forms. When using numerical data, D can be mostly simply represented by the distance of each K form from the central tendency, i.e., the mean of L.

$$D(l_{k_is_i}) = l_{k_is_i} - \frac{\sum_{k_i=1}^{k_{i\max}} \sum_{s_i=1}^{s_{i\max}} l_{k_is_i}}{L} \quad (3)$$

In Eq. (3), we refer to the D as a function of a particular l, and the measure is the difference between a real-valued sample l, for a particular k and s, and the average of all real-valued samples l, for all k and s. Fig. 3a represents this idea for the data set in Fig. 2a. If we produce the sum of squares for all $D(l_{k_is_i})$, then we have the variance in the set.

If using ordinal, nominal, categorical, or another type of data, a difference matrix would need to be constructed in a manner similar to that outlined in Krus and Ceurvorst (1979) in order to find D. In these cases, D for each sample does not represent the distance from the mean, as there can be no mean value when using categories that are not relatable by real numbers. Fig. 3b is analogous to Fig. 3a, in that it details the meaningfulness of D in these situations, yet since there can be no mean, we must measure the absolute difference between the categories K. For example, if the categories are colors, then how can we measure a mean value for red, green, and blue? We cannot. In such a situation, a difference matrix must be constructed to find D for each possible combination of the categories. This is demonstrated below for a variable with $L = 4$ samples, a sequence of Blue, Red, Red, and Green:

	Blue	Red	Red	Green
Blue	0	1	1	1
Red	1	0	0	1
Red	1	0	0	1
Green	1	1	1	0

The total difference and our measure of variability D would be 5/8 after partitioning the symmetric matrix in half along the diagonal, summing the squares of each matrix value, and then dividing this sum by the square of L; this type of variance metric was the preferred method by early statisticians before the more common definition became the convention (Krus and Wilkinson, 1986). In this example, there are $K = 3$ categories, with $S = 1$ for k_1 (blue), $S = 2$ for k_2 (red), and $S = 1$ for k_3 (green).

Let us take Fig. 1a as another example. First, assume that the total number of samples of the distribution remains constant where $L = 256$. Add a

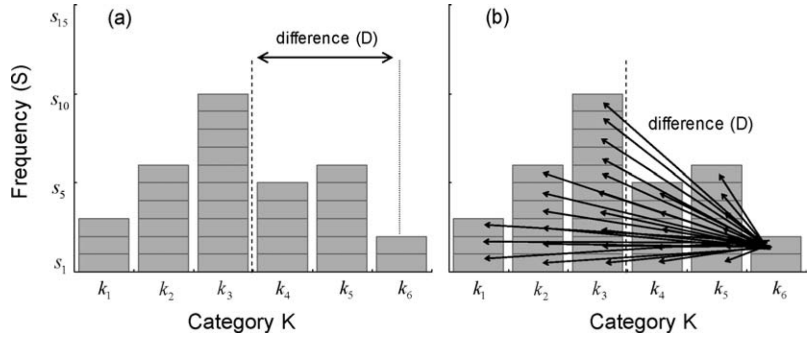

Fig. 3. Histogram representation of the hypothetical distribution from Fig. 2a. A measure of difference, D, of a sample from the rest of the set is most easily represented by the distance from the mean for real-valued data sets (a) and by either ranking or absolute difference for any two samples which are not similar in categorical or ordinal data sets (b). Estimates of variance can then be calculated for the whole set (see text). Abscissa is number of unique categories, K, and ordinate is the frequency of each category, S. The dotted line represents the mean for the real-valued data set in (a), while the solid lines represent D in both the real-valued set (a) and the categorical set (b) for the sample $l_{k_i s_i}$.

new category, orange. In the difference matrix, white (0), black (1), and orange (2) are each different from one another. For every sample along L, we check to see its difference from all other samples; if it is different, then we value this difference as 1. After checking every sample against every other sample, we divide the matrix along the diagonal, sum the squared results of the matrix, and divide by the square of L. This provides a measure of the total variance within the set.

Given that the number of samples is finite and dependent upon L, then the original two categories must decrease in frequency when we add one additional category, orange. Thus, the difference matrix strategy works for directly comparing two data sets only if the number of categories is the same in each.

2.5. Scaling: R

Given a scale R, we can re-sample L according to R, re-create K based upon the new combinations within each R scaled sample, re-calculate S and K, and build a new distribution for that scale R. With this view, a quantitative measure for each scale would be:

$$\frac{L}{R} \propto \sum_{l_{k_i s_i}=1}^{l_{k_i s_i}\max} (D_{l_{k_i s_i}})^2 \tag{4}$$

This formula is essentially a measure of variance that is scale-dependent. If the right-hand side of the equation is maximal relative to the L/R scale, then the distribution would look like Fig. 1a. If the result is minimal or zero, then the distribution would look like Fig. 1b. Fig. 2b shows the earlier data set as it is re-sampled according to R, which here is based upon a spatial proximity rule.

2.6. Commonalities among concepts of heterogeneity and homogeneity

Most common understandings of homogeneity and heterogeneity can fit within this conceptual outline. For example, one component of heterogeneity can be seen as contrast (Kotliar and Wiens, 1990), or as a measure of difference between patch types. To measure contrast, we could construct a difference matrix for the categorical patch types only, thereby obtaining D for each K and S, and calculating the total variation in the set. Heterogeneity can also be seen as a function of aggregation (Kotliar and Wiens, 1990), where it is dependent upon a spatial representation of K, S, D, and a re-sampling of the distribution at scale R across the spatial extent of L.

In another example, heterogeneity has been described as a function of contagion (Li and Reynolds, 1994) where only K and S matter and the calculation is essentially a measure of evenness like Eq. (2). Yet, this measure has an implicit L/R sampling regime based upon adjacency of samples, so it too is a function of L, K, S, D, and R.

2.7. Commonalities among metrics of heterogeneity and homogeneity

Moreover, most specific metrics of variance can be conceptualized as attempting to gain understanding about the variance of the distribution via L, K, S, D, and R. For example, the Fisher-inspired F_{ST} statistic (Fisher, 1918; Wright, 1931, 1951) maps gene flow for two alleles at a locus, and defines populations according to within and between neighborhoods (Slatkin and Voelm, 1991), using a function of the standardized variance,

$$F_{\mathrm{ST}} = \frac{\sigma^2}{p_i(1 - p_i)} \tag{5}$$

as given in Slatkin (1987). The numerator is obtained by sampling L at scale R and calculating the variance of those samples. Note that this calculation is the same as the right-hand side of Eq. (4); we are searching for D and calculating variance. The denominator is a simple function where $p_i = S/L$ for the first of the two K allele forms. Note that by multiplying p_i for one of the alleles by a value that is representative of the other allele's p_i, a different measure of the distribution is arrived at: a measure of the evenness, although in not as complex form as the Shannon entropy of Eq. (2). The entire equation is aimed at estimating whether one or two populations exist, such as in the example of the mountain-climbing ecologist, with values solely dependent upon L, K, S, D, and R.

In another example, the spatial statistic lacunarity (Plotnick et al., 1993, 1996) is very similar in construction:

$$\Lambda(r) = \frac{Z_Q^2(r)}{(Z_Q^1(r))^2} \tag{6}$$

with the second moment $Z_Q^2(r)$ divided by the square of the first moment $Z_Q^1(r)$ on the right-hand side. The second moment, or variance, is the same as the right-hand side of Eq. (4); it is a function of D. The first moment is the mean, which can be found by averaging all $l_{k_i s_i}$. As Plotnick et al. (1996) points out, this mean is also relatable to p_i when the data is composed of two categories K. Note that this mean is squared in Eq. (6), which yields similar results as the denominator in Eq. (5).

Essentially, we have the historical quadrat-variance method used by plant ecologists (Grieg-Smith, 1957; Kershaw, 1957; Usher, 1975), where the variance/mean ratio is dependent upon the sampling scale R. Lacunarity, a measure of gappiness, and the quadrat-variance method, a measure of spatial variance, could be both utilized to quantify the relative heterogeneity/homogeneity in the shoreline photo example.

For lacunarity, we can use our conceptual understanding to explore beyond the simple case where there are only two categories. Interestingly for lacunarity, when the data is categorical and there are only two forms of the variable (0 or 1), the denominator is a function of p_i at $R = 1$ (Plotnick et al., 1996) and the lacunarity value at all other R scales can be normalized by the lacunarity value at $R = 1$. Normalization reduces the effect of changing p_i to low or high values (With and King, 1999), as shown in Feagin (2003).

However, two problems exist with lacunarity: (1) if there are more than two K forms of the variable and the data is composed of real numbers, then the mean is no longer directly relatable to p_i, and (2) the results for a data set that has two categories may not be directly relatable to another data set with three categories even if the underlying distribution is similar, because there is more possible difference in the set in terms of the matrices. In the earlier example of categorical data using Fig. 1a, when the orange (2) category was added, now $K = 3$. The result is that the new third category must appear with some frequency on the histogram in Fig. 1a, and this reduces S for each of the original two K categories. Then, because S/L is reduced for each K, p_i is reduced for each K. Changing K and changing S are inversely related when the number of samples L/R is held constant.

Thus, we should take the total variability in the set as measured by Eq. (4) and divide by some function of the p_i's. This is exactly the purpose of the denominator in both the lacunarity equation and the F_{ST} equation; it scales the variance relative to the number of categories (or in the case of lacunarity with real numbers, it scales by the mean). One interesting idea for a metric of heterogeneity/homogeneity is to divide a measure of variance such as Eq. (4) by a measure of p_i evenness such as Eq. (2).

While some research has been done into the relationship of variance and Shannon entropy (Garner and McGill, 1956; Borzadaran, 2001), much more is needed.

The metrics that have been mentioned are utilized for completely different purposes, yet the components are dependent upon L, K, S, D, and R. Other conceptual understandings and other metrics may be used, but nearly all of them can be condensed into an attempt to gain knowledge about the variability of the distribution (L, K, S, D, R), and thus heterogeneity and homogeneity. As ecologists, we have put a great deal of emphasis upon heterogeneity, and subsequently upon our investigations into measures of variance.

3. Mathematical and physical theory

3.1. Variance and Brownian motion

The classical idea of statistical variance as introduced by Fisher (1915, 1918, 1922) is intimately related to another classical idea, Brownian motion. Traditionally, Brownian motion describes the random movement of a particle about a center locus over time. It is known that the distance that the particle deviates from the locus is proportional to the square root of the time elapsed since the particle was released (Einstein, 1956; Okubo, 1981; Addison, 1997). More formally:

$$s \propto t^{0.5} \tag{7}$$

where s is standard increment deviations from the mean center locus and t is time elapsed.

We can also view this idea in terms of our conceptual understanding, where we build a distribution that is composed of descriptions of the location of the particle at specific times. For the time elapsed, we have L. The distribution is composed of K different locational distances for the particle that we find S times each, where D is the distance from the center locus. With standard Brownian motion, the relationships of K, S, and D are strictly proportional to the time elapsed L. In other words, if given L we can find the variance of the Brownian process.

3.2. Hurst correlation

Fractional Brownian motion (fBM) (Mandelbrot, 1983; Peitgen et al., 1992) can also work under this notation, but it allows generalization to multiple scales of R. Moreover, it allows us to alter the strict proportionality of K, S, and D relative to L and R.

Standard Brownian motion is equivalent to fractional Brownian motion for a value of $H = 0.5$. Fractional Brownian motion allows the use of the Hurst or Hölder exponent H, where the variance of the process may have a different proportionality to the time elapsed. Specifically:

$$E[x_1 - x_\mu]^2 \propto [t_1 - t_2]^{2H} \tag{8}$$

where the left-hand side is the squared difference of the expected values x_1 from the mean x_μ, and the right-hand side is the time elapsed between time one and time two, taken to the power $2H$. For $H = 0.5$, Eq. (8) reduces to the standard Brownian motion in Eq. (7).

The Hurst exponent is a measure of the correlation of the increments in Eq. (8). At $H = 0.5$, there is no correlation between the increment of the distance at time one and time two. The probability that the next increment will deviate further or closer to the mean locus is 0.5. Note that the Brownian $H = 0.5$ is the integration of white noise over the time elapsed. In other words, by taking a stationary random process where the local probability of change is equal for all alternative states (0 or 1 here, referring to closer or farther away from the mean locus) and integrating it over time, a multi-scaled representation of randomness is produced. This representation is statistically self-similar, scale-invariant, and self-affine at $1/f^2$ scales. Regardless of the measurement unit, the expected value is equiprobable (here, for either 0 or 1).

Although the primary interest is with $H = 0.5$, other values of H are possible. The random, but correlated $0 < H < 0.5$ is termed anti-persistent and $0.5 < H < 1$ is termed persistent. Further, one can extrapolate the Hurst concept to the non-random extremes of absolute positive or negative correlation over time (Mandelbrot, 1999).

The concept of Brownian motion can be applied to either space or time. For quantifying spatial variance, Eqs. (7) and (8) now relate the variance of any process s^2 to the distance d elapsed. The implication is that the similarity between two objects decays exponentially with distance, specifically to the square root of distance. This decay in similarity or relatedness is what would be expected for a random process, integrated across distance. Spatial autocorrelation, as the first law of geography, states that ‘‘things closer together are more similar than things far apart’’ (Tobler, 1970).

Explicitly, multi-scaled randomness is defined by $H = 0.5$, where invariance defines the scaling relation between the distribution of a variable (as defined by K, S, and D) and the scale R at which L is sampled. Moreover, the collections of the samples have variance s^2 about the mean locus for every scale. Scale-invariant phenomena with a consistent second moment s^2 implies that the distribution is Gaussian at every scale. Regardless of the scale of resolution at which a limit-Gaussian distribution is sampled, it will still appear Gaussian; no threshold is crossed that divides the sample distribution into distinct phases or states.

3.3. Scaling in physical surfaces: H does not measure criticality

The ubiquity of scale-invariant phenomena, where the second moment remains constant across scale, has led to the development of universality classes and critical scaling exponents in physical systems research (Vicsek, 2001). These exponents define the rate at which self-affinity is expressed across scale, which is dependent upon a scaling relation such as H, as well as upon the parameters of the system.

When we talk about critical exponents, we are indeed discussing scale-invariant distributions that follow a power-law like H, but they are also part of a system of equations that describe the behavior of a model, such as the scaling of growing interfaces in the Kardar–Parisi–Zhang (KPZ) equation (Kardar et al., 1986; Family and Vicsek, 1991), here with exponents roughness α, growth β, and saturation γ. The theoretical results of both the KPZ model and the Eden model (Czirok, 2001) support the argument that the scaling relation $H = 0.5$ for a random distribution is valid in any dimension (i.e., one- or two-dimensional surfaces), although the fluctuation-dissipation theorem leading to the critical exponent in the KPZ equation, $\alpha = 0.5$ as a roughness parameter of a pattern that is growing, is valid in only one-dimension (Barabasi and Stanley, 1995). We find can other values for the critical exponents $0.2 < \alpha < 0.4$ and $0.08 < \alpha < 0.33$ as theoretical results for two- and three-dimensional growth surfaces, respectively (Barabasi and Stanley, 1995). The scaling relation $H = 0.5$ in conjunction with the critical exponent $\alpha = 0.5$ have been found in the spatial domain in bacterial colonies (Czirok, 2001) and small cracks in surfaces (Parisi and Caldarelli, 2000), although different values where $\alpha \neq 0.5$, have been found in the spatio-temporal domain where the speed at which the cracks propagate is also a factor, in correlated models where $H \neq 0.5$, in models where there is surface tension, as well as for large cracks in surfaces.

In codifying heterogeneity and homogeneity, we do not view H as a critical exponent that describes a critical state where catastrophic changes can quickly occur. Nor do we view it as a measure of a growing front, dependent upon roughness α, growth β, and saturation γ. Rather, we view $H = 0.5$ only as a scaling relation where the distribution of values is Gaussian random at all scales, but this is not a necessary nor sufficient condition for criticality. Our view of heterogeneity and homogeneity is much broader, looking at distributions, and simultaneously not restricted to modeling random walk processes on surfaces with specific physical parameters.

4. Methods

As a visual analogue to the abstract theory of variance, Brownian motion and Hurst correlation, we first generated scale-invariant representations of heterogeneity and homogeneity with a known variance and Hurst exponent. We then quantified the relative heterogeneity and homogeneity for patterns from three data sets, where the relative variance and Hurst exponent were scale-covariant.

4.1. Scale-invariant representations

The inverse Fast Fourier Transform method (Keitt, 2000) was utilized in a MATLAB routine to create these visual representations with specific statistical properties, when given H as input. This procedure creates a pattern with a known variance that scales consistently; hence, it is scale-invariant. Of course, we provide the known variance by providing the input parameter H.

We created both one- and two-dimensional patterns. In the one-dimensional case, we entered a known value of H and then took the zero-set (Keitt, 2000) of the output trace. This procedure created an array of binary digits that was composed of two possible values, 0 or 1. For visualization purposes, we represent 0's as white and 1's as black in the graphs. These arrays were 256 segments long where $L = 256$. The procedures in the two-dimensional case were similar, where we used a matrix rather than an array.

We then created histograms to depict the distribution of the values at specific scales, where a moving-window was passed over the array at a specific resolution, R. At each scale of resolution, the moving-window was placed over the array and the number of black "boxes", or black segments of the array that fit within the window were counted. The moving-window was then moved one box-length to the right, and the number of black boxes was again counted. Thus, the histograms depicted the distribution of the box "masses" (Allain and Cloitre, 1991) of each array. Each unique box mass category that was encountered along the array is representable by K and its frequency of occurrence by S. D is the distance of any value from the central tendency of the distribution, which we represented by a dotted line in the graphs. We present this analysis at two scales: $R = 4$ boxes wide and $R = 16$ boxes wide.

4.2. Scale-covariant representations

We then took three data sets with unknown variance and H parameters and quantified them with the $H(r)$ method (Feagin, 2003). The $H(r)$ graph method is a variant of lacunarity that adjusts the results for both percent cover p_i and variance relative to scale. The graphs show the relative amount of variance with H as a function of a specific scale, r, when compared with the $H = 0.5$ multi-scaled, Gaussian distribution. Thus, we obtain a graph that can detail the scale-covariance within a pattern.

The three data sets were: (1) spatial patterns of plants growing on a sand dune, contained within a 35 m^2 plots on Galveston Island, TX, USA that were sampled with 0.01 m^2 quadrats, where one pattern is composed of pioneer plant species in December 2000 and the other of a sand dune plant community in June 2000, (2) spatial patterns of woody versus non-woody landscape types, contained within 300 m $\times$ 426 m photographs taken at 1 m^2 resolution of a "tiger-stripe" landscape in Niger, Africa, in both 1960 and 1992 (Wu et al., 2000), and (3) a temporal sequence of water wave height values recorded by a laser sensor in a wave tank experiment at Texas A&M University, with the readings taken every 0.01 s for a duration of 20.48 s, where the experiment occurred in early 2001.

5. Results

5.1. Scale-invariant representations of heterogeneity, homogeneity, and a transitional state

Spatial patterns and histograms that have a known variance and correlation parameter H are offered in Fig. 4. Table 1 links the visual results of Fig. 4 with the conceptual and mathematical theory.

Fig. 4a shows a pattern that is at the limit of heterogeneity at all scales. Each of the histograms in Fig. 4b and c depicts the distribution of the limit-heterogeneous pattern at a specific scale, where K represents each unique box mass total and S represents the frequency of obtaining that K. For example in Fig. 4b, when the moving-window was $R = 4$ boxes wide, there were $K = 5$ possible box counts (0–4), or number of black boxes that were occupied within the moving-window. As the histogram shows, the distribution of the black boxes along the array was heavy-tailed with box masses being either 0 or 5 implying a high variance from the mean of 2.5 at this scale. As S was high for the 0 and 5 box mass samples, D is maximal for each of those samples; the variance is high. Fig. 4c shows the distribution when the moving-window was 16 boxes wide; again, it is at the limit of heterogeneity. This situation remains the same for all R (except in the limit, where $R = L$ and the pattern becomes limit-homogeneous). Fig. 4d shows a two-dimensional example of the limit-heterogeneous pattern.

Fig. 4e shows a heterogeneous pattern ($H > 0.5$) that fits between the limit of heterogeneity and a transitional state ($H = 0.5$). The pattern appears aggregated. The correlation between one spatial location and another location is so high that there is a strong likelihood that an adjacent pixel is similar. This concept is relevant to scaling, because the integration of the individual increment correlation

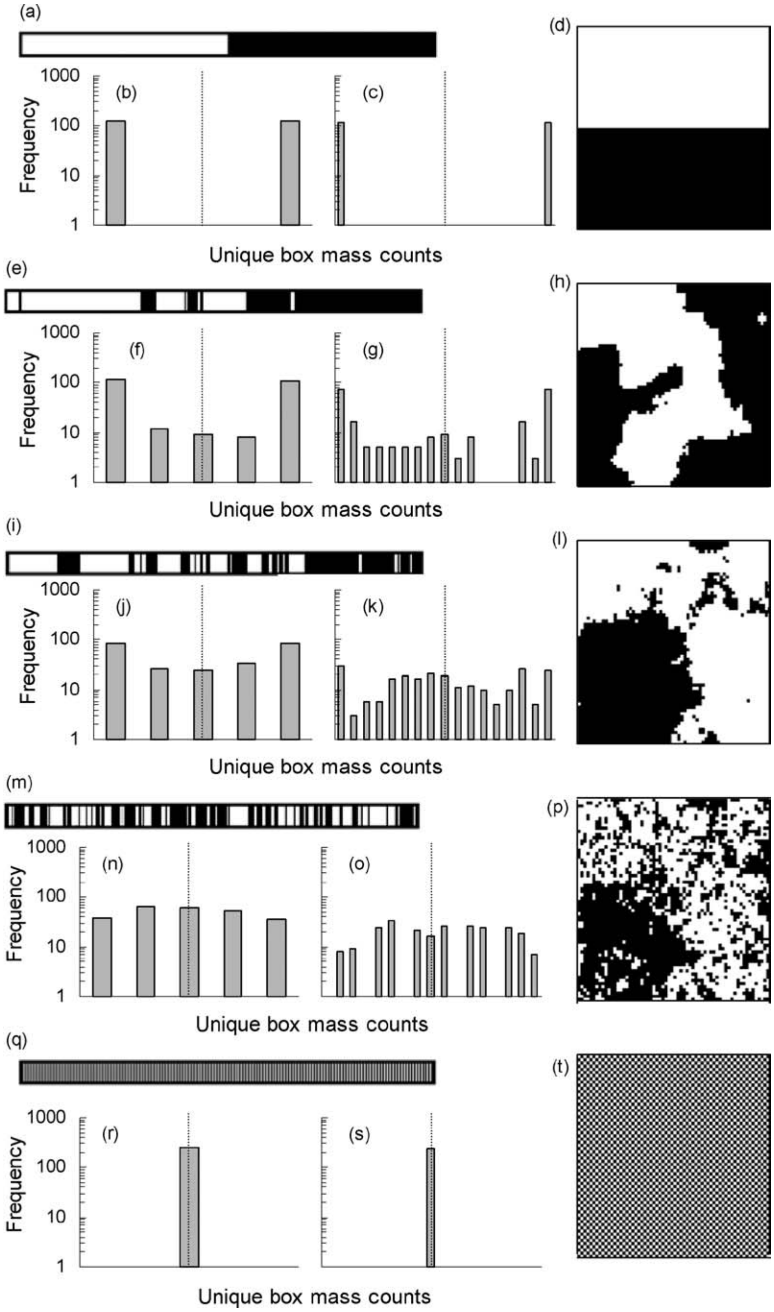

Fig. 4. Limit-heterogeneous (a–d), heterogeneous (e–h), critical (i–l), homogeneous (m–p), and limit-homogeneous (q–t) states. Each state differs in visual appearance, arrangement, and statistical description, yet all have the same p_i-values. For example, the limit-heterogeneous state is depicted as a one-dimensional trace (a) that is composed of $L = 256$ segments, or boxes, with $p_i = 0.05$ for both black and white. The histograms (b and c) display the frequency, S, of each unique ‘‘box mass’’ count, K, for a moving-window of $R = 4$ boxes wide (b) and $R = 16$ boxes wide (c), respectively. Abscissa is categories K, with respect to the mean ‘‘box mass’’ value, as represented by the center dotted line. Ordinate is frequency, S. A two-dimensional version of a limit-heterogeneous pattern is also shown (d). The other states are represented in a similar manner. Table 1 is cross-referenced with this figure and provides look-up information on each state.

Table 1
Summary of the classification system for the five distributions: limit-heterogeneous, heterogeneous, transitional, homogeneous, and limit-homogeneous

Distribution	Conceptual	Mathematical
Limit-heterogeneous (Fig. 4a–d)	S is maximal for each K D is maximal (except at $R = L$)	Bifurcated at L/K
Heterogeneous (Fig. 4e–h)	S is high for K with relatively high D S is low for K with relatively low D	$0.0 < H < 0.5$
Transitional (Fig. 4i–l)	$S \times K$ is maximal when counting the number of unique values that occur for each, D is moderate	$H = 0.5$
Homogeneous (Fig. 4m–p)	S is high for K with relatively low D S is low for K with relatively high D	$0.5 < H < 1.0$
Limit-homogeneous (Fig. 4q–t)	$K = 1$, $S = L$, $D = 0$ (except at $R = 1$)	Translational invariant

Each distribution class has unique conceptual (based upon L, K, S, D, R) and mathematical (based upon the correlation H) attributes. Fig. 4 is cross-referenced with this table and provides visual examples for each distribution.

across the measurement unit defines the relative spread of the variance. The high correlation occurs across the sequence except at the bifurcation of the distribution, where the other alternative (black or white, 0 or 1) begins to occur with increasing frequency. Thus, once the box masses have been calculated, there is a strong deviation from the mean. The result is that the histograms in Fig. 4f and g appear to have most of their values towards the right and left tails of the distribution. Again, S is high for those K with relatively large D, while S is low for those K with low D; the variance is still high. Fig. 4h depicts a two-dimensional example.

There is a transitional state where the variance is proportional to $H = 0.5$ at a particular scale. At the transitional state, the pattern is scale-invariant with Gaussian distribution. The transitional state is neither heterogeneous nor homogeneous. It represents the dividing line between these two oppositional states.

Fig. 4i shows a pattern that is transitional at all scales. The combination of unique values for both S and K are represented maximally in Fig. 4k, where there are many different values for each, resulting in moderate D; the sum of all D and the variance are moderate as well. The fact that the distribution in Fig. 4j appears to be non-Gaussian is an artifact due to the finite length of the array, with the stipulation that acquiring randomness at every separate scale requires an array with infinite length (Martin-Lof, 1966; Chaitin, 1998). However, when integrated across the unit of resolution, the distribution of the array in Fig. 4i is Gaussian with $H = 0.5$. Fig. 4l depicts the two-dimensional version.

Homogeneity ($H < 0.5$) is defined as having variance lower than the transitional state at a particular scale. Fig. 4m–p shows patterns that are homogeneous at all scales. In homogeneous distributions, S is high for those K with low D, while S is low for those K with high D.

At the limit of homogeneity, there is absolute negative correlation between increments, resulting in translational invariance as Fig. 4q–t show. The box masses are the same for all samples, leading to $K = 1$, $S = \mathrm{L}$, and $D = 0$. There is no variance and the histograms have a single value at all scales (except in the limit at $R = 1$, where the pattern becomes limit-heterogeneous). These patterns appear regular.

5.2. Scale-dependent variance and scale-covariance

Some patterns are not self-affine; their relative variance changes according to the scale at which the variables are sampled. For example, although the distribution of pioneer plant species within a plot in December 2000 in Fig. 5a and b was heterogeneous at all scales $0.1 < R < L$ (m), the relative amount of heterogeneity varied according to scale. The distribution of a sand dune plant community in June 2000 at fine scales in Fig. 5c and d had a homogeneous distribution. This pattern exhibited homogeneity at all scales $0.1 < R < L$ (m), although the variance covaried with scale.

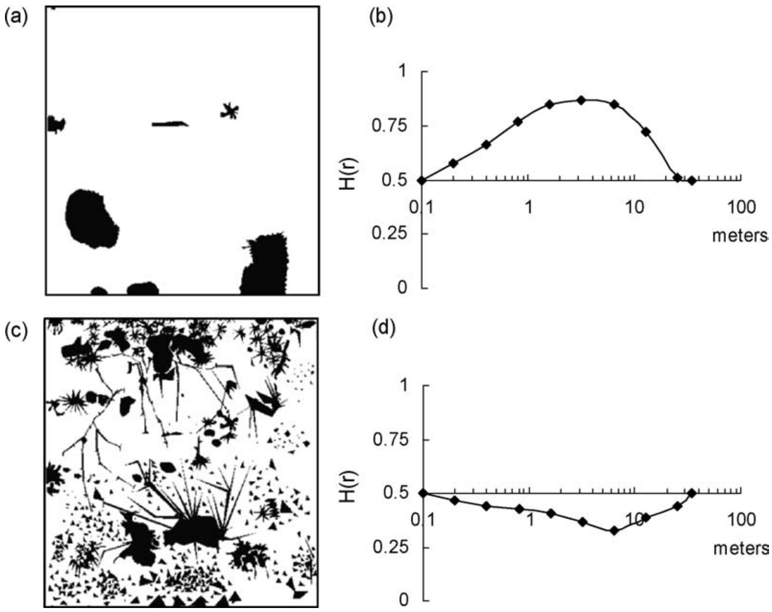

Fig. 5. Two spatial distributions that were scale-covariant. The pioneer plant species' pattern within a plot (a and b) was heterogeneous (high D, $H > 0.5$) at all scales $0.1 < R < 35$ m. The sand dune plant community's pattern within a plot (c and d) was homogeneous (low D, $H < 0.5$) at all scales $0.1 < R < 35$ m. Both plots were 35 m^2 in extent with a resolution of 0.1 m^2.

Some patterns may exhibit both heterogeneity and homogeneity at different scales. In the tigerstripe landscape example in Fig. 6a–c, the pattern was heterogeneous at fine scales in both 1960 and 1992. At coarse scales in both 1960 and 1992, the pattern was homogeneous. At about 25 m of resolution in 1960 and about 65 m in 1992, the pattern approximated a transitional state. Between the periods of 1960 and 1992, the pattern became more heterogeneous at increasingly coarse scales, as the transitional state shifted towards these coarser scales.

Temporal and non-spatial patterns may also exhibit scale-covariance in heterogeneity and homogeneity. Fig. 7a and b shows the temporal record of propagating waves, where the distribution was both heterogeneous and homogeneous with the transitional state approximated at about 1.5 s.

For non-self-affine patterns, the second moment or the variance, changes with scale. Thus, as many ecologists have pointed out, heterogeneity is a scale-dependent phenomenon. An object, attribute, or pattern that appears heterogeneous (high D, $H > 0.5$) at one scale can appear homogeneous (low D, $H < 0.5$) at another scale, or vice-versa. Moreover, a transition in the relative appearance from heterogeneity to homogeneity via a resolution change implies that the pattern crosses the transitional state (moderate D, $H = 0.5$) at some scale.

The scale at which a pattern crosses over the transitional state may provide some definition of the system. At the transitional state, the pattern is neither heterogeneous nor homogeneous. Rather, the variation from one location to another location is equally probable to be either relatively similar or relatively different (if t_1 is black, then t_2 is equally probable to be black or white) when compared against the mean of all samples, and this happens regardless of the distance between samples, implying a likely shift from local (heterogeneous) to global (homogeneous) dynamics. The presence of the transitional state simply indicates that the system is randomly structured when viewed from a certain perspective or scaling regime.

6. Discussion

Conceptually, heterogeneity and homogeneity are based upon difference and similarity, respectively. L,

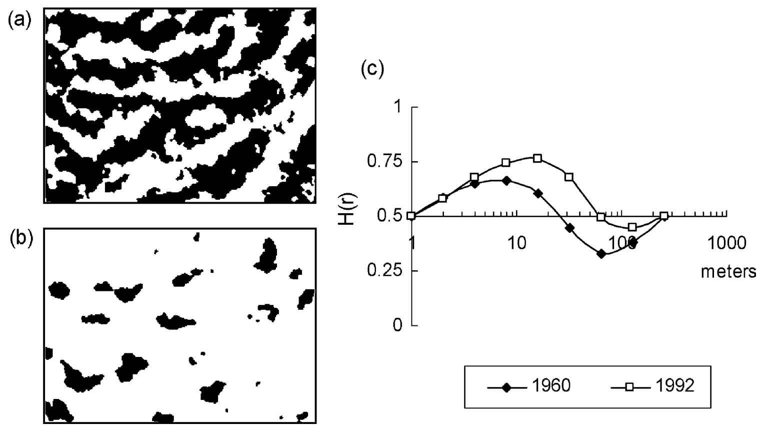

Fig. 6. Heterogeneity and homogeneity in the same pattern. The 1960 aerial photo of the tigerstripe landscape (a) was heterogeneous (high D, $H > 0.5$) at fine scales, but homogeneous (low D, $H < 0.5$) at course scales, as (c) shows. The transitional state was crossed at 25 m. In the 1992 aerial photo of the same location, the landscape (b) was also heterogeneous at fine scales and homogeneous at coarse scales. However, the transitional state had shifted to 65 m, indicating that the landscape had moved towards heterogeneity at increasingly coarse scales during the period. Both photos were 300 m × 426 m in extent with a resolution of 1 m^2, as adapted from Wu et al. (2000).

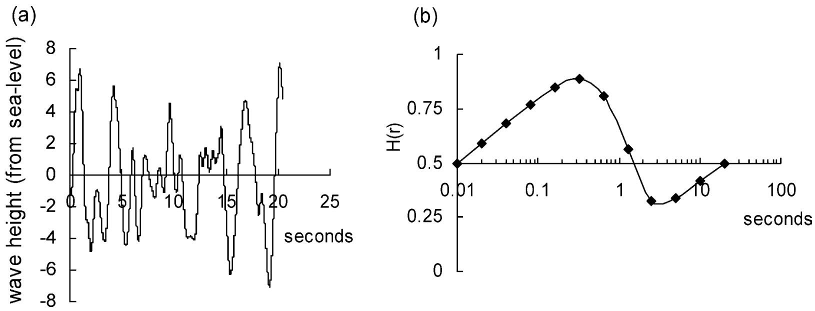

Fig. 7. A temporal pattern that exhibited scale-covariance. The propagating waves (a) were heterogeneously distributed (high D, $H > 0.5$) at fine scales, but homogeneously distributed (low D, $H < 0.5$) at coarse scales, as (b) shows. The transitional state (D moderate, $H = 0.5$) was crossed at 1.5 m. The wave height values were recorded with a laser sensor every 0.01 s for a total of 20.48 s.

K, S, D, and R help us to partition the many components of these two states into discrete functions; this is important because, with the conceptual basis, extrapolation is possible to the multifarious cases where the distribution is composed of an unexpected form. For example, we can discuss the relative amount of heterogeneity in a set even if we are limited to knowledge about how many samples are in the set, L, and how many different types are in the set, K, for we still have a measuring system. Of course, as we gain more knowledge about the distribution, we can begin to quantify it in all of the terms discussed with the conceptual basis. Moreover, we have a measurement system for spatial, temporal, or vector-less data sets even when the data is composed of ordinal, numerical, or categorical data.

The mathematical and physical basis, as outlined in the theories of Brownian motion, scale-invariant fBm, and the Hurst exponent simply allow us to quantify D at multiple scales of R. A distribution with maximum variance and maximum H is heterogeneous, while a distribution with minimum variance and minimum H is homogeneous. A median level of variance and $H = 0.5$ represent a transitional random state that separates heterogeneity from homogeneity. Regardless of specific method, one can define which of these three states exists at a specific scale, when comparing the variance or correlation to the multi-scaled Gaussian distribution.

The visual representations of the black and white cells in Fig. 4 allowed us to depict the concepts in the simplest way: two possible categories, with both at $p_i = 0.5$. However, we can still utilize the mathematical basis, even when there are more than two K, when $p_i \neq 0.5$, or when the data is of an unexpected type, by deferring to the conceptual basis. The relationship between the conceptual theory and mathematical theory was summarized in Table 1. The model of the Hurst exponent is instructive in that it divides heterogeneity from homogeneity with the transitional state, based upon H, which is a function of D. The same procedure can be done with L, K, S, R, and any other function of the distribution, by using the Gaussian distribution as a dividing line.

6.1. The apparent degree of heterogeneity versus homogeneity leads to set categorization

Ecologists have often remarked that variance is high "between" sets of variables, but low "within" sets. As an example, if two contiguous areas were sampled for land cover type in the close-up photo of the shoreline and they were very different (land and water) like the right versus the left halves of Fig. 4a or e, then there would be a relatively high variance "between" the two areas. The extent of what one would see, the two land and water areas, would be heterogeneous.

Or, in another example assume that two populations were sampled by the mountain-climbing ecologist along the transect for their genetic type (black coded for male, white coded for female) and that they are represented by the right half and left half of Fig. 4m. Each half of the transect, or population, has 128 adjacent individuals for a total of 256. Then, if there were not large differences between the two populations in the distribution of the genetic type, then there would be a relatively low variance and perhaps the populations should be redefined as a single population. The single population would be relatively homogeneous for that genetic type.

However, if the variance "within" the single population was relatively high at a finer scale, for example "between" the one male (black) and one female (white) adjacent to one another out of the transect of 256 individuals like Fig. 4q, then one might say that there is heterogeneity, at that finer scale. Perhaps, they could be classified as "families" at that fine scale. Note that at this fine scale, Fig. 4q appears to be 128 copies of the limit-heterogeneous Fig. 4a stacked end to end, albeit at a reduced scale. Indeed, if referred to with an $H(r)$ graph, the transitional state would be crossed just before this fine scale. Alternately, one may imagine that the histogram of the box masses would abruptly change from the form contained in Fig. 4r and s to that of Fig. 4b and c for that $R = 1$ fine scale.

Hypothesis. The transitional state defines sets, systems, and hierarchies.

It has been suggested that the scales at which power-laws abruptly shift can be indicative of the discrete scales at which a system is organized (Allen and Holling, 2002; Gardner, 1998). Some have called these scale breaks or domains of scale.

We propose a hypothesis that the scale at which a system is heterogeneous, homogeneous, or transitional is the salient criteria for marking important differences in scale-covariance, in terms of potential organization.

Heterogeneity (high D values, $H > 0.5$) should exist between sub-systems within a system. Yet, each sub-system is homogeneous (low D values, $H < 0.5$) within itself, relative to other sub-systems. This hypothesis invokes nothing more than the foundation upon which ANOVA procedures were originally based.

The point at which two sub-systems become no longer distinguishable as statistically heterogeneous elements, like any two parts of Fig. 4i, the transitional state (moderate D values, $H = 0.5$) is crossed and they compose one homogeneous system. The transitional state defines the delineation of the two states.

The transitional state may also define the separation of systems and sub-systems into hierarchical levels at the crossover scale. Whenever the transitional state is crossed at a particular scale, such as 11.5 m for the 1960 tigerstripe data in Fig. 6c or 1.5 s for the wave data in Fig. 7b, a hierarchical boundary is defined. At that scale, the division between one hierarchical level and another level is drawn. In the example of Fig. 6c, the plant community exists at the larger homogeneous scales, yet individual plants exist at the smaller heterogeneous scales. O'Neill et al. (1991) discuss this same notion using aerial photos of a Kansas landscape. In Fig. 7b, above 1.5 m the water waves were distributed homogeneously, while below 1.5 m the waves ceased to be part of the global pattern and were localized, irregular events.

6.2. Problems in hierarchy theory

One drawback to hierarchy theory (Allen and Starr, 1982) and general systems theory (von Bertalanffy, 1968) is that few studies have been able to coordinate sets that have been categorized by measurements of variability with ecologically meaningful hierarchies in the real world (King, 1997).

One problem is that as the extent of the observed data set is changed, the apparent variance, and relative heterogeneity and homogeneity, changes (O'Neill and King, 1998). This occurs because as we sample beyond the extent L, we may pick up very different patterns, which will change the relative view of the original pattern in unexpected ways. Wiens (1989) and Kemp et al. (2001) have argued that it is reasonable to expect that the original pattern will look increasingly homogeneous relative to the rest of the new pattern as L is enlarged; this presupposes that the new pattern is quite different from the original pattern and that the world is heterogeneous as we enlarge it. The opposite can be expected to occur if the new pattern is composed of exact replicas of the original pattern; the world would be more stationary that we expect, at those originally sampled scales of R. It seems the only way to keep the relative variance the same for the original pattern is for the rest of the new pattern to obey $H = 0.5$, a wild notion.

We cannot know the relative scaling rate of the outside world if we have not sampled it in terms of L, K, S, D, and R, unless we refer to an external probability distribution. An argument can certainly be made that there is no way to know what this distribution looks like, so there is value in making our measurements of distributions self-referential, and bounded by L.

If our terms K, S, D, and R are quantified relative to the maximum and minimum possible in L, then our measures are portable and self-referential. They do not rely upon L. This can be viewed as an advantage, such as it has been in measures of Kolmogorov entropy relative to Shannon entropy (Kolmogorov, 1968; Li and Vitanyi, 1993). The $H(r)$ graphs of Figs. 5–7 share this property of portability. Our definitions of hierarchical levels in a data set may still shift when we shift our point of view, but is knowledge of the world relative or objective? Research is needed for an extension of the conceptual theory to include a new term that measures the effect of transformations in L.

Interestingly, the fact that K, S, D, and R change when L changes leads to a potential definition of a habitat. We can adjust L and move the extent of our view such that we cover a set of interest. Following Kolasa and Waltho (1998), a habitat can be considered a homogeneous distribution of particular variable that is necessary for a species. As has been proposed and modeled (With and King, 1999; Feagin et al., 2005), the scales at which species interact with a resource must match the scale at which the resource is distributed for persistence to occur. Thus, we would aim to configure L in size and shape such that the distribution of the variable would be homogeneous at the scale of interaction. For example, in Fig. 5a, we could reconfigure the extent such that it covers only one of the black patches, invoking $H < 0.5$ at a small scale for an organism that can only inhabit black patches. More complex descriptions of habitat are possible, where distance to a particular resource could also be determinant, rather than simply whether that resource exists or not at a particular location.

A second problem is that ecologically meaningful hierarchies are not simple artifacts that can be defined by sampling a single variable. We may define the hierarchy in a plant community differently based upon sampling of canopy height, nitrogen consumption, species composition, or an infinite amount of variables and combinations of them. Often, the variance structure of a particular variable does not match our pre-conceived notions of levels in the ecological hierarchy (King, 1997). Similiar to the KPZ model for surface growth, the power-law structure that the system displays in terms of the critical parameters often does not match the theoretical scaling relations that approximate random processes in a system, such as $H = 0.5$. There are many other equations and potentially critical exponents, who would be determinant of the important power-laws, which govern the dynamics of the system. However, it does seem reasonable to assume that by measuring the variance structure of the primary driving variables in the system, we should be able to extricate ecologically meaningful results as long as L matches our preconceived notions of the relevant system size.

Hierarchy theory, and its relation to a conceptual and mathematical basis for heterogeneity and homogeneity, does indeed have deficiencies when it comes to real-world ecological distributions. Still, it can produce testable hypotheses and provide a baseline model for data-driven explorations.

7. Conclusions

With the conceptual basis, we can reduce heterogeneity and homogeneity to some function of L, K, S, D, and R. D is of prime importance as many measures of ecological distributions focus upon the variance. Standard Brownian motion provides a mathematical and physical model for understanding the distribution, particularly in terms of D. Once we begin viewing distributions as multi-scaled objects, R becomes important and the Hurst exponent allows us to include it within the mathematical and physical framework.

Heterogeneity and homogeneity are divided from one another by a transitional state, which approximates a multi-scaled Gaussian distribution. Ecological distributions can be heterogeneous, homogeneous, or both as dependent upon scale.

In patterns where we expect hierarchical levels, we expect more than simple scale-covariance; we expect division by the transitional state at some scale. Although there are problems in hierarchy theory, especially in matching pattern-defined hierarchies with expected hierarchical levels from ecological theory, the theory still provides a potent example of the importance of heterogeneous versus homogeneous distributions at multiple scales.

References

Addison, P.S., 1997. Fractals and Chaos. Institute of Physics Publishing, Bristol.

Allain, C., Cloitre, M., 1991. Characterizing the lacunarity of random and deterministic fractal sets. Phys. Rev. A 44, 3552–3558.

Allen, A.P., Li, B.L., Charnov, E.L., 2001. Population fluctuations, power laws, and mixtures of lognormal distributions. Ecol. Lett. 4, 1–3.

Allen, C.R., Holling, C.S., 2002. Cross-scale structure and scale breaks in ecosystems and other complex systems. Ecosystems 5, 315–318.

Allen, T.F.H., Starr, T.B., 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago.

Anand, M., Li, B.L., 2001. Spatiotemporal dynamics in a transition zone: patchiness, scale, and an emergent property. Community Ecol. 2, 161–169.

Barabasi, A.L., Stanley, H.E., 1995. Fractal Concepts in Surface Growth. Cambridge University Press, Great Britain.

Barkman, J.J., 1989. A critical evaluation of minimum area concepts. Vegetatio 85, 89–104.

Borzadaran, G.R.M., 2001. Relationship between entropies, variance and Fisher information. In: Mohammad-Djafari, A. (Ed.), Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 20th International Workshop, American Institute of Physics, USA, pp. 139–144.

Bradshaw, G.A., Spies, T.A., 1992. Characterizing canopy gap structure in forests using wavelet analysis. J. Ecol. 80, 205–215.

Braun-Blanquet, J., 1932. In: Fuller, G.D., Conard, H.S. (translation), Plant Sociology. McGraw-Hill, New York.

Cadenasso, M.L., Pickett, S.T.A., Weathers, K.C., Bell, S.S., Benning, T.L., Carreiro, M.M., Dawson, T.E., 2003a. An interdisciplinary and synthetic approach to ecological boundaries. Bioscience 53, 717–722.

Cadenasso, M.L., Pickett, S.T.A., Weathers, K.C., Jones, C.G., 2003b. A framework for a theory of ecological boundaries. Bioscience 53, 750–758.

Chaitin, G.J., 1998. The Limits of Mathematics. Springer–Verlag, Singapore.

Cornelius, J.M., Reynolds, J.F., 1991. On determining the statistical significance of discontinuities within ordered ecological data. Ecology 72, 2057–2070.

Czirok, A., 2001. Patterns and correlations. In: Vicsek, T. (Ed.), Fluctuations and Scaling in Biology. Oxford University Press, Oxford, pp. 48–116.

Einstein, A., 1956. Investigations on the Theory of the Brownian Motion. Dover Press, USA.

Epperson, B.K., 2003. Geographical Genetics. Princeton University Press, Princeton, NJ.

Family, F., Vicsek, T., 1991. Dynamics of Fractal Surfaces. World Scientific, New Jersey.

Feagin, R.A., 2003. Relationship of second-order lacunarity, Hurst exponent, Brownian motion, and pattern organization. Phys. A 328, 315–321.

Feagin, R.A., Wu, X.B., Smeins, F.E., Whisenant, S.G., Grant, W.E., 2005. Individual versus community level processes and pattern formation in a model of sand dune succession. Ecol. Model. 183, 435–449.

Fisher, R.A., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10, 507–521.

Fisher, R.A., 1918. The correlation between relatives on the supposition of Mendelian inheritance. Philos. Trans. R. Soc. Edin. 52, 399–433.

Fisher, R.A., 1922. On the mathematical foundations of theoretical statistics. Philos. Trans. R. Soc. Lond. A222, 309–368.

Forman, R.T.T., 1997. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, Cambridge, UK.

Fortin, M.-J., 1994. Edge detection algorithms for two-dimensional ecological data. Ecology 75, 956–965.

Fortin, M.-J., Olson, R.J., Ferson, S., Iverson, L., Hunsaker, C., Edwards, G., Levine, D., Butera, K., Klemas, V., 2000. Issues related to the detection of boundaries. Landscape Ecol. 15, 453–466.

Gardner, R.H., 1998. Pattern, process, and the analysis of spatial scales. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale: Theory and Applications. Columbia University Press, New York, pp. 17–34.

Garner, W.R., McGill, W.J., 1956. The relation between information and variance analyses. Psychometrika 21, 219–228.

Gosz, J.R., 1993. Ecotone hierarchies. Ecol. Appl. 3, 369–376.

Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4, 379–391.

Grieg-Smith, P., 1957. Quantitative Plant Ecology. Butterworth, London.

Hulbert, S.H., 1990. Spatial distribution of the montane unicorn. Oikos 58, 257–271.

Kardar, M., Parisi, G., Zhang, Y., 1986. Dynamic scaling of growing interfaces. Phys. Rev. Lett. 56, 889–892.

Kareiva, P., 1990. Population-dynamics in spatially complex environments: theory and data. Proc. R. Soc. Lond.,Ser. B: Biol. Sci. 330, 175–190.

Keitt, T.H., 2000. Spectral representation of neutral landscapes. Landscape Ecol. 15, 479–493.

Keitt, T.H., Marquet, P.A., 1996. The introduced Hawaiian avifauna reconsidered: evidence for self-organized criticality? J. Theor. Biol. 182, 161–167.

Kemp, W.M., Peterson, J.E., Gardner, R.H., 2001. Scale-dependence and the problem of extrapolation: implications for experimental and natural coastal ecosystems. In: Gardner, R.H., Kemp, W.M., Kennedy, V.S., Peterson, J.E. (Eds.), Scaling Relations in Experimental Ecology. Columbia University Press, New York, pp. 3–57.

Kershaw, K.A., 1957. The use of cover and frequency in the detection of pattern in plant communities. Ecology 38, 291–299.

King, A.W., 1997. Hierarchy theory: a guide to system structure for wildlife biologists. In: Bissonette, J.A. (Ed.), Wildlife and Landscape Ecology: Effects of Pattern and Scale. Springer, New York, pp. 185–212.

Kolasa, J., Zalewski, M., 1995. Notes on ecotone attributes and functions. Hydrobiologia 303, 1–7.

Kolasa, J., Waltho, N., 1998. A hierarchical view of habitat and its relationship to species abundance. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale: Theory and Applications. Columbia University Press, New York, pp. 55–76.

Kolmogorov, A.N., 1968. Logical basis for information theory and probability theory. IEEE Trans. Inform. Theory IT 14, 662–664.

Kotliar, N.B., Wiens, J.A., 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. Oikos 59, 253–260.

Krus, D.J., Ceurvorst, R.W., 1979. Dominance, information, and hierarchical scaling of variance space. Appl. Psych. Meas. 3, 515–527.

Krus, D.J., Wilkinson, S.M., 1986. Matrix differencing as a concise expression of test variance. Educ. Psychol. Meas. 46, 179–183.

Li, B.-L., 2000. Fractal geometry applications in description and analysis of patch patterns and patch dynamics. Ecol. Model. 137, 33–50.

Li, B.-L., 2002. A theoretical framework of ecological phase transitions for characterizing tree-grass dynamics. Acta Biotheor. 50, 141–154.

Li, B.-L., Loehle, C., 1995. Wavelet analysis of multiscale permeabilities in the subsurface. Geophys. Res. Lett. 22, 3123–3126.

Li, H., Reynolds, J.F., 1994. A simulation experiment to quantify spatial heterogeneity in categorical maps. Ecology 75, 2446–2455.

Li, H., Reynolds, J.F., 1995. On definition and quantification of heterogeneity. Oikos 73, 280–284.

Li, M., Vitanyi, P.M.B., 1993. An Introduction to Kolmogorov Complexity and its Applications. Springer–Verlag, New York.

Loehle, C., 2004. Challenges of ecological complexity. Ecol. Complexity 1, 3–6.

Ludwig, J.A., Cornelius, J.M., 1987. Locating discontinuities along ecological gradients. Ecology 68, 448–450.

Mandelbrot, B.B., 1983. The Fractal Geometry of Nature. W.H. Freeman, New York.

Mandelbrot, B.B., 1999. Multifractals and 1/f Noise. Springer–Verlag, New York.

Martin-Lof, P., 1966. The definition of random sequences. Inform. Control 9, 602–619.

Maurer, B.A., 2005. Statistical mechanics of complex ecological aggregates. Ecol. Complexity 2, 71–85.

Milne, B.T., Johnson, A.R., Keitt, T.H., Hatfield, C.A., David, J., Hraber, P., 1996. Detection of critical densities associated with pinon-juniper woodland ecotones. Ecology 77, 805–821.

Moellering, H., Tobler, W., 1972. Geographical variances. Geogr. Anal. 4, 34–50.

Okubo, A., 1981. Diffusion and Ecological Problems: Mathematical Models. Springer–Verlag, New York.

O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H., Graham, R.L., 1988. Indices of landscape pattern. Landscape Ecol. 1, 153–162.

O'Neill, R.V., Gardner, R.H., Milne, B.T., Turner, M.G., Jackson, B., 1991. Heterogeneity and spatial hierarchies. In: Kolasa, J.K., Pickett, S.T.A. (Eds.), Ecological Heterogeneity. Springer–Verlag, New York, pp. 84–96.

O'Neill, R.V., King, A.W., 1998. Homage to St. Michael; or, why are there so many books on scale? In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale: Theory and Applications. Columbia University Press, New York, pp. 3–15.

Orloci, L., Beshir, M.E., 1976. A heuristic test for homogeneity in species populations. Vegetatio 31, 141–145.

Parisi, A., Caldarelli, G., 2000. Self-affine properties of fractures in brittle materials. Phys. A 280, 161–165.

Peitgen, H., Jurgens, H., Saupe, D., 1992. Chaos and Fractals: New Frontiers of Science. Springer–Verlag, New York.

Plotnick, R.E., Gardner, R.H., O'Neill, R.V., 1993. Lacunarity indices as measures of landscape texture. Landscape Ecol. 8, 201–211.

Plotnick, R.E., Gardner, R.H., Hargrove, W.W., Prestegaard, K., Perlmutter, M., 1996. Lacunarity analysis—a general technique for the analysis of spatial patterns. Phys. Rev. A 53, 5461–5468.

Rauch, E.M., Sayama, H., Bar-Yam, Y., 2003. Dynamics and genealogy of strains in spatially extended host pathogen models. J. Theor. Biol. 221, 655–664.

Ricotta, C., Avena, G., Marchetti, M., 1999. The flaming sandpile: self-organized criticality and wildfires. Ecol. Model. 119, 73–77.

Saunders, S.C., Chen, J., Drummer, T.D., Gustafson, E.J., Brosofske, K.D., 2005. Identifying scales of pattern in ecological data: a comparison of lacunarity, spectral and wavelet analyses. Ecol. Complexity 2, 87–105.

Schneider, D.C., 2001. The rise of the concept of scale in ecology. Bioscience 51, 545–553.

Sole, R.V., Manrubia, S.C., 1995. Are rainforests self-organized in a critical state? J. Theor. Biol. 173, 31–40.

Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. Science 236, 787–792.

Slatkin, M., Voelm, L., 1991. Fst in a hierarchical island model. Genetics 127, 627–629.

Stanley, H.E., Amaral, L.A.N., Gopikrishnan, P., Ivanov, P.Ch., Keitt, T.H., Plerou, V., 2000. Scale invariance and universality: organizing principles in complex systems. Phys. A 281, 60–68.

Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46, 234–240.

Turner, M., Gardner, R.H., O'Neill, R.V., 2001. Landscape Ecology in Theory and Practice: Pattern and Process. Springer–Verlag, New York.

Usher, M.B., 1975. Analysis of pattern in real and artificial plant populations. J. Ecol. 63, 569–585.

van der Maarel, E., 1976. On the establishment of plant community boundaries. Berichte der Deutschen Botanischen Gesellschaft 89, 415–443.

Vicsek, T., 2001. Introduction to complex patterns, fluctuations, and scaling. In: Vicsek, T. (Ed.), Fluctuations and Scaling in Biology. Oxford University Press, Oxford, pp. 15–37.

von Bertalanffy, L., 1968. General Systems Theory. Braziller, New York.

Wiens, J.A., 1989. Spatial scaling in ecology. Funct. Ecol. 3, 385–397.

With, K.A., Crist, T.O., 1995. Critical thresholds in species' responses to landscape structure. Ecology 76, 2446–2459.

With, K.A., King, A.W., 1999. Dispersal success on fractal landscapes: a consequence of lacunarity thresholds. Landscape Ecol. 14, 73–82.

Wright, S., 1931. Evolution in Mendelian populations. Genetics 16, 97–159.

Wright, S., 1951. The genetical structure of populations. Ann. Eugenics 15, 323–354.

Wu, X.B., Thurow, T.L., Whisenant, S.G., 2000. Fragmentation and changes in hydrologic function of tiger bush landscapes, southwest Niger. J. Ecol. 88, 790–800.