3 A COMPARISON OF NESTED ANALYSIS OF VARIANCE AND GEOSTATISTICS FOR CHARACTERIZING THE SPATIAL STRUCTURE OF PATCHY GRASSLAND LANDSCAPES

3.0 Citation

Davidson, A and Csillag, F. *A comparison of Nested ANOVA and Geostatistics for characterizing the spatial structure of patchy grassland under a limited sampling budget: a study using simulated landscapes*. Canadian Journal of Remote Sensing (in review).

3.1 Abstract

Ecological processes act upon grassland landscapes over a wide range of spatial and temporal extents, and interact to create complex and ever-changing landscape patterns. The challenge is to monitor and measure changing pattern in a way that can shed light on the ecological processes responsible for this change. While remote sensing provides such an approach, commercially available satellite sensors are unable to resolve spatial pattern to the detail required for finer scale analyses, such as the linking of biogeochemical cycles to community-level processes (e.g. patch formation). In such cases, it is necessary to sample heterogeneity directly on the ground, often using ground-based radiometers. We propose that the sampling methodology and the associated statistical assumptions used to characterize spatio-temporal heterogeneity will strongly influence perceived landscape structure. We hypothesize that nested sampling schemes based upon random-effects analysis of variance (ANOVA) will provide a more reliable characterization of spatial structure than that provided by spatial dependence models of geostatistics whose basic assumption of stationarity is often violated over patchy landscapes. To test this hypothesis, we simulated a variety of stationary and non-stationary grassland landscapes whose spatial structures were known, then compared the spatial structures as estimated by nested ANOVA and geostatistics under a limited sampling budget. Our results showed that nested ANOVA consistently provided a more stable characterization of structure than that provided by geostatistics for all non-stationary and stationary landscapes. Furthermore, the statistical independence of the components of nested ANOVA allowed for successful significance testing between sampling resolutions, and hence, the identification of potential spatial scale hierarchy. We also showed that semivariograms provided highly variable estimates of spatial structure, and this was dependent on the sampling scheme utilized but independent of the underlying spatial properties of the landscape (i.e. non-stationarity vs. stationarity). However, despite the implications of our results for those wishing to characterize grassland spatial structure through ground-based sampling, three limitations must be borne in mind. These are (a) our limited ability to simulate “real” grasslands landscapes, (b) our testing of nested ANOVA using only few variants of our nested sampling approach, and (c) the need for similar comparisons under conditions where sampling budget is unlimited. Such issues need to be further investigated.

3.2 Introduction

Physical and chemical processes act upon grasslands over a wide range of temporal and spatial extents. These processes, both anthropogenic and natural, interact to create complex and ever-changing landscape patterns (Wu and Levin, 1994). Spatial pattern is an important property of grasslands because it not only results from dynamic ecological processes (Paine and Levin, 1981; Urban et al., 1987; Wu and Levin, 1997), but can also directly enhance or constrain these processes (O'Neill, 1996; Riera et al., 1998; Lobo et al., 1998). This dual relationship makes the monitoring and measurement of grassland heterogeneity key to understanding concomitant ecological processes (Turner, 1989a; Turner and Gardner, 1991; Henebry, 1993; Wu and Levin, 1994; Gustafson, 1998).

Spatial pattern (also often synonymously referred to as *spatial heterogeneity* or *patchiness*) can be generally defined as the amount and variability of a system property in space and time. *Spatial structure* is a subset of the concept of spatial pattern, and refers to the spatial components of heterogeneity (i.e. the way in which total landscape variability is partitioned across spatial scales) (Gustafson, 1998). The simplest cases of spatial pattern and structure occur when a landscape variable shows two discrete and unchanging states at one level of spatial resolution (i.e. scale of observation) (Kolasa and Rollo, 1991). Natural grassland systems, in comparison, are far more complex and tend to be patchy on virtually every level of space and time (Caswell and Cohen, 1991; Levin, 1993). Such landscapes are often usefully visualized as a series of “shifting mosaics” (Bormann and Likens, 1979). The challenge, however, is to monitor and measure changing pattern in a way that can shed light on the ecological processes responsible for that change (Henebry, 1993). This requires the characterization of both the spatial and temporal components of grassland heterogeneity (Gustafson, 1998; Turner, 1989a, 1989b) – particularly the spatial extents over which change is most prominent – and an understanding, at least conceptually, of how pattern and process interact. The knowledge of pattern leads to prediction, and prediction involves the basis for testing theory, and for taking action (Hawksworth and Kalin-Arroyo, 1995).

The various pattern-generating processes operating on the landscape, and the range of extents over which they operate determine the dominant spatial and temporal scales over which grassland heterogeneity is manifest (Picket and Cadenasso, 1995). In general, these operational scales covary; processes that operate over short temporal scales (i.e. fine temporal resolutions) also tend to operate over small spatial scales (i.e. fine spatial resolutions, short distances, small grains, small extents), while processes operating at long temporal scales also tend to do so at large spatial scales (Goodchild and Quattrochi, 1997). For instance, over short distances, the spatial heterogeneity of vegetation might be related to microenvironments or competitive processes; at medium distances it might be related to disturbances such as fire and/or grazing; and at long distances it may be related to climate (Urban et al., 1987; Cullinan et al., 1997; Riera et al., 1998). Sometimes, however, the relationship between process and pattern is less clear. The observed heterogeneity at a given scale may be a function of processes and constraints operating at higher and lower organizational levels which combine to influence pattern at intermediate resolutions (O'Neill et al., 1991). For example, increasing evidence suggests that small-scale spatio-temporal variation and differentiation in plant communities plays a key role in maintaining vegetation structure at larger spatial and temporal scales (Shmida and Ellner, 1989; Czárán and Bartha, 1992; Herben et al., 1993a, 1993b; Bartha et al., 1995).

The observed scale-dependency between pattern and process means that grasslands may appear homogeneous or heterogeneous, depending on the resolution of observation (Davis et al., 1991; Ehleringer and Field, 1993; Foody and Curran, 1994). Thus, the measurement of landscape change is not only reliant on some quantification of observed pattern (Turner and Gardner, 1991; Henebry, 1993), but also on the spatial and temporal resolutions over which pattern is observed (Riera et al., 1998). Resolution is important because it defines the limits to our observations of the earth (Goodchild and Quattrochi, 1997), and the explicit consideration of its effects on pattern is arguably as important as recognizing heterogeneity itself (McIntosh, 1991).

Remote sensing technologies offer a unique perspective for the observation and measurement of various terrestrial biophysical parameters over large grassland regions. As a result, these technologies have become one of the main tools used by landscape ecologists interested in quantifying the spatial pattern of vegetation (Wickham and Riitters, 1995). However, the ability to detect and measure vegetation heterogeneity using remote sensing is constrained by both the spatial and temporal resolutions of the sensor(s) utilized. Remote sensors are incapable of resolving landscape variability at spatial resolutions finer than that of the sensor itself, and thus, at sub-pixel resolutions, satellite-derived landscapes are spatially homogeneous. While the limits of sensor resolution on pattern identification and measurement are most severe where coarse-resolution satellite data (e.g. NOAA-derived AVHRR (1 to 4km)) are used, similar problems may also be encountered for finer-resolution imagery (e.g. Landsat MSS (50m), Landsat TM (30m), SPOT-P (10m)) where significant heterogeneity occurs at very localized extents. Furthermore, the lengthy repeat cycles of many satellites mean that it is rarely possible to collect sufficient data at fine enough resolutions to adequately understand the temporal behavior of the observed landscape. This is especially problematic where grasslands undergo rapid and significant changes in local heterogeneity, or where the investigator requires the explicit elucidation of finer-resolution pattern.

Because of these limitations, ground observations are also generally required to characterize vegetation pattern fom remote sensing data (Steven, 1987; Atkinson et al., 2000). These observations often take the form of spectral measurements, as derived from ground-based radiometry. Ground radiometers have been widely utilized within grassland research over the last twenty years (e.g. see Tucker, 1979; Girard, 1982; Asrar et al., 1986; Weiser et al., 1986; Turner et al., 1992; Pickup et al., 1993; Goodin and Henebry, 1997; Goodin and Henebry, 1998; Davidson and Csillag, 2001), and provide a number of advantages over other measurement techniques. First, their small fields of view (FOV) allow spectral data to be collected at spatial resolutions as fine as $1m^2$ when mounted at heights of only a few meters above the ground (see Webster et al. (1989) and others cited above). Second, their portability and ability to sample various biophysical parameters non-destructively allow for rapid and repeatable data collection over "fixed" sample points at times when data is most needed. Third, in comparison to satellite remote sensors that measure atmospherically attenuated radiation, the spectrally reflected signal measured by ground radiometers is free from atmospheric influences. Fourth, the calibration accuracy of most radiometers is known and maintained over the entire operational environmental range experienced in the field. Nevertheless, ground-based radiometry also has its limitations. Small fields of view mean that when ground radiometers are used to measure the mean spectral characteristics (or estimate the mean biophysical characteristics) of a region larger than a few tens of square meters, or to map their variation, some kind of sampling must be used because providing complete spatial coverage is impractical (Webster et al., 1989).

In such situations it is necessary to identify and implement effective sampling methodologies that are most suitable for the task(s) at hand. Several techniques have been developed to quantify the spatial heterogeneity of landscapes from sampled (point) information (see Legendre and Fortin, 1989; Levin, 1992; Bellehumeur and Legendre, 1998; Gustafson, 1998; Fortin, 1999 for a comprehensive review). Two such methods are those based upon random-effects nested analysis of variance (ANOVA) (see Lewis, 1978; Ludwig and Goodall, 1978; Webster, 1979; Oliver and Webster, 1986; Legendre and Fortin, 1989; Webster and Oliver, 1990a) and spatial dependence models of geostatistics (see Cressie, 1991; Ver Hoef et al., 1993; Burrough and McDonnell, 1998; Bellehumeur and Legendre, 1998; Gustafson, 1998; Fortin, 1999). A previous comparison of these two approaches (Bellehumeur and Legendre, 1998) suggests that they provide equivalent means for characterizing spatial structure under assumptions of landscape *stationarity* (see section 3.3.2 for theoretical details). However, data collection and analysis methods are generally designed to meet certain statistical criteria, and perform with specific statistical properties (Graniero, 1999). Our particular concerns are that (a) the variogram model of geostatistics implies certain stationarity criteria that can produce misleading results when violated, particularly in ecological applications over patchy landscapes, and (b) because a large number of samples is normally required to adequately characterize the variogram, nested sampling may be more efficient, especially where a sampling budget is limited (e.g. in terms of time or cost). As a result, we propose that sampling approaches based on random-effects nested ANOVA, which require fewer samples and less stringent assumptions than those of variogram models, may be more appropriate for sampling over non-stationary landscapes such as grasslands, where many nested levels of patchiness often occur.

This study compares random-effects ANOVA and geostatistics for characterizing grassland landscape heterogeneity over both stationary and non-stationary landscapes. Particularly, our aim is to compare both methodologies under a fixed (limited) sampling budget. To achieve this, we simulate a variety of grassland landscapes whose spatial properties are known, then compare the relative abilities of each sampling approach to characterize the spatial structure of each landscape. Although the contextual focus of our research is geared primarily towards the field sampling of spectral (and spectrally-derived biophysical) information, the same structural context can be applied to field sampling in general, such as that for various soil parameters.

3.3 Nested ANOVA and Geostatistics: Two measures of spatial heterogeneity

3.3.1 Spatially nested sampling and random-effects Analysis of Variance

The random-effects ANOVA approach to characterizing spatial heterogeneity applies classical ANOVA to a nested (hierarchical) design for the purpose of partitioning the total variance of a study variable into various components, each corresponding to a different spatial resolution of interest. In effect, this approach provides a good measure of landscape "patchiness" across sampled resolutions (Webster, 1979; Sokal and Rohlf, 1981; Oliver and Webster, 1986; Webster and Oliver, 1990b; Ver Hoef et al., 1993; Bellehumeur and Legendre, 1998). Spatial resolutions of interest are identified prior to field sampling, and are dependent on the surveyor's understanding of the resolutions at which most spatial variation occurs. To date, such surveys have been developed and used in a variety of disciplines, particularly the soil (e.g. Nortcliff, 1978; Webster, 1979; Oliver and Webster, 1986; Bringmark and Bringmark, 1998) and vegetation sciences (e.g. Greig-Smith, 1952; Ludwig and Goodall, 1978).

The application of the random-effects nested ANOVA model to spatial heterogeneity characterization is based on the idea that the spatial components of a population can be subdivided into distinct levels (i.e. resolutions), and that observations are viewed as the result of the nested contributions of these levels (Webster, 1979; Sokal and Rohlf, 1981; Oliver and Webster, 1986; Bellehumeur and Legendre, 1998). To illustrate this, consider a balanced spatially nested sampling scheme of *L* levels, where each unit in level-[1] is partitioned into *n* sub-units at level-[2] which, in turn, are further partitioned until level-[*L*] is reached (i.e. level-[*L*] is nested within level-[*L*-1] which is nested within level-[(*L*-1)-1], etc…). The overall average of a sampled landscape attribute ($\bar{x}$, an unbiased estimate of the mean), its total sum of squares (SSQ_{total}) and its total mean square (MSQ_{total}, an unbiased estimate of the variance) can be written as:

$$\bar{x} = \frac{1}{n_1 n_2 \ldots n_L} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \ldots \sum_{l=1}^{n_L} x_{ij\ldots l} \qquad (3.1)$$

$$SSQ_{total} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}...\sum_{l=1}^{n_L}(x_{ij...l} - \overline{x})^2 \tag{3.2}$$

$$MSQ_{total} = \frac{SSQ_{total}}{df_{total}} \tag{3.3}$$

where n_1, n_2 and n_L correspond to the number of subdivisions within each partition of the level above, x corresponds to the value of the landscape attribute at each sampled location, and df_{total} corresponds to the total degrees of freedom,

$$df_{total} = n_1 n_2 ... n_L - 1. \tag{3.4}$$

The SSQ_{total} (Equation (3.2)) can also be expressed in terms of the individual contributions from each level of the hierarchy:

$$SSQ_{total} = \sum_{i=1}^{n_1} n_2 n_3 ... n_L (\overline{x}_i - \overline{x})^2 + \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} n_3 ... n_L (\overline{x}_{ij} - \overline{x}_i)^2 + ... \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}...\sum_{L=1}^{n_L}(\overline{x}_{ij...l} - \overline{x}_{ij...l-1})^2 \tag{3.5}$$

where the three terms on the right hand side of the equation correspond to the *SSQ*s of level-[1], level-[2] and level-[*L*], respectively (Csillag and Kabos, 1996). The leftmost of these three terms gives the individual *SSQ* of at the coarsest level of sampling, while the rightmost term gives the individual *SSQ* at the finest scale (or often, replication error). For each partition at any given level, the difference between its mean and the mean of the partition to which it belongs at the level immediately above is squared and multiplied by the total number of observations that make up that partition (Webster, 1979). The stochastic independence of the individual *SSQ* terms under the usual assumptions of ANOVA (Scheffé, 1959; Webster, 1979; Csillag and Kabos, 1996) allows significance testing between any two consecutive levels by computing their *F* ratio:

$$F = \frac{MSQ_{level-[b]}}{MSQ_{level-[b+1]}} \tag{3.6}$$

where *b* corresponds to the “upper” (i.e. coarser) level of interest. Where the variance of a given level is significantly high, we reject the hypothesis that the means of this level and the level it is being compared to are equal. In effect, *F*-testing allows for statistically significant differences in patchiness between sampling resolutions to be identified.

3.3.2 Spatial dependence models of Geostatistics: The semivariogram

Spatial dependence models of geostatistics also allow the quantification of landscape spatial structure from point-sampled data. One such model that has received much attention is the semivariogram (see Webster and Oliver, 1990b; Cressie, 1991; Ver Hoef et al., 1993; Bellehumeur and Legendre, 1998; Gustafson, 1998; Fortin, 1999). The semivariogram reveals the randomness and structured aspects of the spatial dispersion of a given variable, and is a plot of the average square difference between the values of a spatial variable at pairs of points separated by a lag distance, against the lag (Figure 3.1). The *empirical semivariogram* is estimated by:

$$g(h) = \frac{1}{2n_h} \sum_{i=!}^{n} (z_i - z_{i+h})^2 \qquad (3.7)$$

where $\gamma(h)$ is the *semivariance*; z_i and z_{i+h} are measurements of a given random variable at locations *i* and *i*+*h*, separated by the distance *h*; and n_h is the number of pairs of sampling units considered in the given distance class. This calculation is then repeated for different intervals of *h*.

The shape of the empirical semivariogram obtained with sample data (often, for the sake of brevity, referred to as simply the variogram, $2\gamma(h)$) allows the description of the overall spatial pattern of sample data (Fortin, 1999). Of particular interest are the estimates of three parameters – the *sill*, the *range* and the *nugget variance* – which define the spatial autocorrelation structure of the sampled landscape (Figure 3.1). The rate of increase of $\gamma(h)$ allows one to characterize the spatial continuity of the variable (Bellehumeur and Legendre, 1998). In landscapes where there is some degree of spatial dependence, $\gamma(h)$ generally increases with *h* because pairs of sample points become increasingly different as larger distances separate them. As *h* increases further, sample points eventually become unrelated to one another so that $\gamma(h)$ levels off at a *sill* that is equal to the average variance of all samples. The distance at which this occurs is called the *range*, and at distances greater than this, pairs of sample

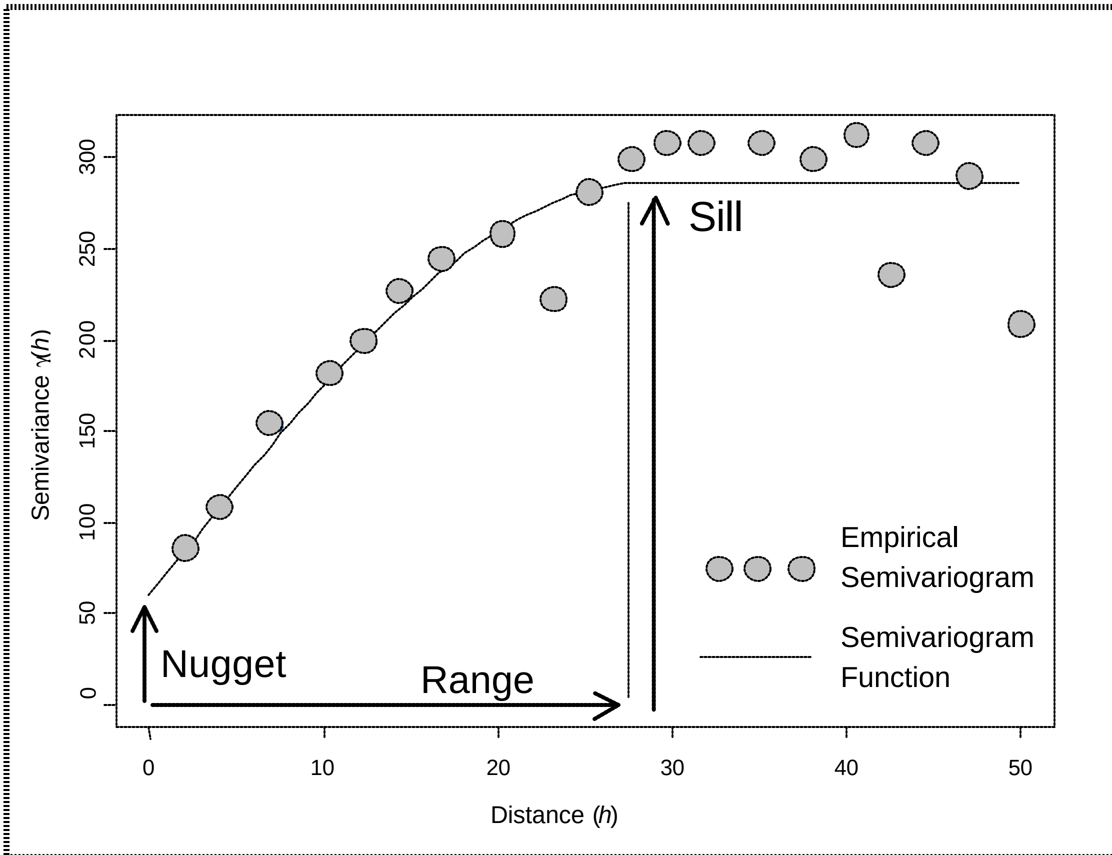

Figure 3.1. An empirical variogram (showing the nugget, range and sill) to which a spherical generalized least squares model has been fit (data derived from the transect sampling of a non-stationary simulations used in this study).

points can no longer be considered spatially correlated. The *nugget effect* refers to the non-zero intercept of the semivariogram, and is an overall estimate of error caused by measurement inaccuracy and environmental variability occurring at fine enough scales to be unresolved by the sampling interval (Deutsch and Journel, 1992; Bellehumeur and Legendre, 1998; Burrough and McDonnell, 1998; Fortin, 1999).

A *variogram function* is a theoretical model chosen to describe the spatial structure of a landscape attribute and to estimate $\gamma(h)$ at unknown values of h (Figure 3.1). However, this inference is valid only under the statistical assumption of *intrinsic stationarity*; that is, that the data are normally distributed around the same mean, with constant variance and isotropy across the entire study area (see Webster and Oliver, 1990b; Bellehumeur and Legendre, 1998). Such strict statistical requirements are often problematic, especially for ecological data, which are rarely stationary (Fortin, 1999). Fortunately, recent advances in this field have led to the development of more flexible techniques that rely on less stringent stationarity assumptions.

A variety of theoretical variogram models may be used to describe spatial structure. Of these, the most commonly used are the linear, spherical, exponential, gaussian and power models (see Journel and Huilbregts, 1978; McBratney and Webster, 1986; Webster and Oliver, 1990b; Deutsch and Journel, 1992; Burrough and McDonnell, 1998). Because most of the spatial dependence signal is encompassed in the first part of the variogram (i.e. up to the spatial range), good parameter fitting for short values of h are important (Burrough and McDonnell, 1998; Fortin, 1999). Such fitting is made easier by the development of various optimization techniques for variogram modeling. These numerical fitting procedures are either based on the likelihood principle (e.g. Maximum Likelihood; Conditional Likelihood), or least squares (e.g. Generalized Least Squares, Weighted Least Squares), and allow the elimination of many subjective elements from the analysis (Cressie, 1991; Fortin and Jacquez, 2000). However, subjectivity is not eliminated altogether; certain model features (e.g. stationarity decision; anisotropic/isotropic models; number of model components; initial starting parameters for non-interactive model fitting) must still be explicitly chosen by the investigator (Atkinson, 1999).

3.4 Methods

3.4.1 Landscape simulation

To properly evaluate the effectiveness of each approach for characterizing spatial heterogeneity, one must compare conclusions derived from field sample data to the “reality” of a completely known landscape. Graniero (1999) refers to this problem as the “paradox of self evaluation”; that is, we are only able to understand the landscape from field sampling itself, while our “truthing” knowledge is limited to the very techniques we wish to test. When faced with such a problem, investigators tend to choose a sampling technique based upon their expertise in the field, and their perception of how the landscape is structured. However, since sampling approaches (and their associated analysis methods) cannot be developed solely in the hope that their statistical assumptions meet that of the landscape under scrutiny, and since such approaches cannot be physically tested on the landscape itself (Graniero, 1999), alternative strategies must be implemented. One such approach is the use of computer-simulated landscapes whose spatial heterogeneity through space and time are fully known and controlled by the investigator. In this study, we use HQ-simulation and geostatistical simulation, respectively, to simulate a variety of non-stationary and stationary grassland landscapes.

3.4.1.1 Generation of non-stationary landscapes using HQ-simulation

Haar-Quadtree (HQ) simulation (see Englund, 1993; Csillag and Kabos, 1996) is a nested simulation implemented by wavelets, and a flexible and computationally efficient tool for simulation experiments. We used an HQ-simulation algorithm (Kertész and Csillag, unpublished) implemented within a GIS (Geographical Resource Analysis Support Systems (GRASS), Version 4.1) to generate a variety of non-stationary landscapes (hereon referred to as *scenarios*), each with a different spatial structure. The stochastic nature of this algorithm allowed us to then simulate an additional 9 equally likely landscapes (hereon referred to as *simulations*) for each given scenario; that is, landscapes that were the same in terms of their overall structural properties, but differed in terms of the relative spatial arrangement of that structure. This gave a total of 10 simulations for each scenario. The output of each simulation took the form of a regular grid measuring 256 rows by 256 columns (65536 cells). These dimensions allowed our

HQ-simulation algorithm to generate landscapes with up to eight levels of patchiness (Figure 3.2). Each grid was considered a surrogate of a “real” landscape measuring 128m by 128m at a spatial resolution of 0.5m, where the patchiness of the landscape attribute under scrutiny (say, for example, radiance, NDVI, biomass) was explicitly controlled at local (0.5m, 1m, 2m), intermediate (4m, 8m, 16m) and coarse (32m, 64m) resolutions. The non-stationary landscape scenarios generated in this study ranged in complexity, from landscapes where no spatial structure was evident, to simple and complex structures (Figure 3.3, Tables 3.1(a), (b), and (c)). In our first experiment, we generated a landscape scenario with no spatial structure (Table 3.1(a), Figure 3.3(a)). In our second experiment, we generated landscape scenarios with one, two and three distinct levels of patchiness (Table 3.1(b), Figure 3.3(b)). In our third experiment, we generated a variety of landscape scenarios whose multiple (i.e. > 4) levels of patchiness closely corresponded to the observed spatial structure of various upland mixed-grass prairie communities, as described by Csillag et al. (1997) and Davidson (Unpublished data) (Table 3.1(c), Figure 3.3(c)).

3.4.1.2 Generation of stationary landscapes using geostatistical (CAR) simulation

Geostatistical landscape simulation based on conditional autoregressive (CAR) models rely on conditional probability distributions and the Fast Fourier Transform (see Cressie, 1993; Dungan, 1999). We used a CAR-based simulator (see Cressie, 1993; Csillag et al., Unpublished Manuscript; Kabos and Csillag, Unpublished Manuscript) implemented within the S-PLUS statistical package (Version 4.5, MathSoft Inc., 1998) to simulate three stationary landscape scenarios, whose spatial autocorrelation structures were defined by exponential semivariograms of short, medium and long correlation lengths (Experiment 4; Figures 3.4(a), (b) and (c)). In total, we generated 10 different landscape simulations for each given scenario. The data structure of each geostatistical simulation was identical to those described for the non-stationary scenarios (i.e. a regular grid of 256 rows by 256 columns representing a “real” landscape of 128m by 128m at 0.5m resolution).

3.4.2 Sampling Design

We used four different sampling schemes (Figures 3.5(a) to (d)) to extract information from each

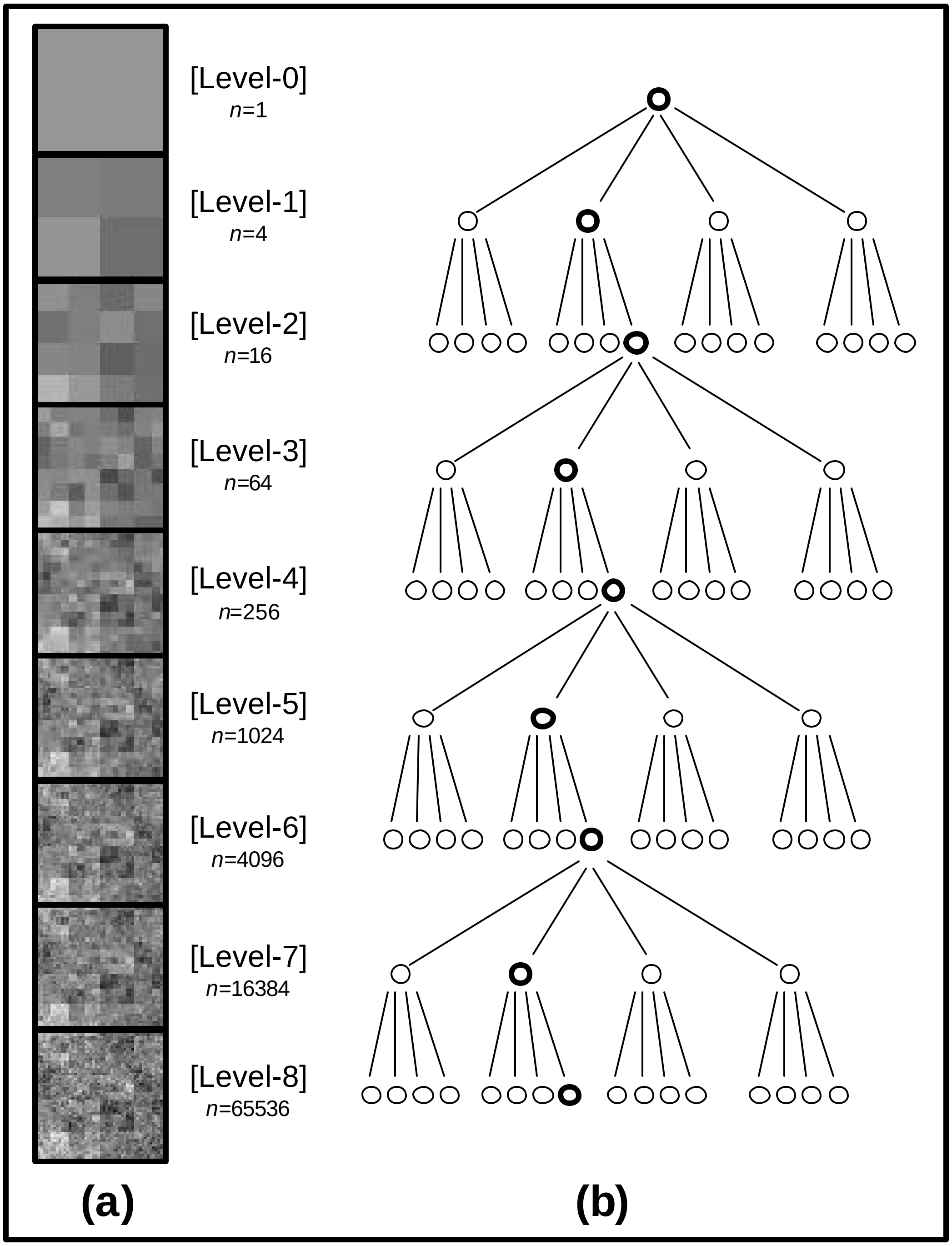

Figure 3.2. Using HQ-simulation to generate a patchy landscape. Level-0 corresponds to a homogeneous field that is partitioned (levels-1 to –8) according to the variance desired at each level of the hierarchy [(a)]. Decomposition proceeds along the “leaves” of the quadtree [(b)].

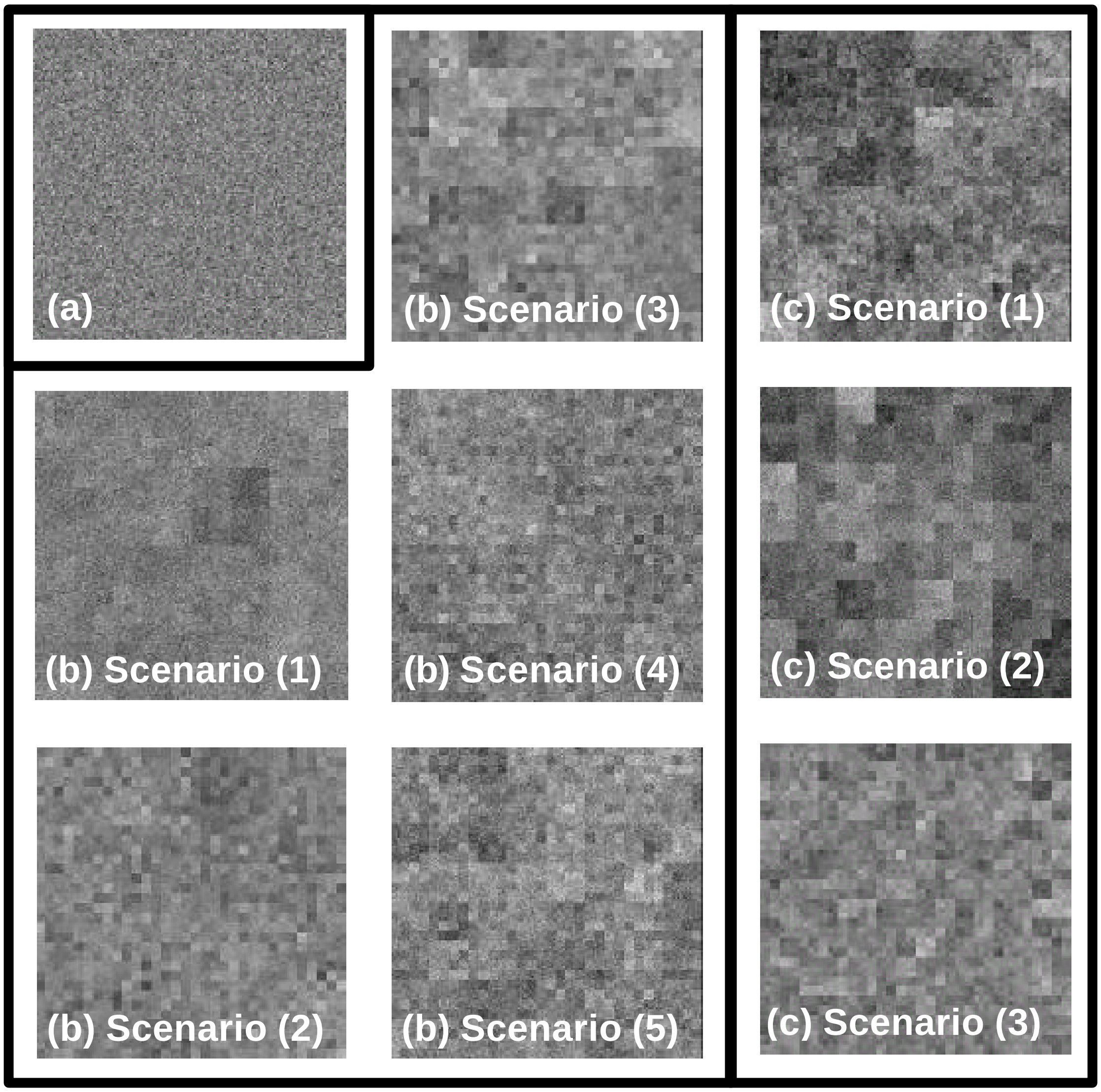

Figure 3.3. Non-stationary simulated landscapes used in (a) Experiment 1 (no spatial structure), (b) Experiment 2 (simple spatial structures), and (c) Experiment 3 (complex spatial structures; eight distinct levels of patchiness). Scenarios (b)(1) and (2) show one distinct level of patchiness; scenarios (b)(3) and (4) show two distinct levels of patchiness; scenario (b)(5) shows three distinct levels of patchiness. The structural characteristics of these simulations are given in Table 3.1.

(a)	Level	(64m)	(32m)	(16m)	(8m)	(4m)	(2m)	(1m)	(0.5m)
Expt. 1	%VAR	0.004	0.018	0.073	0.292	1.171	4.687	18.750	75.001
	MAX	100.1	100.2	101.9	102.5	105.4	110.9	122.5	157.2
	MIN	99.9	99.6	99.2	98.2	95.5	90.1	77.0	53.6

(b)		Level	(64m)	(32m)	(16m)	(8m)	(4m)	(2m)	(1m)	(0.5m)
Expt. 2	Scen. (1)	%VAR	6.7	6.7	6.7	6.7	6.7	6.7	6.7	53.3
		MAX	103.2	104.9	108.8	111.9	116.5	118.2	122.1	136.1
		MIN	96.9	95.3	90.6	85.5	84.8	79.9	77.8	61.5
	Scen. (2)	%VAR	6.7	6.7	6.7	6.7	6.7	53.3	6.7	6.7
		MAX	103.7	104.7	107.5	111.5	115.9	125.7	129.6	131.9
		MIN	97.5	94.8	92.0	87.9	86.6	70.0	65.1	61.8
	Scen. (3)	%VAR	4.5	4.5	4.5	36.4	4.5	36.4	4.5	4.5
		MAX	102.7	105.9	108.6	119.9	123.7	137.5	142.7	146.3
		MIN	96.4	93.9	88.7	80.8	78.6	66.2	64.6	61.4
	Scen. (4)	%VAR	4.5	4.5	4.5	4.5	4.5	36.4	4.5	36.4
		MAX	103.0	106.2	109.0	111.1	114.6	128.8	134.8	145.4
		MIN	97.1	94.3	90.1	88.4	86.1	67.4	64.8	55.2
	Scen. (5)	%VAR	3.4	3.4	3.4	27.6	3.4	27.6	3.4	27.6
		MAX	102.5	106.8	107.5	123.6	124.9	136.6	138.2	150.9
		MIN	96.6	95.8	91.5	75.2	72.9	59.7	58.3	48.4

(c)		Level	(64m)	(32m)	(16m)	(8m)	(4m)	(2m)	(1m)	(0.5m)
Expt. 3	Scen. (1)	%VAR	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5
		MAX	103.3	115.4	122.5	128.4	134.3	143.0	146.7	154.9
		MIN	92.6	88.5	83.1	78.5	72.1	67.5	58.2	51.3
	Scen. (2)	%VAR	18.35	17.75	15.94	14.28	10.61	4.72	0.001	97.30
		MAX	113.1	122.6	131.9	144.2	159.1	170.8	170.9	198.1
		MIN	85.5	79.2	60.5	49.7	33.4	23.9	23.8	3.3
	Scen. (3)	%VAR	0.003	0.211	1.63	8.47	43.38	31.96	14.33	0.003
		MAX	100.1	101.2	105.1	113.4	135.6	146.2	157.6	157.5
		MIN	99.9	98.6	93.6	88.8	65.4	48.4	44.3	44.1

Tables 3.1(a), (b) and (c). Structural characteristics of simulated landscape scenarios used in Experiments 1, 2 and 3. The variance partitioned to each level (resolution) is described in terms of its percentage of total variance (%VAR). The range of values simulated at each resolution is presented as maximum (MAX) and minimum (MIN) values.

of the non-stationary and stationary simulated landscapes described in section 3.4.1. These samples were then used to provide estimates of spatial structure using ANOVA and geostatistical approaches. Sampling schemes conformed to (a) nested, (b) random, (c) systematic and (d) transect designs. In order to provide a direct comparison between approaches, the sampling budget of each scheme was restricted to n = 72 samples. In the nested sampling design, sample points were spatially nested at distances of 60m, 10m, 2.5m and 0.5m. These distances were selected based on *a priori* knowledge of the patchiness of upland grassland landscapes, and to capture measurement error (0.5m) and fine (2.5m) to coarse (60m) resolution patchiness. In the systematic sampling design, sample points were separated by distances of 12.5m (X direction) and 14m (Y direction). In the transect sampling design, sample points were situated 0.5m apart. Twenty-four variants of each sampling scheme were utilized in the study. For the nested, systematic and transect designs, each variant was created by randomly shifting the origin of each sampling frame relative to that of the simulated landscape. For the random sampling design, each variant was created by simply creating a new random sample. Sample data from each simulation were then extracted using each of these variants. This approach gave a total of 240 "sampling attempts" for each combination of sampling scheme and landscape scenario (i.e. 10 simulations for each scenario x 24 sampling variants). For each of these combinations, nested ANOVA was then used to provide 240 estimates of spatial structure using scheme (a), while semivariogram analysis was used to provide 120 estimates of spatial structure from each of schemes (b), (c) and (d).

3.4.3 Statistical Methods

We compared each of the above estimates of spatial structure with the known spatial properties of the landscapes from which they were sampled. However, because these known structures were described in different ways (i.e. non-stationary fields were described by the explicit partitioning of total variance among various levels of the hierarchy (Table 3.1), while stationary fields were described by a smooth semivariogram function (Figure 3.4)), we used two different approaches to assess the overall ability of each sampling scheme to correctly characterize landscape spatial structure. In both approaches, all statistical analyses were conducted using the S-PLUS statistical package (Version 4.5, MathSoft Inc.,

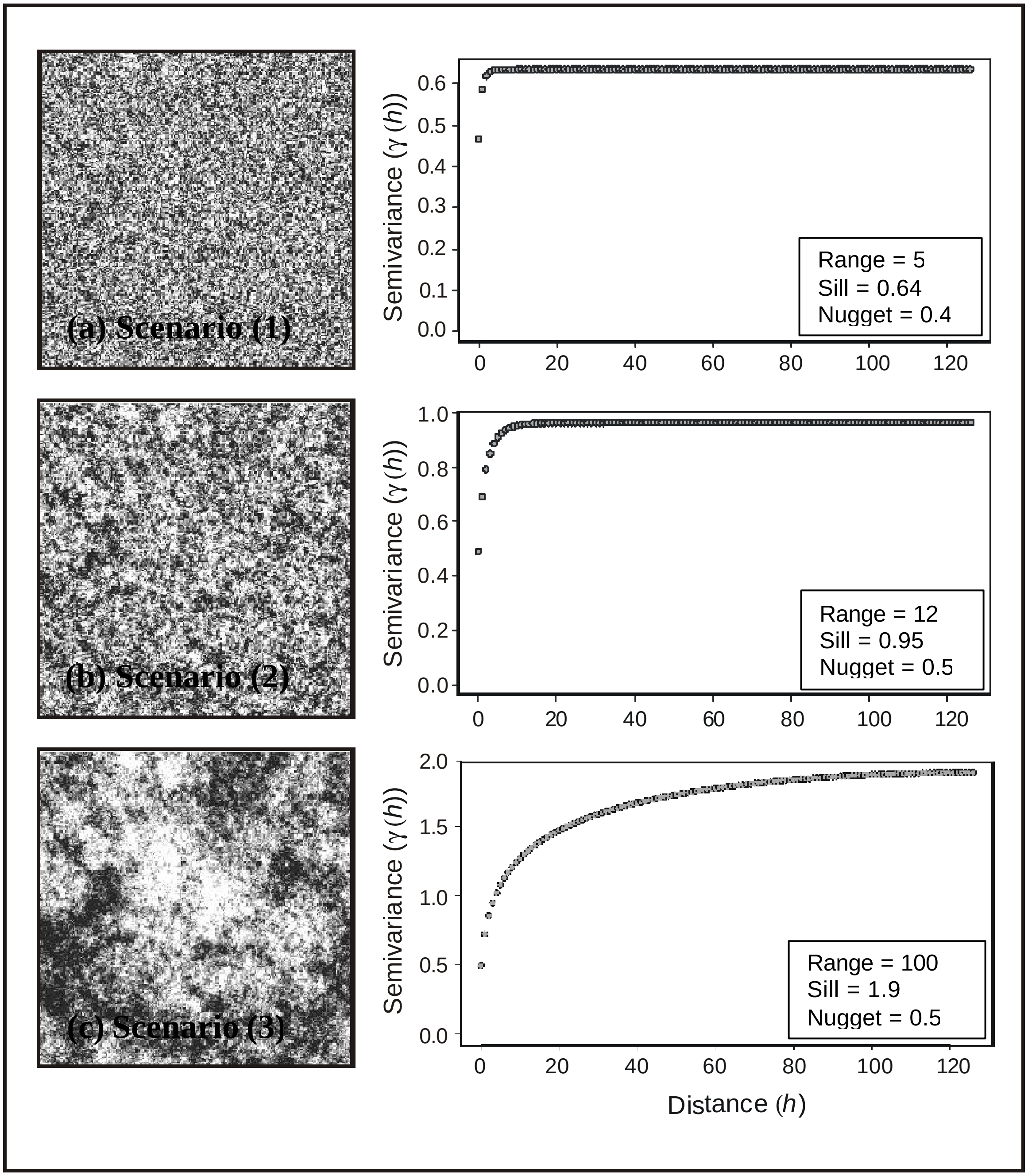

Figure 3.4. Simulated stationary landscapes used in Experiment 4. Landscapes were simulated with (a) short, (b) medium, and (c) long correlation lengths. The exponential variograms used to parameterize each simulation, and their associated sills, ranges and nuggets are also given (see text for details).

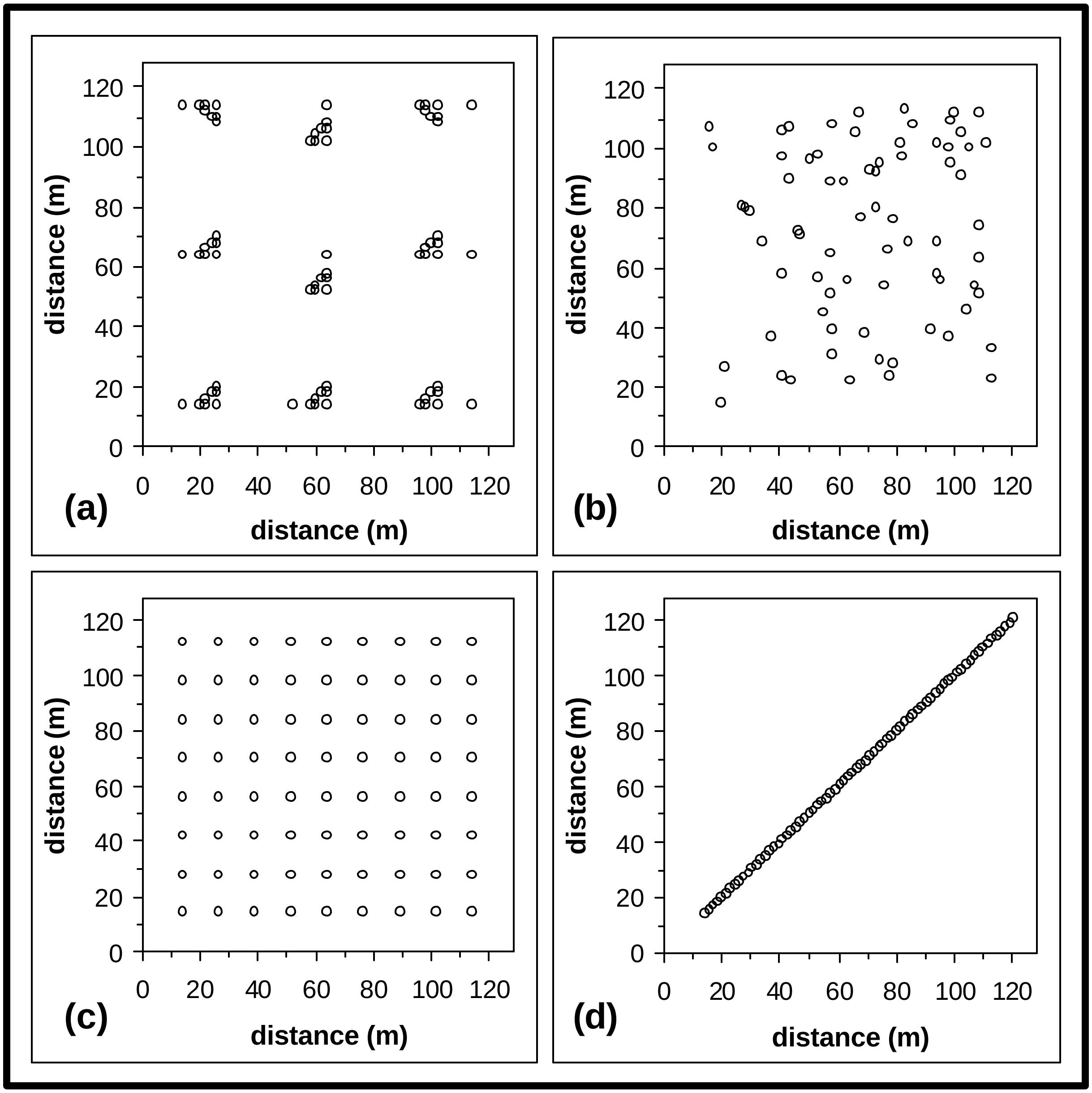

Figure 3.5. Sampling schemes used to extract information from each simulated landscape. These are (a) nested sampling, (b) random sampling, (c) systematic sampling, and (d) transect sampling. Nested Analysis of Variance (ANOVA) was used to estimate spatial structure from data sampled using scheme (a), while variogram analysis was used to estimate spatial structure using schemes (b), (c) and (d) (see text for details).

1998). All geostatistical (semivariogram) analyses were conducted using the S-PLUS SpatialStats Module (Version 1.0, MathSoft Inc., 1996).

3.4.3.1 Comparing estimated and known spatial structures of non-stationary scenarios

We applied nested ANOVA (random-effects model) to our nested sampling data derived from Experiments 1, 2 and 3. This partitioned the total sampled variance of each non-stationary landscape simulation into its various spatial components (i.e. estimated landscape patchiness at resolutions of 60m, 10m, 2.5m, 0.5m), and yielded 240 ANOVA-derived estimates of spatial structure for each landscape scenario. We then used Equation 3.6 (using both $P < 0.05$ and $P < 0.01$) to estimate whether significant differences in variances (i.e. significant differences in patchiness) existed between consecutive sample resolutions (i.e. between 60m and 10m, 10m and 2.5m, 2.5m and 0.5m). These results were then compared directly to the known characteristics of each landscape scenario (see Figure 3.3, Table 3.1). For each of these scenarios, the ability of nested ANOVA to correctly estimate expected significant or non-significant differences between consecutive sample resolutions was expressed as a percentage (n = 240).

We then used a robust estimator (Cressie and Hawkins, 1980) to generate 720 empirical semivariograms for each non-stationary landscape scenario (i.e. 240 semivariograms for each of the random, systematic and transect sampling schemes). The advantage of using a robust estimator was that the effects of outliers were reduced, without removing specific data points from the data set. Semivariance was computed for distances only where greater than 30 pairs of points existed (see Fortin, 1999). We then chose a random subset (n = 120) of semivariograms for each combination of scenario and sampling scheme in order to evaluate their overall ability to detect patchiness at the various expected spatial scales. A subset of this size was chosen because of the interactive (and thus, time-consuming) nature of semivariogram interpretation and fitting. A theoretical variogram function was then fit to each empirical semivariogram in this subset using a spherical model by generalized least squares (see Cressie, 1985). The generalized least squares approach has been shown, in most cases, to be as good as many of the more complicated and computationally intensive variogram fitting methods (Zimmerman and Zimmerman, 1991). While our approach provided us with only half of the possible landscape structure characterizations for each combination of landscape scenario and sampling strategy, the sample size was large enough to allow us, with confidence, to estimate the potential variation in estimated variogram parameters (i.e. sill, range, nugget) for each scenario. This variation was then compared directly to the known landscape structures outlined in Table 3.1.

3.4.3.2 Comparing estimated and known spatial structures of stationary scenarios

We were unable to directly compare ANOVA-derived estimates of the spatial structure of stationary landscapes (Experiment 4) with their corresponding known (simulated) structures. This limitation was a direct consequence of the way in which our geostatistical simulator described landscape spatial structure, particularly the individual components of variance. For semivariograms, semivariances at different distances are statistically dependent on each other, and thus significance testing between components – such as that possible with nested ANOVA – could not be undertaken. This meant that it was not possible to determine whether the differences in semivariance between distance classes, as illustrated by the semivariogram, were statistically significant. As a result, our consideration of the ability of nested ANOVA to characterize the spatial structure of stationary landscapes is restricted to the overall "stability" of the approach; that is, the consistency to which it estimates significant / non-significant differences between sample resolutions, rather than the degree to which such estimates are correct.

In comparison, we were able to directly compare semivariograms generated through the random, systematic and transect sampling of stationary landscapes with the semivariograms used to simulate and describe each landscape. Our creation of empirical and theoretical semivariograms from these sampled fields followed the procedures outlined in section 3.4.3.1, but with one important exception. While we utilized spherical semivariogram models to characterize the spatial structure of non-stationary landscapes, the spatial autocorrelation structure of stationary scenarios was expressed by exponential semivariogram models (Figure 3.4). Thus, to provide a direct comparison, we fitted exponential semivariograms, instead of spherical semivariograms, to the sample data derived from stationary landscapes.

3.5 Results

3.5.1 Nested sampling of non-stationary landscape simulations

We used nested sampling and random-effects ANOVA to estimate whether significant differences in variance existed between consecutive pairs of sample levels (i.e. 60m and 10m, 10m and 2.5m, 2.5m and 0.5m) for each of the non-stationary simulated landscapes used in Experiments 1, 2 and 3. The results of these analyses are illustrated in Tables 3.2(a) (for $P < 0.01$) and (b) (for $P < 0.05$). To avoid simply repeating the information presented in Table 3.2, our following discussions serve to highlight only the most extreme results. Unless stated otherwise, results refer to conservative significance testing using $P = 0.01$.

Table 3.2 shows that, for the most part, nested ANOVA is able to correctly estimate both significant and non-significant differences in variance between pairs of levels. In Experiment 1 (no spatial structure; random fields), the expected non-significant differences between 60m and 10m, 10m and 2.5m, and 2.5m and 0.5m are estimated correctly by nested ANOVA in 98%, 100% and 85% of all sampling attempts, respectively. The use of a less conservative test (i.e. $P = 0.05$; Table 3.2(b)) results in only slightly lower success (90%, 99% and 75%, respectively). In Experiment 2 (simple landscape structure), the expected non-significant and significant differences between pairs of levels are also estimated consistently well, and this is independent of landscape complexity. Here, the expected non-significant differences between 60m and 10m are estimated correctly in 85% of all sampling attempts at worst (scenario (2)), and in 93% of all sampling attempts at best (scenario (4)). Estimates of the non-significant differences between 10m and 2.5m are highly accurate, and are estimated correctly in 99% of all sampling attempts at worst (scenarios (2), (3) and (4)), and in 100% of all sampling attempts at best (scenarios (1) and (5)). Estimates of non-significant differences between 2.5m and 0.5m are less accurate (ranging from 81% (scenario (3)) to 90% of all sampling attempts (scenario (5)), but estimates of the expected significant differences between these two levels are high (ranging from 93% (scenario (1)) to 94% of all sampling attempts (scenario (2)). Similarly high predictive capabilities were also found in Experiment 3 (multiple levels of distinct patchiness) and, like Experiment 2, these trends are independent

(a) Levels	Expt. 1	Expt. 2					Expt. 3		
	Scen. (1)	Scen. (1)	Scen. (2)	Scen. (3)	Scen. (4)	Scen. (5)	Scen. (1)	Scen. (2)	Scen. (3)
60m & 10m	98	92	85	88	93	90	78	82	98
10m & 2.5m	100	100	99	99	99	100	98	64	100
2.5m & 0.5m	85	94	93	81	90	80	78	95	93

(b) Levels	Expt. 1	Expt. 2					Expt. 3		
	Scen. (1)	Scen. (1)	Scen. (2)	Scen. (3)	Scen. (4)	Scen. (5)	Scen. (1)	Scen. (2)	Scen. (3)
60m & 10m	90	73	65	66	75	73	60	61	82
10m & 2.5m	99	99	96	98	98	99	97	75	97
2.5m & 0.5m	75	98	97	71	78	64	67	97	96

Table 3.2. The ability of nested ANOVA to correctly estimate significant or non-significant differences in variance (patchiness) between 60m, 10m, 2.5m and 0.5m resolutions from the non-stationary simulated landscapes used in Experiments 1, 2 and 3. Figures correspond to the percentage (%) of correctly estimated differences between each pair of levels (P = 0.01 (a), P = 0.05 (b)). Shaded cells indicate levels between which significant differences (P << 0.01) in variance were expected. Unshaded cells indicate levels between which non-significant differences (P > 0.05) in variance were expected.

of landscape complexity. However, here the worst estimates of non-significant and significant differences were found (78% of all sample attempts between both 60m and 10m and 10m and 2.5m (scenario (1)), and 64% of all sample attempts between 10m and 2.5m (scenario (2)), respectively). In all Experiments, with the exception of pairs of levels between which significant differences were expected, the use of a more conservative test (i.e. P = 0.01) yielded more accurate estimates of nonsignificant differences between levels by nested ANOVA. Conversely, where a less conservative test was used (i.e. P = 0.05), a higher success rate was found for the estimation of significant differences between levels (see corresponding cells in Tables 3.2(a) and (b)).

3.5.2 Semivariograms of non-stationary landscape simulations

3.5.2.1 Empirical Semivariograms

Prior to fitting spherical semivariogram functions to our sampled data, we analyzed a subset of empirical semivariograms (n=120 for each combination of sampling approach and scenario) for their overall ability to detect patchiness at the various expected scales of variation. “Typical” examples of these semivariograms are provided in Figure 3.6. In most cases, the sampling of random fields (Experiment 1) produces “noisy” semivariograms, where the semivariance comprises mostly nugget variance and is independent of the distance between sample points (Figures 3.6(1), (2) and (3)). In comparison, the trends in semivariance derived from the sampling of non-stationary structured fields (Experiments 2 and 3) are less consistent both within and among sampling schemes, but are often more evident. For the sake of brevity, we do not provide a detailed discussion of the trends derived for each scenario here. Instead, we focus on two examples, one with a simple landscape structure (Experiment 2) and one with a complex landscape structure (Experiment 3), to illustrate the most apparent trends.

Figures 3.6(4), (5) and (6) show empirical semivariograms generated by random, systematic and nested sampling, respectively, for a simple landscape structure (Experiment 2, scenario (2), one level of distinct patchiness at 2m resolution). While each semivariogram (particularly that derived from transect sampling) indicate the presence of patchiness at longer distances (inferred by the semi-periodic signals

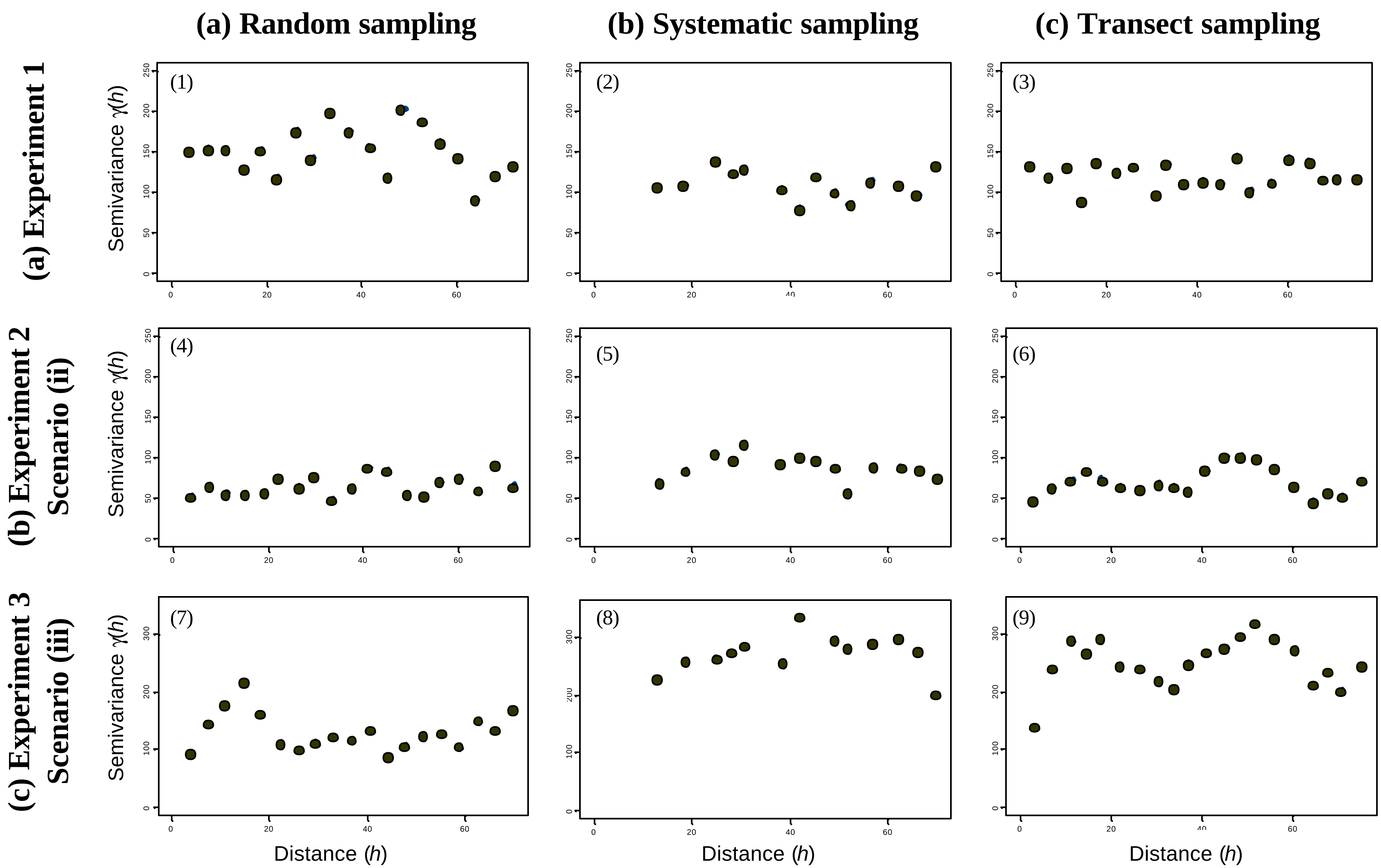

Figure 3.6. Empirical semivariograms derived from various combinations of sampling (random, systematic and transect) and non-stationary simulated landscapes of various complexity (Experiment 1(no structure); Experiment 2, Scenario (2) (one level of distinct patchiness); Experiment 3, Scenario (3) (8 levels of distinct patchiness)). See Tables 3.1(a) to (c) for detailed structural characteristics of these landscapes.

within the semivariogram), none detect the presence of considerable local variation at a distance of *h* = 2m. Figures 3.6(7), (8) and (9) also show empirical semivariograms generated by random, systematic and nested sampling, but for a more complex landscape structure (Experiment 3, Scenario (3), multiple levels of distinct patchiness). Here, a more pronounced pattern is detected at short distances by random and transect sampling, and at longer distances by all sampling approaches. However, the expected local variation at distances of *h* = 2m, *h* = 4m and *h* = 8m remains undetected. Note that empirical semivariograms derived from transect and random sampling are able to provide semivariances across a wider range of distances (minimum distances are *h* = 3m and *h* = 4m, respectively), compared to that of systematic sampling (minimum distance is *h* = 15m).

3.5.2.2 Theoretical (Spherical) Model Fitting

We used spherical model fitting by generalized least squares to provide 120 estimates of spatial structure for each combination of non-stationary landscape scenario and sampling strategy. The results of these analyses are summarized in Tables 3.3(a) to (c), and provided graphically in Figures 3.7, 3.8 and 3.9. While we acknowledge that other theoretical models (e.g. linear, exponential, may be a more appropriate fit for some of our data, we believe that spherical models are sufficient for illustrative purposes.

Figures 3.7(a) to (c) show semivariograms created by fitting a spherical model to empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of simulated fields with no spatial structure (Experiment 1). In most cases, and particularly for transect sampling, the spherical model fits semivariograms with very small ranges (i.e. << 1) to our data. (Note: where Figures 3.7, 3.8 and 3.9 display theoretical model fits as horizontal lines, they are, in actuality, estimated to be spherically -shaped, but with a very small range. This range, however, falls outside the range of our empirical data and is thus not plotted by our statistical software package). The most consistent characterization of the random fields is provided by transect sampling, where greater than 99% of the spherical fits estimate a range of less than 0.00001 (Figure 3.7(c)). Semivariograms generated using random samples provide the

(a) Random		Nugget			Range			Sill		
		Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.
Expt. 1	(1)	6.4 (7.9)	0	37	2.6 (6.8)	0	30	127.9 (28.5)	62	224
Expt. 2	(1)	13.9 (18.1)	0	71	11.1 (13.1)	0	75	53.0 (16.0)	14	101
	(2)	11.5 (14.1)	0	65	10.7 (13)	0	75	96.7 (27.0)	41	188
	(3)	11.4 (24.0)	0	135	4.3 (10.0)	0	50	140.3 (27.3)	98	206
	(4)	17.1 (19.1)	0	69	18.3 (24.9)	0	134	53.5 (26.5)	0	100
	(5)	12.3 (16.7)	0	80	5.5 (15.1)	0	97	95.3 (28.5)	30	161
Expt. 3	(1)	21.9 (26.6)	0	117	31.2 (138.1)	0	924	123.1 (84.0)	38	612
	(2)*	53.1 (85.0)	0	338	24.8 (32.1)	0	173	379.3 (110)	167	728
	(3)*	8.1 (26.5)	0	143	2.9 (8.1)	0	36	219.9 (52)	91	365

(b) Systematic		Nugget			Range			Sill		
		Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.
Expt. 1	(1)	14.0 (23.4)	0	110	11.6 (12.2)	0	75	118.7 (31.6)	0	194
Expt. 2	(1)	15.1 (18.2)	0	73	24.0 (13.1)	3	60	52.9 (15.1)	15	110
	(2)	20.4 (34.2)	0	119	24.5 (28.0)	0	161	88.3 (35.1)	0	162
	(3)	9.0 (26.1)	0	109	13.0 (10.0)	0	53	136.2 (33.5)	10	200
	(4)*	8.9 (14.5)	0	54	22.3 (10.5)	0	80	59.3 (17.5)	2	104
	(5)*	15.6 (28.2)	0	107	21.5 (17.6)	0	93	92.6 (23.3)	0	154
Expt. 3	(1)	25.1 (23.1)	0	91	14.5 (12.7)	2	65	114.1 (31.0)	15	190
	(2)*	84.0 (82.6)	0	302	31.6 (25.5)	0	131	367.9 (126)	100	914
	(3)*	18.4 (43.1)	0	197	8.6 (12.6)	0	64	207 (53.6)	68	316

(c) Transect		Nugget			Range			Sill		
		Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.
Expt. 1	(1)	14.1 (4.9)	5	58	0.3 (1.2)	0	6	120.7 (26.9)	51	207
Expt. 2	(1)	16.2 (15.1)	0	65	15.1 (21.1)	0	82	45.3 (18.1)	23	99
	(2)	23.3 (16.9)	0	75	16.7 (20.2)	0	164	30.1 (86.5)	37	183
	(3)	25.0 (20.0)	0	101	9.0 (11.8)	0	42	120.1 (43.5)	41	217
	(4)*	23.3 (14.7)	9	71	22.0 (60.0)	0	436	49.5 (36.0)	22	255
	(5)*	15.3 (8.5)	10	86	3.5 (5.5)	0	27	87.9 (19.0)	48	141
Expt. 3	(1)	30.1 (32.1)	0	134	40.1 (25.1)	0	214	117.0 (90.4)	46	521
	(2)*	55.0 (49.8)	0	221	22.1 (22.9)	0	141	316 (140.0)	16	804
	(3)*	20.1 (29.0)	0	136	6.8 (7.7)	0	25	204.0 (50)	108	320

Tables 3.3(a), (b) and (c). The mean, standard deviation (StD), minimum (Min.) and maximum (Max.) values of variogram parameters (Nugget, Range, Sill) resulting from fitting a spherical model by generalized least squares to sample data derived from non-stationary fields using different sampling designs (Random sampling (a); Systematic sampling (b); Transect sampling (c)). * indicate where extreme values were excluded from the summary statistics (see text for details).

next most consistent characterization of random spatial structure. Here, 78% of the spherical fits estimate a range of less than 0.00001 (Figure 3.7(b)). Systematic sampling provides the least consistent characterization of random fields where, in comparison to the other sampling approaches, a significant proportion (84%) of the spherical fits estimate some degree of spatial structure (i.e. estimate ranges > 5m)(Figure 3.7(b)). Also of interest is the general variation in sill estimates. We kept the sample variance constant throughout our simulation of random fields. However, spherical semivariogram fitting estimates a wide range in total sample variances (denoted by the variation in the estimated sills (Tables 3.3(a) to (c); Figures 3.7, 3.8 and 3.9)). This variation is independent of sampling strategy (mean sills and their standard deviations are relatively similar between sampling approaches).

Figures 3.8(a) to (c) show semivariograms created from fitting a spherical model to empirical semivariograms derived from non-stationary simulated fields with simple spatial structure (Experiment 2, scenario (2)). Contrary to Experiment 1, the parameter estimates generated by model fitting are dependent upon the sampling strategy utilized (Tables 3.3(a) to (c)). For all sampling strategies, the estimation of some spatial structure (i.e. the identification of a range > 1m) is more frequent than that estimated for the random fields. However, the expected local variation at 2m is still only estimated correctly in very few cases (in 8% of model fits using random sampling; in 12% of model fits using systematic sampling, and in 4% of model fits using transect sampling).

Figures 3.9(a) to (c) show semivariograms created from fitting a spherical model to empirical semivariograms derived from non-stationary simulated fields with complex spatial structure (Experiment 3, scenario (3)). Similar to that of the simply structured fields, parameter estimates generated by model fitting are dependent upon the sampling strategy utilized (Tables 3.3(a) to (c)). Here, random sampling overwhelmingly estimated semivariograms with very short ranges (82% of estimates). Systematic and transect sampling estimated very short ranges in 36% and 42% of all cases, respectively. In all cases, spherical model fitting indicated less frequent spatial structure at short distances than that estimated in Experiment 2 for the simple landscapes.

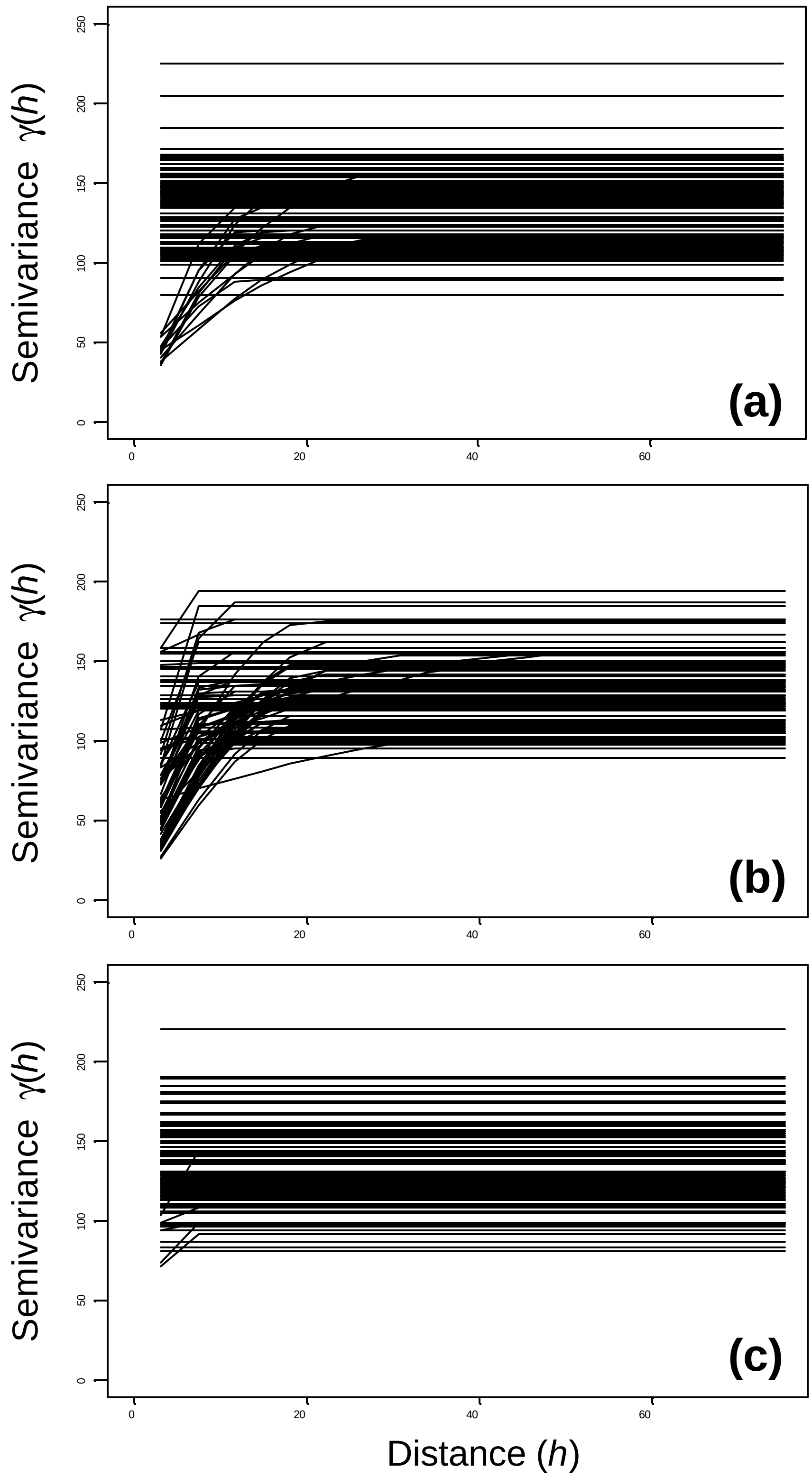

Figure 3.7. Theoretical semivariograms created from fitting a spherical model by generalized least squares to the empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of non-stationary simulated fields with no spatial structure (Experiment 1).

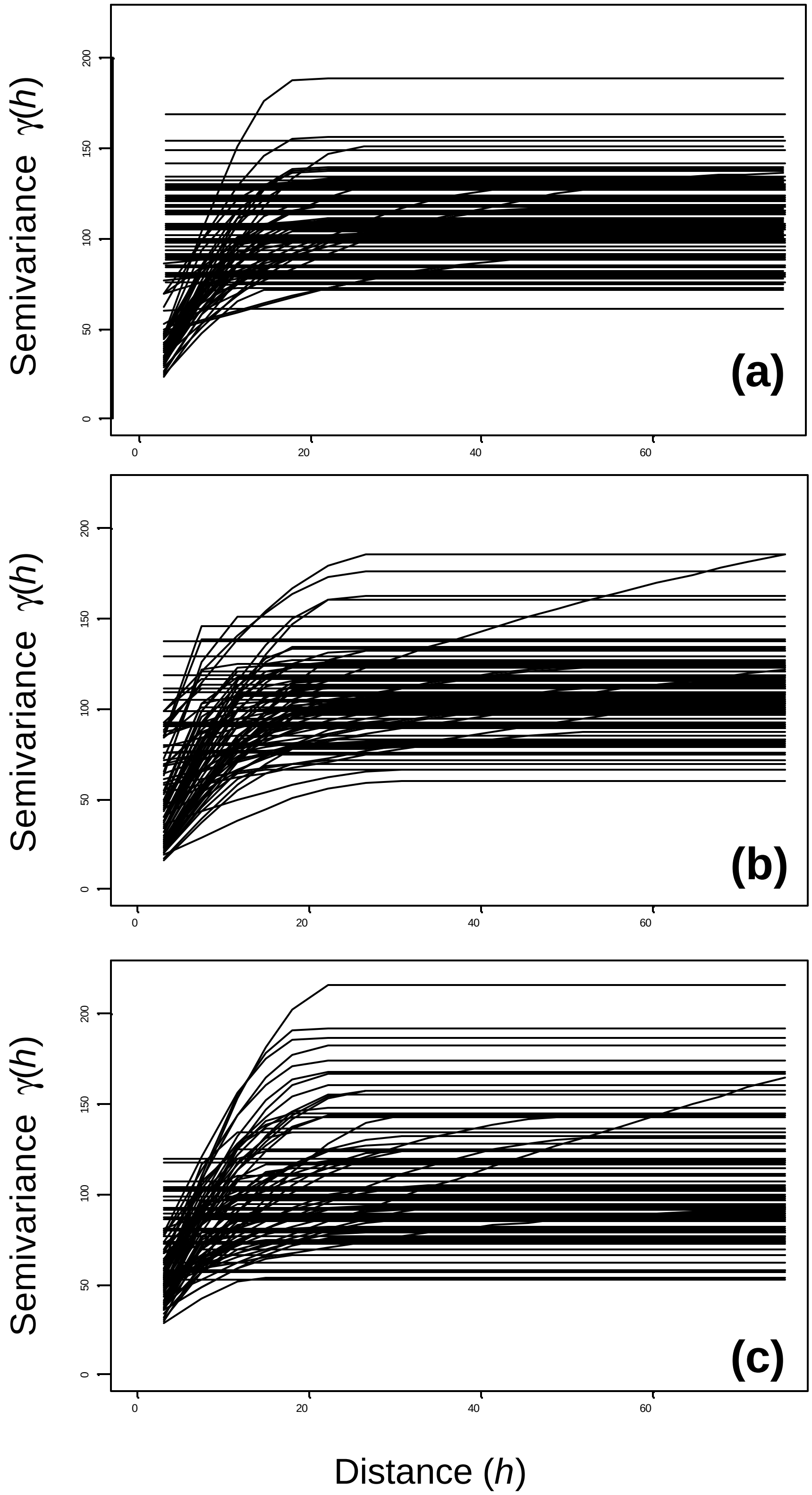

Figure 3.8. Theoretical Semivariograms created from fitting a spherical model by generalized least squares to the empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of non-stationary simulated fields with one distinct level of patchiness (Experiment 2; Scenario (2)).

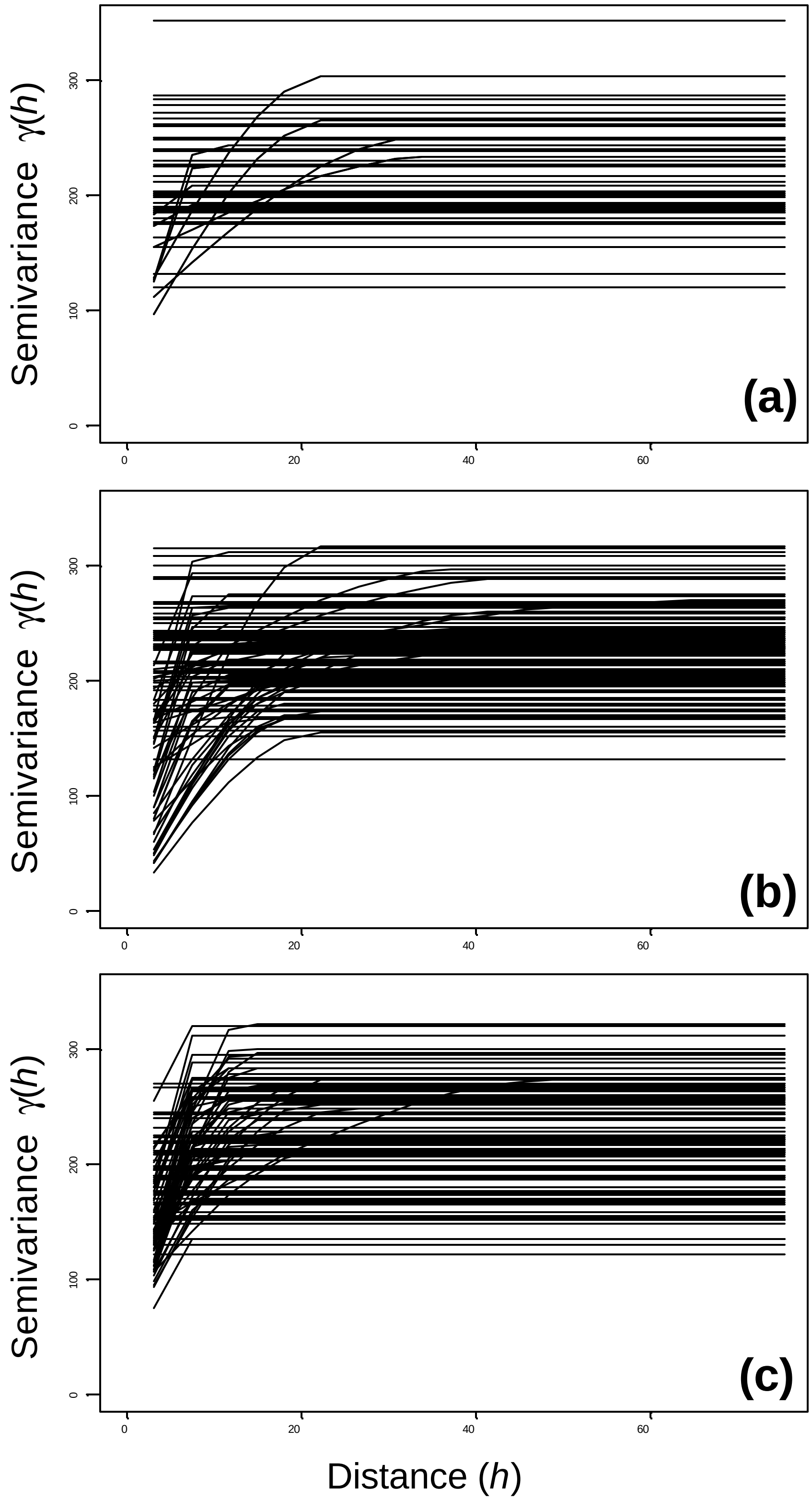

Figure 3.9. Theoretical Semivariograms created from fitting a spherical model by generalized least squares to the empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of non-stationary simulated fields with multiple distinct level of patchiness (Experiment 3; Scenario (3)).

3.5.3 Nested sampling of stationary landscape simulations

We used nested sampling and random-effects ANOVA to estimate whether significant and nonsignificant differences in variance existed between consecutive pairs of sample levels (i.e. 60m and 10m, 10m and 2.5m, 2.5m and 0.5m) for each of the stationary simulated landscapes used in Experiment 4. The results of these analyses are illustrated in Tables 3.4(a) (for $P < 0.01$) and (b) (for $P < 0.05$). Note that in comparison to our treatment of nested ANOVA's application to non-stationary landscapes (Table 3.2), we were unable, in this case, to assess the degree to which significant and nonsignificant differences in variance were correctly predicted. Thus, our analysis of the application of nested ANOVA to stationary landscapes is restricted to an assessment of the "stability" of the sampling approach; that is, the consistency with which nested ANOVA predicted significant / non-significant differences in variance between pairs of sampling resolutions. Unless stated otherwise, results refer to conservative significance testing using $P = 0.01$.

Table 3.4 shows that, for the most part, nested ANOVA consistently estimated differences in variance between pairs of sample resolutions to be nonsignificant. The exception to this trend occurred between 10m and 2.5m sample resolutions for scenario (2), where significant differences were consistently estimated. Overall, the most consistent predictions occurred between 10m and 2.5m, where 98% (scenario (1)), 97% (scenario (2)) and 99% (scenario (3)) of all differences were estimated as nonsignificant. The least-consistent estimates occurred between resolutions of 10m and 2.5m for scenario (2), where 60% of all differences were estimated as significant. The use of a more conservative test (i.e. $P = 0.01$) yielded more consistent estimates of nonsignificant differences between levels by nested ANOVA. Conversely, where a less conservative test was used (i.e. $P = 0.05$), significant differences were more consistently estimated.

(a) Levels	Expt. 4		
	Scen. (1)	Scen. (2)	Scen. (3)
60m & 10m	95	95	82
10m & 2.5m	98	97	99
2.5m & 0.5m	85	60	84

(b) Levels	Expt. 4		
	Scen. (1)	Scen. (2)	Scen. (3)
60m & 10m	87	81	68
10m & 2.5m	97	93	99
2.5m & 0.5m	80	70	70

Table 3.4. The consistency to which nested ANOVA estimates significant and nonsignificant differences in variance (patchiness) between 60m, 10m, 2.5m and 0.5m resolutions from the stationary simulated landscapes used in Experiment 4. Figures in unshaded cells correspond to the percentage (%) of estimates in which a nonsignificant difference between pairs of levels ($P > 0.01$ (a), $P > 0.05$ (b)) was calculated. Figures in shaded cells correspond to the percentage (%) of estimates in which a significant difference between pairs of levels ($P < 0.01$ (a), $P < 0.05$ (b)) was calculated.

3.5.4 Semivariograms of stationary landscape simulations

3.5.4.1 Empirical Semivariograms

Prior to fitting exponential semivariogram functions to our sampled data, we analyzed a subset of empirical semivariograms (*n*=120 for each combination of sampling approach and scenario) for their overall ability to detect patchiness at the various expected spatial extents of variation. Examples of these semivariograms are provided in Figure 3.10. In most cases, sampling with the random and systematic designs produced “noisy” empirical semivariograms. This was especially apparent for fields of short correlation length (Experiment 4, scenario (i)), and only slightly less so for fields of medium correlation length (scenario (ii)). However, both these sampling approaches were slightly more successful at identifying spatial structure for fields of longer correlation lengths (scenario (iii)). Transect sampling also produced mostly noisy semivariograms for fields with short correlation lengths, although this approach was more consistent in its ability to detect spatial structure for fields simulated with medium and long correlation lengths. Indeed, in most cases, transect sampling was able to detect the expected increases in semivariance to longer distances for fields with longer correlation lengths.

3.5.4.2 Theoretical (Exponential) model fitting

We used an exponential model fitting by generalized least squares to provide 120 estimates of spatial structure for each combination of stationary landscape scenario and sampling strategy. These estimates were then checked for “unreasonable” (extreme) parameter values, which were not included in our overall summary statistics (see section 3.5.2.2). The results of these analyses are illustrated in Tables 3.5(a) to (c), and graphically in Figures 3.11, 3.12 and 3.13.

Figures 3.11(a) to (c) show semivariograms created by fitting an exponential model to empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of fields with short correlation length. In most cases, and particularly for transect sampling, the exponential model fitted semivariograms of either pure nugget variance (horizontal lines) of very short ranges (i.e. range < 0.00001). Of the various sampling schemes, transect sampling provided the most consistent

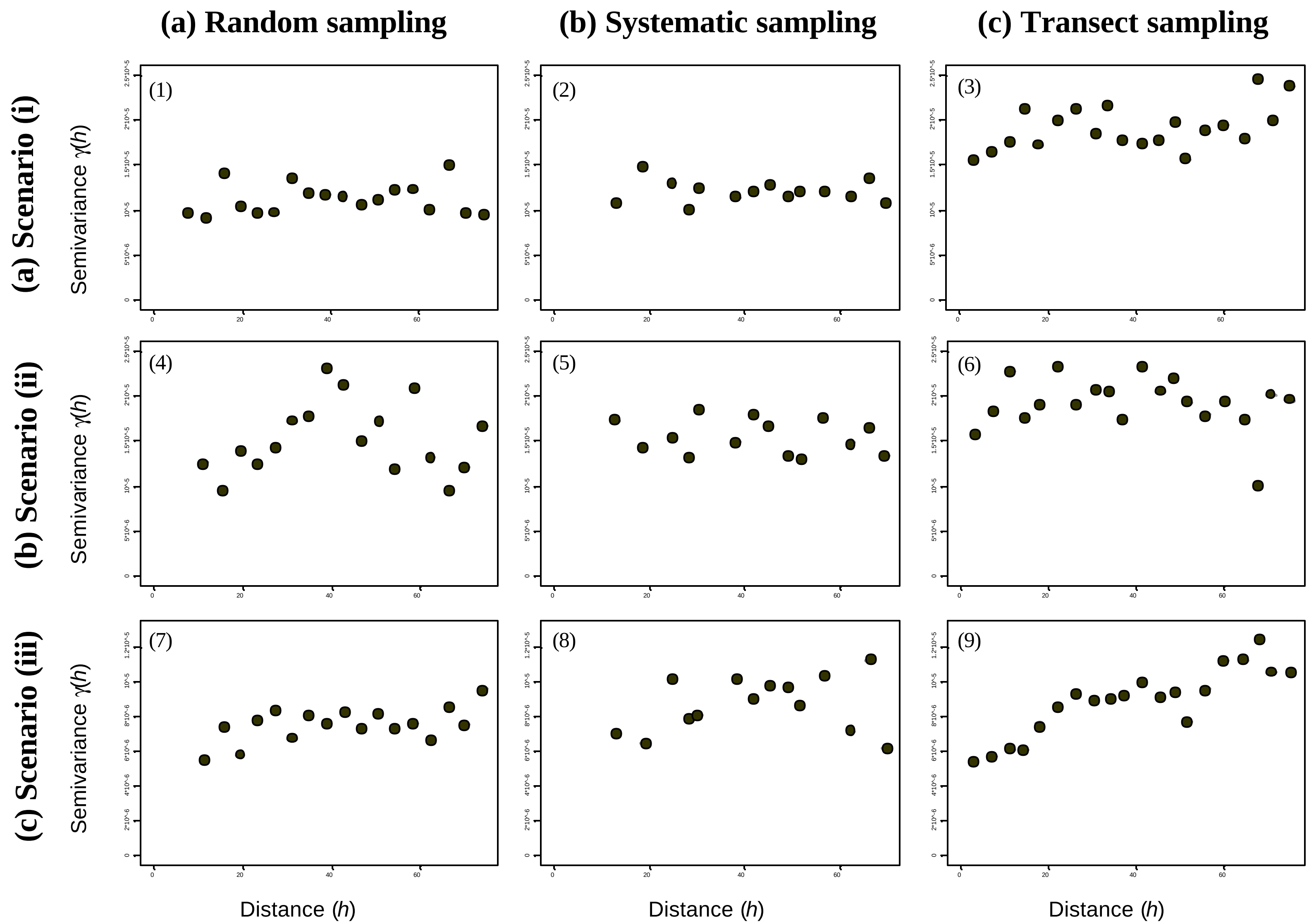

Figure 3.10. Empirical semivariograms derived from various combinations of sampling (random, systematic and transect) and stationary simulated landscapes of short (a), medium (b) and long (c) autocorrelation lengths (Experiment 4). See Figure 3.2 for detailed structural characteristics of these landscapes.

characterization of stationary fields with short correlation length (in 99% of cases, the exponential model fitted a range of less than 0.00001 (Figure 3.11(c)). Although random and systematic sampling provided, on average, similar estimates of ranges (33.5m and 31.3m, respectively), the estimates derived from systematic sampling were more variable (Table 3.5).

Figures 3.12(a) to (c) show semivariograms created by fitting an exponential model to empirical semivariograms derived from fields with medium correlation length. The resultant semivariograms showed similar patterns to those illustrated for short correlation lengths. Transect sampling consistently estimated no spatial structure, while systematic and random sampling estimate similar but variable ranges.

Figures 3.13(a) to (c) show semivariograms created by fitting an exponential model to empirical semivariograms derived from fields with long correlation length. While transect sampling began to detect the spatial autocorrelation structure at longer distances, random sampling were more variable in comparison to estimates derived for shorter length fields while, conversely, parameter estimates derived from systematic sampling were less variable (Table 3.5).

3.6 Discussion

Our results show that nested sampling and nested ANOVA provide relatively accurate and consistent estimates of spatial structure (in terms of significant or non-significant differences in patchiness between sampled levels) over non-stationary simulated landscapes whose patchiness is nested at a variety of spatial scales. Our results also indicate that when nested ANOVA is applied to data sampled from stationary landscapes, consistent (but not necessarily accurate) estimates of spatial structure result. We have also shown that under a similar sampling budget (in our case, n = 72), theoretical semivariogram fitting provides highly variable estimates of landscape structure, and that while these derived from estimates are dependent upon the choice of sampling strategy utilized, they are independent of the underlying statistical properties of the sampled landscape (i.e.non-stationarity vs. stationarity). Here, we discuss these results and, in doing so, evaluate the relative merits of nested ANOVA and semivariogram fitting for characterizing the spatial structure of grassland landscapes.

(a) Random		Nugget			Range			Sill		
		Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.
Expt. 4	(1)*	0.000016 (0.00006)	0	0.000033	33.5 (10.2)	3.7	53.0	0.000003 (0.000007)	0	0.000036
	(2)*	0.000016 (0.000006)	0	0.000030	30.2 (10.5)	5.1	39.1	0.000002 (0.000005)	0	0.000029
	(3)*	0.000002 (0.000002)	0	0.000010	30.0 (16.2)	2.8	113.4	0.000004 (0.000006)	0	0.000010

(b) Systematic		Nugget			Range			Sill		
		Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.
Expt. 4	(1)*	0.000012 (0.000008)	0	0.000033	31.3 (18.9)	6	112.0	0.000012 (0.000008)	0	0.000033
	(2)*	0.000013 (0.000008)	0	0.000028	32.3 (27.3)	4.9	184.9	0.000005 (0.000009)	0	0.000033
	(3)*	0.000003 (0.000005)	0	0.000026	28.8 (14.4)	5.4	93.8	0.000004 (0.000005)	0	0.000002

(a) Transect		Nugget			Range			Sill		
		Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.	Mean (StD)	Min.	Max.
Expt. 4	(1)*	0.000016 (0.000004)	0	0.000032	0.0 (0.1)	0	1.0	0.000001 (0.000001)	0	0.000008
	(2)*	0.000015 (0.000004)	0	0.000029	0.0 (0.1)	0	1.0	0.000015 (0.000004)	0	0.000029
	(3)*	0.000017 (0.000111)	0	0.000984	19.7 (20.2)	0	74.4	0.000003 (0.000004)	0	0.000023

Tables 3.5(a), (b) and (c). The mean, standard deviation (StD), minimum (Min.) and maximum (Max.) values of variogram parameters (Nugget, Range, Sill) resulting from fitting an exponential model by generalized least squares to sample data derived from stationary fields using different sampling designs (Random sampling (a); Systematic sampling (b); Transect sampling (c)). * indicate where extreme values were excluded from the summary statistics (see text for details).

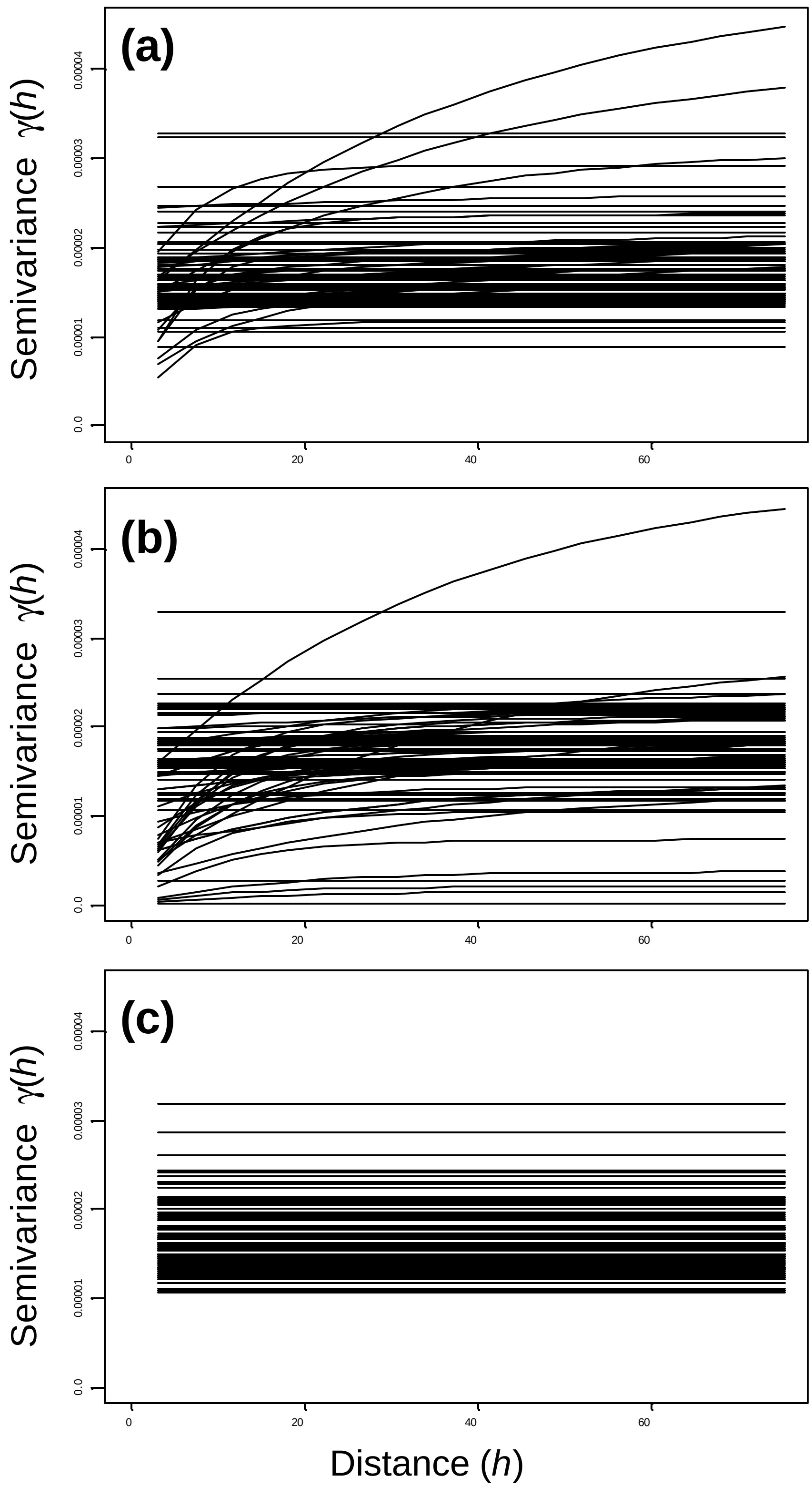

Figure 3.11. Theoretical Semivariograms created from fitting an exponential model by generalized least squares to the empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of stationary simulated fields with short spatial autocorrelation length (Experiment 4; Scenario (1)).

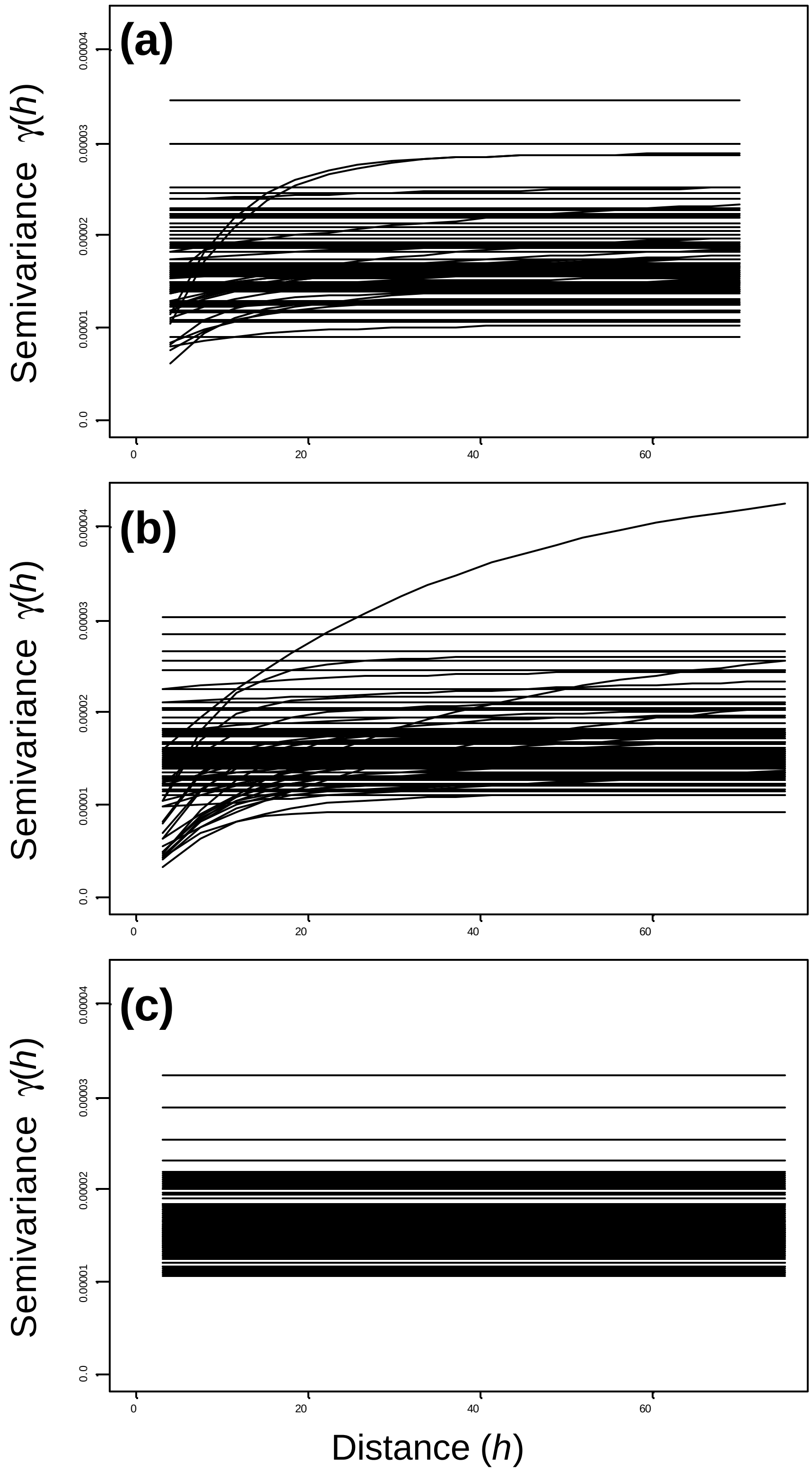

Figure 3.12. Theoretical Semivariograms created from fitting an exponential model by generalized least squares to the empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of stationary simulated fields with medium spatial autocorrelation length (Experiment 4; Scenario (2)).

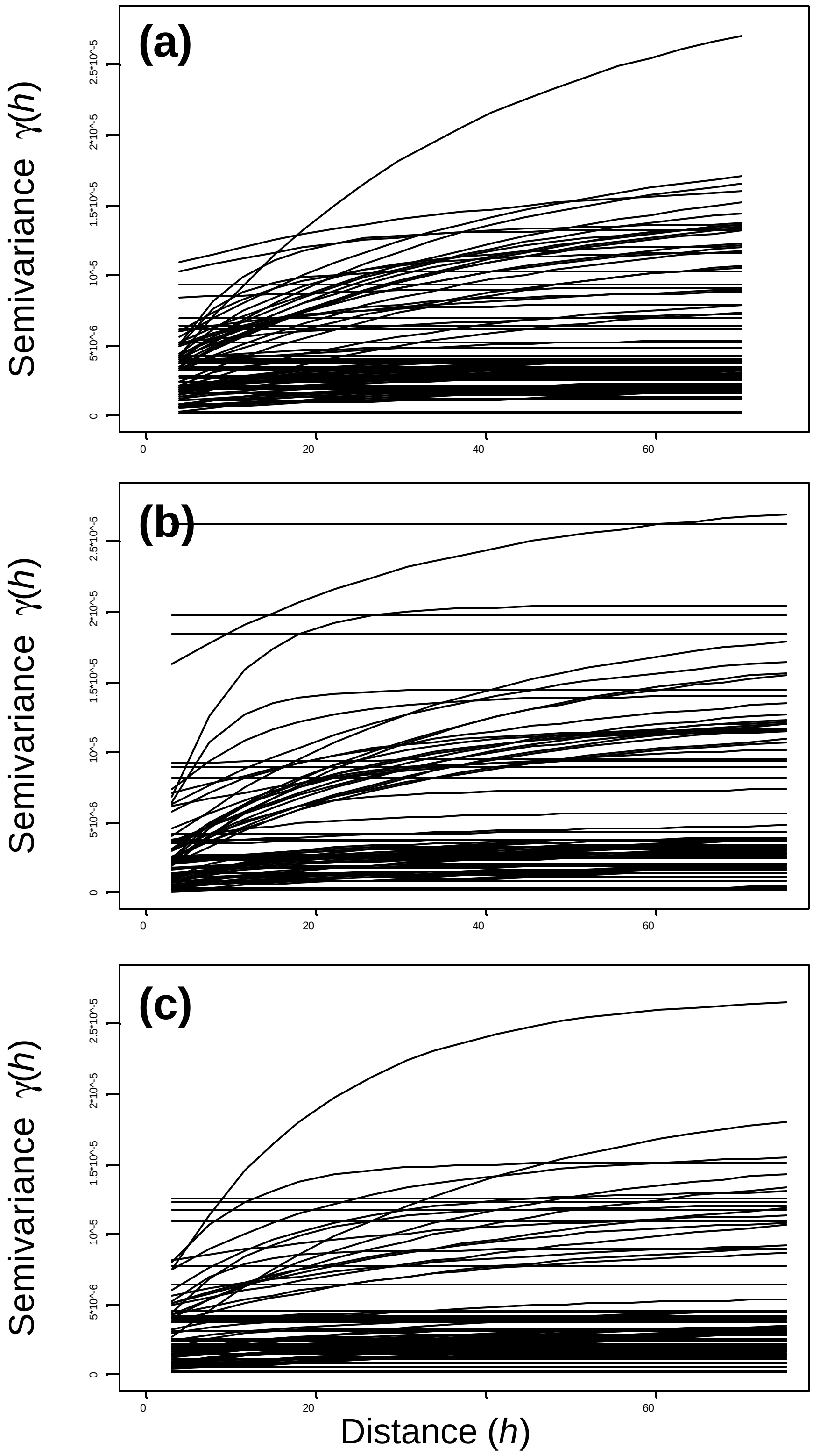

Figure 3.13. Theoretical Semivariograms created from fitting an exponential model by generalized least squares to the empirical semivariograms derived from (a) random, (b) systematic and (c) transect sampling of stationary simulated fields with long spatial autocorrelation length (Experiment 4; Scenario (3)).

The ability of nested ANOVA to estimate nonsignificant and significant differences in patchiness between pairs of sample resolutions is, as expected, dependent upon the *P*-values used in significance testing (i.e. $P = 0.05$ or $P = 0.01$). The successful estimation of nonsignificant differences between pairs of sample levels is greater where a conservative test is used (i.e. $P = 0.01$) in preference to one that is less conservative (i.e. $P = 0.05$). Conversely, where a less conservative test is used, the successful estimation of significant differences between pairs of sample levels is greater than where a conservative test is used (i.e. it becomes easier to reject the null hypothesis of no significant differences between variances). The choice of *P*-value utilized thus depends on whether the investigator is more willing to accept errors of commission (Type I error) or omission (Type II error).

The shapes of the empirical and theoretically fitted semivariograms can be explained as follows. In Experiment 1, the random, systematic and transect sampling of structureless (random) fields provide, for the most part, noisy empirical semivariograms. One may expect such results because, by definition, the absence of spatial structure implies independence between semivariance ($\gamma(h)$) and distance (h). Thus, the ability to characterize these trends becomes possible even where relatively long distances exist between sample pairs (as, for example, in the case of our systematic sampling strategy). The adequate characterization of random fields by each sampling method is further reinforced when spherical semivariogram models are fit to these empirical data. Our expectation of semivariogram shape for a completely random field is a horizontal sill, with zero range, whose height is determined by the total sample variance (i.e. the intensity of the “noise”). However, when a spherical function is fit to such a field, the fitting algorithm forces the calculation of a range, nugget and sill. Thus, if such a model is utilized under these conditions (e.g. when an investigator has no *a priori* knowledge of the structural characteristics of the landscape, and assumes some spatial dependence when, in fact, none exists), one may expect a resultant semivariogram whose range is close to zero, and whose sill corresponds to the intensity of the noise. This trend indeed occurs with many of our spherical model fits to data derived from the sampling of structureless fields, and may be interpreted as an artifact of an inappropriate choice of semivariogram model, rather than as the identification of spatial variation at very short distances. In these situations, the fitting of linear semivariogram models to empirical data may be more appropriate. It should also be noted here that the generation of a noisy semivariogram does not necessarily indicate the presence of a truly random field. Indeed, spatial structure may actually exist, but at finer resolutions (shorter distances) to those which can be resolved by the sampling scheme utilized. Thus, semivariogram interpretation and model fitting should always proceed while keeping the minimum sampled lag distance firmly in mind.

In Experiments 2, 3 and 4, while spatial structure at longer distances is characterized (to varying degrees) by semivariograms derived using each sampling approach, none are able to capture the expected spatial variation at short distances. In Experiments 2 and 3, the semi-periodic signals within semivariograms at longer distances indicate that samples have been derived from non-stationary fields. In comparison, empirical semivariograms derived from stationary fields (Experiment 4) tend not to show such signals. The inability of each approach to characterize spatial variation at short distances is largely due to the inherent limitations of each sampling design – particularly those of the systematic and random approaches – that are imposed by our restricted sampling budget. For n = 72, and a field of 128m by 128m, the systematic sampling approach provides pairs of points only at long distances (>12.5m), and thus short-range landscape variation cannot be characterized. While the random sampling approach consistently provides a number of points separated by extremely short and long distances, these are, in most cases, too few in number (i.e. n < 30, see section 3.3.4) to adequately estimate semivariance. We are again unable to explain the inability of transect sampling to resolve short range spatial variation to a consistent degree, particularly since measurements were separated by very short distances (0.5m).

3.6.1 Choosing an approach: Nested ANOVA or geostatistics?

3.6.1.1 Considerations

Of practical concern to the investigator is whether one should choose nested ANOVA or semivariograms for analyzing spatial data. While both approaches may be used to describe the data, the decision ultimately rests on a number of factors that influence whether spatial heterogeneity should be modeled as variation in nested blocks of deterministic mean structure assuming independent residuals (nested ANOVA), or modeled as autocorrelated random error with constant mean structure (variogram) (Ver Hoef et al., 1993). These factors include (a) the overall objectives of the investigator (e.g.

identification of patchiness vs. interpolation), (b) the perceived spatial properties of the landscape in question, and hence, the model assumptions that seem most reasonable (e.g. landscape stationarity vs. non-stationarity), and (c) sampling budget (e.g. limited vs. unlimited time and/or cost).

3.6.1.2 Advantages and disadvantages of approaches

For characterizing landscape structure, nested ANOVA enjoys a number of advantages over semivariogram analysis, especially under a limited sampling budget. First, in terms of sample size, nested ANOVA requires fewer samples for adequate coverage of the study area. Transect sampling allows points from even a relatively small sample size to be located close together, and has the added advantage of simplicity in planning and implementation. For these reasons, transect sampling is an attractive spatial sampling technique, as is illustrated by its frequent use in the soil sciences (see Webster, 1978; Oliver and Webster, 1986). However, the spatial variation in many landscape variables is often *anisotropic* (i.e. is spatially autocorrelated to various degrees depending on direction) because pattern is often produced by directional geophysical phenomena (Gustafson, 1998). In such cases, a grid sampling design would be more effective than transect sampling given its ability to provide insights into landscape spatial variation in more than just one direction (Oliver and Webster, 1986; Bringmark and Bringmark, 1998). Unfortunately, as we have shown, if the sampling budget remains fixed, sampling in two dimensions results in a much larger minimum distance between sample points compared to that of transect (one-dimensional) sampling. If the detection of fine-scale spatial structure in two dimensions is dependent on maintaining the sampling distance used for one-dimensional sampling, then the number of samples needed to satisfactorily estimate the semivariogram becomes equal to the square of the number of samples used in the transect. Thus, a grid sample requires considerably more points to characterize spatial structure. Webster and Oliver (1992) examined the spatial structure of various soil properties, and found that semivariograms generated using less than 50 grid points are of little value, 150 points are often satisfactory, while 225 points are usually reliable. Thus, regardless of measurement simplicity and low-cost attribute selection (e.g. spectral estimates of plant biophysical characteristics), sample-size requirements can render semivariogram analysis to be an expensive choice of model. Where circumstances do not permit such a large sample size (as in our sampling experiments), we argue that the reliable characterization of landscape structure in one direction is preferable to an unreliable characterization in two dimensions. However, in such cases, it must be kept in mind that any semivariogram-derived structure may be directionally dependent, and not necessarily characteristic of the landscape as a whole.

The second advantage that nested ANOVA enjoys over semivariograms is its ability to test for significant differences in variance between hierarchical levels (Equation 3.6) (Webster, 1979; Webster and Oliver, 1990a; Csillag and Kabos, 1996). This statistical property allows the investigator to assess whether significant differences in patchiness exist between sample resolutions (i.e. identify potential spatial scale hierarchy), and provides a useful tool for scaling studies, where the scale-dependence of variables (or the relationships between them) are of interest. In comparison, semivariances at different distances are statistically dependent upon each other (Cressie, 1985). Because of this, significance testing between variance components cannot be undertaken through semivariogram analysis.

Third, semivariograms rely on the statistical assumption of intrinsic stationarity. While more flexible techniques have been developed that avoid such stringent assumptions (see section 3.3.2), this remains a problem, particularly for ecological data, which are rarely stationary (Bailey and Gatrell, 1995).

Despite the relative advantages outlined above, nested sampling is also limited in a number of respects. First, the implementation of a nested sample design relies on the investigator's *a priori* knowledge of the spatial scales of the phenomenon of interest (Webster, 1979; Bellehumeur and Legendre, 1998). Without such knowledge, an investigator may sample at scales of interest where minimal variation occurs. It is thus important that the investigator understand, at least conceptually, the local biological and environmental factors controlling the distribution of the attribute of interest before sampling is implemented. We suggest that (a) the samples at the top level of the spatial hierarchy (i.e. coarsest sampling resolution; or the location of replicate sample grids) be located ay distances greater than the homogeneous sample area, (b) that samples at intermediate levels be located at distances that reflect a visual perception of average patch size, and (c) that samples at the lowest level (i.e. finest sampling resolution) be located at distances that allow the sampling size effect to be tested. Second, nested sampling schemes are normally designed according to a balanced sampling scheme; that is, the number of sampling units in each partition of any level of the hierarchy is the same (although it may vary from level to level) (Bellehumeur and Legendre, 1998). For each additional level in the hierarchy, sample size increases by a factor of ≥ 2, resulting in large sample sizes where many levels are utilized. In such instances, the problems with balanced nested sampling become comparable to those problems arising from attempts at estimating the semivariogram in two dimensions (see previous discussion). The use of unbalanced sampling may provide a partial solution to this problem instigated by large sample sizes. However, the penalty for the lack of balance may be further complexity in estimation and interpretation (see Webster and Oliver, 1990a).

3.6.1.3 Similarities between approaches

Although nested ANOVA and semivariograms partition the total sample variance differently, they are similar in a few respects. First, both approaches are based on calculating the sum of squared differences. Nested ANOVA computes these differences for pairs of values sorted by nested partitions, while the empirical semivariogram computes these differences for pairs sorted by distance (Csillag and Kabos, 1996) (see Equations 3.5 and 3.7). Second, Miesch (1975) showed that the accumulation of nested components, starting with the smallest spacing, were equivalent to the semivariances derived over the same range of distances. However, Webster (1990b) notes that, in practice, these are only rough estimates of the true semi-variances because each is based on few degrees of freedom. Furthermore, Bellehumeur and Legendre (1998) caution that Miesch's assertion is only true for the sample design he specified, and would not hold true for a transect of contiguous quadrats.

A variety of authors suggest that the choice of nested ANOVA or semivariogram need not be mutually exclusive. Webster (1990a) outlines a two-step approach that incorporates both techniques. First, nested sampling is used in the preliminary stages of an investigation to create a crude semivariogram spanning several orders of magnitude of distance. Second, sampling to estimate the variogram can then be concentrated in the range of distance that encompasses most of the spatial variation. In comparison, Ver Hoef et al. (1993) outline a theoretical approach where various statistical properties of nested ANOVA and geostatistics are merged to form one “hybrid” technique. This approach takes the form of spatially nested effects models where the random errors are independent among blocks, but correlated within blocks. Using this approach, the spatially nested effects could be removed within nested ANOVA and the residuals analyzed by a variogram. Although such approaches need further study, preliminary analyses (Davidson, Unpublished data) suggest that where sample size is small and limited, semivariograms created in this way suffer from many of the problems discussed earlier.

3.6.2 Limitations of study

While these results potentially have large implications for those wishing to characterize grassland structure through ground-based sampling, there are several unresolved issues. First, landscape simulation through HQ-simulation generates artificial landscapes that are somewhat “blocky” in nature (see Figure 3.3). Furthermore, for our simulated landscapes of 128m by 128m at 0.5m resolution, this approach only allows for the explicit control of patchiness at eight resolutions (0.5m, 1m, 2m, 4m…64m; see section 3.3.1 and Figure 3.2). Thus, while we simulated landscapes of varying complexity and structure, our realizations cannot be considered as “true” surrogates of upland grassland systems. Although we feel this limitation does not detract significantly from our overall results or their implications for field sampling, these issues need to be addressed. Second, our results, and thus conclusions, are based solely on the sampling of spatial structure at 60m, 10m, 2.5m and 0.5m resolutions. The ability of nested ANOVA to characterize landscape structure with consistent accuracy across a variety of resolutions needs to be further investigated. However, through subsequent analyses, we have since addressed this issue in part. These results (not shown here (Davidson, Unpublished data)) indicate that variants of the nested sampling scheme used in this study (e.g. sample points located at distances of 60m, 8m, 1m, 05m; 60m, 16m, 4m, 2m) produce similarly accurate and consistent estimates of significant and nonsignificant differences in patchiness between sampled resolutions for both non-stationary and stationary landscape simulations. Third, we do not attempt to compare nested ANOVA and geostatistics under conditions where sample budget is unlimited. Under such circumstances, semivariogram analysis may be a more appropriate choice for characterizing landscape structure than nested ANOVA. This issue warrants further investigation.

3.7 Conclusions

The work presented here compares the ability of two approaches – nested ANOVA and geostatistics – for characterizing the spatial structure of various non-stationary and stationary grassland landscapes. Our results support our original expectation that the sampling methodology used to characterize spatio-temporal heterogeneity (and its associated statistical assumptions) strongly influence perceived landscape structure. We have shown that (a) nested ANOVA provides a more reliable characterization of landscape structure than that provided by semivariogram analysis under conditions of a limited sampling budget, (b) the ability of nested ANOVA to accurately characterize landscape structure is largely independent of landscape complexity, and (c) semivariogram analysis provides highly variable estimates of spatial structure, even for landscapes where the statistical assumptions of stationarity are maintained. Furthermore, the independence of variance components under nested ANOVA allows for significance testing between levels, and thus, the identification of potential spatial scale hierarchy. Significance testing between variance components is not possible with semivariogram analysis. However, while these results potentially have large implications for those wishing to characterize grassland structure through ground-based sampling, there are several unresolved issues. These include (a) our limited ability to simulate “real” grassland landscapes, and (b) our testing of nested ANOVA using few variants of our nested sampling scheme, and (c) the need for the comparison of sampling approaches under conditions where sample budget is unlimited. These issues need to be further investigated.