Towards a Coherent Framework for the Multi-scale Analysis of Spatial Observational Data: Linking Concepts, Statistical Tools and Ecological Understanding

Guillaume Larocque

Department of Natural Resource Sciences,

Macdonald Campus of McGill University, Montréal.

December 2008

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of DOCTOR OF PHILOSOPHY.

Contents

Abstract

Recent technological advances facilitating the acquisition of spatial observational data and an increasing awareness of issues of spatial pattern and scale have fostered the development and use of statistical methods for multi-scale analysis. These methods can be interesting tools to improve our understanding of natural systems, but their use must be guided by a good comprehension of the statistics and their assumptions. This thesis is an effort to develop a coherent framework for multi-scale analysis and to identify theoretical, statistical and practical issues and solutions. After defining terminology and concepts, several methods are compared using a common dataset in Chapter 2. The geostatistical method of regionalized multivariate analysis is identified as possessing several advantages, but shortcomings are identified, discussed and addressed in two manuscripts. In the first one (Chapter 3), a mathematical formalism is presented to characterize the spatial uncertainty of cokriged regionalized components and an approach is proposed for the conditional Gaussian co-simulation of regionalized components. In the second manuscript (Chapter 4), the theory underlying coregionalization analysis is discussed and its robustness and limits are assessed through a theoretical and mathematical framework. The assumptions underlying the method and the high levels of uncertainty associated with its use highlight problems with the interpretation of results, and issues with the application of probabilistic models in a spatial context (Chapter 5). Coregionalization analysis with a drift (CRAD), presented in detail in two co-authored publications, is proposed as a sensible alternative for multi-scale analysis. In Chapter 6, CRAD is used in an application to discuss the role of scale in site-specific agricultural management and study the relationships between spatial structure and temporal heterogeneity in soil variables. In Chapter 7, the use of CRAD is extended to the multi-scale causal modelling of relationships between physical factors, tree species distribution and soil variables in a forest ecosystem. These applications show the great potential of multi-scale analysis to facilitate ecological understanding, but highlight the need for further development of ecological theories to generate precise expectations about process-pattern linkages within and across scales.

Résumé

Des avancées technologiques récentes facilitant l'acquisition de données observationnelles spatiales, et la conscientisation grandissante des chercheurs aux problèmes d'échelles, ont favorisé le développement et l'utilisation de méthodes statistiques d'analyse multi-échelles. Ces méthodes peuvent être des outils intéressants pour améliorer notre compréhension des écosystèmes, mais leur usage nécessite une bonne connaissance des théories statistiques et des hypothèses qui leur sont sous-jacentes. Cette thèse a pour but de contribuer au développement d'un cadre conceptuel cohérent pour l'analyse multi-échelles en identifiant des problématiques et des solutions théoriques, statistiques et pratiques. La terminologie et les concepts appropriés sont d'abord définis. Ensuite, dans le chapitre 2, plusieurs méthodes sont comparées en utilisant un jeu de données commun. La méthode géostatistique d'analyse multivariable régionalisée semble offrir plusieurs avantages, mais certains problèmes sont identifiés, discutés et traités dans deux articles publiés. Le premier (chapitre 3) présente un formalisme mathématique servant à caractériser l'incertitude spatiale des composantes régionalisées cokrigées et propose une approche pour la co-simulation conditionnelle Gaussienne de composantes régionalisées. Dans le deuxième article (chapitre 4), l'analyse de corégionalisation est discutée et la robustesse ainsi que les limites de cette méthode sont évaluées selon une analyse théorique et mathématique. Les postulats de cette méthode et le haut niveau d'incertitude lié à son utilisation mettent en évidence des problèmes associés à l'interprétation des résultats et à l'application des modèles probabilistes dans un contexte spatial (chapitre 5). Pour remédier à ces problèmes, l'analyse de corégionalisation avec une dérive (CRAD), détaillée dans deux publications, est proposée comme méthode alternative pour l'analyse multi-échelles. Au chapitre 6, la méthode CRAD est utilisée pour discuter des questions d'échelle en agriculture de précision et des liens entre l'hétérogénéité spatiale et temporelle dans les variables du sol. Au chapitre 7, l'utilisation de CRAD est étendue à la modélisation causale multi-échelles des relations entre les facteurs physiques, la distribution des espèces d'arbres et les variables du sol dans un écosystème forestier. Ces applications démontrent le grand potentiel de l'analyse multi-échelles pour faciliter la compréhension des écosystèmes. Par contre,

les résultats révèlent également la nécessité de poursuivre le développement de théories écologiques pour permettre l'identification des processus expliquant les patrons spatiaux à différentes échelles.

Contributions of the authors

Chapters 1, 2, 5, 6, 7 and 8, were written by me and are primarily the result of my personal research efforts. Collaborators (Drs. Pierre Dutilleul, James W. Fyles, Bernard Pelletier) provided some guidance and advice, helped to the development of the ideas, and acted as reviewers. Chapters 3 and 4 were co-authored by me and Drs. Dutilleul, Pelletier and Fyles. I acted as the main author, proposed the main ideas and established most of the problems and solutions. These two chapters were extensively reviewed by the co-authors. Dr. Dutilleul collaborated closely to the mathematical development, to the consistency of the notations and the veracity of the statistical discussion throught the thesis, but particularly in Chapters 3 and 4. I undertook the majority of the Matlab programming used in Chapters 2, 3 and 4. The dataset used in Chapter 7 was collected by me and the laboratory analysis was undertaken by me, or under my supervision. The dataset used in Chapter 6 was collected by me and Dr. Caroline Begg, and by other helpers under our supervision. All the figures presented in the chapters of this thesis were created by me.

Two other publications were partially derived from this research (Pelletier et al., 2009a,b) and are the result of a collaborative effort between the co-authors, with Drs. Pelletier and Dutilleul being the main authors. I provided some advice and ideas, I acted as a reviewer and I collaborated to the Matlab programming of the CRAD method used in these publications and in Chapters 6 and 7. The soil dataset used in these publications was collected and analyzed by me and field assistants under my supervision, while the species inventory was undertaken under the supervision of Dr. Marcia Waterway. I collaborated to the localization and demarcation of the sampling design and analyzed the LIDAR dataset and obtained the topographical variables used in the analysis.

Chapter 1

Introduction

Spatial structure is an essential functional component of all ecosystems. For hundreds of years, scientists have divided the environment into similar climatic zones, geomorphic setting, forest type or stand characteristics. Structuring of the environment allows us to make predictions as to the types of vegetation, the volume of precipitation, the frequency of fires or the amount of nitrogen in the soil. Without it, not only would ecosystems be dysfunctional, but ecological science would be impossible, or at least irrelevant. In the realms of classical statistics, however, spatial structuring of the environment has been considered as a problem, preventing independence of field observations. For several decades, ecological research has been carried without regard to spatial structure (Legendre, 1993).

Recent concerns over phenomena such as global warming, loss of biodiversity, food production or timber supply have enhanced the need to estimate ecological variables at the landscape or planetary scale. This is made very difficult by the fact that typical ecological field experiments are carried at a variety of scales determined by the needs and tradition of each discipline. In order to integrate scales of research and scale-up from field studies to global ecosystem modelling, there is a growing need to understand the spatial patterns of ecosystem properties, to see what processes govern those patterns and to assess the robustness of our theories across scales (Levin, 1992).

Fortunately, recent technologies such as global positioning systems, remote sensing and automated sampling devices have made available spatially intensive and extensive datasets. The parallel development of geographic information systems, spatial statistics and the exponential increase in computational power have made the analysis of those datasets much easier. These new capabilities allow scientists to ask questions about the spatial structure or scaling behaviour of natural systems that were completely out of reach before the emergence of such tools. However, several ecological theories have to be revised or re-invented to account for space and scale (Wiens, 1999) and considerable

effort is required by the scientific community to understand the new tools and measure their usefulness for answering specific questions about ecosystem behaviour. Those urgent needs have contributed to the recent emergence of the disciplines of ecosystem and landscape ecology.

One of the key challenges of these new disciplines is the need to deal with the overwhelming complexity of ecosystems (Wiens, 1999; Jorgensen and Muller, 2000). The presence of multiple causal factors, scale-dependencies and spatial dependence makes the study of single variables and relationships extremely difficult. Several authors have criticized the traditional 'reductionist' approach to ecosystem studies in which single factors are isolated and studied independently of the whole in favour of a more pluralistic and holistic approach, which embraces complexity rather than imposing simplicity upon ecosystems (Allen and Starr, 1982; Levins and Lewontin, 1985; O'Neill, 1986; Jorgensen, 1997; Wiens, 1999). On the other hand, such ecosystem studies are often based on observational data and are considered by many to be too descriptive, phenomenological and lacking general frameworks. Contributions from various fields of ecology have thus called for logically designed, rigorous frameworks for understanding ecosystem processes from observational data (Vepsäläinen and Spence, 2000).

Forested and agronomic systems are in the midst of this new scientific, environmental and social context. In the fields of vegetation science, soil science and agronomy, there has been a gradual shift in the scales of study, the hypothesis formulated and analytical methods used. Each of these fields has adapted or developed methods to address the issues of pattern and scale. Very few studies, however, have studied empirically the interactions between vegetation, environment and soil through space and scale. Such studies are essential before we can start answering questions pertaining to pressing environmental issues in terrestrial ecosystems.

In the section that follows, I will briefly review the published literature on the concepts of space and scale in ecology, on the causes of spatial pattern in soil-plant systems, on the methods of spatial analysis, on the conceptual approaches to address the concept of scale, and on the problem of understanding natural systems in the presence of complexity. In the next section, and in the rest of this doctoral thesis, the focus will be switched to the multi-scale analysis of spatial observational data. The particular issues related to multi-scale analysis will be discussed and the objectives of this doctoral research project will be presented.

1.1 Space and scale in ecology

The concepts of space and scale have extremely varied definitions. Spatial structure, spatial pattern or spatial autocorrelation are often used interchangeably. Griffith (1992) outlines that there is no universal and simple definition of spatial autocorrelation and he describes the following nine meanings found in the literature : 1) Self-correlation attributable to the geographical ordering of data, 2) a descriptor of the nature and degree of certain types of map pattern, 3) an index of the information content latent in georeferenced data, 4) a diagnostic tool for spatial model misspecification, 5) a surrogate for unobserved geographic variables, 6) a nuisance in applying conventional statistical methodology to spatial data series, 7) an indicator of the appropriateness of, and possibly an artifact of, areal unit demarcation, 8) a spatial process mechanism, and 9) a spatial spillover effect.

In a more technical statistical definition, the term serves to designate the property of random variables that take values, at pairs of sites a given distance apart, that are more similar (positive) or less similar (negative) than expected for randomly associated pairs of observations (Legendre, 1993). Legendre and Legendre (1998) and several others differentiate between autocorrelation and spatial dependence. Autocorrelation arises from the direct influence of an organism on its surrounding environment. For instance, the presence of a fern at a given location may be due to the fact that the parent from which it comes is located a short distance from it. Spatial dependence, or false autocorrelation, describes the spatial structure in one variable that is due to its interaction with other spatially distributed variables. An example may involve the presence of patches of certain species of ferns on zones of different moistures.

The importance of spatial structure in the functioning of ecological systems has been recognized for several decades. Nevertheless, from the 1950s to the 1970s, the dominant paradigm was that ecological dynamics took place in spatially homogeneous local habitats (Wiens, 1999). The extensive use of controlled experiments and quadrats of specific sizes in ecological research and standard homogeneity assumptions (Wiens, 1989) have limited the possibilities of explicitly accounting for spatial pattern. The incorporation of spatial heterogeneity into ecological concepts and studies only started in the late 1970s and became widespread by the mid 1980s. The drivers behind this shift in attention were numerous, from the increasing availability of spatially explicit datasets to the lack of success of traditional approaches in spatially heterogeneous environments.

Scale has also acquired several common technical meanings (reviewed in Schneider, 2001). In its simplest form, however, the concept of scale is extremely simple: it refers

to physical dimensions of observed entities and phenomenon. It should be possible to assign dimension and units (cm, kg, km^2, etc.) to measurements of scale (O'Neill and King, 1998). As mentioned in numerous other places, scale should not be used to describe levels which are concepts without precise physical dimensions, such as a stand, an ecosystem or a landscape (see Allen and Hoekstra, 1992 for a thorough discussion).

Pickett et al. (1994) make an interesting distinction between what they call 'thing' ecology and 'stuff' ecology. For them, ecological studies of discrete entities are closer to 'thing' ecology, while the analysis of continuous variables is closer to 'stuff' ecology. Things are not present everywhere on the landscape, while Stuff refers to pools or fluxes that have values everywhere on the landscape (Fig. 1.1). Things are tangible entities such as animals, plants, agricultural fields and lakes. Concentrations of soil nutrients, air temperature, water depth or terrain attributes represent 'stuff'. For 'things', the application of the concept of scale is rather straightforward. They have spatially discrete boundaries, and their scale may refer to the dimensions (height, width, weight, etc.) of those boundaries. Once we define a convention to measure the dimensions of 'things', each 'thing' can only be attributed a given scale. For example, the scale of a given tree may be defined by its diameter at breast height or by its height. Similarly, an agricultural field is a 'thing' to which we can attribute a dimension such as a width or area. However, the application of the concept of scale to 'stuff' is more problematic because the lack of defined boundaries precludes a measurement and thereby an attribution of scale. It requires a certain discretization of the variable's spatial structure into *spatial features* or 'things' (patches, bumps, ridges, waves, etc) that can be measured. The separation between 'things' and 'stuff' is also a matter of scale. For example, plant abundance can be treated as a continuous variable at larger scales although plants are 'things' at small scales. Similarly, soil variables can be considered 'stuff' at small scales and be classified into discrete zones ('things) in soil survey maps.

In the terminology used throughout this thesis, I separate the scale concepts into broad categories. The *scales of observation*, also referred to as the measurement or sampling scales, are the scales at which the data is collected and is composed of the grain size, sampling distance and the extent. The grain size represents the size of the sampling units (quadrats, cores, etc), while the sampling distance is the typical distance between sampling units. Soil scientists often use a very small grain size and a much larger sampling distance while vegetation scientists require a larger grain size. *Scales of variation* is a general term to describe sizes of spatial features, independent of the scale of observation. Scales of analysis are completely determined by the framework of the analytical methods. They are the sizes and shapes of spatial features as depicted by

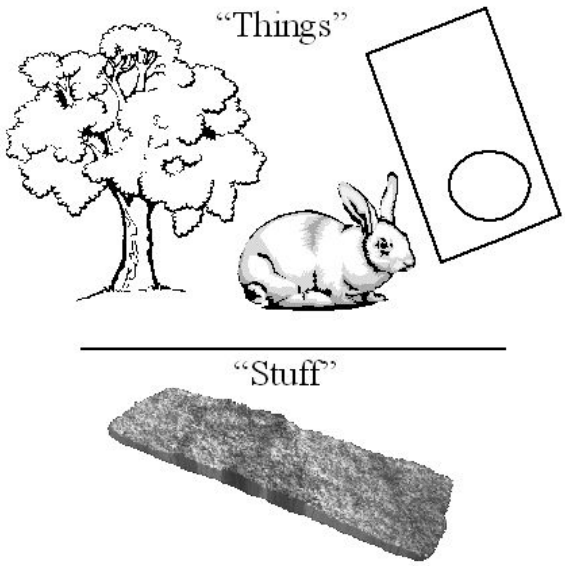

Figure 1.1: Things are entities that are not present everywhere on the landscape. Stuff refers to pools or fluxes that have values everywhere on the landscape.

each statistical method. Obviously, the scales of observation impose constraints on the scales of analysis. The lower bounds are set by the grain size and sampling distance, while the upper bounds are set in part by the sampling extent.

We can define intrinsic, or *characteristic scales* as properties of the 'true' population, or the scales of variation describing the patterns in the natural system (Wu and Li, 2006), independent of the observer. Whether such intrinsic scales are actually accessible in practice, or even 'exist', can be questioned and depends on the philosophical position one takes.

Before the 1980s, very little attention was given to the concept of scale in ecology (Schneider, 2001)but since then, its importance has been gradually endorsed by ecologists (Wiens, 1989; Levin, 1992; Mac Nally and Quinn, 1998; Thompson et al., 2001). The problem of scale has two components (modified from Schneider (2001)): first, most ecological information is collected at scales smaller than that of modern day environmental issues. Although some technologies, such as remote sensing, are available for collecting information over large areas, they measure few ecosystem parameters, such as reflectance. Because of the traditions of each discipline and the sizes of living organisms, most ecological research has been conducted at fairly small scales (Wiens, 1989). First attempts at transposing such studies to broader scales to generate estimates of pools and fluxes at the global scale have shown that this is a very daunting task. The second component of the problem of scale is that the processes operating at a given scale may not operate at smaller or larger scale. Such scale-specific processes leading

to scale-dependent patterns have been found in aquatic systems (Levin, 1992), in soils (Goovaerts and Webster, 1994; Dobermann et al., 1997; Crawley and Harral, 2001) and in vegetation assemblages (Crawley and Harral, 2001).

For 'things', the concept of scale-dependence is easily understood. Entities have definable spatial boundaries imposing a mutual exclusion with other entities of the same type at scales equal to or slightly greater than the size of those boundaries. For example, when sampling trees on quadrats of 10 m^2 in a mature coastal forest where trees are large and widely spaced, we do not expect to measure more than a few trees, if any. This can be perceived as a carrying capacity of the environment, not in terms of resources but simply in terms of space. Trees species that tend to occur in the same stand may still be negatively correlated at small scales because of this, and may exhibit a scale-dependent relationship. The effect that biological entities (organisms) have on their surroundings and on other organisms may generate also scale-dependencies. A classic example is the interactions between predators and prey. At small scales, a negative association results from the fact that both are rarely found in close proximity while at larger scales both species inhabit the same area and are therefore positively associated (Rose and Leggett, 1990). This small scale interaction does not need to be negative however, some processes positively influence the presence of other organisms, at least of the same species, in their surroundings (e.g. dispersal, clonal growth, etc.). As long as we remember that the identification of scale requires the discretization of the pattern into spatial features, the concept of scale-dependence can be seen as applying to 'stuff' as it applies to 'things'. However, there is no obligate negative small scale relationship between variables representing 'stuff' and spatial features defined from 'stuff' can be nested within larger spatial features. The issue of scale-dependence is thus much more complex for 'stuff' than it is for things.

In one of the landmark articles in the development of the understanding of scale in ecology, Simon Levin (1992) argued that the problem of pattern and scale is the central unifying problem in ecology. There is, according to him, a growing need to study and understand how pattern and variability change with the scale of description and to develop laws for simplification, aggregation, and scaling (Levin, 1992). Thompson et al. (2001) refer to this as an ecological topology, "how our scales of information [...] affect our understanding of ecological processes."

1.2 Causes of spatial patterns in soil-plant systems

Spatial structure is an essential functional component of all ecosystems. Ultimately, every pattern can be considered 'spatial' depending on the scale of observation, and every ecological process could be considered as either creating or preventing the creation of spatial pattern. At the scales typical of ecological research, spatial patterns arise either from the morphological characteristics of organisms, from biotic interactions between organisms, from environmental controls or from disturbance.

Obviously, not all organisms have the same shape and size. The morphology of a given individual or species induces one or several scales of pattern (Kershaw, 1963). In plants, morphological characteristics of a species may determine the density of individuals and the pattern of influence of individual plants on the surrounding environment. In forests, the effect of the morphological characteristics of trees is particularly pronounced due to the high variability in the structure of different tree species, and their long range of influence. Examples of ways by which trees can affect their surrounding environments include stemflow (Gersper and Holowaychuk, 1971), litter fall properties (Ferrari and Sugita, 1996), localized uptake of nutrients and water or by generating microclimates through shading and moisture retention (Binkley and Giardina, 1998). Individual trees have been documented to affect spatial patterns in understory vegetation (Crozier and Boerner, 1984), soil chemical properties (Zinke, 1962; Lodhi, 1977; Boettcher and Kalisz, 1990), soil micro-organisms (Nantel and Neumann, 1992) and earthworm abundance (Boettcher and Kalisz, 1990).

Biotic interactions between plants and species groups may help to generate spatially structured, or patchy, plant communities. Clonal growth and other modes of vegetative reproduction may produce patches of varying sizes in which other species tend to be excluded (Rackham, 1991). In trees, species such as Fagus grandifolia, Ulmus sp. and Populus sp. have been shown to form clones by means of suckering. Tilman (1994) suggests that dispersal and competitive interactions may also allow many species to survive together on physically homogeneous sites, even if one of the species has a competitive advantage over other on resource utilization. With sample and simulated data, Frelich et al. (1993) showed that spatial patterns in a sugar maple and hemlock stand could be created on sites with uniform soil and topography. They hypothesized that long-term interactions between the two species, perhaps related to the creation of proper seedbed conditions by the parent trees for understory germination and to the invasion pattern of hemlock, could be sufficient to explain the patchy structure of both species. Other mechanisms, such as allelopathy, may create or prevent the creation of spatial patterns.

Long-term geological and geomorphological history has created a cross-scale mosaic of different rock types, deposits and topographical features on the earth's surface. Although we often think of these processes as mostly operating at large scales and varying more between ecosystems than within ecosystems, a number of geomorphological processes may induce micro to meso-scale variation. At these scales (<10 km), patterns are normally constituted of landforms derived from exogenic processes, such as moraine ridges, fluvial sand bars, small glacier valleys or aeolian sand deposits. According to Wilding et al. (1994), we should expect the spatial variability in soils to increase with the nature of the parent material in the following order: loess<till<fluvial deposits<pyroclastic and tectonic materials<drastically disturbed materials. A number of studies have shown that microtopography can have an effect on the depth of the forest floor, the distribution of tree species, the forest floor temperature and moisture and bacterial populations (Dwyer and Merriam, 1982; Beatty, 1984; Ruel et al., 1988). On much larger spatial scales in humic soils of France, Arrouays et al. (1998) found good correlations between the topographical features extracted from a digital elevation model and soil organic carbon. The effect of geomorphological heterogeneity has also been related to vascular plant species richness at both the patch scale (Burnett et al., 1998) and the landscape scale (Nichols et al., 1998).

Considerable research has been devoted to the effect of disturbance on the spatial structure of terrestrial ecosystems. At the local scale, patterns in soils may be generated due to tree falls (Beatty, 1984), whereas on broader scales the catastrophic blowdown created by a tornado (Peterson and Pickett, 1995), the effect of historical landuse (Motzkin et al., 1999), insects, diseases (Chokkalingam and White, 2001) or fires (Rebertus et al., 1989) may generate a wide array of spatial patterns in both soil and vegetation. Disturbances may induce patterns that are extremely gradual but they are more often patchy or discrete.

1.3 Methods of spatial analysis

Spatial structure has different meanings depending on the technique with which it is described or studied. The approaches developed to characterize and study spatial structure can broadly be classified in three categories: point pattern, surface pattern and landscape metrics. The first two serve to study spatial structure from sampled data whereas the latter helps to describe the spatial arrangement of mapped data (modified from classifications of Legendre, 1993 and Fortin, 1999).

Point pattern analysis aims to describe the distribution of individual objects, such as

trees, in space. Surface pattern analysis comprises a whole array of methods, mostly centered on the description of spatial autocorrelation and mapping. Correlograms and variograms are direct tools for characterizing spatial autocorrelation. Other surface pattern methods include spectral analysis, fractal analysis, trend surface analysis and neighbourhood matrices.

The other approach to studying spatial structure is to transform mapped/population data into a mosaic of patches. Landscape metrics then serve to describe the spatial arrangement and geometrical properties of those patches. These methods are often carried within Geographic Information Systems.

French and South African schools of mining and geology preceded ecologists in their attention to spatial patterns. Matheron (1963a, 1965) based on earlier work from Krige (1951) and others (Cressie, 1990), laid out the foundations of geostatistics. The theory of regionalized variables, at the core of geostatistics, states that any spatially distributed (regionalized) variable possesses two components: a local, random aspect and a structured aspect that requires a functional representation (Journel and Huijbregts, 1978). Based on this theory, several geostatistical methods have been developed for describing the spatial continuity of variables (variography), optimal spatial estimation (kriging) and uncertainty assessment (simulation). Although the ability of geostatistics to describe spatial patterns is widely recognized, most of the attention of geostatistical research has been focused on methods of estimation, simulation and uncertainty analysis. The use of geostatistics for exploratory data analysis is fairly limited. Despite now widespread use of geostatistics in several fields, the fundaments of geostatistical theory have been developed with mining and geological applications in mind.

Spatial structure has always been a central theme for soil scientists. Approaches to studying soil spatial variability can be divided in two principal categories: soil classification and geostatistics, the former of which largely predated the latter in history (this discussion follows Heuvelink and Webster, 2001). In soil classification the landscape is seen as a mosaic of patches of different soil types within which the soil is considered relatively uniform. These patches are separated by sharp boundaries where the observed differences are greater than elsewhere. The variability of each soil type is usually described in vague terms and qualitatively. This methodology dominated soil science research for several decades. It was not until the late 1960's that soil scientists began studying soil variation in a systematic way. First studies looked at the accuracy of maps and saw spatial variation as a nuisance that reduced map quality (Burrough et al., 1994). Accumulation of evidence showed the continuous nature of numerous soil properties and the inadequacy of the discrete model of soil variation for describing

continuous landscapes.

The first uses of kriging and geostatistics in soil science date from the early 1980's (e.g. Burgess and Webster (1980)). Since then, geostatistics have become a widespread tool for observing the spatial variability of soil variables and generating continuous soil maps (reviews can be found in Burrough et al., 1994; Goovaerts, 1999; McBratney et al., 2000; Heuvelink and Webster, 2001). The term 'pedometrics' was coined to categorize quantitative methods of analysis in soil science (Heuvelink and Webster, 2001). The development of geostatistical applications in soils has contributed to a gradual change in the way soil scientists perceive soils. In many applications, spatial variability is no longer seen as a nuisance but as a useful tool for understanding soils and soil forming processes.

Although the two approaches described above seem like a clear dichotomy, attempts have been made to merge classification and geostatistics, taking advantage of the ability of geostatistics to deal with continuous variation and of the power of classification in discrete soil landscapes (Heuvelink and Webster, 2001). Examples of hybrid techniques are stratified variograms and fuzzy set theory (McBratney and de Gruijter, 1992). de Gruijter et al. (1997) show that fuzzy set theory can be merged with geostatistics in order to obtain continuous maps of data-derived soil classes.

1.4 Conceptual approaches to scale

A number of conceptual frameworks have been developed in recent years to address the question of scale in ecology. Among, them hierarchy theory has probably been the most influential. Hierarchy theory suggests that the organization of a system into a hierarchy can be the natural consequence of the evolution of complex systems (O'Neill, 1986; Allen and Starr, 1982). Although much of the discussion is based on rate differences and "temporal hierarchies", hierarchies can be defined based on a variety of criteria and are, therefore, dependent on the observation set. For example, O'Neill (1986) point out that hierarchies defined by functions or processes (i.e. in an ecosystem approach) can be of little relevance to define a hierarchy of species (i.e. a population approach) and vice-versa. This is due to the fact that different species can achieve the same functions, and multiple functions can be achieved by the same species. The theory states that many complex systems can be divided into levels of organization, assembled together in a hierarchy. Most interactions occur within levels and interactions between levels are limited. Levels higher than the focal level, which are larger in scale and slower in time, set boundary conditions or constraints for lower levels and lower levels "aggregate" to

influence the focal level. Dynamics from lower levels are considered too rapid for the problem of interest and dynamics at higher levels are considered too slow. Therefore, it is assumed that interactions within level can be studied without consideration to higher and lower levels. An important prediction of the theory is that a hierarchy of process rates should translate into multiple scales of pattern on the landscape. However, the detection of multiple scales of variation does not necessarily indicate that the system is hierarchically organized (O'Neill, 1986).

Although hierarchy theory generates interesting questions about the complicating effects of scale, it has still been unable to provide efficient analytical methods for addressing those issues (Wiens, 1999). Moreover, the theory mentions that the hierarchies observed are not universal in nature, that they depend on the problem and observation set. This has been considered a major limitation to the generalization of the concept. Nevertheless, several studies have been conducted to test some assumptions of hierarchy theory and its implications for the structure of ecosystems. Using transect data and four different analytical procedures, O'Neill et al. (1991)found multiple scales of pattern on a semiarid grassland, on a shrub-steppe system and in a deciduous forest. Holling (1992) compiled data from a wide array of ecosystem types and observed clumps in the body-mass distribution of animals. He suggests that this is caused by a small number of plants, animals and abiotic processes which structure ecosystems and establish dominant frequencies and associated discontinuous spatial structures. Cullinan et al. (1997) used postulates of hierarchy theory to combine the scales of pattern of each individual species within a community to estimate the scales of pattern obtained through remote sensing. For some, hierarchy theory is a framework that has very direct applications; for others, it is a heuristic framework for thinking about ecosystems.

Even outside of the well developed framework of hierarchy theory, we may still recognize that there are levels, or domains of scales, within which patterns do not change or change monotonically with changes in scale (Wiens, 1989). Between those domains, there may be sharp transitions from one set of controlling factors, to another set of factors, in which interactions may be more chaotic or unpredictable. Whether or not ecosystems are hierarchically structured has great significance for ecological research. A strong hierarchical organization would allow processes to be studied in a scale-specific manner or within domains of scale. A lack of hierarchical organization would mean that interactions between components of the system occur equally at all scales. This would prevent studies from choosing scales of research that correspond with true levels of species-environment assemblages.

Numerous authors have suggested general processes to explain the scale-dependent

nature of biotic assemblages. In some studies, patterns at large scales have been postulated to be caused by the underlying physical environment, whereas smaller scale patterns were postulated to arise from biotic interactions (Brown, 1984; Levin, 1992; Cullinan et al., 1997). A change in scale may also be associated with a change in predictability of patterns. The stochastic nature of small scale patterns may make them unpredictable, while the more regular and slower nature of larger scale processes may make large scale patterns more predictable (Levin, 1992). Reed et al. (1993) found that the relationship between environment and species composition increased with the grain size (quadrat size) but decreased with the extent of sampling. Another widespread hypothesis in vegetation science is the idea that scales of variation are related to stage of succession. For example, plant communities in earlier succession could be more heterogeneous than late successional stages (Greig-Smith, 1952). Schaefer (1993) could not validate this hypothesis in a taiga plant community regenerating after fire.

Fractal geometry Mandelbrot (1983) has been extremely influential in the recent concern over the problem of scale (Xu et al., 1993; Halley et al., 2004). This theory demonstrates that very complex, multi-scaled patterns can arise from simple mechanisms. Such mechanisms may occur at small scales and still structure larger scale processes. Fractals thus emphasize the possible prevalence of cross-scale mechanisms, whereas hierarchy theory highlights scale-specific processes and interactions. Burrough (1983) explains the potential use of fractal concepts in soil sciences. He describes soil variables as fractals because increasing the scale of mapping continues to reveal more and more detail. However, these fractals are not 'ideal' because soils generally lack properties of self-similarity. Moreover, fractal models generally have unlimited a priori variances. Although contributions have indicated that this could be the case for some soil variables (Bell et al., 1993), this goes against some general conceptions of variability in soils and geostatistical assumptions of stationarity (Webster, 2000).

1.5 The challenge of understanding in the presence of complexity

In a landmark paper in ecology, Levin (1992) suggested that "the key to prediction and understanding lies in the elucidation of mechanisms underlying observed patterns". Understanding is defined by (Pickett et al., 1994) as "an objectively determined, empirical match between some set of confirmable, observable phenomena in the natural world and a conceptual construct."

When simple problems are assessed (i.e. simple conceptual constructs and observable phenomena with few elements), traditional falsificationist and experimental models may provide the best approach towards understanding. From the classification described in Diamond (1986), spatial observational studies can be considered 'natural snapshot experiments', in which "one observes only a final steady state [...], but not the trajectory that lead to it". However, the word 'experiment' may be avoided because they lack control and repeatability (Wiens, 2001). The context in which such studies are conducted, much like landscape ecology studies, is in sharp contrast to that of traditional experimental studies. Inadequacies of the falsificationist and reductionist approaches when dealing with complex ecosystems have been described by many (Levins and Lewontin, 1985; Vepsäläinen and Spence, 2000; Taylor and Haila, 2001). "[...] the potential process-pattern linkages are likely to be sensitive to the details of the spatial configuration of a mosaic and [...] simple experiments are likely to omit key components among a web of interacting elements."(Wiens, 1999). For Quinn and Dunham (1983), a strong adherence to the hypothetico-deductive research framework may be counterproductive when multiple causal factors are interacting.

The appreciation of the overwhelming complexity of ecosystems has lead to the development of ecosystem ecology and of new tools and theories that attempt to formally address complexity instead of ignoring it. Landscape ecology later developed in the same general philosophy as ecosystem ecology but with a strong emphasis on spatial patterns and scaling issues and with a close association with technologies such as geographic information systems and remote sensing. According to Wiens (1999), the mantra of landscape ecologists could be 'embracing complexity'.

One of the greatest challenges of ecosystem and landscape ecology is generating understanding in the presence of complexity. Quinn and Dunham (1983) suggest that an objective of biological research should be to assess the relative contributions of a number of causal agents operating simultaneously. Vepsäläinen and Spence (2000) recommend that ecologists dealing with complex systems should not try to find simple universal laws or theories which are not put in the context of the local contingencies, because these are doomed to fail. They argue that we should instead develop general explanatory frameworks, that are "road maps to solutions rather than solutions themselves". The goal of these frameworks should be to reach bounded generalizations about the elements of the system. Patterns and processes that we wish to study establish the focal level of the generalization. It is then necessary to delineate the upper-level boundary conditions and the lower level conditions that determine the possible states of the focal level.

Thus, the requirement for explicit, coherent and logically constructed theoretical

frameworks does not imply that these are the goals of ecological inquiry. The goal still needs to be obtaining useful generalizations. Cooper (1998) suggests that generalizations should be categorized in a continuous space between theoretical principles, causal generalizations and phenomenological patterns. With spatial observational studies, it may be very difficult to obtain purely causal generalizations, or universal laws applicable to very broad systems. However, we must avoid to remain completely descriptive, blocked in the contingency of the local system. Rather, the objective should be to go as far as possible along the continuum from pure contingency to pure cause, from the description of phenomena to the identification of processes.

1.6 Multi-scale analysis as a tool to generate understanding

1.6.1 What is multi-scale analysis?

Vegetation scientists generally use sampling quadrats that range in size from a few centimeters to several meters, depending on the types of species sampled. Sampling cores used in soil science are rarely more than a few centimeters in size. To study the spatial structures of soils or vegetation, scientists thus have to collect samples on transects or quadrats. I define here spatial observational studies as any study for which data is recorded on a fixed small scale support at spatially referenced locations and for which spatial location is inherently taken into account in the analysis. Typically, spatial 'observations' are recorded on a series of quadrats, adjacent or not, or on a quasi-punctual support at locations on a grid or transect. In spatial observational studies, all sampled variables are treated as 'stuff'. This definition thereby excludes methods based on the discretization of spatial patterns into patches, such as landscape metrics (Fortin, 1999).

Methods of multi-scale analysis are defined here as statistical procedures that attempt to separate the variation in total variables sampled in a spatial observational study into scale-specific components of variation (this definition is expanded in Chapter 2).

1.6.2 Why multi-scale analysis?

One of the main postulates of this doctoral thesis is that methods of multi-scale analysis are important tools that will help us reach generalizations and facilitate understanding of natural systems in spatial observational studies. Multi-scale analysis can be

conducted with a variety of objectives in mind, some of which are summarized below.

Assess same-scale relationships. Multi-scale analysis may allow focusing on variables with similar spatial structuring and with stronger relationships at each scale while reducing the attention on variables without significant contributions at those scales. This may facilitate the identification of possible cause-effect relationships. The patterns and relationships at larger scales may help us identify the boundary conditions affecting the focal scale. The patterns and relationships at smaller scales can define the possible states at the focal scale (based on Vepsäläinen and Spence, 2000).

Define and summarize the spatial structure of studied variables. The description of the spatial structure of variables with complex spatial patterns can be very difficult. Maps showing zones of high or low abundance may prove inadequate to describe properly the spatial structures. Methods for multi-scale analysis may be used to provide a summary of the spatial structure of each variable by quantifying the amount of variation found at different scales of observation.

Separate the contingent variation from the consistent patterns. At the heart of any method of exploratory data analysis is the desire to separate the variation that is contingent on the observation set from the patterns that are consistent. For achieving this, the notion of repetition is crucial. In classical exploratory analysis, the criteria for consistency can be extremely varied, whether we look for binary relationships, gradients, groups in multiple species, etc. In spatial analysis, the consistency describes the spatial entities defined from the data. Is the size, shape and structure of the entities the same throughout the sampling site or is it different for different sections of the site? Small scale spatial features may repeate themselves more on a sampling site of fixed extent than larger scale ones. Thus, scale can be used as a surrogate for consistency and can be used to identify patterns that are more likely to be generalizable. This idea will be a central point of interest in this thesis.

Isolate scale-specific components. There may be a discrepancy between the scale at which information is required and the scale at which it is available. In such cases, multi-scale analysis may be used to isolate or filter out scale-specific components occurring within the grain size and the extent of the study. For example, a large scale component extracted by multi-scale analysis could be used as a layer in a larger scale model.

1.6.3 Challenges in multi-scale analysis

Ultimately, the focus of inference in multi-scale analysis are the characteristic scales. They are the inherent properties of the natural systems around which our research objectives are defined. Characteristic scales are never directly accessible to the scientists

and, as was outlined above, their existence can even be questioned. When dealing with 'things', this debate can be somewhat futile because the definition of the entity to which a scale is attributed is generally clear (though not always, see Allen and Hoekstra, 1992). However, when dealing with groups of 'things' or with 'stuff', the choices of concepts, definitions, scales of observation and methods of analysis will have a greater influence on the final results of the analysis simply because the exact definition of what it is that we must measure is not as clear. Since spatial observational studies treat all the sampled variables as 'stuff', these issues are crucial here.

The scales of observation and the scales of analysis act as filters separating the scientist from characteristic scales (Fig. 1.2). The scales of observation act as a first filter that greatly affects how the studied natural system will ultimately be perceived by the scientist. The problems related to the contingency of the observation set are a central issue in multi-scale analysis. The extent of the sampling grid imposes a limit to the largest observable scales while the grain size and sampling distance impose limits to the smallest scales.

The scales of analysis act as a second filter separating the scientist and the natural system. The greatest difficulty of multi-scale analysis lies in the fact that the decomposition of total variables into scale-specific components is not unique. This idea is conveyed in the geographical context by the Modifiable areal unit problem (MAUP, Openshaw, 1985; Jelinski and Wu, 1996). The MAUP recognizes that inferences or statistical measures may change according to the boundaries of the zones chosen to describe the data (zoning problem) and according to the scale at which the data is aggregated (aggregation or scale effect). Different methods of multi-scale analysis will have a different approach to identify spatial features in the data (different zoning) and to determine the scales at which variation is to be decomposed (scale effect).

In recent years, methods of multi-scale analysis have undergone intensive development and have been applied extensively in ecology and agronomy. Several of these methods have been developed in other fields of research, such a mining or image processing, before being used in ecological or agronomic applications. Even within their originating fields, the development of methods of statistical analysis is often carried in the laboratories of mathematical or statistical theoreticians and the results are published in specialized journals. The mathematical formalism and the heavy use of specialized jargon often makes those publications inaccessible to the practitioner. Perhaps more importantly, those methods can be developed with objectives that can be unrelated to the objectives of the practitioners. The result of this process is that statistical methods are often misunderstood by practitioners and can be misadapted to their needs.

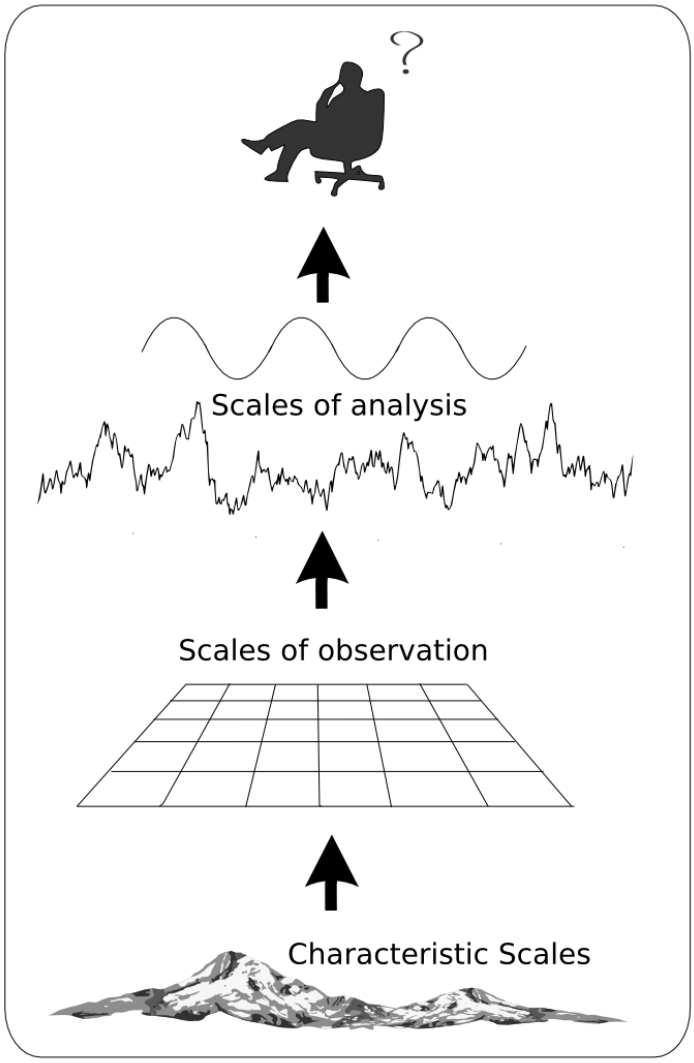

Figure 1.2: Scales perceived by the researcher are filtered by the scales of observation and the scales of analysis specific to each method.

Geostatistics provides a good example of the problems related to these issues. Developed in the fields of mining and geology, the theory of geostatistics was laid with objectives specific to those fields and with a jargon particular to those fields. The theory of geostatistics is probably the statistical approach that possesses the most elaborate frameworks for addressing practical and theoretical problems related to the characterization of spatial structures, spatial estimation, spatial uncertainty and scale. Although geostatistical methods have been increasingly applied to multiple fields of research, to this day, the theory is often misunderstood. One particular method of geostatistics, regionalized multivariate analysis, was proposed as a powerful method for multi-scale analysis. It offers several advantages over existing methods and combines the abilities

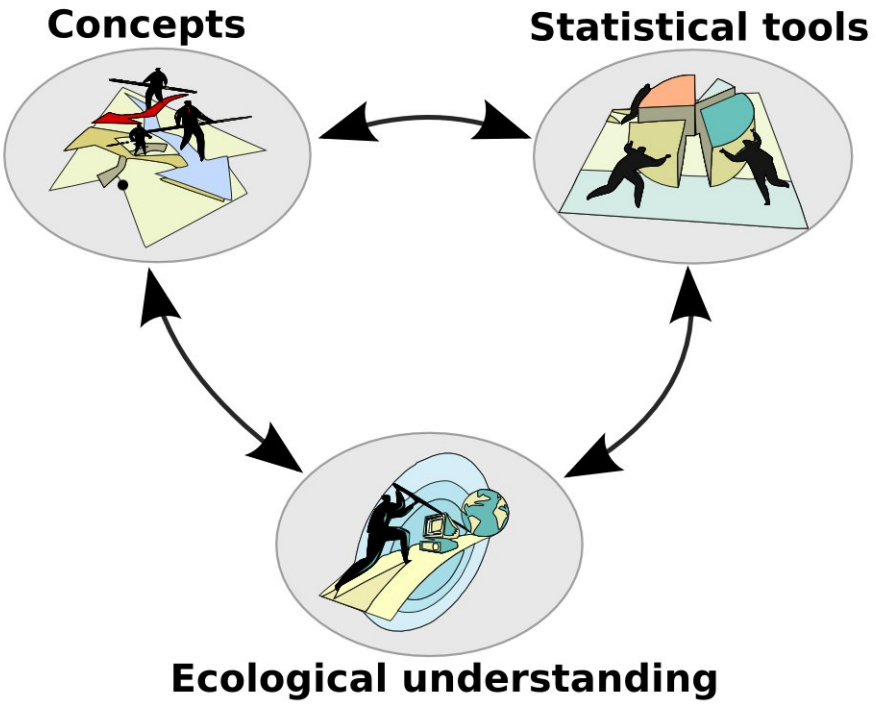

Figure 1.3: Establishing better linkages between concepts and statistical tools are fundamental to generate better ecological understanding. A better understanding can help develop new concepts and improve statistical methods.

to detect dominant scales of variation, to assess scale-dependent relationships between variables and to estimate scale-specific components of variation.

While scientists increasingly collect spatial datasets in an attempt to improve their understanding of spatial pattern and scale in natural systems, there is an increasing need for better linkages between concepts, statistical tools and ecological understanding (Fig. 1.3). Coherent conceptual frameworks are needed that address the issues of spatial pattern and scale. Methods of multi-scale analysis have to be consistent with the conceptual frameworks and with the research objectives of the practitioners. The use of statistical methods that are coherently developed and properly used by practitioners should help the generation of ecological understanding. Based on the results of the analysis it may be possible to make different choices for the analysis, or even modify the existing statistical methods, and obtain new results. The process can be repeated until a satisfactory solution is found.

1.7 Objectives and outline

The main objective of this doctoral research project is to provide better linkages between concepts, statistical methods of multi-scale analysis and ecological understanding, by focusing more particularly on the geostatistical method of regionalized multivariate analysis. Chapter 2 will provide an overview of methods available for multi-scale analysis. The analysis of the same dataset with all methods, as well as the use of a common language to describe the features of each method will facilitate the comparison between methods. This chapter will identify regionalized multivariate analysis as possessing several advantages over other methods of multi-scale analysis, but potential problems with the method will be identified and addressed this chapter and in the two following chapters. In Chapter 3, the use of kriging of regionalized components for providing scale-specific components of variation that are consistent with the linear model of coregionalization is questioned and an alternative based on conditional co-simulation is proposed. In Chapter 4, the question of uncertainty in coregionalization analysis is addressed. A theoretical framework is developed to characterize the different aspects pertaining to the uncertainty and a mathematical formalism is presented to quantify the uncertainty. The use of some simple scenarios highlights the high levels of uncertainty associated with coregionalization analysis. In Chapter 5, a framework for discussing the problem of inference in spatial and multi-scale analysis is presented and used to provide a critical analysis of coregionalization analysis. The logic leading to the development of a new approach for multi-scale analysis, called coregionalization analysis with a drift (CRAD), is presented and the approach is briefly described. The approach is described in detail in two co-authored research articles (Pelletier et al., 2009a,b). In Chapter 6, CRAD is applied in an agronomic setting to discuss the problems of inference and study the relationships between spatial structure and temporal heterogeneity. In Chapter 7, the use of CRAD is extended to perform multi-scale causal modelling of the relationships between trees and soils in a forest ecosystem. The Conclusion will return to the important points of the thesis and highlight new research needs.

Chapter 2

Comparison of methods for multi-scale analysis

The increasing awareness of the issue of scale in all areas of environmental sciences has fostered the development and use of numerous statistical methods for scaling and for multi-scale analysis. These methods originated from various field of research such as geography, signal processing, image processing, ecology, geology and mining. Their application has led to new findings and has helped the development of theories in all areas of research. The variety of sources has created a very diverse language with different definitions for concepts such as scale, scale-dependence, autocorrelation and spatial pattern. The true meaning of concepts can get blurred by mathematical formalism or evasive statistical concepts. Also, practitioners don't always have clear and precise objectives before conducting the analysis. It can then become very hard to match the world views and objectives of the practitioner to those implicit in the statistical methods. The development and use of methods for multi-scale analysis must thus be guided by coherent conceptual and analytical frameworks.

A particular type of exploratory multi-scale analysis focuses on the separation of variables from a spatial observational dataset into multiple scales of variation. These datasets are generally sampled on a grid of points or quadrats covering a single region and are rarely replicated. Studied variables are generally considered to be continuous (i.e. 'stuff') and their spatial distribution is often complex enough so that the decomposition of a total variable into multiple scales of variation does not have a unique solution.

In this Chapter, a selection of methods proposed for the multi-scale analysis of spatial observational data in the fields of ecology and agronomy will be described and compared. All methods are broadly classified as spatial deterministic or probabilistic. Deterministic methods are further classified into drift estimation methods, graphical

methods and orthogonal decomposition methods. The probabilistic approach of regionalized multivariate analysis is the discussed. To facilitate the comparison, each will be used to analyze the same dataset, comprised of tree height data on two transects in the Morgan Arboretum. Although the capabilities of each method for multivariate analysis will be described and discussed, the focus in this chapter will be on the identification of dominant scales of variation and the separation of the total variables into scale-specific components in the univariate case. The objective is to provide an overview of the variety of methods for multi-scale analysis and to compare the different approaches in simple terms and on a common ground. The intention is not to provide an exhaustive description of each method, as this is published elsewhere.

2.1 Description of dataset

The dataset used below to compare the different procedures constitutes of LIDAR (Laser Detection And Ranging) elevation data covering the Morgan Arboretum, a forested area of 250 hectares on the western part of the island of Montréal in Southern Québec. The forest of the Arboretum is dominated by deciduous vegetation on a variety of geomorphological types. The LIDAR data was acquired by Lasermap Image Plus in May 2002 from an airborne sensus. It was separated by the company into two layers, one covering ground points and another for non-ground points containing primarily reflections from tree and shrub branches. The complete dataset is comprised of over 1.9 million points for the ground layer and three million points for the canopy layer. For the purposes of this study, two transects were generated from the canopy layer with 512 points separated by two m for a total of 1022 m for each transect (Fig. 2.1). A dyadic number of sampling locations (512) was chosen to facilitate the use of some of the methods described below.

The first transect (Transect I), was selected to cover a relatively uniform section that is dominated by American beech, eastern hemlock and red maple on a fluvial deposit of variable thickness. This transect was selected to represent a dataset without large scale pattern but with some apparent small scale variation. Transect II was selected to cover a catena that ranges from a fluvial ridge to a glacial till along an elevation gradient. The associated vegetation varies from a American beech-red maple forest to a sugar maple-hickory forest along the same gradient. Each location along the transects was assigned the average value of the LIDAR elevations within a one m radius around this location. For Transect I, this gave an average of 3.32 and 1.05 points per location for the canopy and ground layers, respectively, and 3.41 and 2.43 for Transect II. The

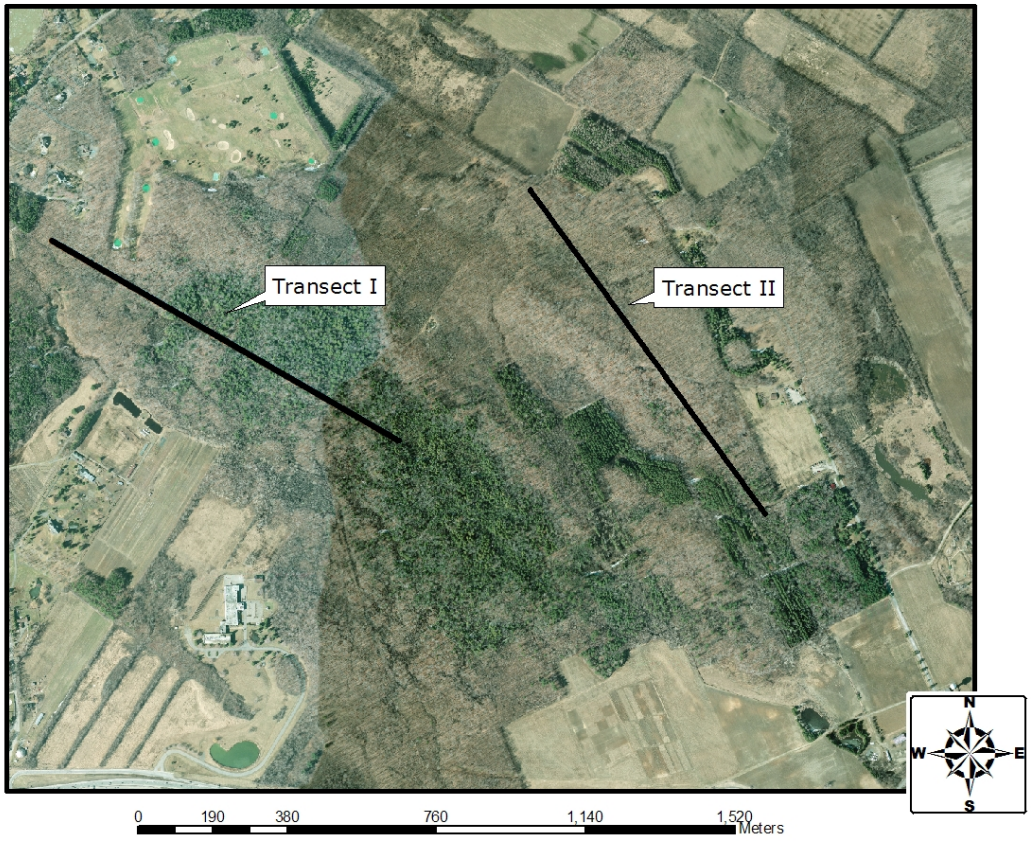

Figure 2.1: Location of the two transects in the Morgan Arboretum.

canopy height data used in the following comparisons was obtained by subtracting the ground layer values from the canopy layer values along the transect (unstandardized values are plotted in Fig. 2.2, standardized values are used in all subsequent analyses).

In order to facilitate the interpretation of results, 100 simulated random variables were generated on a transect of the same length as the others. These variables consisted of non-spatially distributed numbers with a normal distribution ($N(0,1)$, white noise). Where relevant, the methods presented below were also performed on the simulated random transects for comparison purposes.

2.2 Classification of methods

2.2.1 Spatial deterministic methods

In the deterministic formalism, a variable is seen as possessing a fixed value at each sampling location. In the non-spatial case, the sample data are analyzed with no

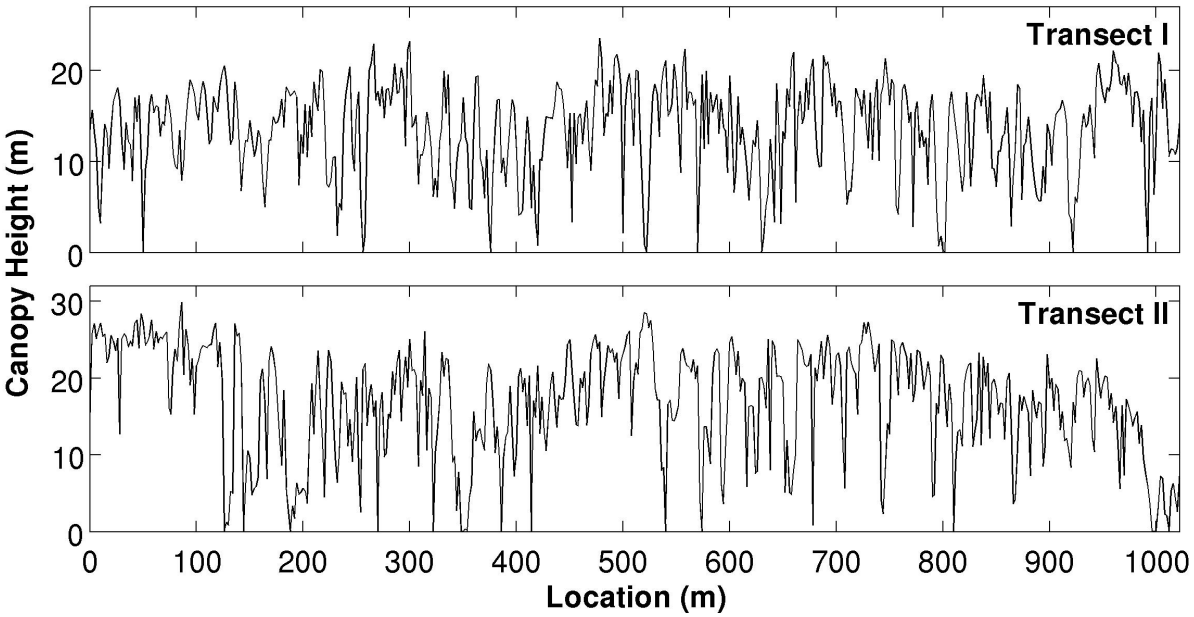

Figure 2.2: Unstandardized canopy height data along Transects I and II

reference to the notion of random variables or to the location of the samples. Thus, all measures are global properties of the pooled samples. This can include simple statistics, such as a sample mean or a (pseudo) sample variance, or graphical representations such as histograms (Wiens, 1999). In deterministic methods of spatial analysis, the location of the samples is used in the analysis, but the notion of random variables is not used.

In the context of multi-scale analysis, we will define three types of spatial deterministic models. A first group of methods separates the data into spatial and residual components:

$$z(u) = m(u) + e(u)$$

where $m(u)$ is a deterministic spatial, or drift, component at the spatial location u and $e(u)$ is a residual (non-spatial and non-random) term. These are not generally considered suitable for multi-scale analysis since they only extract one spatial component corresponding to a unique scale from the spatial dataset. However, two of them will be discussed here, trend surface analysis and local drift estimation procedures, because they will be important in the subsequent discussion in later chapters. Single scale methods, especially the local procedures, are extensively used in Geographic Information Systems when the datasets are in the form of rasters. Other methods exist that perform the same type of decomposition, including smoothing splines (Silverman, 1985) and median polish (Cressie, 1986).

A second group of methods decompose the total variable into a number (S) of scale-specific spatial components $z_s(u)$:

$$z(u) = \sum_{s=1}^{S} z_s(u)$$

Because the components are generally defined as orthogonal (their scale product is zero, $\sum_u z_s(u)z_{s'}(u) = 0$), the sums of squares for the S components add up to the total sum of squares. This orthogonality allows the addition of multiple scale-specific components to obtain components comprising "bands of scales", for example: small, medium and large scale variation. These methods will be referred to as orthogonal decomposition methods and three of them will be described here: the discrete Fourier transform (DFT), principal coordinate analysis of neighbourhood matrices (PCNM, DBEM), and the discrete wavelet transform (DWT).

A third group of deterministic methods for multi-scale analysis attempts to change the scale of observation in order to infer the scales of variation in the data. Unlike the first type of methods, such methods do not decompose the total variables into scale-specific components. Rather, they generally provide a representation of the variance as a function of scale of observation. This group comprises a large variety of methods, but four of them will be described here: the blocked quadrat variance methods (TTLQV, 3TLQV), lacunarity analysis, the continuous wavelet transform (CWT) and experimental variograms.

2.2.2 Spatial probabilistic approaches

With probabilistic methods, the sample values are used to describe properties of a larger (theoretical) population that is only partially sampled. In classical (non-spatial) probabilistic methods, sample values are assumed independent and each sample value contributes equally to the estimation of the population parameters. Generally speaking, the objective of probabilistic methods is inference or generalization beyond the sample data. Geotatistics was originally developed in a mining context by Danie Krige (1951), Georges Matheron (1962; 1963b; 1963a) and others (Cressie, 1990) as a response to the inadequacy of classical statistical methods for analyzing complex spatial data. At the core of the geostatistical methods is the probabilistic theory of random functions, which puts together the notions of randomness and spatial structuring. To understand random functions, we can imagine an infinite number of spatial locations in an infinite space. At each location, an imaginary variable possesses a unique value. A given field of finite shape and dimension can be located at random within the infinite space, giving

rise to a realization of the variable within the field. This variable is random because there is an infinite number of possible realizations depending on the location of the field, yet it is also spatially structured. The type of spatial structure is expressed through the autocovariance function, or variogram function (depending on the type of stationarity), characterizing the random function. The variogram and autocovariance functions are not spatial laws linking specific values of the realizations at different locations, they are parameters of the random function that describe expected values, or averages over an infinite number of values in an infinite space.

Regionalized multivariate analysis comprises a set of geostatistical techniques for multi-scale analysis based on the linear model of coregionalization (LMC), which sees total variables as the result of superimposed random functions. These methods allow the separation of the variance of total variables into scale-specific variance components and the estimation of scale-specific spatial components corresponding to the scales of the LMC. These methods will be extensively discussed here and in the subsequent chapters.

2.3 Drift estimation procedures

2.3.1 Trend surface analysis

Trend surface analysis (TSA) is a simple method of spatial analysis in which a polynomial function of a chosen order is fitted to the sample data by a regression (least-squares) technique (Gittins, 1968). Depending on the order chosen, the scale of the variation captured by the function will change, with lower orders describing larger scale variation. Nevertheless, the number of terms of the polynomial function increases rapidly as the order increases. Thus, the use of TSA is generally limited to the extraction of the larger scale spatial structure and the order of the polynomial rarely exceeds four or five.

2.3.2 Local drift estimation procedures

In contrast to TSA, local drift estimation procedures give a smoothed representation of the sequence of data rather than a global model linking values at each point. The simplest case is a moving average approach in which a window of a given size centered on each sampling point is specified and all sample values within the window are averaged. The average value is assigned to the center unit and this calculation is repeated for each sampling location. The appearance of the smoothed representation is thus entirely based on the data sequence and the window size. Larger windows will provide smoother

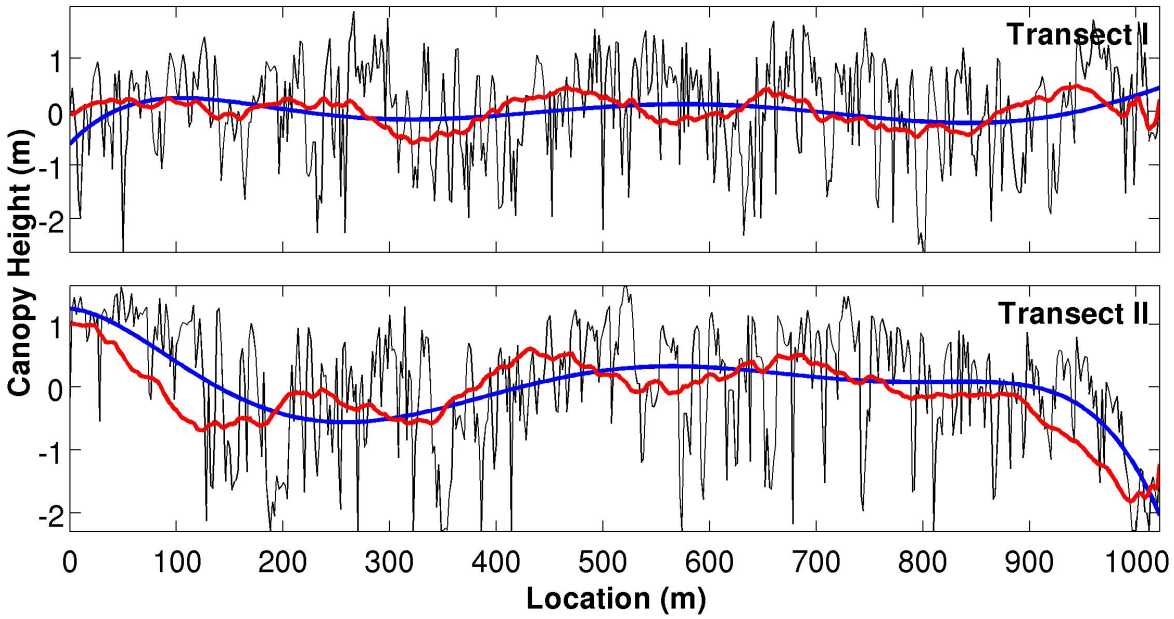

Figure 2.3: Comparison of original data (black line) and single scale spatial components obtained with trend surface analysis (blue line) and moving average (red line) on the three transects.

estimated drifts while the complexity of the total variable will be gradually retained as the window size decreases. Instead of using the average, a polynomial of low order can be fitted inside each window to capture more elaborate spatial structures.

2.3.3 Comparison of methods

TSA was used to extract spatial components on the two transects with a polynomial of order six (Fig. 2.3, blue lines). The trend surface only captures 2.4 percent of the variation in Transect I, while it captures 24.6 percent on Transect II. The moving average approach was used on the three transects (Fig. 2.3, red lines) with a window size of 75 m, chosen to capture variation at a scale similar to TSA. The drift component captures 6.4 and 30 percent of the variation in the total variable on Transects I and II, respectively. With the two methods, the mean spatial component calculated over the 100 simulated random transects consist of straight lines with an overall sample mean near zero).

The two methods capture very small amounts of the variation on Transect I indicating the inability of those methods to capture small scale variation. The drift components captured by TSA are smoother than those captured by the local procedures. Local procedures have the advantage over TSA of being 'model-free', meaning that there is

no predefined model for the spatial structure they capture in the total variable. Because of this lack of model, the scale corresponding to a drift obtained with the local procedure may be difficult to infer. This method should however be more flexible for capturing more varied types of spatial variation. The complexity of the component obtained with the local procedures is expected to increase with decreasing window size. Thus, both of these methods are better suited to the extract large scale variation in a dataset.

2.4 Orthogonal decomposition methods

2.4.1 Discrete Fourier transform

Methods for spectral analysis were originally developed for time series analysis, signal processing and physics (Diggle, 1990), and were used in several ecological applications (Hill, 1973; Platt, 1975; Ford and Renshaw, 1984; Renshaw, 1997; Denny et al., 2004; Saunders et al., 2005). The discrete Fourier transform (DFT), also called harmonic analysis, is a method of spectral analysis that can be used for multi-scale analysis by decomposing the signal (total variable) into a series of orthogonal cosine waves of specific frequencies called Fourier frequencies. In the context of multi-scale analysis, the frequency, or equivalently the wavelength, characterizing each wave corresponds to a specific scale of variation. Each wave can be expressed in the form

$$z_s(u) = A_s cos(\omega_s u + \phi_s)$$

where ω_s is the fixed frequency corresponding to each wave, while the amplitude A_s and phase ϕ_s are adjusted globally to the data. The periodogram, or auto-spectrum, describes in graphic form the amplitude (y axis) as a function of frequency (x axis) characterizing each wave. For a transect of even length L with N regularly spaced sampling points, only $N/2$ waves (plus an intercept) are needed to express the total variation; the first wave then has a wavelength of L, the second of $L/2$, and so forth until the $N/2$-th wave, corresponding to the Nyquist frequency that has a wavelength of $2L/N$. Although individual cosine waves show clear patterns of periodicity, the sum of multiple cosine waves can give rise to any type of spatial pattern. In fact, the sum of all waves described in the periodogram is always equal to the total variable. This means that any total variable on a transect with N sampling points can be perfectly reconstructed from $N/2$ waves.

The use of DFT for bivariate and multivariate analysis remains somewhat limited. Nevertheless, for bivariate and multivariate analysis, the cross-periodogram between

two variables is calculated in a way similar to the auto-spectra (Priestley, 1981, p. 694). It highlights scales of periodic fluctuations that are shared between two variables. Another approach for multivariate analysis would be to independently separate each total variables into scale-specific components with the DFT and to use standard methods of multivariate analysis on those components.

The DFT can be readily used for analyzing regularly spaced two-dimensional data (Renshaw and Ford, 1983) and it is possible to modify the method for use with irregular grids. However, without modification, the use of the DFT on irregular grids will yield imperfectly formed waves and the scale of the spatial features represented by such 'waves' can be somewhat ambiguous. The method can be modified to yield perfectly formed waves, but this comes at the cost of loosing the orthogonality between waves.

2.4.2 PCNM and DBEM

Principal coordinate analysis of neighbourhood matrices (PCNM) was first presented by Borcard and Legendre (2002) as a method for the multi-scale analysis of multivariate spatial data. It has since been classified by Dray et al. (2006) into a broader family of methods called distance-based eigenvector maps (DBEM). An empirical comparison of PCNM and geostatistics can be found in Bellier et al. (2007). In PCNM, a number of orthogonal wave-like spatial components representing specific scales of variation (PCNM variables) are obtained through a decomposition of a truncated distance matrix. In the DBEM approach, a similarity matrix, or spatial weighing matrix, is used as the basis for the decomposition instead of the truncated distance matrix. Similarly to the DFT, the amount of variation explained by those waves is determined by a least-squares fitting of the PCNM variable to the total variable.

When applied to datasets with regularly spaced observations, it can be shown that the PCNM variables are a series of sine waves and the method is very similar to spectral analysis. Each PCNM variable can be described by a function of the form

$$z_s(u) = a_s + b_s \sin(c_s + d_s u)$$

where a_s, b_s, c_s and d_s are constants specific to each PCNM variable and u is location along the transect. In fact, with our example dataset (see below), all odd PCNM variables nearly correspond to the Fourier frequencies while the even variables correspond to fractional frequencies. In contrast to spectral analysis, in PCNM, the exact frequency and phase of the PCNM variables are entirely determined by the sampling grid configuration. Relative to DFT, the lack of phase adjustment in PCNM allows a

larger number of orthogonal sine waves to be obtained from a given sampling configuration. Although it was originally proposed by Borcard and Legendre (2002) that one should only retain the PCNM variables that correspond to positive eigenvalues, it can be shown that the N PCNM variables are needed, including the ones with negative eigenvalues, to capture the total sample variance. In other words, twice as many waves are needed in PCNM to capture the same variation as in the DFT. A smaller number of scale-specific components can be obtained by adding bands of scales corresponding, for example, to small, medium and large scales. For doing this, permutation testing can be performed to determine which PCNM variables contribute significantly to the variation in the dataset (Borcard and Legendre, 2002).

In PCNM and DBEM, the decomposition of each total variable into PCNM variables is performed and the resulting scale-specific components can be subjected to any standard method of multivariate analysis (Borcard and Legendre, 2002; Borcard et al., 2004). PCNM does not require special modifications for datasets with irregularly spaced sampling locations. However, the PCNM variables do not describe perfect sine waves when applied to irregular sampling schemes and the scale(s) depicted by such wave-like functions can be somewhat ambiguous. Dray et al. (2006) suggested that the DBEM approach allows more regular waves to be obtained on irregular grids if a proper similarity matrix is used. This method is highly dependent on the similarity matrix used to perform the decomposition.

2.4.3 Discrete Wavelet transform

One of the main limitations of DFT and PCNM is the requirement that the a function be fitted to the whole signal at each scale. This implies that such methods are more appropriate for datasets exhibiting periodicity and a rather regular form, without isolated peaks or troughs. Wavelet analysis comprises a family of methods that is better suited for datasets exhibiting such irregularities. It was developed mainly in the areas of signal and image processing (Torrence and Compo, 1998; Hubbard, 1998), and it has been used considerably in ecology (Dale and Mah, 1998; Csillag and Kabos, 2002; Keitt and Urban, 2005; Saunders et al., 2005). The mathematical and theoretical formulation of wavelet analysis is somewhat complicated; only the important points will therefore be discussed here. Instead of using a sine wave as a basis for the multi-scale decomposition of the data sequence, the wavelet transform uses a dilated and translated basic local function, called a wavelet (Lark and Webster, 2001; Csillag and Kabos, 2002; Keitt and Urban, 2005). Two forms of decomposition can be distinguished, the discrete wavelet transform (DWT, also called multi-resolution analysis) and the contin-

uous wavelet transform (CWT). The latter will be discussed in the next section. In the simplest formulation of DWT, a data sequence of length $N = 2^S$ (a dyadic length) is divided into S levels of decomposition. At the s-th level, the data sequence is divided into $N/2^s$ windows, each spanning 2^s sampling locations. Within each window, the amplitude of the dilated wavelet function is adjusted to the data. Wavelet coefficients reflect the amplitude of the wavelet function fitted to the data sequence within the windows. A high positive or negative value indicates a good match between the function pointing, respectively, upwards or downwards and the data sequence. In the context just described, the choice of wavelets for DWT is generally limited to those providing an orthogonal basis for the decomposition. These include, the 'Haar', 'Daubechies' and 'Coiflet' families.

Similarly to PCNM and DFT, DWT results in a separation of the data sequence into a hierarchy of orthogonal scale-specific spatial components. Unlike the other methods, however, DWT yields scale-specific components with juxtaposed spatial features of the same size, but of different amplitudes. The 'phase' of each wavelet is determined by the window locations and is therefore generally not adjusted to the data. Modifications to the method have been proposed to prevent artefacts caused by this lack of phase adjustment (Keitt and Urban, 2005). In DWT, the number and scales of the scale-specific components are imposed by the number of sampling locations.

In order to assess scale-dependent relationships, the wavelet covariance can be calculated from the product of the wavelet coefficients for two variables (Keitt and Urban, 2005). Another approach for multivariate analysis would be to independently separate each total variable into scale-specific components and to use standard methods of multivariate analysis on those components, as in PCNM and DFT. The application of DWT to irregular transects and grids poses similar challenges as the DFT. Its use is thus generally restricted to regular transects and grids.

2.4.4 Analysis of transect data

The DFT, PCNM and DWT were applied to Transects I and II. For PCNM, code was written in the Matlab language (The Mathworks, 2008) to implement the original method described in Borcard and Legendre (2002). For the DWT, the Coiflet wavelet (Fig. 2.4) was used as the basis for the decomposition because it provides a relatively smooth function. The 'ribbon plot' in Figure 2.5 depicts the first 25 DFT waves, or scales, fitted to the data sequence of Transect II. The same was done for the first 50 scales of PCNM (Fig. 2.6) and the 7 scales arising from the DWT decomposition (Fig. 2.7). The periodograms for Transect I (Fig. 2.8) and II (Fig. 2.9) are shown for DFT

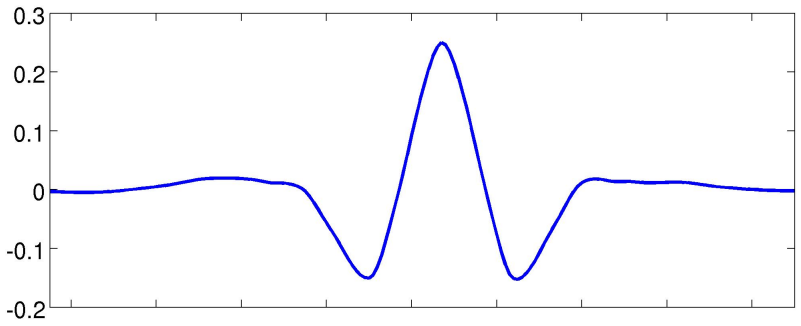

Figure 2.4: Coiflet wavelet basis used in the CWT analysis of transects.

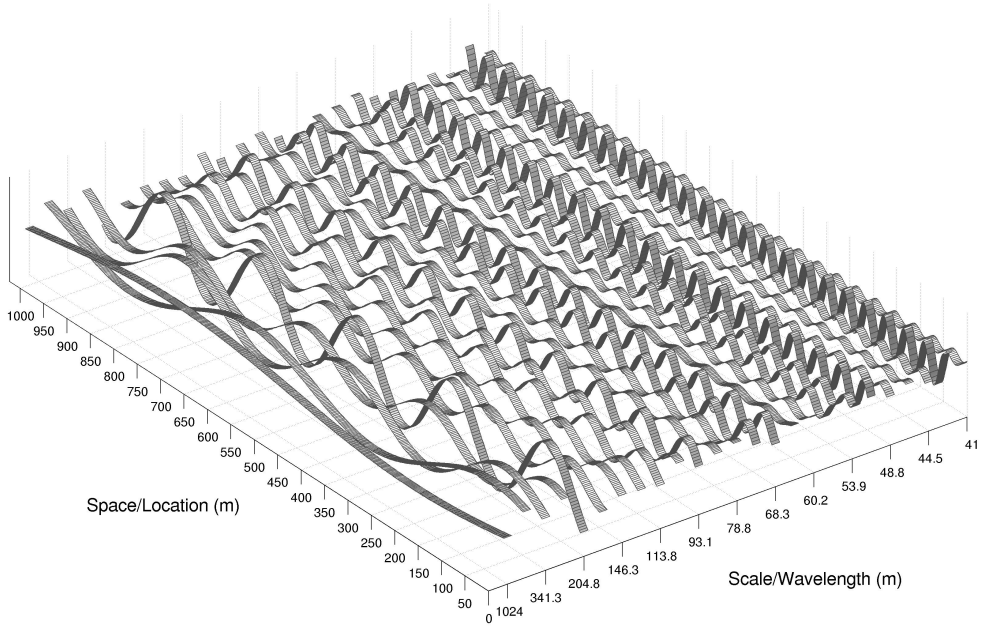

Figure 2.5: Ribbon plot showing the first 25 scales of the DFT decomposition of the total variable on Transect II.

and PCNM. To facilitate the interpretation of the results and the comparison with other methods, wavelength is used on the x-axis of periodograms and variance on the y axis. Scale-specific components were grouped (summed) to perform a decomposition of the total variables at large scale, medium scale, small scale, and non-spatial. Waves 1-3, 4-10, 11-30 and 31-255 were added accordingly for DFT; Waves 1-7, 8-19, 20-63 and 64-255 for PCNM; 1-2, 3-4, 5 and 6-7 for DWT. The spatial components are shown for Transects I (Fig. 2.10) and II (Fig. 2.11).

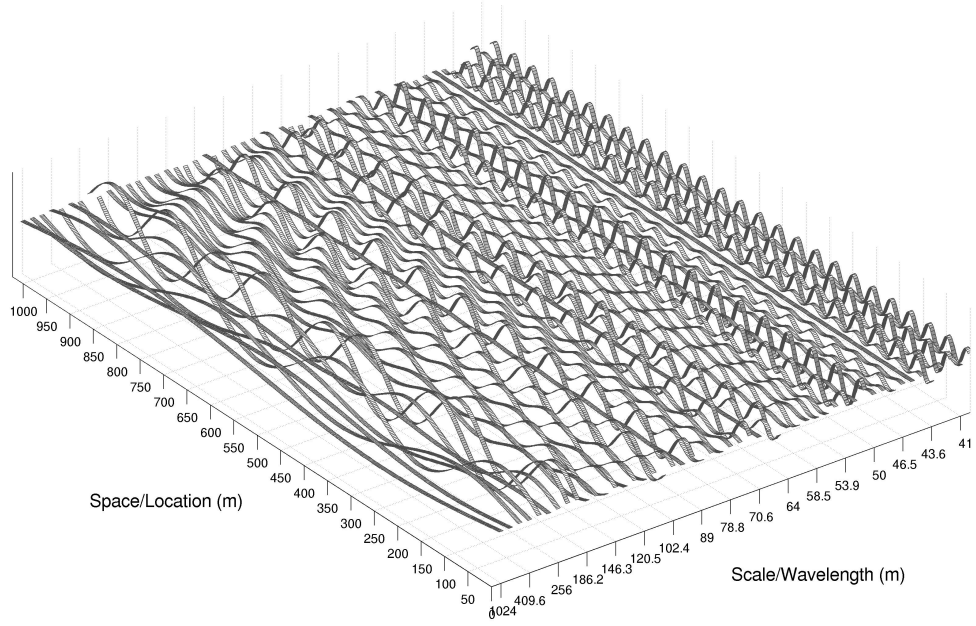

Figure 2.6: Ribbon plot showing the first 50 scales of the PCNM decomposition of the total variable on Transect II.

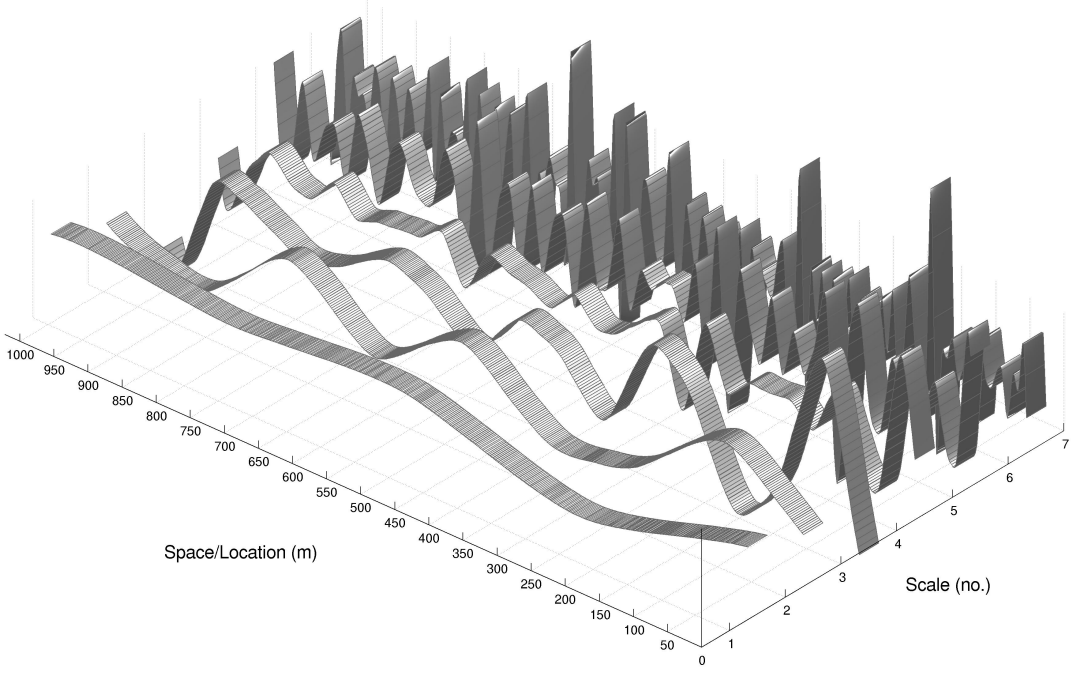

Figure 2.7: Ribbon plot showing the seven scales of the DWT decomposition of the total variable on Transect II.

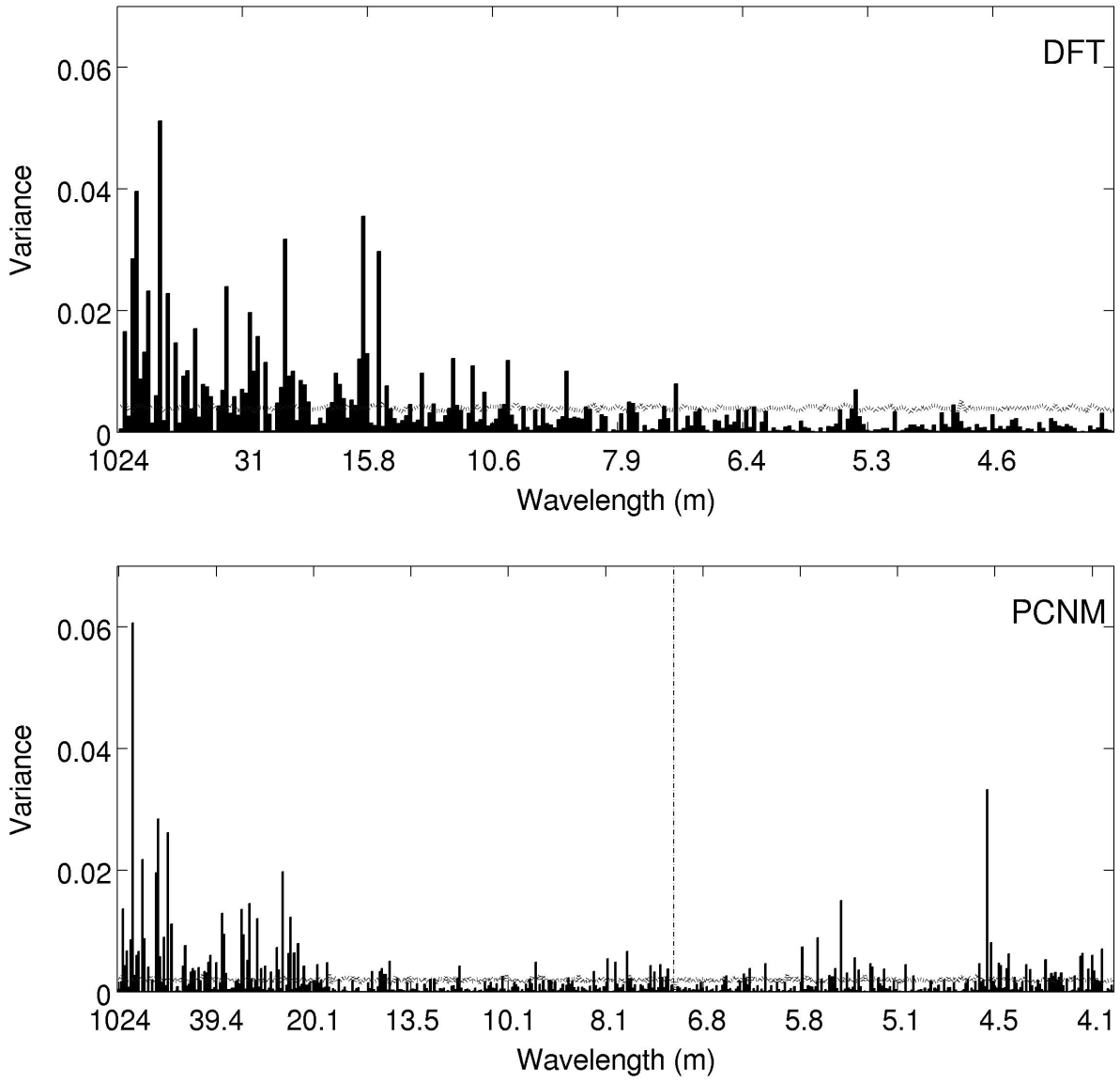

Figure 2.8: Periodogram for DFT and an equivalent representation for PCNM, showing variance as a function of wavelength for Transect I. The dashed vertical line in the lower graph represents the separation of positive (to the left) and negative (to the right) eigenvalues for PCNM. The horizontal lines represent the average periodograms computed over the 100 simulated random transects.

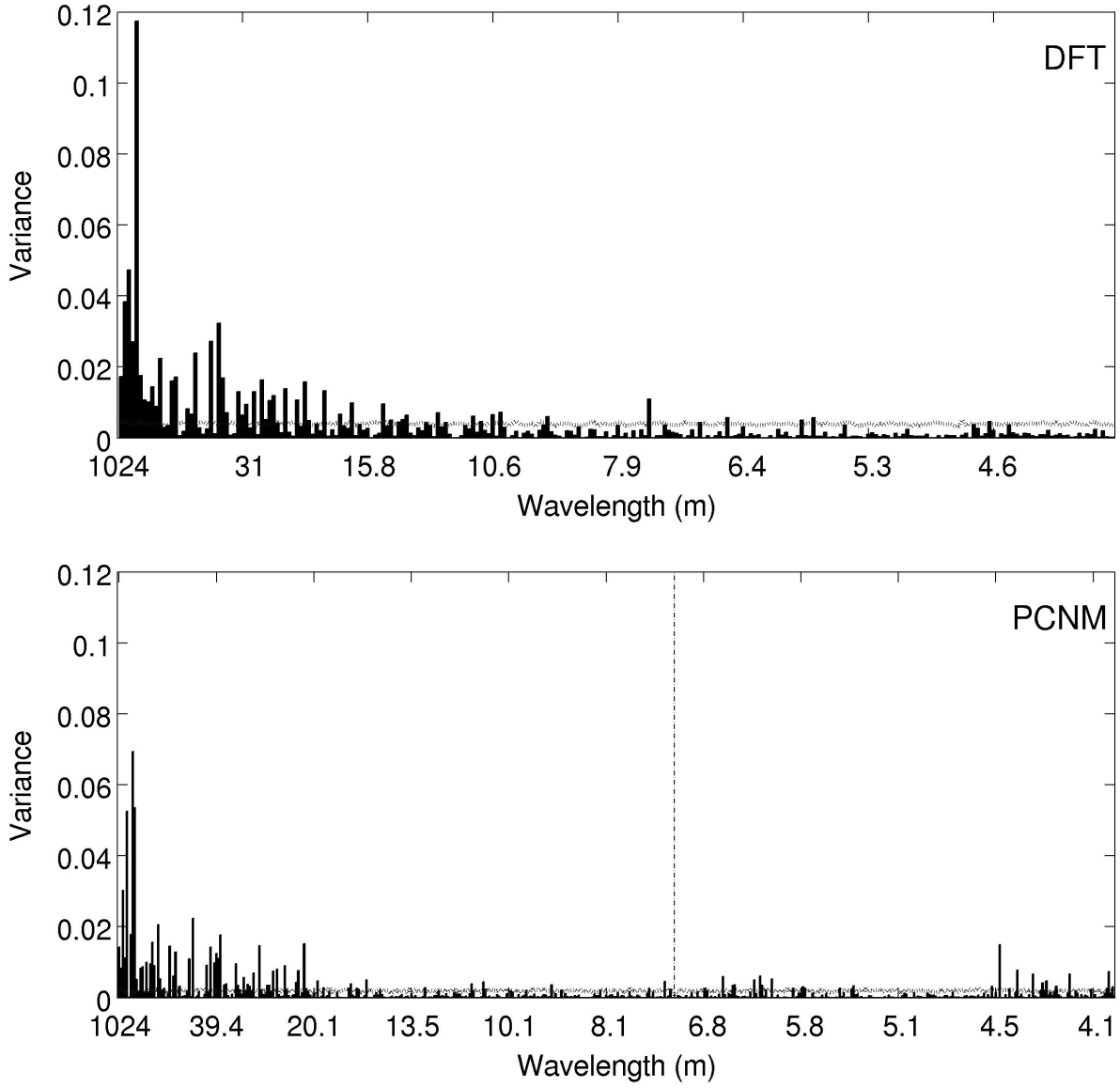

Figure 2.9: Periodogram for DFT and an equivalent representation for PCNM, showing variance as a function of wavelength for Transect II.

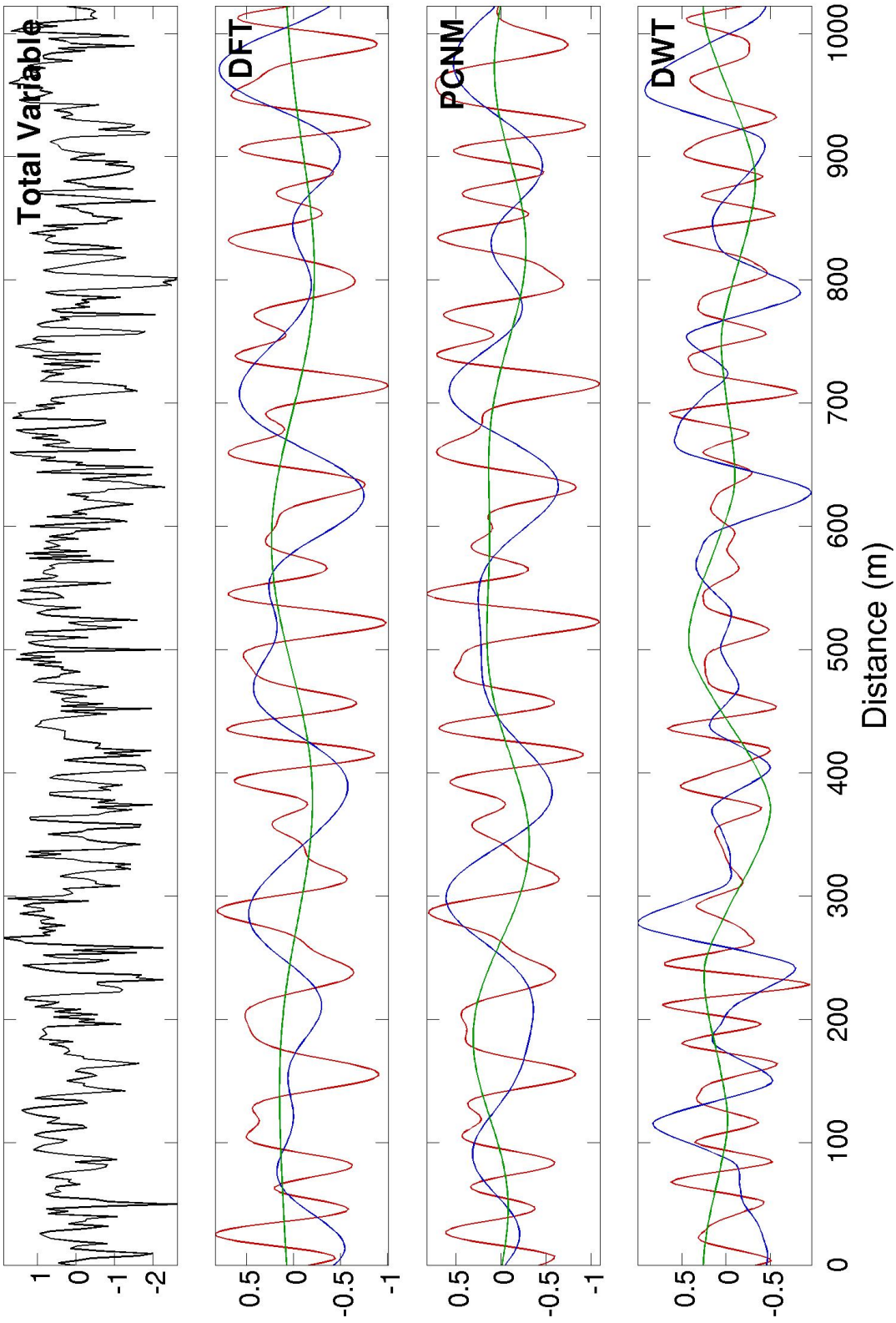

Figure 2.10: Decomposition of the total variable into three scale-specific spatial components corresponding to large scale (green line), medium scale (blue line) and small scale (red line) for DFT, PCNM and DWT performed on Transect I.

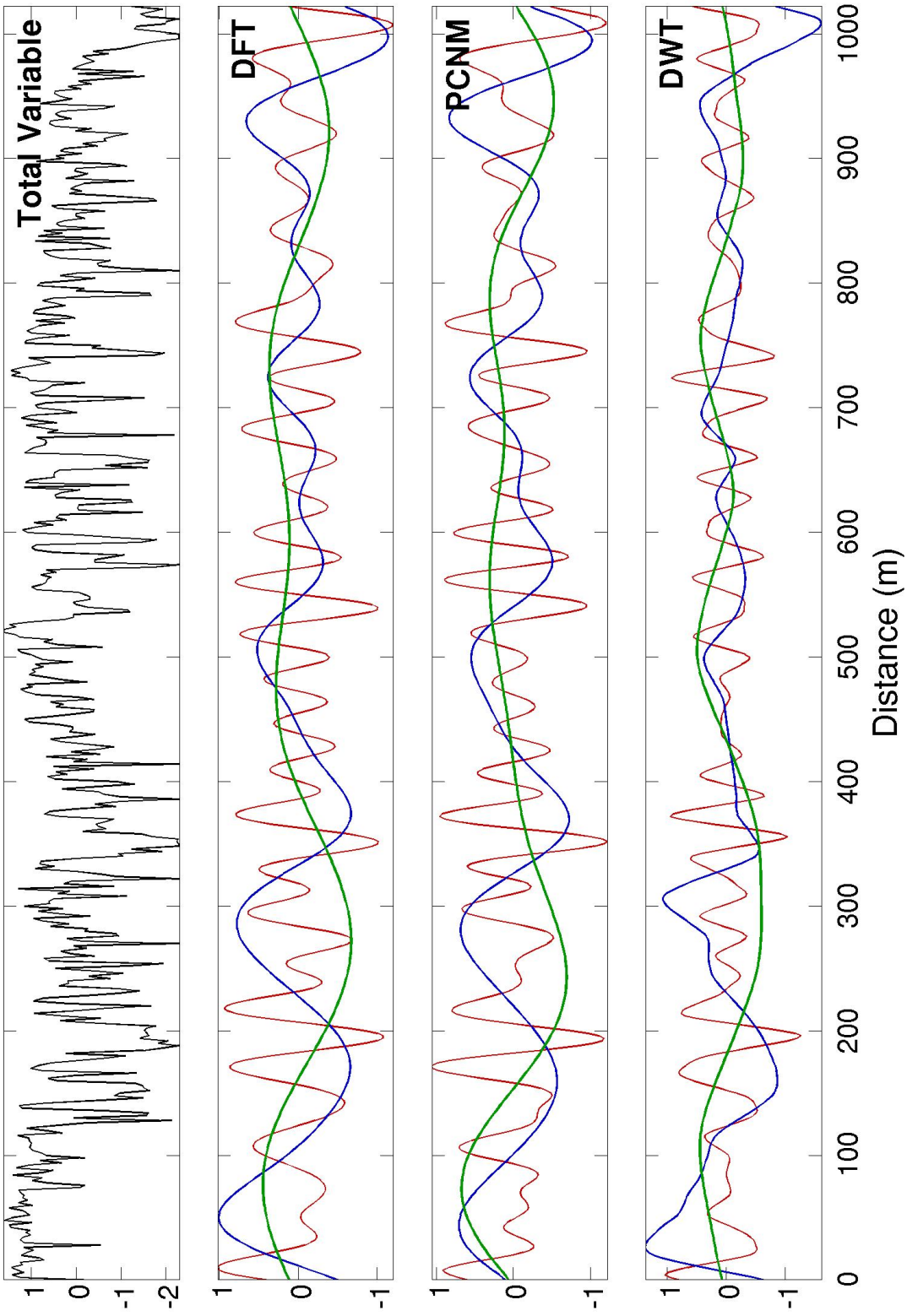

Figure 2.11: Decomposition of the total variable into three scale-specific spatial components corresponding to large scale (green line), medium scale (blue line) and small scale (red line) for DFT, PCNM and DWT performed on Transect II.

2.4.5 Comparison of orthogonal decomposition methods

A general scheme emerges from the description of the orthogonal decomposition methods. A global adjustment of the phase, or a local adjustment of the amplitude of the spatial functions, comes at the cost of a smaller number of orthogonal levels available for the decomposition. This problem is partly related to the Heisenberg uncertainty principle, which, if adapted to our context, states that a better description of the local features (e.g. with a local adjustment) implies a worst description of the scale (fewer levels), and vice-versa (Hubbard, 1998, p. 54). When the objective is to identify precisely scales of variation, or periodicities, fewer levels in the decomposition implies that spatial features may be present in the dataset at scales that are in between the scales of the decomposition, and can be described improperly or incompletely by the analysis. A lack of phase adjustment implies that spatial features with the same size and shape as the features depicted by the spatial components, but with a different phase, can be missed partially, or even completely. When the amplitude is adjusted globally, repeated spatial features of the same shape and size as the spatial function, but with varying amplitude, may be improperly described. PCNM provides the greatest number of levels, but a lack of phase and local amplitude adjustment (as is apparent in Fig. 2.6, all waves start at the beginning of the transect). The DWT allows an adjustment of the local amplitude, but supplies fewer levels in the decomposition (Fig. 2.7). Note that none of the orthogonal decomposition methods adjusts both the phase and the local amplitude.

The mean periodogram, and equivalent representation for PCNM, over the 100 simulated random transects consists of a straight line at a low variance value (Figs. 2.8 and 2.9). This does not necessarily imply, however, that each random variable is described by such a periodogram. In fact the simulated random transects have periodograms containing peaks at randomly distributed wavelength values. By contrast, the periodograms for both transects have higher peaks at long wavelengths indicating that the total variables are spatially structured. There appears to be a larger amount of large scale variation (approximately 30 m and more) for Transect II than Transect I.

From Fig. 2.10 and 2.11, we can see however that despite the important differences between methods, the medium and large scale components are similar for all three methods, whereas the small scale components for PCNM and DFT are similar but noticeably different from that of DWT. From these decompositions, it is obvious that the shapes of scale-specific components is constrained by the shapes of spatial functions specific to each method (sinusoidal waves for PCNM and DFT, wavelet functions for DWT). Another choice of wavelet function in DWT would have provided a different

decomposition of the signal. When the actual spatial features differ from the nice rounded shapes of the sinusoidal waves or of the Coiflet wavelet, the goodness of fit between the two may be poor. Also, the small scale components seem to contain some extraneous variation that does not coincide with the spatial features found on the transects. This problem is less apparent for DWT due to the local amplitude adjustment.

For all three methods, the choice of what constitute small scale spatial and non-spatial variation remains subjective and troublesome. Note that for this particular dataset, the use of contiguous neighbourhoods for obtaining the averages representing values along the transect alleviated problems related to aliasing (Diggle, 1990). In spectral analysis, aliasing is the interference of spatial variation at scales smaller than that of the Nyquist frequency with variations at larger scales. If a strictly punctual support had been used, spatial variations at scales smaller than the sampling distance would have appeared as larger scale variation due to aliasing. This issue amplifies the difficulties related to identifying spatial from non-spatial variation with orthogonal decomposition methods.

In all orthogonal decomposition approaches, it is possible to use significance tests to identify scale-specific components that contribute more importantly to the variation in the total variables. This was suggested as being an important aspect of PCNM (Borcard et al., 2004). Results not shown indicate that for both PCNM and DFT, scale-specific components at larger scales are much more frequently found to be significant by permutation tests than components at small scales. Significance tests have been proposed as a means to obtain a parsimonious model (Borcard et al., 2004; Bellier et al., 2007). However, performing significance tests on each scale-specific components before adding them to form 'bands' (e.g. small scale, large scale, etc) may not make the model more parsimonious since each 'band' can be considered a (single) scale-specific component. The focus of a parsimonious model should perhaps be to keep the number of bands small.

2.5 Graphical methods

2.5.1 Blocked quadrat variance methods

Blocked quadrat variance methods were among the first to be proposed for the analysis of spatial patterns in ecology. The first method originally presented by Greig-Smith (Greig-Smith, 1952) was based on the blocking of a series of contiguous quadrats into successively larger blocks, followed by a nested analysis of variance at each block size.

The relatively strict requirements of this method and criticisms directed towards it have led to the development of modified blocking technique, such as the two-term local quadrat variance (TTLQV) and the three-term local quadrat variance (3TLQV), developed by Hill (1973). A more recent extensive discussion can be found in Dale (1999) and Ludwig and Goodall (1978), and a comparison with wavelets in Dale and Mah (1998). With TTLQV, sampling units are summed into larger and larger blocks, and the average squared difference between adjacent units is calculated for each block size. This is then used to express variance as a function of block size. Peaks in this plot are interpreted as dominant scales of the spatial pattern. The 3TLQV method is very similar, but its computation involves groups of three adjacent quadrats instead of two.

2.5.2 Lacunarity

Developed by Mandelbrot in the context of fractal geometry (Mandelbrot, 1983), the concept of lacunarity was first applied to binary data. It was later applied to ecological data and adapted for the analysis of quantitative variables (Plotnick et al., 1996; Larsen and Bliss, 1998; Dale, 2000; Saunders et al., 2005). Lacunarity at a given scale measures translational invariance, or how similar different regions of a geometric object are to each other, with high lacunarity indicating higher dissimilarity between regions. As in blocked quadrat variance methods, sampling units are summed into successively larger block sizes. For each block size, lacunarity is calculated as a ratio involving the variance of the summed units and the square of their mean (Saunders et al., 2005):

$$\Lambda(s) = \frac{1 + variance(w_s)}{mean(w_s)^2}$$

where w_s denotes the set of sums of the sampling units located within blocks of size s. Results obtained from this 'gliding box' approach can suffer from edge effects, especially as the block size increases (Feagin et al., 2007). In the context of multi-scale analysis, as lacunarity generally decreases with increasing block size, dominant scales in the data are inferred from inflection points, or abrupt changes in the relationship between block size and lacunarity displayed on a log-log scale.

2.5.3 Continuous wavelet transform

The Continuous Wavelet Transform (CWT) combines elements of spatial functions and graphical methods. The principle is essentially the same as that of the DWT: a dilated and translated wavelet function is adjusted to the data to yield scale-specific

components (Torrence and Compo, 1998). However, CWT does not aim to perform a decomposition of the signal into orthogonal components. Instead, an arbitrary (theoretically unlimited) number of scales is chosen for the decomposition. At each scale, a given wavelet function centered on each sampling location is adjusted over the data sequence. Because the wavelets are overlapping and the scales are not orthogonal, the CWT therefore yields a lot of redundant information and the interpretation of results in CWT is different than in DWT. A typical output of CWT is a wavelet spectrum in which the variance captured by wavelets is expressed as a function of position and scale. The wavelet spectrum is the continuous equivalent to the ribbon plots shown above. If one sums the square of the wavelet coefficients at a given scale, a wavelet variance diagram, or global wavelet spectrum, is obtained. An advantage of CWT over other graphical methods described above is that there is a direct correspondence between values on the x-axis and scales of variability. The cross-wavelet transform calculated from the products of the CWT coefficients for two variables can be used for multivariate analysis (Torrence and Compo, 1998; Grinsted et al., 2004).

2.5.4 Experimental variograms

The experimental semi-variogram (variogram) is closely related to a variety of methods that rely on the squared differences of cross-products of samples as a function of distance (e.g. Geary's c and Moran's I correlograms). In the univariate case, the direct variogram represents graphically the semi-variance statistic $\tilde{\gamma}(h)$, defined as the average squared difference between values at all possible pairs of sampling units a given distance h apart, as a function of distance. That is,

$$\tilde{\gamma}(h) = \frac{1}{2N(h)} \sum_{u=1}^{N} (z(u) - z(u+h))^2$$

where u is sampling location and $N(h)$ is the number of pairs at a distance lag h. If the variable is spatially structured, its values at locations close together will be more similar and the difference between these values will be smaller than values at locations far apart. In such cases, semi-variances are low at small distances and they increase with increasing distance. Values of $\tilde{\gamma}$ generally tend to become equally similar after a certain distance and the direct variogram attains a plateau. The larger are the spatial features in the data, the more distance there will be before this plateau is reached. If the size of spatial features approach or surpass the size of the sampling domain, a plateau will not be reached. For a non spatially structured variable, semivariance values will be almost constant at all lags and the variogram then approximates a straight horizontal

line.

Experimental variograms can also be used to perform bivariate and multivariate analyses. The cross variogram $\widetilde{\gamma}_{ij}(h)$, used to represent the spatial relationship between two variables (z_i, z_j), is calculated in a way similar to the direct variogram:

$$\tilde{\gamma}_{ij}(h) = \frac{1}{2N(h)} \sum_{u=1}^{N} (z_i(u) - z_i(u+h))(z_j(u) - z_j(u+h))$$

However, its values may be positive of negative, depending on the sign of the relationship betwen the variables as a function of distance. The codispersion coefficient defined as $\frac{\widetilde{\gamma}_{ij}(h)}{\sqrt{\widetilde{\gamma}_i(h)\widetilde{\gamma}_j(h)}}$ can be slightly easier to interpret since its values range from -1 to 1. Because the strength or the sign of the cross semi-variance or codispersion coefficient may change significantly with distance, these tools are often used to show whether the relationship between two variables changes as a function of scale (i.e. is scale-dependent).

For multivariate analysis, variance-covariance matrices may be calculated at each lag from all possible combinations of direct and cross variograms. These matrices can then be used to perform various distance-based ordination methods, such as principal components analysis. This approach was described in various contexts (Noy-Meir and Anderson, 1971; Ver Hoef and Glenn-Lewin, 1989; Goovaerts, 1992) and was recently formalized and popularized by Wagner (2003) into the framework of 'multi-scale ordination'. Wagner (2004) also presented a procedure to perform direct multi-scale ordination through canonical correspondence analysis. A generalization of the approach was proposed by Couteron and Ollier (2005) to allow an application with more varied types of ordination methods.

2.5.5 Analysis of transect data

Both TTLQV and 3TLQV were applied to Transects I, II, and the 100 simulated random transects (Figs. 2.12 and 2.13) for block size (side lengths) ranging from the smallest distance between points on the transect (2 m) to a fourth of the transect length (256 m). The average variance over the 100 random transects is shown in the two figures. For Transect I, a dominant peak can be seen at approximately 115 m on the TTLQV graph, while the 3TLQV shows a large peak at 125 m with smaller peaks at 25, 185 and 235 m. For Transect II, a dominant peak at 100 m with a significant increase at larger block sizes based on the TTLQV analysis and a dominant peak at 100 m with smaller peaks at 175 m and a sharp increase at large block sizes based on the 3TLQV approach.

Lacunarity analysis was performed on the data of Transects I and II, as well as

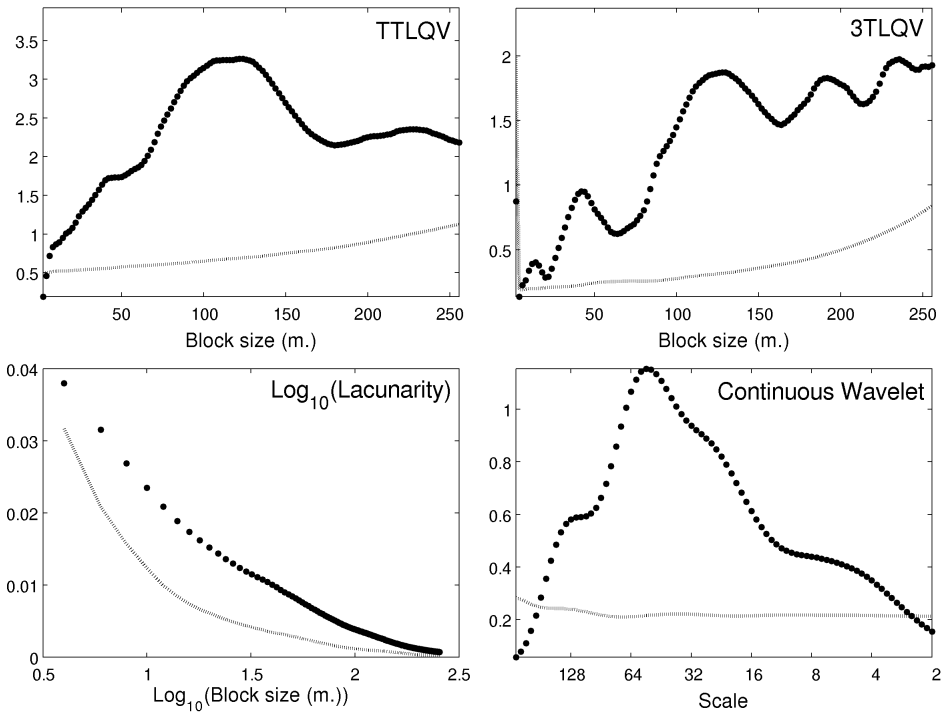

Figure 2.12: Graphical representations of variance as a function of observation scale for the graphical methods for Transect I. The dashed line represents the mean statistic over the 100 simulated random transects.

the 100 simulated random transects, for block sizes ranging from 2 to 256 m. As the method only works properly with positive numbers, the data were transformed to be in the [0-1] range prior to analysis. The lacunarity graphs show a generally decreasing pattern for both transects (Figs. 2.12 and 2.13). Lacunarity is higher for Transect II than Transect I and is consistently higher for both transects than for the mean over the random transects. This indicates a greater amout of aggregated patches of various sizes (more spatial structure) for Transect II. Inflection points are not easily discernable for either transects, making the identification of dominant scales of variation difficult.

The Matlab wavelet toolbox (The Mathworks, 2008) was used to perform CWT on the Lidar dataset. The Mexican hat (Sombrero) wavelet was used to separate the data into 84 scales. The local wavelet spectra are represented in three-dimensional form for Transects I and II (Figs. 2.14 and 2.15), while the global wavelet spectra are in shown in Figs. 2.12 and 2.13. The global spectrum for Transect I shows dominant scales of variation at approximately 50 m with less large scale and small scale variation. Transect II has more variation at large scale with a peak at 55 m. The local wavelet spectra indicate that fluctuations on Transect I are, in general, more regular than those on Transect II, which are more localized.

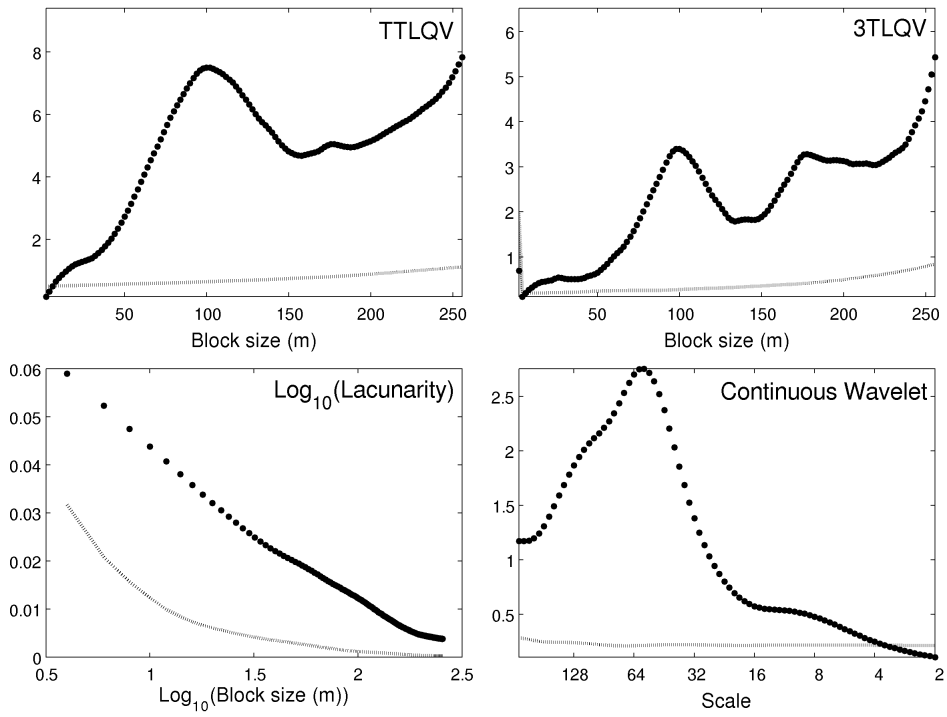

Figure 2.13: Graphical representations of variance as a function of observation scale for the various graphical methods for Transect II.

Experimental direct variograms were computed for distance classes ranging from 10 to 250 m with 10 m spacing every with extra lags at 2.5, 5 and 7.5 m to measure semi-variances at short distances (Fig. 2.16). The experimental variograms for Transect I shows smaller semi-variances at short distances (<10 m), with a slight gradual increase thereafter. A similar pattern is present on the semi-variogram for Transect II with even smaller semi-variances at short distances (<15 m) and a more pronounced increase at larger distances. This would indicate that small scale spatial features are present in both transects, but the amount of large scale variation is greater on Transect II. As expected, the mean variogram for the simulated random transects consists of a straight line ($\tilde{\gamma} = 1$) indicating no spatial variation.

2.5.6 Comparison of graphical methods

Except for CWT, the x-axis in the graphs showing variance as a function of 'scale' (Figs 2.12 and 2.13) cannot be directly associated with a scale of variation. A variance value at a given 'scale' in the graphical methods results from the combined effect of spatial features of all sizes located throughout the site. Conversely, spatial features of a given size (or scale) present on the site will not affect values at a unique distance, but at all

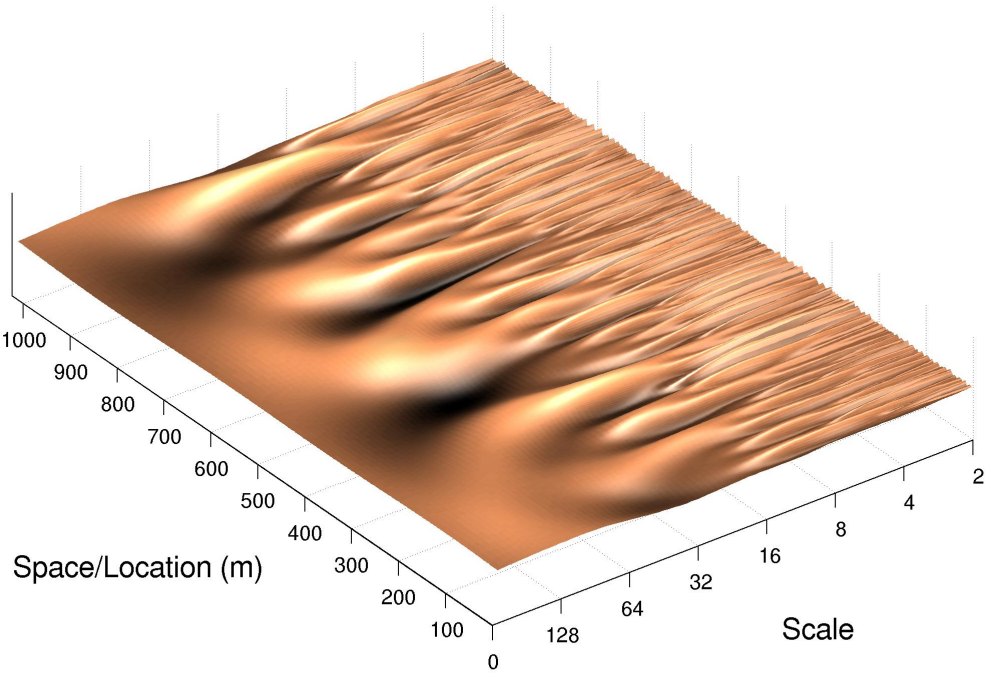

Figure 2.14: Three-dimensional representation of the CWT local wavelet power spectrum for Transect I.

distances. This is partly discussed for the experimental variogram in Goovaerts (1992) and Wackernagel (2003). To see the implications of this, imagine a data sequence made of one large bump followed by a series of small bumps. Small and large bumps will affect the variance at each distance or block size. Thus, graphs calculated for this transect using TTLQV, 3TLQV, lacunarity and the experimental variogram will not be able to identify the lack of medium scale spatial features. In fact, their interpretation might even lead one to conclude that medium scale spatial features are dominant on the site.

The analysis of the LIDAR data revealed that the interpretation of results from blocked quadrat variance methods and lacunarity can be somewhat difficult. For a trained eye, these graphical representations may give useful information about the spatial structure of the variables, but they seem to be unstatisfactory tools for analyzing patterns of variation at multiple scales.

The CWT offers the advantages of providing a quantification of variance as a function of scale of variation, which is easily interpretable, and as a function of location, making

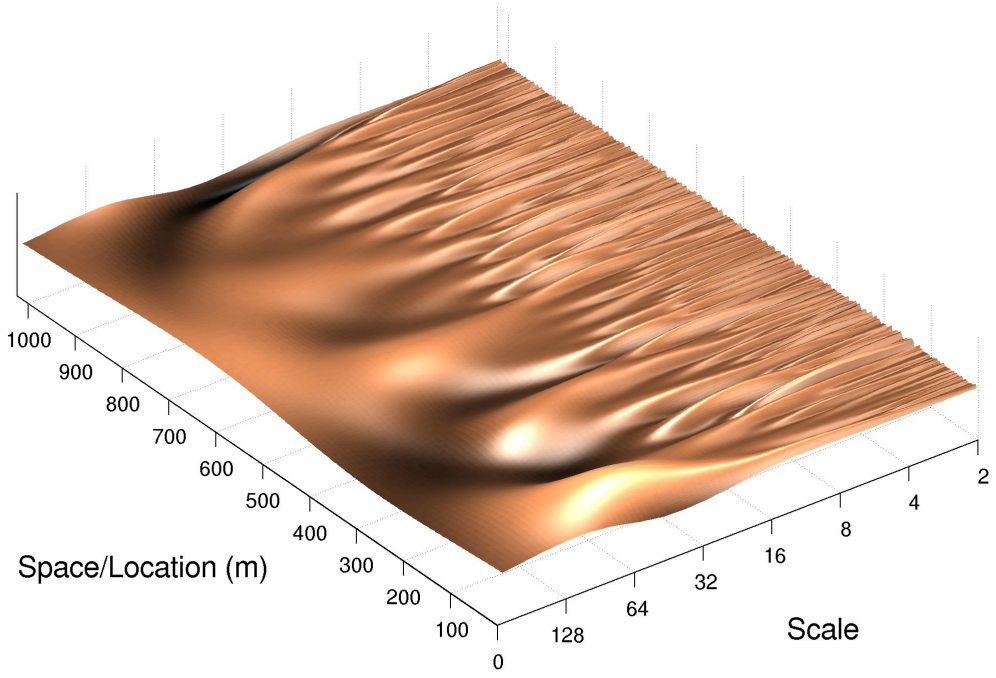

Figure 2.15: Three-dimensional representation of the CWT local wavelet power spectrum for Transect II.

it suitable for transects that show irregularities. However, this method remains limited to one-dimensional and regularly spaced datasets. While other graphical methods can be applied to two-dimensional grids, only the experimental variogram can be easily applied to irregular datasets in one, two or three dimensions.

2.6 Regionalized multivariate analysis

Matheron (1982) identified the difficulty of applying methods of spectral analysis with irregular grids in more than one dimension. He proposed a geostatistical alternative for estimating scale-specific components for single variables and 'factors' describing variation at specific scales shared between variables. From the French 'analyse krigeante', the approach was translated to 'factorial kriging analysis' (Goovaerts, 1992). It was later popularized and incorporated into the larger framework of regionalized multivariate analysis (Wackernagel et al., 1989; Goulard, 1989). In the terminology used

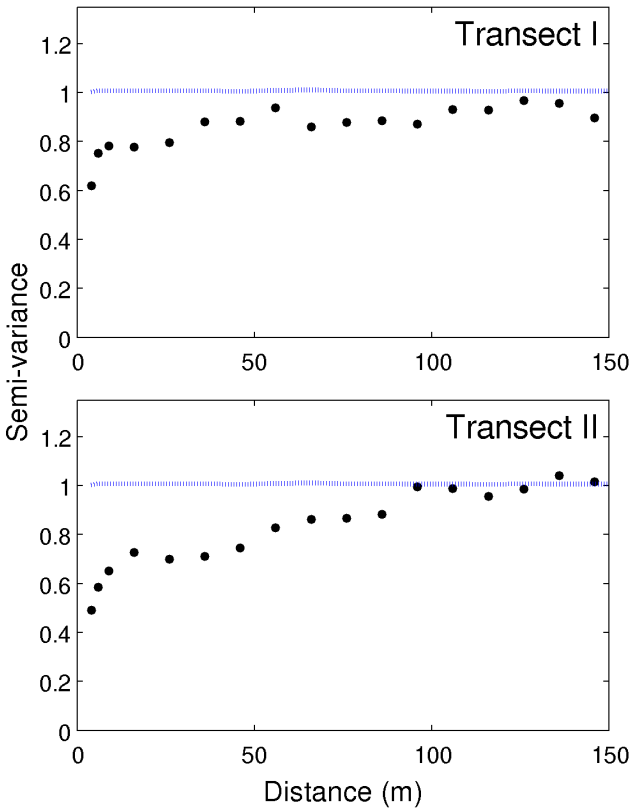

Figure 2.16: Experimental direct variograms for Transects I and II. The blue horizontal line is the mean variogram computed over the 100 random transects.

here and in the subsequent chapters, regionalized multivariate analysis is comprised of two methods: coregionalization analysis (CRA) and cokriging of regionalized components. The former is needed to perform the latter. The approach has been applied, for example, in soil science (Goovaerts, 1994; Webster et al., 1994; Bocchi et al., 2000; Castrignano et al., 2000a; Lin, 2002; Bourennane et al., 2003; Lark and Papritz, 2003), hydrology (Goovaerts, 1993) and remote sensing (Van Meirvenne and Goovaerts, 2002), but applications in ecology remain limited.

2.6.1 Coregionalization analysis

Coregionalization analysis comprises a set of geostatistical techniques for multi-scale analysis based on the linear model of (co)regionalization (LMR, LMC). In the LMR, a total random function is described as resulting from the superimposition of independent stationary random functions ($Z = \sum_{s=1}^{S} Z_s$), each characterized by different variogram functions,

$$\gamma(h) = \sum_{s=1}^{S} \gamma_s(h) = \sum_{s=1}^{S} b_s g_s(h, \varphi_s)$$

where b_s is the sill associated with the basic variogram function g_s characterized by a model and range φ_s. The range corresponds to the value of distance beyond which the values from the realizations are expected to be equally dissimilar, and the sill corresponds to the expected value of the squared difference between such values. For any stationary random function, the variogram function is a mirror image of the autocovariance function. Not all continuous functions expressing covariance as a function of distance can correspond to mathematically possible spatial structures, a particular class of 'allowed' autocovariance functions is generally used, including the spherical, exponential and Gaussian models. Variogram functions start at a value of zero at the zero distance, and attain a fixed semi-variance value (sill) at the distance corresponding to the range. An exception is the nugget effect, used to define non-spatially distributed RF, which has a variogram consisting of a straight line (a constant), with no range value.

In the linear model of coregionalization (LMC, Journel and Huijbregts, 1978; Wackernagel, 2003), the same variogram functions, with same models and ranges, are associated with all possible combinations of direct and cross variograms giving rise to sill matrices, also called coregionalization matrices, ($\mathbf{B}_s$) specific to each variogram function.

$$\mathbf{\Gamma}(h) = \sum_{s=1}^{S} \mathbf{\Gamma}_s(h) = \sum_{s=1}^{S} \mathbf{B}_s g_s(h, \varphi_s)$$

In practical applications, experimental variograms of the sampled total variables are used as estimators of the variograms of the multivariate total random functions. Fitting the same set of basic variogram functions (the LMC) to all possible combinations of direct and cross experimental variograms yields a set of sill matrix estimates $\hat{B}_s$:

$$\hat{\mathbf{\Gamma}}(h) = \sum_{s=1}^{S} \hat{\mathbf{\Gamma}}(h) = \sum_{s=1}^{S} \hat{\mathbf{B}}_s g_s(h, \varphi_s)$$

The sill matrix estimates $\hat{\mathbf{B}}_s$ can then be treated as variance-covariance matrices and be used in various methods of multivariate analysis (e.g. correlations, principal components analysis, redundancy analysis). The analysis of coregionalization matrices referred here as coregionalization analysis (Wackernagel, 2003) is often misleadingly referred to as factorial kriging analysis even when kriging is not performed (Goovaerts, 1992; Bocchi et al., 2000; Castrignano et al., 2000a).

Table 2.1: Range and sill estimates resulting from the LMR fitted to experimental variograms for transect I and II.

	Nugget	**Spherical**	**Gaussian**
Ranges			
Transect I	-	20 m	160 m
Transect II	-	20 m	170 m
Sills			
Transect I	0.61	0.22	0.08
Transect II	0.41	0.31	0.4

For the LMR and LMC to make sense in practical applications, it must be reasonable to treat the total variables as resulting from superimposed first and second-order stationary random functions. This implies that the mean and variance of the total variables should be reasonably similar throughout the study site. As we will see in subsequent Chapters, the stationarity assumption does not imply that practitioners must 'believe' that the studied variables are the result of random functions, it is merely a requirement for the interpretation of results and application of the method to be sensible.

A LMR comprised of a nugget effect, a spherical model and a Gaussian model was fitted to direct experimental variograms for Transects I and II. These models were chosen to properly capture the pattern found in the experimental variograms for both transects. Variogram ranges were estimated by least-squares (Pelletier et al., 2009b) and sills were estimated by the use of the Generalized Least-squares algorithm described in Pelletier et al. (2004) (Fig. 2.17, Table 2.1). Range estimates indicate that similar scales of spatial variation are present in both transects. However, sill estimates for the spherical and Gaussian models are higher for Transect II than for Transect I, indicating stronger spatial structure as well as a greater variation at larger scales for the former. The larger nugget effect for Transect I indicates more micro scale spatial and non-spatial variation.

2.6.2 Cokriging of regionalized components

In classical geostatistical applications, the modelling of the experimental variograms is rarely the ultimate objective of the analysis. It is often a preliminary step before the

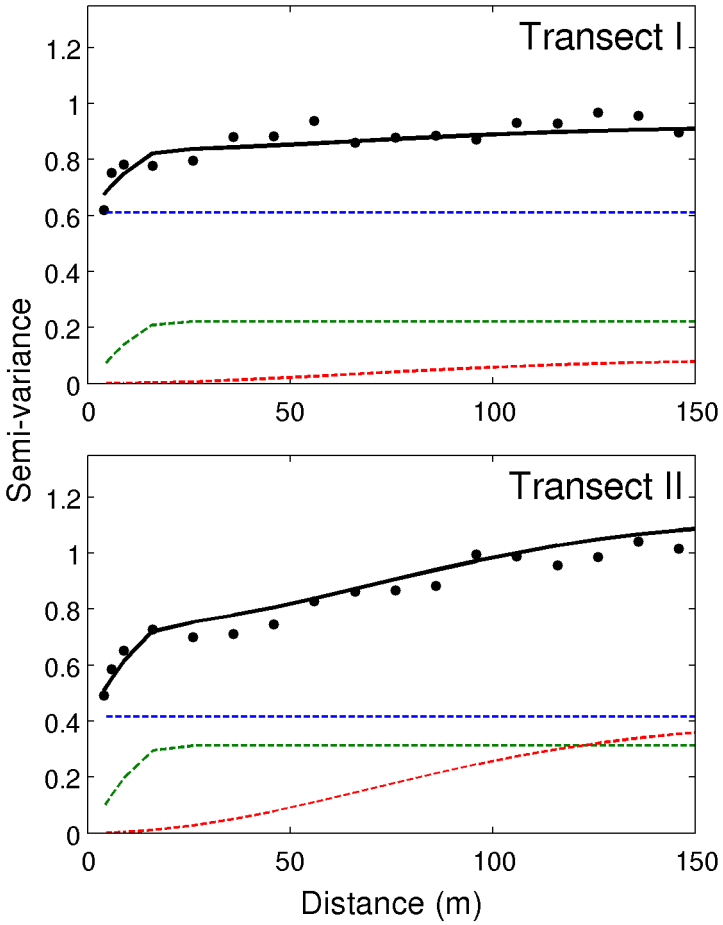

Figure 2.17: Linear model of regionalization fitted to direct experimental variograms for Transects I and II. The nugget effect model (dashed blue line), the spherical model (green line), the Gaussian model (red line), and the sum of the three models (solid black line) are shown.

estimation of values at unsampled locations or the estimation of spatial means. This is carried with methods of kriging, which are geostatistical procedures of interpolation that use a form of weighted local averaging, in which weights are obtained to minimize the estimation error and to provide an unbiased estimator (Isaaks and Srivastava, 1989). The standard kriging equations make use of the summed direct variogram functions obtained from the fiited LMR, to provide estimates of the total variable. In cokriging, the auto and cross variogram functions of the LMC are used in a system of equations to obtain weights for all the total variables in order to improve the estimation of each variable.

Matheron (1982) proposed a modification of the cokriging system to obtain estimates of the spatial components, or regionalized components, from the LMR and LMC for

each variable at both sampled and unsampled locations. The method also allows the estimation of regionalized factors that are used to represent spatially the variation shared by groups of variables at the scales corresponding to the variogram functions. It was then adapted and popularized by Goovaerts (1992), Wackernagel (2003) and others. The kriging system used for regionalized components and factors is different from the standard kriging system in that it uses the total variable to 'condition' the estimation of spatial components at each scale (equations are presented in Chapter 3). While the standard kriging system provides an exact estimator (the values estimated at the sampling locations coincide with the observed values of the total variable) regionalized components are not observed directly, and therefore cannot be estimated exactly. Nevertheless, the sum of regionalized component estimates at the sampling gives back the observed value of the total variable. Since they are scale-specific components, regionalized components for each variable could be used to perform statistical analyses in place of coregionalization matrices.

For the analysis of LIDAR data, the same variogram functions, sills and ranges obtained above in the LMR were used in a program written in Matlab to perform cokriging of regionalized components (Fig. 2.18). The ranges of Gaussian models (160 and 170 m) were used to define the radii of local neighbourhoods for the estimation.

2.6.3 Discussion of regionalized multivariate analysis

2.6.3.1 Scales of the variogram and autocovariance functions

It is common practice in applications of regionalized multivariate analysis to use the variogram range as a descriptor of the scales of variation found in the studied spatial variables. To assess the relevance of this practice, partial realizations from three variogram functions (Gaussian, spherical, exponential) with a unit sill and two different ranges (20 and 40 units) were generated on a transect with the same random seed (Fig. 2.19). Clearly, different types of variogram functions with the same sill and range give rise to partial realizations with significantly different spatial structures. For a given range value, the partial realization of the Gaussian model is much smoother than either realization of the spherical or exponential model, which both contain significantly more small scale variability. It is interesting to note that the difference between realizations drawn from the same variogram function but different ranges (20 vs. 40 units) varies a lot with the type of function. The difference is more pronounced for the Gaussian model, and almost unnoticeable for the exponential function, the spherical function falling in-between.

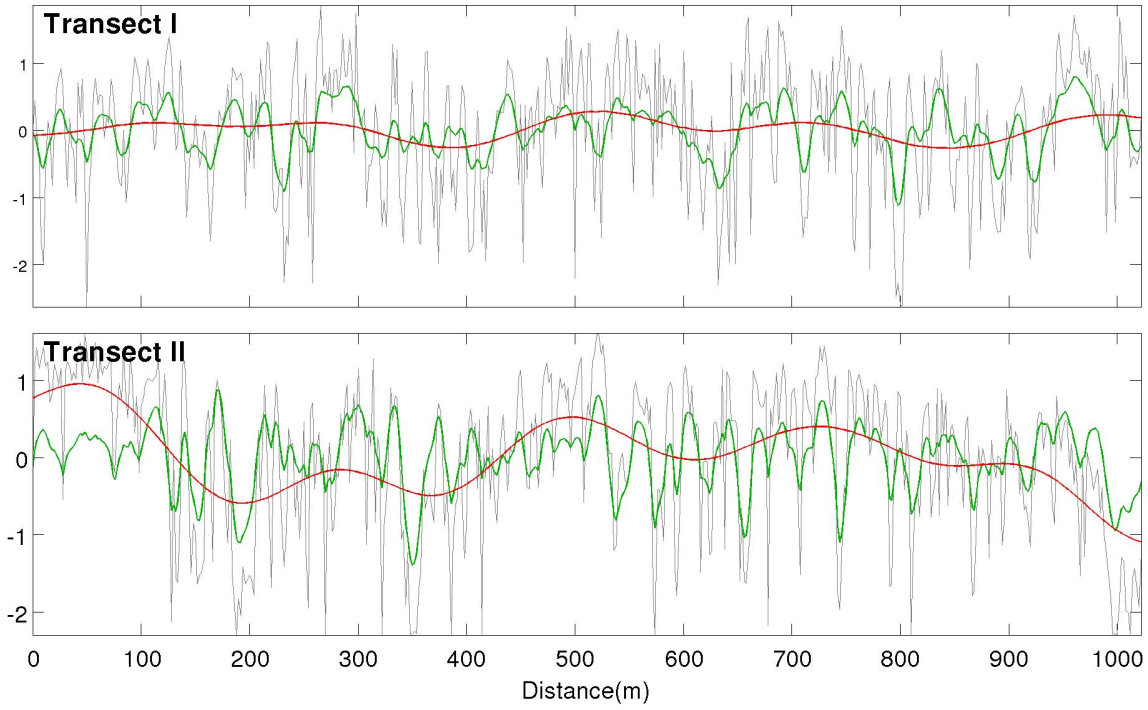

Figure 2.18: Cokriged regionalized components corresponding to the spherical model (green line) and the Gaussian model (red line) for Transects I and II.

From this example, it is very clear that, to describe scales of variation of random functions, 1) the variogram range is inappropriate, if used alone, and 2) the type of variogram model is extremely important. Simulated realizations can help us perform a qualitative classification of which models possess more small or large scale variability, but how can we express quantitatively the scales of variation corresponding to different models?

To help us achieve this, we can turn to the probabilistic methods of spectral analysis. Any random function can be seen as the superimposition of an infinite number of sinusoidal waves, each characterized by a specific frequency, or wavelength (Priestley, 1981; Diggle, 1990; Denny et al., 2004). As we have seen with DFT, although each wave is clearly periodic, realizations derived from their superimposition do not need to be periodic. In fact, combinations of sinusoidal waves can be used to give rise to any spatially structured patterns. Under the stationarity assumption, random functions can be defined equivalently from their autocovariance function or their power spectrum (spectral density function, density spectrum), which describes the energy, or variance, of a random function as a function of frequency. In fact, the autocovariance function is the cosine Fourier transform of the power spectrum. As frequencies range from infinitely

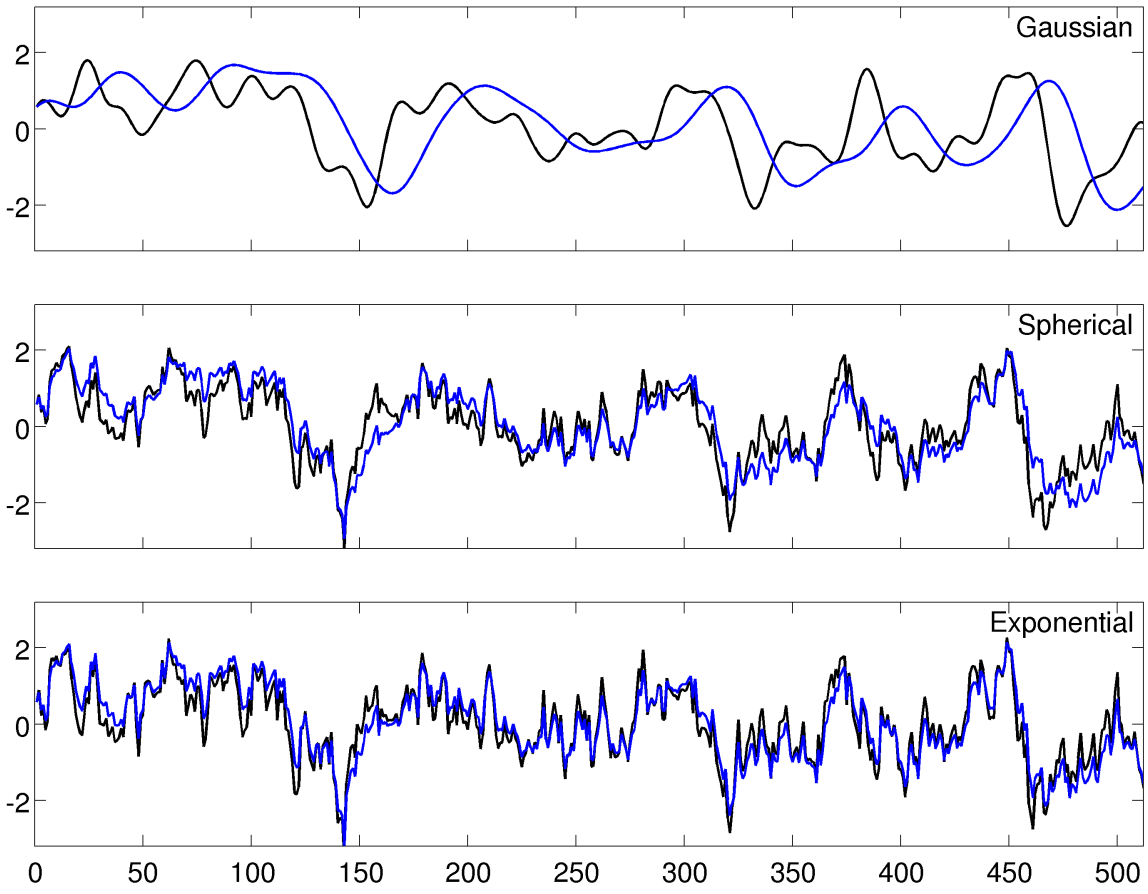

Figure 2.19: Comparison of realizations derived from a spherical, Gaussian and Exponential variogram functions with a unit sill and a range of 20 units (black line) and 40 units (blue line). All realizations were obtained with the same random seed.

small to infinitely large values on the power spectrum, wavelengths range from infinitely large to infinitely small values. Power spectra corresponding to Gaussian, spherical and exponential variogram functions are shown in Figure 2.20 (Equations can be found in Schabenberger and Gotway (2005), p.199). The advantage of the power spectrum over the autocovariance function is that sinusoidal waves can be easily interpreted in terms of scale of variation, through their wavelength.

However, the scales of variation are difficult to infer from the power spectrum due to the non-linearity of scale if wavelength is reported on the x-axis. For this reason, I have developed an expression to express the spectral density function as a linear function of wavelength (a 'wavelength spectrum'):

$$\zeta(\lambda) = \frac{1}{\lambda^2}\psi(\frac{1}{\lambda})$$

where $\psi(\frac{1}{\lambda})$ is the spectral density function expressed in terms of wavelength. Like the density spectrum, the area under $\zeta(\lambda)$ is equal to the semi-variogram sill. This graph, with the wavelength of sinusoidal functions expressed linearly on the abscissa, can be more easily interpreted and gives a more intuitive depiction of scales of variation. Using this representation with the same functions as above (Fig. 2.20), we can reiterate our earlier statement that the range of a variogram function is not a satisfactory measure of scales of variation. For a given range, Gaussian functions possess much fewer small scale spatial features than either the spherical or exponential model. For the Gaussian and spherical functions, the distribution peaks at approximately 0.3 times the value of the range, while the exponential function peaks close to zero. From each peak value, the variance corresponding to spatial features of larger sizes decreases, and eventually reaches zero asymptotically (for an infinitely large wavelength). Realizations from the three models thus possess spatial features much smaller and much larger than the range. The pattern of distribution is greatly dependent of the type of variogram function used.

This clearly indicates that linking variogram range and scale of variation warrants extreme caution. In practical applications of the LMR and LMC, the type of variogram models used should greatly affect the results of the analysis. If simulated realizations derived from known variogram functions contain variation at multiple scales, it implies that in practical applications, a given variogram function will capture variation in total variables over multiple scales also. We may then wonder whether it is sensible to use superimposed functions to capture variation in a total variable.

Several contributions have considered the basic premices of the LMC, that the spatial pattern in all total variables can be characterized by the same set of basic variogram functions, to be too rigid (Ver Hoef and Barry, 1998; Stein, 1999; Lark and Webster,

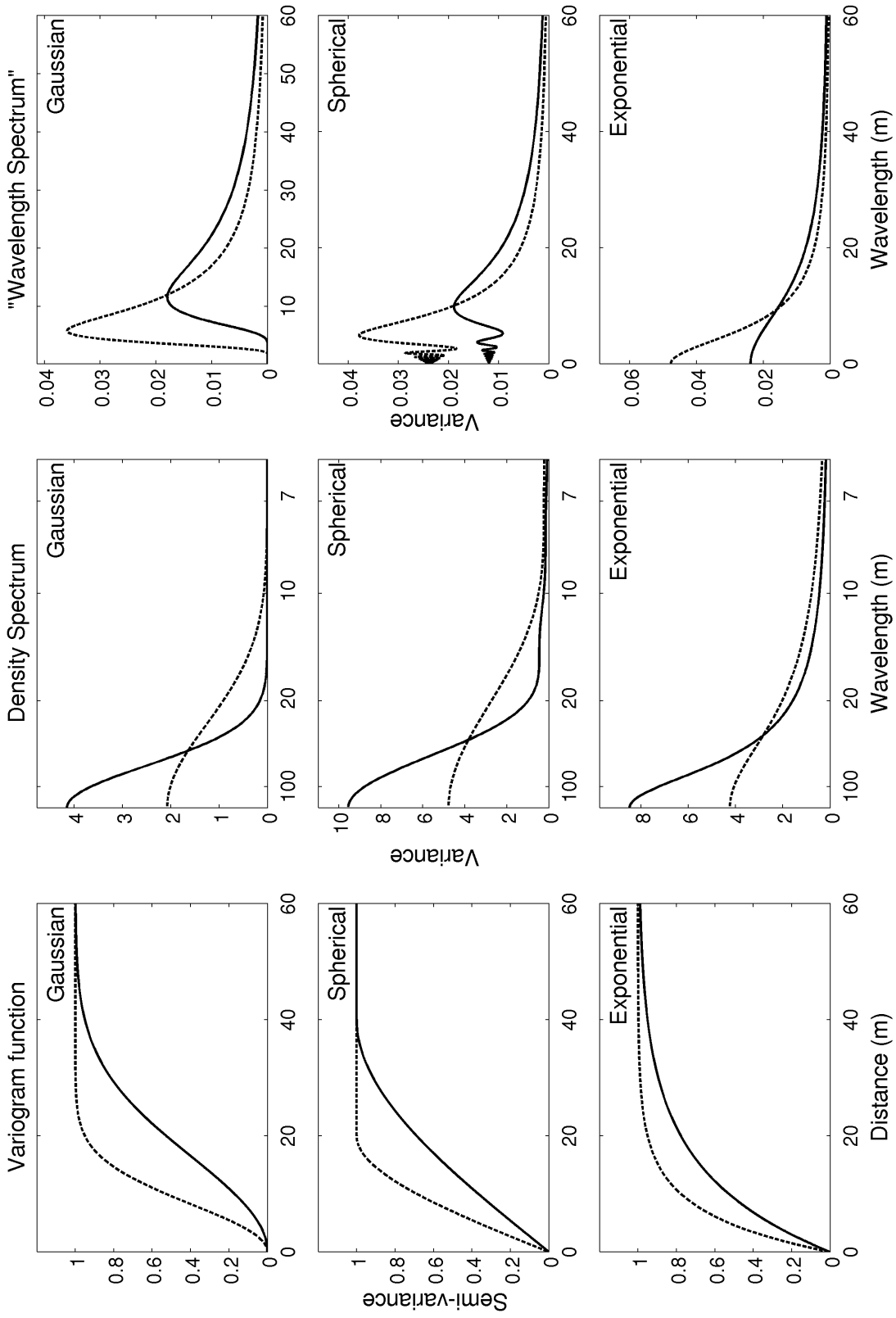

Figure 2.20: Variogram functions and the corresponding density spectrum and 'wavelength spectrum' for Gaussian, spherical and exponential models with a unit sill and a range of 20 (dashed line) and 40 (solid line) units.

2001). Those criticisms are certainly valid if the variogram functions are simply seen as means of capturing variation at the specific scale corresponding to their range. From the results presented here, the variogram functions in the LMC rather captures bands of scales. A small number of properly choosen variogram functions should then be able to capture the variation found in a wide variety of total variables. From this perspective, the assumptions of the LMC seem less stringent.

2.6.3.2 Use of regionalized component estimates

Being based on weighted averages, kriging resembles local drift estimation procedures described above. However, the scale of the spatial structuring found in the regionalized components is imposed by the autocovariance function defined in the LMC. This allows the different types of spatial structuring on Transects I and II to be nicely captured by the regionalized component estimates (Fig. 2.18), while still corresponding to specific scales of variation. It may then be appealing to use regionalized components to perform multivariate methods of analysis in place of the coregionalization matrices.

Before considering this option, however, one must understand the effects of a phenomenon known as the 'smoothing effect' of kriging in geostatistics. Due to this effect, the properties of kriging estimates never correspond exactly with the properties of the random function model, as the kriging system aims to minimize the error variance and to provide an unbiased estimate of the total variable. This system is not designed to reproduce the exact structure of the total variable as defined by the random function model. To obtain maps that look like realistic representations of the total variables under this model, conditional (co)simulation can be performed by adding to the kriged estimates a simulated random component that compensates for the smoothing effect (Davis, 1987; Alabert, 1987; Goovaerts, 1997, 2001). This way, several simulated maps can be obtained that share the pattern of kriging estimates while differing by the added simulated random deviation.

The kriging system for regionalized components is unique in that the components corresponding to each structure are all obtained from the same total variables. Thus, the smoothing effect in this context has the additional effect that regionalized component estimates for different structures are not orthogonal, although they are orthogonal in the random function model (Matheron, 1982). The analyses at the scales corresponding to the basic variogram functions may therefore provide redundant information about the structure of the observed total variables. The exact effects of this redundancy, along with the characteristics of the smoothing effect in factorial kriging have not been discussed in the literature. Also, these effects prevent the direct application of

the standard conditional (co)simulation algorithms for the simulation of regionalized components. Alternatives for the simulation of regionalization components have not yet been proposed.

2.7 Summary

An important point for the comparison of methods of multi-scale analysis are the possibilities and objectives of each method. Orthogonal decomposition methods focus on the separation of total variables into scale-specific components. This makes them suitable both for the detection of dominant scales of variation and for conducting separate and independent analyses on scale-specific components. Graphical methods provide means to detect dominant scales of variation but do not separate patterns into scale-specific components, making them unsuited for independent analyses at multiple scales. Regionalized multivariate analysis can be used both for the detection of dominant scales and the separation into scale-specific components.

The different methods of multi-scale analysis cannot all be used on the same type of dataset. In the way they have been described here, DFT, DWT, TTLQV, 3TLQV, lacunarity and CWT are restricted to either regular one-dimensional transects or regular two-dimensional grids. For some of them, such as DFT and DWT, modifications can be applied to allow their use on irregular grids, but these modifications further complicate the analysis and interpretation of results. Although PCNM can be directly used with irregular grids, results can be nonsensical on highly irregular grid patterns. Experimental variograms, regionalized multivariate analysis, DBEM, trend surface analysis and local drift estimation procedures are all directly applicable to regular and irregular transects and grids.

All the approaches discussed except DWT and CWT are global approaches that provide an overall quantification of dominant scales of variation in the dataset. They are therefore more applicable for datasets with relatively regular patterns throughout the study sites. The DWT and the CWT appear more suitable for datasets containing more local irregularities.

The interpretation of the results from several of the methods arrived at similar conclusions pertaining to the large scale structure of the two transects, with Transect II containing considerably more large scale variation. With the blocked quadrat variance methods and lacunarity, it was more difficult to infer scales of variation. The quantity and structure of the small scale variation differed more between the different approaches. This is likely due to the lack, in most approaches, of an objective procedure to identify

and quantify the non-spatial and spatial variation in the dataset. In regionalized multivariate analysis, the nugget effect modelling provides such a procedure, conferring a distinct advantage to this method over the other methods.

Overall, it seems that regionalized multivariate analysis offers several advantages over the other methods. It can be used with various regular or irregular datasets, it can be used to quantify the variation at multiple scales and to generate scale-specific components.

The comparative analysis of methods for multi-scale analysis presented in this chapter demontrated that each method must be used with a proper understanding of its assumptions, mathematics and capabilities. Each represents a different conception of spatial structure, spatial features and scale and will highlight different aspects of the patterns found in the dataset. The methods described here are thus not equivalent tools to perform multi-scale analysis, although some may prove to be complementary.

Linking paragraph

The following two chapters present co-authored manuscripts that were previously published in scientific journals. They cover the uncertainty associated with the two 'steps' involved in regionalized multivariate analysis: coregionalization analysis and cokriging of regionalized components. Chapter 3, published in Geoderma (Larocque et al., 2006), addresses the issue of spatial uncertainty of (co)kriged regionalized components, and presents a new procedure to perform the joint conditional Gaussian co-simulation of regionalized components. Note that throughout this chapter, the linear model of coregionalization will be considered appropriate and adequate. Chapter 4 was published in Mathematical Geology (now Mathematical Geosciences, Larocque et al., 2007) and provides an assessment of the limits and robustness of coregionalization analysis. Chapter 5 revisits the content of those chapters.

Larocque, G. Dutilleul, P., Pelletier, B, and J. W. Fyles. 2006. Conditional Gaussian co-simulation of regionalized components of soil variation. Geoderma 134(1), 1-16.

Larocque, G. Dutilleul, P., Pelletier, B, and J. W. Fyles. 2007. Characterization and quantification of uncertainty in coregionalization analysis. Mathematical Geology 39, 263-288.

Chapter 3

Conditional Gaussian co-simulation of regionalized components of soil variation

3.1 Abstract

Stochastic simulations are increasingly used to represent and characterize the spatial structure and uncertainty of soil properties. Due to the potential presence of scale dependencies, simulations of the total variables can represent a mixture of spatial components operating at different scales, which may be better interpreted separately. While coregionalization analysis and factorial kriging provide means to characterize and estimate scale-specific components of variation, no methods are available that allow a proper representation of their spatial structure and an assessment of their spatial uncertainty. In this paper, the formulation of cokriging of regionalized components and regionalized factors is first reviewed, after which a method for the conditional Gaussian co-simulation of regionalized components and regionalized factors is presented. We highlight the need for performing conditional simulations for all structures jointly to reduce the correlation between components for different structures and avoid any bias on the sum of simulated components. Simulations obtained with this method adequately represent both the specific features of, and the uncertainty associated with, each scale of variation, as modeled in a coregionalization analysis. The method is applied to an agronomic dataset to characterize the spatial uncertainty of regionalized components of plant available phosphorous and potassium in the soil and illustrate advantages of this new simulation approach.

Keywords: Factorial kriging, coregionalization analysis, spatial uncertainty, scale dependence, conditional stochastic simulations.

3.2 Introduction

Geostatistical methods are now an essential part of the toolbox of most pedometricians and of many soil scientists. While the variogram and kriging have been preferred tools to characterize the spatial distribution of soil variables and obtain optimal maps, the importance of uncertainty assessment is also increasingly recognized (see Goovaerts, 2001). Conditional stochastic simulations are among the methods used to assess the uncertainty associated with spatial estimates (i.e., the spatial uncertainty). Kriging and cokriging estimates are known to minimize the local error variance, but fail to reproduce histograms of original variables, modeled variograms or correlations between variables (Goovaerts, 2000). In contrast, the objective of conditional stochastic simulation is to generate realizations that reproduce those global statistics, conditional to the data values, and realistically represent the spatial structure of the variables. In the simulation approach, multiple realizations are needed for a complete assessment of the spatial uncertainty.

Until now, conditional simulation methods have aimed at reproducing characteristics of total variables, as determined from the variogram modelling. However, thanks to methods such as coregionalization analysis, it has become clear that spatial patterns of variables and relationships between variables are often scale-dependent because of different processes operating at different scales. Numerous applications of coregionalization analysis in soil science have confirmed the presence of such scale-dependencies (Goulard and Voltz, 1992; Goovaerts et al., 1993; Goovaerts and Webster, 1994; Webster et al., 1994; Dobermann et al., 1995, 1997; Bocchi et al., 2000; Castrignano et al., 2000a; Lin, 2002; Bourennane et al., 2003). Estimated or simulated maps of total variables can then represent a mixture of components operating at different scales, which may be better interpreted separately. With methods grouped under the heading 'factorial kriging', it is possible to obtain scale-specific spatial estimates for each variable, called regionalized components, and for components of variation shared by a group of variables, called regionalized factors (Goovaerts, 1992). Despite their ability to deal with scale dependencies, these methods do not provide estimates possessing characteristics of the fitted variogram models, such as the autocovariance structure and the structural correlations between variables.

It is also well known that kriging estimates are mostly affected by the fitting of the

variogram model at short distances (Isaaks and Srivastava, 1989, pp. 301-308). For this reason, it is often advised to obtain soil samples on a regular grid as well as on a number of randomly chosen locations at relatively short distances from points on the regular grid. Such a sampling design can provide reliable variogram estimates at short distances and ensure that the nugget effect estimated from the experimental variogram is predominantly comprised of measurement error, as opposed to micro-scale variation. In such cases, or whenever the nugget is thought to arise from measurement error, conditional co-simulations performed on the total variable inappropriately represent the true spatial structure of the variable. Note that Marcotte (1995) proposed a post-processing approach based on factorial kriging that removes the nugget effect from a given realization following the conditional simulation of the total variable, and improvements were suggested by Bourgault (1996).

To our knowledge, no published study has attempted to develop a mathematically rigorous framework for performing conditional simulations of regionalized components or regionalized factors. Such simulations would adequately represent both the specific features of, and the uncertainty associated with, each scale of variation. In this paper, the formulation of cokriging of regionalized components is reviewed prior to the presentation of a method for the conditional Gaussian (co-)simulation of regionalized components. The method is applied to an agronomic dataset to characterize the spatial uncertainty of regionalized components of plant available phosphorous and potassium in the soil. A corresponding method for the conditional Gaussian co-simulation of regionalized factors is also presented.

3.3 Cokriging of regionalized components

Coregionalization analysis is a necessary step before regionalized components can be estimated by cokriging (Goovaerts, 1992, 1997; Wackernagel, 2003). It is entirely based on variogram modelling through the linear model of regionalization (LMR, univariate case) or linear model of coregionalization (LMC, multivariate case). In the LMC, the same number and types of basic variogram functions are fitted to all direct and cross experimental variograms computed for a number of variables (Journel and Huijbregts, 1978). Each function in the nested model corresponds to a scale-specific component of variation that is assumed orthogonal to components at other scales. The matrices of coefficients (i.e., sills) of those functions, called coregionalization matrices, are then used as variance-covariance matrices to perform multivariate analyses, such as principal component analysis.

Factorial kriging methods aim at estimating components corresponding to variogram functions in the LMR or LMC (original work by Matheron, 1982, or more accessibly, see Goovaerts, 1997). Two methods can be distinguished: (co)kriging of regionalized components and factorial cokriging. In the univariate case, both are identical and use a modified kriging system to estimate regionalized components corresponding to variogram functions of the LMR. In the multivariate case, cokriging of regionalized components shares this same objective but uses cross-covariances, as modeled in the LMC, in a cokriging system. In factorial cokriging, also called multivariate factorial kriging, coefficients resulting from the factorization of coregionalization matrices are used to estimate regionalized factors (Wackernagel, 2003). Below, we focus on regionalized components, but a description of factorial cokriging and conditional Gaussian co-simulation of regionalized factors is presented in the Appendix.

In what follows, it is assumed that coregionalization matrices corresponding to S (> 1) structural variogram models are known (in practice, they are estimated in the LMC) prior to estimating or simulating regionalized components for p variables, whose distribution is multi-Gaussian, under the second-order stationarity assumption. It will also be assumed that $p > 1$, although the development is valid for $p = 1$. When the regionalized components are estimated by simple cokriging, estimated values of a regionalized component on N_o nodes are expressed as linear combinations of the total variables observed on N_n sampling points:

$$\mathbf{z}_s^* = \mathbf{W}_s^T \mathbf{z} \tag{3.1}$$

where $\mathbf{z}$ is the $N_n p$ x 1 vector of 'total observations' and $\mathbf{W}_s$, the $N_n p$ x $N_o p$ matrix of simple cokriging weights, is obtained by solving

$$\mathbf{K}\mathbf{W}_s = \mathbf{k}_s \tag{3.2}$$

$$\text{with } \mathbf{K} = \begin{bmatrix} \mathbf{K}_{11} & \cdots & \mathbf{K}_{1p} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_{p1} & \cdots & \mathbf{K}_{pp} \end{bmatrix} \text{ and } \mathbf{k}_s = \begin{bmatrix} \mathbf{k}_{11,s} & \cdots & \mathbf{k}_{1p,s} \\ \vdots & \ddots & \vdots \\ \mathbf{k}_{p1,s} & \cdots & \mathbf{k}_{pp,s} \end{bmatrix},$$

where $\mathbf{K}_{jj'}$ (j, j'=1, ..., p) is the N_n x N_n matrix of total covariances between variables j and j', calculated from the sills $\beta_{jj',s}$ ($s = 1, \ldots, S$) and the $N_n \times N_n$ autocorrelation matrix $\boldsymbol{\rho}_{(n),s}$:

$$\mathbf{K}_{jj'} = \sum_{s=1}^{S} \beta_{jj',s} \boldsymbol{\rho}_{(n),s}; \tag{3.3}$$

$\mathbf{k}_{jj',s}$ is the N_n x N_o matrix of structural covariances between variable j at the sampling points and variable j' at the nodes, calculated from the sill $\beta_{jj',s}$ and the N_n x N_o matrix of autocorrelation $\boldsymbol{\rho}_{(n,o),s}$ at structure s:

$$\mathbf{k}_{jj',s} = \beta_{jj',s}\boldsymbol{\rho}_{(n,o),s}\,(s = 1, \ldots, S) \tag{3.4}$$

The cokriging system in Eq. (3.2) aims at locally minimizing the variance of the difference between estimated and true regionalized components, so the cokriging estimator is the best linear unbiased estimator (BLUE). This property results in a well-known smoothing effect for the cokriging estimator (Goovaerts, 1997), which has two important consequences. First, the estimated regionalized components do not have the same autocovariance structure as that modeled in the LMC (Marcotte, 1995). Second, the structural correlations between variables calculated from estimated regionalized components can be biased with respect to the ones calculated from coregionalization matrix estimates. In fact, the variance-covariance matrix of the regionalized components at structure s, as modeled in the LMC, is

$$\mathrm{var}(\mathbf{z}_s) = \mathbf{K}_{(o),s}, \tag{3.5}$$

where $\mathbf{K}_{(o),s}$ is the $N_o p$ x $N_o p$ matrix of node-to-node covariances calculated from the Kronecker product ($\otimes$) of the p x p coregionalization matrix $\mathbf{B}_s$ and the N_o x N_o node-to-node autocorrelation matrix $\boldsymbol{\rho}_{(o),\mathrm{s}}$:

$$\mathbf{K}_{(o),s} = \mathbf{B}_s \otimes \boldsymbol{\rho}_{(o),\mathrm{s}}. \tag{3.6}$$

Note that $\mathbf{K}_{(o),s}$=$\mathbf{k}_s$ if cokriging is performed on the sampling points. From Eqs. (3.1) and (3.2), it follows that the variance-covariance matrix of the cokriging estimator of the regionalized components at structure s is given by

$$\mathrm{var}(\mathbf{z}_s^*) = \mathbf{W}_s^T\mathbf{K}\mathbf{W}_s = \mathbf{k}_s^T\mathbf{K}^{-1}\mathbf{k}_s \neq \mathbf{K}_{(o),s} \tag{3.7}$$

and the variance-covariance matrix of cokriging errors is

$$\boldsymbol{\Sigma}_s^E = \mathbf{K}_{(o),s} - \mathbf{k}_s^T\mathbf{K}^{-1}\mathbf{k}_s. \tag{3.8}$$

Similarly, the covariance between the cokriging estimator of regionalized components at two different structures ($s \neq s'$) is given by

$$\mathrm{cov}(\mathbf{z}_s^*, \mathbf{z}_{s'}^*) = \mathbf{W}_s^T\mathbf{K}\mathbf{W}_{s'} = \mathbf{k}_s^T\mathbf{K}^{-1}\mathbf{k}_{s'}. \tag{3.9}$$

Therefore, although the modeled value of the covariance between regionalized components from different structures is zero in the LMC, it follows from Eq. (3.9) that the covariance between estimated components at two different structures ($s \neq s'$) is not null. Consequently, we have:

$$\mathbf{\Sigma}^E_{s,s'} = -\mathbf{k}_s^T \, \mathbf{K}^{-1}\mathbf{k}_{s'}. \tag{3.10}$$

It must be underlined that even when cokriging of regionalized components is performed on the sampling points (i.e., $\mathbf{K}_{(o),s}$=$\mathbf{k}_s$), the quantities $\mathbf{\Sigma}^E_s$ and $\mathbf{\Sigma}^E_{s,s'}$ are not zero and a smoothing effect affects the cokriging estimates. For cokriging estimates of a nugget effect component, negative autocorrelation is induced over very short distances producing what could be called a 'roughing effect'.

3.4 Conditional Gaussian co-simulation of regionalized components

Hereafter, we extend the theory of cokriging of regionalized components to generate conditional simulations of regionalized components. Classically, conditional realizations are simulated by adding a pseudo-random component with a variance-covariance matrix equal to that of cokriging errors. For regionalized components, this implies:

$$\mathbf{z}^{cs}_s = \mathbf{z}^*_s + \mathbf{z}^E_s, \tag{3.11}$$

where $\mathbf{z}^E_s = \mathbf{L}^E_s \boldsymbol{\varepsilon}_s$, with $\mathbf{L}^E_s$, the lower triangular matrix arising from the LU decomposition (e.g., a Cholesky factorization) of $\mathbf{\Sigma}^E_s$ ($\mathbf{\Sigma}^E_s = \mathbf{L}^E_s[\mathbf{L}^E_s]^T$), and $\boldsymbol{\varepsilon}_s$, a vector of $N(0,1)$ pseudo-random numbers. Note that the following equation is equivalent to Eq. (3.11) :

$$\mathbf{z}^{cs}_s = \mathbf{L}_{21,s}\mathbf{L}^{-1}_{11,s}\mathbf{z} + \mathbf{L}_{22,s}\boldsymbol{\varepsilon}_s, \tag{3.12}$$

where $\mathbf{L}_{11,s}$,$\mathbf{L}_{22,s}$ and $\mathbf{L}_{21,s}$ arise from the following LU decomposition:

$$\begin{pmatrix} \mathbf{L}_{11,s} & \mathbf{0} \\ \mathbf{L}_{21,s} & \mathbf{L}_{22,s} \end{pmatrix} \begin{pmatrix} \mathbf{U}_{11,s} & \mathbf{U}_{12,s} \\ \mathbf{0} & \mathbf{U}_{22,s} \end{pmatrix} = \begin{pmatrix} \mathbf{K} & \mathbf{k}_s \\ \mathbf{k}_s^T & \mathbf{K}_{(0),s} \end{pmatrix}. \tag{3.13}$$

Equations (3.12) and (3.13) are derived from the adaptation of the LU decomposition method described in Davis (1987) and Alabert (1987), to the conditional simulation of regionalized components.

However, when Eq. (3.11) or Eq. (3.12) is used to simulate regionalized components for each structure separately (more specifically, $\mathbf{z}_s^E$ is simulated independently of $\mathbf{z}_{s'}^E$, hence $\text{cov}(\mathbf{z}_s^E, \mathbf{z}_{s'}^E) = 0$), the variance-covariance matrix of the sum of simulated regionalized components is

$$\begin{aligned}
\text{var}(\sum_{s=1}^{S} \mathbf{z}_s^{cs}) &= \sum_{s=1}^{S} \text{var}(\mathbf{z}_s^*) + 2 \sum_{s=1}^{S-1} \sum_{s'=s+1}^{S} \text{cov}(\mathbf{z}_s^*, \mathbf{z}_{s'}^*) + \sum_{s=1}^{S} \text{var}(\mathbf{z}_s^E) \\
&\quad + 2 \sum_{s=1}^{S-1} \sum_{s'=s+1}^{S} \text{cov}(\mathbf{z}_s^E, \mathbf{z}_{s'}^E) + \sum_{s=1}^{S} \sum_{s'=1}^{S} \text{cov}(\mathbf{z}_s^*, \mathbf{z}_{s'}^E) \\
&= \sum_{s=1}^{S} \mathbf{k}_s{}^T \mathbf{K}^{-1}\, \mathbf{k}_s + 2 \sum_{s=1}^{S-1} \sum_{s'=s+1}^{S} \mathbf{k}_s{}^T \mathbf{K}^{-1}\, \mathbf{k}_{s'} \\
&\quad + \sum_{s=1}^{S} \left(\mathbf{K}_{(o),s} - \mathbf{k}_s{}^T \mathbf{K}^{-1}\, \mathbf{k}_s \right) + 0 + 0 \\
&= \sum_{s=1}^{S} \mathbf{K}_{(o),s} + 2 \sum_{s=1}^{S-1} \sum_{s'=s+1}^{S} \mathbf{k}_s{}^T \mathbf{K}^{-1}\, \mathbf{k}_{s'} \\
&\neq \sum_{s=1}^{S} \mathbf{K}_{(o),s}. \qquad (3.14)
\end{aligned}$$

It follows that the variance of the simulated total variables obtained by the addition of simulated regionalized components is likely to be greater than that observed on the sampled variables (i.e., on the total data values). Therefore, we suggest instead to work with all $\mathbf{z}_s^E$ (s = 1, ..., S) simultaneously, which provides an $N_o pS$x 1 random vector $\underset{\cdot}{\mathbf{z}}^E$ with variance-covariance matrix

$$\text{var}(\underset{\cdot}{\mathbf{z}}^E) = \begin{pmatrix} \boldsymbol{\Sigma}_1^E & \cdots & \boldsymbol{\Sigma}_{1,S}^E \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{S,1}^E & \cdots & \boldsymbol{\Sigma}_S^E \end{pmatrix}, \qquad (3.15)$$

where $\boldsymbol{\Sigma}_s^E$ (s = 1, ..., S) are given by Eq. (3.8) and $\boldsymbol{\Sigma}_{s,s'}^E$ by Eq. (3.10). Eq. 3.14 then becomes

$$\begin{aligned}\text{var}(\sum_{s=1}^{S}\mathbf{z}_s^{cs}) &= \sum_{s=1}^{S}\mathbf{k}_s^T\mathbf{K}^{-1}\,\mathbf{k}_s + 2\sum_{s=1}^{S}\sum_{s'=s+1}^{S}\mathbf{k}_s^T\mathbf{K}^{-1}\,\mathbf{k}_{s'} \\ &\quad + \sum_{s=1}^{S}\left(\mathbf{K}_{(o),s} - \mathbf{k}_s^T\mathbf{K}^{-1}\,\mathbf{k}_s\right) + +2\sum_{s=1}^{S-1}\sum_{s'=s+1}^{S} -\mathbf{k}_s^T\mathbf{K}^{-1}\,\mathbf{k}_{s'} + 0 \\ &= \sum_{s=1}^{S}\mathbf{K}_{(o),s}. \end{aligned} \tag{3.16}$$

This way, all regionalized components can be simulated simultaneously by

$$\underset{\cdot}{\mathbf{z}}^{cs} = \underset{\cdot}{\mathbf{z}}^{*} + \underset{\cdot}{\mathbf{z}}^{E} = \underset{\cdot}{\mathbf{z}}^{*} + \mathbf{L}^E\boldsymbol{\varepsilon}, \tag{3.17}$$

where $\mathbf{L}^E$ is the lower triangular matrix arising from the Cholesky factorization of $\text{var}(\underset{\cdot}{\mathbf{z}}^E)$ ($\mathbf{L}^E[\mathbf{L}^E]^T = \text{var}(\underset{\cdot}{\mathbf{z}}^E)$), $\boldsymbol{\varepsilon}$ is an N_opS x 1 vector of $N(0,1)$ pseudo-random numbers and $\underset{\cdot}{\mathbf{z}}^*$ is the vector containing the cokriging estimates of regionalized components for all variables and all structures. From Eq. (3.16), it appears that the auto- and cross-covariance structures that are modeled in the LMC, within and among structures as well as for the total variables, are preserved by conditional simulations obtained with Eq. (3.17).

Conditional simulations can be obtained equivalently by

$$\underset{\cdot}{\mathbf{z}}^{cs} = \mathbf{L}_{21}\mathbf{L}_{11}^{-1}\,(\mathbf{1}\otimes\mathbf{z}) + \mathbf{L}_{22}\boldsymbol{\varepsilon}, \tag{3.18}$$

where $\mathbf{1}$ is the S x 1 vector of ones, and $\mathbf{L}_{11}$, $\mathbf{L}_{22}$ and $\mathbf{L}_{21}$ arise from the following LU decomposition:

$$\begin{pmatrix}\mathbf{L}_{11} & \mathbf{0}\\ \mathbf{L}_{21} & \mathbf{L}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{U}_{11} & \mathbf{U}_{12}\\ \mathbf{0} & \mathbf{U}_{22}\end{pmatrix} = \begin{pmatrix}\mathbf{K} & \mathbf{k}_1 & \mathbf{k}_2 & \cdots & \mathbf{k}_S\\ \mathbf{k}_1^T & \mathbf{K}_{(o),1} & 0 & \cdots & 0\\ \mathbf{k}_2^T & 0 & \mathbf{K}_{(o),2} & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & 0\\ \mathbf{k}_S^T & 0 & \cdots & 0 & \mathbf{K}_{(o),S}\end{pmatrix}. \tag{3.19}$$

Note that Eq. (3.19) can be rewritten so that the LU decomposition of one matrix of size N_np x N_np and that of one matrix of size N_opS x N_opS are needed instead of the decomposition above that involves an $(N_np + N_opS)$ x $(N_np + N_opS)$ matrix (see Alabert, 1987 for details).

It should be mentioned that with datasets comprising a large number of points or variables, the matrix $\text{var}(\mathbf{z}^E)$ may become very large and its LU decomposition may become unworkable. In such cases, the field can either be separated into partially overlapping zones, as proposed by Alabert (1987), or the algorithm developed by Vargas-Guzman and Dimitrakopoulos (2002) may be adapted. Moreover, the conditional simulation of regionalized components in a sequential simulation framework, in which simulations would be performed for each grid node separately, is made difficult by the fact that simulated values are generated for each regionalized component, while the original conditioning data are only available for the total variables. Thus, a sequential simulation would have to use two types of conditioning data with different spatial properties: the total variables on the original sampling points and the previously simulated regionalized components on the grid nodes. Instead, when simulations of regionalized components are to be obtained on a large number of nodes outside of the original sampling points, one can proceed in two steps: 1) Obtain simulated regionalized components on the original sampling points with the method described above; 2) for each structure separately, use the regionalized components simulated on the original sampling points as conditioning data in a procedure such as sequential conditional Gaussian co-simulation (Gomez-Hernandez and Journel, 1994) or direct sequential co-simulation (Soares, 2001). This is justified because regionalized components simulated for different structures on the original sampling points are generated to be independent of one another.

Strictly speaking, the co-simulation approach requires a multi-Gaussian distribution of the original data (Journel and Huijbregts, 1978). It is generally recommended to transform all variables to normal scores before performing the simulations and to back-transform simulated values to ensure a good reproduction of the histograms of original data (Goovaerts, 1997). In the co-simulation of regionalized components, a back-transformation is impossible because there are no sampled data that correspond to the simulated data. Furthermore, cokriging estimates obtained on the sampling points cannot be used as a basis for normal score transformation because their distribution is altered by the smoothing effect discussed above. We are thus left with two options: 1) Transform each variable to a normal score without back-transforming the simulated regionalized components; 2) Assume an approximate multi-Gaussian distribution for the sampled data without transforming them to normal scores. The true regionalized components are, by definition, never accessible from the sampled data. There is thus little ground on which to speculate about their exact distribution in practice. If the distribution of the total observations is close to normality, our inclination is to suggest option 2), which will lead to simulated regionalized components whose sum will have a

distribution close to that of the sampled data.

3.5 Case study

3.5.1 Objectives

In this section, the method described above for the conditional Gaussian co-simulation of regionalized components is applied to an agronomic dataset to characterize the spatial variation in soil phosphorus and potassium and obtain maps that adequately represent the scale-specific spatial structure and uncertainty of those variables. We will use this case study to illustrate how scale-specific structures and correlations between variables, as modeled in the LMC, are reproduced on simulated regionalized components, and how the uncertainty of those structures and correlations can be examined from simulated regionalized components.

3.5.2 Site description

The experiment was carried in Ste-Brigide, Québec, Canada (73^{o}03'W, 45^{o}20'N), in two adjacent agricultural fields with similar management history. The two fields had been under a corn-barley-soybean-alfalfa crop rotation for several years, and had been fertilized with uniform applications of inorganic fertilizers and dairy cattle manure.

3.5.3 Soil sampling and analysis

The soil sampling followed from the combination of a staggered regular grid (266 points) with spacing of 20 m in the width orientation of the fields and a random grid (58 points) with points at various distances from their nearest neighbors on the regular grid, providing a total of 324 points (Fig. 3.1). In November 2000, the spatial coordinates of grid points were recorded with a Global Positioning System with differential corrections from a radio beacon. A composite soil sample was taken from the top 15 cm over a 1-m^2 surface surrounding each point. Samples were air-dried and ground to pass through a 2-mm sieve. Extractable phosphorous and potassium levels (now referred to as P and K) were determined from soil samples using the Mehlich-III extractant (Carter, 1993).

3.5.4 Statistical analysis

To improve normality, a Box-Cox transformation (Box and Cox, 1964) was applied to the data (i.e., total observations) for each soil variable separately. Each set of

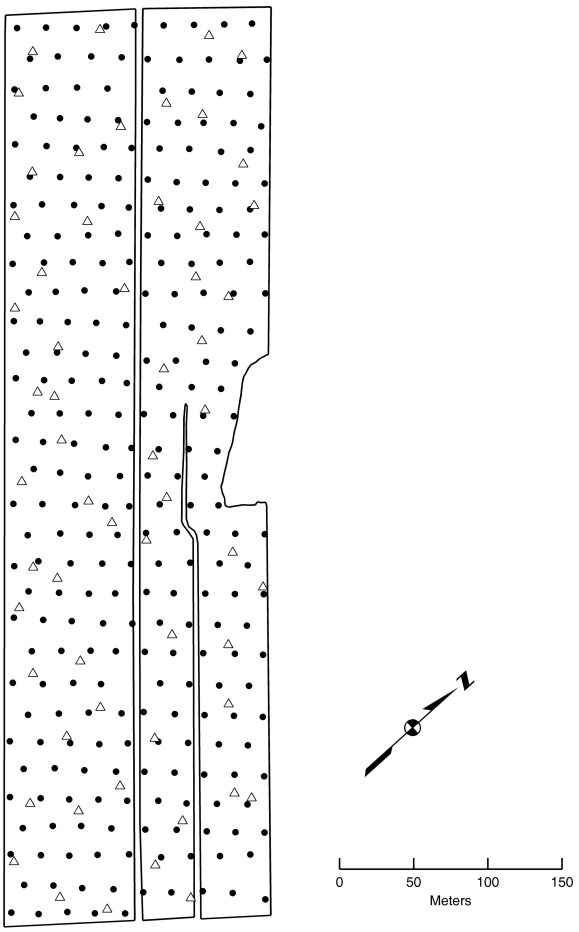

Figure 3.1: Map of the sampling scheme, made of the combination of 266 points on a staggered regular grid (•) and 58 randomly located points (Δ).

transformed data was standardized to a zero mean and a unit variance prior to further statistical analysis. Isotropic direct and cross experimental variograms were computed using distance classes (15-m intervals), in addition to one lag located at 7.5 m to evaluate the semi-variance at short distances (Fig. 3.2). After close examination, it was determined that the combination of one nugget effect, one short-range (30 m) spherical model and one long-range (150 m) Gaussian model would be appropriate to model all variograms. The LMC was fitted using the generalized least-squares algorithm developed by Pelletier et al. (2004).

The Gstat software (Pebesma and Wesseling, 1998)was used with parameter values following from the fitting of the LMC, to perform simple cokriging of the total variables on a 1-m resolution raster (Fig. 3.3A). Programs were written in the Matlab language (The Mathworks, 2002) to compute simple cokriging estimates (Fig. 3.3B-C) and vari-

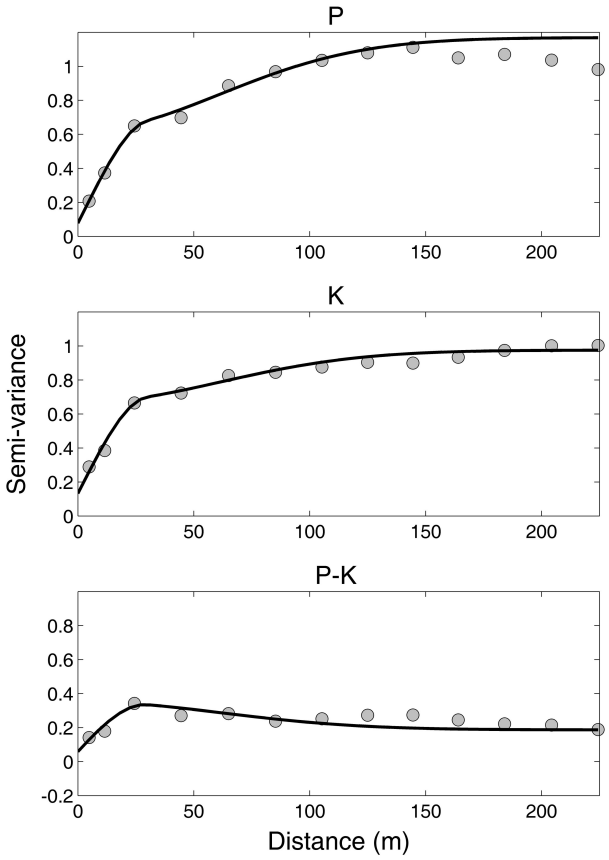

Figure 3.2: Direct and cross experimental variograms and models fitted by generalized least squares.

ances of cokriging errors (Fig. 3.4) of the regionalized components for each variable on a raster. Simple cokriging and conditional Gaussian co-simulation of regionalized components (Eq. (3.17)) were also performed in Matlab to obtain estimates and 1000 conditional co-simulations of regionalized components on the original sampling points. Experimental variograms evaluated from these estimates and simulations are displayed in Fig. 3.5, while variances, covariances and correlation coefficients calculated from them are reported in Table 3.1. The first two simulations were used in the Gstat conditional sequential Gaussian co-simulation routine with parameters associated with each structure, to obtain simulations of regionalized components on a raster (Figs. 3.6B-C and 3.7B-C). The three maps of regionalized components (including the nugget effect) for a given conditional co-simulation and variable were then summed to obtain one simulation of the total variable (Figs. 3.6A and 3.7A). When estimating or simulating onto a raster, a circular neighbourhood with a radius equal to the maximum range of the variogram model(s) involved was used.

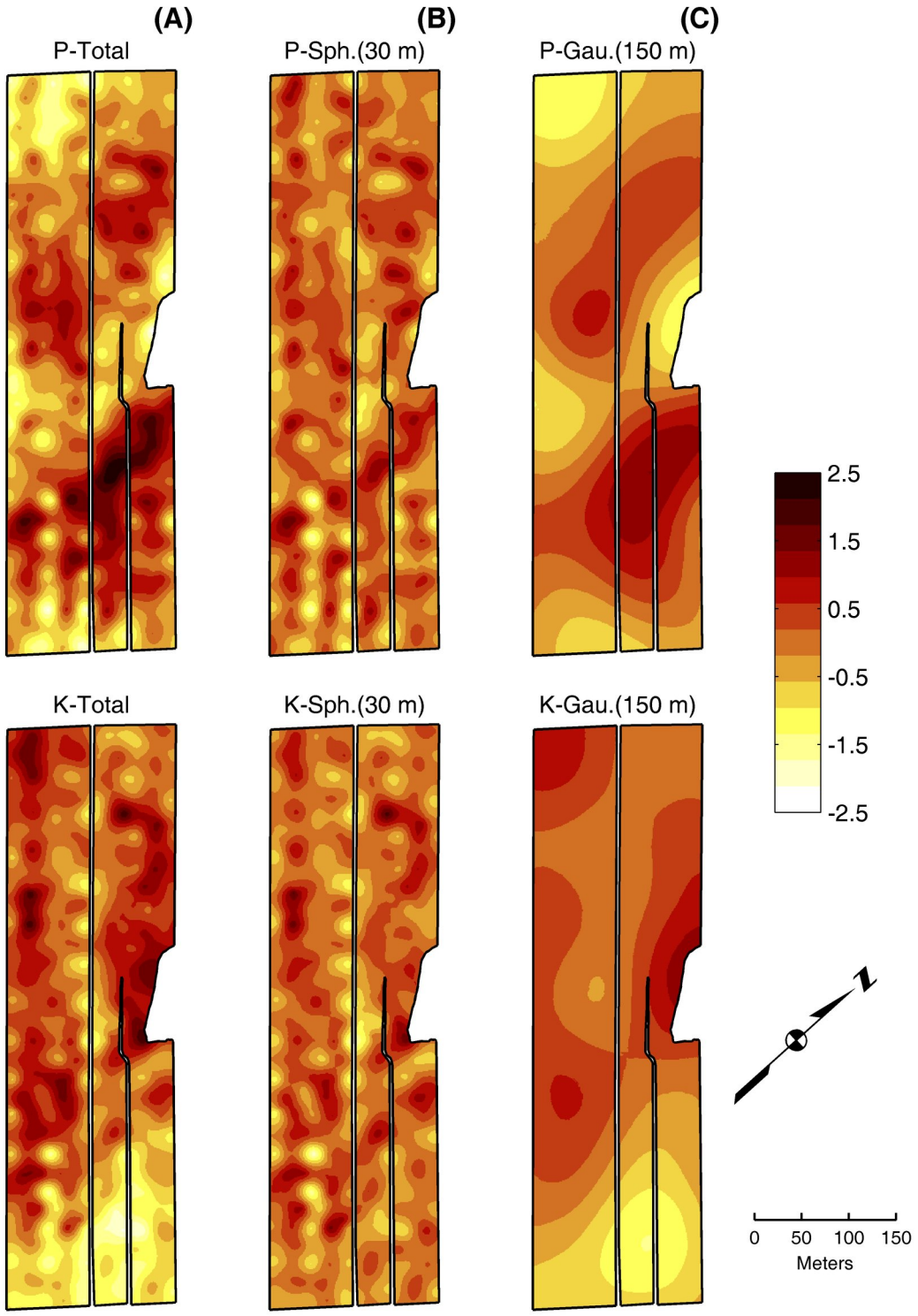

Figure 3.3: Maps obtained by simple cokriging of the two total variables (**A**) and simple cokriging of regionalized components for the short-range spherical (**B**) and long-range Gaussian (**C**) components.

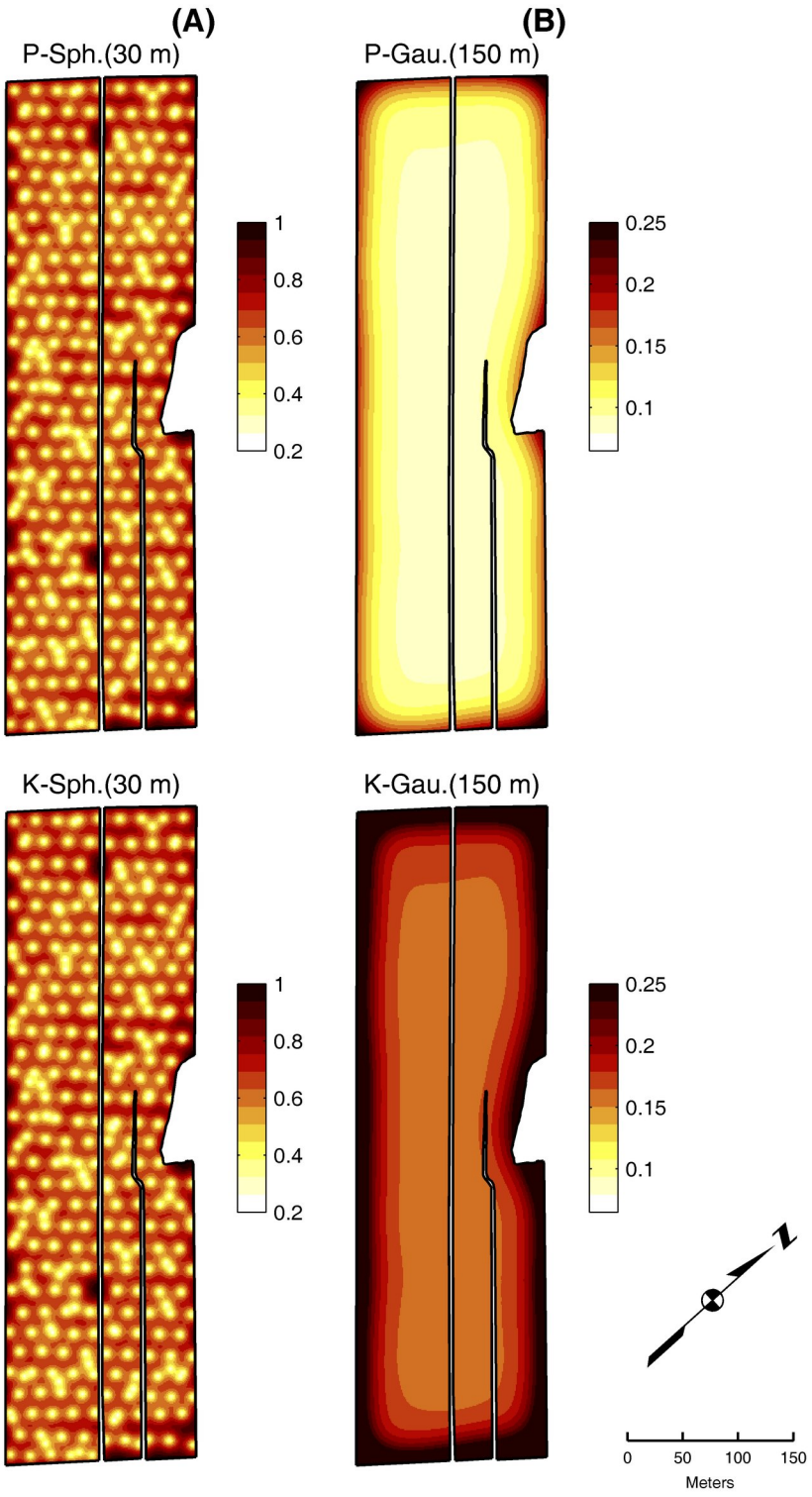

Figure 3.4: Maps of the variance of the cokriging error of the short-range spherical (**A**) and long-range Gaussian (**B**) components for P and K. Variances were divided by the corresponding direct variogram sill to obtain a proportional quantification of the variability of cokriging errors.

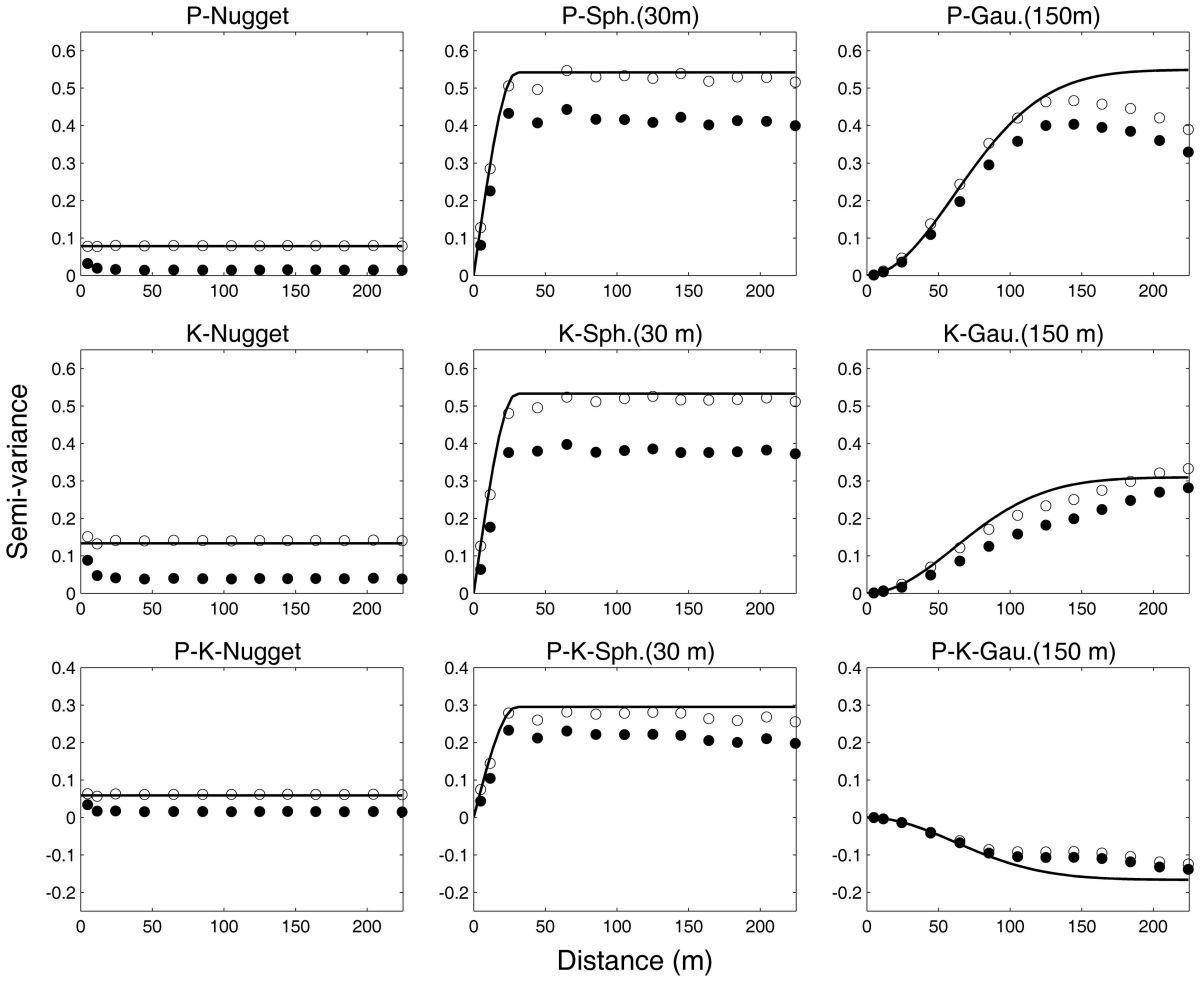

Figure 3.5: Direct and cross experimental variograms for the cokriging estimates of regionalized components on the sampling points (•) and the average experimental variograms of 1000 conditional co-simulations of regionalized components on the sampling points (∘). The solid line represents the variogram model for each structure as fitted in the LMC.

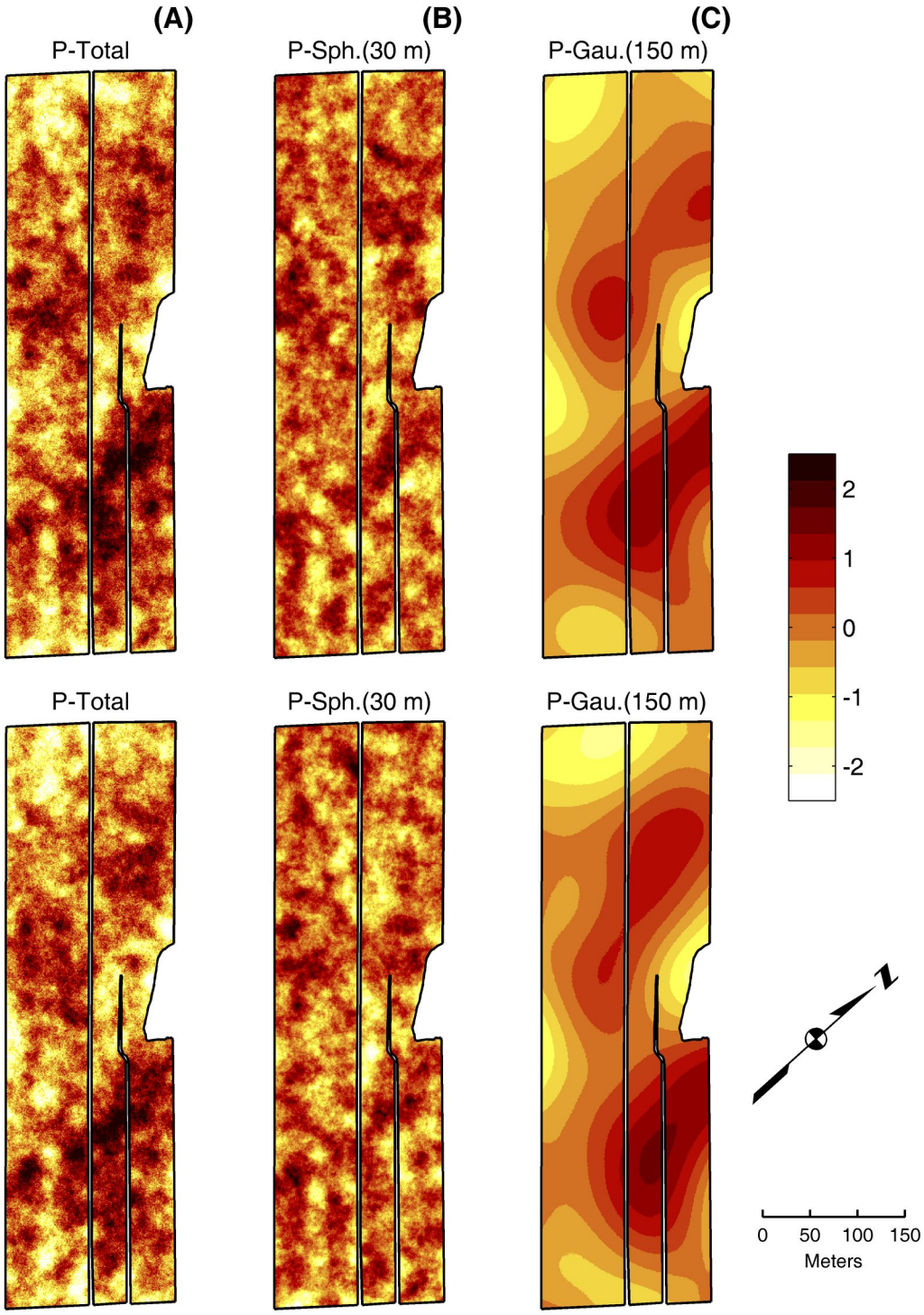

Figure 3.6: Maps of two simulations for P obtained by conditional co-simulation for the total variable (**A**), and the short-range spherical (**B**) and long-range Gaussian (**C**) regionalized components.

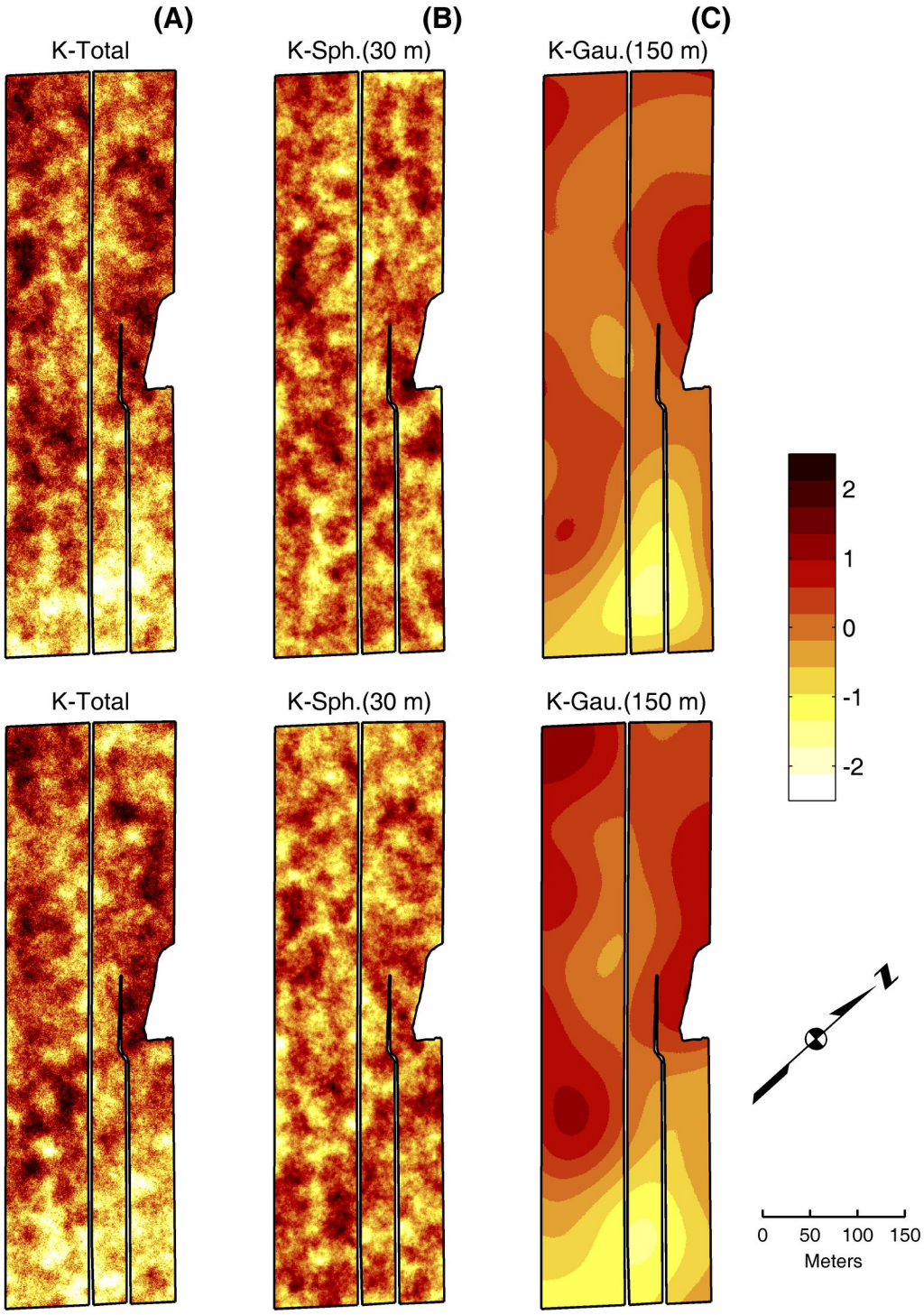

Figure 3.7: Maps of two simulations for K obtained by conditional co-simulation for the total variable (**A**), and the short-range spherical (**B**) and long-range Gaussian (**C**) regionalized components.

3.5.5 Results and discussion

The use of a random grid in this study allowed variogram estimates to be obtained at short distances (Fig. 3.2). Therefore, it may be reasonable to assume that the nugget effect fitted in the LMC (Table 3.1) arises from experimental error rather than micro-scale variation. Consequently, conditional simulations of the total variables generated without a nugget effect should more appropriately represent the true spatial characteristics of the variables. They can simply be obtained by summing the simulated short- and long-range regionalized components for each variable (Figs. 3.6B-C and 3.7B-C).

Focusing on spatially structured components, the correlation between K and P (Table 3.1) is moderately positive at the short-range spherical structure (0.55), but is negative at the long-range Gaussian structure (-0.40). This is illustrated by the cokriged and simulated maps of the regionalized components (Figs. 3.3, 3.6 and 3.7): maps for K and P have similar patterns for the short-range component and generally opposite patterns of peaks and troughs for the long-range component. Such scale dependence is usually interpreted as reflecting the presence of different controlling factors acting on the variables at the scales described by each variogram model. However, due to the smoothing effect of kriging, the spatial structure of short-range components is inadequately represented by the cokriged maps (Fig. 3.3). In fact, the comparison of experimental variograms (Fig. 3.5) shows that the smoothing effect is clearly noticeable in the cokriging estimates for the nugget effect component (for which negative autocorrelation is also observed at very short distances) and short-range components, even when cokriging is performed on the sampling points. Therefore, simulated short-range components (Figs. 3.6B and 3.7B) better represent the complexity of the small scale spatial structure. By contrast, variograms (Fig. 3.5) and maps of cokriging estimates (Fig. 3.3) and conditional co-simulations for the long-range component (Figs. 3.6C and 3.7C) are generally similar because the Gaussian model describing the long-range component captures virtually no variation at scales smaller than the distance between adjacent grid points.

The proportion of variance of cokriging errors for the short-range components of K and P is large, but is nearly identical for both variables (Fig. 3.4). In contrast, this proportion is much smaller for the long-range components, but significantly higher for K than for P. Consequently, the short-range components of both variables are quite dissimilar from one simulated realization to the next (Figs. 3.6B and 3.7B), and although realizations of long-range components are generally similar, they tend to differ more for K than for P (Figs. 3.6C and 3.7C). In this context, an interesting use of a series of such simulated realization would be to identify potentially high or low values in

Table 3.1: A comparison between the GLS sill estimates obtained by fitting the LMC, the variances and covariances calculated from the cokriging estimates of regionalized components on the original sampling points, and the mean values and empirical 95 percent confidence intervals (lower and upper bounds in parentheses) of the variances and covariances computed from 1000 conditional co-simulations of regionalized components on the original sampling points. The corresponding correlation coefficients are shown in bold above the main diagonals.

		Nugget		*Spherical (30 m)*		*Gaussian (150 m)*	
		P	***K***	***P***	***K***	***P***	***K***
Sill estimates							
Nugget	***P***	*0.08*	***0.57***	-	-	-	-
	K	*0.06*	*0.13*	-	-	-	-
Spherical (30 m)	***P***	-	-	*0.54*	***0.55***	-	-
	K	-	-	*0.29*	*0.53*	-	-
Gaussian (150 m)	***P***	-	-	-	-	*0.55*	***-0.40***
	K	-	-	-	-	*-0.17*	*0.31*
Cokriging estimates							
Nugget	***P***	*0.01*	***0.65***	***0.84***	***0.64***	***0.06***	***0.04***
	K	*0.02*	*0.04*	***0.39***	***0.85***	***0.00***	***0.08***
Spherical (30 m)	***P***	*0.07*	*0.05*	*0.42*	***0.57***	***0.12***	***0.07***
	K	*0.05*	*0.10*	*0.22*	*0.37*	***0.04***	***0.16***
Gaussian (150 m)	***P***	*0.00*	*0.00*	*0.05*	*0.01*	*0.33*	***-0.40***
	K	*0.00*	*0.01*	*0.02*	*0.05*	*-0.12*	*0.27*
Conditional simulations							
Nugget	***P***	*0.08 (0.07, 0.09)*	***0.58 (0.51, 0.65)***	***0.02 (-0.07, 0.12)***	***0.02 (-0.09, 0.11)***	***0.00 (-0.10, 0.11)***	***0.01 (-0.11, 0.12)***
	K	*0.06 (0.05, 0.07)*	*0.14 (0.12, 0.16)*	***0.01 (-0.08, 0.11)***	***0.02 (-0.06, 0.11)***	***0.00 (-0.11, 0.10)***	***0.01 (-0.09, 0.12)***
Spherical (30 m)	***P***	*0.00 (-0.01, 0.02)*	*0.00 (-0.02, 0.03)*	*0.53 (0.48, 0.59)*	***0.54 (0.48, 0.59)***	***-0.01 (-0.14, 0.12)***	***0.03 (-0.11, 0.17)***
	K	*0.00 (-0.02, 0.02)*	*0.01 (-0.02, 0.03)*	*0.28 (0.23, 0.33)*	*0.51 (0.45, 0.57)*	***0.00 (-0.14, 0.12)***	***0.03 (-0.10, 0.16)***
Gaussian (150 m)	***P***	*0.00 (-0.02, 0.02)*	*0.00 (-0.03, 0.02)*	*-0.01 (-0.07, 0.05)*	*0.00 (-0.06, 0.05)*	*0.39 (0.28, 0.52)*	***-0.31 (-0.09, -0.51)***
	K	*0.00 (-0.02, 0.02)*	*0.00 (-0.02, 0.02)*	*0.01 (-0.04, 0.06)*	*0.01 (-0.05, 0.05)*	*-0.11 (-0.18, -0.03)*	*0.32 (0.22, 0.44)*

scale-specific patterns, by building empirical confidence intervals on regionalized component estimates at each node. On average, however, these are not expected to differ significantly from confidence intervals calculated from the cokriging estimates (for the center) and the variance of cokriging errors (for the width).

The variances and covariances calculated from the cokriging estimates of regionalized components are smaller than the corresponding sill estimates (Table 3.1). There is a much better agreement between the mean values of variances and covariances of the co-simulated regionalized components and the sill estimates, especially for the nugget effect and the short-range spherical structure. However, in this case study, the correlations evaluated from the cokriging estimates of regionalized components are relatively close to those calculated directly from the sill estimates. As expected from Eq. (3.9), the correlation between cokriging estimators of regionalized components for different structures is larger than for their co-simulated counterparts (Table 3.1). With one exception (i.e., the long-range Gaussian structure in the direct variogram of P), the sill estimates fall within the empirical 95 percent confidence intervals built for the variances and covariances calculated from conditional co-simulations. This also holds true, in all cases, for correlation coefficients. Therefore, while providing a truly independent analysis at each scale, conditional co-simulations of regionalized components are more appropriate descriptors of the variation of, and relation between, variables than the cokriging estimators.

3.6 Conclusions

To our knowledge, this study is the first detailed attempt at expressing the spatial uncertainty of regionalized components. One must keep in mind that the interpretation of regionalized components, both estimated and simulated, is greatly dependent on the validity of the results from coregionalization analysis. Thus, a complete assessment of the spatial uncertainty of regionalized components will need to incorporate the effect that the uncertainty of variogram model parameter estimates has on the cokriged estimates. Nevertheless, this study clearly shows that co-simulations of regionalized components and factors provide a useful addition to the already well-developed methods of coregionalization analysis and factorial kriging, by representing both the specific features of, and the uncertainty associated with, each scale of variation. Any application that can make use of stochastic simulations can potentially make use of conditional stochastic co-simulations of regionalized components. The addition of scale to the conditional simulation framework could prove extremely useful for generating maps without exper-

imental error, studying scale-specific propagation of uncertainty in a variety of transfer functions (e.g., hydrological models or crop models), and studying the spatial uncertainty in multi-scale multivariate analyses.

3.7 Acknowledgements

The authors would like to thank the producer for allowing access to his fields and supplying essential information to properly conduct this experiment. We are also grateful to Caroline Begg for assistance, comments and supervision of field activities, and to numerous students who have helped with sampling activities. This study was funded in part by the Potash and Phosphate Institute of Canada and the Natural Sciences and Engineering Research Council of Canada (NSERC). The computer resources necessary to complete this research were purchased thanks to a Canada Foundation for Innovation grant. During his doctoral work, the first author was supported by scholarships from the Fond Québécois de la Recherche sur la Nature et la Technologie and NSERC. The second and fourth authors acknowledge NSERC Individual Discovery grants. We are grateful to two anonymous reviewers for providing useful comments on a previous version of this manuscript.

3.8 Appendix

The system of equations for conditional Gaussian co-simulation of regionalized factors is very similar to that described for regionalized components in the body of this paper. Factorial cokriging first requires the decomposition of the coregionalization matrices with, among other possibilities, principal component analysis (Goovaerts, 1992):

$$\mathbf{B}_s = \mathbf{Q}_s \mathbf{\Lambda}_s \mathbf{Q}_s^T = \mathbf{A}_s \mathbf{A}_s^T, \tag{3.20}$$

with

$$\mathbf{A}_s = \mathbf{Q}_s \mathbf{\Lambda}_s^{1/2} \tag{3.21}$$

where $\mathbf{Q}_s$ is the matrix of eigenvectors and $\mathbf{\Lambda}_s$ is the diagonal matrix of eigenvalues. Estimators of regionalized factors are then provided by

$$\mathbf{y}_s^* = \mathbf{W}_{\mathrm{F},s}^T \mathbf{z}, \tag{3.22}$$

where the matrix of factorial cokriging weights $\mathbf{W}_{\mathrm{F},s}$ is obtained by solving

$$\mathbf{K}\mathbf{W}_{\mathrm{F},s} = \mathbf{A}_s \otimes \boldsymbol{\rho}_{(n,o),s}. \tag{3.23}$$

The variance-covariance matrices of the estimates of regionalized factors and the associated cokriging errors are given respectively by

$$\mathrm{var}(\mathbf{y}_s^*) = \mathbf{W}_{\mathrm{F},s}^T \mathbf{K}\mathbf{W}_{\mathrm{F,s}} = \mathbf{A}_s^T \otimes \boldsymbol{\rho}_{(n,o),s}^T \mathbf{K}^{-1} \mathbf{A}_s \otimes \boldsymbol{\rho}_{(n,o),s} \tag{3.24}$$

and

$$\boldsymbol{\Sigma}_{\mathrm{F},s}^E = \mathbf{A}_s \otimes \boldsymbol{\rho}_{(o),s} - \mathbf{A}_s^T \otimes \boldsymbol{\rho}_{(n,o),s}^T \mathbf{K}^{-1} \mathbf{A}_s \otimes \boldsymbol{\rho}_{(n,o),s} \tag{3.25}$$

From Eq. (3.24), it appears that, although regionalized factors are defined to be orthogonal in Eq. (3.20), their cokriging estimates are not guaranteed to be orthogonal and hence, their variance-covariance matrix and autocorrelation structure are not maintained.

Similarly, the covariance between the estimators of two regionalized factors and that between the corresponding cokriging errors are respectively given by

$$\mathrm{cov}(\mathbf{y}_s^*, \mathbf{y}_{s'}^*) = \mathbf{W}_{\mathrm{F},s}^T \mathbf{K}\mathbf{W}_{\mathrm{F,s'}} = \mathbf{A}_\mathbf{s}^\mathbf{T} \otimes \boldsymbol{\rho}_{(n,o),s}^T \mathbf{K}^{-1} \mathbf{A}_{s'} \otimes \boldsymbol{\rho}_{(n,o),s'} \tag{3.26}$$

and

$$\boldsymbol{\Sigma}_{\mathrm{F},(s,s')}^E = -\mathbf{A}_s^T \otimes \boldsymbol{\rho}_{(n,o),s}^T \mathbf{K}^{-1} \mathbf{A}_{s'} \otimes \boldsymbol{\rho}_{(n,o),s'}. \tag{3.27}$$

A matrix $\mathrm{var}(\underset{\cdot}{\mathbf{y}}^E)$, similar to $\mathrm{var}(\underset{\cdot}{\mathbf{z}}^E)$ in Eq. (3.27), can then be built from all matrices $\boldsymbol{\Sigma}_{\mathrm{F},s}^E$ and $\boldsymbol{\Sigma}_{\mathrm{F},(s,s')}^E$. Following the same logic as for the simulation of regionalized components, conditional simulations of all regionalized factors can be obtained simultaneously by

$$\underset{\cdot}{\mathbf{y}}^{cs} = \underset{\cdot}{\mathbf{y}}^* + \underset{\cdot}{\mathbf{y}}^E = \underset{\cdot}{\mathbf{y}}^* + \mathbf{L}_\mathrm{F}^E \boldsymbol{\varepsilon}, \tag{3.28}$$

where $\underset{\cdot}{\mathbf{y}}^*$ is the vector containing all regionalized factors $\mathbf{y}_s^*$ estimated by simple factorial cokriging and $\mathbf{L}_\mathrm{F}^E$ is the lower triangular matrix arising from the LU decomposition of $\mathrm{var}(\underset{\cdot}{\mathbf{y}}^E)$.

A fundamental feature of regionalized factors simulated this way is that their variance-covariance matrix is the identity matrix, so their orthogonality and autocovariance structure are maintained.

Chapter 4

Characterization and quantification of uncertainty in coregionalization analysis

4.1 Abstract

Coregionalization analysis has been presented as a method of multi-scale analysis for multivariate spatial data. Despite an increasing use of this method in environmental and earth sciences, the uncertainty associated with the estimation of parameters in coregionalization analysis (e.g., sills and functions of sills) is potentially high and has not yet been characterized. This article aims to discuss the theory underlying coregionalization analysis and assess the robustness and limits of the method. A theoretical framework is developed to calculate the ergodic and fluctuation variance-covariance matrices of least-squares estimators of sills in the linear model of coregionalization. To adjust for the positive semidefiniteness constraint on estimated coregionalization matrices, a confidence interval estimation procedure for sills and functions of sills is presented. Thereafter, the relative importance of uncertainty measures (bias and variance) for sills and structural coefficients of correlation and determination is assessed under different scenarios to identify factors controlling their uncertainty. Our results show that the sampling grid density, the choice of the least-squares estimator of sills, the positive semidefiniteness constraint, the presence of scale dependence in the correlations, and the number and range of variogram models, all affect the level of uncertainty, sometimes through multiple interactions. The asymptotic properties of variogram model parameter estimators in a bounded sampling domain impose a theoretical limit to their accuracy and precision. Because of this limit, the uncertainty was found to be high for several scenarios, especially with three variogram models, and was often more depen-

dent on the ratio of variogram range to domain extent than on the sampling grid density. In practice, in the coregionalization analysis of a real dataset, the circular requirement for sill estimates in the calculation of uncertainty measures makes the quantification of uncertainty very problematic, if not impossible. The use of coregionalization analysis must be made with due knowledge of the uncertainty levels and limits of the method.
Key words: least-squares estimators of sills, direct and cross variograms, multi-scale analysis, positive semidefiniteness, ergodic, estimation and fluctuation variances, structural correlations and coefficients of determination.

4.2 Introduction

Increasing awareness about the existence of scale dependencies in spatial data sets has motivated the development of analytical methods to separate scales of variation and describe relationships between variables at multiple scales. Coregionalization analysis, sometimes referred to as 'factorial kriging analysis', comprises a set of geostatistical techniques designed for multi-scale analysis of multivariate spatial data (Wackernagel, 2003, chap. 28). These techniques rely on sill matrices (coregionalization matrices) estimated in a nested variogram model. In regionalized multivariate analysis, coregionalization matrix estimates are submitted to methods of data analysis such as structural correlations, principal component analysis (PCA) and redundancy analysis (RDA). To date, most applications have used structural correlations or PCA, or both (Goovaerts, 1992; Goulard and Voltz, 1992; Goovaerts et al., 1993; Goovaerts and Webster, 1994; Webster et al., 1994; Dobermann et al., 1997; Bocchi et al., 2000; Castrignano et al., 2000a,b; Lin, 2002; Bourennane et al., 2003). To our knowledge, the only application of regionalized RDA is by Goovaerts (1994). Variogram parameter estimates can also be used in a modified cokriging system to estimate regionalized components or factors (Matheron, 1982; Goovaerts, 1997, chaps. 5 and 6), or used to perform conditional Gaussian co-simulation of regionalized components and factors (Larocque et al., 2006).

Despite the increasing use of these methods, the different aspects of the uncertainty associated with the estimation of sills in coregionalization analysis have not yet been documented. As variogram modelling is often seen as a preliminary step towards kriging, more attention has been paid to characterizing the uncertainty of kriged values than that of variogram model parameter estimates. The choice of variogram model parameter values that minimize uncertainty is often carried out indirectly by comparing the observed and kriged values (e.g., cross-validation; Isaaks and Srivastava, 1989, pp. 351-368). Most studies on variogram uncertainty have focused on the ergodic

variance, defined as the variance of the difference between the experimental variogram and the variogram of the random function (theoretical variogram). In this context, Pardo-Igúzquiza and Dowd (2001) and Ortiz and Deutsch (2002) have shown how the uncertainty associated with a single direct experimental variogram can be quantified theoretically or through simulations, and how this uncertainty can be transferred to variogram model parameter estimates. They have shown that in applications, the evaluation of the uncertainty associated with variogram model parameter estimates must make use of the model parameter estimates themselves. This circular problem is due to the fact that variogram uncertainty is not only dependent on the characteristics of the sampling scheme, including the number of sites and their spatial location (Warrick and Myers, 1987; Webster and Oliver, 1992; Pettitt and McBratney, 1993; Müller and Zimmerman, 1999), but also on the spatial structures and scales of variation of the processes and the estimation methods used to fit the variogram model (Müller and Zimmerman, 1999; Ortiz and Deutsch, 2002). Pardo-Igúzquiza and Dowd (2003) propose an assessment of variogram parameter uncertainty based on the maximum likelihood approach.

Following Matheron's (1989; original French version, 1978) theory and terminology, model parameters defined within a spatially bounded sampling domain can be considered objective only if they satisfy certain criteria. One of those criteria is that they must have a deterministic counterpart (a 'regional magnitude'), which could be obtained without modelling if we had complete knowledge of the regionalized variable within the sampling domain. Another criterion concerns the asymptotic properties of model parameter estimators. From a modelling point of view, parameters said to be 'micro-ergodic' can be rigorously determined by the knowledge of one realization of the random function within the sampling domain. An internal evaluation of the models (a 'sorting out' operation) is needed to identify micro-ergodic parameters. Only those can be considered objective.

For example, the regional variogram is a regional magnitude that can be unambiguously defined from a complete knowledge of the regionalized variable within the sampling domain. When an increasing number of samples are collected within a given domain (this is referred to as 'infill asymptotics' by Cressie (1993)), the variogram model parameter estimates do not necessarily converge to the parameter values of the theoretical model. Instead, they converge to the values provided by the fitting of the model to the regional variogram. The fluctuation variance, defined as the variance of the difference between regional and theoretical variograms (Matheron, 1965, p. 227; Chilès and Delfiner, 1999), can be used to assess the objectivity of the parameters defining a

variogram model: a low fluctuation variance implies an objective meaning, whereas a high fluctuation variance implies no more than a 'conventional' meaning. According to Matheron (1989), parameters with a conventional meaning should disappear from the practical application of models and the final formulation of conclusions.

The variance of the difference between experimental and regional variograms, in the model, is the estimation variance of the variogram (Matheron, 1965). When this estimation variance is low, the model fitted to the experimental variogram is expected to summarize adequately the spatial structure of the variable within the sampling domain, whether the fluctuation variance of the variogram is low or high. In a slightly different context (i.e., using a simulated region from which many samples are collected), Marchant and Lark (2004) present a procedure to calculate the 'non-ergodic error' (Matheron's estimation variance), assuming the 'ergodic variogram' (theoretical variogram) is known. Their results emphasize the key distinction that must be made between the regional variogram and the theoretical variogram when characterizing variogram uncertainty.

In this study, we assess theoretically the uncertainties associated with the estimation of sills and functions of sills in the framework of the probabilistic model that serves as a basis for coregionalization analysis, that is, the linear model of coregionalization (LMC). The main hypothesis formulated in the LMC is that the total variables result from the superimposition of scale-specific structural components and that all the variables share the same scales of variation. That superimposition represents a potentially important source of uncertainty in the model due to the 'interference' between structural components. Also, at the center of coregionalization analysis are variogram model parameters (i.e., sills), for which there are no deterministic counterparts that can be defined without having recourse to the LMC. In other words, sills and functions derived from them do not exist independently of a model and can never be considered 'objective' in Matheron (1989) sense. Due to this, we will below define the notion of fluctuation variance of sills, but it will be used as a heuristic tool to discuss the hypotheses and limits of the models rather than as an indicator of the objectivity of model parameters. Comparing the fluctuation and ergodic variances will allow us to separate the effects of the sampling scheme from other factors controlling the uncertainties of sills and functions of sills. Also, the positive semidefiniteness constraint on coregionalization matrix estimates and the iterative nature of LMC fitting algorithms (Goulard, 1989; Pelletier et al., 2004) are expected to affect the distribution of sill and correlation estimators, and will thus require some numerical approximations.

The main objective of this study is to discuss the theory underlying coregionaliza-

tion analysis and assess the robustness and limits of the method. To achieve this, we first establish a framework to characterize uncertainty in the context of coregionalization analysis. Secondly, we present theoretical calculations of ergodic and fluctuation variance-covariance matrices of the direct and cross variograms and sills in the LMC. Thirdly, we combine those theoretical calculations with a simulation procedure to adjust for the positive semidefiniteness constraint and obtain confidence intervals for sills and derived functions (i.e., structural correlations and coefficients of determination). Finally, we use those confidence intervals to quantify uncertainty under different theoretical scenarios with known parameters. An approach based on theoretical scenarios is privileged because of practical problems posed by the recursive requirement for variogram model parameters in the evaluation of their own uncertainty. This will allow us to identify factors controlling uncertainty and to study the plausible levels of uncertainty in practical situations.

4.3 Uncertainty in the framework of the LMC

Before fitting a model to direct and cross experimental variograms in the context of coregionalization analysis, one must first choose the number, type and range of basic variogram models as well as the variogram estimator. Once these choices are made, all the variables are represented, in the model, as the result of the superimposition of S random structural components:

$$\mathbf{Z}(\mathbf{u}) = \sum_{s=1}^{S} \mathbf{Z}_s(\mathbf{u}) \tag{4.1}$$

where $\{\mathbf{Z}(\mathbf{u}) \mid \mathbf{u} \in \mathbb{D}\}$ is a second-order stationary vector random function, with $\mathbf{Z}(\mathbf{u}) = (Z_1(\mathbf{u}), \ldots, Z_p(\mathbf{u}))^T$, $\mathbb{D} \subset \mathbb{R}^d$ and $d \geq 1$. For a given structure s, the permissible basic variogram function g_s characterizes $Z_{is}(\mathbf{u})$ $(i = 1, \ldots, p)$ and the positive semidefinite coregionalization matrix $\mathbf{B}_s$ defines the variance-covariance matrix of the p x 1 random vector $\mathbf{Z}_s(\mathbf{u})$ for any $\mathbf{u}$. Because structural components $\mathbf{Z}_s(\mathbf{u})$ and $\mathbf{Z}_{s'}(\mathbf{u})$ for $s \neq s'$ are assumed orthogonal (i.e., statistically independent), the variance-covariance matrix of $\mathbf{Z}(\mathbf{u})$ is given by $\mathbf{V} = \sum_{s=1}^{S} \mathbf{B}_s$. Similarly, the $p \times p$ matrix containing the values of direct and cross (theoretical) variogram model values at lag $\mathbf{h}$, denoted $\boldsymbol{\Gamma}(\mathbf{h})$ and referred to as 'theoretical variogram' hereafter, is given by:

$$\boldsymbol{\Gamma}(\mathbf{h}) = \sum_{s=1}^{S} \boldsymbol{\Gamma}_s(\mathbf{h}) \tag{4.2}$$

or

$$\boldsymbol{\Gamma}(\mathbf{h}) = \sum_{s=1}^{S} \mathbf{B}_s g_s(\mathbf{h}) \tag{4.3}$$

where $\boldsymbol{\Gamma}_s(\mathbf{h})$ ($s = 1,\ldots, S$) denote matrices of values for the theoretical structural variogram models. Equation (4.3) defines the linear model of coregionalization (Journel and Huijbregts, 1978, p. 172). It must be noted that Vargas-Guzman et al. (2002) have proposed a modification of the linear model of regionalization ($p = 1$) in which the random structural components are not assumed orthogonal. Though appealing, this approach could become very difficult with more than two structures or in the multivariate framework of the LMC, due to the uncertainty associated with supplementary models that would have to be fitted.

A given realization of the structural random functions that would entirely cover a bounded sampling domain D would provide a set of structural components characterized by regional structural variogram matrices $\boldsymbol{\Gamma}_s^R(\mathbf{h})$. Such structural components would not be perfectly orthogonal because the values of cross regional variograms between components of different structures, denoted $\boldsymbol{\Gamma}_{s,s'}^R(\mathbf{h})$ ($s \neq s'$), would not be exactly zero. It follows that regional total variogram matrices resulting from their superimposition are given by:

$$\boldsymbol{\Gamma}^R(\mathbf{h}) = \sum_{s=1}^{S} \sum_{s'=1}^{S} \boldsymbol{\Gamma}_{s,s'}^R(\mathbf{h}). \tag{4.4}$$

The variance of the difference between regional and theoretical variograms,

$$\mathrm{Var}\left(\boldsymbol{\Gamma}^R(\mathbf{h}) - \boldsymbol{\Gamma}(\mathbf{h})\right)$$

is defined as the fluctuation variance of the variogram (Matheron, 1965, p. 227).

In practice, we only have access to experimental variogram matrices $\boldsymbol{\Gamma}^*(\mathbf{h})$ evaluated from a finite sampling scheme. The variance of the difference between experimental and theoretical variograms, $\mathrm{Var}\left(\boldsymbol{\Gamma}^*(\mathbf{h}) - \boldsymbol{\Gamma}(\mathbf{h})\right)$, defines the ergodic variance of the variogram, while the variance of the difference between experimental and regional variograms (in the model), $\mathrm{Var}\left(\boldsymbol{\Gamma}^*(\mathbf{h}) - \boldsymbol{\Gamma}^R(\mathbf{h})\right)$, is the estimation variance of the variogram.

An important point to note is that the fluctuation variance of the variogram can be expressed, in theory, at any lag within the limits of the sampling domain, whereas the choice of lags for the evaluation of the ergodic variance is limited by the sampling scheme. Nonetheless, at the lags at which the experimental variogram is evaluated, the ergodic variance is equal to the sum of the estimation and fluctuation variances:

$$\mathrm{Var}\left(\boldsymbol{\Gamma}^*(\mathbf{h}) - \boldsymbol{\Gamma}(\mathbf{h})\right) = \mathrm{Var}\left(\boldsymbol{\Gamma}^*(\mathbf{h}) - \boldsymbol{\Gamma}^R(\mathbf{h})\right) + \mathrm{Var}\left(\boldsymbol{\Gamma}^R(\mathbf{h}) - \boldsymbol{\Gamma}(\mathbf{h})\right). \tag{4.5}$$

Thus, similar values for the ergodic and fluctuation variances of the variogram imply a small estimation variance and hence, an experimental variogram close to the regional variogram.

Until recently, most studies on variogram model parameter uncertainty have characterized the ergodic variance. For sills, this is the variance of the difference between experimental and theoretical sills, $\mathrm{Var}\left(\hat{\mathbf{B}}_s - \mathbf{B}_s\right)$. Let $\hat{\mathbf{B}}_s^R$ ($s = 1,\ldots, S$) denote the sill matrices that would be obtained by fitting the LMC to regional total variograms. These regional sills are not likely equal to the theoretical sills $\mathbf{B}_s$. We define the variance of the difference between regional and theoretical sills, $\mathrm{Var}\left(\hat{\mathbf{B}}_s^R - \mathbf{B}_s\right)$, as the fluctuation variance of the sills. This fluctuation variance can provide insight into the minimum uncertainty that can be achieved within a given sampling domain. A similar reasoning applies to functions of sills, such as structural correlations or coefficients of determination. The equality above (Eq.(4.5)) does not extend to sills because sill estimates ($\hat{\mathbf{B}}_s$) are dependent on the lags at which the experimental variogram is evaluated. This poses a problem for defining estimation variances for sills and functions of sills. For this reason, we will focus on the calculation of ergodic and fluctuation variances in the following sections.

4.4 Theoretical characterization of uncertainty in the LMC

Below, we give the theoretical expressions for the ergodic and fluctuation variances and covariances of direct and cross variograms, as well as the mean value and the ergodic and fluctuation variance-covariance matrices of the ordinary least-squares (OLS) and generalized least-squares (GLS) estimators of sills in the context of coregionalization analysis. We show calculations for both estimators, and will present results for them in a later section, to facilitate the identification of factors controlling the uncertainty. Also, Pelletier et al. (2004) demonstrated that those sill estimators were the least efficient (OLS) and the most efficient (GLS). Even though we work under the assumption that the basic variogram functions $g_s(\mathbf{h})$ ($s = 1,\ldots, S$) in Equation (4.3) are correctly specified, while the coregionalization matrices $\mathbf{B}_s$ ($s = 1,\ldots, S$) need to be estimated, the calculations presented below require prior knowledge of all variogram model parameters, including $\mathbf{B}_s$. In practice, this leads to a well-known circular problem, which will be discussed in a later section. Any uncertainty related to the identification of the number, type and range of variogram models could be the object of another contribu-

tion.

4.4.1 Ergodic variances and covariances of variograms

In its most general form, the classical variogram estimator can be expressed as a bilinear form

$$\gamma_{ij}^{*}(\mathbf{h}) = \mathbf{z}_i^T \mathbf{A}(\mathbf{h}) \mathbf{z}_j \tag{4.6}$$

where $\mathbf{z} = (z(\mathbf{u}_1),\ldots, z(\mathbf{u}_n))^T$ is the data vector for the random function Z at a set of n sampling locations $\{\mathbf{u}_1,\ldots, \mathbf{u}_n\}$ and $\mathbf{A}(\mathbf{h})$ is the spatial design matrix

$$\mathbf{A}(\mathbf{h}) = \frac{[\boldsymbol{\eta}(\mathbf{h}) - \mathbf{M}(\mathbf{h})]}{2N(\mathbf{h})} \tag{4.7}$$

where $\mathbf{M}(\mathbf{h})$ is the n x n binary matrix $(m_{uu'})$ with $m_{uu'} = 1$ if the vector linking sampling locations u and u' corresponds to lag $\mathbf{h}$ and $m_{uu'} = 0$ otherwise, $\boldsymbol{\eta}(\mathbf{h})$ is the n x n diagonal matrix (η_{uu}) with $\eta_{uu} = \sum_{u'=1}^{n} m_{uu'}$ (the number of neighbors for sampling location u at lag $\mathbf{h}$), and $N(\mathbf{h})$ is the number of pairs of observations at lag $\mathbf{h}$.

Let $\boldsymbol{\gamma}_{ij}^{*} = \left(\gamma_{ij}^{*}(h_1), ..., \gamma_{ij}^{*}(h_K)\right)^T$ denote an experimental variogram (direct if $i = j$ and cross otherwise) at distance lags $\{h_1,\ldots, h_K\}$. Pelletier et al. (2004) showed that the entry $(h_k,\ h_{k'})$ of the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_{ij}^{*}}$ of the experimental variogram involving variables i and j can be calculated as follows

$$\mathrm{Cov}\left(\gamma_{ij}^{*}(h_k), \gamma_{ij}^{*}(h_{k'})\right) = \sum_{s=1}^{S}\sum_{s'=1}^{S} \left[(\beta_{ij,s}\beta_{ij,s'} + \beta_{ii,s}\beta_{jj,s'})\ \mathrm{tr}\left(\boldsymbol{\rho}_s \mathbf{A}(h_k) \boldsymbol{\rho}_{s'} \mathbf{A}(h_{k'})\right)\right] \tag{4.8}$$

where $(\beta_{ij,s}) = \mathbf{B}_s$, and $\boldsymbol{\rho}_s$ is the $n \times n$ matrix of values of the basic correlation function at structure s for the n sampling locations. From Equation (4.8), it follows that the variance of the experimental variogram at a given distance lag is smaller when 1) the number of pairs of observations involved in its computation is larger, and 2) when the spatial correlation between those pairs is weaker.

Equation (4.8) can be generalized to evaluate the entries of the K x K covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_{ij}^{*}\boldsymbol{\gamma}_{i'j'}^{*}}$, which contains the covariances between the cross experimental variogram involving variables i and j at lag h_k and the cross experimental variogram involving different variables i' and j' at lag $h_{k'}$:

$$\begin{aligned}&\mathrm{Cov}\left(\gamma_{ij}^{*}(h_k), \gamma_{i'j'}^{*}(h_{k'})\right) = \\ &\textstyle\sum_{s=1}^{S}\sum_{s'=1}^{S} \left[0.25\left(\beta_{ii',s}\beta_{jj',s'} + \beta_{jj',s}\beta_{ii',s'} + \beta_{ij',s}\beta_{ji',s'} + \beta_{ji',s}\beta_{ij',s'}\right) \mathrm{tr}\left(\boldsymbol{\rho}_s \boldsymbol{A}(h_k) \boldsymbol{\rho}_{s'} \boldsymbol{A}(h_{k'})\right)\right]\end{aligned} \tag{4.9}$$

This formula simplifies to calculate the covariances between a direct and a cross experimental variogram ($i = j$ and $i' \neq j'$), between direct experimental variograms ($i = j$ and $i' = j'$), for a single cross experimental variogram ($i = i'$ and $j = j'$), and for a single direct experimental variogram ($i = i' = j = j'$).

4.4.2 Ergodic variances and covariances of sills

To simplify things, but without loss of generality, consider a situation in which p=2, so there are $\frac{p(p+1)}{2} = 3$ experimental variograms, γ^*_{11}, γ^*_{12} and γ^*_{22}, and one vector of least-squares estimators of sills for each of them, $\hat{\beta}_{11}$, $\hat{\beta}_{12}$ and $\hat{\beta}_{22}$. The simultaneous least-squares fitting of all direct and cross variogram models gives

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\Phi}\boldsymbol{\gamma}^* \tag{4.10}$$

where $\widehat{\beta}$ is the $3S$ x 1 random vector $(\hat{\boldsymbol{\beta}}^T_{11},\ \hat{\boldsymbol{\beta}}^T_{12},\ \hat{\boldsymbol{\beta}}^T_{22})^T$ and Φ ($3S$ x $3K$) can be seen as a 3 x 3 diagonal matrix with $\left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T$ or $\left(\mathbf{G}^T\boldsymbol{\Sigma}^{-1}_{\gamma^*_{ij}}\mathbf{G}\right)^{-1}\mathbf{G}^T\boldsymbol{\Sigma}^{-1}_{\gamma^*_{ij}}$ (i, j = 1, 2) as the diagonal S x K block entries in the case of an OLS fitting or a GLS fitting, respectively; $\mathbf{G}$ denotes the K x S matrix of values of the basic variogram functions.

We now make use of Equation (4.9) to obtain the theoretical covariances between all possible combinations of least-squares estimators of sills for direct and cross experimental variograms. In this case, the theoretical covariance matrices $\mathrm{Cov}(\hat{\boldsymbol{\beta}}_{11},\hat{\boldsymbol{\beta}}_{12})$, $\mathrm{Cov}(\hat{\boldsymbol{\beta}}_{11},\hat{\boldsymbol{\beta}}_{22})$ and $\mathrm{Cov}(\hat{\boldsymbol{\beta}}_{12},\hat{\boldsymbol{\beta}}_{22})$ can be obtained as follows. Let $\boldsymbol{\Omega}$ denote the variance-covariance matrix of the $3K$ x 1 random vector $\boldsymbol{g}^* = ({\boldsymbol{\gamma}^*_{11}}^T,{\boldsymbol{\gamma}^*_{12}}^T,{\boldsymbol{\gamma}^*_{22}}^T)^T$, namely

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Sigma}_{\gamma^*_{11}} & \boldsymbol{\Sigma}_{\gamma^*_{11}\,\gamma^*_{12}} & \boldsymbol{\Sigma}_{\gamma^*_{11}\,\gamma^*_{22}} \\ \boldsymbol{\Sigma}_{\gamma^*_{11}\,\gamma^*_{12}} & \boldsymbol{\Sigma}_{\gamma^*_{12}} & \boldsymbol{\Sigma}_{\gamma^*_{12}\,\gamma^*_{22}} \\ \boldsymbol{\Sigma}_{\gamma^*_{11}\,\gamma^*_{22}} & \boldsymbol{\Sigma}_{\gamma^*_{12}\,\gamma^*_{22}} & \boldsymbol{\Sigma}_{\gamma^*_{22}} \end{pmatrix}. \tag{4.11}$$

The $3S$ x $3S$ ergodic variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$ and its $3S$ x 1 expected value are given by

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\Phi}\boldsymbol{\Omega}\boldsymbol{\Phi}^T \tag{4.12}$$

and

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}. \tag{4.13}$$

Equations (4.12) and (4.13) do not require the Gaussian assumption on $\boldsymbol{\gamma}^*$, and readily follow from the application of matrices in statistics (Searle, 1971) and the fact that the random vector $\hat{\boldsymbol{\beta}}$ in Equation (4.10) is a linear function of $\boldsymbol{\gamma}^*$. Equation (4.12) clearly shows how the uncertainty of experimental variograms transfers to sills estimates, since $\mathrm{Var}(\hat{\boldsymbol{\beta}})$ is a function of $\boldsymbol{\Omega}$. Equation (4.13) confirms that in the absence of constraint of positive semidefiniteness on sill matrix estimates, there is no bias in theory in any of the least-squares estimators of sills.

As we will see in later sections, knowing the theoretical variances of, and covariances between, sill estimators for direct and cross experimental variograms at the same structure and at different structures is important for (i) assessing the effect of constraining

sill matrix estimates to be positive semidefinite in the LMC and (ii) evaluating the uncertainties associated with the resulting sill estimates and the derived structural correlations and coefficients of determination.

4.4.3 Fluctuation variances and covariances of variograms

By considering the regional variogram at a given lag as a random variable, Matheron (1965) obtains a theoretical expression for the fluctuation variance of the direct variogram:

$$\mathrm{Var}[\gamma^R(\mathbf{h}) - \gamma(\mathbf{h})] = \mathrm{Var}[\gamma^R(\mathbf{h})] = \frac{1}{2K(\mathbf{h})^2} \int_{D\cap D_{-\mathbf{h}}} \int_{D\cap D_{-\mathbf{h}}} G(\mathbf{x},\mathbf{x}';\mathbf{h})\mathbf{dxdx}' \quad (4.14)$$

with, under the Gaussian assumption for the distribution of increments of the random function (e.g., $\mathbf{z}(\mathbf{x}+\mathbf{h}) - \mathbf{z}(\mathbf{x})$),

$$G(\mathbf{x},\mathbf{x}';\mathbf{h}) = [\gamma(\mathbf{x}-\mathbf{x}'-\mathbf{h}) + \gamma(\mathbf{x}-\mathbf{x}'+\mathbf{h}) - 2\gamma(\mathbf{x}-\mathbf{x}')]^2. \quad (4.15)$$

In the two-dimensional space, the value of the geometric covariogram $K(\mathbf{h})$ is the area of the intersection of the domain D with its translation by $\mathbf{h}$, denoted $|D \cap D_{-\mathbf{h}}|$. The double integral (Eq.(4.14)) involves two points $\mathbf{x}$ and $\mathbf{x}'$ sweeping independently $D \cap D_{-\mathbf{h}}$. Therefore, it corresponds in reality to a quadruple integral involving the coordinates of points $\mathbf{x}$ and $\mathbf{x}'$.

In the omni-directional case, with $\mathbf{h}$= (h,θ) where $h = \| \boldsymbol{h} \|$ is the Euclidean norm of $\mathbf{h}$ and θ denotes the angle that $\mathbf{h}$ forms with some reference axis, an average value of the regional direct variogram at distance lag h is given by

$$\overline{\gamma^R(h)} = \frac{1}{2\pi} \int_0^{2\pi} \gamma^R(h,\theta)\mathrm{d}\theta \quad (4.16)$$

so that

$$\mathrm{Var}[\overline{\gamma^R(h)}] = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{1}{2\left|D \cap D_{-(h,\theta)}\right|^2} \int_{D\cap D_{-(\mathrm{h},\theta)}} \int_{D\cap D_{-(\mathrm{h},\theta)}} G(\mathbf{x},\mathbf{x}';h,\theta)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' \right) \mathrm{d}\theta. \quad (4.17)$$

We now generalize that expression to obtain the fluctuation covariance between the cross variogram involving variables i and j and the cross-variogram involving variables i' and j' at any pair of distance lags.

$$Cov[\overline{\gamma^R_{ij}(h)}, \overline{\gamma^R_{i'j'}(h')}] = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \frac{1}{4\left|D \cap D_{-(h,\theta)}\right| \left|D \cap D_{-(h',\theta')}\right|}$$
$$\times \left(\int_{D\cap D_{-(h,\theta)}} \int_{D\cap D_{-(h',\theta')}} \left(G_{ii'}(\mathbf{x},\mathbf{x}';\mathbf{h},\mathbf{h}')G_{jj'}(\mathbf{x},\mathbf{x}';\mathbf{h},\mathbf{h}') + G_{ij'}(\mathbf{x},\mathbf{x}';\mathbf{h},\mathbf{h}')G_{ji'}(\mathbf{x},\mathbf{x}';\mathbf{h},\mathbf{h}') \right) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' \right) \mathrm{d}\theta\mathrm{d}\theta' \quad (4.18)$$

where

$$G_{ij}(\mathbf{x},\mathbf{x}';\mathbf{h},\mathbf{h}') = [\gamma_{ij}(\|\mathbf{x}-\mathbf{x}'-\mathbf{h}\|)+\gamma_{ij}(\|\mathbf{x}-\mathbf{x}'+\mathbf{h}'\|)-\gamma_{ij}(\|\mathbf{x}-\mathbf{x}'\|)-\gamma_{ij}(\|\mathbf{x}-\mathbf{x}'+\mathbf{h}'-\mathbf{h}\|)].$$

Like Equation (4.9), Equation (4.18) simplifies to calculate the variance of, or covariance between, any direct or cross variogram at any distance lag. Equation (4.18) could be used in theory to represent $\mathrm{Cov}[\overline{\gamma_{ij}^R(h)}, \overline{\gamma_{i'j'}^R(h')}]$ as a continuous surface. However, the exact evaluation of this function, either analytically or numerically, seems rather impractical. Instead, we have developed a relatively simple simulation algorithm that approximates the value of $\mathrm{Cov}[\overline{\gamma_{ij}^R(h)}, \overline{\gamma_{i'j'}^R(h')}]$ for different combinations of lags h and h' (see Appendix 1), in order to build covariance matrices $\mathbf{\Sigma}_{\gamma_{ij}^R\gamma_{i'j'}^R}$ (i, i', j, $j' = 1$, ..., p) between regional variograms. Results not shown here indicate a tight correspondence between the fluctuation variances obtained with our simulation algorithm and the ergodic variances obtained with theoretical calculations (Eq.(4.9)) for very dense grids within an equivalent sampling domain.

4.4.4 Fluctuation variances and covariances of sills

The regional sills $\hat{\beta}_{ij,s}^R$ represent the sill estimates that would be obtained by fitting a model to the regional variograms evaluated at K lags from an infinite number of sampling points in D. The fitting of a given model to all direct and cross regional variograms simultaneously would provide

$$\hat{\boldsymbol{\beta}}^R = \mathbf{\Phi}^R\boldsymbol{\gamma}^R \tag{4.19}$$

where $\mathbf{\Phi}^R$ is the block diagonal matrix made of matrices

$$\left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T \text{ or } \left(\mathbf{G}^T\mathbf{\Sigma}_{\gamma_{ij}^R}^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{\Sigma}_{\gamma_{ij}^R}^{-1}$$

for an OLS or a GLS fitting, respectively. To obtain the fluctuation variance-covariance matrix of sills, the regional version of the variance-covariance matrix $\mathbf{\Omega}$ (Eq.(4.11)), denoted $\mathbf{\Omega}^R$, must be built from all $\mathbf{\Sigma}_{\gamma_{ij}^R\gamma_{i'j'}^R}$. The fluctuation variance-covariance matrix involving all sills is then provided by

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}^R - \boldsymbol{\beta}) = Var(\hat{\boldsymbol{\beta}}^R) = \mathbf{\Phi}^R\mathbf{\Omega}^R[\mathbf{\Phi}^R]^T \tag{4.20}$$

and their expected value is

$$E(\hat{\boldsymbol{\beta}}^R) = \boldsymbol{\beta}. \tag{4.21}$$

When $S > 1$, the fluctuation variance (Eq. (4.20)) originates both from the superimposition of structural components and from the fluctuation variance of variograms before superimposition. For a given structure, the latter variance alone can be quantified by setting $S = 1$ in the calculations above, that is, by using Equation (4.20) in the

absence of the other structures.

4.5 Accounting for the positive semidefiniteness constraint

The classical variogram estimator is known to be unbiased (Cressie, 1993, p. 71), and so are least-squares estimators of slopes in linear models, including sills in the LMC when the ranges are correctly specified (Searle, 1971; Pelletier et al., 2004). This is, however, before sill matrix estimates are submitted to the constraint of positive semidefiniteness; thereafter, some bias may be expected. The constraint may also have an effect on the variance of sill estimators. We present below a confidence interval estimation procedure to evaluate uncertainty from simulated sill matrices constrained to be positive semidefinite. The center and length of a confidence interval will provide information about the bias and dispersion of the point estimator on which it is built.

Consider the $\frac{p(p+1)}{2}S$ x 1 multivariate normal distribution with mean $\boldsymbol{\beta}$ and variance-covariance matrix $\boldsymbol{\Phi\Omega\Phi}^T$ for sill estimates obtained from the fitting of a model to experimental variograms by least squares (the Gaussian assumption was verified empirically on unconstrained sill estimates). In what follows, the matrix $\boldsymbol{\Phi\Omega\Phi}^T$ can be replaced by $\boldsymbol{\Phi}^R\boldsymbol{\Omega}^R[\boldsymbol{\Phi}^R]^T$ to obtain confidence intervals for regional sills and functions derived from them.

Step 1- Simulate a $\frac{p(p+1)}{2}S$ x 1 vector of sills from the multivariate normal distribution $N(\boldsymbol{\beta}, \boldsymbol{\Phi\Omega\Phi}^T)$:

$$\boldsymbol{\beta}^* = L\boldsymbol{\varepsilon} + \boldsymbol{\beta} \tag{4.22}$$

where $\mathbf{L}$ is the lower triangular matrix arising from the Cholesky decomposition of $\boldsymbol{\Phi\Omega\Phi}^T$ and ε is a $\frac{p(p+1)}{2}$ S x 1 vector of $N(0, 1)$ pseudo-random numbers.

Step 2- Use the vector $\boldsymbol{\beta}^*$ to build $p \times p$ sill matrices $\mathbf{B}_\mathrm{s}^*$ and perform the spectral decomposition of $\mathbf{B}_\mathrm{s}^* = \mathbf{Q}_s\boldsymbol{\Lambda}_s\boldsymbol{Q}_s^\mathrm{T}$, where $\mathbf{Q}_s$ contains the eigenvectors and $\boldsymbol{\Lambda}_s$ is the diagonal matrix of eigenvalues ($s = 1,\ldots, S$).

Step 3- For each structure s, set any negative eigenvalue in $\boldsymbol{\Lambda}_s$ at zero to obtain $\boldsymbol{\Lambda}_s^+$. New, positive semidefinite simulated matrices of sills are then given by $\mathbf{B}_s^{*+} = \mathbf{Q}_s\boldsymbol{\Lambda}_s^+\mathbf{Q}_s^\mathrm{T}$ ($s = 1,\ldots, S$).

Step 4- Different functions of the sills can be evaluated from $\mathbf{B}_\mathrm{s}^{*+}$, including structural correlation coefficients, ρ_s^*, defined as $\frac{\beta_{\mathrm{ij,s}}^{*+}}{\sqrt{\beta_{\mathrm{ii,s}}^{*+}\,\beta_{\mathrm{jj,s}}^{*+}}}$, and structural coefficients of determination (ρ_s^{*2} if p=2).

Step 5- Repeat Steps 1-4 a large number of times (e.g., 10,000).
Step 6- For each structure separately, sort in ascending order the sills obtained for direct and cross variograms at Step 3 and the functions of sills evaluated at Step 4. Confidence intervals are then obtained by calculating the percentiles corresponding to a given confidence level. With 10,000 simulations and a confidence level of 0.9, the 500th and 9,500th sorted values of sills and structural coefficients provide the lower and upper bounds of the confidence intervals, while the arithmetic mean is used as center.

Results not presented here revealed an excellent agreement between confidence intervals obtained with this procedure and those obtained empirically through simulated partial realizations of the associated vector random function. Because of this agreement, we rely on theoretical calculations adjusted for the positive semidefiniteness constraint for the quantification of uncertainty in the following section.

4.6 Quantification of uncertainty under different scenarios

We will now use the theoretical calculations of previous sections under different scenarios to quantify the respective uncertainties associated with the estimation of sills and functions of sills. By doing this, one of our objectives is to assess the relative importance of factors potentially controlling uncertainty, including the positive semidefiniteness constraint on sill matrix estimates, the number of structures or variogram models (S), the number of sampling locations on a square regular grid (n), the ratio of variogram range to sampling domain extent, and the presence of scale dependence in the correlation between variables. We define eight scenarios combining different values of S (1, 2, or 3) and various patterns of scale dependence (Table 4.1). For each scenario, grid sizes are 12 by 12 ($n = 144$), 15 by 15 ($n = 225$), 20 by 20 ($n = 400$), 30 by 30 ($n = 900$), and 50 by 50 ($n = 2500$). The extent of the sampling domain corresponds to the size of the sampling grid in distance units (e.g., the 12 by 12 grid size is used in a 12 units by 12 units domain), so that the horizontal and vertical distances between adjacent nodes on the grid are always slightly greater than one unit. In all cases, the number of variables (p) is fixed to two.

Variograms were evaluated up to the 4-unit distance lag for scenarios A, B, D, and E, and up to the 8-unit distance lag for scenarios C, F, G, and H. Distance lag increments of 0.5 and 0.25 units were used, respectively, to calculate the ergodic and fluctuation

Table 4.1: Description of Scenarios in Terms of the Theoretical Values of the Parameters

Scenario	**Variogram model (structure s)**		**Range**	**Direct variogram sills** $(\beta_{11,s},\beta_{22,s})$	**Cross variogram sill** $(\beta_{12,s})$	**Structural correlation** (ρ_s)	**Structural coefficient of determination** (ρ_s^2)
A	**1**	Nugget	-	0.3333	0.2357	$\sqrt{0.5}$	0.5
B	**2**	Spherical	3	0.3333	0.2357	$\sqrt{0.5}$	0.5
C	**3**	Spherical	6	0.3333	0.2357	$\sqrt{0.5}$	0.5
D	**1**	Nugget	-	0.3333	0.2357	$\sqrt{0.5}$	0.5
	2	Spherical	3	0.3333	0.2357	$\sqrt{0.5}$	0.5
E	**1**	Nugget	-	0.3333	0.2357	$\sqrt{0.5}$	0.5
	2	Spherical	3	0.3333	-0.2357	$-\sqrt{0.5}$	0.5
F	**1**	Nugget	-	0.3333	0.2357	$\sqrt{0.5}$	0.5
	2	Spherical	3	0.3333	0.2357	$\sqrt{0.5}$	0.5
	3	Spherical	6	0.3333	0.2357	$\sqrt{0.5}$	0.5
G	**1**	Nugget	-	0.3333	0	0	0
	2	Spherical	3	0.3333	0.1826	$\sqrt{0.3}$	0.3
	3	Spherical	6	0.3333	0.2582	$\sqrt{0.6}$	0.6
H	**1**	Nugget	-	0.3333	0.2357	$\sqrt{0.5}$	0.5
	2	Spherical	3	0.3333	-0.2357	$-\sqrt{0.5}$	0.5
	3	Spherical	6	0.3333	0.2357	$\sqrt{0.5}$	0.5

variance-covariance matrices. For each scenario, coregionalization matrices $\mathbf{B}_s$ and the ergodic variance-covariance matrices were used in the algorithm for the adjustment for positive semidefiniteness, to build 0.9 confidence intervals from 10,000 $\mathbf{B}_s^{*+}$, ρ_s^* and ρ_s^{*2}. Coregionalization matrices $\mathbf{B}_s$ and the fluctuation variance-covariance matrices were used in the algorithm for the adjustment for positive semidefiniteness, to build 0.9 confidence intervals from 10,000 $\mathbf{B}_s^{R*+}$, ρ_s^{R*} and ρ_s^{R*2}.

4.7 Results and discussion

Entries of the ergodic and fluctuation variance-covariance matrices of direct and cross variograms are presented graphically in Figures 4.1 and 4.2 for scenario F and in Figures 4.3 and 4.4 for scenario H. Ergodic variance-covariance matrices of GLS sill estimators, calculated for the 15 x 15 grid size under the scenarios with $S = 2$ (Table 4.2) and $S = 3$ (Table 4.3) are presented without and with adjustment for the positive semidefiniteness constraint. Confidence intervals built on the OLS and GLS estimators of sills in direct

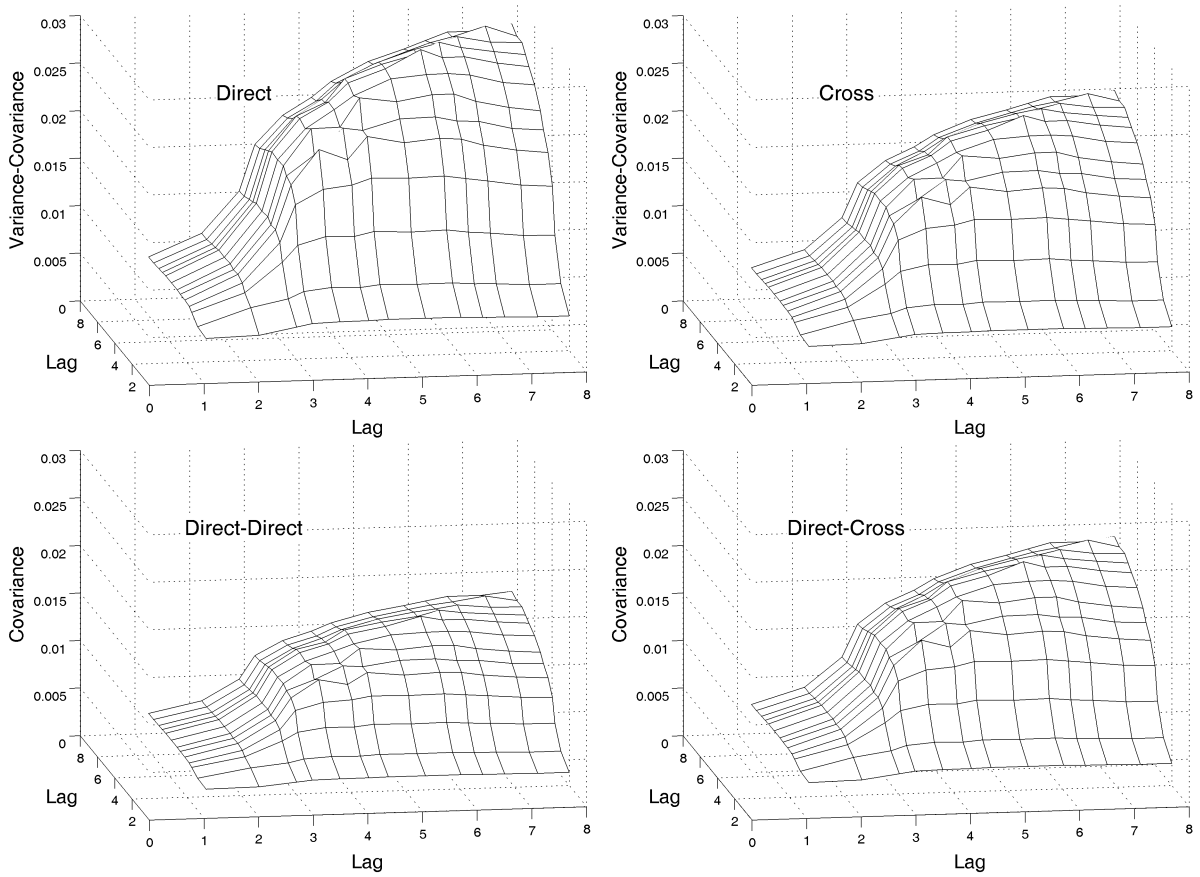

Figure 4.1: Ergodic variance-covariance matrices of the direct and cross variograms (above) and their ergodic covariance matrices (below) for the 15 x 15 grid under scenario F described in Table 4.1.

and cross experimental and regional variograms are presented in Figure 4.5 for scenarios A, B and C with one structure, in Figure 4.6 for scenarios D and E with two structures, and in Figure 4.9 for scenarios F, G and H with three structures. Results for structural correlations and coefficients of determination are presented in Figures 4.5 ($S = 1$), 4.7 and 4.8 ($S = 2$), and 4.10 and 4.11 ($S = 3$). While the uncertainty measures are represented graphically through confidence intervals (CI), to simplify our discussion we will loosely use the term 'bias' to describe the discrepancy between the theoretical value of the parameter and the empirical mean of the estimator and the term 'variance' to describe the length of the CI.

Effect of the Positive Semidefiniteness Constraint

The effect of the positive semidefiniteness constraint can be observed by comparing the variances of sill estimators without and with the adjustment for the constraint

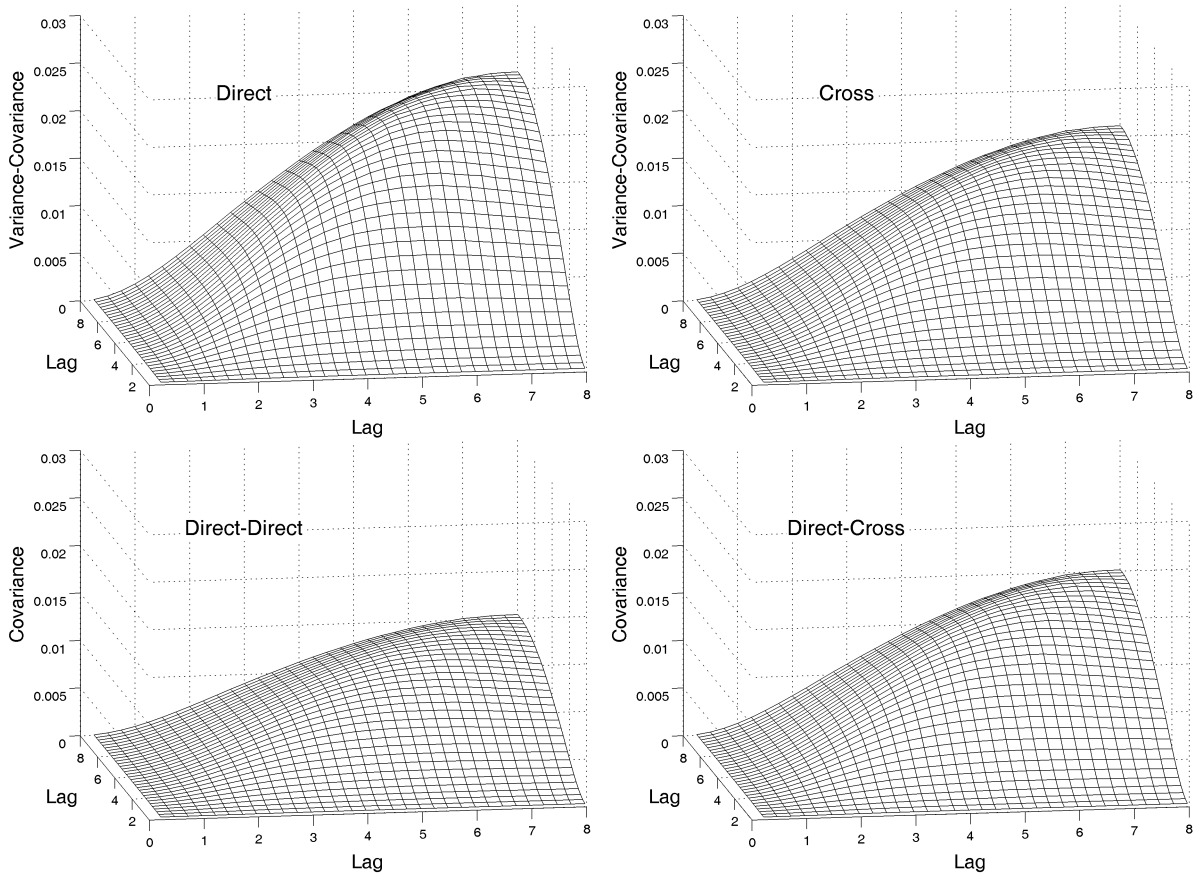

Figure 4.2: Fluctuation variance-covariance matrices of the direct and cross variograms (above) and their fluctuation covariance matrices (below) for the 15 x 15 grid under scenario F described in Table 4.1.

(Tables 4.2 and 4.3). In each of our scenarios, the theoretical sill matrices were chosen to be positive semidefinite, while the structural coefficients of determination were defined to be moderate (Table 4.1). Accordingly, a larger variance of sill estimators before adjustment implies a lower proportion of simulated sill matrices that are positive semidefinite and a more pronounced effect of the constraint. The constraint then results in a decrease of the variance of direct variogram sill estimators, since it forces the estimates to be non-negative ($\hat{\beta}_{11,s} \geq 0$, $\hat{\beta}_{22,s} \geq 0$), and of the variance of cross variogram sill estimators through the Cauchy-Schwarz inequality ($\left|\hat{\beta}_{12,s}\right| \leq \left|\hat{\beta}_{11,s}\ \hat{\beta}_{22,s}\right|$). The observed effect is more important when three variogram models are used in the LMC, in particular for spatially autocorrelated structures (Table 4.3).

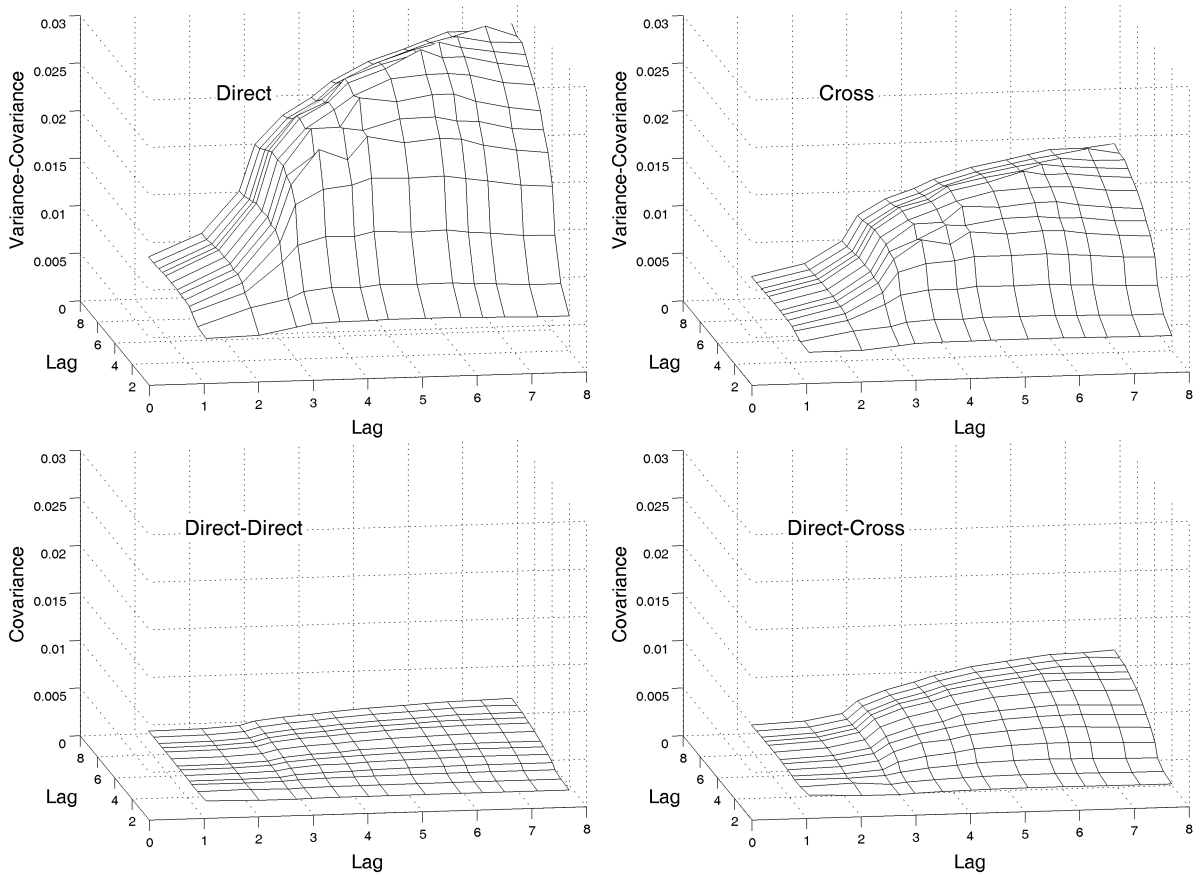

Figure 4.3: Ergodic variance-covariance matrices of the direct and cross variograms (above) and their ergodic covariance matrices (below) for the 15 x 15 grid under scenario H described in Table 4.1.

4.7.1 Factors controlling uncertainty

Below, we discuss the effects of five main factors influencing the uncertainties of the variogram, sill estimates and structural coefficients of correlation and determination in the context of coregionalization analysis: the range of autocorrelation relative to the sampling domain extent, the number of superimposed structural components, the sampling scheme (the grid density and configuration within the sampling domain), the sill estimator, and the presence of scale dependence in the correlations between variables. Perhaps the most important elements of discussion are the interactions between these factors.

By definition, fluctuation variances depend on the size and shape of the sampling domain, but not on the sampling scheme. For the nugget effect when $S = 1$, the fluctuation variance of the sill is exactly zero because the variance of the sample variance for

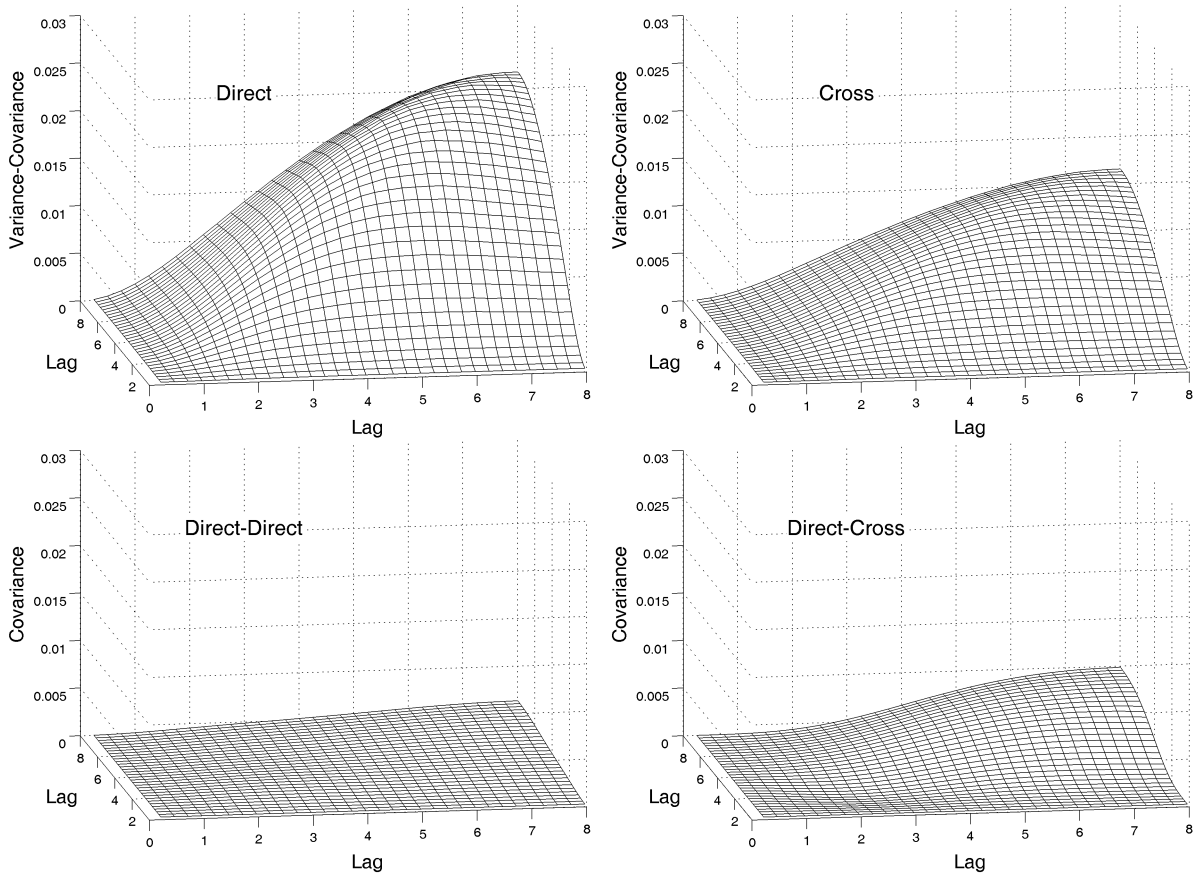

Figure 4.4: Fluctuation variance-covariance matrices of the direct and cross variograms (above) and their fluctuation covariance matrices (below) for the 15 x 15 grid under scenario H described in Table 4.1.

non-autocorrelated data is asymptotically zero, regardless of the extent of the sampling domain. Concerning spatially autocorrelated structures (spherical structures, $S = 1$, 2 or 3), the fluctuation variance of sills and structural coefficients of correlation and determination is proportional to the ratio of variogram range to sampling domain extent, a smaller ratio allowing a greater repetition of the spatial features within the domain, thereby providing a better basis for the estimation of model parameters (with lesser uncertainty). This issue, which has been previously discussed (e.g., Matheron, 1989, p. 84; Marchant and Lark (2004)), explains the gradual decrease of the fluctuation variances as the extent of the sampling domain increases. It also explains the smaller fluctuation variances for the short-range spherical structure compared to the long-range spherical structure for a given sampling domain ($S = 1$ or $S = 3$).

The portion of fluctuation variances that is caused by the superimposition of struc-

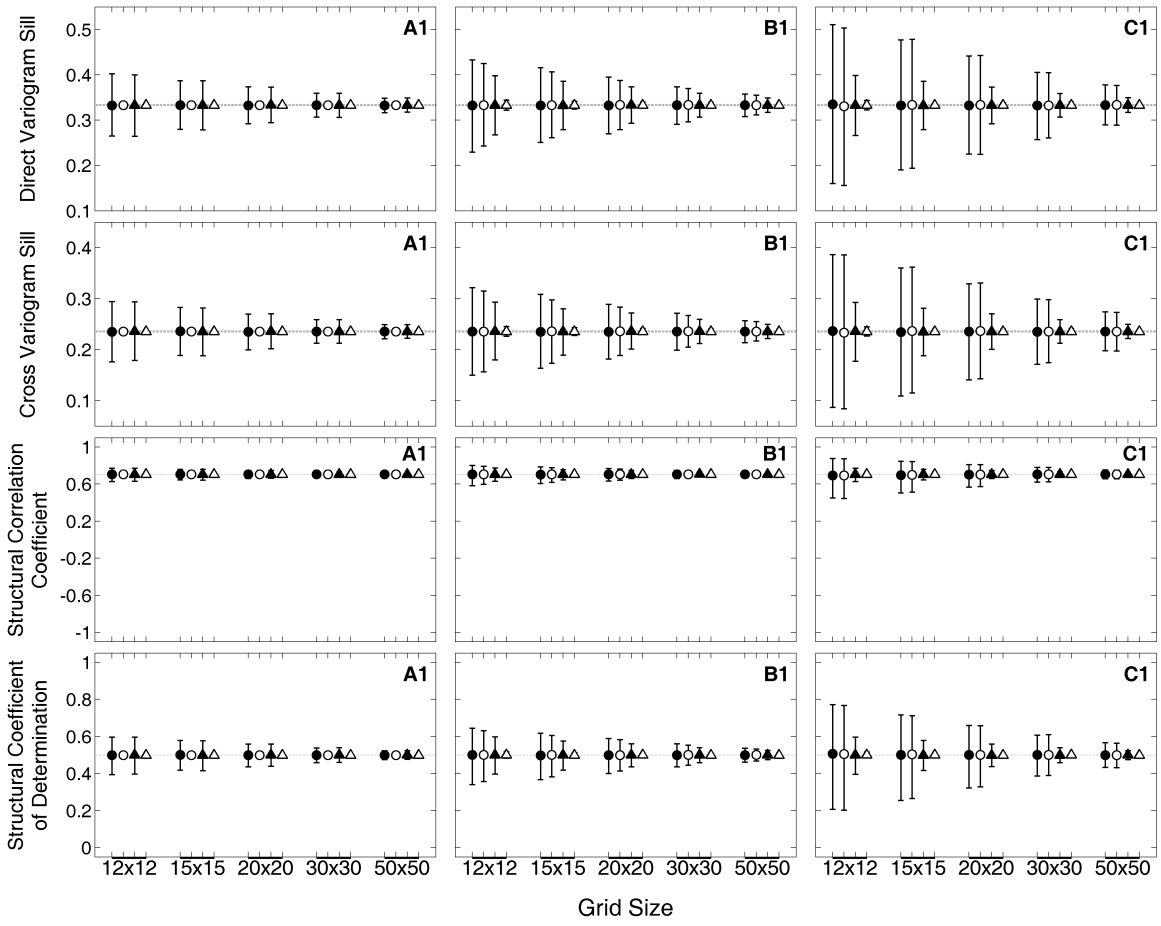

Figure 4.5: 90% confidence intervals for direct and cross variogram sills, structural correlations and structural coefficients of determination for the nugget effect (scenario A1), short-range spherical (B1) and the long-range spherical (C1) variogram models. Each scenario is described in Table 1. Confidence intervals were obtained, respectively, from ergodic variances calculated for the OLS estimator (●), fluctuation variances calculated for the OLS estimator (○), ergodic variances calculated for the GLS estimator (△), and fluctuation variances calculated for the GLS estimator (▲). All intervals were generated with adjustment for the positive semidefiniteness constraint. The dotted lines represent the theoretical values.

Table 4.2: Entries of the Ergodic Variance-Covariance Matrix of GLS Sill Estimators, without and with the Adjustement for the Positive Semidefiniteness Constraint, Calculated for the 15 x 15 Grid under Scenarios D and E Described in Table 4.1

	Unadjusted		**Adjusted**	
Scenarios	**D**	**E**	**D**	**E**
Direct				
Nugget	0.0047	0.0047	0.0047	0.0046
Nugget-Sph(3)[a]	-0.0052	-0.0052	-0.0051	-0.0048
Sph(3)	0.0115	0.0115	0.0114	0.0107
Direct-Cross				
Nugget-Nugget	0.0033	0.0010	0.0033	0.0011
Nugget-Sph(3)	-0.0036	0.0002	-0.0036	0.0004
Sph(3)-Sph(3)	0.0081	-0.0023	0.0080	-0.0028
Cross				
Nugget	0.0035	0.0027	0.0035	0.0026
Nugget-Sph(3)	-0.0039	-0.0029	-0.0038	-0.0026
Sph(3)	0.0086	0.0063	0.0085	0.0058
Direct-Direct				
Nugget	0.0024	0.0007	0.0024	0.0006
Nugget-Sph(3)	-0.0026	-0.0007	-0.0026	-0.0004
Sph(3)	0.0057	0.0011	0.0057	0.0007

[a]Sph(3): spherical variogram model with range 3.

tural components can be evaluated by comparing values for the same structural component at $S = 3$ (scenario F), $S = 2$ (D, excluding the long-range spherical model) and $S = 1$ (A, B, C). This comparison shows that the fluctuation variances of sills and structural coefficients of correlation and determination are directly proportional to the number of basic functions used to fit the LMC to total variograms. Even with an infinite number of sampling locations within a given domain, the estimation of sills for a given structure is complicated by the presence of other structures, due to the cross covariances between structural components (Eq.(4.4)). The large difference between fluctuation variances at $S = 1$ and $S = 3$ may suggest that the cross covariances between structural components (superimposition effect) have a far greater role to play in controlling the uncertainty when $S = 3$ than the ratio of variogram range to sampling domain extent. However, the

Table 4.3: Entries of the Ergodic Variance-Covariance Matrix of GLS Sill Estimators, without and with the Adjustement for the Positive Semidefiniteness Constraint, Calculated for the 15 x 15 Grid under Scenarios F, G and H Described in Table 4.1

	Unadjusted			**Adjusted**		
Scenarios	**F**	**G**	**H**	**F**	**G**	**H**
Direct						
Nugget	0.0068	0.0068	0.0068	0.0068	0.0067	0.0066
Nugget-Sph(3)[a]	-0.0096	-0.0096	-0.0096	-0.0090	-0.0087	-0.0078
Nugget-Sph(6)[a]	0.0013	0.0013	0.0013	0.0012	0.0012	0.0010
Sph(3)	0.0384	0.0384	0.0384	0.0344	0.0338	0.0289
Sph(3)-Sph(6)	-0.0326	-0.0326	-0.0326	-0.0271	-0.0266	-0.0210
Sph(6)	0.0616	0.0616	0.0616	0.0503	0.0500	0.0456
Direct-Cross						
Nugget-Nugget	0.0048	0.0018	0.0021	0.0048	0.0018	0.0021
Nugget-Sph(3)	-0.0068	-0.0029	-0.0012	-0.0064	-0.0028	-0.0002
Nugget-Sph(6)	0.0009	-0.0005	-0.0002	0.0009	-0.0005	0.0000
Sph(3)-Sph(3)	0.0271	0.0161	0.0030	0.0246	0.0149	-0.0027
Sph(3)-Sph(6)	-0.0230	-0.0165	-0.0057	-0.0198	-0.0147	-0.0031
Sph(6)-Sph(6)	0.0435	0.0365	0.0154	0.0365	0.0319	0.0171
Cross						
Nugget	0.0051	0.0037	0.0039	0.0050	0.0036	0.0038
Nugget-Sph(3)	-0.0072	-0.0053	-0.0052	-0.0067	-0.0048	-0.0040
Nugget-Sph(6)	0.0010	0.0003	0.0007	0.0008	0.0001	0.0005
Sph(3)	0.0288	0.0229	0.0200	0.0263	0.0208	0.0141
Sph(3)-Sph(6)	-0.0244	-0.0208	-0.0173	-0.0211	-0.0178	-0.0106
Sph(6)	0.0462	0.0421	0.0333	0.0396	0.0363	0.0268
Direct-Direct						
Nugget-Nugget	0.0034	0.0006	0.0010	0.0034	0.0005	0.0009
Nugget-Sph(3)	-0.0048	-0.0010	-0.0008	-0.0045	-0.0009	0.0000
Nugget-Sph(6)	0.0006	-0.0007	0.0001	0.0005	-0.0006	0.0002
Sph(3)-Sph(3)	0.0192	0.0075	0.0017	0.0168	0.0061	-0.0014
Sph(3)-Sph(6)	-0.0163	-0.0093	-0.0022	-0.0133	-0.0072	0.0010
Sph(6)-Sph(6)	0.0308	0.0229	0.0051	0.0246	0.0187	0.0042

[a]Sph(3) and Sph(6): spherical variogram models with ranges 3 and 6.

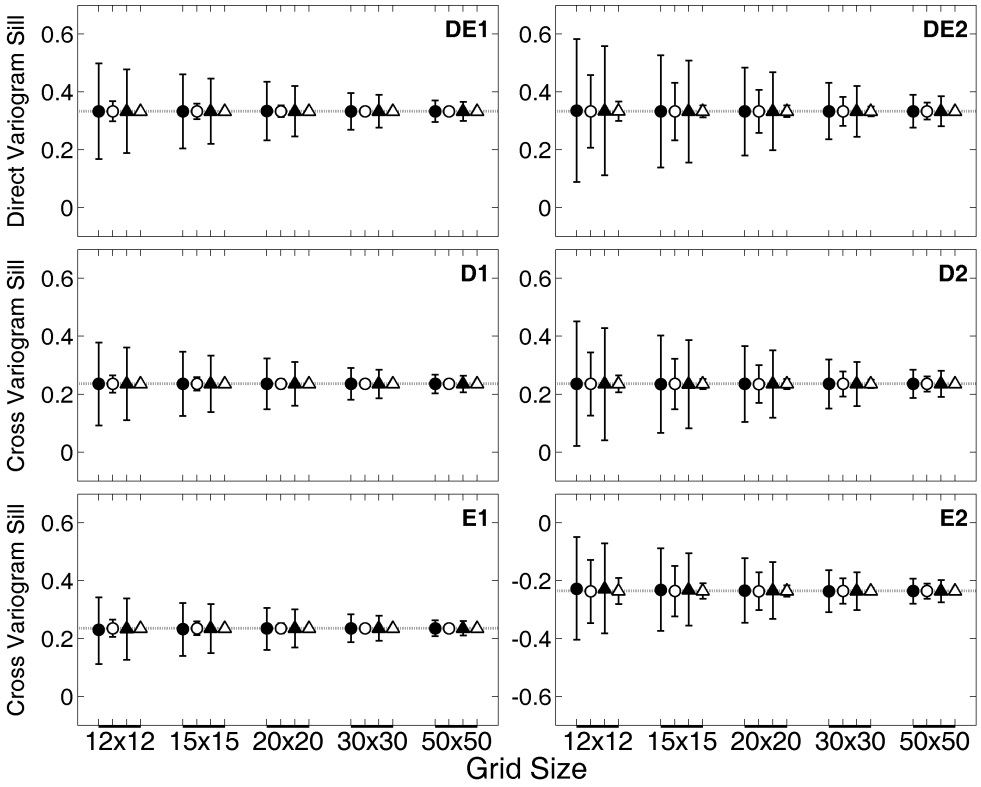

Figure 4.6: 90% confidence intervals for direct and cross variogram sills of the nugget effect (D1 and E1) and the short-range spherical (D2 and E2) variogram models under scenarios D and E described in Table 4.1. Results for direct variograms are identical for scenarios C and D and are therefore presented once. See caption of Figure 4.5 for the meaning of symbols.

sharp increase of fluctuation variances when $S = 3$ as the range of autocorrelation of the structural component increases entails that the superimposition effect predominantly exacerbates the effect of the ratio of variogram range to sampling domain extent. This implies that achieving a low fluctuation variance in coregionalization analysis with $S > 2$ requires a ratio of variogram range to sampling domain extent much lower than what is usually recommended in standard variography. For example, Journel and Huijbregts (1978, p. 94) advise that variogram analysis should not include distances exceeding half of the longest side of the sampling domain.

Although we did not quantify the estimation variance explicitly, the effect of the sampling scheme can be studied by comparing the fluctuation and ergodic variances. On the basis of the results for the GLS estimator of sills with $S = 1$ and $S = 2$, we see that in most cases the ergodic variance could be lowered by having a denser grid within the same domain. In particular, this would provide a better characterization

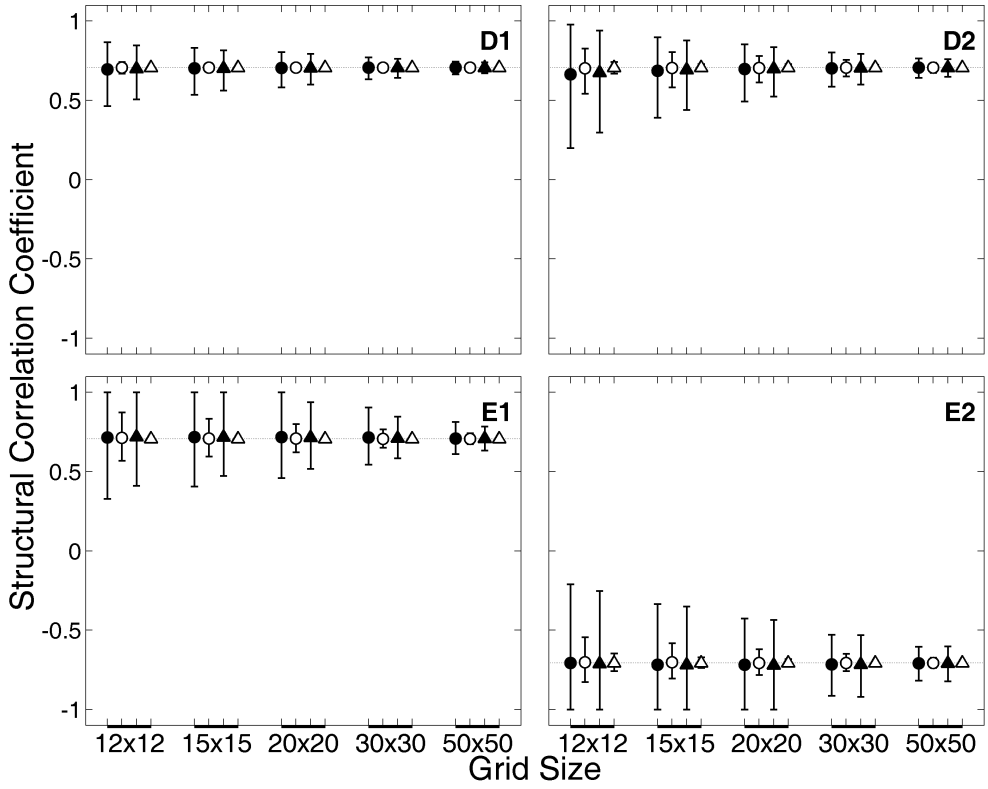

Figure 4.7: 90% confidence intervals for structural correlations for the nugget effect (D1 and E1) and the short-range spherical (D2 and E2) variogram models under scenarios D and E described in Table 4.1. See caption of Figure 4.5 for the meaning of symbols.

of lags at short distances, which is important for the GLS estimator (see below). At $S = 3$, however, a denser sampling scheme would help reduce the ergodic variance for the nugget effect and, to a lesser degree, for the short-range spherical structure, but would not be of much use for improving estimation for the long-range spherical structure because of the high fluctuation variance. When designing a sampling scheme with the objective of performing coregionalization analysis, a compromise must thus be found between a sampling scheme that allows a characterization of the variogram at distances smaller than the range of autocorrelation, and a small ratio of variogram range to sampling domain extent. Thus, the ideal sampling scheme for performing coregionalization analysis in the presence of multiple structures should be both intensive and extensive, which is often difficult to achieve in practice.

By definition, the GLS estimator of sills uses the variance-covariance matrices of the variograms in the fitting of the LMC. By doing so, it generally gives more weight to

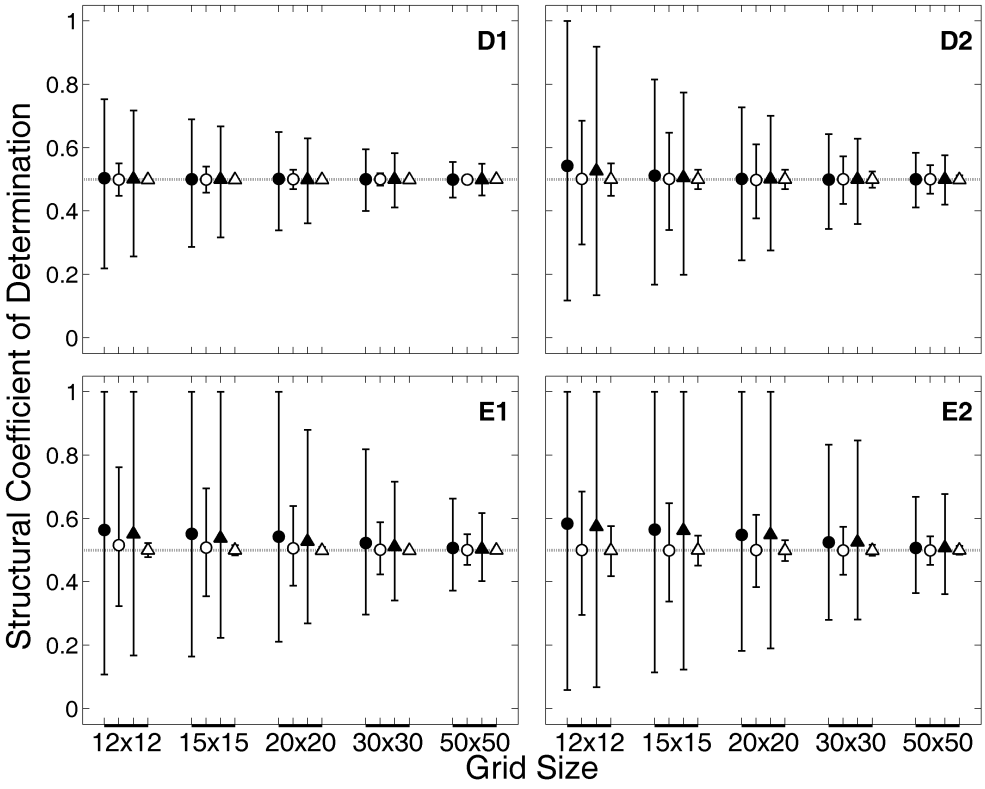

Figure 4.8: 90% confidence intervals for structural coefficients of determination for the nugget effect (D1 and E1) and the short-range spherical (D2 and E2) variogram models under scenarios D and E described in Table 4.1. See caption of Figure 4.5 for the meaning of symbols.

variogram ordinates at short distances, which have a smaller fluctuation variance and smaller fluctuation covariances with ordinates at other lags, thereby reducing the effect of the high fluctuation variance of the variogram at longer distances (Figs. 4.2 and 4.4). This explains the very small fluctuation variances of the GLS sill estimator and the derived coefficients of correlation and determination when $S = 1$ and 2, compared to the OLS estimator. However, the GLS estimator cannot correct for uncertainty that is present at all lags in the variogram, as a consequence of the sampling scheme or of the superimposition of the structural spatial components. When the uncertainty of sill estimates originates from those factors, the differences between the two estimators are reduced. For example, for the long-range spherical structure with $S = 3$, the fluctuation variances of sills for the two estimators are close, suggesting that in practice, the high variance of sills due to the superimposition effect will prevent the proper estimation of the variance-covariance matrix used in the GLS estimator. Thus, the efficiency of

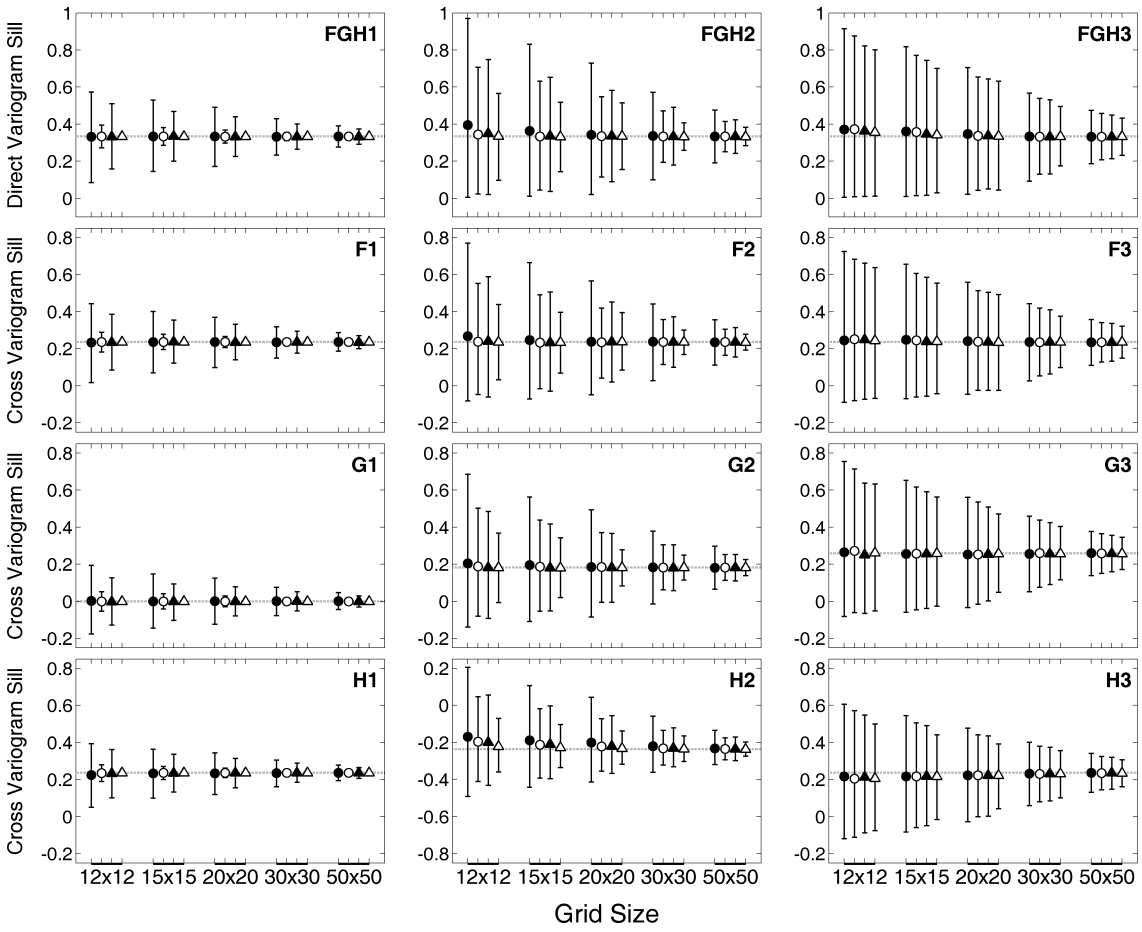

Figure 4.9: 90% confidence intervals for direct and cross variogram sills for the nugget effect (F1, G1 and H1), short-range spherical (F2, G2 and H2) and long-range spherical (F3, G3 and H3) variogram models under scenarios F, G and H described in Table 4.1. Results for direct variograms are identical for scenarios F, G and H, and are therefore presented once. See caption of Figure 4.5 for the meaning of symbols.

the GLS estimator of sills is affected by the presence of multiple structures, although it generally remains more efficient than the OLS estimator, or equally efficient at worst.

Concerning the pattern of correlation between variables, the ergodic variance of sills for cross variograms is larger in scenario F (with intrinsic correlations) than in scenarios G and H (with scale-dependent correlations), whereas the ergodic variance of sills for direct variograms is unaffected (Fig. 4.9). These results are consistent with those obtained by Pelletier et al. (2004), and originate from Equations (4.8) and (4.9), in which sills for different structures are combined in the calculations of $\mathbf{\Sigma}_{\gamma^*_{ij}}$ and $\mathbf{\Sigma}_{\gamma^*_{ij}\,\gamma^*_{i'j'}}$. Perhaps surprisingly, the variances of estimated structural coefficients are generally larger

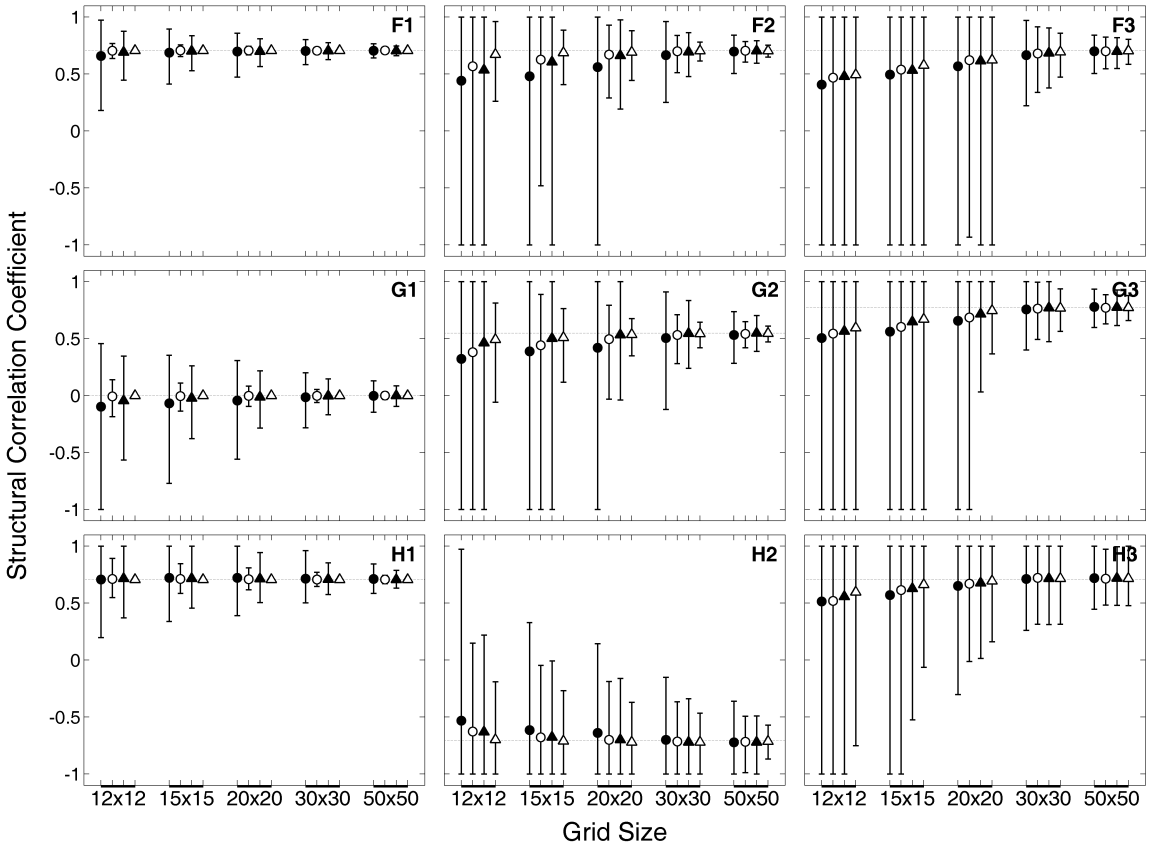

Figure 4.10: 90% confidence intervals for structural correlations for the nugget effect (F1, G1 and H1), short-range spherical (F2, G2 and H2) and long-range spherical (F3, G3 and H3) variogram models under scenarios F, G and H described in Table 4.1. See caption of Figure 4.5 for the meaning of symbols.

in scenarios G and H (with scale-dependent correlations) than in intrinsic scenario F. However, we must bear in mind that, in practice, the estimated structural coefficients of correlation and determination are calculated from $\widehat{\beta}_{11,\mathrm{s}}$, $\widehat{\beta}_{22,\mathrm{s}}$ and $\widehat{\beta}_{12,s}$, so their variances are dependent on the covariances between sill estimators (Tables 4.2 and 4.3). Furthermore, the cross experimental variogram is computed from the same two vectors of observations that are used to compute the two direct experimental variograms. The resulting covariance between direct and cross empirical variograms, which is related to the strength of the corresponding structural correlation (Figs. 4.1-4.4), affects the covariance between sill estimators at a given structure. From Table 4.2, it is possible to calculate, for example, that the correlation between direct and cross variogram sill estimators at the spherical structure is greater in scenario D (0.8165) than in scenario

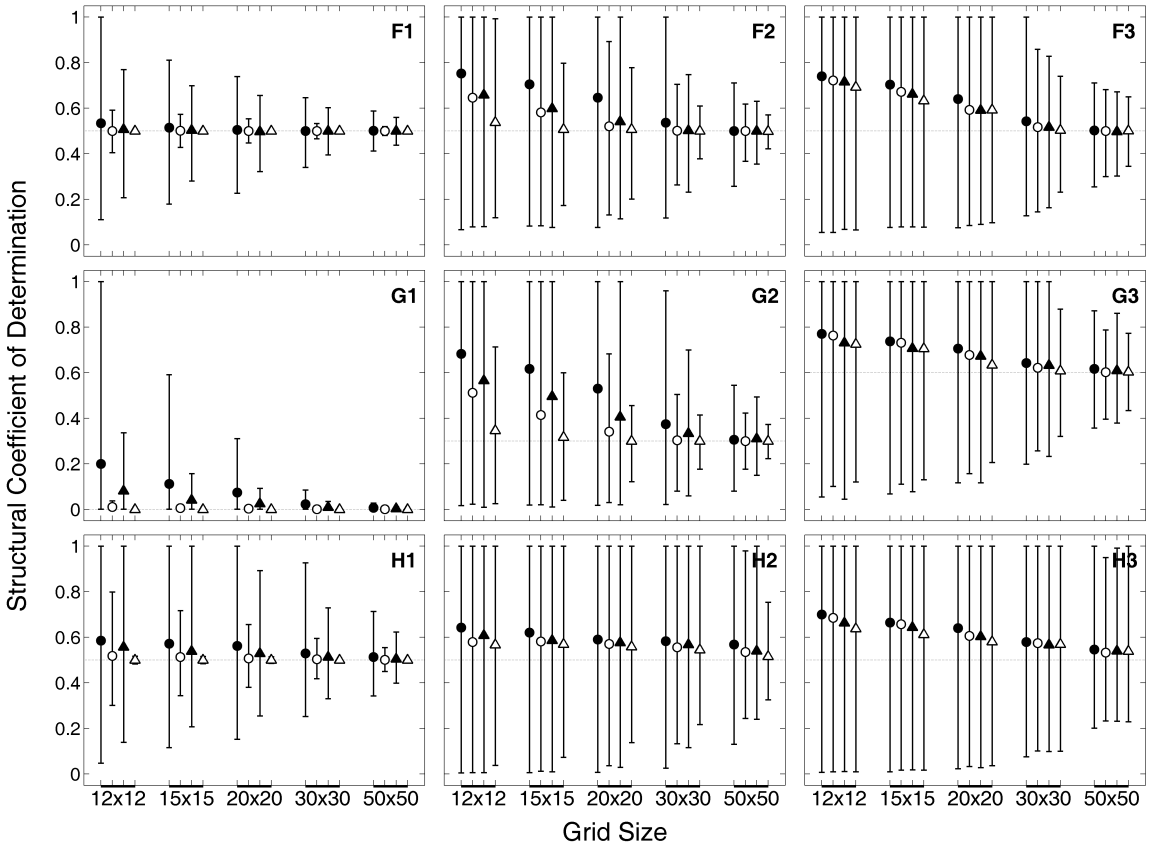

Figure 4.11: 90% confidence intervals for structural coefficients of determination for the nugget effect (F1, G1 and H1), short-range spherical (F2, G2 and H2) and long-range spherical (F3, G3 and H3) variogram models under scenarios F, G and H described in Table 4.1. See caption of Figure 4.5 for the meaning of symbols.

E (-0.2689). Scale-dependent correlations thus reduce the 'coherence' between direct and cross variogram sill estimates, leading to a higher uncertainty in the functions derived from them. This result can be of major significance in regionalized multivariate analysis, and has, to our knowledge, never been discussed in this context.

The structural coefficients of determination and, to a lesser degree, the structural correlations tend to be biased upwards in all scenarios involving two or three structures (Figs. 4.7-4.8 and 4.10-4.11). This bias seems to result from complex effects of the multiple interactions between factors controlling the uncertainty of sill estimates, and its strength appears to be proportional to the variance of sills estimators. Only the coefficients of correlation and determination for the nugget effect in scenarios D and F do not present such a bias.

To summarize, in scenarios with one structure the use of the GLS estimator in the fitting of the LMC corrects for the fluctuation variance and the uncertainty is predominantly controlled by the sampling scheme. In scenarios with one nugget effect and one spatially autocorrelated structure ($S = 2$), the uncertainty increases due to the superimposition effect, but the sampling scheme remains the dominant factor. When there are more than one spatially autocorrelated structure ($S = 3$), the uncertainty is the combined outcome of the fluctuation variance, the superimposition of structural components and, for the structural coefficients of correlation and determination, the presence of scale dependence in the correlations between variables.

4.7.2 Practical Implications

In the coregionalization analysis of a real dataset, the definition of an acceptable level of uncertainty depends on the judgment and objectives of the user. Our theoretical results provide a general idea of the levels of uncertainty for different scenarios that can be compared to practical situations.

From a purely probabilistic perspective, the uncertainty assessment must be made on the basis of the ergodic variances. Figure 4.9 shows that at smaller grid sizes ($n = 144$ or $n = 225$), structural correlations computed from sill estimates give little information about the relationship between variables at different scales in the scenarios with 3 structures. At an intermediate grid size ($n = 400$), one can detect the sign of the theoretical structural correlation but not its magnitude, and larger grids ($n > 900$) are needed to identify its sign and general strength with confidence when $S = 3$. The uncertainty of estimated structural coefficients of determination in scenarios with 3 structures is even more problematic, since only a general idea of the strength of the theoretical coefficients can be gathered with larger grids. By comparison, estimated structural correlations in scenarios with 2 structures can be interpreted with some confidence at most grid sizes; only the CIs for structural coefficients of determination with smaller grids cover a large proportion of their possible range ($[0, 1]$) when $S = 2$. It follows that results from previous uncertainty assessments made in the contexts of standard variography and kriging are likely not applicable to coregionalization analysis. The number of sampling locations and the extent of the sampling domain required in coregionalization analysis could be higher than in other kinds of applications. For example, Webster and Oliver (1992) recommended $n = 100$ as minimum and preferably $n = 150$ or 225, to keep variogram uncertainty within reasonable limits for standard variography. Many published applications of regionalized multivariate analysis have reported structural correlations computed from sill matrices estimated in a LMC with

three variogram models (including one nugget effect), using relatively small grids ($n < 300$). Our results strongly suggest that the uncertainties under many of such scenarios would be high enough that little confidence could be attributed to the interpretation of the structural correlations reported.

Assuming that the type and range of the basic variogram models represent adequately the regional variogram, the relative proximity of the ergodic and fluctuation variances of sills can be seen as an indicator of the agreement between the model fitted to the experimental variogram and the regional variogram, regardless of the size of the fluctuation variance. However, ergodic and fluctuation variances with similar values does not imply that, in practice, sill estimates will necessarily describe something that is relevant about the scales of variation of the regionalized variable within the sampling domain. As mentioned in the Introduction, the sills and functions of sills do not have an unambiguous meaning outside of the LMC. Therefore, a high fluctuation variance of sills necessarily implies that the parameters of the model are not micro-ergodic. In practice, this means that the probabilistic model that constitutes the LMC is likely not the tool of choice to separate multiple spatial scales of variation unless the sample data cover regions much larger than the scales of variation of the phenomena under study.

To what degree can sill estimates be used in practice to quantify their own uncertainty and that of derived functions? As shown by the theoretical equations above, sill matrices are a key component of the quantification of uncertainty in coregionalization analysis. In practice, these matrices can only be estimated, so that our ability to evaluate uncertainty from a real dataset is impeded by the uncertainty of sill estimates. The circularity of the problem therefore makes the precise evaluation of the structure-specific uncertainty from a real dataset very problematic, if not impossible. Alternatively, to investigate the uncertainty levels from a given sampling scheme and domain, the user could choose different combinations of theoretical parameters corresponding to plausible, best-case (e.g., short variogram range, one or two structures, no scale dependence) and worst-case (e.g., longer ranges, three or more structures, strong scale dependence) scenarios. From those, confidence intervals could be generated with theoretical calculations adjusted for the positive semidefiniteness constraint, and used to determine whether the uncertainty levels appear acceptable.

4.7.3 Closing remarks

Since its beginnings, the primary focus of geostatistical research has been on local estimation and the associated uncertainty of local estimates. In this framework, it is certainly sound to use nested variogram models in order to improve local estimation

through a better variogram fitting. In coregionalization analysis, however, sill estimates are submitted to a method of data analysis, which puts the variogram models, their interpretation and the uncertainty associated with the estimation of their parameters at the forefront. In a LMC, the variables are viewed as sums of spatially superimposed, random structural components characterized by variogram models. In this article, we have shown that a proper identification of the characteristics of such nested models is a difficult task. We have quantified and discussed the uncertainties for sills and functions of sills from bivariate data. All other things being equal, our results suggest that uncertainty in regionalized PCA, where structural correlations are used, could be lower than in regionalized RDA, where structural coefficients of determination are analyzed. Because of the high level of uncertainty that we found when the observed spatial patterns are assumed to result from superimposed random structural components, alternative approaches must be sought that identify scale-specific characteristics in multivariate spatial data sets, while reducing uncertainty. Nevertheless, the uncertainty obtained in the context of coregionalization analysis may prove acceptable when one variogram model (with or without a nugget effect) is sufficient to properly characterize the spatial structure of the dataset or when the available data is abundant and covers extensive regions (e.g., in remote sensing applications).

Acknowledgements

The computer resources necessary to complete this research were purchased thanks to a Canada Foundation for Innovation (CFI) grant. During his doctoral work, the first author was supported by scholarships from the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) and the Natural Sciences and Engineering Research Council of Canada (NSERC). The second and fourth authors acknowledge individual NSERC Individual Discovery grants.

Appendix 1

A relatively simple algorithm was developed to approximate the values of fluctuation covariances between a cross variogram involving variables i and j and another cross variogram involving variables i' and j'. This provides a numerical approximation to the integral in Equation (4.18).

For a given combination of distance lags h and h':

1. At each step (τ), simulate a pair of vectors, one of length h made of points $\mathbf{x}$ and

$\mathbf{x}_e$ ($\mathbf{x}_e = \mathbf{x} + h$) and one of length h' made of points $\mathbf{x}'$ and $\mathbf{x}'_e$ ($\mathbf{x}'_e = \mathbf{x}' + \mathbf{h}'$), each randomly located within the domain D.

2. From the simulated vectors $\mathbf{x}$, $\mathbf{x}_e$, $\mathbf{x}'$, $\mathbf{x}'_e$, evaluate $G^{\tau}_{ii'}, G^{\tau}_{jj'}, G^{\tau}_{ij'},$ and $G^{\tau}_{ji'}$, with

$$G^{\tau}_{ij} = [\gamma_{ij}(\|\mathbf{x}_e - \mathbf{x}'\|) + \gamma_{ij}(\|\mathbf{x}'_e - \mathbf{x}\|) - \gamma_{ij}(\|\mathbf{x}_e - \mathbf{x}'_e\|) - \gamma_{ij}(\|\mathbf{x} - \mathbf{x}'\|)]$$

where $\|\mathbf{x} - \mathbf{x}'\|$ represents the Euclidean distance separating points $\mathbf{x}$ and $\mathbf{x}'$.

3. Repeat steps 1 and 2 a large number (N_{sim}) of times.

4. The approximate value of the fluctuation covariance at distance lags h and h' is provided by

$$\mathrm{Cov}[\overline{\gamma^R_{ij}(h)}, \overline{\gamma^R_{i'j'}(h')}] \approx \frac{1}{4N_{sim}} \sum_{\tau=1}^{N_{sim}} G^{\tau}_{ii'} G^{\tau}_{jj'} + G^{\tau}_{ij'} G^{\tau}_{ji'}.$$

Fluctuation covariances approximated with this algorithm have the advantage of being totally independent of any sampling scheme within a given domain. For the results presented in this paper, we used $N_{sim} = 2.5 \times 10^6$.

Chapter 5

Inference in spatial observational studies

5.1 Introduction

Inference is at the core of any scientific research. Most science aims to identify from a given dataset, patterns that are consistent or generalizable. Inferential statistics is "the process of drawing conclusions about the nature of some system on the basis of data subject to random variation" (Upton and Cook, 2006). In classical non-spatial inferential statistics, the sample data is treated as a random outcome from a reference theoretical population. In what could be called an 'orthodox' view of inferential statistics, the properties of the sampled variables are used as estimators of the 'true' population parameters. For example, the sample mean is seen as an estimator of the 'true' population mean. Implicit in this view is the idea that the data are only an imperfect approximation of some idealized reality.

Matheron (1978, 1989) proposes a reversal of perspectives from this orthodox approach in his 'Estimating and Choosing', an essay on the objectivity of probabilistic statements in the geostatistical context. In Matheron's view, a strict adherence to the classical statistical approach can lead scientists to have more faith in the model than in the data. The data should be considered as the reference for truth, not the model. Models are subjective and should be considered 'chosen'. The choice of proper inferential (or probabilistic) models is particularly important and challenging since, underlying their use, comes a set of assumptions pertaining to the population that should be described. For Matheron, there is no randomness in nature; it lies in the models that we 'choose' to describe nature. In the same perspective, Webster (2000) asks 'is soil variation random?'. He concludes that soil variation is definitely not random but that we may not be far wrong when using models that assume that it is. The real concerns

should be about the appropriateness of our models for achieving specific purposes.

Unfortunately, the inference focus, or the exact view of the population that must be described in practice, is not always stated explicitely in scientific studies, or not even reflected upon. Practitioners often use statistical methods as 'ready-to-use' packages without examining their underlying assumptions or the inference focus implicit in them. Researchers should clarify their research objectives and understand properly the chosen models, rather than having a blind faith in them. The view of the population and the inference focus implicit in those models have to comply with the research objectives. We must thus strive to choose the appropriate models that will be suitable for achieving specific research objectives. Following the principles of Ockham's razor, or parsimony principle, these models should be as simple as possible, but as complicated as necessary.

The role of models and the question of infence is even more problematic in the context of spatial analysis. The classical paradigms of inferential statistics were developed in a context where spatial structure was ignored, or was considered a nuisance. In spatial analysis, the localized aspect of the sampling and the presence of spatial structuring render the application of the 'classical' paradigms of statistical inference difficult, if not problematic. In this Chapter, I demonstrate that these issues are particularly crucial in the context of spatial observational studies and multi-scale analysis and examine the specific aspects related to inference that are most important in this context. In a first part, I discuss the important concept of stationarity and differentiate the concepts of ergodicity and micro-ergodicity. In the second part, I illustrate the importance of those concepts using the example of the simple linear correlation coefficient. In the third part, I summarize the important results of previous chapters and discuss the question of inference in regionalized multivariate analysis. I then describe a new approach for multi-scale analysis, coregionalization analysis with a drift (CRAD), that was developed to provide a coherent and consistent alternative to existing methods.

5.2 Stationarity, erdogicity and micro-ergodicity

Statistical concepts such as stationarity and ergodicity are often misunderstood. For one, they are intimidating concepts that are not always directly relevant to the applications of the methods in practice (Chilès and Delfiner, 1999). Also, such statements as "non-stationary data" are often found in the literature and can be quite misleading. It must be emphasized here that these concepts characterize the statistical models, not the data. In Matheron's sense, we generally *choose* models that are stationary and ergodic to describe the data. Asking whether a given dataset is stationary or ergodic

is simply senseless (Myers, 1989).

Stationarity is a property of random function models defined in an infinite space. For a second-order stationary random function, the expected value at any location is equal to the mean of the random function. Likewise, the covariance between values of the random function at two spatial locations is only a function of the distance separating them (Chilès and Delfiner, 1999, p. 16). In practical situations, if we choose to describe a given dataset with a stationary random function, it has to be reasonable to assume that the studied phenomenon is relatively homogeneous in space. Its mean value has to be relatively constant over the study area and its spatial structure has to be reasonably repeated within the study area. Such assumptions are clearly unreasonable if large scale spatial structures or localized peaks are perceived in the data or if the types of spatial structuring are perceived as being clearly different in different regions of the study area.

Ergodicity is also a property of models. It assures a coherence between the sample statistics evaluated from realizations and the population parameters. With standard (non-spatial) statistical models, the statistics evaluated from an infinite number of values from a given realization generally correspond with the parameters of the random variable (their expected values), regardless of the zone covered by the realization. Such a model can thus be said to be ergodic. For example, the mean obtained from the infinite number of values for a given realization derived from an identically and independenly distributed random variable model (i.i.d. or nugget effect model) is exactly equal to the mean parameter of the random variable. This is because each value of the random variable is independent from the other values, regardless of its location. So even within a bounded sampling domain, such models are internally consistent and ergodic.

Stationary random functions defined in an infinite sampling domain are also generally ergodic (although there can be exceptions, Chilès and Delfiner, 1999). If the size of the sampling domain in which realizations of a random function are found is increased until it becomes infinite (increasing domain asymptotics), the mean obtained from the realizations will converge to the mean of the random function and the average covariance between values obtained from the realizations will converge to zero.

In practice, however, we generally only have access to one bounded sampling domain and the notion of ergodicity may not be very useful. Matheron (1971) coined the term micro-ergodicity for discussing what happens when the ergodicty concept is applied to random functions within a limited sampling domain. To understand this simply, we must appreciate that it is impossible to fit an infinite number of spatial features with a non-negligible size within a fixed area. Therefore, the characteristics of the combination of the spatial features within the area will differ, although perhaps only

slightly, from the parameters calculated from the infinite number of spatial features. We can easily recognize that this difference will be stronger for larger, and potentially less numerous, spatial features than for smaller, and therefore potentially more numerous, spatial features. Imagine an infinite number of marbles of different sizes forming a 'population', with a size distribution following a fixed probabililty distribution. Now imagine that we randomly select enough marbles to fill a bowl of a specific size. The bowl will contain marbles of various sizes. The average size of the marbles within the bowl will always be different, perhaps only slightly, from the mean size given by the probability distribution. Doing this for multiple bowls, we can calculate the average difference between the mean size of the marbles within the bowls and the mean of the infinite number of marbles. This difference will be higher if the marbles are, on average, larger and if the bowls are smaller. A model is said to be micro-ergodic if the statistics measured from realizations within a limited sampling domain (the marbles filling the bowl) are always equal to the model parameters. If it is micro-ergodic, it is then necessarily ergodic. Just like the marbles in the bowl, a random function model limited within a bounded sampling domain cannot be strictly micro-ergodic.

In Chapter 4, the notion of fluctuation variance was used to highlight that if variogram model parameters are ergodic they are generally not micro-ergodic. Even if an infinite number of samples within the bounded sampling domain were taken, some uncertainty would still remain on the variogram model parameters. This was shown by the fact that a good proportion of the ergodic variance (variance of the difference between the estimates and the random function parameters) is actually composed of fluctuation variance (variance of the difference between the estimates obtained on realizations and the random function parameters). Matheron (1978, 1989) recognized this fact but emphasized that the parameters that pose problems are the ones defining the behaviour of the variogram function at large distances. At small distances, the variogram model parameters have a much smaller fluctuation variance and they could be considered micro-ergodic to a reasonable degree. Matheron used this to justify the use of the random function approach for local estimation purposes, even in the context of a limited sampling domain. He noted that it is the parameters of the variogram at short distances that have the greatest impact on the estimation uncertainty. Those same parameters also happen to be micro-ergodic.

The concept of micro-ergodicity highlights an interesting contradiction. Micro-ergodicity is defined by applying the notion of egodicity within a bounded sampling domain. Model parameters that are reasonably micro-ergodic are more robustly estimated within the sampling domain because they are obtained from more repetitions (more marbles in

the bowl). It is therefore more likely that they correspond to actual patterns that extend beyond the limits of the domain. Thus, by choosing models with micro-ergodic parameters to describe a dataset, we obtain, in practice, parameter estimates that are likely more generalizable outside of the limits of this domain.

5.3 Assessing correlations within a bounded sampling domain

Imagine that we wish to describe the correlations between two sampled variables at the scale of the sampling domain. This amounts to determining the average correlations between the sample values at any location within this given domain. In geostatistical terms, the focus of inference in this case are the properties of the regionalized variables, which is a conceptualization of continuous variables within a localized area (the domain). For example, the total amount of iron ore in a given deposit, or carbon in an agricultural field, represents the total pool of the regionalized variable. In this view, there is nothing beyond the limits of the domain. The variable corresponds to a physical reality that could be completely accessed if the field was completely sampled and it is independent of any model (Matheron, 1970, 1978). This reference point is understandable in the mining and geological context where geostatistics were developed, since ore deposits are always limited in extent and mining companies don't have access to unlimited territories for mineral exploration. It is often referred to as the estimation of spatial means. Because it is considered continuous, there is an infinite number of possible values for the variable, even within the limited extent of the study area. Therefore, any measure characterizing regionalized variables has to be defined with spatial integrals. For example, the average linear correlation between two regionalized variables i and j would be given by:

$$\overline{corr^R(z_i, z_j)} = \frac{1}{D}\int_D \frac{(z_i(u) - \overline{z_i^R})(z_j(u) - \overline{z_j^R})}{\sqrt{(z_i(u) - \overline{z_i^R})^2 (z_j(u) - \overline{z_i^R})^2}} du$$

where $z_i(u)$ is the value of variable i at location u and $\overline{z_i^R}$ is the mean of the regionalized variable j over the field of area D ($\overline{z_i^R} = \frac{1}{D}\int_D z_i(u)du$).

Obviously, it is impossible in practice to have access to the regionalized variable completely. This would involve sampling the entire volume of the domain on a punctual support (infinitely small points). Therefore, the properties of regionalized variables have to be estimated from fragmentary samples. Alternatives for achieving such a task are

examined below.

The first option, the simplest one, is to follow a completely deterministic approach in which the sampling scheme is assumed to exhaustively capture the variability within the field. Descriptive statistical measures are then calculated without recourse to probabilistic models. For example, the simple linear correlation between sample values is

$$\overline{corr(z_i, z_j)} = \frac{1}{N} \sum_{u=1}^{N} \frac{(z_i(u) - \overline{z_i})(z_j(u) - \overline{z_j})}{\sqrt{(z_i(u) - \overline{z_i})^2 (z_j(u) - \overline{z_i})^2}}$$

where $z_i(u)$ is the value of variable i at sampling location u and $\overline{z_i}$ is the sample mean for a sample of size N. The results obtained with this approach are entirely dependent on the adequacy of the sampling design. However, no assumptions of independence between samples are required, and thus, spatial autocorrelation has no effect. The type and scale of spatial variability found within the sampling domain have no effect on the estimation of this model parameter. Discussions pertaining to stationarity, normality or ergodicity are simply irrelevant in this case.

These discussions become important if we hope to obtain uncertainty measures on the correlation coefficients describing how far the statistics calculated from the fragmentary data are to the same statistics measured on the regionalized variables. This is the realm of inferential statistics. Uncertainty describes how far sample statistics calculated from the data are, on average, to the 'real' population parameters, as seen through the model. Since the population to describe is tied to a model, uncertainty is also necessarily tied to a model. Uncertainty measures are thus greatly dependent on the proper choice of models to perform specific tasks. For example, we can choose to obtain a probability of significance of the correlation coefficients by using the 'Student' t-test. For doing this correctly, we would need to assume that the sample is random and i.i.d. (not spatially structure; nugget effect). Strictly speaking, this model is internally consistent because it is ergodic and micro-ergodic (see above). However, its assumptions are clearly not appropriate for characterizing spatially structured data. The presence of spatial structure in the variables makes the assumption of independence between sample data unreasonable.

Alternative tests have been devised that account for the spatial structure (autocorrelation) in the variables. For example, Dutilleul's modified t-test (Dutilleul, 1993) obtains the probability of significance of the correlation between variables by adjusting the sample size to account for autocorrelation. The assumptions of such models are certainly more reasonable with spatially structured data. However, their focus of inference is the characterization of the properties of the random function, instead of

the properties of a realization limited within the sampling domain, it. As we have seen above, a stationary random function model is ergodic but not micro-ergodic (except the nugget effect mode). In fact, through infill asymptotics (by letting the number of samples become infinite within a fixed area), the probability of significance of the 'Student' t-test will converge to zero as the number of samples increases. However, by accounting for the effect of the spatial structure, the probability of significance obtained with the modified t-test will converge, on average, to a fixed non-zero value.

Within a bounded sampling domain, we are thus faced with having to choose one model that has unreasonable assumptions but that is internally consistent, or one that has reasonable assumptions but that is not internally consistent. This does not mean that either or both approaches are wrong. In the context of a bounded sampling domain presented in this section, significance testing is perhaps not even necessary. Switching the focus of inference to the generalization of patterns outside of the limits of the sampling domain, the modified test would become the obvious choice.

An alternative would be the use of the design-based approach (de Gruijter and ter Braak, 1990; Brus and de Gruijter, 1997; Gregoire, 1998), in which a probabilistic model is built by assuming that the sampling was performed randomly, while the values of the variables, and the populations, are regarded as fixed. Subsamples are repeatedly taken from the total set of samples to generate a reference distribution (Gregoire, 1998). The probabilities of significance and uncertainty measures are derived by using this reference distribution as a benchmark. The design-based approach can be considered micro-ergodic, leading to coherent uncertainty measures consistent with the focus of inference. However, the design-based approach cannot be extended to perform spatial or multi-scale analysis in the context of spatial observational studies, so it will not be further discussed here.

5.4 Inference in coregionalization analysis

When dealing with a single dataset, the definition of spatial structure and the decomposition of spatial variables into multiple scales of variation is not unique and is inherently subjective, each method of analysis carrying with it a different conceptualization of scale and a different approach for separating scales (Chapter 2). Thus, on top of the localized aspect and the presence of spatial structuring, the question of inference in multi-scale analysis is made even more problematic by the subjectivity of the scale concept.

As explained in Chapter 2, the scales depicted by experimental variograms do not correspond to scales of variation. Each lag contains contributions from spatial features

of all sizes located within the sampling site. Withouth modelling, a decomposition of the variation in the total variable into scales of variation cannot be obtained. Thus, although for a trained eye, experimental variograms may give a wealth of qualitative information concerning the spatial structure of each variable and the spatial relationships between them, they are not suited for multi-scale analysis.

In Chapter 2, it was shown that a correspondence can be made between variogram functions derived from the modelling of the experimental variogram and scales of variation. However, it was also shown that the range a variogram functions does not correspond to a single scale of spatial variation. Each function actually represents a probability distribution of scales of variation (Fig. 2.20 on page 54) with a peak value smaller than the range. Different types of variogram functions with the same range have different probability distributions. Accordingly, the range is not an adequate descriptor of the scales of variation found in a variable. In the context of spatial observational studies and multi-scale analysis, the choice of variogram model as a whole has important implications for the results of the analysis. They are even more crucial for the application of the LMC, or when superposed variogram functions are used to characterize the total variables.

In Chapter 3, it was shown that the properties of kriged regionalized components are greatly dependent on, but differ from, the variogram model parameter estimates. Due to the smoothing effect of kriging, regionalized components for different structures are not orthogonal. Also, they generally contain less small scale spatial variation than realizations derived from their corresponding variogram functions, and the correlations between kriged regionalized components for different variables are not equal to the correlations obtained from sill matrix estimates. Uncertainty measures were thus obtained on kriged local estimates corresponding to regionalized components, thereby allowing the conditional Gaussian co-simulation of regionalized components that have properties consistent with the variogram model parameters. These approaches are appropriate insofar as the application of the LMC is considered appropriate.

In Chapter 4, it was made evident that the problems related to micro-ergodicity are aggravated when multiple variogram functions are superimposed. When two or more spatial variogram functions are used to characterize a variogram, it was shown that if the model chosen to describe the data (LMC) happened to be the right one, we wouldn't be able to estimate its parameters properly. That is, even if the sampling intensity was extremely large. This was highlighted by the high estimation variances when two spatial variogram functions are used. We can only hope to retrieve the 'true' model parameters with reasonable uncertainty if the sampling domain extent is extremely

large when compared to the variogram ranges. Only then can the fluctuation variances in the LMC attain reasonably low values. The factors that have greatest effects on the fluctuation variance are the number of variogram models and the ratio of variogram range to domain extent. The sampling density had a greater effect on the estimation variance allowing, most notably, a better estimation of variogram model parameters at short distances.

The issues pertaining to micro-ergodicity and uncertainty are even more problematic for the correlation coefficients and coefficients of determination used to assess the relationships between variables. Results not presented in Chapter 4 and my own experience with the use of the LMC with real datasets showed that, with multivariate datasets, the estimated coregionalized matrices are often poorly conditioned and are very close to not being positive definite. This is especially problematic when the number of variables is large. As an effect of this, coefficients of multiple determination in a multiple regression or redundancy analysis involving multiple independent variables are often very close to one. This effect is caused in part by the high uncertainty levels associated with the LMC. Also, if the studied variables differ in spatial structuring and scales of variation, their experimental direct and cross variograms will have varied shapes. When fitting the variogram models, this may lead to the LMC algorithm more frequently having to force coregionalization matrices to be positive definite resulting in a poorly conditioned matrices. This effect could be seen as a symptom of the inadequacy of the LMC model for a given dataset.

The model that underlies coregionalization analysis (the LMC) is thus not appropriate for multi-scale analysis in many circumstances. The most important assumptions underlying its use, such as stationarity and superimposed spatial processes with multiple scales of variation, are often not reasonable and it is not internally consistent as highlighted by the lack of micro-ergodicity of the model. In most geostatistical applications, such as the characterization of ore deposits, problems related to micro-ergodicity are ignored because the variogram model parameter estimates are ultimately used to perform local estimation. By opposition, the lack of micro-ergodicity in the LMC is a major concern for multi-scale analysis since variogram model parameters are, ultimately, the focus of inference.

5.4.1 Coregionalization analysis with a drift

Coregionalization analysis with a drift (Pelletier et al., 2009a,b, Appendix) was developed to overcome the shortcomings of the LMC and to obtain an approach for multi-scale analysis that is both coherent and internally consistent. The basic princi-

ple is to use a combination of deterministic and probabilistic approaches, using each where they are most appropriate. As was seen in Chapter 2, drift estimation procedures are appropriate and flexible for capturing variation at larger scales. Orthogonal decomposition methods were equally efficient on regular transects and grids, but their application to irregular datasets was more problematic. Conversely, the application of the random function approach is more problematic at larger scales due to uncertainty and microergodicity problems. As discussed in Chapter 2, through variogram modelling of the nugget effect, the LMC is the only approach that provides a coherent procedure for separating the non-spatially structured variation from the small scale variation.

Coregionalization analysis with a drift combines a random function approach (LMC) for capturing small scale and non-spatial variation and a deterministic approach for capturing larger scale variation. In the CRAD model, the total variable is described as

$$Z = m + \sum_{s=1}^{S} Z_s$$

where m is a deterministic drift (trend) component and Z_s ($s = 1, ..., S$) are random functions associated with different variogram functions. For parsimony and for avoiding unreasonable assumptions related to the superimposition of multiple spatial random functions, the number of random components is generally restricted to two: one nugget effect component and one spatially structured component (e.g. spherical, exponential, Gaussian). In a first step in CRAD, each sampled variable is decomposed into a deterministic drift and a residual component (Pelletier et al., 2009a). As described in Chapter 2, the deterministic drift component can be obtained by adjusting global functions, such as low order polynomials, or by using a local procedure where the drift is estimated at each sampling location using a moving window. The local means, for the local procedures, and the parameters of the polynomial model, for global procedures, are estimated for each variable separately by iterative Generalized Least Squares. In this way, the spatial stucture of residuals and the spatial configuration of the sampling locations are taken into account when obtaining weights. In the second step, an LMC is fitted to all combinations of direct and cross variograms involving residuals of all variables jointly (Pelletier et al., 2009b). Ranges and sills of the variogram functions are estimated by iterative Generalized Least Squares, yielding estimated sill matrices associated with the non-spatial (nugget effect) and small-scale spatial components (e.g. spherical model), as well as an estimated range for the spatially structured variogram function that is common to all variables.

The simulation experiments by Pelletier et al. (2009a) have shown that local proce-

dures are more efficient at retrieving complex, patchier, drifts than the global procedures. In a univariate context, they have shown that parameters of the spatial autocovariance of the residuals were retrieved with good success. Multivariate coefficients in a bivariate (simple correlation or regression) or multivariate (redundancy analysis) analysis also had lower uncertainty in CRAD when compared with CRA (Pelletier et al., 2009b).

The method of conditional Gaussian co-simulation of regionalized components presented in Chapter 3 can be used in CRAD to obtain spatial representations that have properties corresponding to the variogram model characterizing the small scale spatially structured component, and to the nugget effect.

Coregionalization analysis with a drift is a method of analysis that is coherent and internally consistent, and that has reasonable assumptions in the context of the multiscale analysis of spatial observational studies. By restricting the application of the probabilistic approach to small scales, the model underlying CRAD at small scales is both ergodic and reasonably micro-ergodic. Thus, small scale patterns obtained in this manner are likely to be more generalizable outside of the limits of the sampling domain. In contrast, characteristics of the larger scale components are seen as being limited to the sampling domain. They are thus coherently captured by a deterministic model that takes into account the probabilistic nature of the model used for small scale components in its adjustment by Generalized Least Squares.

5.5 Concluding remarks

Statistical models should not be regarded as being true or false, nor right or wrong, but should be regarded as being useful, or not, for achieving specific objectives. Practitioners generally choose statistical methods that will allow them to attain their objectives, but not enough attention is generally paid to the assumptions underlying inferential statistical models. Probabilistic models, which are at the core of inferential statistics, have a focus of inference that may not correspond to the practitioner's objectives. In this Chapter, I have demonstrated that this issue is of significant importance in the context of spatial observational studies, even for simple statistics such as the linear correlation coefficient.

Matheron (1978) argues that although statistical models are always subjective and should be considered 'chosen', we must attempt to minimize the effect of their choice on the final results of the analysis. Through my experiences with the use of a LMC involving more than one spatial variogram functions with real datasets, I noticed that

small changes in the model choices (variogram functions, variogram range, lag distances, etc.) often had large effects on the final results of the analysis. From the discussions in this chapter and the preceding one, it appears that this 'instability' of the LMC can be attributed to the interrelated issues of uncertainty and inference. Since the LMC has consistently high levels of uncertainty with realistic scenarios and sampling designs, its focus of inference and assumptions are likely not appropriate for practical applications. Despite its extensive use in several fields of studies, these questions had not been thoroughly examined in the literature on coregionalization analysis.

CRAD was developed with the objective of providing a method for multi-scale analysis that would be less dependent on model choices and that would have reasonable assumptions. Altough CRAD was shown to be coherent and internally consistent in this chapter and in Pelletier et al. (2009a,b), we still need to assess whether the method can be useful and adaptable in practical applications. The next two chapters provide examples of the use of CRAD in agronomic and forest ecology applications.

Chapter 6

The issue of scale in site-specific agricultural management of soil nutrients: a case study

Site-specific agricultural management (SSM), also called precision agriculture, is the management of soil fertility and agricultural crops at a scale smaller than that of the whole field (Plant, 2001). Several factors explain the rapid development of SSM in recent years. Firstly, technological advances such as combines equipped with yield monitors, global positioning systems, variable-rate fertilizer systems and high-resolution sensors of various kinds are allowing unprecedented capabilities for mapping soil and crop variability and monitoring changes through time. Secondly, due to financial pressures, producers are increasingly interested in minimizing fertilizer inputs while maximizing crop yield. Thirdly, ground water pollution, greenhouse gas emissions as well as other environmental problems have brought strong incentives to reduce nutrient burden in agricultural soils.

Although technologies for mapping the within-field variability in yield and for the spatial application of fertilizers are already available to producers for most crops and in most regions of the world, other conditions must be met if SSM is to become widely accepted. First, the economic and environmental benefits of using SSM must outweigh the costs. Significant increases in yield or decreases in fertilizer inputs must result from the change in strategy compared to the uniform management system. Plant (2001) outlines three criteria that must be satisfied in order to justify this change to SSM: 1) significant within-field spatial variability must exist in factors that influence crop yield, 2) it must be possible to identify and measure causes of the variability and 3) the information from these measurements must be useable to modify crop management practices to increase profit or reduce environmental impact. This logic follows very

closely that of Whelan and McBratney (2000).

Even if the soil properties and abiotic factors responsible for the within-field patterns in yield have been identified, not all of those causal factors can be managed. Very few studies have been published that distinguish the manageable variables from the unmanageable and the conditions under which variation can or cannot be controlled. In industrial agriculture, the soil properties that are normally managed include: N, P and K through fertilizer inputs; soil organic matter, through manure amendments; and soil pH, partially controllable with lime applications. However, it is very clear that a single site-specific application of inputs cannot completely control within-field variability. It is well known that controlling variables like soil pH and soil organic matter can only be done partially and in the long term. If yield is limited by soil water content, soil texture or topography, it will be very hard, if not impossible, to develop a SSM plan to have an effect on yield.

Although it has received little explicit attention in the literature, the issue of scale is of central importance for assessing the viability of SSM. Cahn et al. (1994) outline that the benefits of variable rate fertilizer applications may be highest for nutrients exhibiting larger scale variation and spatial patterns that are temporally correlated. The high cost of soil sampling and analysis renders the generation of detailed soil maps impractical. Soil maps can be generated from fewer sampling points for soil properties with smoother spatial patterns and larger scale variation. Also, it can be hypothesized that soil nutrients with larger patches or zones are easier to manage with modern farming equipment than nutrients with smaller patches. Soil nutrients that have spatial patterns that are consistent from year to year will be easier to manage than those that are temporally heterogeneous.

The use of 'indicator' variables that are easy to measure has also received a lot of attention in SSM. For example, electrical conductivity can be measured with on-the-go sensors and has been shown to be correlated to a number of soil properties (Johnson et al., 2001; Heiniger et al., 2003). The use of remote sensing (Moran et al., 1997) and topography (Fraisse et al., 2001) for delineating zones for SSM have also been examined.

By its very definition, the application of SSM is restricted to the management of within field variation. Thus, the questions pertaining to assumptions and inference within a bounded sampling domain described in Chapter 5 are particularly relevant for the choice of the statistical methods for data analysis. In this case study, a dataset collected in two adjacent agricultural fields will be used to study the within- and between-year relationships between soil nutrients and the relationships between scales of variation and temporal heterogeneity. The field-scale relationships will be studied

and contrasted with the within-field spatial relationships. The CRAD method will be used for multi-scale analysis to identify variables that contain variation at a manageable scale and that are temporally correlated. The usefulness of electrical conductivity measurements from a VERIS sensor and elevation obtained from a digital elevation model as indicator variables for other nutrients will also be addressed.

6.1 Study area and dataset

The field study was carried out in Ste-Brigide, Québec, Canada (73° 03', 45° 20' N), in two adjacent agricultural fields with similar management history. The two fields had been under a corn–barley–soybean–alfalfa crop rotation for several years, and had been fertilized with spatially uniform applications of inorganic fertilizers and dairy cattle manure. The crops in 2000, 2001 and 2002 were soybean, corn and corn, respectively. In 2000, the fields received about 30 $^{t}/_{ha}$ of solid dairy manure. Basic field operations were standard: the fields were moldboard plowed every year when the fields were in annual crops. Inorganic nitrogen, phosphorus and zinc were applied at seeding and urea was applied later in the growing season.

The soil sampling grid was designed as a combination of a staggered regular grid (266 points) with spacing of 20 m in the width orientation of the fields and a random grid (58 points) with points at various distances from their nearest neighbors on the regular grid, providing a total of 324 points (Fig. 6.1). In November 2000, the spatial coordinates of grid points were recorded with a Global Positioning System (GPS) with differential corrections from a radio beacon. In June 2002, the GPS was used to navigate back to the same coordinate locations. For each of the two years, a composite soil sample comprising three sub-samples within an area of a 1-m^2 surrounding each point was taken from the top 15 cm. A moist soil subsample was taken and stored at 4 °C until analysis for inorganic nitrate using a KCl extraction procedure (Maynard and Kalra, 1993; now referred to as NO_3). The remainder of each sample was air-dried and ground to pass through a 2-mm mesh sieve. Extractable phosphorus, potassium, magnesium, calcium and aluminium levels (now referred to as P, K, Mg, Ca and Al) were determined from soil samples using the Mehlich-III extractant (Sen Tran and Simard, 1993; P was determined with Lachat autoanalyzer, other nutrients with atomic absorption flame emission). Percent organic matter (OM) was determined by loss on ignition (Tiessen and Moir, 1993). Soil sand, silt and clay content were measured by the hydrometer method (Sheldrick and Wang, 1993). To avoid collinearity, only sand content was included in statistical analyses. In the spring of 2001, a Veris 3100 on-the-go

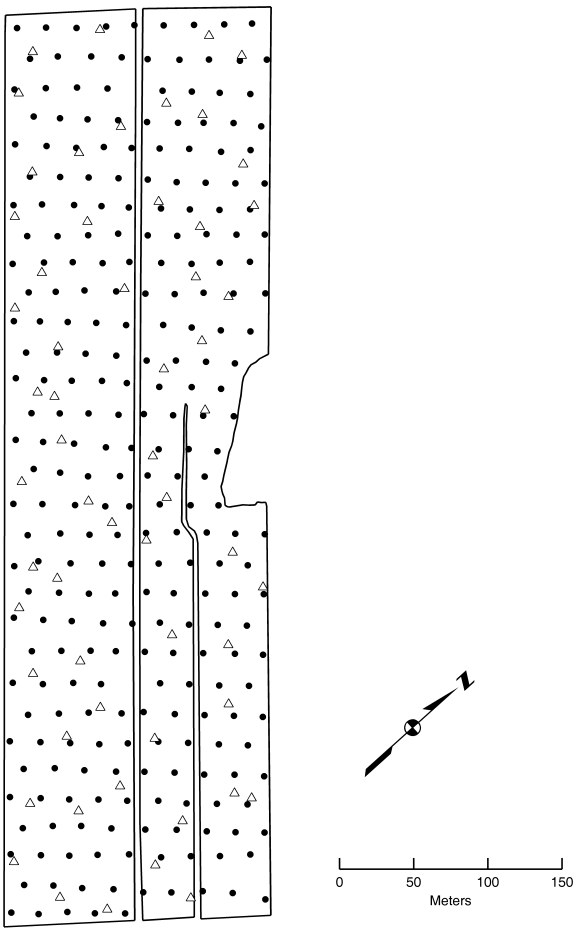

Figure 6.1: Map of the sampling scheme, made of the combination of 266 points on a staggered regular grid (•) and 58 randomly located points (Δ).

sensor linked to a differential GPS for intensive measurements of electrical conductivity was used to sense the whole study site in the spring of 2001, giving a total of 4491 electrical conductivity readings. The average electrical conductivity value (EC) of all Veris sensing points within a 5-m radius around each point on the soil sampling grid was used in the analysis. Elevation values were surveyed at 457 locations regularly spaced throughout the field with the use of a 'total station'. A digital elevation model with grid points every meter was generated with spline interpolation from the surveyed values (Fig. 6.2). Then, each location on the soil sampling grid was attributed the closest value on the elevation model (Elev).

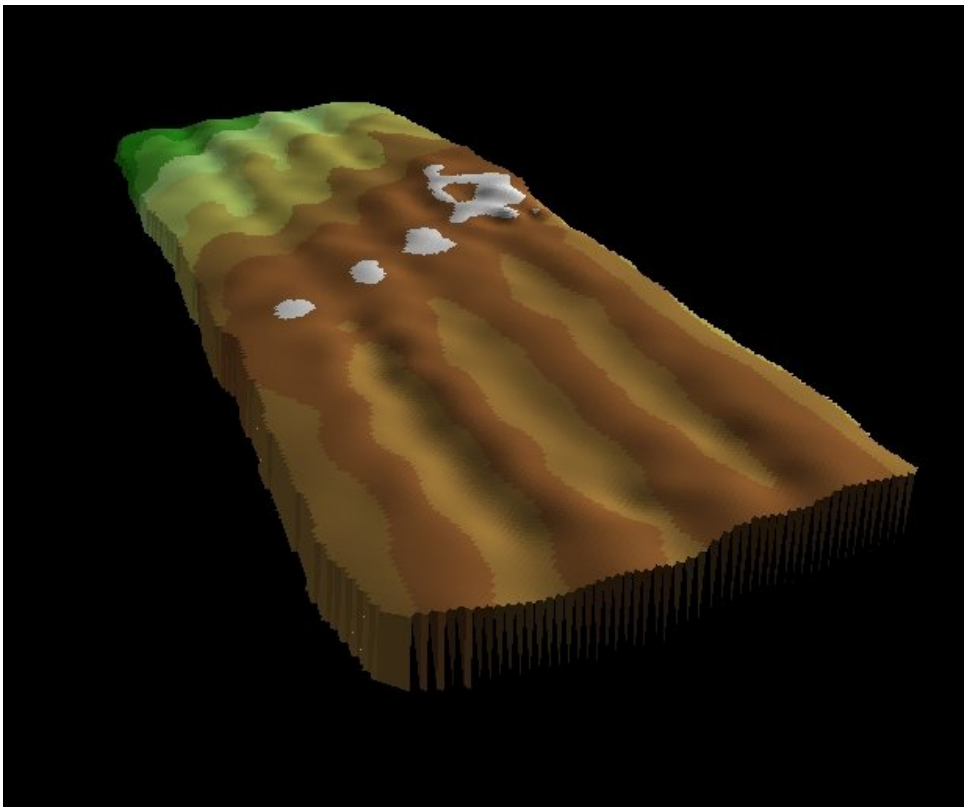

Figure 6.2: Three-dimensional perspective of the digital elevation model of the two adjacent agricultural fields looking North-North-East. The highest difference is 5 m between low elevations (green) and higher elevation (gray).

6.2 Statistical analysis

The simple linear correlations for soil properties were calculated between variables sampled in 2000, as well as EC (Table 6.1), and between variables sampled in 2000 and the same variables sampled in 2002 (Table 6.2).

All possible combinations of experimental direct ($\hat{\gamma}_i(h)$) and cross variograms ($\hat{\gamma}_{ij}(h)$) were calculated at distances from 0 to 240 m with lag distances every 15 m. Experimental direct variograms for soil variables sampled in 2000 are shown in Fig. 6.3) and associated cross variograms are shown in Fig. 6.4. To facilitate the interpretation of the strength of relationships, hulls of perfect correlations, calculated as $\pm\sqrt{\hat{\gamma}_i(h)\hat{\gamma}_j(h)}$, are also shown. These represent the highest and lowest possible values of the cross semi-variances possible between two variables, given the direct variograms for each variable.

The CRAD method (Pelletier et al., 2009a,b, Chapter 5) was used to perform multi-scale analysis on all soil variables using the local drift estimation procedure with a linear

Table 6.1: Simple linear correlation coefficients between variables sampled in 2000.

	pH	**OM**	**P**	**K**	**Mg**	**Ca**	**Al**	**NO_3**	**Sand**	**EC**	**Elev**
pH	1	-0.14	-0.11	-0.38	0.49	0.72	-0.62	-0.27	-0.01	0.31	0.07
OM		1	0.32	0.28	-0.08	0	0.22	0.39	0.12	-0.09	0.26
P			1	0.17	-0.24	-0.09	0.04	0.3	0.09	0.00	0.23
K				1	-0.13	-0.1	0.44	0.46	-0.08	-0.03	-0.11
Mg					1	0.76	-0.7	-0.15	-0.37	0.57	-0.28
Ca						1	-0.65	-0.11	-0.40	0.58	-0.31
Al							1	0.26	0.22	-0.48	0.09
NO_3								1	0.14	-0.04	0.13
Sand									1	-0.42	0.74
EC										1	-0.35
Elev											1

Table 6.2: Correlation coefficients between years.

	Correlation Coefficients
pH-pH	0.8495
OM-OM	0.4441
P-P	0.7533
K-K	0.5755
Mg-Mg	0.8158
Ca-Ca	0.8572
Al-Al	0.9229
NO_3-NO_3	0.2018

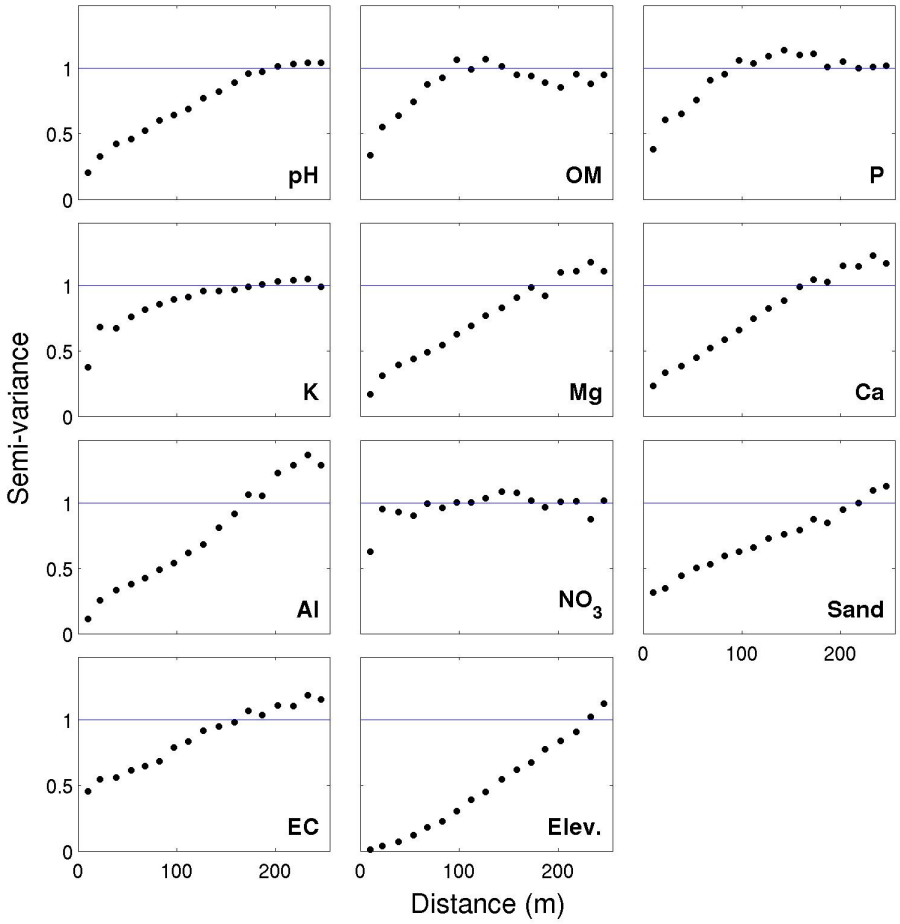

Figure 6.3: Experimental direct variograms for total soil variables sampled in 2000. The blue line is the sample variance (1 for standardized variables)

function fitted within the windows. The window sizes used for all variables were selected in values between 100 and 300 m in radius. Variograms of residuals were modeled with a spherical model with an estimated range of 35 m and a nugget effect. The modelled variograms of residuals for variables sampled in 2000 are shown in Fig 6.5. Maps of small scale (spherical model) regionalized components obtained by kriging (Chapter 3) and maps of large scale (drift) components represented with delaunay triangulation for variables sampled in 2000 are shown in Figs 6.6 and 6.7. Scale-specific correlation matrices corresponding to the nugget effect, the spherical model and the drift (pseudo correlations in this case) are shown in Fig. 6.3. Figure 6.8 illustrates the proportion of variance at each scale for the same variables sampled in 2000 and 2002 as well as the covariance between those variables.

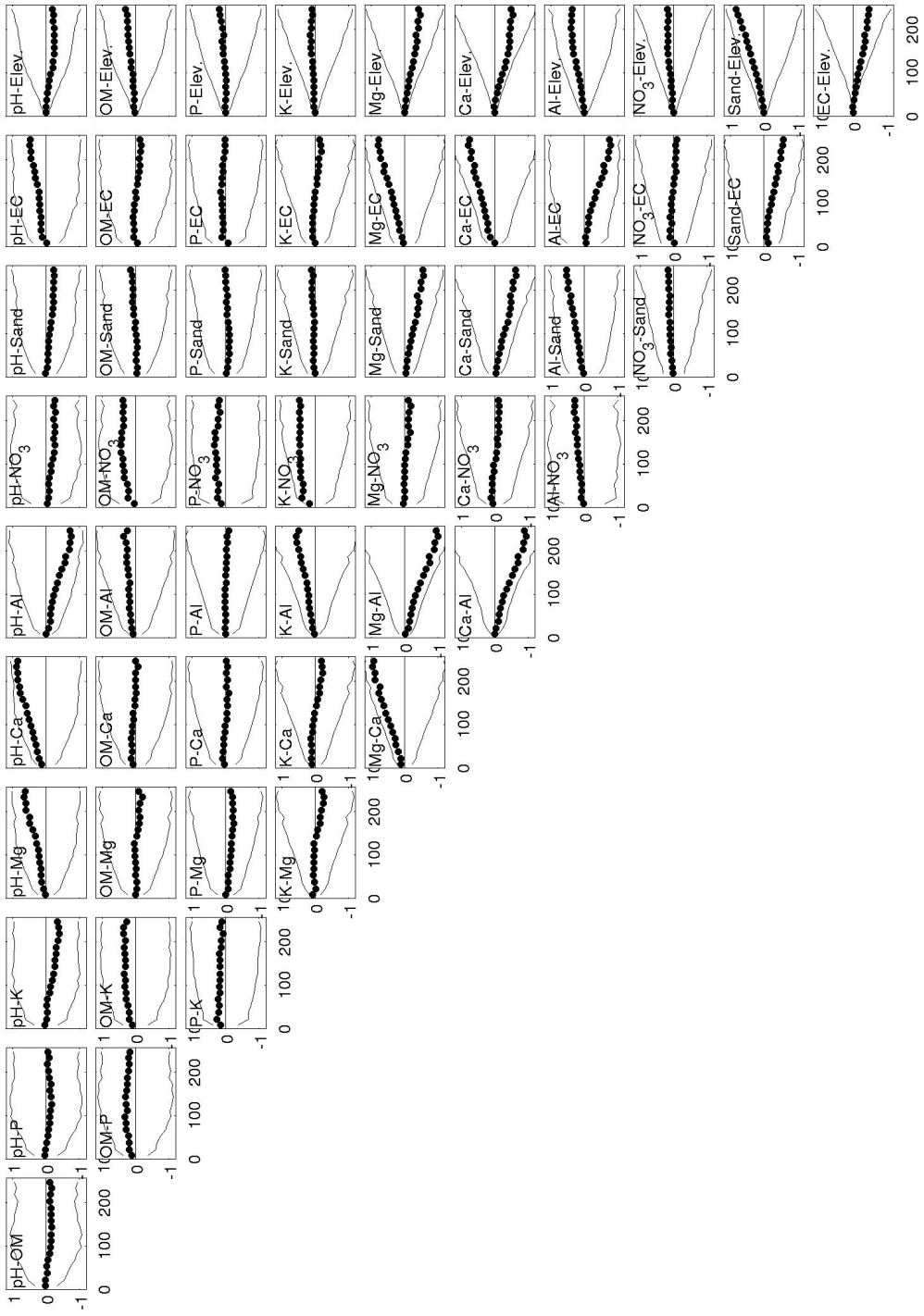

Figure 6.4: Experimental cross variograms between total soil variables sampled in 2000.

Table 6.3: Scale-specific correlation matrices obtained with CRAD for soil variables sampled in 2000 and electrical conductivity. (-) indicates the absence of correlations due to the lack of variation at this structure for the variable.

Nugget effect

	pH	**OM**	**P**	**K**	**Mg**	**Ca**	**Al**	**NO_3**	**Sand**	**EC**	**Elev**
pH	1	0.33	-	0.37	-	0.63	-	-0.19	0.23	-0.43	-
OM		1	-	0.34	-	0.32	-	0.21	-0.26	-0.32	-
P			-	-	-	-	-	-	-	-	-
K				1	-	0.59	-	0.53	-0.02	-0.08	-
Mg					-	-	-	-	-	-	-
Ca						1	-	0.32	0.18	-0.43	-
Al							-	-	-	-	-
NO_3								1	0.02	-0.04	-
Sand									1	-0.48	-
EC										1	-
											-

Spherical

	pH	**OM**	**P**	**K**	**Mg**	**Ca**	**Al**	**NO_3**	**Sand**	**EC**	**Elev**
pH	1	0.06	0.07	-0.12	0.27	0.68	-0.25	0.07	-0.2	0.55	-
OM		1	0.43	0.3	0.07	0.4	0.31	0.76	-0.07	0.51	-
P			1	0.43	0.03	0.28	-0.01	0.61	-0.22	0.21	-
K				1	0.2	0.14	0.27	0.54	0.09	0.24	-
Mg					1	0.54	-0.42	0.22	-0.3	0.36	-
Ca						1	-0.25	0.28	-0.41	0.74	-
Al							1	0.33	0.27	-0.19	-
NO_3								1	0.07	0.41	-
Sand									1	-0.03	-
EC										1	-
Elev											-

Drift

	pH	**OM**	**P**	**K**	**Mg**	**Ca**	**Al**	**NO_3**	**Sand**	**EC**	**Elev**
pH	1	-0.53	-0.24	-0.82	0.68	0.78	-0.83	-0.84	0.1	0.34	0.10
OM		1	0.32	0.49	-0.46	-0.5	0.5	0.73	0.23	-0.55	0.22
P			1	-0.1	-0.55	-0.59	0.25	0.38	0.65	-0.5	0.53
K				1	-0.47	-0.49	0.79	0.58	-0.37	-0.19	-0.37
Mg					1	0.96	-0.89	-0.78	-0.43	0.85	-0.36
Ca						1	-0.85	-0.84	-0.46	0.78	-0.42
Al							1	0.78	0.07	-0.66	0.03
NO_3								1	0.19	-0.65	0.15
Sand									1	-0.66	0.96
EC										1	-0.61
Elev											1

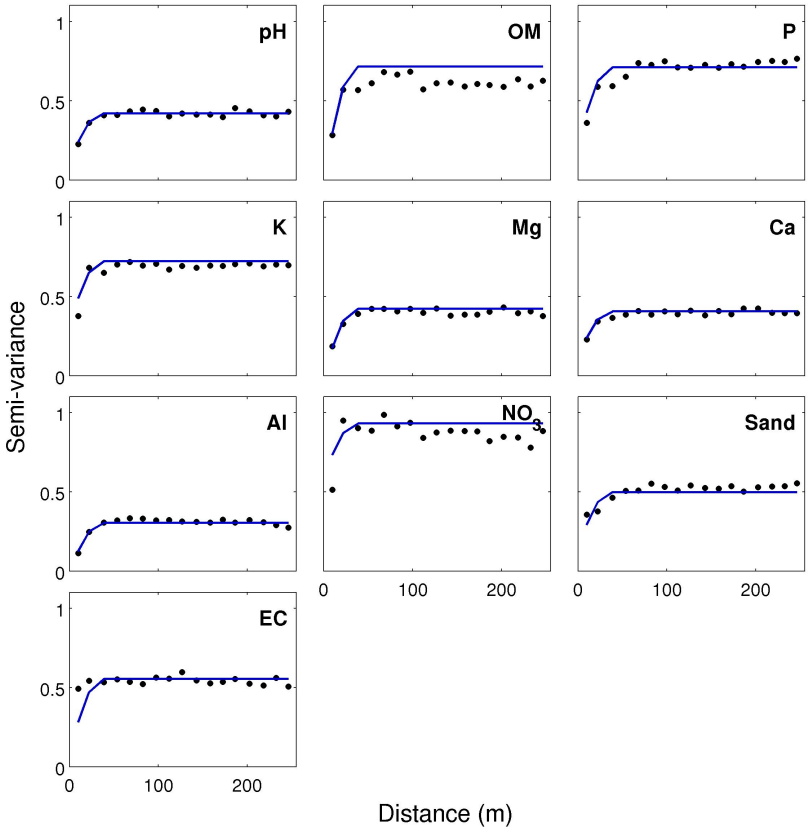

Figure 6.5: Variograms of residuals in CRAD modelled with a nugget effect and a spherical function with a range of 35 m for the variables sampled in 2000 and EC.

6.3 Results and discussion

6.3.1 Field-scale correlations

In traditional (non site-specific) farming practices, fertilizers are applied uniformly to the whole field and yield is assessed on a field-by-field basis. In standard agronomic practices, a few samples are taken from the entire field to obtain a composited average for specific soil properties. Such values comprise variation at all scales smaller than the field. In this case study, an average of 31 samples per hectare were taken, far exceeding the standard recommendations for obtaining field-scale averages used to determine fertilizer applications. For example, in Ontario the nutrient management regulations specify that a single soil sample cannot represent more than 10 ha (Reid, 2007).

Due to the high density of sampling, it is reasonable to assume that the field-scale correlations (Tables 6.1 and 6.2) represent adequately the field-scale relationships without a recourse to statistical inference. Electrical conductivity is moderately correlated with

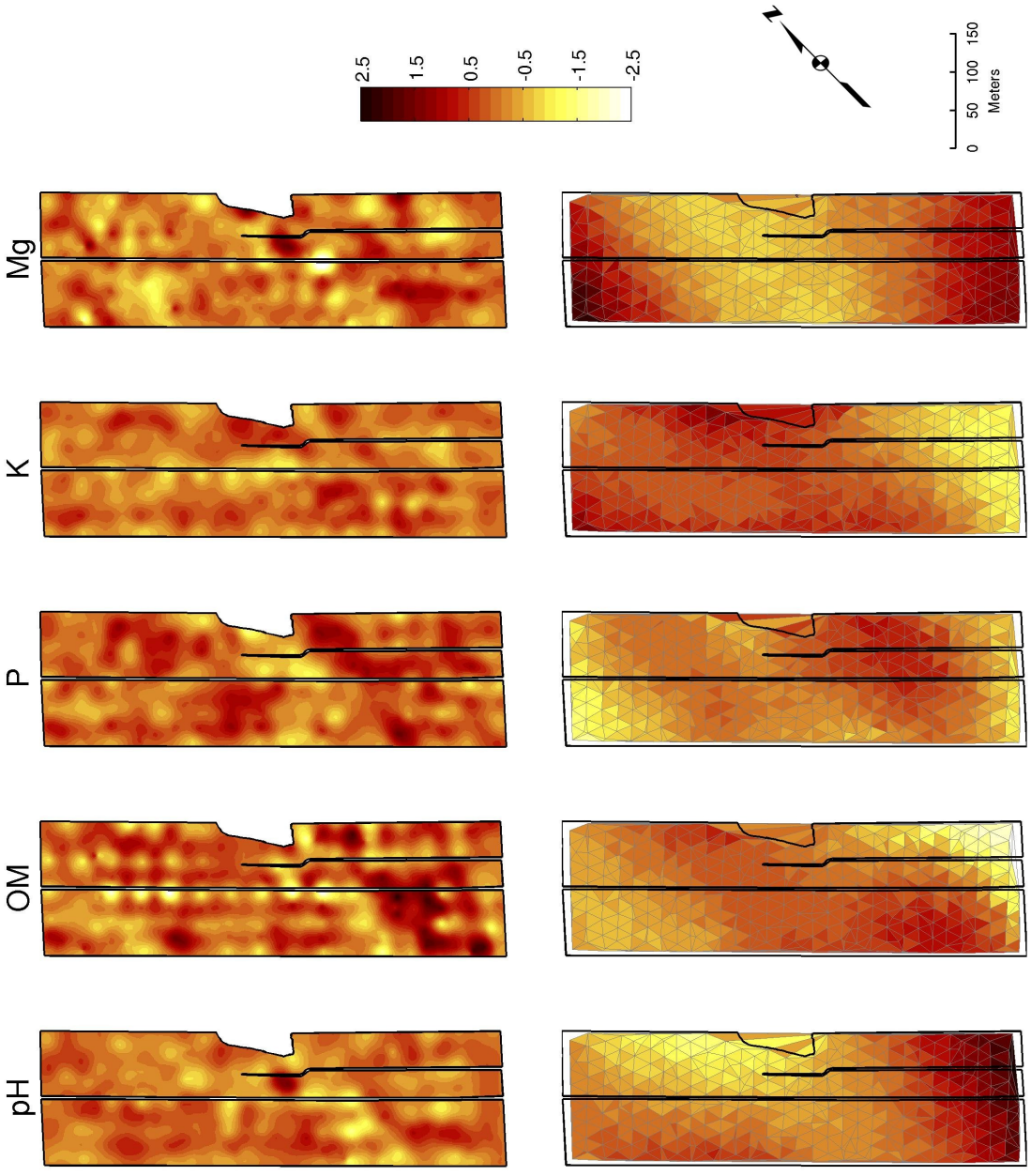

Figure 6.6: Maps of small scale (spherical model, upper panels) regionalized components obtained by cokriging and large scale (drift, lower panels) components represented by Delaunay triangulation.

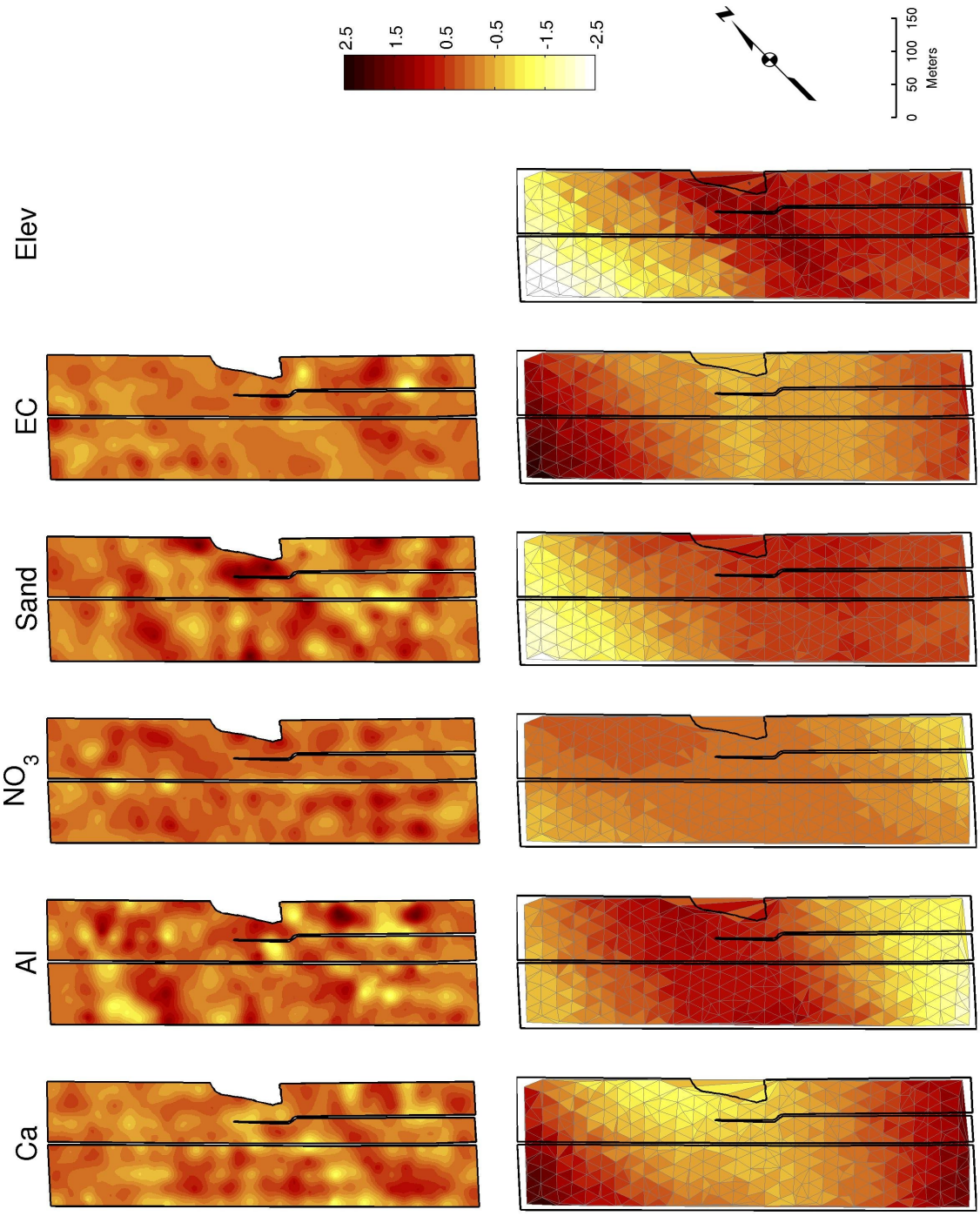

Figure 6.7: Maps of small scale (spherical model, upper panels) regionalized components obtained by cokriging and large scale (drift, lower panels) components represented by Delaunay triangulation.

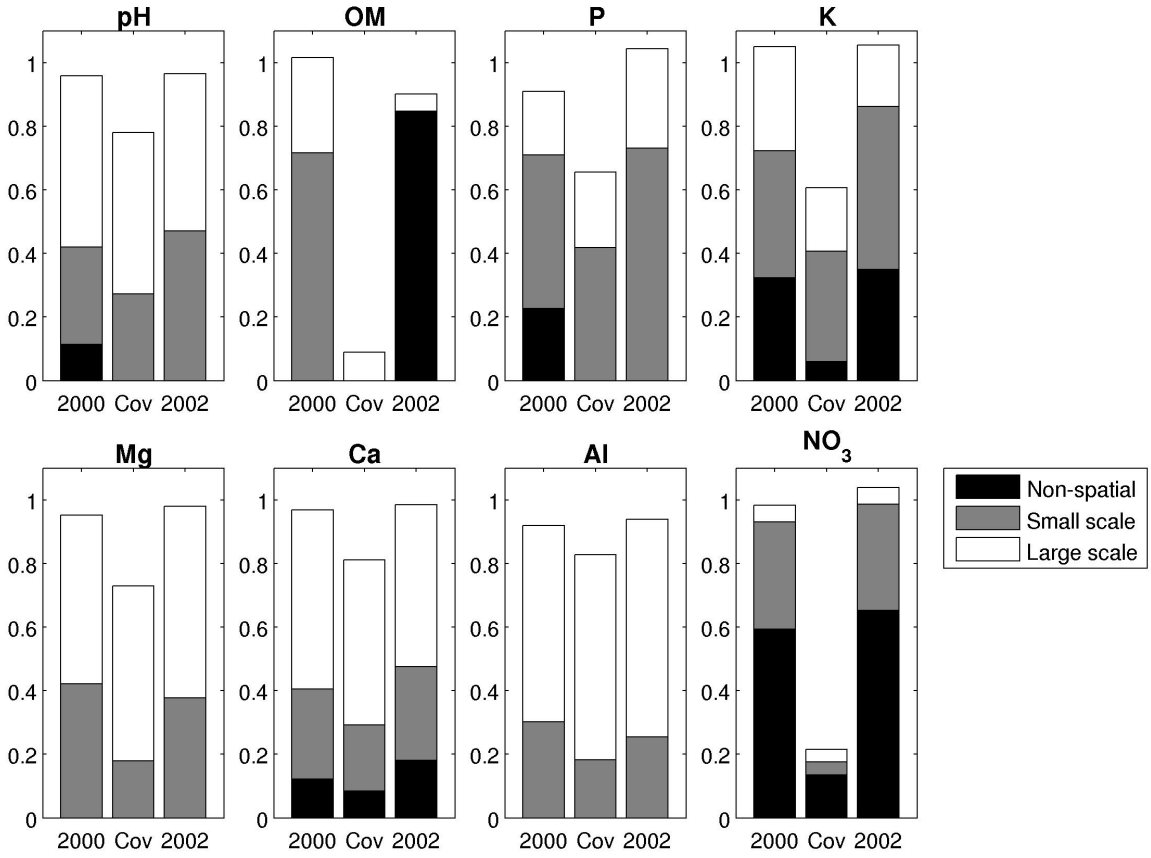

Figure 6.8: Scale-specific proportions of variance for the same variables sampled in 2000 and 2002 for the non-spatial (nugget effect), small scale (spherical model) and large scale (drift) components.

Mg, Ca and pH (positively) and with sand and Al (negatively). Between-year correlations are relatively weak for NO_3 and OM, but moderate to strong for other variables (Table 6.2). For producers, these field-scale relationships may ultimately determine the controls on the total yield that will be obtained for a given field.

6.3.2 Spatial patterns and relationships

The direct variograms for total variables sampled in 2000 (Fig. 6.3) show that three main types of patterns can be defined: semi-variances for pH, Mg, Ca, Al, Sand, EC and elevation increase in a generally linear fashion beyond the value of the sample variance (value of 1) without reaching a plateau; semi-variances for NO_3 are nearly constant, except at the first lag; OM, P and K have intermediate variograms with values increasing gradually to reach a plateau close to 1. For the first group of variables, spatial structures

Table 6.4: Correlations between scale-specific components for the same variables sampled in 2000 and 2002.

	Non-spatial	**Small scale**	**Large scale**
pH-pH	-	0.68	0.98
OM-OM	0.74	-	0.71
P-P	-	0.73	0.95
K-K	0.18	0.77	0.79
Mg-Mg	-	0.48	0.97
Ca-Ca	0.38	0.80	0.97
Al-Al	-	0.68	0.99
NO_3-NO_3	0.24	0.08	0.76

are likely present on the site with a scale of variation approaching or exceeding the size of the field. Using a variogram model with a sill for them would be problematic because they do not reach a plateau. Variogram functions without a sill would need to be used instead, leading to the false perception that these variables contain unlimited variance. The assumption that these total variables can be treated as the result of a stationary random function model is therefore not reasonable.

An important point to make is that there are no necessary relationships between field-scale spatial means for single variables and the spatial distribution of variables. For example, one could imagine two regionalized variables with the same mean and variance, but completely different spatial distributions. However, there is an obligate relationship between the field-scale correlations between variables and their scale-specific correlations. When two variables have a correlation close to 1 (or -1) at the field scale, they necessarily have the same (or opposite) spatial patterns and there can be no scale dependency in the correlations between them. Conversely, a weak correlation between two variables at the field scale may indicate strong correlations of different signs at different scales. This is evident from the observation of field scale correlations (Table 6.1) and cross variograms (Fig. 6.4). Variables that have a strong correlation at the field scale, such as Mg and Ca, also have similar variograms and their cross variogram have values close to the hulls of perfect correlation. Other combinations of variables, such as P and K, or K and NO_3 have weaker field-scale correlations and cross variograms that do not follow the patterns shown by the hulls. This could be an indication of scale-dependent correlations.

6.3.3 Multi-scale patterns and relationships

From the generally smooth aspect of maps of large scale components in Figs 6.6 and 6.7 (lower panels), it is clear that the local drift estimation procedure adequately captured large scale variation in the data. The large scale components have a finite sample variance that represent the magnitude of the variation at this scale within the limits of the field. The residual component can then be reasonably treated as the result of a second-order stationary random function and be modeled by a short-range spherical variogram function and a nugget effect (Fig. 6.5).

From the CRAD results, we can observe that the same groupings of variables defined from the patterns in the variograms can also be observed from the proportion of variance captured at each scale (Fig. 6.8). Variables with experimental variograms that did not seem to reach a plateau (pH, Al, Mg, Ca) have a large, but finite, amount of variation at large scale, with less than 40 % small scale variation and very small amounts of non-spatially distributed variation. EC and Sand had similar proportions with 34, 24 and 44, and 15, 35 and 50 %, respectively, of non-spatial, small and large scales variations. Elevation was entirely captured by the large scale drift component. Results for pH and Sand are in general agreement with the review of several studies by Whelan and McBratney (2000) which shows that soil texture had generally a small nugget effect (21 %, on average) and variograms with a short range (63 m) and pH had a very small nugget effect (12 %) and a range of 105 m.

Phosphorus has a small (in 2000) or absent (in 2002) nugget effect, together with a large proportion of small scale variation and a moderate amount of large scale variation. Potassium has a similar proportion of variation at the three scales. Likewise, in Whelan and McBratney (2000), semi-variograms computed for P and K had, on average, nugget effects representing respectively 71 and 78 % of the total within-field variation, but fairly long autocorrelation ranges (180 and 157 m).

Nitrate is characterized by a large proportion of variation at the nugget effect (approximately 60 %) and at the small scale (35 %) with very small amounts of large scale variation. The variation in organic matter is also predominantly composed of non-spatial and small scale variation, although the results for the two years differ in the respective proportions of each. We note that the very large amount of non-spatial variation for OM in 2002 may suggest an unidentified source of errors, potentially in the laboratory procedures. Hence, we are cautious about using and basing interpretation on these data. Cahn et al. (1994) also found nitrate to contain a lot of small scale variation but the soil organic carbon in their fields in Illinois was predominantly composed of large scale variation.

In modern farming practices, soil nutrients that are generally managed with soil amendments are phosphorus, potassium, nitrogen and, to some extent, organic matter. Interestingly, in our results these are the variables that contain the most variation at scales smaller than what can be managed by site specific amendment applications. Zones that would be delineated from maps of small scale components (Figs. 6.6 and 6.7) would be smaller than the scale of operation of modern machinery, or at least smaller than what could be economically viable to manage. Nitrate and organic matter are also characterized by very weak temporal correlations (Fig. 6.8). For nitrate, both the lack of large scale variation and the lack of temporal correlation could be attributed to the great mobility of nitrate in soils. The lack of temporal correlation in organic matter may be attributed to tilling practices, manure applications or inconsistencies in laboratory analysis.

At the large scale, higher elevations have a higher sand content, lower concentrations of base cations, and lower electrical conductivity and phosphorus (Table 6.3). This general large scale pattern would likely be difficult to manage with site-specific applications since it appears related to the geomorphological conditions of the site. Sites in the region are of fluvial origin and the presence of a river less than 1 km away from the fields would tend to indicate that fluvial sand 'bars' were deposited at higher elevations in the central portions of the field. Thus, variables with more variation at larger scales such as Ca, Mg, Al, Elev and Sand would be very difficult to manage.

At small scales, correlations between EC, Mg, Ca, pH and OM are moderately positive. From Figs. 6.6 and 6.7, it appears that small scale patterns in those variables are at least partly related to the longitudinal furrows apparent on the digital elevation model (Fig. 6.2). The furrows are remnants of a previous parallel surface drainage system. The observed nutrient patterns could be attributed to the long term accumulation of cations and organic matter in the furrows. Although the furrows are clearly apparent in the digital elevation model, their effect cannot be assessed by using elevation since the drift component completely captured the elevation gradient present in the field.

Overall, electrical conductivity is well correlated with several soil properties, indicating that it may be a good indicator of a soil's nutrient regime. However, in this study, EC was shown to be mostly related to variables that are harder to manage such as Ca, Mg, Al and Sand content. Its large scale component was moderately correlated to that of nitrate and organic matter, but these variables had limited amounts of large scale variation. For the two fields studied here, topography was a good indicator of the large scale variation in many variables, especially sand content. Those relationships could be tied to the the specific geomorphology of the area and may not be generalizable beyond

this site.

An interesting result arising from the CRAD analysis is that there is a stable relationship between scales of variation and temporal heterogeneity. Variables that contain a greater amount of large scale variation are more temporally correlated. Conversely, variables with more non-spatial and small scale variation have spatial patterns that are more variable from year to year. This could be explained in part by the greater mobility of those nutrients in the soil. Interestingly, a small proportion of variation in NO_3 was at large scales. If those results were found to be general patterns in agricultural fields, this would imply that spatial heterogeneity and scales of variation could be used as surrogates for temporal correlations. Variables that contain a large amount of variation at small scales are less likely to be temporally correlated. It is therefore not worthwhile investing in the site-specific management of those variables. The focus should instead be on variables with greater amounts of large scale variation, since they are thus likely to have patterns that are stable through time.

6.4 Concluding remarks

The issue of scale appears to be a crucial aspect for the assessment of the viability of SSM. Results presented here have highlighted interesting relationships between the manageability of nutrients, the scales of variation, the correlations between variables and the temporal correlations of variables.

The patterns at a manageable (large) scale were more coherent and easier to explain. The large scale component in all variables was significantly more temporally correlated than the small scale and non-spatial components. Overall, the relationships between variables were stronger at larger scales, even for variables that had small proportions of large scale variation. Topography was a good indicator of the large scale distribution in soil sand content. Electrical conductivity showed good correlations with several soil nutrients, especially base cations, with the strength of correlations increasing with scale.

The soil sampling grid used in this study is much denser than grids typically used for whole-field or site-specific management. A sampling grid that is significantly less dense would not allow the use of multi-scale analysis but may allow the generation of interpolated maps. Such maps would only reveal, with less accuracy, the larger scale patterns. Conveniently, those patterns are at a manageable scale and show stronger relationships. It may thus not be worthwhile investing time and money for the collection of soil samples at distances smaller than manageable scales.

Unfortunately, the manageable nutrients (NO_3, P and K) had smaller amounts of

large scale variation than the soil properties that are difficult to manage, and had spatial patterns that were less correlated temporally. In the fields sampled in this study, it may thus not be worthwhile investing in site-specific fertilizer applications because the zones that would be defined from large scale patterns in NO_3, P or K would encompass only a small proportion of the total variation in these nutrients. The effects such site-specific applications would have on yield, or on nutrient leaching, would therefore be marginal. It may nevertheless be worthwhile applying lime preferentially in more elevated areas of the fields to increase calcium content and pH in those areas.

Ultimately, the assessment of the viability of SSM is not complete without the investigation of patterns in yield and of correlations between soil nutrients and yield. If yield does not exhibit large scale patterns, the site-specific management or soil nutrients may not have an effect on yield. In this study, inconsistent yield data were provided by the producer, thereby preventing the inclusion of yield data in the statistical analysis. Some partial analyses have revealed an overall weak correlation between yield and most soil nutrients. Even without good correlations between soil nutrients and yield, it may still be worthwile conducting site-specific management to prevent problems related to over-fertilization (e.g. to prevent nitrate leaching Delgado et al., 2005).

In this study, CRAD proved to be an efficient method to decompose the variation in variables with various spatial patterns into scale-specific components. The stationarity assumptions were clearly not reasonable for several variables in this dataset, thereby making difficult the application of many methods described in Chapter 2, including CRA. In the specific context of SSM, where the interest is restricted to within-field variation, the use of a deterministic drift component with finite variance for describing large scale variation in CRAD was certainly more sensible than the use of a random function model. Thus, the assumptions underlying CRAD were reasonable and allowed us to perform multi-scale analysis in a framework that is coherent and internally consistent.

Chapter 7

Scale-specific causal modelling for understanding tree-soil interactions in a southern Québec forest

Forest ecosystems are characterized by two seemingly contradictory aspects that are related to scale. On the one hand, at large scales, they are spatially structured, organized and predictable. Tree species are generally found in relatively predictable associations, or forest cover types, on specific kinds of deposits as determined by their silvicultural characteristics. By knowing the characteristics of a given soil, one can often predict with reasonable confidence, from the pool of regionally available species, the species that will be found in a given stand. On the other hand, the complexity of the interrelations between and among trees and soils at small scales appears overwhelming. The large number of variables that are interacting and interdependent, and the role of chance, imply that predicting which tree will grow where and understanding the tree-soil interactions at small scales is extremely difficult, if not impossible.

Datasets collected in forest ecosystems often lead to poor predictive power and non-significant relationships because the total variation is included in the statistical analysis. The unstructured small scale variation may add a large amount of noise that interfere with the more organized large scale variation. The explicit scale-specific quantification of predictable components of variation in trees species distributions and soil variables has generally been lacking in forest ecology.

In recent years, the metacommunity concept has been suggested as a useful way to think about linkages between environmental heterogeneity biotic processes and spatial scales in ecology. The theoretical and empirical work related to metacommunities has been classified into four overall paradigms based on the importance each gives to the effects of envionmental heterogeneity and dispersal processes (Leibold et al., 2004).

The 'neutral model' and the 'patch dynamic' paradigms assume that strong spatial structuring of species is the result of dispersal limitations or to local extinction and colonization, and not the result of environmental heterogeneity. In the 'species sorting' paradigm, species are seen as not limited by dispersal and are distributed along environmental gradients based on associated niche differences. The 'mass effect' paradigm recognizes that the differences in population sizes between different patches may lead to flows of infividuals through dispersal between patches, leading to some species being found outside of their optimal habitat.

In reality, we can appreciate that any of these theories may apply best to specific regions or scales, at least to some degree. 'Neutral models' may be more appropriate in areas, or scales, where the environmental variations are more uniform while 'species sorting' would apply better in areas, or scales, where large environmental variation is present. Also, population ecologists recognize that certain species may behave differently depending on their niche requirements. In a forest setting, we could hypothesize that tree species that are specialized to specific areas may be well sorted along environmental gradients, while generalists may be more dispersal limited.

While the metacommunity concept could be useful in forest ecosystems for understanding the drivers underlying the spatial structure of tree communities, few theories are available that address the spatial structure in soil properties and the interrelationships between trees and soils. Since trees are known to have siginificant effects on soils (Binkley and Giardina, 1998), we could expect that the community structure of the trees will leave a footprint on the soil properties. Soil properties have been shown to be spatially structured through geomorphic processes and trees have been shown to respond to variations in soil nutrients (Burnett et al., 1998). Community dynamics can therefore be influenced by the spatial structure of the soil properties.

Causal modelling is an approach of exploratory statistical analysis in which the hypothesized directions of interactions between variables are described and, using various statistical methods, the relative effects of the different predictors are assessed. In this framework, causality refers to the hypothesis that changes occurring in one variable have an effect on changes in another (causality resides in the hypothesis only (Legendre and Legendre, 1998)). The causal modelling framework has been used to assess the respective effects of space and environmental factors on a set of dependent variables (e.g. plant species abundances). In this framework, variation in the dependent variables explained by environmental factors is often referred to as 'spatial dependence', while the unexplained spatial variation is considered 'spatial autocorrelation' (Legendre and Legendre, 1998). The former is seen as a deterministic influence of environmen-

tal factors while the latter is assumed to arise from biotic processes associated with the measured variable, such as dispersal. The approach presented by Borcard et al. (1992) combines partial canonical analysis with trend surface analysis. The spatially distributed variation in each variable is captured by a trend surface model fitted to the data. The variation in the dependent variables is partitioned between that explained by the environmental variables, the spatial matrix and the combined effects of environment and space. One of the drawbacks of this method is that the trend surface analysis only captures the larger scale spatial variation (Meot et al., 1998). Pelletier et al. (1999) used both trend surface analysis and neighbourhood matrices to capture the spatial distribution in forest floor variables. They showed that although there was an overlap between the variation explained by these two methods, the former tended to capture the large-scale patterns whereas the latter was better able to capture the micro-scale patterns.

More recently, principal coordinate analysis of neighbourhood matrices (PCNM Borcard and Legendre, 2002; Borcard et al., 2004) and distance-based eigenvector maps (DBEM, Dray et al. (2006)) were proposed as extensions of the canonical partitioning approach to include multiple scales of analysis. The advantages and drawbacks of this method are discussed in Chapter 2. Coregionalization analysis with a drift (CRAD) should provide an interesting alternative for causal modelling at multiple scales that overcomes some of the shortcomings of PCNM and DBEM. Most notably, the model used in CRAD remains parsimonious by restricting the number of scales to three, while PCNM provides a large number of scales. The separation of non-spatial and spatial variation, that was shown to be a problematic aspect of PCNM and DBEM, is conducted explicitely in CRAD by the modelling of the nugget effect. Also, CRAD can be readily used on regular or irregular transects and grids, while the interpretation of scales on irregular transects in PCNM, and to a certain extent in DBEM, can be problematic.

In this study, the combination of CRAD and causal modelling through partial redundancy analysis will be used to study the scale-specific interrelationships between tree species distribution, soil variables and physical factors in a temperate deciduous forest of southern Québec. The sampling grid used in this study was selected to cover a topographical and geomorphological gradient associated with large variations in soil conditions and tree species composition. Because of the strong variations occurring over the relatively small area covered by the grid, the study site could be seen as a 'scale model' of landscape level spatial variations. The causal model used is a very simple one where the physical factors have unidirectional influences on soil properties

and tree species, while the influences between trees and soils are bi-directional. The use of such scale-specific causal models should provide some insights into the complicating or facilitating effects of scale for understanding the functioning of the forest ecosystem.

7.1 Study area and dataset

The dataset used in this study was collected in the Morgan Arboretum, a forested area of 245 hectares located in the West Island of Montreal, Québec. In the fall of 1998, a systematic sampling grid composed of 124 points spaced by 50 m was laid down using a compass and hip chain. Geographical coordinates of the grid points were subsequently recorded with a global positioning system with differential corrections from a radio beacon. The sampling grid was localized to cover a gentle topographical gradient associated with different geomorphological conditions and forest cover types. The more elevated areas are characterized by a sub-mesic to sub-hygric fluvial sand deposit of variable thickness (approximately 0.2 to 4 m) underlain by a lacustrine clay layer. The soil profile has a moder humus form with small amount of organic matter in the top layer of mineral soil (A_{he} horizon). It is covered by a natural forest dominated by American beech (*Fagus grandifolia*) and red maple (*Acer rubrum*). The lower lying areas are dominated by a mesic glacial till deposit derived from a calcareous bedrock with a mature forest dominated by sugar maple (*Acer saccharum*) with a smaller abundance of white ash (*Fraxinus americana*), bitternut hickory (*Carya ovata*) and american basswood (*Tilia americana*). The central portion of the grid is characterized by a sub-hygric lacustrine clay deposit with a diverse forest composed of species such as red maple, silver maple (*Acer saccharinum*), shagbark hickory (*Carya cordiformis*) and black ash (*Fraxinus nigra*). Areas of the grid dominated with sugar maple are managed as a sugar bush while most of the forest encompassed by the grid is partially managed for firewood by selective cutting.

At each sampling point, the basal area value for each tree species was inventoried using a factor two prism. While a total of 35 tree species were identified in this inventory, only species that were present at more than 10 sampling locations were included in the statistical analysis (Table 7.1). Fixed volumes of mineral soil were collected with an aluminum core at depths of 0-15 and 15-30 cm and, where present, forest floor (LFH) was collected using a fixed area template. The weight of soil within the core and the volume of the core were used for the bulk density calculations of mineral soil samples, while the area of template, the depth of the LFH layer and the weight of samples were used for forest floor samples. The mineral soil was ground and sieved to pass through

a 2 mm mesh. Larger woody debris were removed from the forest floor samples prior to grinding in a coffee grinder. Laboratory analysis was carried out on all the soil and forest floor samples for percent organic carbon content by a wet oxidation and back titration method (Tiessen and Moir, 1993). These values were then converted to total soil organic carbon content by adjusting for bulk density. On the 0-15 cm samples only, Mehlich-III extracted K, P, Mg and Ca (Sen Tran and Simard, 1993; P was determined with Lachat autoanalyzer, other nutrients with atomic absorption flame emission), as well as pH in water were measured (Hendershot et al., 1993). Magnesium was not included in the statistical analysis due to its very high correlation with calcium. A KCl extraction procedure was used to determine the available nitrate and ammonium concentrations prior to and after a 28 day incubation (Hart et al., 1994; with Lachat autoanalyzer). The difference between these values was used to provide an indicator of nitrification and ammonification rates in the soil. On the 15-30 cm samples only, particle size analysis was carried for determination of sand, silt and clay content with a pre-treatment for organic matter removal. To avoid problems related to collinearity, only sand content was included in the analysis.

The elevation and slope at each sampling location was extracted from a one meter resolution digital elevation model generated from LIDAR (Laser detection and ranging) data acquired by Lasermap Image Plus in May 2002 from an airborne census (same dataset as Chapter 2). Table 7.2 provides a summary of the soil variables and physical factors used in the statistical analysis.

7.2 Statistical analysis

All the variables were standardized to a zero mean and unit variance prior to the statistical analysis. Experimental variograms of the total variables were calculated for distances from 50 to 300 m with lags every 50 m (Fig. 7.3). Coregionalization analysis with a drift (CRAD, Pelletier et al. (2009a,b)) was used to separate the variation in the total variables into scale-specific components of variation. The large-scale drift components (Fig. 7.5-7.8) were estimated with the local procedure in which a polynomial of order 1 (plane) is used to model the drift within the window. Optimal window sizes for each variable ranged between 100 and 200 m. The sample covariance matrix between drift components was used to describe large scale variation in the subsequent analyses. All possible combinations of direct and cross-experimental variograms of the residual components (Fig. 7.4) were modeled with a spherical variogram model with an estimated range of 150 m and a nugget effect. The resulting sill matrix estimates

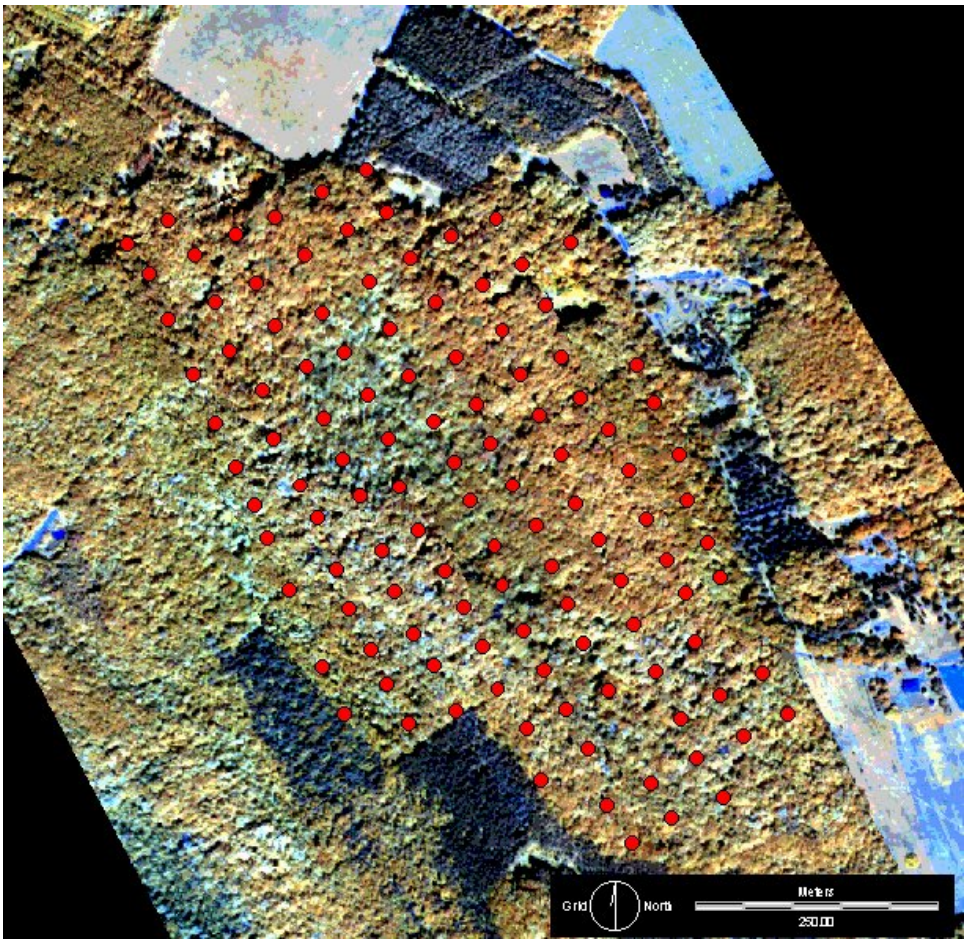

Figure 7.1: False color composite aerial image of the Morgan Arboretum showing the systematic sampling grid (red dots) used in this study.

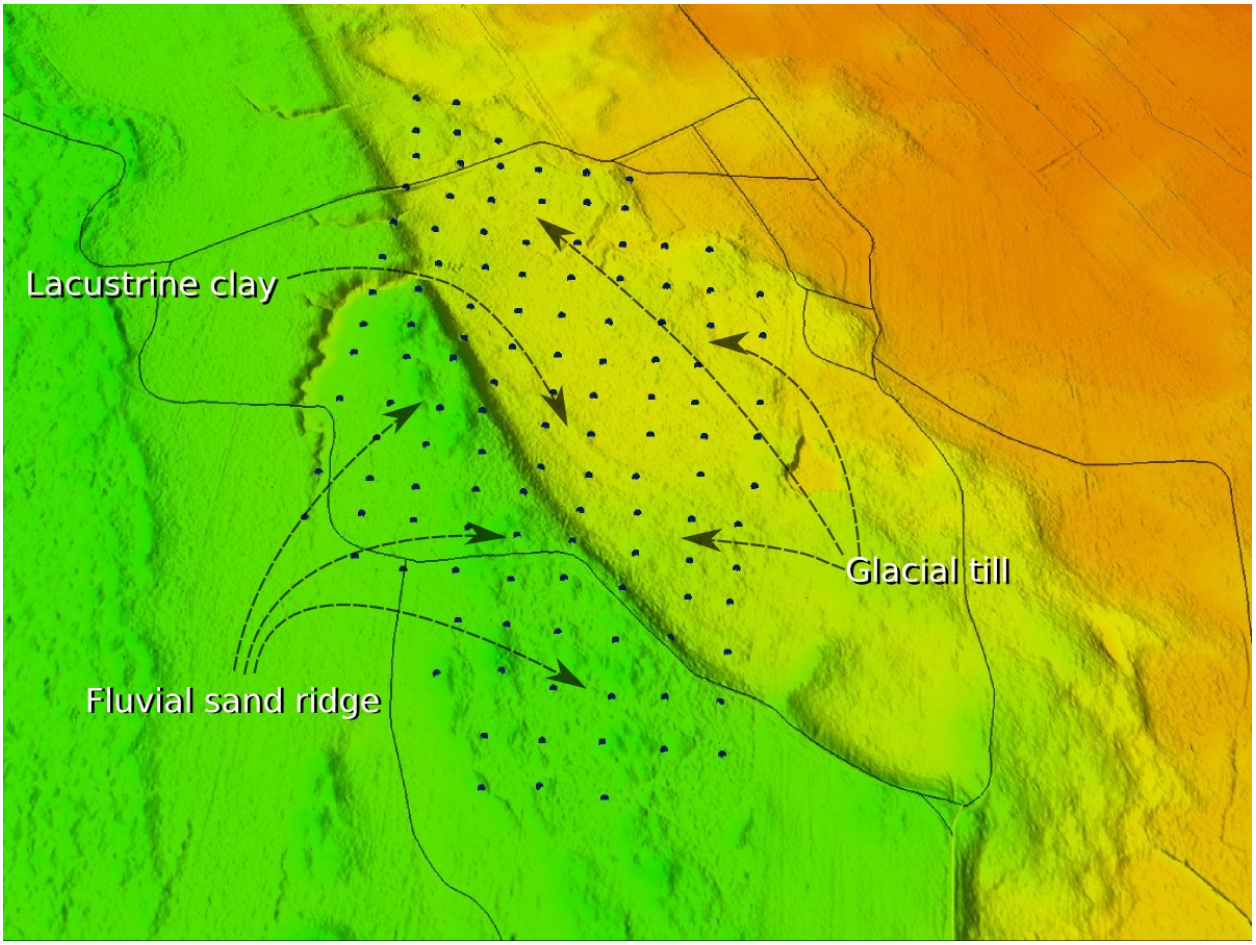

Figure 7.2: Three dimensional perspective of the ground surface of the Morgan Arboretum obtained with LIDAR technology showing the systematic sampling grid (blue dots). The largest difference between higher elevations (green) and lower elevations (yellow) on the area covered by the grid was 15.5 m.

Table 7.1: Most frequent species found in the tree inventory that were included in the statistical analyses.

English name	**Latin name**	**Acronym**	**Average basal area (m^2/ha)**	**Number of sites where found**
Sugar Maple	*Acer saccharum* Marsh.	AcSa	8.8	83
American Beech	*Fagus grandifolia* Ehrh.	FaGr	6.3	77
Red Maple	*Acer rubrum* L.	AcRu	3.9	70
American Basswood	*Tilia americana* L.	TiAm	2.4	61
White Ash	*Fraxinus americana* L.	FrAm	2.0	60
Bitternut Hickory	*Carya cordiformis* (Wangenh.) K. Koch	CaCo	1.6	38
Red Oak	*Quercus rubra* L.	QuRu	1.3	44
Shargbark Hickory	*Carya ovata* (Mill.) K. Koch	CaOv	1.1	29
Silver Maple	*Acer saccharinum* L.	AcSi	0.7	12
Yellow Birch	*Betula alleghaniensis* Britton	BeAl	0.5	16
Black Ash	*Fraxinus nigra* Marsh.	FrNi	0.4	12

Table 7.2: Description of soil variables and physical factors

Variable	**Short Name**	**Mean**	**Standard deviation**
Soil variables			
Total soil organic carbon (kg/m^2)	tC	6.54	1.36
Total soil organic carbon 0-15 cm (kg/m^2)	tC(0-15)	4.22	1.02
Total soil organic carbon 15-30 cm (kg/m^2)	tC(15-30)	2.00	0.79
Total organic carbon in forest floor (kg/m^2)	tC(FF)	0.34	0.68
Extractable potassium (ug/g)	K	98.30	71.00
Extractable phosphorus (ug/g)	P	23.93	20.32
Extractable calcium (ug/g)	Ca	1389.42	1209.07
Mineralized nitrate (ug N-NO_3/g)	dNO_3	19.41	33.95
Mineralized ammonium(ug N-NH_4/g)	dNH_4	120.56	85.11
pH in water	pH	5.70	0.62
Physical factors			
Sand content (%)	Sand	53.64	24.53
Elevation (m above sea level)	Elev	41.86	4.87
Slope (%)	Slope	3.43	3.13

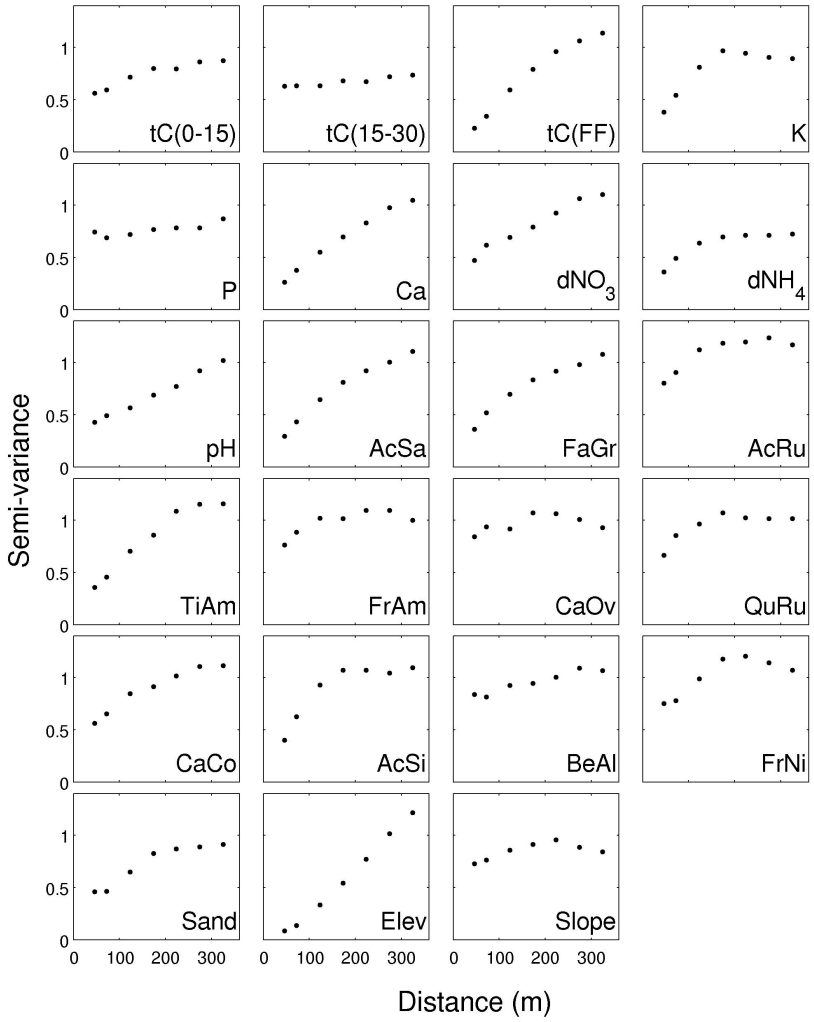

Figure 7.3: Experimental variograms of total variables.

were used in the multivariate analysis to describe the non-spatial (nugget effect) and medium scale (spherical model) components. Maps of the medium scale components were obtained by cokriging of regionalized components (Figs. 7.5-7.8, see Chapter 3 for methodology).

Covariance matrices corresponding to each scale of variation were used to perform partial correlation analyses and multivariate causal modelling. The total matrices (C^s) were divided into physical factors (C^s_{PF}), soil variables (C^s_{SV}), trees (C^s_{TR}) and other covariance components as follows:

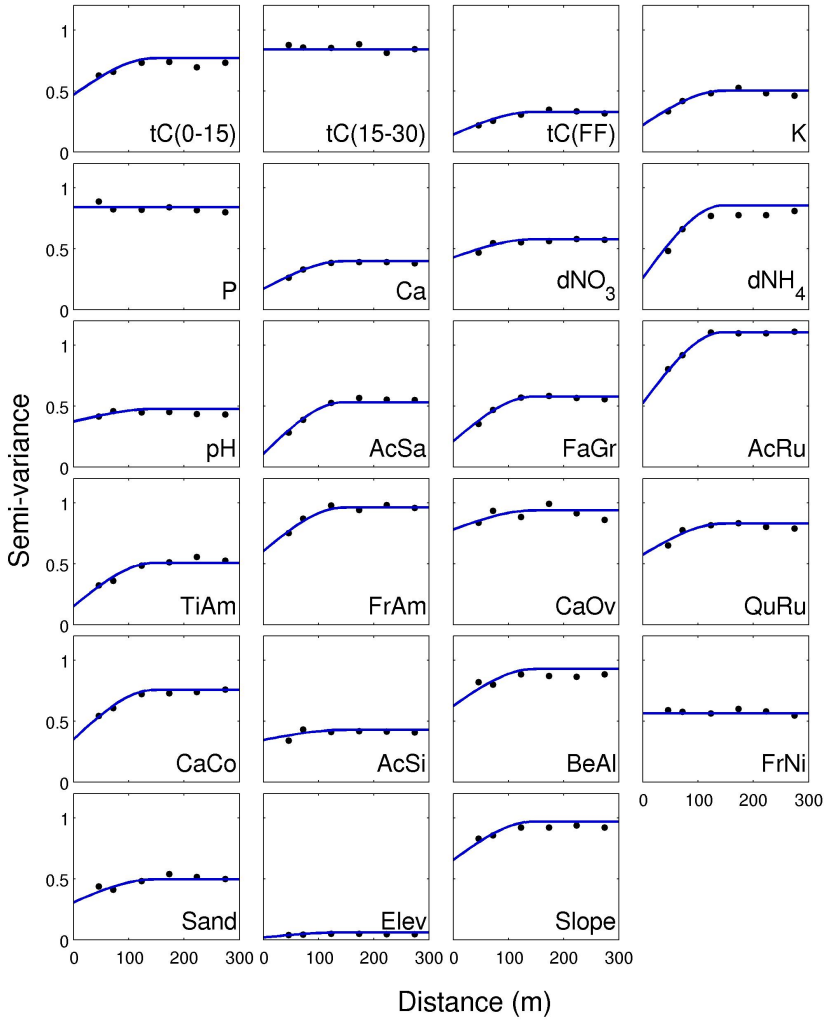

Figure 7.4: Variograms of residuals in CRAD fitted with a nugget effect and a spherical component (solid line).

$$C^s = \begin{Bmatrix} C^s_{SV} & C^s_{SV\,TR} & C^s_{SV\,PF} \\ & C^s_{TR} & C^s_{TR\,PF} \\ & & C^s_{PF} \end{Bmatrix}$$

After transforming the scale-specific covariance matrices to correlation matrices (R^s), the partial correlations between and among soil nutrients and trees adjusted for the effects of physical factors were calculated as follows:

$$R^s_{SV\cup TR|PF} = R^s_{SV\cup TR} - R^s_{SV\cup TR\,PF}\left[R^s_{PF}\right]^{-1}\left[R^s_{SV\cup TR\,PF}\right]^T$$

Partial correlations calculated in this way are shown in Tables 7.3, 7.4 and 7.5.

Direct and partial redundancy analyses (van den Wollenberg, 1977; Legendre and Legendre, 1998) were used to assess the multivariate total and scale-specific relationships between physical factors, soil nutrients and trees. Assymmetrical (unidirectional) effects of physical factors on soil nutrients and trees were assessed with direct redundancy analysis. For example, the covariance of soil nutrients predicted by physical factors was given by

$$C^s_{\hat{SV}(PF)} = [C^s_{SV\,PF}]\,[C^s_{PF}]^{-1}\,[C^s_{SV\,PF}]^T$$

and the r-square statistic was given by

$$r^{2,s}_{\hat{SV}(PF)} = \frac{tr(C^s_{\hat{SV}(PF)})}{tr(C^s_{SV})}$$

where tr is the trace operator. The amount of variation for each variable at each scale and the proportion explained by physical factors ($r^{2,s}_{\hat{SV}(PF)}$ and $r^{2,s}_{\hat{TR}(PF)}$) are shown in the form of bar charts in Figure 7.9. The interactions between soil nutrients and tree species interactions bi-directional and assessed with partial redundancy analysis. For example, the covariance of soil nutrients predicted by trees while adjusting for the physical factors was given by

$$\begin{aligned} C^s_{\hat{SV}(TR|PF)} &= \begin{bmatrix} C^s_{SV\,TR} & C^s_{SV\,PF} \end{bmatrix} \begin{bmatrix} C^s_{TR} & C^s_{TR\,PF} \\ & C^s_{PF} \end{bmatrix}^{-1} \begin{bmatrix} C^s_{SV\,TR} \\ C^s_{SV\,PF} \end{bmatrix} - \\ & [C^s_{SV\,PF}]\,[C^s_{PF}]^{-1}\,[C^s_{SV\,PF}]^T \end{aligned}$$

and the r-square statistic was given by

$$r^{2,s}_{\hat{SV}(TR|PF)} = \frac{tr(C^s_{\hat{SV}(TR|PF)})}{tr(C^s_{SV})}$$

Total and scale-specific multivariate path diagrams shown in Figure 7.10 were generated by combining direct and partial redundancy analyses. Note that although the arrows representing the relationships between trees and soils are depicted as unidirectional, they really indicate a two-way interaction between trees and soils. However, the strength of the tree effects on soils and of soil effects on trees are not equal, hence the use of two unidirectional arrows instead of one two-directional arrow. To avoid collinearity between carbon at the different depths and total carbon, the latter (tC) was not included

in this partial redundancy analysis.

7.3 Results and discussion

From the maps of large scale drift components and the variograms of residuals (Fig. 7.4), we can conclude that the CRAD method with the local drift estimation procedure adequately separated the total variables into smoothly varying large-scale drift components and residual components that share similar small and medium scale spatial structures. The assumption that the residuals for all variables can be viewed as the result of second-order stationary random functions sharing the same variogram functions (nugget and spherical model) is more reasonable than the same assumption on the total variables.

Large scale patterns

The large scale components of variation in soil variables and trees are better predicted by physical factors than the non-spatial or medium scale components (Fig. 7.9). The large scale direct effects (adjusted for the effects of physical factors) of trees on soil variables or soil variables on trees are relatively weak. The soil variables and trees species distribution at large scales in the area covered by the grid are thus tightly associated with the gradient in topography and soil texture. This is summarized in the multivariate path diagram (Fig. 7.10).

By observing maps of drift components (Figs. 7.5-7.8) we can observe that more elevated areas are on sandy soils (fluvial sand deposit) with an accumulation of carbon in the forest floor and comparatively less carbon in the mineral soil. The soil pH, the availability of calcium and potassium, and nitrification in those areas is relatively low. These areas are dominated by *Fagus grandifolia* and *Acer rubrum* with lower abundances of *Betula alleghaniensis.* Areas on the lower end of the topographical gradient are have an intermediate sand content (glacial till deposit, loamy texture). The soils have a higher pH as well as a higher calcium and potassium content, and are characterized by a higher nitrification rate and a mull humus form with an accumulation of organic carbon in the top 15 cm of mineral soil. These areas are dominated by *Acer Saccharum* with a smaller abundance of *Carya cordiformis, Fraxinus americana* and *Tilia americana.* The dominance of *Acer Saccharum* in those sites may be partly attributed to the management of the area as a sugar bush.

Another pattern clearly visible at the large scale but not directly tied to the elevation

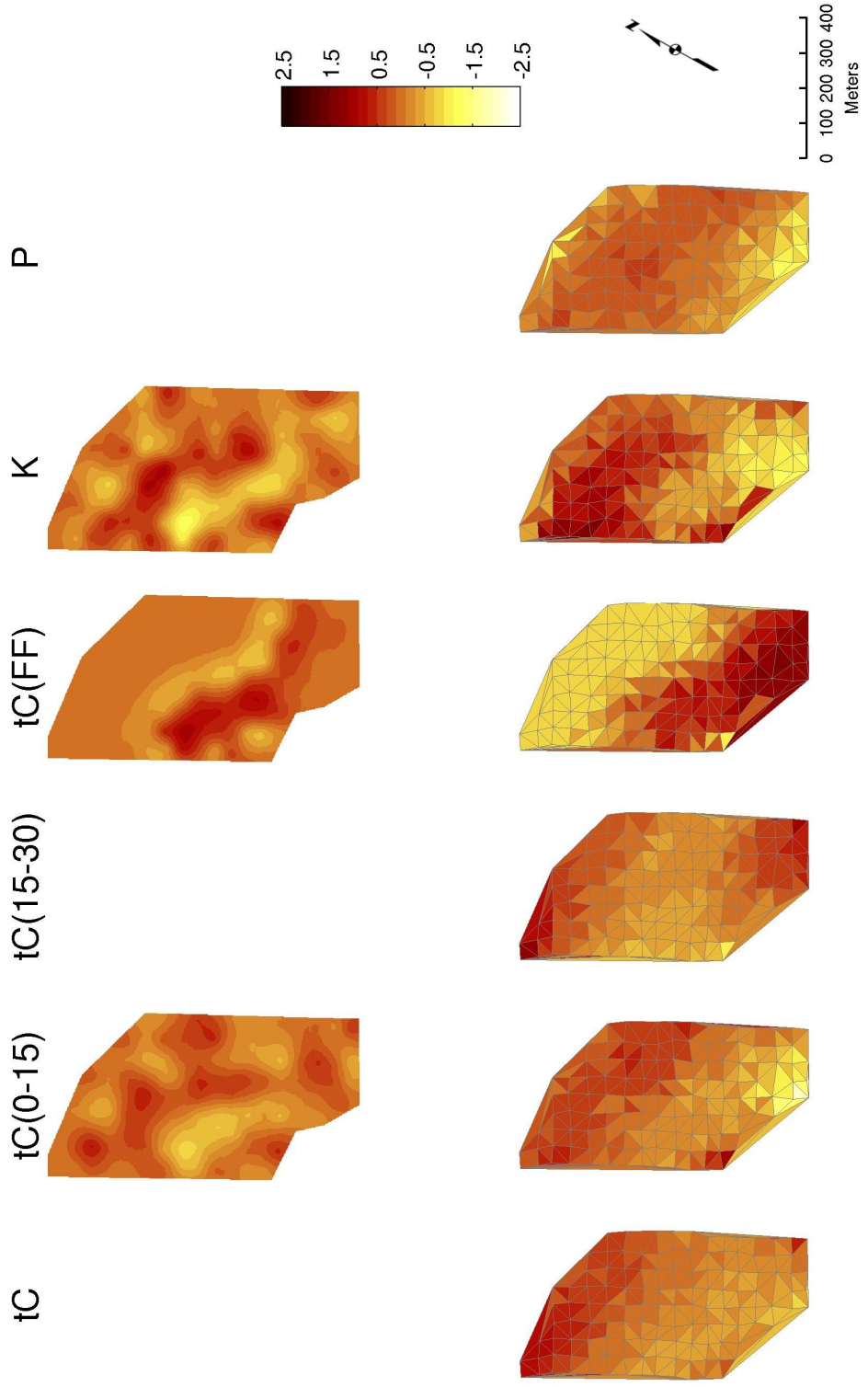

Figure 7.5: Maps of medium scale (spherical model, top) regionalized components obtained by cokriging and large scale (drift, bottom) components represented by delaunay triangulation.

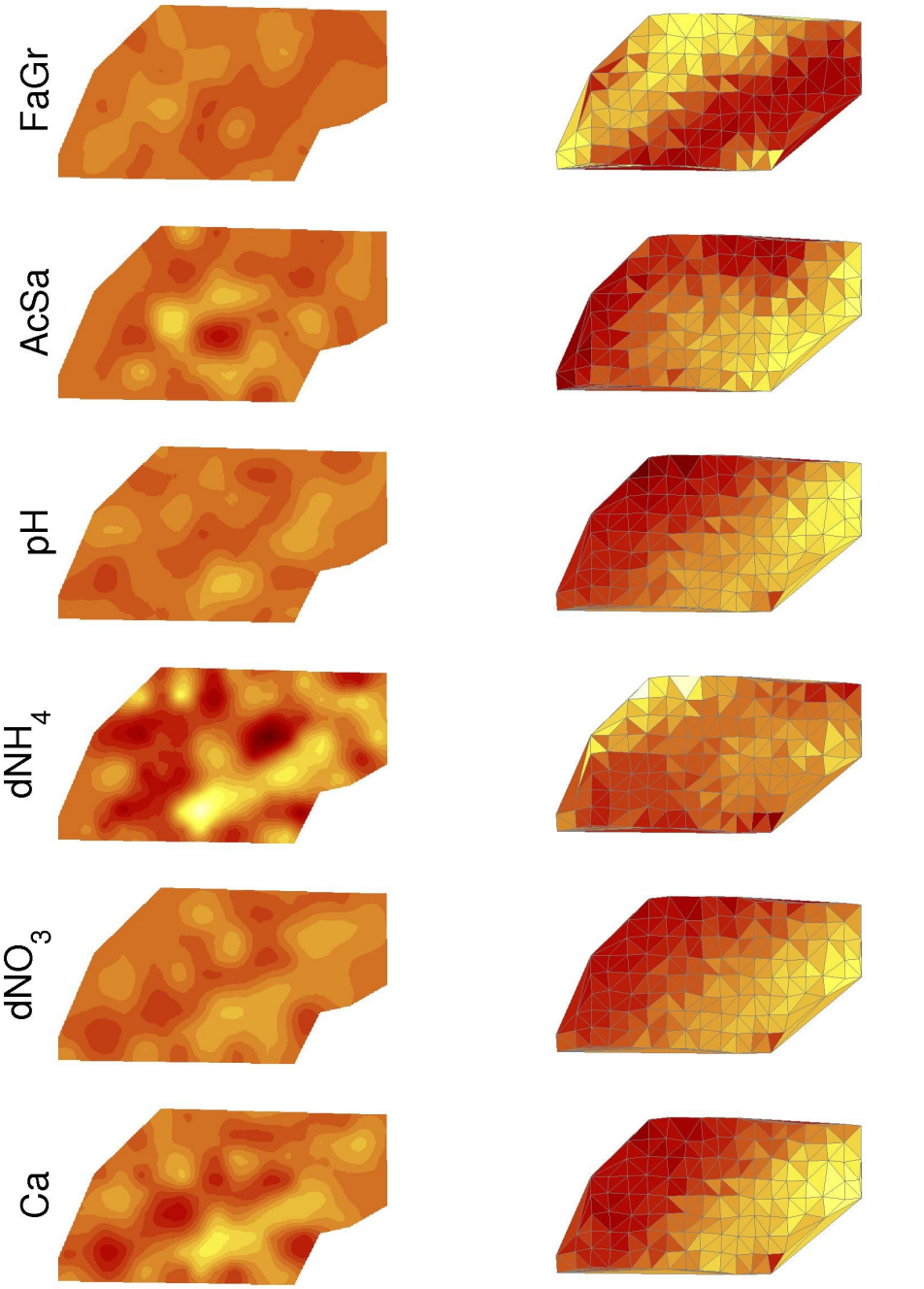

Figure 7.6: Maps of medium scale (spherical model, top) regionalized components obtained by cokriging and large scale (drift, bottom) components represented by delaunay triangulation.

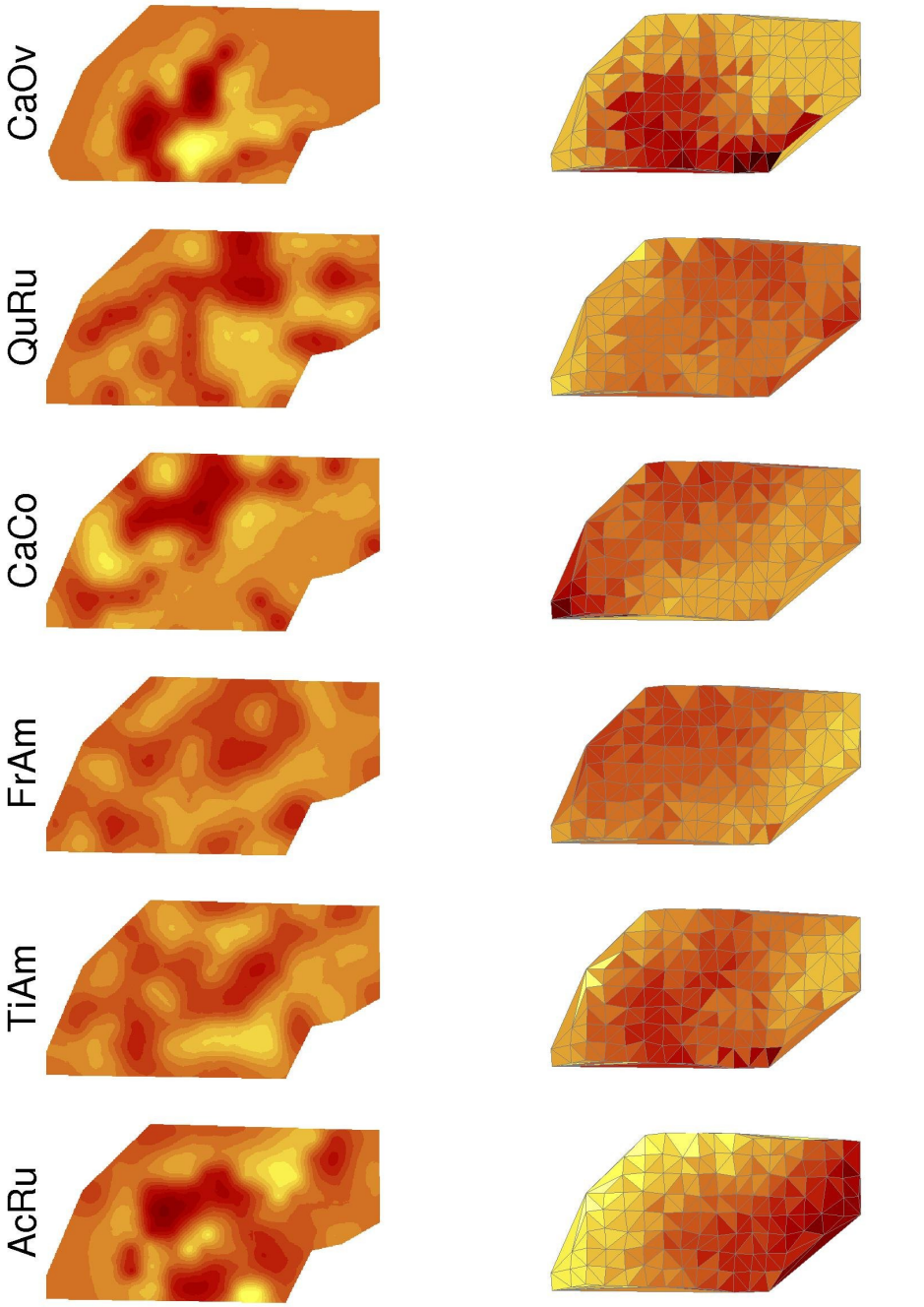

Figure 7.7: Maps of medium scale (spherical model, top) regionalized components obtained by cokriging and large scale (drift, bottom) components represented by delaunay triangulation.

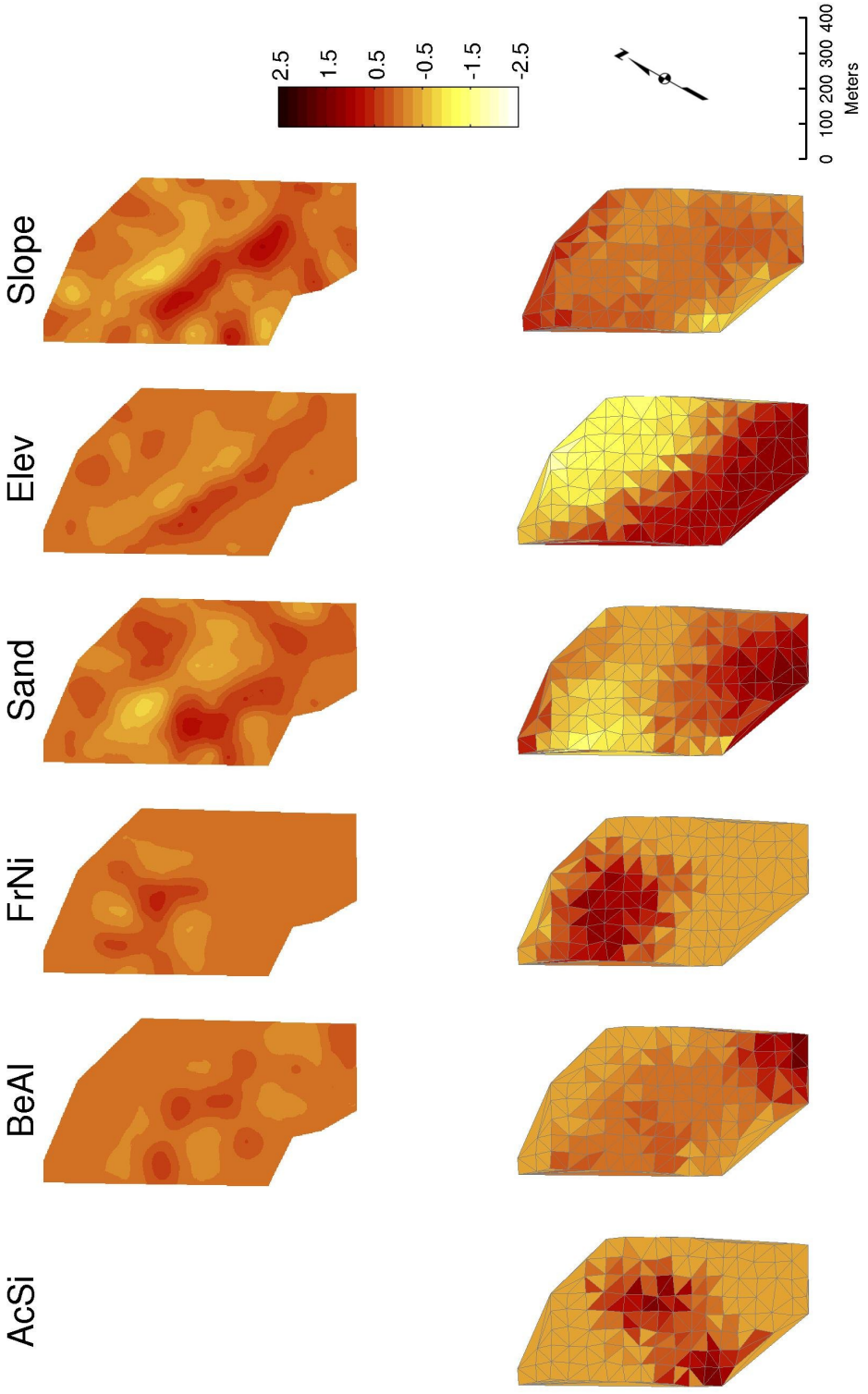

Figure 7.8: Maps of medium scale (spherical model, top) regionalized components obtained by cokriging and large scale (drift, bottom) components represented by delaunay triangulation.

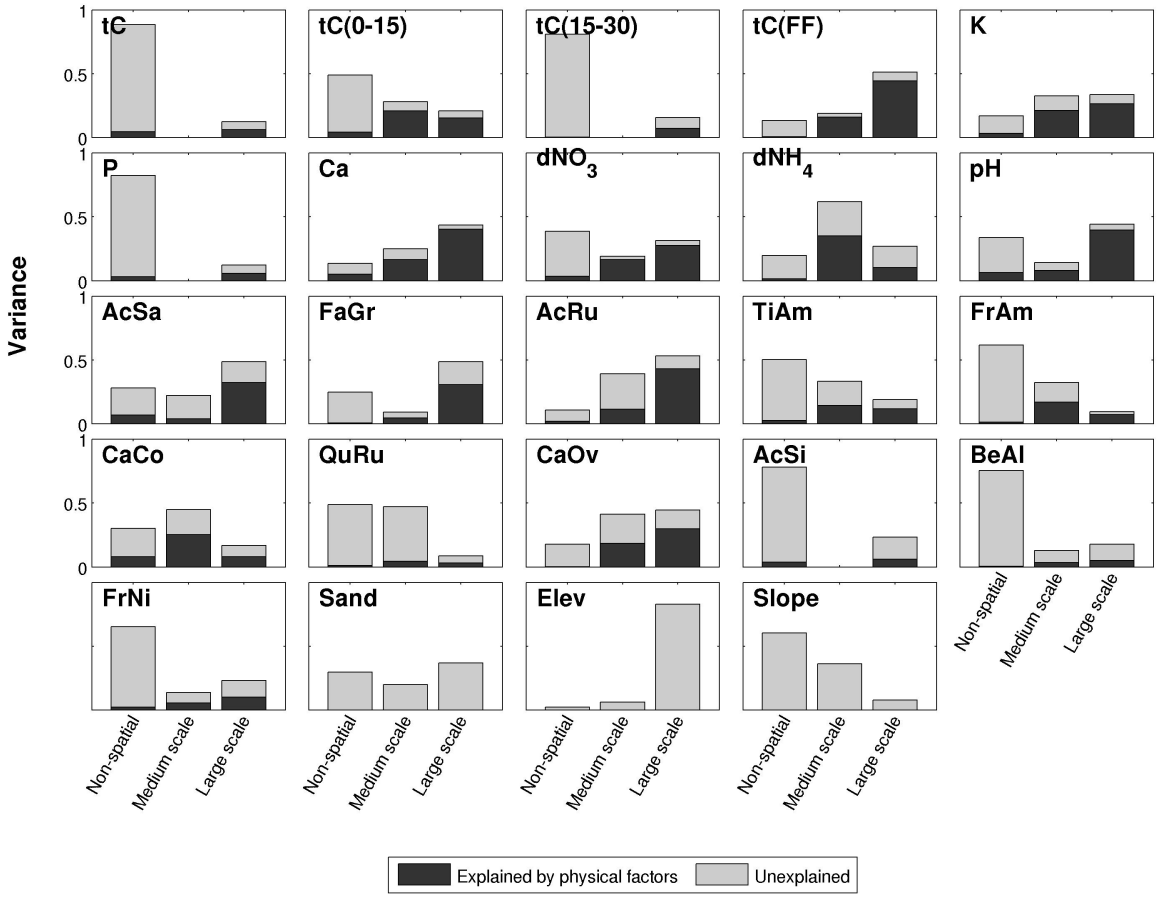

Figure 7.9: Proportion of variation at each scale for each variable. The height of the bar corresponds to the variance for that variable at that scale. The dark grey portion corresponds to the variation explained by physical factors.

Table 7.3: Table of correlations (on and above main diagonal) and partial correlations adjusting for the effects of physical factors (below main diagonal) for the non-spatial (nugget effect) component.

	tC	**tC(0-15)**	**tC(15-30)**	**tC(FF)**	**K**	**P**	**Ca**	**dNO_3**	**dNH_4**	**pH**	**AcSa**	**FaGr**	**AcRu**	**TiAm**	**FrAm**	**CaCo**	**QuRu**	**CaOv**	**AcSi**	**BeAl**	**FrNi**	**Sand**	**Elev**	**Slope**
tC	1	0.76	0.48	-0.21	0.44	0.06	0.55	0.42	0.08	0.07	-0.12	0	0.06	-0.01	0.1	0.26	-0.05	0.07	0.05	-0.03	0.23	-0.21	0	0.1
tC(0-15)	0.69	1	0.18	-0.13	0.15	-0.02	0.41	0.33	0.02	0.09	0.15	0.11	-0.06	-0.01	-0.03	0.16	-0.1	-0.14	0.04	-0.05	0.08	-0.15	0.01	0.26
tC(15-30)	0.49	0.18	1	0.03	0.03	0.03	0.09	0.16	-0.15	-0.21	0.02	-0.12	-0.03	-0.13	-0.03	0.26	-0.08	-0.05	0.1	-0.1	-0.07	0.02	0.03	0.02
tC(FF)	-0.22	-0.09	0.04	1	0.02	-0.2	0.03	-0.11	-0.14	-0.04	0.1	0.07	0.01	-0.09	0.1	-0.24	-0.36	-0.13	0.02	-0.04	-0.05	0.02	-0.19	-0.2
K	0.36	0.1	0.04	0.02	1	0.33	0.53	0.18	0.23	0.09	-0.29	0.06	0.08	0.14	0.17	0.25	-0.03	-0.09	0.11	-0.02	0.22	-0.42	0.24	0
P	0.07	0.02	0.03	-0.21	0.28	1	0.05	0.07	0.24	-0.2	-0.24	-0.05	0.08	0.01	-0.03	-0.1	0.13	0.19	-0.03	-0.06	0.11	-0.08	0.15	-0.17
Ca	0.41	0.24	0.09	0.08	0.37	0.1	1	0.53	-0.01	0.54	-0.05	-0.13	0.07	-0.04	0.08	0.23	0.13	-0.15	-0.01	0.12	0.12	-0.44	0.1	0.45
dNO_3	0.38	0.25	0.16	-0.07	0.2	0.14	0.4	1	-0.29	0.34	-0.22	0.26	0.27	0.18	-0.05	0.12	0.25	-0.01	-0.1	0.06	0.2	0	-0.08	0.29
dNH_4	0.08	0	-0.16	-0.13	0.32	0.3	-0.02	-0.36	1	-0.37	-0.01	-0.02	-0.18	0.01	-0.28	0.12	-0.31	0.21	0	0.13	0.16	0.18	-0.2	0.17
pH	-0.02	-0.03	-0.21	0.03	-0.02	-0.17	0.28	0.25	-0.37	1	-0.14	0.01	0.17	-0.01	0.11	-0.09	0.36	-0.18	-0.15	0.08	0	-0.3	0.14	0.35
AcSa	-0.1	0.09	0	0.19	-0.16	-0.15	-0.1	-0.34	-0.13	-0.2	1	-0.12	-0.28	-0.04	0.1	0.11	-0.08	-0.04	-0.14	0.1	-0.43	0.25	-0.09	0.4
FaGr	0.02	0.09	-0.13	0.15	0.04	-0.06	-0.15	0.25	0	-0.03	-0.18	1	0.08	0.1	0.01	-0.07	0.03	0.22	-0.3	-0.01	-0.09	-0.02	0.25	0.13
AcRu	-0.05	-0.11	0	-0.08	-0.08	0.06	-0.1	0.28	-0.12	0.09	-0.11	0.17	1	-0.13	-0.33	-0.04	-0.1	0.22	0.33	0.06	0.37	-0.4	-0.07	-0.09
TiAm	0.01	-0.02	-0.13	-0.07	0.22	0.05	-0.02	0.13	-0.06	0	-0.14	0.1	-0.06	1	0.23	0.52	0.32	-0.23	-0.17	-0.05	0.01	0.18	-0.13	0.12
FrAm	0.15	0	-0.05	0.15	0.22	-0.05	0.16	-0.03	-0.28	0.15	0.04	-0.05	-0.2	0.22	1	0.25	0.11	0.05	0.06	-0.18	0.01	0.14	0.12	0.01
CaCo	0.2	0.03	0.25	-0.12	0.21	-0.03	-0.02	-0.02	0.06	-0.29	-0.07	-0.16	0	0.48	0.24	1	-0.01	-0.02	0.25	-0.11	-0.03	-0.14	0.14	0.51
QuRu	-0.05	-0.13	-0.09	-0.29	-0.08	0.11	0.07	0.24	-0.28	0.31	-0.11	-0.04	-0.05	0.33	0.07	-0.09	1	0	-0.11	-0.06	-0.08	-0.1	0.25	0.11
CaOv	0.08	-0.13	-0.05	-0.13	-0.05	0.2	-0.12	-0.01	0.19	-0.15	-0.07	0.22	0.26	-0.25	0.03	-0.01	0	1	0.07	-0.39	0.09	0.09	-0.04	0.01
AcSi	0.04	0.07	0.11	-0.04	0.07	-0.06	0.03	-0.05	0.04	-0.1	-0.03	-0.25	0.24	-0.14	0.1	0.36	-0.07	0.08	1	-0.28	0.01	-0.06	-0.06	-0.2
BeAl	-0.02	-0.06	-0.1	-0.02	-0.01	-0.05	0.12	0.04	0.12	0.07	0.06	-0.03	0.09	-0.07	-0.19	-0.14	-0.08	-0.4	-0.26	1	-0.07	0.03	0.02	0.06
FrNi	0.17	0.03	-0.05	-0.08	0.15	0.12	0.01	0.19	0.17	-0.06	-0.38	-0.05	0.25	0.03	0.07	-0.04	-0.06	0.11	-0.02	-0.06	1	-0.18	-0.05	0.02
Sand	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-0.51	-0.1
Elev	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.06
Slope	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Table 7.4: Table of correlations (on and above main diagonal) and partial correlations adjusting for the effects of physical factors (below main diagonal) for the medium scale (spherical) component.

	tC	tC(0-15)	tC(15-30)	tC(FF)	K	P	Ca	dNO_3	dNH_4	pH	AcSa	FaGr	AcRu	TiAm	FrAm	CaCo	QuRu	CaOv	AcSi	BeAl	FrNi	Sand	Elev	Slope
tC	1	0.76	0.48	-0.21	0.44	0.06	0.55	0.42	0.08	0.07	-0.12	0	0.06	-0.01	0.1	0.26	-0.05	0.07	0.05	-0.03	0.23	NaN	NaN	NaN
tC(0-15)	0.69	1	0.18	-0.13	0.15	-0.02	0.41	0.33	0.02	0.09	0.15	0.11	-0.06	-0.01	-0.03	0.16	-0.1	-0.14	0.04	-0.05	0.08	-0.82	-0.57	-0.71
tC(15-30)	0.49	0.18	1	0.03	0.03	0.03	0.09	0.16	-0.15	-0.21	0.02	-0.12	-0.03	-0.13	-0.03	0.26	-0.08	-0.05	0.1	-0.1	-0.07	NaN	NaN	NaN
tC(FF)	-0.22	-0.09	0.04	1	0.02	-0.2	0.03	-0.11	-0.14	-0.04	0.1	0.07	0.01	-0.09	0.1	-0.24	-0.36	-0.13	0.02	-0.04	-0.05	0.91	0.73	0.52
K	0.36	0.1	0.04	0.02	1	0.33	0.53	0.18	0.23	0.09	-0.29	0.06	0.08	0.14	0.17	0.25	-0.03	-0.09	0.11	-0.02	0.22	-0.8	-0.54	-0.45
P	0.07	0.02	0.03	-0.21	0.28	1	0.05	0.07	0.24	-0.2	-0.24	-0.05	0.08	0.01	-0.03	-0.1	0.13	0.19	-0.03	-0.06	0.11	NaN	NaN	NaN
Ca	0.41	0.24	0.09	0.08	0.37	0.1	1	0.53	-0.01	0.54	-0.05	-0.13	0.07	-0.04	0.08	0.23	0.13	-0.15	-0.01	0.12	0.12	-0.8	-0.54	-0.6
dNO_3	0.38	0.25	0.16	-0.07	0.2	0.14	0.4	1	-0.29	0.34	-0.22	0.26	0.27	0.18	-0.05	0.12	0.25	-0.01	-0.1	0.06	0.2	-0.91	-0.6	-0.48
dNH_4	0.08	0	-0.16	-0.13	0.32	0.3	-0.02	-0.36	1	-0.37	-0.01	-0.02	-0.18	0.01	-0.28	0.12	-0.31	0.21	0	0.13	0.16	-0.72	-0.41	-0.37
pH	-0.02	-0.03	-0.21	0.03	-0.02	-0.17	0.28	0.25	-0.37	1	-0.14	0.01	0.17	-0.01	0.11	-0.09	0.36	-0.18	-0.15	0.08	0	-0.75	-0.6	-0.43
AcSa	-0.1	0.09	0	0.19	-0.16	-0.15	-0.1	-0.34	-0.13	-0.2	1	-0.12	-0.28	-0.04	0.1	0.11	-0.08	-0.04	-0.14	0.1	-0.43	0.07	0.12	0.32
FaGr	0.02	0.09	-0.13	0.15	0.04	-0.06	-0.15	0.25	0	-0.03	-0.18	1	0.08	0.1	0.01	-0.07	0.03	0.22	-0.3	-0.01	-0.09	0.48	0.06	0.54
AcRu	-0.05	-0.11	0	-0.08	-0.08	0.06	-0.1	0.28	-0.12	0.09	-0.11	0.17	1	-0.13	-0.33	-0.04	-0.1	0.22	0.33	0.06	0.37	0.04	0.17	-0.4
TiAm	0.01	-0.02	-0.13	-0.07	0.22	0.05	-0.02	0.13	-0.06	0	-0.14	0.1	-0.06	1	0.23	0.52	0.32	-0.23	-0.17	-0.05	0.01	-0.56	-0.31	-0.12
FrAm	0.15	0	-0.05	0.15	0.22	-0.05	0.16	-0.03	-0.28	0.15	0.04	-0.05	-0.2	0.22	1	0.25	0.11	0.05	0.06	-0.18	0.01	-0.7	-0.45	-0.31
CaCo	0.2	0.03	0.25	-0.12	0.21	-0.03	-0.02	-0.02	0.06	-0.29	-0.07	-0.16	0	0.48	0.24	1	-0.01	-0.02	0.25	-0.11	-0.03	-0.11	-0.07	-0.65
QuRu	-0.05	-0.13	-0.09	-0.29	-0.08	0.11	0.07	0.24	-0.28	0.31	-0.11	-0.04	-0.05	0.33	0.07	-0.09	1	0	-0.11	-0.06	-0.08	-0.15	-0.23	-0.21
CaOv	0.08	-0.13	-0.05	-0.13	-0.05	0.2	-0.12	-0.01	0.19	-0.15	-0.07	0.22	0.26	-0.25	0.03	-0.01	0	1	0.07	-0.39	0.09	-0.63	-0.34	-0.53
AcSi	0.04	0.07	0.11	-0.04	0.07	-0.06	0.03	-0.05	0.04	-0.1	-0.03	-0.25	0.24	-0.14	0.1	0.36	-0.07	0.08	1	-0.28	0.01	NaN	NaN	NaN
BeAl	-0.02	-0.06	-0.1	-0.02	-0.01	-0.05	0.12	0.04	0.12	0.07	0.06	-0.03	0.09	-0.07	-0.19	-0.14	-0.08	-0.4	-0.26	1	-0.07	-0.44	-0.49	-0.19
FrNi	0.17	0.03	-0.05	-0.08	0.15	0.12	0.01	0.19	0.17	-0.06	-0.38	-0.05	0.25	0.03	0.07	-0.04	-0.06	0.11	-0.02	-0.06	1	-0.61	-0.55	-0.44
Sand	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.74	0.59
Elev	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.32
Slope	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Table 7.5: Table of (pseudo) correlations (on and above main diagonal) and partial (pseudo) correlations adjusting for the effects of physical factors (below main diagonal) for the large scale (drift) component.

	tC	**tC(0-15)**	**tC(15-30)**	**tC(FF)**	**K**	**P**	**Ca**	**dNO$_3$**	**dNH$_4$**	**pH**	**AcSa**	**FaGr**	**AcRu**	**TiAm**	**FrAm**	**CaCo**	**QuRu**	**CaOv**	**AcSi**	**BeAl**	**FrNi**	**Sand**	**Elev**	**Slope**
tC	1	0.76	0.48	-0.21	0.44	0.06	0.55	0.42	0.08	0.07	-0.12	0	0.06	-0.01	0.1	0.26	-0.05	0.07	0.05	-0.03	0.23	-0.47	-0.68	0.3
tC(0-15)	0.69	1	0.18	-0.13	0.15	-0.02	0.41	0.33	0.02	0.09	0.15	0.11	-0.06	-0.01	-0.03	0.16	-0.1	-0.14	0.04	-0.05	0.08	-0.8	-0.65	-0.11
tC(15-30)	0.49	0.18	1	0.03	0.03	0.03	0.09	0.16	-0.15	-0.21	0.02	-0.12	-0.03	-0.13	-0.03	0.26	-0.08	-0.05	0.1	-0.1	-0.07	0.38	-0.16	0.49
tC(FF)	-0.22	-0.09	0.04	1	0.02	-0.2	0.03	-0.11	-0.14	-0.04	0.1	0.07	0.01	-0.09	0.1	-0.24	-0.36	-0.13	0.02	-0.04	-0.05	0.75	0.87	-0.15
K	0.36	0.1	0.04	0.02	1	0.33	0.53	0.18	0.23	0.09	-0.29	0.06	0.08	0.14	0.17	0.25	-0.03	-0.09	0.11	-0.02	0.22	-0.89	-0.44	-0.1
P	0.07	0.02	0.03	-0.21	0.28	1	0.05	0.07	0.24	-0.2	-0.24	-0.05	0.08	0.01	-0.03	-0.1	0.13	0.19	-0.03	-0.06	0.11	-0.67	-0.44	0.07
Ca	0.41	0.24	0.09	0.08	0.37	0.1	1	0.53	-0.01	0.54	-0.05	-0.13	0.07	-0.04	0.08	0.23	0.13	-0.15	-0.01	0.12	0.12	-0.76	-0.91	0.1
dNO$_3$	0.38	0.25	0.16	-0.07	0.2	0.14	0.4	1	-0.29	0.34	-0.22	0.26	0.27	0.18	-0.05	0.12	0.25	-0.01	-0.1	0.06	0.2	-0.65	-0.91	0.12
dNH$_4$	0.08	0	-0.16	-0.13	0.32	0.3	-0.02	-0.36	1	-0.37	-0.01	-0.02	-0.18	0.01	-0.28	0.12	-0.31	0.21	0	0.13	0.16	-0.32	0.26	-0.25
pH	-0.02	-0.03	-0.21	0.03	-0.02	-0.17	0.28	0.25	-0.37	1	-0.14	0.01	0.17	-0.01	0.11	-0.09	0.36	-0.18	-0.15	0.08	0	-0.66	-0.93	0.17
AcSa	-0.1	0.09	0	0.19	-0.16	-0.15	-0.1	-0.34	-0.13	-0.2	1	-0.12	-0.28	-0.04	0.1	0.11	-0.08	-0.04	-0.14	0.1	-0.43	-0.45	-0.79	0.42
FaGr	0.02	0.09	-0.13	0.15	0.04	-0.06	-0.15	0.25	0	-0.03	-0.18	1	0.08	0.1	0.01	-0.07	0.03	0.22	-0.3	-0.01	-0.09	0.46	0.71	0.15
AcRu	-0.05	-0.11	0	-0.08	-0.08	0.06	-0.1	0.28	-0.12	0.09	-0.11	0.17	1	-0.13	-0.33	-0.04	-0.1	0.22	0.33	0.06	0.37	0.58	0.87	-0.4
TiAm	0.01	-0.02	-0.13	-0.07	0.22	0.05	-0.02	0.13	-0.06	0	-0.14	0.1	-0.06	1	0.23	0.52	0.32	-0.23	-0.17	-0.05	0.01	-0.57	0.03	-0.5
FrAm	0.15	0	-0.05	0.15	0.22	-0.05	0.16	-0.03	-0.28	0.15	0.04	-0.05	-0.2	0.22	1	0.25	0.11	0.05	0.06	-0.18	0.01	-0.63	-0.83	0.06
CaCo	0.2	0.03	0.25	-0.12	0.21	-0.03	-0.02	-0.02	0.06	-0.29	-0.07	-0.16	0	0.48	0.24	1	-0.01	-0.02	0.25	-0.11	-0.03	-0.28	-0.66	0.36
QuRu	-0.05	-0.13	-0.09	-0.29	-0.08	0.11	0.07	0.24	-0.28	0.31	-0.11	-0.04	-0.05	0.33	0.07	-0.09	1	0	-0.11	-0.06	-0.08	0.27	0.44	-0.47
CaOv	0.08	-0.13	-0.05	-0.13	-0.05	0.2	-0.12	-0.01	0.19	-0.15	-0.07	0.22	0.26	-0.25	0.03	-0.01	0	1	0.07	-0.39	0.09	-0.52	0.2	-0.46
AcSi	0.04	0.07	0.11	-0.04	0.07	-0.06	0.03	-0.05	0.04	-0.1	-0.03	-0.25	0.24	-0.14	0.1	0.36	-0.07	0.08	1	-0.28	0.01	-0.18	0.04	-0.48
BeAl	-0.02	-0.06	-0.1	-0.02	-0.01	-0.05	0.12	0.04	0.12	0.07	0.06	-0.03	0.09	-0.07	-0.19	-0.14	-0.08	-0.4	-0.26	1	-0.07	0.41	0.5	-0.1
FrNi	0.17	0.03	-0.05	-0.08	0.15	0.12	0.01	0.19	0.17	-0.06	-0.38	-0.05	0.25	0.03	0.07	-0.04	-0.06	0.11	-0.02	-0.06	1	-0.62	-0.32	0.13
Sand	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.56	0.04
Elev	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-0.28
Slope	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

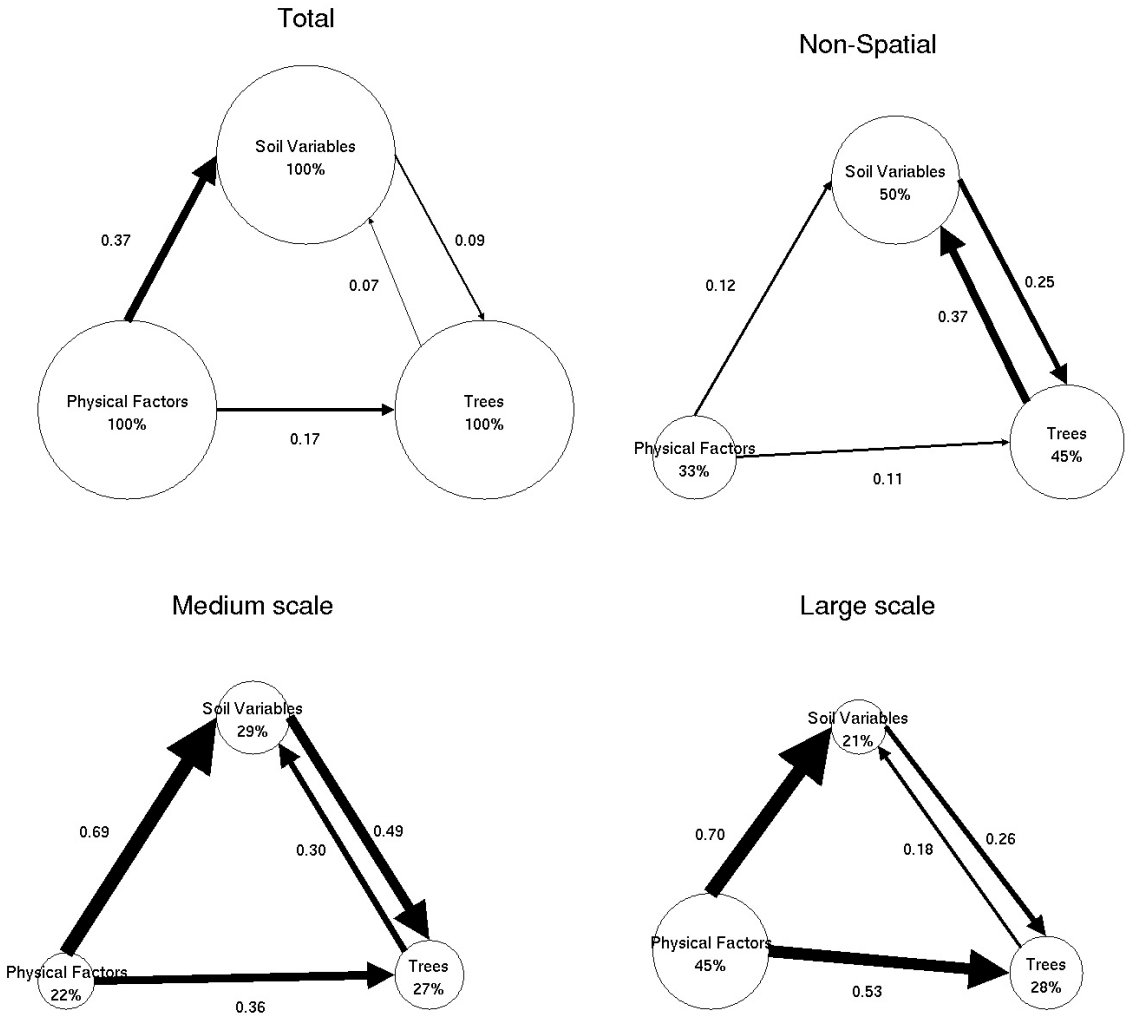

Figure 7.10: Multivariate path diagrams from the partial redundancy analysis showing the interrelationships between physical factors, soil variables and trees for the total variables and for scale-specific components of variation. The size of the circles for each group of variable at each scale is proportional to the percentage of variation for this group at this scale (percentages shown in the circles). The size of the arrows is proportional to the r-square statistic associated with the redundancy or partial redundancy depicted by the arrow (values are shown beside the arrows).

gradient is the association between species such as *Acer saccharinum*, *Fraxinus nigra*, *Carya ovata*, and *Tilia americana* in two patches located near the middle of the grid (see the large scale map for *Acer saccharinum* in Fig. 7.8). These are areas of the Morgan Arboretum that are on a poorly drained lacustrine clay deposit (patch in the middle) or on a very shallow fluvial sand deposit over lacustrine clay (smaller patch near the lower-left edge of the grid). They are characterized by a low sand content and a relatively high ammonification rate. In the large scale pattern of sand content, these patches are superposed with the elevation-sand gradient described above. This explains why, for example, the large scale pattern in *Acer saccharinum* is only partially explained by physical factors (Fig. 7.9).

Thus, in the area covered by the grid, the geomorphology of the site represented by the gradients in elevation and soil texture appears to be the overriding dominant factor that control the pH in the soil, the abundance of calcium and potassium, and the allocation pattern of carbon in the forest floor or in the top layer of mineral soil. Tree species that are specialized for the site conditions described by the gradient show clear patterns of association with the physical factors at large scales. Tree species that occur on different ends of this gradient thus have negative pseudo correlations. For example, *Acer saccharum* and *Acer rubrum* have clearly opposite large scale spatial patterns (Fig. 7.6 and 7.7) as is shown by the strong negative a correlation (-0.9) at large scales. On the contrary, species co-occurring on the same end of the gradient have positive pseudo correlations at large scales. For example, *Acer saccharum* has strong positive pseudo correlations with *Carya cordiformis* (0.67) and *Fraxinus americana* (0.71). These species thus have maps of large scale components with similar patterns (Fig. 7.6 and 7.7).

An interesting aspect of the large scale relationships is that few variables contain a large amount of large scale variation that is not explained by physical factors. The small amount of unexplained large scale variation in the tree species indicates that processes such as dispersal or competitive interactions may not be significantly controlling the distribution at this scale. In other words, the large scale pattern in trees can be considered spatially dependent as opposed to spatially autocorrelated.

Medium scale patterns

Even if less variation in physical factors is captured by the medium scale component, the effects of physical factors on tree species distribution and soil variables are still relatively strong at the medium scale (Figs. 7.9 and 7.10). The direct effects of soil variables on tree species are also noticeably stronger than at the large scale, while the

direct tree effects on soil variables have a similar strength as at the large scale. The medium scale component captured no variation in total carbon (tC), total carbon in the 15-30 cm layer (tC(15-30)), phosphorus (P) and *Acer saccharinum*.

Almost all the medium scale variation in the soil variables (0.8+0.19=0.99) is accounted for by the combination of physical factors and tree species (Fig. 7.10 and 7.9), while a great percentage of the variation in trees (0.85) is accounted for by physical factors and soil variables. Since the interaction between trees and soils is bi-directional, the r-squares linking them should not be seen as indicative of causal relationships. Rather, it indicates the overall concordance between the medium scale patterns in the two groups of variables.

The smaller amount of variation explained by physical factors at the medium scale is evidence that either unmeasured physical factors, management or disturbance related factors or biotic processes such as competitive interactions and dispersal are partly responsible for spatial structuring at this scale. The greater strength of the tree-soil interactions at the medium scale show that the processes that determine the tree species distribution also determine, to a certain extent, the distribution of soil variables at this scale. If biotic processes are controlling the medium scale tree species distribution, then trees are altering the conditions of the soil in return, to induce spatial structuring. Alternatively, if unmeasured physical, management or disturbance related factors affect the medium scale distribution of the soil variables, those factors and the soil variables are jointly controlling the tree species distribution. In any event, it is impossible to assess the exact direction of influence of trees on soils and soils on trees. Also, tree-soil interactions may be exacerbated by the presence of feedback loops, such as those involved in the nutrient cycle between plants and soil.

The interesting patterns occurring within the glacial till areas, where the medium scale patterns are noticeably different than the large scale patterns, are a good example of the complexity of the interactions involved. For example, the correlations between *Acer saccharum* and other species that were strong and positive at large scales are weak (*Carya cordiformis,* 0.05) or negative (*Fraxinus americana*, -0.14) at the medium scale. The correlations between *Acer saccharum* and calcium and total carbon in the 0-15 cm layer that were positive at large scales (Ca, 0.76 and tC(0-15), 0.61) are negative at medium scales (Ca, -0.06 and tC(0-15), -0.47), while the correlations between those variables and *Fraxinus americana* (Ca, 0.62 and tC(0-15), 0.68) remain positive. This would indicate that, within the glacial till, zones that are dominated by *Acer saccharum* have a lower calcium content and have less organic carbon than zones where *Acer saccharum* co-occurs with other species such as *Fraxinus americana* or *Carya cordi-*

formis. In a study conducted in the Morgan Arboretum, sugar maple has been shown to have acidic litter that may lead to gradual deterioration of site conditions (Côté and Fyles, 1994). Although the slightly negative correlation with calcium would seem to provide evidence for this, the positive partial correlation of *Acer saccharum* and pH (0.36) does not support this effect at this scale on our study site. This species has also been shown to favor nitrification and nitrate leaching, which may gradually lead to depletion of the soils' nitrogen status (Lovett and Mitchell, 2004). This is supported by the positive partial correlation with nitrification (0.26). Thus, our results show clear evidence of a non-uniform spatial distribution of species within the glacial till sites that is partly associated with differences in soil conditions. However, the complexity of the interrelationships between and among soil conditions and tree species prevent the direct identification of causal mechanisms.

Non-spatial patterns

The variation that was considered as non-spatial, and that was captured by the nugget effect, in this study is comprised of sampling errors, analytical errors, identification errors for tree species and uncertainty in the global positioning system. The nugget effect also captures a component of the spatial varation in the variables at a scale smaller than the minimum distance between the sampling locations (50 m). Because of its linear behavior at short distances, the spherical model used for modelling the experimental variograms of residuals likely provides a conservative estimate of the amount of spatial small and medium scale variation and an overestimate of the amount of non-spatial varation. Soil variables have often been shown to have a fractal-type of behavior (Burrough, 1983), with denser and denser sampling grids revealing more and more spatially structured small scale variation. A recent study reveals that soil organic matter is spatially structured at the nanometer scale (Lehmann et al., 2008). Thus, we would expect that variables measured without error from a very intensive dataset would contain almost no nugget effect. The spherical model is a conservative compromise that attempts to extrapolate the behavior of the variogram at short distances from the behavior at larger distances. It does captures some, but not all the small scale variation in the dataset. Thus, although the analysis used in this study has distinguished three scales (non-spatial, medium and large scale), in the discussion that follows, 'small scale' will refer to the (unmeasurable) component of the nugget effect that is spatially structured.

Although the non-spatial variation comprises nearly half of the total variation in soil variables and tree species distribution, the amount of variation explained by non-spatial physical factors is small (Fig. 7.10). This may be partly attributed to the lack

of variables in the group of physical factors that account for variations in microsite conditions, such as micro-topography. However, the r-squares between tree species and soil nutrients for the non-spatial variation are larger than at the large scale. This would support the idea that a proportion of the non-spatial variation is actually comprised of small scale spatially structured variation partly shared by tree species and soil variables, given that measurement errors in trees and soils are likely independent.

Trees are spatial entities that have a non-negligible size at small scales (see the discussion on 'Things' in Chapter 1). This simple fact imposes some constraints on the possible small scale correlations observed between trees. As opposed to two soil variables, two trees cannot be found at the same location. The size of the trees imposes a mutual exclusion at scales close to their canopy width which should 'propagate' to small scales. Thus, with a small measurement scale, such as that given by the prism, we cannot expect to obtain a high abundance of multiple tree species at any given spatial location. This is partly evident from the lack of strong partial correlations and the non-spatial partial correlations between trees in Table 7.5.

Cross-scale patterns in trees

As was discussed above, the tree species that are clearly associated with the topographical and geomorphological gradient showed a large amount of variation at large scales. These species could be considered 'specialists' for the site conditions on which they are found. Conversely, species such as *Quercus rubra* and *Betula alleghaniensis* are characterized by a small amount of variation at large scales and were poorly explained by the physical factors. Within the set of conditions covered by the grid, these species could be considered habitat generalists that thrive on sites with varied site conditions, but that are never particularly abundant on any site. Such a pattern is likely generalizable, to some extent, to different sites and different scales. Within the set of conditions covered by a given area at a given scale, the specialist species will be associated with specific site conditions and will likely show clear patterns of spatial structuring, depending on the spatial structuring of the physical factors. Generalist species will have spatial distributions that are more independent of the environmental conditions, and will thus exhibit less spatial structuring.

Cross-scale patterns in soil variables

The amount of spatially structured variation in soil variables appears to be related to the proportion of their overall pools in the physical environment and to the importance

of fluxes between trees and soils. Soil variables that contain a lot of spatially structured variation are tightly linked to the physical factors of the site. For example, calcium and potassium are mineral nutrients available in higher concentrations on glacial tills due to the presence of a calcareous bedrock and the good water holding capacity of the soil that prevent leaching. Variables that are not spatially structured (e.g. tC and P) are greatly affected by biotic processes and are present in low quantities in the physical environment. Soil variables with moderate spatial structuring (e.g. pH, dNO_3) are affected both by biotic processes and by the physical environment.

Soil organic carbon was shown to be related, among other factors, to soil texture (Arrouays et al., 1995), topography (Bergstrom et al., 2001; Arrouays et al., 1998) and tree species (Finzi et al., 1998; Giardina et al., 2001). In the Arboretum, these factors appear more important for determining the spatial structure in the pattern of carbon distribution within the profile than the total amount of carbon in the soil. The combination of topography and soil texture determine to a great extent whether the carbon will be located in the forest floor (mor or moder humus form on sand) or in the top layer of mineral soil (mull humus form on loams and clay soils). The presence of species such as *Fagus grandifolia* that have a litter that is low in nutrients and high in lignin on the moder humus form and of species with more liberal nutrient use strategies, such as *Acer saccharum*, on the mull humus forms help to maintain or enhance this allocation pattern. Thus, the allocation pattern of carbon in the forest floor (tC(FF)) or in the top layer of mineral soil (tC(0-15)) is spatially structured. The lack of spatial structuring in the pooled amounts of carbon in the soil (tC) seem to indicate that, although there can be significant spatial structuring in turnover rates, inputs to and from the soil are in a relative equilibrium throughout the site. The higher decomposition rates in some areas (e.g. on the glacial tills), are necessarily matched by higher inputs of carbon through litterfall, and vice-versa. This is evidence that soil organic carbon pools may be ultimately controlled by factors, such as climate, that operate at scales larger than the Arboretum.

7.4 Concluding remarks

In this chapter, I have demonstrated that the combination of multi-scale analysis (CRAD in particular) and causal modelling can be useful tools to generate understanding in ecology. For one, these methods allow us to separate the total variation in a dataset into components that are intelligible and predictable from components that are not. The large scale variation was shown to be related to physical factors

and correlations between variables were in general stronger. Non-spatial components of variation were weakly correlated and poorly explained by physical factors. Interestingly and perhaps unfortunately, a large proportion of the total variation was found in the non-spatial component.

These results highlight that numerical models that try to predict the quantities of soil nutrients or the abundance of trees species at any location on the landscape with reasonable confidence may fail due to the large amount of non-spatial or micro-scale variation in many variables. Such models however may have more success in explaining the large scale patterns. For total soil organic carbon even large scale modelling may be prove difficult due to the lack of spatial structure and its lack of relationships with physical factors. However, the pattern of allocation of carbon in the top soil or in the forest floor was much more spatially structured and easily predictable at large scales. Similarly, the spatial distribution of specialist species may be more easily explained than that of generalists. Nevertheless, since soil properties and tree species abundances are more easily sampled with small sampling units, the combination of causal modelling and multi-scale analysis may provide some tools to extract information on large scale patterns from small scale datasets that would be useful to calibrate larger scale models.

The results from this study suggest that different metacommunity paradigms may apply at different scales. At large scales, environmental determinism was the dominant control on both the tree species distribution and the soil variables, bringing support for the species sorting paradigm. At the medium scale, a larger amount of unexplained variation in the tree species would suggest a greater effect of dispersal and a support for the mass effect paradigm. A purely neutral model in which spatial variation in trees would be completely independent of physical factors, was not supported by this study. Our study points out that the unexplained small and medium scale variation in trees and soil variables is relatively tightly linked.

Understanding the interrelationships between trees and soils, and the changes in the nature of those relationships with scale, would be faciliated by coherent and well developed conceptual frameworks that would lead to the formulation of testable hypotheses. New theories are needed that integrate the metacommunity concepts and ecosystem dynamics and tackle the complicating, or facilitating, effects of scale.

Chapter 8

Conclusion

"Statistics are no substitute for judgment." - Henry Clay

When I started this project (which was originally a masters project), I was, similarly to many ecologists, unfamiliar with geostatistics and with most methods of multi-scale analysis. As an ecologist looking for methods of spatial analysis, the published statistical literature proved to be intimidating and somewhat overwhelming. However, geostatistics appeared to offer a very elaborate framework for spatial analysis and regionalized multivariate analysis had been fairly recently presented as an interesting approach for multi-scale analysis. At the time, no statistical software were easily accessible to perform regionalized multivariate analysis. This lead me and a research collaborator (Dr. Bernard Pelletier) to write new Matlab programs, and adapt existing ones, to perform coregionalization analysis. While this task proved difficult and time-consuming, it allowed us to develop an improved knowledge of geostatistical techniques in general and of coregionalization analysis in particular. Unfortunately, when we attempted to analyze ecological datasets with the Matlab programs, the results seemed very sensitive to the particular choices of variogram models, lag distances on the variograms, number of variables in the analysis, and several other factors. Moreover, when we attempted to perform multiple regression or redundancy analysis from sill matrices, we consistently obtained r-squares close to one. As ecologists, this appeared surprising and unsatisfactory.

These results forced us to question the theory underlying regionalized multivariate analysis and to look for alternatives. With the help of an experienced spatial statistician (Dr. Pierre Dutilleul), we explored several avenues in an attempt to identify the problems associated with the methods and to find possible solutions. Originally, this quest was conducted in a somewhat orthodox framework in which the statistical models are ultimately seen as a reference for truth, or as objective depictions of reality. By

doing this we were missing out on two important and interrelated aspects. Firstly, there are no spherical models in nature; statistical models are tools to represent reality, but they are not the reality. Secondly, in multi-scale analysis, the separation of total variables into scale-specific components of variation is not unique and inherently subjective. These two aspects are tied to the questions of uncertainty and inference. Any measure of uncertainty will depend on the choice of models used to perform the analysis, and on the foci of inference implicit in them. Matheron's 'Estimating and choosing' (1978) provided an interesting perspective on these problems by emphasizing the importance of proper choices of models and introducing the concept of micro-ergodicity to discuss the problem of inference in a limited sampling domain. It then appeared that the probabilistic model that constitutes the linear model of coregionalization had an underlying focus of inference that may not be appropriate in practical situations where data is obtained in a limited sampling domain (Chapters 4 and 5).

Another area of exploration in the search for solutions was the use of cokriged regionalized components, or versions of them obtained in modified cokriging systems, in place of coregionalization matrices to perform multi-scale analysis. However, as I have highlighted in Chapter 3, cokriged components for different scales are not independent and provide some redundant information betwen scales. I have also shown that the only way to obtain regionalized components that comply with their representations in the linear model of coregionalization is through the simulation of all components corresponding to the different variogram functions jointly. Although the approach can prove useful for a variety of purposes, the results obtained from it are greatly dependent on the adequacy of the linear model of coregionalization.

This then lead to the development of coregionalization analysis with a drift in an attempt to obtain a method of multi-scale analysis that is coherent, and that has a reasonable focus of inference that should be more appropriate in practical applications. By the use of a deterministic approach to capture the large scale components, CRAD addresses the issues of uncertainty, inference and unreasonable assumptions inherent in the linear model of coregionalization. Yet, the approach retains the ability of the linear model of coregionalization to describe the patterns and relationships at small scales and to separate the small scale spatial and the non-spatial components of variation. These aspects, combined with the possibility of using CRAD with regular or irregular transects and grids, should give the method an advantage over other existing methods of multi-scale analysis (Chapter 2).

Matheron (1978) refers to the external objectivity of models as 'the sanction of practice', or the ability of models to succesfully attain research objectives of practitioners

through repeated uses in a wide variety of situations. Although the use of CRAD has still been limited, we have started to assess the external objectivity of the method. The two case studies (Chapters 6 and 7), have shown that CRAD was able to decompose the variation in total variables showing very diverse types of spatial distributions. The CRAD approach was successfully used to perform scale-specific bivariate correlations in an agronomic setting (Chapter 6), scale-specific redundancy analysis to study the relationships between physical factors and forest biodiversity (Pelletier et al., 2009a,b) and partial redundancy analysis in the framework of multivariate causal modelling to study the interrelationships between physical factors, soil variables and tree species distribution (Chapter 7).

This journey has taught me the importance of improving linkages between concepts and statistical tools (Fig. 1.3). Some of the methods of multi-scale analysis discussed in Chapter 2 appeared to be lacking strong conceptual foundations and their application to ecological datasets may lead to misleading or confusing results. The development and use of methods of analysis should be guided by thorough conceptual frameworks that describe in detail the perceptual filter used by the scientist when attempting to describe and understand natural systems. These frameworks should be organized around precise definitions of concepts and terms. Practitioners should not have a blind faith in statistical methods, and they should not see the principles of statistics as 'dogmas' that apply to every situation. Not only should practitioners have better *knowledge* of statistical methods, they should also strive to improve their *understanding* of models and assumptions underlying them. In fact, statistical methods of analysis could be seen as mere extensions to their conceptual frameworks or even as mathematical representations of them. More collaboration is thus needed between statisticians and practitioners in order to obtain statistical methods that comply with the needs and frameworks of practitioners.

The localized aspect of spatial observational data may seem like an important obstacle for reaching generalizations and for improving our understanding of natural systems. However, we cannot be content with remaining completely descriptive, blocked in the contingency of the local patterns. In this thesis, I have supported the fact that multi-scale analysis, in combination with methods of multivariate analysis, may help us understand the particular conditions under which patterns occur and should allow us to formulate better hypotheses about mechanisms or processes. However, the elaboration of coherent conceptual frameworks in the context of spatial observational studies can be a daunting task. Although issues of space and scale have received tremendous attention in the recent ecological literature, it seems that the development of ecological theories

that address those issues is still in its infancy. The context in which spatial observational studies are carried out, much like landscape ecology studies, forces ecologists to rethink exisiting research paradigms. Until ecological theories are further developed to generate precise expectations on issues of spatial analysis and scale, the use of tools for spatial and multi-scale analysis to generate ecological understanding may remain difficult.

To paraphrase Gomory (1995), multi-scale analysis may prove to be an appropriate statistical mechanic that can help separate the knowable unknown from the truly unknowable. The gradient of 'knowability' may be associated with a gradient in scale. Results from the two case studies in this thesis and the one in the Pelletier et al. (2009a,b) have shown that correlations between variables, or r-squares between groups of variables, are generally greater at the large scale than at the small scale or non-spatial components. This result can likely be generalizable to many types of natural systems. Physical factors operating at larger scales may define the boundary conditions constraining the dynamics of the natural systems and should impose some spatial structuring that will be shared by multiple variables in the system. Small scale processes may be more stochastic in nature and lead to patterns that are harder to predict. While the empirical identification of processes underlying larger scale patterns may be in the realm of the knowable unknown, processes controlling the exact small scale spatial structuring may prove to be truly unknowable. The true challenge of ecologists may be to develop adequate frameworks to explain the variation at intermediate scales, where patterns are structured but controlled by a large number of interrelated factors.

Contributions to knowledge

In the introductory chapter, and throughout this thesis, I have contributed to the development of a conceptual framework for multi-scale analysis that is simple, practical and coherent. I believe that I have contributed to bridging the gap between concepts, statistical tools and ecological understanding.

Statistical contributions

- In Chapter 2, I provide a comparison of multiple methods for multi-scale analysis. Through a simple classification scheme, and a common terminology and dataset, methods are compared on even grounds and in an framework that should be accessible to practitioners.

- I have shown that despite important differences in their mathematical represen-

tations, PCNM and DFT are nearly equivalent methods for performing multi-scale analysis. Procedures classified as orthogonal decomposition methods (DFT, PCNM, DWT) provide similar results when multiple scale-specific components are superimposed to perform decomposition of the total variables into a small number of scales.

- I have shown that inferring scales of variation from the graphical methods of multi-scale analysis can be problematic.

- The representation of the variogram functions as spectral density functions that are a linear function of wavelength was, to my knowledge, presented for the first time in this thesis in Chapter 2. Through this intuitive representation, I have shown that the range of the variogram functions are inadequate descriptors of the scales of variation. Variograms represent a probability distribution of scales that is greatly dependent on the type of variogram function and that always has a mode lower than the range. This was the first account of the exact interpretation of the variogram in terms of scales of variation.

- In Chapter 3, I present the first detailed published method for expressing the spatial uncertainty of regionalized components and I use it to develop the first procedure for simulating regionalized components.

- While methods of coregionalization analysis have been extensively used in several fields of research, the characterization and quantification of uncertainty associated with the method had never been undertaken before. In Chapter 4, I provide the first detailed account of the different constituents of uncertainty and present several mathematical expressions that are not found elsewhere in the literature. The high levels of uncertainty highlighted in this Chapter may have important implications for the future use of coregionalization analysis.

- In Chapter 5, I have shown that the issue of inference should be a particular area of concern when using spatial probabilistic models within a limited sampling domain. The notions of ergodicity and micro-ergodicity can be useful concepts for discussing this issue.

- I have contributed to the development of coregionalization analysis with a drift, which is probably the most carefully elaborated method of multi-scale analysis available today. This method is based on strong statistical and conceptual foun-

dations and provides a flexible approach for multi-scale analysis that should be easy to use by practitioners.

Contributions in ecology and agronomy

- To my knowledge, the agronomic case study in Chapter 6 presents the first use of methods of multi-scale analysis to study the relationship between scales of variation in soil variables and their temporal correlations. The main result from this study was that variables that contain a large (small) amount of variation at large scales tend to have patterns that are more (less) correlated in time. This result is likely generalizable to many types of natural systems.
- The forest ecology case study was a first attempt to study the interactions between physical factors, soil variables and tree species distributions at multiple scales through the use of scale-specific multivariate causal modelling.

Acknowledgements

I am grateful to a large number of people who have made this thesis possible. Many thanks to my supervisor Dr. Jim Fyles for giving me the freedom to let my ideas evolve, for helping my ramblings become narratives, for his tremendous patience and for providing financial assistance. Huge thanks to Dr. Bernard Pelletier for the endless hours of discussions that have allowed the emergence and development of many ideas presented in this thesis. Thanks for the Matlab code, for the review work, for the discussions in the car and for being a good friend. Numerous thanks to Dr. Pierre Dutilleul, for helping the statistical and mathematical aspects of this thesis take shape, for his extensive and meticulous review work, and for his patience and understanding. Without the collaboration of these three people, this thesis wouldn't exist.

Un énorme merci à ma merveilleuse Sylviane, pour sa patience, sa joie de vivre, son immense support, et l'aide pour la traduction du résumé. Merci, à mon père Claude, ma mère Bernadette et ma soeur Catherine pour leur support moral et pour n'avoir jamais cessé de croire en moi, même dans les moments extrêmement difficiles. Merci aussi pour les corrections. Merci au reste de ma famille et mes amis pour leur soutien.

Thanks to Hélène Lalande for her assistance in the laboratory and to the following people who have helped for the data collection or for laboratoy analysis (in no particular order): Johanne Lambert, Marie Ducharme, Caroline Dufour, Line Plourde, Simon Bourgeois, Pierre-Olivier Quesnel, David Maneli, Andrea Grant, Ilkka Numiranta, Sophia, Jessica Epstein, Anne Desrochers, Liam, Natalka Iwanicky, Elizabeth, Kerry Knettle, Chantal Duval, Reagan Guilfoyle, Virginie Demers. Thanks to Ernst Bernhardt for allowing access to his fields for the collection of the agronomic dataset.

During my doctoral work, I received fellowships from the Fond Québecois pour la Nature et la Technologie, from the Natural Sciences and Engineering Research Council of Canada, as well as a McGill Majors Fellowship and a Tomlinson Fellowship. The work presented in Chapters 3 and 6 was partly funded by the Potash and Phosphate Institute of Canada. The computer resources necessary to complete some of my research were purchased thanks to a Canada Foundation for Innovation (CFI) grant. I am very grateful to those funding sources.

List of acronyms

CI Confidence interval

CRA Coregionalization analysis

CRAD Coregionalizationa analysis with a drift

CWT Continuous wavelet transform

DBEM Distance based eigenvector maps

DFT Discrete Fourier transform

DWT Discrete wavelet transform

GLS Generalized least squares

iid Identically and independently distributed

LIDAR Laser detection and ranging

LMC Linear model of coregionalization

LMR Linear model of regionalization

MAUP Modifiable areal unit problem

OLS Ordinary least squares

PCA Principal components analysis

PCNM Principal coordinate analysis of neighbourhood matrices

RDA Redundancy analysis

SSM Site-specific management

TSA Trend surface analysis

TTLQV Two-term local quadrat variance

3TLQV Three-term local quadrat variance

Bibliography

Alabert, F., 1987. The practice of fast conditional simulations through the LU decomposition of the covariance-matrix. Mathematical Geology 19 (5), 369–386.

Allen, T. F. H., Hoekstra, T. W., 1992. Toward a Unified Ecology. Columbia University Press.

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

Arrouays, D., Daroussin, J., Kicin, J., Hassika, P., 1998. Improving topsoil carbon storage prediction using a digital elevation model in temperate forest soils of france. Soil Science 163 (2), 103–108.

Arrouays, D., Vion, I., Kicin, J., 1995. Spatial analysis and modeling of topsoil carbon storage in temperate forest humic loamy soils of france. Soil Science 159 (3), 191–198.

Beatty, S. W., 1984. Influence of microtopography and canopy species on spatial patterns of forest understory plants. Ecology 65 (5), 1406–1419.

Bell, G., Lechowicz, M., Appenzeller, A., Chandler, M., Deblois, E., Jackson, L., Mackenzie, B., Preziosi, M., Schallenberg, M., Tinker, N., 1993. The spatial structure of the physical environment. Oecologia 96, 114–121.

Bellier, E., Monestiez, P., Durbec, J.-P., Candau, J.-N., 2007. Identifying spatial relationships at multiple scales: principal coordinates of neighbour matrices (PCNM) and geostatistical approaches. Ecography 30 (3), 285–399.

Bergstrom, D. W., Monreal, C. M., St. Jacques, E., 2001. Spatial dependence of soil organic carbon mass and its relationship to soil series and topography. Canadian Journal of Soil Science 81, 53–62.

Binkley, D., Giardina, C., 1998. Why do tree species affect soils? The warp and woof of tree-soil interactions. Biogeochemistry 42, 89–106.

Bocchi, S., Castrignano, A., Fornaro, F., Maggiore, T., 2000. Application of factorial kriging for mapping soil variation at field scale. European Journal of Agronomy 13 (4), 295–308.

Boettcher, S. E., Kalisz, P. J., 1990. Single-tree influence on soil properties in the mountains of easter kentucky. Ecology 71 (4), 1365–1372.

Borcard, D., Legendre, P., 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153 (1-2), 51–68.

Borcard, D., Legendre, P., Avois-Jacquet, C., Tuomisto, H., 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology 85 (7), 1826–1832.

Borcard, D., Legendre, P., Drapeau, P., 1992. Partialling out the spatial component of ecological variation. Ecology 73 (3), 1045–1055.

Bourennane, H., Salvador-Blanes, S., Cornu, S., King, D., 2003. Scale of spatial dependence between chemical properties of topsoil and subsoil over a geologically contrasted area (massif central, france). Geoderma 112 (3-4), 235–251.

Bourgault, G., 1996. Probability field for the post-processing of stochastic simulations. Mathematical Geology 28 (6), 723–734.

Box, G. E. P., Cox, D. R., 1964. An analysis of transformations. Journal of the Royal Statistical Society Series B-Statistical Methodology 26 (2), 211–252.

Brown, J. H., 1984. On the relationship between abundance and distribution of species. The American Naturalist 124 (2), 255–279.

Brus, D. J., de Gruijter, J. J., 1997. Random sampling or geostatistical modelling? choosing between design-based and model-based sampling strategies for soil. Geoderma 80 (1-2), 1–44.

Burgess, T., Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties; I The semi-variogram and punctual kriging. European Journal of Soil Science 31 (2), 315–331.

Burnett, M., August, P., J.H., J. B., Killingbeck, K., 1998. The influence of geomorphological heterogeneity on biodiversity. i. a patch-scale perspective. Conservation Biology 12 (2), 363–370.

Burrough, P. A., 1983. Multiscale sources of spatial variation in soil. i. the application of fractal concepts to nested levels of soil variation. Journal of Soil Science 34 (3), 577–597.

Burrough, P. A., Bouma, J., Yates, S. R., 1994. The state of the art in pedometrics. Geoderma 62, 311–326.

Cahn, M., Hummel, J., Brouer, B., 1994. Spatial analysis of soil fertility for site-specific crop management. Soil Science Society of America Journal 58, 1240–1248.

Carter, M. R., 1993. Soil Sampling and Methods of Analysis. Lewis Publishers.

Castrignano, A., Giugliarini, L., Risaliti, R., Martinelli, N., 2000a. Study of spatial relationships among some soil physico-chemical properties of a field in central italy using multivariate geostatistics. Geoderma 97 (1-2), 39–60.

Castrignano, A., Goovaerts, P., Lulli, L., Bragato, G., 2000b. A geostatistical approach to estimate probability of occurrence of tuber melanosporum in relation to some soil properties. Geoderma 98 (3-4), 95–113.

Chilès, J.-P., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley, New York.

Chokkalingam, U., White, A., 2001. Structure and spatial patterns of trees in old-growth northern hardwood and mixed forests of northern maine [review]. Plant Ecology 156 (2), 139–160.

Cooper, G., 1998. Generalizations in ecology: a philosophical taxonomy. Biology and Philosophy 13, 555–586.

Côté, B., Fyles, J. W., 1994. Nutrient concentration and acid-base status of leaf litter of tree species characteristic of the hardwood forest of southern Quebec. Canadian Journal of Forest Research 24, 192–196.

Couteron, P., Ollier, S., 2005. A generalized, variogram-based framework for multi-scale ordination. Ecology 86 (4), 828–834.

Crawley, M. J., Harral, J. E., 2001. Scale dependence in plant biodiversity. Science 291 (5505), 864–868.

Cressie, N. A., 1986. Kriging nonstationary data. Journals of the American Statistical Association 81 (395), 625–634.

Cressie, N. A., 1990. The origins of kriging. Mathematical Geology 22 (3), 239–252.

Cressie, N. A., 1993. Statistics for Spatial Data. Wiley, New York.

Crozier, C. R., Boerner, R. E. J., 1984. Correlations of understory herb distribution patterns with microhabitats under different tree species in a mixed mesophytic forest. Oecologia 62, 337–343.

Csillag, F., Kabos, S., 2002. Wavelets, boundaries, and the spatial analysis of landscape pattern. Ecoscience 9 (2), 177–190.

Cullinan, V. I., Simmons, M. A., Thomas, J. M., 1997. A bayesian test of hierarchy theory: scaling up variability in plant cover from field to remotely sensed data. Landscape Ecology 12, 273–285.

Dale, M. R. T., 1999. Spatial Pattern Analysis in Plant Ecology. Vol. 326 of Cambridge studies in ecology. Cambridge University Press.

Dale, M. R. T., 2000. Lacunarity analysis of spatial pattern: a comparison. Landscape Ecology 15, 467–478.

Dale, M. R. T., Mah, M., 1998. The use of wavelets for spatial pattern analysis in ecology. Journal of Vegetation Science 9 (6), 805–814.

Davis, M. W., 1987. Production of conditional simulations via the lu triangular decomposition of the covariance-matrix. Mathematical Geology 19 (2), 91–98.

de Gruijter, J. J., ter Braak, C. J. F., 1990. Model-free estimation from spatial samples: A reappraisal of classical sampling theory. Mathematical Geology 22 (4), 407–415.

de Gruijter, J. J., Walvoort, D. J. J., van Gaans, P. F. M., 1997. Continuous soil maps - a fuzzy set approach to bridge the gap between aggregation levels of process and distribution models. Geoderma 77, 169–195.

Delgado, J., Khosla, R., Bausch, W., Westfall, D., Inman, D., 2005. Nitrogen fertilizer management based on site-specific management zones reduces potential for nitrate leaching. Journal of Soil and Water Conservation 60 (6), 402–410.

Denny, M. W., Helmuth, B., Leonard, G. H., Harley, C. D. G., Hunt, L. J. H., Nelson, E. K., 2004. Quantifying scale in ecology: Lessons from a wave-swept shore. Ecological Monographs 74 (3), 513–532.

Diamond, J. M., 1986. Overview : laboratory experiments, field experiments and natural experiments. In: Diamond, J. M., Case, T. J. (Eds.), Community Ecology. Harper & Row, pp. 3–22.

Diggle, P. J., 1990. Time Series: a Biostatistical Introduction. Oxford University Press, New York.

Dobermann, A., Goovaerts, P., George, T., 1995. Sources of soil variation in an acid Ultisol of the Philippines. Geoderma 68 (3), 173–191.

Dobermann, A., Goovaerts, P., Neue, H., 1997. Scale-dependent correlations among soil properties in two tropical lowland rice fields. Soil Science Society of America Journal 61 (5), 1483–1496.

Dray, S., Legendre, P., Peres-Neto, P. R., 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196, 483–493.

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

Dwyer, L. M., Merriam, G., 1982. Influence of topographic heterogeneity on deciduous litter decomposition. Oikos 37, 228–237.

Feagin, R. A., Wu, X. B., Feagin, T., 2007. Edge effects in lacunarity analysis. Ecological Modelling 201 (3-4), 262–268.

Ferrari, J. B., Sugita, S., 1996. A spatially explicit model of leaf litter fall in hemlock-hardwood forests. Canadian Journal of Forest Research 26 (11), 1905–1913.

Finzi, C., Breemen, N., Canham, C., 1998. Canopy tree-soil interactions within temperate forests: species effects on soil carbon and nitrogen. Ecological Applications 8 (2), 440–454.

Ford, E. D., Renshaw, E., 1984. The interpretation of process from pattern using two-dimensional spectral-analysis - modeling single species patterns in vegetation. Vegetatio 56 (2), 113–123.

Fortin, M.-J., 1999. Spatial statistic in landscape ecology. In: Klopatek, J. M., Gardner, R. H. (Eds.), Landscape Ecological Analysis : Issues and Applications. Springer, pp. 253–279.

Fraisse, C. W., Sudduth, K. A., Kitchen, N. R., 2001. Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. Transactions of the ASAE 44 (1), 155–166.

Frelich, L. E., Calcote, R. R., Davis, M. B., Pastor, J., 1993. Patch formation and maintenance in an old-growth hemlock-hardwood forest. Ecology 74 (2), 513–527.

Gersper, P. L., Holowaychuk, N., 1971. Some effects of stem flow from forest canopy trees on chemical properties of soils. Ecology 52 (4), 691–702.

Giardina, C. P., Ryan, M. G., Hubbard, R. M., Binkley, D., 2001. Tree species and soil textural controls on carbon and nitrogen mineralization rates. Soil Science Society of America Journal 65 (4), 1272–1279.

Gittins, R., 1968. Trend-surface analysis of ecological data. The Journal of Ecology 56 (3), 845–869.

Gomez-Hernandez, J. J., Journel, A. G., 1994. Joint sequential simulation of multi-Gaussian fields. Kluwer Academic Press, Dordrecht, pp. 85–94.

Gomory, R. E., 1995. The known, the unknown and the unknowable. Scientific American 272 (6), 120–120.

Goovaerts, P., 1992. Factorial kriging analysis: a useful tool for exploring the structure of multivariate spatial soil information. European Journal of Soil Science 43 (4), 597 – 619.

Goovaerts, P., 1993. Factorial kriging analysis of springwater contents in the dyle river basin, belgium. Water Resources Research 23 (7), 2115–2125.

Goovaerts, P., 1994. Study of spatial relationships between two sets of variables using multivariate geostatistics. Geoderma 62 (1-3), 93–107.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89 (1-2), 1–45.

Goovaerts, P., 2000. Estimation or simulation of soil properties? an optimization problem with conflicting criteria. Geoderma 97 (3-4), 165–186.

Goovaerts, P., 2001. Geostatistical modelling of uncertainty in soil science. Geoderma 103 (1-2), 3–26.

Goovaerts, P., Sonnet, P., Navarre, A., 1993. Factorial kriging analysis of springwater contents in the Dyle River basin, Belgium. Water Resoures Research 29 (7), 2115–2125.

Goovaerts, P., Webster, R., 1994. Scale-dependent correlation between topsoil copper and cobalt concentrations in scotland. European Journal of Soil Science 45 (1), 79–95.

Goulard, M., 1989. Inference in a coregionalization model. In: Armstrong, M. (Ed.), Geostatistics, Vol. 1. Vol. 1. Kluwer Academic Publishers, pp. 397–408.

Goulard, M., Voltz, M., 1992. Linear coregionalization model - tools for estimation and choice of cross-variogram matrix. Mathematical Geology 24 (3), 269–286.

Gregoire, T. G., 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research 28 (10), 1429–1447.

Greig-Smith, P., 1952. The use of random and contiguous quadrats in the study of the structure of plant communities. Annals of Botany 16 (2), 293–316.

Griffith, D., 1992. What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. L'Espace géographique 3, 265–280.

Grinsted, A., Moore, J. C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes in Geophysics 11 (5-6), 561–566.

Halley, J. M., Hartley, S., Kallimanis, A. S., Kunin, W. E., Lennon, J. J., Sgardelis, S. P., 2004. Uses and abuses of fractal methodology in ecology. Ecology Letters 7 (3), 254–271.

Hart, S. C., Stark, J. M., Davidson, E. A., Firestone, M. K., 1994. Nitrogen mineralization, immobilization, and nitrification. In: Mickelson, S. H., Bigham, J. M. (Eds.), Methods of Soil Analysis; Part 2 Microbial and Biochemical Properties. Ch. 42, pp. 985–1018.

Heiniger, R. W., McBride, R. G., Clay, D. E., 2003. Using soil electrical conductivity to improve nutrient management. Agronomy Journal 95 (2003), 508 – 519.

Hendershot, W. H., Lalande, H., Duquette, M., 1993. Soil reaction and exchangeable acidity. In: Carter, M. R. (Ed.), Soil Sampling and Methods of Analysis. Lewis Publishers, pp. 141–145.

Heuvelink, G. B. M., Webster, R., 2001. Modelling soil variation: past, present, and future [review]. Geoderma 100 (3-4), 269–301.

Hill, M. O., 1973. The intensity of spatial pattern in plant communities. The Journal of Ecology 61 (1), 225–235.

Holling, C. S., 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. Ecological Monographs 62 (4), 447–502.

Hubbard, B. B., 1998. The World According to Wavelets: the Story of a Mathematical Technique in the Making. A.K. Peters.

Isaaks, E. H., Srivastava, R. M., 1989. Applied Geostatistics. Oxford University Press, New York.

Jelinski, D. E., Wu, J. G., 1996. The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology 11 (3), 129–140.

Johnson, C. K., Doran, J. W., Duke, H. R., Wienhold, B. J., Eskridge, K. M., Shanahan, J. F., 2001. Field-scale electrical conductivity mapping for delineating soil condition. Soil Science Society of America Journal 65, 1829 – 1837.

Jorgensen, S. E., 1997. Ecological Modelling. Vol. 100. Elsevier Science.

Jorgensen, S. E., Muller, F., 2000. Handbook of Ecosystem Theories and Management. Environmental and ecological modeling. Lewis Publishers.

Journel, A. G., Huijbregts, C., 1978. Mining Geostatistics. Academic Press.

Keitt, T. H., Urban, D. L., 2005. Scale-specific inference using wavelets. Ecology 86, 2497 – 2504.

Kershaw, K. A., 1963. Pattern in vegetation and its causality. Ecology 44 (2), 377–388.

Krige, D. G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. J. of the Chem., Metal. and Mining Soc. of South Africa 52 (6), 119–139.

Lark, R. M., Papritz, A., 2003. Fitting a linear model of coregionalization for soil properties using simulated annealing. Geoderma 115 (3-4), 245–260.

Lark, R. M., Webster, R., 2001. Changes in variance and correlation of soil properties with scale and location: analysis using an adapted maximal overlap discrete wavelet transform. European Journal of Soil Science 52 (4), 547–562.

Larocque, G., Dutilleul, P., Pelletier, B., Fyles, J. W., 2006. Conditional Gaussian co-simulation of regionalized component of soil variation. Geoderma 134 (1), 1–16.

Larocque, G., Dutilleul, P., Pelletier, B., Fyles, J. W., 2007. Characterization and quantification of uncertainty in coregionalization analysis. Mathematical Geology 39, 263–288.

Larsen, B., Bliss, L. C., 1998. An analysis of structure of tree seedling populations on a Lahar. Landscape Ecology 13 (5), 307–323.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm. Ecology 74 (6), 1659–1673.

Legendre, P., Legendre, L., 1998. Numerical Ecology, 2nd Edition. Elsevier.

Lehmann, J., Solomon, D., Kinyangi, J., Dathe, L., Wirick, S., Jacobsen, C., 2008. Spatial complexity of soil organic matter forms at nanometer scales. Nature Geosciences 1, 238–242.

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., Gonzalez, A., 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7 (7), 601 – 613.

Levin, S., 1992. The problem of pattern and scale in ecology. Ecology 73 (6), 1943–1967.

Levins, R., Lewontin, R. C., 1985. The Dialectical Biologist. Harvard University Press.

Lin, Y. P., 2002. Multivariate geostatistical methods to identify and map spatial variations of soil heavy metals. Environmental Geology 42 (1), 1–10.

Lodhi, M. A., 1977. The influence and comparison of individual forest trees on soil properties and possible inhibition of nitrification due to intact vegetation. American Journal of Botany 64 (3), 260–264.

Lovett, G. M., Mitchell, M. J., 2004. Sugar maple and nitrogen cycling in the forests of eastern North America. Frontiers in Ecology and the Environment 2 (2), 81–88.

Ludwig, J. A., Goodall, D. W., 1978. A comparison of paired- with blocked-quadrat variance methods for the analysis of spatial pattern. Vegetatio 38 (1), 49–59.

Mac Nally, R., Quinn, G., 1998. Symposium introduction: The importance of scale in ecology. Australian Journal of Ecology 23, 1–7.

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

Marchant, B. P., Lark, R. M., 2004. Estimating variogram uncertainty. Mathematical Geology 36 (8), 867–898.

Marcotte, D., 1995. Conditional simulation with data subject to measurement error - post-simulation filtering with modified factorial kriging. Mathematical Geology 27 (6), 749–762.

Matheron, G., 1962. Traité de Géostatistique Appliquée. Tome 1. Mémoires du Bureau de recherches géologiques et minières ; no 14.

Matheron, G., 1963a. Principles of geostatistics. Economic Geology 58 (8), 1246–1266.

Matheron, G., 1963b. Traité de Géostatistique Appliquée. Tome 2, le Krigeage. Mémoires du Bureau de recherches géologiques et minières ; no 24.

Matheron, G., 1965. Les Variables Régionalisées et leur Estimation : Une Application de la Théorie des Fonctions Aléatoires aux Sciences de la Nature. Masson.

Matheron, G., 1970. La théorie des variables régionalisées, et ses applications. Les Cahiers du Centre de morphologie mathématique de Fontainebleau ; fasc. 5. Ecole nationale supérieure des mines.

Matheron, G., 1971. The theory of regionalized variables and its applications. Les cahiers du Centre de Morphologie Mathématique de Fontainebleau; No. 5. Ecole nationale supérieure des mines.

Matheron, G., 1978. Estimer et choisir : essai sur la pratique des probalités. Ecole nationale supérieure des mines de Paris.

Matheron, G., 1982. Pour une Analyse Krigeante des Données Régionalisées. Rapport N-732, Centre de géostatistiques de Fontainebleau, France.

Matheron, G., 1989. Estimating and Choosing : an Essay on Probability in Practice. Springer-Verlag.

Maynard, D. G., Kalra, Y. P., 1993. Nitrate and exchangeable ammonium nitrogen. In: Carter, M. R. (Ed.), Soil Sampling and Methods of Analysis. Lewis Publishers, pp. 25–38.

McBratney, A. B., de Gruijter, J. J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science 43, 159–175.

McBratney, A. B., Odeh, I. O. A., Bishop, T. F. A., Dunbar, M. S., Shatar, T. M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97 (3-4), 293–327.

Meot, A., Legendre, P., Borcard, D., 1998. Partialling out the spatial component of ecological variation: questions and propositions in the linear modelling framework. Environmental and Ecological Statistics 5 (1), 1–27.

Moran, M. S., Inoue, Y., Barnes, E. M., 1997. Opportunities and limitations for image-based remote sensing in precision crop management. Remote Sensing of Environment 61, 319–346.

Motzkin, G., Wilson, P., Foster, D. R., Allen, A., 1999. Vegetation patterns in heterogeneous landscapes: The importance of history and environment. Journal of Vegetation Science 10 (6), 903–920.

Müller, W. G., Zimmerman, D. L., 1999. Optimal designs for variogram estimation. Environmetrics 10 (1), 23–27.

Myers, D., 1989. To be or not to be stationary: That is the question. Mathematical Geology 21, 347–362.

Nantel, P., Neumann, P., 1992. Ecology of ectomycorrhizal-basidiomycete communities on a local vegetation gradient. Ecology 73 (1), 99–117.

Nichols, W., Killingbeck, K., August, P., 1998. The influence of geomorphological heterogeneity on biodiversity. ii. a landscape perspective. Conservation Biology 12 (2), 371–379.

Noy-Meir, I., Anderson, D., 1971. Multiple pattern analysis or multiscale ordination: towards a vegetation hologram. In: Patil, G. P., Pielou, E. C., Water, E. W. (Eds.), Statistical ecology: populations, ecosystems, and systems analysis. Pennsylvania State University Press, University Park, Pennsylvania, USA, pp. 207–232.

O'Neill, R. V., 1986. A Hierarchical Concept of Ecosystems. Monographs in population biology ; no. 23. Princeton University Press.

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, pp. 3–15.

O'Neill, R. V., Turner, S. J., Cullinan, V. I., Coffin, D. P., Cook, T., Conley, W., Brunt, J., Thomas, J. M., Conley, M. R., Gosz, J., 1991. Multiple landscape scales: An intersite comparison. Landscape Ecology 5 (3), 137–144.

Openshaw, S., 1985. The modifiable areal unit problem. CAT-MOG 38. GeoBooks.

Ortiz, C. J., Deutsch, C. V., 2002. Calculation of uncertainty in the variogram. Mathematical Geology 34 (2), 169–183.

Pardo-Igúzquiza, E., Dowd, P., 2001. Variance-covariance matrix of the experimental variogram: Assessing variogram uncertainty. Mathematical Geology 33 (4), 397–419.

Pardo-Igúzquiza, E., Dowd, P., 2003. Assessment of the uncertainty of spatial covariance parameters of soil properties and its use in applications. Soil Science 168 (11), 769–781.

Pebesma, E. J., Wesseling, C. G., 1998. Gstat: A program for geostatistical modelling, prediction and simulation. Computers & Geosciences 24 (1), 17–31.

Pelletier, B., Dutilleul, P., Larocque, G., Fyles, J. W., 2004. Fitting the linear model of coregionalization by generalized least squares. Mathematical Geology 36 (3), 323–343.

Pelletier, B., Dutilleul, P., Larocque, G., Fyles, J. W., 2009a. Coregionalization analysis with a drift for multi-scale assessment of spatial relationships between ecological variables 1. Estimation of drift and random components. Environmental and Ecological Statistics In Press.

Pelletier, B., Dutilleul, P., Larocque, G., Fyles, J. W., 2009b. Coregionalization analysis with a drift for multi-scale assessment of spatial relationships between ecological variables 2. Estimation of correlations and coefficients of determination. Environmental and Ecological Statistics In Press.

Pelletier, B., Fyles, J., Dutilleul, P., 1999. Tree species control and spatial structure of forest floor properties in a mixed-species stand. Écoscience 6 (1), 79–91.

Peterson, C. J., Pickett, S. T. A., 1995. Forest reorganization: A case study in an old-growth forest catastrophic blowdown. Ecology 76 (3), 763–774.

Pettitt, A. N., McBratney, A. B., 1993. Sampling designs for estimating spatial variance components. Applied Statistics-Journal of the Royal Statistical Society Series C 42 (1), 185–209.

Pickett, S. T., Kolasa, J., Jones, C. G., 1994. Ecological Understanding. Academic Press.

Plant, R. E., 2001. Site-specific management: the application of information technology to crop production. Computer and Electronics in Agriculture 30, 9–29.

Platt, T., 1975. Spectral analysis in ecology. Annual review of ecology and systematics 6, 189–210.

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. Physical Review E 53 (5), 5461–5468.

Priestley, M. B., 1981. Spectral analysis and time series. Academic Press, London; New York.

Quinn, J. F., Dunham, A. E., 1983. On hypothesis testing in ecology and evolution. The American Naturalist 122 (5), 602–617.

Rackham, O., 1991. Mixtures, mosaics and clones: the distribution of trees within european woods and forests. In: Cannell, M. G. R., Malcolm, D. C., Robertson, P. A. (Eds.), The Ecology of Mixed Species (Stands of Trees).

Rebertus, A. J., Wiliamson, G. B., Moser, E. B., 1989. Fire-induced changes in quercus laevis spatial pattern in florida sandhills. Journal of Ecology 77 (3), 638–650.

Reed, R. A., Peet, R. K., Palmer, M. W., White, P. S., 1993. Scale dependence of vegetation-environment correlations: A case study of a north carolina piedmont woodland. Journal of Vegetation Science 4, 329–340.

Reid, K., 2007. How many soil samples do I need? http://www.omafra.gov.on.ca/english/crops/field/news/croptalk/2007/ct-0907a3.htm.

Renshaw, E., 1997. Spectral techniques in spatial analysis. Forest Ecology and Management 94 (1-3), 165–174.

Renshaw, E., Ford, E. D., 1983. The interpretation of process from pattern using two-dimensional spectral analysis: method and problems of interpretation. Applied Statistics 32 (1), 51–63.

Rose, G. A., Leggett, W. C., 1990. The importance of scale to predator-prey spatial correlations: An example of atlantic fishes. Ecology 71 (1), 33–43.

Ruel, J.-C., Loustau, D., Marius, P., 1988. Relations entre la microtopographie, les caractéristiques de la couverture morte et la répartitions des essences dans une érablière à bouleau jaune. Canadian Journal of Forest Research 18, 1196–1202.

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. Ecological Complexity 2 (1), 87.

Schabenberger, O., Gotway, C. A., 2005. Statistical Methods for Spatial Data Analysis. Chapman and Hall, Boca Raton.

Schaefer, J., 1993. Spatial patterns in taiga plant communities following fire. Canadian Journal of Botany 71, 1568–1573.

Schneider, D. C., 2001. The rise of the concept of scale in ecology. Bioscience 51 (7), 545–553.

Searle, S. R., 1971. Linear models. Wiley, New York.

Sen Tran, T., Simard, R. R., 1993. Mehlich III-extractable elements. In: Carter, M. R. (Ed.), Soil sampling and methods of analysis. Lewis Publishers, pp. 43–49.

Sheldrick, B. H., Wang, C., 1993. Particle size distribution. In: Carter, M. R. (Ed.), Soil Sampling and Methods of Analysis. Lewis Publishers, pp. 499–511.

Silverman, B. W., 1985. Some aspects of the Spline smoothing approach to non-parametric regression curve fitting. Journal of the Royal Statistical Society. Series B 47 (1), 1–52.

Soares, A., 2001. Direct sequential simulation and cosimulation. Mathematical Geology 33 (8), 911–926.

Stein, M. L., 1999. Interpolation of Spatial Data: Some Theory for Kriging. Springer-Verlag, New York.

Taylor, P., Haila, Y., 2001. Situatedness and problematic boundaries: Conceptualizing life's complex ecological context. Biology & Philosophy 16 (4), 521–532.

The Mathworks, 2002. Matlab Version 6.5. The Mathworks Inc., Natick, MA.

The Mathworks, 2008. Matlab Version 2008a. The Mathworks Inc., Natick, MA.

Thompson, J. N., Reichman, O. J., Morin, P. J., Polis, G. A., Power, M. E., Sterner, R. W., Couch, C. A., Gough, L., Holt, R., Hooper, D. U., Keesing, F., Lovell, C. R., Milne, B. T., Molles, M. C., Roberts, D. W., Strauss, S. Y., 2001. Frontiers of ecology. Bioscience 51 (1), 15–24.

Tiessen, H., Moir, J. O., 1993. Total and organic carbon. In: Carter., M. R. (Ed.), Soil Sampling and Methods of Analysis. Lewis Publishers.

Tilman, D., 1994. Competition and biodiversity in spatially structured habitats. Ecology 75 (1), 2–16.

Torrence, C., Compo, G. P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79 (1), 61–78.

Upton, G., Cook, I., 2006. A Dictionary of Statistics. Oxford University Press.

van den Wollenberg, A., 1977. Redundancy analysis an alternative for canonical correlation analysis. Psychometrika 42 (2), 207–219.

Van Meirvenne, M., Goovaerts, P., 2002. Accounting for spatial dependence in the processing of multi-temporal sar images using factorial kriging. International Journal of Remote Sensing 23 (2), 371–387.

Vargas-Guzman, J. A., Dimitrakopoulos, R., 2002. Conditional simulation of random fields by successive residuals. Mathematical Geology 34 (5), 597–611.

Vargas-Guzman, J. A., Warrick, A. W., Myers, D. E., 2002. Coregionalization by linear combination of nonorthogonal components. Mathematical Geology 34 (4), 401–415.

Vepsäläinen, K., Spence, J. R., 2000. Generalization in ecology and evolutionary biology: from hypothesis to paradigm. Biology and Philosophy 15, 211–238.

Ver Hoef, J. M., Barry, R. P., 1998. Constructing and fitting models for cokriging and multivariable spatial prediction. Journal of Statistical Planning and Inference 69, 275–294.

Ver Hoef, J. M., Glenn-Lewin, D. C., 1989. Multiscale ordination - a method for detecting pattern at several scales. Vegetatio 82 (1), 59–67.

Wackernagel, H., 2003. Multivariate Geostatistics : an Introduction with Applications, 3rd Edition. Springer.

Wackernagel, H., Petitgas, P., Touffait, Y., 1989. Overview of methods for coregionalization analysis. In: Armstrong, M. (Ed.), Geostatistics, vol. 1. Vol. 1. Kluwer Academic Publishers, pp. 409–420.

Wagner, H. H., 2003. Spatial covariance in plant communities: Integrating ordination, geostatistics, and variance testing. Ecology 84 (4), 1045–1057.

Wagner, H. H., 2004. Direct multi-scale ordination with canonical correspondence analysis. Ecology 85 (2), 342–351.

Warrick, A. W., Myers, D., 1987. Optimization of sampling locations for variogram calculations. Water Resources Research 23 (3), 496–500.

Webster, R., 2000. Is soil variation random? Geoderma 97 (3-4), 149–163.

Webster, R., Atteia, O., Dubois, J. P., 1994. Coregionalization of trace-metals in the soil in the Swiss Jura. European Journal of Soil Science 45 (2), 205–218.

Webster, R., Oliver, M., 1992. Sample adequately to estimate variograms of soil properties. Journal of Soil Science 43, 177–192.

Whelan, B. M., McBratney, A. B., 2000. The "null hypothesis" of precision agriculture management. Precision Agriculture 2, 265–279.

Wiens, J. A., 1989. Spatial scaling in ecology. Functional Ecology 3, 385–397.

Wiens, J. A., 1999. The science and practice of landscape ecology. In: Klopatek, J. M., Gardner, R. H. (Eds.), Landscape Ecological Analysis : Issues and Applications. Springer, pp. xv, 536.

Wiens, J. A., 2001. Understanding the problem of scale in experimental ecology. In: Gardner, R. H. (Ed.), Scaling Relations in Experimental Ecology. Complexity in ecological systems series. Columbia University Press, pp. 61–88.

Wilding, L. P., Bouma, J., Goss, D. W., 1994. Impact of spatial variability on interpretive modeling. In: Bryant, R. B., Arnold, R. W. (Eds.), Quantitative Modeling of Soil Forming Processes : Proceedings of a Symposium Sponsored by Divisions S-5 and S-9 of the Soil Science Society of America in Minneapolis, Minnesota, 2 Nov. 1992. Soil Science Society of America, Inc., Madison, Wis., USA.

Wu, J., Li, H., 2006. Concepts of scale and scaling. In: Wu, J., Jones, K. B., Li, H., Loucks, O. L. (Eds.), Scaling and Uncertainty Analysis in Ecology Methods and Applications. Vol. XVIII. Springer, p. 351.

Xu, T., Moore, I. D., Gallant, J. C., 1993. Fractals, fractal dimensions and landscapes – a review. Geomorphology 8 (4), 245.

Zinke, P. J., 1962. The pattern of influence of individual forest trees on soil properties. Ecology 43 (1), 130–133.