DISSERTATION

SAMPLE DESIGNS FOR MULTI-OBJECTIVE ENVIRONMENTAL SURVEYS

Submitted by

Michael Scott Williams

Department of Forest, Rangeland, and Watershed Stewardship

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2006

COLORADO STATE UNIVERSITY

March 23, 2006

WE HEREBY RECOMMEND THAT THE DISSERTATION SAMPLE DESIGNS FOR MULTI-OBJECTIVE ENVIRONMENTAL SURVEYS PREPARED UNDER OUR SUPERVISION BY MICHAEL SCOTT WILLIAMS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Committee Member (John E. Lundquist)

Committee Member (H. Todd Mowrer)

Committee Member (F. Jay Breidt)

Adviser (Robin M. Reich)

Department Head (N. Thompson Hobbs)

ABSTRACT OF DISSERTATION

SAMPLE DESIGNS FOR MULTI-OBJECTIVE ENVIRONMENTAL SURVEYS

Large-scale forest and range surveys have been collecting ground data for the better part of the last century. From their inception, these surveys were designed for the estimation of population level means and totals. While these survey programs have been effectively meeting this objective, users of these data are often interested in analytical issues other than population level estimation. One of the most widely requested products are maps displaying the spatial extent of the resource in question. The production of these maps requires the prediction of the attributes at unobserved locations. The concern for a number of surveys programs is that data collected for estimation purposes is often poorly suited for spatial prediction. While it is always possible to collect additional data for the purpose of spatial prediction, this solution is often too costly to be practical. The purpose of this study to determine effective methods for simultaneously meeting both estimation and prediction objectives. This is achieved by first determining the most appropriate sample design for the purpose of estimating population level means and totals and then by adapting this design for the purpose of spatial prediction.

Michael Scott Williams
Department of Forest, Rangeland, and Watershed Stewardship
Colorado State University
Fort Collins, Colorado 80523
Summer 2006

ACKNOWLEDGEMENTS

The author would like to thank the Canadian Forest Service, Department of Fisheries and Forestry for the Canadian data used in the first study. The support of my committee members, friends, family and co-workers.

CONTENTS

0.1 Overview

There are a number of comprehensive environmental surveys in the U.S. where ground data are collected at sample of locations within the study area. Examples include the US Environmental Protection Agency's Environmental Monitoring and Assessment Program (Messer et al. 1991), the Forest Inventory and Analysis (FIA) survey (Frayer and Furnival 1999), and the National Resources Inventory (NRI) (Goebel and George 1990, Nusser et al. 1998). These types of surveys are often described as *multi-resource* because they are design to simultaneously address more than one resource issue. While these survey programs are often very successful at providing information for a number of different resources, users of these data are not only interested in multiple resources, they also have a number of different analytical objectives. For example, the data from these surveys are often used to generate descriptive statistics (e.g., means, population totals, and variance estimates) and maps illustrating the spatial extent of the resource. At the same time these data may be used for analytical objectives, such as searching for significant correlations between the presence of a disease and other attributes measured at the same location. These correlations can be used to identify possible causal relationships, though these relationships cannot be used to establish causality (Overton and Stehman 1995, Olsen and Schreuder 1997). For these examples, there are design issues related both the survey and modeling objectives that must be considered. Given the broad array of analytical objectives, it may be more appropriate to refer to these surveys as both multi-resource and multi-objective.

At the heart of any areal survey is the observation of responses $Z(s)$ at different locations, indexed by s, which are spread across the study area. An important issue for these programs, and all areal surveys in general, is that a design for the collection of data that is appropriate and efficient for one objective is often inefficient or even completely inappropriate for another objective. The underlying problem is that some analytical objectives focus on the estimation of the mean response over the area (e.g., estimation of total carbon sequestered in coarse woody debris), while other objectives focus on the prediction of the response at unobserved locations. An example of this second objective is the mapping of basal area or volume over an area. The problem faced by any survey program is that it

is very difficult, if not impossible, to simultaneously meet the data requirements for these two objectives with the same sampling design and protocol.

The underlying problem is rooted in the contrast between estimation of the mean versus prediction of an observation. To illustrate, suppose that you are given the choice of observing two different samples of size 3 that are both drawn from a $N(\mu, \sigma^2)$ distribution. Let the first sample be drawn in such a way that $Z_1, Z_2, Z_3 \sim N(\mu, \sigma^2)$, with $Cor[Z_i, Z_{i'}] = 0$ with the second sample drawn so that $Z'_1, Z'_2, Z'_3 \sim N(\mu, \sigma^2)$, with $Cor[Z'_i, Z'_{i'}] = 1$. Now if you are given the choice between the two different samples, and the objective is to estimate μ, then the mean of the first sample will provide the most precise estimate because it is the best linear unbiased estimator of μ. On the other hand, if one of the observations is not observed and the objective is to predict the missing observation, then the perfect correlation between Z'_1, Z'_2, Z'_3 ensures that the missing observation is known exactly by observing either of the remaining values. From this example, it can be concluded that any sample design for a multi-objective survey is likely to be, at best, a compromise.

The concern with many of the current environmental surveys is that they are often designed for meeting estimation objectives. Thus, it would seem that the data collected from such a survey would be of limited value for addressing objectives related to prediction, which in this study will be spatial prediction. However, there are a number of subtle practical issues involved in the design of an environmental survey. Thus, the underlying goal of this study is two-fold, with the first goal being to carefully examine all of the factors that influence the design of an environmental survey when estimation is the objective. The second goal is to exploit some of the practical limitation of these surveys when prediction is the objective. This study has been broken into three related studies. A summary of the studies is as follows;

- Study I. Using established guidelines for environmental surveys, determine the appropriate sampling and response design for estimating the mean or total of an attribute. In addition, the factors that determine the selection of the appropriate design are identified and ranked according to their level of influence.

- Study II. Determine how the constraints imposed by the sample survey objectives impact spatial modeling objectives and introduce a new sampling method that allows for very precise estimation of the variogram for spatial modeling within the imposed constraints.

- Study III. Develop and test new kriging procedures based on the methodology outlined in the previous study.

Each of three studies builds on the results of the previous study and is presented in a separate chapter. Therefore, the results of each study builds on results of the previous study in a more or less linear fashion. However, each chapter is structured so that it also serves a stand-alone document. While this facilitates publication of the individual studies, the reason for using this approach is that it allows for an extensive literature review to document the approach taken for each study. The motivation for providing individual literature reviews is that results from numerous fields are incorporated and the relevance of some results is not clear without the results of the previous studies.

Chapter 1

CHOICE OF GROUND PLOTS CONFIGURATIONS FOR CONTINUOUS POPULATION SURVEYS OVER A FRAGMENTED LANDSCAPE

Abstract

Cluster plots are ubiquitous in many forest and environmental surveys. For small-scale surveys, such as for a single forest stand, a cluster sample based on visiting a single point is often used. However, for surveys over large areas, such as a state or eco-region, a cluster plot consisting of anywhere from 3 and 10 points or plots is common. The sample design and estimation for a large-scale survey can be approached by either using a fixed area frame that tessellates the landbase into a finite collection of area sample units, or by viewing the population as an infinite set of points, which will be referred to as continuous population sampling. Plot configurations consisting of multiple points or plots are know to be efficient under the tessellated population approach. However, the factors influencing the choice of an appropriate plot configuration in the continuous population setting have not been specified. This study identifies the boundary correction technique, within-cluster variability, and the average area of the plot that covers the population as the three factors that should be considered when choosing a ground plot configuration. Each of these factors is influenced by the level of fragmentation in the target population. Tests of these factors on two forest populations suggests a cluster sample consisting of a single point or plot, coupled with the walkthrough boundary correction technique, will often be the most efficient and practical solution.

Key Words: Infinite population survey, boundary correction technique, large-scale survey, point sample

1.1 Introduction

Many large-scale environmental surveys in the U.S. rely on a combination of ground and remotely sensed data. Examples include the US Environmental Protection Agency's Environmental Monitoring and Assessment Program (Messer et al. 1991), the Forest Inventory and Analysis (FIA) (Frayer and Furnival 1999), and the National Resources Inventory (NRI) (Goebel and George 1990, Nusser et al. 1998). These surveys are designed to accomodate multiple objectives. For example, the data are often used to generate descriptive statistics such as means, population totals, and variance estimates. Another objective of these surveys is the production of maps illustrating the spatial extent of the resource. Thus, there is both a traditional sample survey and spatial modeling component that should be considered during the design phase of an inventory. However, this study will only focus on the estimation of descriptive statistics (e.g., total carbon sequestered, number of standing dead trees).

In the past, large-scale forest surveys relied on sampling designs that utilized a widely-spaced systematic grid of ground plot locations. At each grid location a ground plot was installed, and estimation followed using the results of Cochran (1977), by assuming the sample was a single or multi-stage cluster sample drawn from a finite collection of areal sample units. This approach to forest sampling has been used extensively in research publications and virtually every forest inventory and survey sampling text. Examples include Bickford (1952), Bickford et al. (1963), de Vries (1986), Thompson (1992, p. 23), and Shiver and Borders (1996). The distinguishing feature of these inventories is the need to establish an area frame (Sarndal et al. 1992, p. 12). A simple example for an area of 1 square kilometer is given in Figure 1, where an area of forest and non-forested land is divided into four primary sampling units (PSUs) with nine secondary sampling units in each PSU. The area frame provides a mechanism for drawing samples from the population and is essential for the construction of the reference or randomization distribution (Gregoire 1998). While this is the approach used by NRI, many forest surveys have never actually established an area frame. Instead they assume that an area frame exists and the ground plot "samples a hectare or acre" (Bickford (1952), Spada 1960, Bickford et al. (1964),

Arvanitis and O'Regan 1972). As Gregoire (1998) points out: "If the region A truly were tessellated by N plots, then this approach might be defensible. Its appropriateness to current practice is questionable, as A rarely comprises a set of N nonoverlapping plots of land." This will be referred to as the *tessellated population sampling* approach.

Concerns over the difficulties associated with the construction and maintenance of an area frame is one of the factors that has led to a change from viewing the ground sample as an application of finite population cluster sampling to viewing each sample element as a point drawn from an infinite set of possible locations. Under this approach, no frame with a finite number of units is assumed to exist. Instead, points within the boundary of the population are visited and estimates are derived from the randomly located points. The term **continuous population** sampling will be used to distinguish this approach from the tessellated population sampling approach described above. To formalize the design and estimator, define the continuous population A, of area $|A|$,which forms the set of points over which the resource of interest exists. A sample point $s \in A \subset U$ denotes the point s, which is in the subset A of the universe U. At each point s within A, let $z(s)$ denote a real-valued function defined on A. This function describes an attribute of the resource and it is a fixed prior to sampling when inference is with respect to the design. The population parameter to be estimated is

$$\tau_z = \int_A z(s)ds, \qquad [1]$$

where for this study $z(s)$ is a measure of abundance per unit area at the point s.

For a forest inventory, the set U could consist of all points in a region (e.g., country, state, or land owned by a corporation) and the set A might contain only those points that cover forested land. A simple example is given in Figure 2, where roughly 42% of the area in the graph is defined as forested. Comparing Figures 1 and 2 shows that one potential advantage of the continuous population approach is that it may be possible to carry out the inventory without ever visiting points that fall outside of the forested area (A), while 3 out of the 4 PSU's may need to be visited in order to estimate the attributes of the forested area under the tessellated population approach.

The continuous population approach covers a very rich collection of literature. Examples include variable-radius and fixed-area plot sampling as described by Roesch et al.

(1993) and Eriksson (1995), quadrat sampling (Ripley 1981 Chapter 6, Diggle 1983, Upton and Fingelton 1985, and Stoyan et al. 1995), spatial sampling (Ripley 1981 Chapter 3), line-intersect sampling (Kaiser 1983), sampling of continuous populations (Bartlett 1986, Cordy 1993), crude monte carlo and importance sampling (Valentive et al. 2001), and sampling based on random tessellations (Steven 1997, Stevens and Olsen 2003, 2004).

While many environmental surveys now approach the sampling and estimation problem from the continuous population perspective, many aspects of large-scale inventory programs still rely on results that were derived under the previously used tessellated population approach. One aspect is the configuration of the plot that is used to collect the ground data, specifically the use of open-cluster plots. These are plot configurations that comprise a systematically arranged group of smaller plots or points, which will be referred to as *sub-plots*. De Vries (1986 p. 167), Frayer and Furnival (1999), and Labau et al. (2005) give examples of the open-cluster plots used in various national inventory programs in the US and Europe. The motivation for open-cluster plots comes from the finite population sample survey literature on two-stage cluster sampling, where it is felt that spreading the cluster sample over a larger area sufficiently increases the within-cluster variability so that the overall variance of the estimator will be reduced for a given sampling effort (Spada 1960, Cochran 1977, Schreuder et al. 1993, Scott 1993). A number of studies, based on the early agricultural field trial work of Smith (1938), have concluded that open-cluster plots are significantly more efficient that single plots (e.g., Spada 1960, Freese 1961, Scott 1993). Thus, open-cluster plots are used in many large-scale environmental inventories. Some of the factors involved in choosing an open-cluster plot are the size and shape of each sub-plot, the number of sub-plots, and the separation distance between the sub-plots. However, what has not been investigated is whether the reduction in variance associated with the use of open-cluster plots holds under the continuous population sampling approach.

Given that there is currently no justification for the use of open-cluster plots, the goals of this study are to:

- Describe the necessary conditions for achieving a design-unbiased estimator in a continuous population inventory, particularly when open-cluster plots are used.

- Identify and examine the factors that affect the selection of plot configurations for a large-scale environmental inventory when the goal of the inventory is to estimate z_T.
- Test three different sampling methods across the range of these factors to determine the relative performance of the sampling methods.

1.2 Plot Configurations Considered

Given the multitude of different areal sampling techniques, an exhaustive test of plot configurations is not possible. However, the majority of current large-scale forest inventories use some type of circular fixed-area plot. Two different fixed-area plot configurations will be considered in this study. The performance of the single fixed-area plot will be compared to an open cluster of circular fixed-area sub-plots. The open clusters comprise a center sub-plot surrounded by an equally spaced constellation of sub-plots (Figure 4). Each of these sub-plots will have a fixed radius of $r = 5$ meters. The number of sub-plots will be varied from $o = 2$ to 7 and the separation distance between sub-plots will vary from $d = 10 - 40$ meters in 5 meter increments (Figure 4). The single plot used in this study is a circular fixed-area plot with radius r. The radius of this plot will vary from $r = 2 - 20$ meters in one meter increments (Figure 5).

1.3 Review of Sampling Strategies for Unbiased Estimation

Within A there can be a large number of attributes of interest. For some attributes the response, $z(s)$, can be observed at a point (e.g., soil temperature, humidity), while other attributes must be measured with a plot located about the point (e.g., $z(s)$=the number of trees per hectare). The attribute of interest can then be deduced by observing $z(s)$ at a sample of points $s_j, j = 1, 2...m$. For example, if A is a watershed and $z(s)$ denotes the depth of snow at the point s, then by definition eq. [1] gives the total volume of snow in the watershed. However, many ecological inventories focus on estimating attributes associated with trees and other organisms. Thus, A is the area over which N organisms are located. Associated with each organism are one or more attributes z_i, (e.g., tree volume, species, presence of disease, number of nesting cavities etc.). When producing descriptive statistics,

one objective is to estimate the total abundance of a resource ($\tau_z = N$) or the aggregate value of an attribute $\tau_z = \sum_{i=1}^{N} z_i$. Estimation of total tree basal area will be the focus of this study.

The concern with continuous population surveys is that it is often assumed that

$$\int_A z(s)ds$$

is equal to the true parameter of interest. While this is true for the depth of snow example given above, it is generally incorrect when a plot is placed at s and information about the vegetation is aggregated to construct $z(s)$. Stevens and Urquhart (2000) refer to sampling strategies that meet this requirement as being *aggregation unbiased*. While a number of estimators exist that are model-unbiased (Williams 2001a and references therein), the goal of this study is to ensure that

$$\int_A z(s)ds = \sum_{i=1}^{N} z_i,$$

(i.e., the estimators are design-unbiased) because design-based inference is often preferred (Overton and Stehman 1993, Olsen and Schreuder 1997). However, this is not possible without careful consideration of the construction of $z(s)$. One potential source of bias is a consequence of sampling along the boundaries of A. However, first a general discussion of the methods for constructing $z(s)$ are necessary.

The object-centered concept is often used to develop plot-based sampling methods (see Husch et al. 1982, pp. 224-225). The fundamental idea of the object-centered approach is that visiting a sample point with coordinates s and selecting objects that fall within the boundary of the plot is equivalent to an imaginary plot being drawn about each object and each element is selected into the sample any time a sample point s falls within the object-centered plot. This concept is illustrated in Figure 6 and this approach will be used to develop the estimators in this study. The idea of an object-center plot has been extended to many other areal sampling applications, such as line-intersect sampling (Kaiser 1983) where the "plots" have no consistent shape or size. The term *inclusion zone* is used to describe the object-centered "plot" about each population element and the element is selected into the sample whenever a sample point falls within the inclusion zone.

Many different approaches have been taken to derive the design-based properties of plot- and point-based sampling methods (Mandallaz (1991), Cordy (1993), Eriksson (1995), Stevens and Urquart (2000) Williams (2001b), Valentine et al. (2001), and Stevens and Olsen (2004)). However, the simplest treats the estimator, $\hat{Z}$, as a function of the random variable that determines the location of a sample point. For $A \subset \Re^2$, this response surface is essentially a map of every possible sample outcome over A. The 3-dimensional response surface, assuming a circular plot with area $a = \pi r^2$, can be constructed across the population using the following algorithm:

- For every population element, indexed by i, center the circular plot about the element.
- Use the circular plot boundary to construct a 3-dimensional solid of height $z_i|A|/a_i^T$, where a_i^T is the area of the inclusion zone that falls within A (i.e., $a^T < a$ if the object-centered plot extends beyond the boundary).
- Define the height of the surface at any coordinate, s, as the sum of the heights of the overlapping solids defined for each element that would be selected, i.e. $\hat{Z}_{TI}(s) = \sum_{i=1}^{n} z_i|A|/a_i^T$.

The subscript TI indicates that the inclusion zone is truncated by the boundary of A. A proof of design-unbiasedness is given in the Appendix.

1.4 Factors Affecting Efficiency

Three factors will be identified that play a role in the selection of a plot configuration in the continuous population setting. However, all of these factors are directly influenced to the degree of fragmentation of A within U. Thus, an understanding of the degree of fragmentation of forest lands will be beneficial in understanding the influence of each factor.

In the context of forest surveys, the population of interest is the collection of forest patches or fragments within the larger mosaic of forested and nonforested lands. Forest patches have no consistent shape or size, so developing meaningful metrics to describe the level of fragmentation is difficult. Some of the simplest metrics are average forest stand size, interior-to-edge ratio, and distance to the edge. Forest fragmentation has been studied

mostly in the context of how fragmentation affects ecological processes. However, there have been a number of studies that determine general levels of fragmentation across large areas. For example, Riitters et al. (2001) studied forest fragmentation for the conterminous United States and found that fragmentation was extensive in virtually all areas where forests occur, with 43.5 % of all forested land falling within 90 m of a forest edge. However, this estimate excludes fragmentation of the forest due to bodies of water, roads, as well as bare rock, and sandy areas that do not support forest cover. Butler et al. (2004) studied forest fragmentation at a smaller scale and found that 13% of all forest lands fall within 30m of a forest edge. Schmidt and Raile (1999) report that 40 % of all forested stands in the Great Lakes region are less than 2.5 ha (10 acres) in size. Thus, fragmentation is a pervasive feature of the landscape.

Given the level of fragmentation, the following three related factors are identified as affecting the efficiency of a continuous population survey.

1.4.1 Method of boundary correction when sampling along the boundary

One factor that influences the efficiency an of estimator is the method of data collection along the boundary between A and U (Schreuder et al. 1993, Steven and Urquhart 2000). The term *edge-effect* has at least two distinctly different meanings. In an ecological context, this term describes differences in the spatial and temporal patterns along the boundary of a population when compared to those of the interior. In a sampling context, the terms edge-effect, *boundary-overlap*, and *slop-over* have been used to describe the problem of design-bias in a continuous population estimator. The problem arises whenever population elements (trees, forbs, coarse woody debris, etc.) within the population are close to the boundary of the population. The potential bias occurs because these elements have a smaller probability of being included in the sample when compared to similar elements that are well within the interior of A (Figure 6). The term boundary-overlap will be used in this study. The reason for this reduced inclusion probability is that a portion of the inclusion zone for these elements falls outside of A, where no samples point are located. This is why the proof of unbiasedness in the Appendix assumes that the area of the inclusion zone, a_i^T, that falls within the boundary of A is known, though this is usually not the case.

The concept of a truncated inclusion zone can be extended to open-cluster plots, but the orientation between the plot and inclusion zone must be reversed (Williams and Eriksson 2002, Valentine et al . 2005).

Three additional approaches have been taken to the boundary-overlap problem. To motivate the different approaches, note the a continuous population estimator, $\hat{Z}(s)$, is comprised of four components; These being plot size, plot shape, total area of the population A, and the attribute z_i. Thus, each of the solutions modifies one of the four components to reduce or eliminate the design-bias.

The first approach is to calculate the true inclusion probabilities of all elements near the boundary by mapping the boundaries in the field and calculating a^T for each element in the sample. This could be accomplished by surveying a number of points along the boundary and fitting a function, such as a spline, to the points. Then the size, shape, and location of each inclusion zone would need to be determined in relation to the boundary. This would be very labor intensive, especially when open-cluster plots are used because the length of the boundary that would need to be mapped would be very long and possibly irregularly shaped, and the location of every tree with respect to the boundary must also be known. Thus, this field-based approach is not practical. However, it may be feasible to use a combination of high-resolution aerial photography and the ground data if the following conditions are meet:

- The center point of the plot can be identified on the aerial photo.
- The boundaries of A can be determined (i.e., drawn) on the photo and the scale of the photo in known.
- The locations of every sample element can be determined in relation to the center-point of each sub-plot.
- The area of the inclusion zone can be determined for the sampling design. Thus, only circular or rectangular fixed-area plots and variable radius plot sampling designs can be reasonably accommodated.

This approach will be referred to as the truncated inclusion (TI) approach. The design-unbiased estimator for this approach is

$$\hat{Z}_{TI}(s) = \sum_{i=1}^{n} z_i |A|/a_i^T.$$

Another approach is to allow the shape of the plot to change along the boundary of the population while maintaining equal plot areas. An example, proposed by Stevens and Urquart (2000), is an "ears" plot configuration. This approach takes the portion of a circular plot that falls outside the boundary, divides the area in half, and "attaches" the ears to the side of the plot. Some concerns are that this approach would not be feasible for some of the more efficient methods of sampling other attributes because the shape and area of the inclusion zone can not realistically be determined. An example is perpendicular distance sampling (Williams and Gove 2003), where the shape and area of the inclusion zone differs for every piece of coarse woody debris. Stevens and Urquart (2000) also note that this approach would be difficult to implement along irregular boundaries and do not provide solutions for open-cluster plots. Another concern with the "ears" plot approach is that it assumes that A is flat, whereas, forest survey plots are slope corrected so that it is the vertical projection of the plots that have equal area, rather than the plots themselves. Thus, both the shape and surface area of plots can differ substantially in steep or uneven terrain. For these reasons this boundary correction technique was not considered in this study.

The third method changes the population of interest A and allows sample points to fall outside the population boundary, but within some region of known area that completely encompasses all inclusion zones that fall outside the boundary of A. Thus if a fixed-area plot of radius r is used, a buffer strip of width r is added along the boundary and sampling is carried out over this new population A_{buff}. An example is given in Figure 7. An open-cluster plot can also be accommodated by increasing the width of the buffer strip to $r + d$, where d is the separation distance between sub-plots in the open-cluster. The design-unbiased estimator of the population total is given by

$$\hat{Z}_{BUFF}(s) = \sum_{i=1}^{n} z_i |A_{buff}|/a_i.$$

This method, which will be referred to as the *buffering* technique, was first described by Masuyama (1954) for agricultural applications, and discussed briefly in a forestry context by Schmid-Haas (1982), Mandallaz (1991), and Ducey et al. (2004). The primary advantage of this method is that it can be used to achieve design-unbiased estimators for any shaped boundary, sampling technique, and plot configuration. However, the practicality of this method is questionable because it requires that a maximum inclusion zone width must be specified for every sample element and sample design, which is nearly impossible for relascope based sampling methods (Bitterlich and Bitterlich 1996). Some loss in precision is also expected under this technique because $z(s)$ will decrease to zero as s approach the boundary of A_{buff}. This will increase the heterogeneity of $z(s)$. Other concerns are the increase in the area that must be sampled and, as pointed out by Ducey et al. (2004), this method requires that points are located outside of A, which is often impractical because a portion of the boundary of A are likely to be either impervious (e.g., rock wall), hazardous (e.g., river, swamp, or cliff), or access can be denied (e.g., private landowners refuse access, military installations).

A number of useful methods have appeared in the literature that replicate the attribute value, z_i, multiple times rather than adjusting the area of the plot (see, Schreuder et al. 1993, pp. 297-301). Of these, the most widely discussed is the mirage method introduced by Schmid-Haas (1969) and later extended to line-intersect sampling by Gregoire and Monkich (1994). However, the inability of this method to deal with irregular boundaries and pocket inclusions, such as roads or landings, presents a severe limitation to its use in a natural resource setting (Ducey et al. 2004).

In an effort to deal with the shortcomings of the mirage method, Ducey et al. (2004) developed an alternative method, called the *walkthrough* technique. This method is based on a modification of the boundary reflection method proposed by Gove et al. (1999). It is easily implemented in the field, the field crew is never required to work outside of A, and it can handle virtually any shaped inclusion zone. The estimator for this technique is given by

$$\hat{Z}_{WT}(s) = \sum_{i=1}^{n} z_i \delta(s)_i |A|/a_i,$$

where $\delta(s) = 1, 2$ depending on the location of each population element with respect to the sample location s. This is illustrated in Figure 8. One concern with this technique is that $\hat{Z}_{WT}(s)$ is not design-unbiased because a small number of situations exist where $\delta(s)$ should be greater than 2. However, these situations are thought to be rare. A concern with this method is that it is only appropriate for when the sample unit is a single plot or point. A modification of this technique for open-cluster plots, referred to as *walkabout* has been developed by Valentine et al. (2005). However, they note that this method is more of an intellectual curiosity than a practical solution.

To summarize; there are many of possible boundary correction techniques for both single- and open-cluster plots. However, with the exception of the walkthrough technique for single cluster plots, the practicality of all of these techniques is somewhat suspect. For this study, it is assumed that the truncated inclusion probabilities (TI), buffering (BUFF), and walkthrough (WT) techniques are feasible for single cluster plots, and that the (TI) and (BUFF) techniques are feasible for open-cluster plots.

1.4.2 Trade-off of within- and between-cluster variability

The premise of cluster sampling is that an efficient design can be obtained by partitioning the population in such a way that each cluster is as similar as possible to the other clusters in the population. For the tessellated population approach, this is equivalent to making the within-cluster variability as large as possible so the between-cluster variability is minimized. In most natural resource populations, this condition is often satisfied by using a systematic sample to select primary sampling units, in which the secondary sampling units are widely spaced as well (Thompson 2002, pp. 134-5). In this setting, a sample design based on widely spaced open-cluster plots will generally have a smaller variance than if the same area were to be sampled by a single plot of equal area because the open-cluster plot will increase the within-cluster variability.

This result is common in the finite population sampling literature (e.g., Cochran 1977, Chap. 10, Thompson 2002, Chap. 12) and is appropriate in the context of a tessellated population sampling design. This same general result has also been verified through the use of point process models where a systematic, rather than a clustered sampling design,

has been shown to be nearly optimal for estimating population means and totals across a wide range of spatial patterns (Bellhouse 1977, Iachan 1985, Stein 1995, 1999).

However, when open-cluster plots are used in the continuous population setting, the notion of within- and between cluster variance has little appeal because there are no actual primary or secondary sampling units. One way to view the problem in the continuous population setting is to assume that if m sample points are selected and an open-cluster plot with o sub-plots is used, then o separate estimates can be generated by viewing the m points as an open-cluster plot. In this setting, the first estimate is derived from the m center points and the remaining estimates derived by systematically moving the locations of the center points by a distance d so that the new points coincide with center of each of the remaining sub-polts in the open-cluster. Thus, one estimate would be derived from the center points, and $o-1$ additional points from the remaining sub-plots. If this estimator is denoted by,

$$\hat{Z}' = \frac{1}{o}\sum_{j=1}^{o}\hat{Z}_j,$$

where $\hat{Z}_j$ is the response from each of the o points in the subplot, then the variance is

$$Var[\hat{Z}'] = \frac{1}{o^2}\sum_{j=1}^{o}Var[\hat{Z}_j] + \frac{1}{o^2}\sum_{j\neq j'}^{o}\rho(Var[\hat{Z}_j]Var[\hat{Z}_{j'}])^{1/2} = \frac{1}{o}Var[\hat{Z}_j] + \frac{o-1}{o}\rho Var[\hat{Z}_j],$$

where ρ is the correlation between each of the o estimators. While it is generally impossible to know the true correlation between the individual estimators (Ripley 1981, Chapter 3), it is not unreasonable to assume that estimators derived by moving the sample locations by only d meters in any direction will be very highly correlated. To illustrate, if $o = 4$ and $\rho = 0.9$, the total reduction in the variance is only 7.5 %. Even for the most heavily fragmented populations, where it might be reasonable to assume that $\rho = 0.5$, the reduction in variance would be only 37.5 %.

For the plot configurations considered here, the most effective way to increase the within-cluster variability is to increase the separation distance between sub-plots in an open cluster plot. A less effective method of increasing within-cluster variability is to increase the radius of a single fixed-area plot.

1.4.3 Average area sampled.

The final factor identified is to note that the average area of A sampled by an open-cluster plot in the continuous population setting will almost always be smaller than the area sampled by a single plot of equal area. This occurs because it is more likely that parts of an open-cluster plot will fall outside A (Figure 9). This is especially true in highly fragmented and linear forest populations where the average stand size is small.

Some insight into the loss of efficiency can be gained by viewing the population as a realization of stationary point process spread across A with intensity λ (Ripley 1981, p. 103-104) and assuming that the goal of the study is to estimate the intensity of the point process. In this setting, two attributes are measured at every point in the population. The first is the sum of the number of points (e.g., trees) that fall within the boundary of the plot. The second is the portion of the plot that falls within the boundary of A.

The construction of the estimator of intensity proceeds as follows:

- For $s \in A$, determine $n(s)$, which is the number of points within a distance r of s.
- For the same $s \in A$, determine $a(s)$, which is the area of the plot falling within the boundary of A.
- The estimator of the intensity is $\hat{\lambda}(s) = n(s)/a(s)$, where $\hat{\lambda}(s)$ estimates the number of point per unit area.

In this setting the variance of the plot count is related to the area of the plot covering the population. If the forest is viewed as a realization of a stationary point process, or more generally if the point process describing the forest has a constant intensity function (Stoyan et al. 1995), the estimator can be rearranged as

$$n(s) = \lambda a(s) + \epsilon, \qquad [5]$$

where $E[\epsilon] = 0$ and $Var[\epsilon] = \sigma^2 g(a(s))$ (Ripley 1981, p. 13, Taylor and Karlin 1994, Sec. 5.2 and 5.5). In this setting the variance of the plot count, given by $\sigma^2 g(a(s))$, is a function of the area of the plot covering the population. Thus, in the continuous population

setting, a cluster plot will only be more efficient if the conditions within A are sufficiently heterogeneous to offset the effect of the reduction in area sampled.

The area of forest sampled by the plot configurations considered here can be increased by increasing the radius of each plot, by adding more sub-plots, or both. Increasing the separation distance between plots will decrease the average area sampled for any realistic fragmented population.

1.5 Data Description

Two different data sets were used in testing. The first covers a small forest stand (5.37 ha = 13.27 acres), while the second data set was constructed by combining the data from three smaller data sets and covers (67.42 ha = 166.5 acres). These data sets will be used to model both fragmented and unfragmented landscape types. Both data sets are comprised of multiple compartments with distinctly different forest conditions in each compartment. This was done in an attempt to maximize the within-cluster heterogeneity of the open-cluster plots because they are more likely to cover more than one stand condition. Whether either data set is actually representative of any particular large-scale landscape pattern is unknown. For example, one would expect that large patches of forest would be more likely to be undisturbed, so their interiors would be more homogeneous than the large data set used in this study. The increase in the area that would have to be sampled under the BUFF boundary-correction technique is also smaller than would be expected for natural stands because these stands tend to be irregularly shaped, whereas these populations are rectangular. For example, Ducey et al. (2004) calculate a 30 % increase in the area sampled for an irregularly shaped 46.5 ha. stand with a 25 m. wide buffer. Thus the differences in the performance of the boundary-correction techniques is probably under represented. Both of these factors will tend to favor the open-cluster plot designs.

1.5.1 The New Jersey Data Set

The data set, referred to as NJ, was collected in central New Jersey in 1985 (Charles Scott, personal communication) on an area of 5.37 ha, which was extracted from the center of two larger forest stands. Two distinctly different forest types are represented, these

being adjacent hardwood stands of roughly equal area containing a total of 4,744 trees. The data were collected using square 15.2 m x 15.2 m (50 x 50 ft) plots. These were laid out and all trees with diameter at breast height greater than 7.5 cm (3 in) were measured and their x and y coordinates were recorded. The data set contains two stands of different structure, one comprised of mature hardwoods of saw-log size and the other of hardwood regeneration. The two stands differed in structure and are separated by a dirt road that divides the population roughly in half and has partially grown over. Conditions within each stand are fairly heterogeneous with areas of visibly different intensity within and between the stands. There is no apparent edge effect along the boundary of the population and none is expected because the data were collected from the interior of a larger stand. Figure 10 depicts the structure of the data set. The diameter of each circle represents the relative size of each tree. An open-cluster plot with a 20m separation distance is also overlayed.

The difference in the area of the population ($|A|$) in relation to the area of the population plus the buffer strip ($|A_{buff}|$) is also of interest. For this data set, the increase in the area that must be sampled ranged from 1.7% to 18% for the single cluster plot ($r = 2, 20$, respectively). The increase in area for the open-cluster plot design ranged from 13.4% to 42.6%, with $d = 10, 40$ respectively.

1.5.2 The Canada Data Set

The large data set, referred to as Canada, was constructed using data from three different forest stands near Sault Sainte Marie, Ontario, Canada and covers a total area of 67.42 hectares (166.5 acres). The stands used were a hardwood, softwood and pine plantation. To create the large data set, two copies of each stand were arranged in a 2×3 matrix so that the heterogeneity of the population would be maximized (Figure 11). A detailed description of the characteristics of each stand can be found in Ek (1969).

The difference in the area of the population ($|A|$) in relation to the area of the population plus the buffer strip ($|A_{buff}|$) for this data set ranged from .6% to 6.2% for the single cluster plot ($r = 2, 20$, respectively) and from 4.7% to 14.2% for the open-cluster plot ($d = 10, 40$, respectively).

1.6 Simulation Studies

Generating artificial forest stands that accurately mimic the size or shape of real forest stands would be nearly impossible. However, the sampling problem considered here does not require a realistically shaped stands because interest is in the properties of the reference distribution (Gregoire 1998) of the estimators (i.e., this sampling application is inherently non-spatial). Thus, all the simulation study has to do is produce estimates from each sample whose reference distribution realistically mimics one that would be derived from a real population.

Rather than use a Monte Carlo simulation, the bias and variance results were determined by calculating the appropriate expected values of each estimator using the approach described in Williams (2001b). The bias and variance for a sample of size $m = 1$ is derived by approximating the integrals in eqs. A.2, A.3 [in the Appendix] with a Riemann sum using a very fine grid spacing over A.

Two different metrics were used to compare the variance of the various plot configurations and boundary-correction techniques. The first was the standard error, given by $SE = \sqrt{(Var[\hat{Z}_*])}$, where $* = BUFF, TI, WT$ denotes the boundary-correction technique. The design effect was also used to facilitate the comparison. This quantity is given by

$$deff(alternative, baseline) = \frac{V[\hat{Z}_{alternative}]}{V[\hat{Z}_{baseline}]},$$

where $\hat{Z}_{baseline}$, is the estimator associated with using either the WT or TI boundary-correction techniques and $\hat{Z}_{alternative}$ will be the estimator based on either the WT or BUFF boundary-correction techniques. A $deff$ greater than one indicates that precision is lost when the alternative boundary-correction technique is used. Conversely, a $deff$ less than one indicates that the baseline boundary-correction technique is more precise. An advantage of using $deff$ to compare estimators is that the difference in the number of points required to achieve equal variance between the two estimators is equal to the $deff$. For example, if the $deff(BUFF, TI) = 2$, then it would take twice as many sample points using the BUFF boundary-correction technique to achieve equal variance as would be required if the TI technique had been used.

The bias of the various techniques is of lesser interest than the variance because, with the exception of the walkthrough technique, the estimators are design-unbiased. However, many environmental surveys completely ignore the importance of boundary-correction, so the magnitude of the bias when boundary-overlap is ignored is reported for a small subset of the simulations. The bias of the estimators is expressed as a the ratio of the expected value of the estimator to the true value, i.e.,

$$B = 100\frac{\hat{Z}_*}{\tau_z},$$

where $*$ depicts the estimator, with $B = 100$ representing unbiasedness.

1.7 Simulation Results and Interpretation

The performance of the various plot configurations and estimators is presented separately for each data set. Dealing with the multitude of plot sizes, configurations, and boundary-corrections techniques is difficult, so results are first presented for the single plots, then for the open-cluster plots, and finally for a subset of all possible configurations.

1.7.1 The NJ Population

Figure 12 summarizes the standard error results for single cluster plot sampling as a function of plot radius and boundary correction technique. When the SE is graphed as a function of plot size and boundary-correction technique, it can be seen that the standard error drops dramatically as r increases from 2 to 5 m for all three techniques. After that point, the rate of reduction in the SE decreases more slowly for both the TI and WT boundary correction techniques. In contrast, the SE for BUFF essentially reaches its minimum by $r = 7$ and then begins to increase for all plot radii greater than $r = 10$. The cause of the increase in the overall variability is that the heterogeneity associated with the sample points falling in the buffer increases more quickly with increasing plot width than the offsetting reduction in the variability associated with sampling a larger area.

Figure 13 summarizes the design effect ($deff$) for single cluster plot sampling as a function of plot radius and boundary correction technique. The $deff$ for all three techniques is essentially equivalent for very small plot sizes ($r = 2, 3\ m$). As the plot radius increases, the

advantage of both the TI and WT boundary-correction techniques over buffering increases rapidly, with a maximum difference in relative efficiency exceeding 1.5 for $r = 6$ and a maximum difference in excess of $deff = 5$ for the largest plot size considered. The graph also shows that the WT technique is slightly less efficient than using the true inclusion probabilities. For example, the $deff = 1.11$ and 1.30 for plot radii of $r = 10$ and 20, respectively. However, this loss in efficiency when using WT is small in comparison to the loss associated with the BUFF boundary-correction technique.

Figure 14 shows the SE for the open-cluster plot designs as a function of the number of sub-plots, the separation distance between plots, and the method of boundary-correction. The results show there were no situations in which $\hat{Z}_{BUFF}$ was as precise as $\hat{Z}_{TI}$. Thus, any increase in the within-cluster variability associated with adding plots or increasing the separation distance was more than offset by the loss in precision due to the boundary-correction technique and average area sampled. An interesting result is that increasing the separation distance for the TI method tended to slightly increase the standard error, so the increase in within-cluster variability is more than offset by the reduction in the average area of the plot that covers the forest.

The design effect results (i.e., $deff(BUFF, TI)$) express the differences in efficiency in a more practical perspective. These are summarized in Figure 15 and show that using the BUFF boundary-correction technique would require increasing the number of sample points anywhere from 3 to 8 times the number needed for the TI technique.

The only way to compare all possible plot designs is by examining the standard errors across all possible sampling strategies. However, because it has already been established that the single plot buffering estimator is substantially less precise than both the TI and WT estimator, only the results for TI estimator are considered. These results are given in Figure 16, where the standard error for the single cluster plots with $\hat{Z}_{TI}$ and plot radii $r = 5, 10, 15, 20$ are superimposed on Figure 14. For this population, $\hat{Z}_{TI}$ with a single $r = 5$ plot was substantially less variable than all of the open-cluster plot designs that used the BUFF boundary-correction technique. Doubling the plot radius of the single cluster plot produced another substantial increase in precision and a single cluster with radius

$r = 15$ was more precise than any of the open-cluster plot designs. Further increasing the plot radius to $r = 20$ resulted in a negligible increase in precision.

Figure 17 depicts the design-bias for the single fixed-area plot with the three boundary-correction techniques, as well as the bias that would result from ignoring boundary correction. The slight design-bias of $\hat{Z}_{WT}$ is essentially imperceptible. In contrast, ignoring the boundary-overlap problem produced biased that increased linearly with the plot radius and exceeded 25% for the largest plots. Bias results for the open-cluster plots will be given for the Canadian data set.

1.7.2 The Canada Population

The design effect results for the single cluster plots are summarized in Figure 18, where it can be seen that the precision of both the TI and WT estimators is superior to those of the BUFF estimators for all plot sizes, though the difference in design effect is not as dramatic as for the New Jersey data set. When the standard errors are compared (Figure 19), the pattern is similar to the New Jersey data set, where the standard error for WT and TI decreases across the range of plot radii, while the SE for the BUFF estimator decreases for plot radii up to approximately $r = 12$ and then begins to increase.

Figure 18 also suggests that for this population the WT technique is nearly as precise as the TI method, with the largest difference in the design effect being 1.04%.

Figure 20 shows the design effect for the open-cluster plot as a function of the number of sub-plots, the separation distance between plots, and the method of boundary-correction (i.e., $deff(BUFF, TI)$). As with the New Jersey data set, the estimator based on the TI boundary-correction technique is always more precise than the BUFF estimator and increasing the separation distance between sub-plots exacerbates the problem. Thus, the increase in within-cluster variability associated with spreading the plot over a larger area is easily offset by the increase in the variance associated with the BUFF technique. One interesting aspect of this data set is that there is no clear pattern in the design effect associated with the number of sub-plots.

An overall comparison of all methods is given in Figure 21, where the results are not as clear cut as for the New Jersey data. For this data set, all but one of the open-cluster plot designs is more precise than the single plot with radius $r = 5$ meters, provided

the separation distance was less than 20-30 meters. However, increasing the plot radius to $r = 10, 15, 20$ meters produces a very similar pattern to the New Jersey data. One interesting aspect of these results is that for the TI estimator, increasing the separation distance between the sub-plot generally reduced the standard error. Thus, the increase in within-cluster variability is greater than the efficiency lost by reducing the average area of the plot that samples the forest.

Figure 22 depicts the design-bias that results from ignoring the boundary-overlap problem. Only the open-cluster plots comprised of $o = 2, 4$, and 6 sub-plots are displayed. The graph illustrates that substantial biases will exist, even in situations where there is very little fragmentation.

1.7.3 Summary of Results for Both Populations

These populations attempt to represent the sampling conditions that would be expected in populations with both a low and high degree of fragmentation. The results suggest that the largest single factor in determining the efficiency of a plot configuration is the method of boundary-correction. Thus, unless the population has a very high interior-to-edge ratio, it is unlikely that the additional variability associated with adding the buffer strip, can realistically be overcome by increasing the size of the plot, the number of sub-plots, or the separation distance between the sub-plots.

1.8 Conclusions

This study identifies the major factors that need to be considered when selecting a fixed-area plot configuration for environmental surveys of fragmented populations. These factors include the method of boundary correction, the level of within-cluster heterogeneity as a function of plot shape and size, and the mean area of the plot that samples A. All of these factors are related to the level of fragmentation in the target population. For the two populations studied, the primary factor determining the efficiency of a sampling strategy was the boundary-correction method. The results indicate that the buffering method substantially reduced the efficiency of all of the sampling strategies considered, with a sampling strategy that used buffering requiring anywhere from 1.05 and 8 times

as many sample points in order to achieve the same variance as a sampling strategy that utilized either the walkthrough or truncated inclusion probability techniques. While the truncated inclusion probability approach was slightly more efficient than the walkthrough approach, the walkthrough approach can also accomodate other sampling methods, such as line-intersect (Kaiser 1983) or perpendicular distance sampling (Williams and Gove 2003), which can not realistically be accommodated by either the buffering or truncated inclusion probability approach. Thus, we concluded that if the goal of the survey is to assess aggregate quantities across a fragmented population, the only practical approach is to use a response design based on a single plot or point and use the walkthrough technique to correct for the unequal probabilities associated with sampling along the boundary.

The results of this paper also illustrate one of the substantial differences between the tessellated and continuous population approaches for large-scale surveys. This being that in contrast to the tessellated population approach, a continuous population sampling strategy based on open-cluster plots will almost surely be the least efficient type of ground plot unless the population has very little fragmentation. This leads to the obvious question of why the use of these plots has been advocated for so many years, particularly in light of the well known results of Smith (1938) on which these results were often based. While it is not possible to deduce the answer from the existing literature, two factors can be identified. The first is that some of the studies were carried out in parts of the country that have the lowest rates of fragmentation. For example, the state of Maine, which as the highest proportion of forested area in the US, is more than 90 % forested. The other factor is that the time frame in which the research advocating open-cluster plots appears also coincides with a period in time in which forest inventory plots that straddled a condition boundary (e.g., forest-nonforest, hardwood-softwood stand boundaries) were moved, or rotated, so that all of the sub-plots would be within the same forest condition (Birdsey 1995). Thus, it appears that the field trials that advocated the use of open-cluster plots were based on studies where data were only collected on plots that fell completely within a single homogeneous area. Thus, the influence of fragmentation was ignored.

While the use of open-cluster plots is questionable for this particular application it is important to realize that the goals of many environmental surveys often include other

applications and the adequacy of the plot configuration must be assessed for these applications as well. Many of these applications require some form of spatial modeling. For example, plot data are often used for growth and yield studies where individual trees on a plot are "grown" forward in time. The models that predict the rate of growth often require information about the competition between the individual trees, which in turn are related to the spacing between an individual tree and its neighbors. A problem with these applications is that trees beyond the boundary of the plot, that contribute to competition, are missing from the data. While various techniques exist to adjust for this boundary problem (Stoyan et al. 1995, Williams 2001 and references therein), for this application it is obvious that a single fixed-area plot would be superior to a number of smaller plots because large plots have a much smaller perimeter to interior ratio. On the other hand, many spatial modeling and mapping applications require the estimation of a variogram. In this setting, open-cluster plots are often thought to be beneficial because they contain information that is beneficial for estimating small-scale spatial variation (e.g., Olea 1984). Thus, the challenge is to select ground plot configurations that effectively meets all of the data needs, rather than focusing on a single objective.

1.9 Literature Cited

Bartlett, R.F., 1986. Estimating the total of a continuous population. J. Stat. Plann. Infer. 13:51-66.

Bellhouse, D. R. 1977. Some optimal designs for sampling in two dimensions. Biometrika 64:605-611.

Bickford, C. A. 1952. The sampling design used in the forest survey of the northeast. J. For. 50:290-293.

Bickford, C. A., Mayer, C. A., Ware, K. D 1963. An efficient sampling design for forest inventory: The northeast perspective. J. For. 61:826-833.

Birdsey, R. A. 1995. A brief history of the "straddler plot" debates. For. Sci. Monogr. 31:7-11.

Bitterlich, C. and Bitterlich, W. 1996. Relascope ideas relative measurements in forestry. C.A.B. International, New York.

Bush, J., House, C. 1995. The area frame: A sampling base for establishment surveys, In: Cox, B.G., Binder, D.A., Chinnappa, B.N., Christianson, A, Colledge, M.J., Kott, P. S. (Eds.) Business Survey Mehtods. Wiley, New York. pp. 335-344.

Butler B. J., Swenson, J. J., and Alig, R. J. 2004. Forest fragmentation in the Pacific Northwest:quantification and correlations. For. Ecol. Manage. 189:363-373.

Cochran, W. G. 1946. Relative accuracy of systematic and stratified samples from a certain class of populations. Ann Mathe. Stat. 17.164-177.

Cochran, W. G., 1977. Sampling Techniques. 3^{rd} ed. J. Wiley & Sons, New York.

Cordy, C. 1993. An extension of the Horvitz-Thompson Theorem to point sampling from a continuous universe. Prob. Statistics Letters. 18:353-362.

de Vries, P.G., 1986. Sampling Theory for Forest Inventory. Springer-Verlag, New York.

Diggle, P. J., 1983. Statistical Analysis of Spatial Point Patterns. Academic Press. London

Ducey, M.J., Gove, J.H., and Valentine, H.T. 2004. A walkthrough solution to the boundary overlap problem. For Sci. 50:427-435.

Dunn, R. and Harrison, A. R. 1993. Two-dimensional systematic sampling of land use. Applied Stat. 42:585-601.

Ek, A. R. 1969. Stem map data from three forest stands in Northern Ontario. Canada Dept. Fisheries and Forestry, Can. Forest. Serv., Inf. Rep. O-X-113. p. 22.

Eriksson, M., 1995. Design based approaches to horizontal-point-sampling. For. Sci. 41(4):890-907.

Essen, P. E. 1994. Tree mortality patterns after experimental fragmentation of an old-growth conifer forest. Biol. Conserv. 68:19-28.

Frayer, W. E., and G. M. Furnival. 1999. Forest survey sampling designs: A history. J. For. 97:4-8.

Freese, F. 1961. Relation of plot size to variability: an approximation. J. For. 59:679

Goebel, J. J., and George, T. A. 1990. Establishing a consistent national base for assessing natural resource issues. *in* proc. 30th International Atlantic Economic Conference. Williams, VA. Oct. 11-14.

Gove, J. H., Ringvall, A., Stahl, G., and Ducey, M. J. 1999. Point relascope sampling of downed coarse woody debris. Can. J. For. Res. 29:1718-1726.

Gregoire, T. G., 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Can. J. For. Res. 28:1429-1447.

Gregoire, T. G., and Monkevich, N. S. 1994. The reflection method of line-intersect sampling to eliminate boundary bias. Environ. Ecol. Stat. 1:219-226.

Husch, B., Miller, C.I., Beers T.W., 1982. Forest Mensuration 3rd edn. J. Wiley and Sons, New York.

Iachan, R. 1985. Plane sampling. Stat. Prob. Letters. 50:151-159.

Kaiser, L., 1983 Unbiased estimation in line-intersect sampling. Biometrics. 39:965-976.

Labau, V. J., Bones, J. T., Kingsley, N. P., Lund, H. G., and Smith, B. W. 2005. A history of the USFS Forest Survey 1900-2003. USDA For. Serv. Gen . Tech. Rep. (in press).

Mandallaz, D. 1991. A unified approach to sampling theory for forest inventory based on infinite population and superpopulation models. Chair of Forest Inventory and Planning, Swiss Federal Institute of Technology, Zurich. 256 p.

Masuyama, M. 1954. On the error in crop cutting experiment due to the bias on the border of the grid. Sankhya 14: 181-186.

Messer, J.J., Linthurst, R. A., and Overton, W. S. 1991. An EPA program for monitoring ecological status and rends. Environ. Mon. Assesst. 20:67-78.

Nusser, S. M., F. J. Breidt, and W. A. Fuller. 1998. Design and estimation for investigating the dynamics of natural resources. Ecol. Applic. 8:234-245.

Olea, R. A., 1984. Sampling design optimization for spatial functions. Mathe. Geol. 16:369-392.

Olsen, A. R., and Schreuder, H. T. 1997. Perspectives on large-scale natural resource surveys where cause-effect is a potential issue. Environ. Ecol. Stat. 4:167-180.

Overton, W. S. and Stehman S. V. 1995. Design implications of anticipated data uses for comprehensive environmental monitoring programmes. Environ. Ecol. Stat. 2:287-303.

Ripley, B. D., 1981. Spatial Statistics. J. Wiley & Sons, New York

Riitters, K. H., Wickman, J. D., O'Neill, R. V., Jones, K.B., Smith, E. R., Coulston, J. W., Wade, T. G. and Smith, J. H. 2002. Fragmentation of continental United States forests. Ecosystems. 5:815-822.

Roesch, F. A., Green, E. J., and Scott, C. T. 1993. An alternative view of forest sampling. Surv. Meth. 19:199-204.

Sarndal, C.E., Swensson, B., Wretman, J., 1992. Model Assisted Survey Sampling. Springer-Verlag, New York.

Schmid-Haas, P. 1982. Sampling at the forest edge. pp. 263-276 in B. Ranneby, ed. Statistics in Theory and Practice: Essays in Honor of Bertil Matern. Swedish University of Agricultural Sciences, Umea.

Schreuder, H.T., Gregoire, T. G., and Wood, G. B. 1993. Sampling methods for multiresource forest inventories. J. Wiley & Sons, New York.

Spada, B. 1960. A test of several designs for sampling an acre to obtain forest survey volume and area statistics and area condition classification data. In-service Report 9/6/60. Portland, OR: US Dept. of Agric, For. Serv. PNWRS

Scott, C. T. 1993. Optimal design of a plot cluster for monitoring. (In) The optimal design of forest experiments and forest surveys. Rennolls, K., and Gertner, G. (eds.). University of Greenwhich Shool of Math. Stat. and Comput. London, UK. pp. 233-244.

Shivers, B. D., Borders, B. E., 1996. Sampling Techniques for Forest Resource Inventory. J. Wiley and Sons, New York.

Smith, H. F. 1938. An empirical law describing heterogeneity in the yield of agricultural crops. J. Agric. Sci. 28:1-23.

Stein, M. L. 1995. Locally lattice sampling designs for isotropic random fields. Ann. Statistics. 23:1991-2012.

Stein, M. L. 1999. Interpolation of spatial data: Some theory for Kriging. Springer, NY

Stevens, D.L., Jr., 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics 8:167-195.

Stevens, D. L. Jr., and Urquhart, N. S. 2000. Response designs and support regions in sampling continuous domains. Environmetrics 11-13-41.

Stevens, D. L. Jr., and Olsen, A. R. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14:593-610.

Stevens, D. L. Jr., and Olsen, A. R. 2004. Spatially balanced sampling of natural resources. J. Am. Stat. Assoc. 99:262-278.

Stoyan, D., Kendall, W. S. and Mecke, J. 1995. Stochastic Geometry and Its Applications, 2^{nd} ed. J. Wiley & Sons, New York.

Taylor, H. T., and S. Karlin. 1994 An Introduction to Stochastic Modeling. Academic Press. New York. 566 p.

Thompson, S.K., 2002. Sampling, 2^{nd} ed. J. Wiley & Sons, New York.

Upton, G. J. G., Fingleton, B., 1985. Spatial Data Analysis by Example. J. Wiley and Sons, New York.

Valentine, H. T., Gove, J. H., and Gregoire, T. G. 2001. Monte Carlo approaches to sampling forested tracts with lines or points. Can. J. For. Res. 31:1410-1424.

Valentine, H. T., Ducey, M. J., and Gove, J. H. 2005. A Walkabout correction for cluster-plot slop. For. Sci. (in press).

Vries, P.G. de. 1986. Sampling theory for forest inventory. A teach yourself course. Springer-Verlag, New York. 399 p.

Williams, M. S. 2001a. Performance of two fixed-area (quadrat) sampling estimators in ecological surveys. Environmetrics 12:421:436

Williams, M. S. 2001b. New approach to areal sampling in ecological surveys. For. Ecol. and Manage. 154:11-22.

Williams, M. S., Williams, M. T., and Mowrer, H. T. 2001. A boundary reconstruction technique for circular fixed-area plots in environmental survey. J. Agric. Biol. Environ. Stat. 6:479-494

Williams, M. S., and Eriksson, M. 2002. Comparison of two paradigms for fixed-area sampling. For. Ecol. and Manage. 168:135-148

Williams, M. S., and Gove, J. H. 2003. Perpendicular distance sampling: Another method of sampling coarse woody debris. Can. J. of For. Res. 33:1564-1579.

1.10 Appendix

With the exception of the shape and size of the inclusion zones, the properties of all continuous population estimators are identical in the context of survey sampling. As suggested by Gregoire (1998), a concise description of the properties of $\hat{Z}_{TI}$ is as follows:

The area of interest is a two-dimensional plane, A, of area $|A|$ and the target parameter is the total of some attribute z_i across all elements within A. Inference follows the design-based framework, which views the population as fixed. The only source of randomness is the selection of sampling locations (points), which are determined by generating the (s) coordinates of m independent sampling locations from a Uniform distribution over A. Thus, the probability density function of the sample locations is $f(s) = 1/|A|$. At each location the elements selected by any of the plot configurations are tallied. These data are then used to determine an estimate of the total at the point (s), denoted by

$$\hat{Z}_{TI}(s) = |A| \sum_{i=1}^{n(s)} z_i / a_i^T, \qquad [A.1]$$

where $n(s)$ is the number of elements in the sample at the point with coordinates (s), a_i^T is the area of the inclusion zone falling within A for element i, and z_i is the parameter of interest. The infinite set of $\hat{Z}_{TI}(s)$ values over A forms the reference set, with the reference distribution being the infinite set of all equally likely samples.

The estimator is design-unbiased with the expected value of $\hat{Z}_{TI}$ given by

$$E[\hat{Z}_{TI}] = \frac{1}{|A|} \int_A \hat{Z}_{TI}(s) ds = \frac{1}{|A|} \sum_{i=1}^{N} (\frac{z_i |A|}{a_i^T}) a_i^T = \sum_{i=1}^{N} z_i = \tau_Z, \qquad [A.2]$$

where N is the number of elements. Assuming that m independent sample points are drawn from $f(s)$, the estimator for these points is $\bar{Z}_{TI} = 1/m \sum_{i=1}^{m} \hat{Z}_{TI}(s_i)$, where $\hat{Z}_{TI}(s_i)$ is the estimator in A.1 for point s_i. The variance is given by

$$Var[\hat{Z}_{TI}] = \frac{1}{|A|} \int_A (\hat{Z}_{TI}(s) - \tau_Z)^2 ds. \qquad [A.3]$$

The sample based variance estimator is derived from the observed variation among the m sample points. Thus, for the m randomly located points, the design-unbiased estimator of the variance is

$$v[\bar{Z}_{TI}] = \frac{1}{m(m-1)} \sum_{i=1}^{m} (\hat{Z}_{TI}(s_i) - \bar{Z}_{TI})^2. \qquad [A.4]$$

1.11 Figure Captions

Figure 1. Example of an area frame imposed on a one square hectare population consisting of forested and nonforested areas. The frame is comprised of four primary sampling units and each primary units is subdivided into nine secondary sample units. Some forested area exits in 3 out of the 4 PSUs

Figure 2. The same one hectare population without the area frame. Only approximately 42 % of the hectare is covered by forest.

Figure 3. The open-cluster plots used in the study ranged from 2 to 7 sub-plots. The radii of each sub-plot was $5m$. The small circle indicates the center point used to locate the plot within the population.

Figure 4. The range of separation distances considered ranged from $d = 10 - 40$m in increments of 5 m. The range of separation distances for the open-cluster plot comprised of 7 sub-plots is given here.

Figure 5. The range of fixed-area plot sizes considered ran from $r = 2 - 20$ m.

Figure 6. The figure on the left show a three meter radius plot located at the origin. The center-point in denoted by the + symbol. Two trees, denoted by the *'s, fall within the boundary of the plot and assume that the population A is the area without the cross-hatching. The figure on the right shows the object-centered inclusion zones centered about each of the two trees. The inclusion zone for the tree on the right is truncated by the boundary.

Figure 7. The buffering solution to the boundary overlap problem extends the boundary of A so that no inclusion zone is truncated by the boundary.

Figure 8. To perform the walkthrough technique, each tree between the boundary of A and the center-point of the plot is checked by walking from the center-point "through" the tree and traveling an equal distance past the tree. If the boundary of A is encountered before the end of the walkthrough path is reached, the tree is counted twice.

Figure 9. If the crossed area denotes a fragment of A, then the average area of A sampled by an open-cluster plot will be reduced when the separation distance between the sub-plots is increased.

Figure 10. Overview of the New Jersey data set with an open-cluster plot of 6 sub-plots and a 20 m separation distance.

Figure 11. The Canada data set.

Figure 12. Standard errors for the New Jersey population when single fixed-area plots with radii ranging from $r = 2 - 20$ are used.

Figure 13. Design effect results for the New Jersey population when single fixed-area plots with radii ranging from $r = 2 - 20$ are used.

Figure 14. Standard errors for the New Jersey population across the range of open-cluster plot design. The results on the top are for $\hat{Z}_{BUFF}$ and the lowe cluster of results are for $\hat{Z}_{TI}$

Figure 15. Design effect results for the New Jersey population across the range of open-cluster plot design. The design effect is given by $Var[\hat{Z}_{BUFF}]/Var[\hat{Z}_{TI}]$.

Figure 16. Overall results for the New Jersey data set.

Figure 17. Ratio of the expected value of the estimator to the true total for the three boundary-correction techniques and to an estimator that is uncorrected.

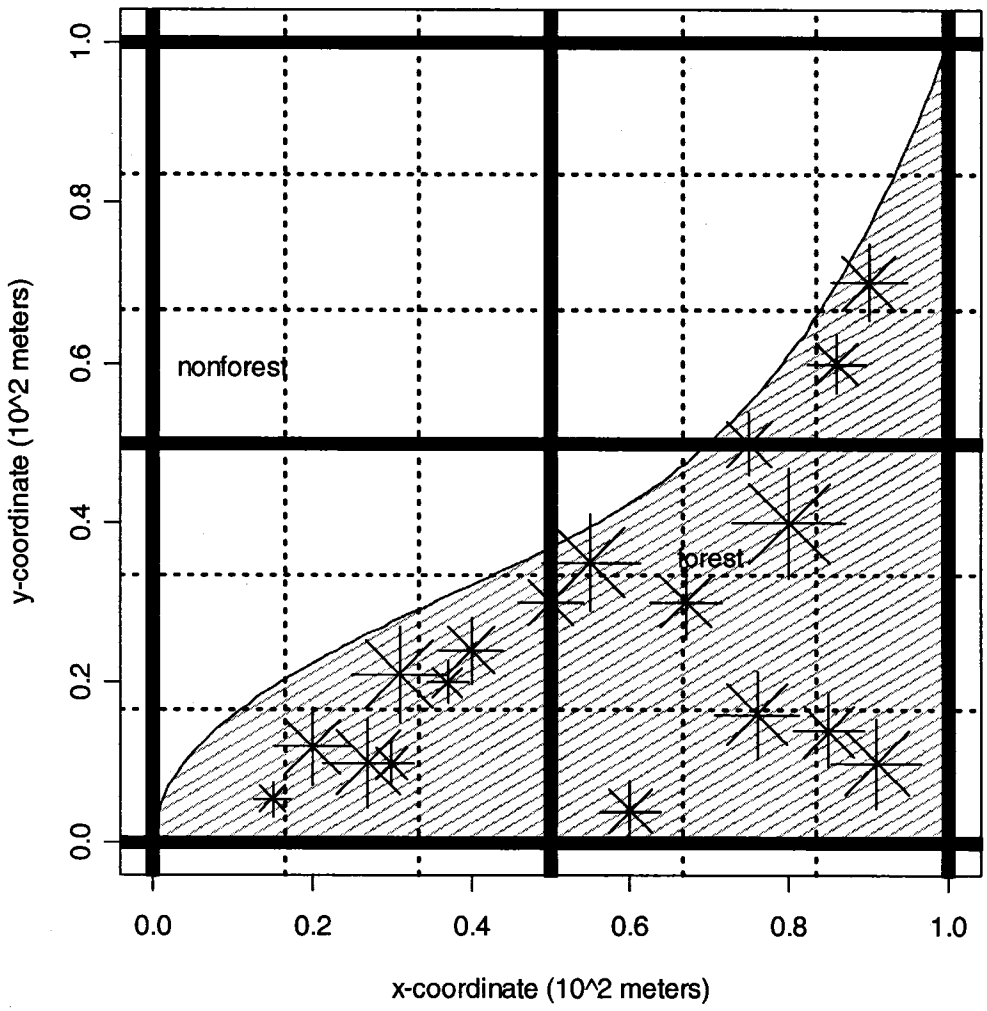
Figure 18. Design effect results for the Canada data set when single fixed-area plots with radii ranging from $r = 2 - 20$ are used.

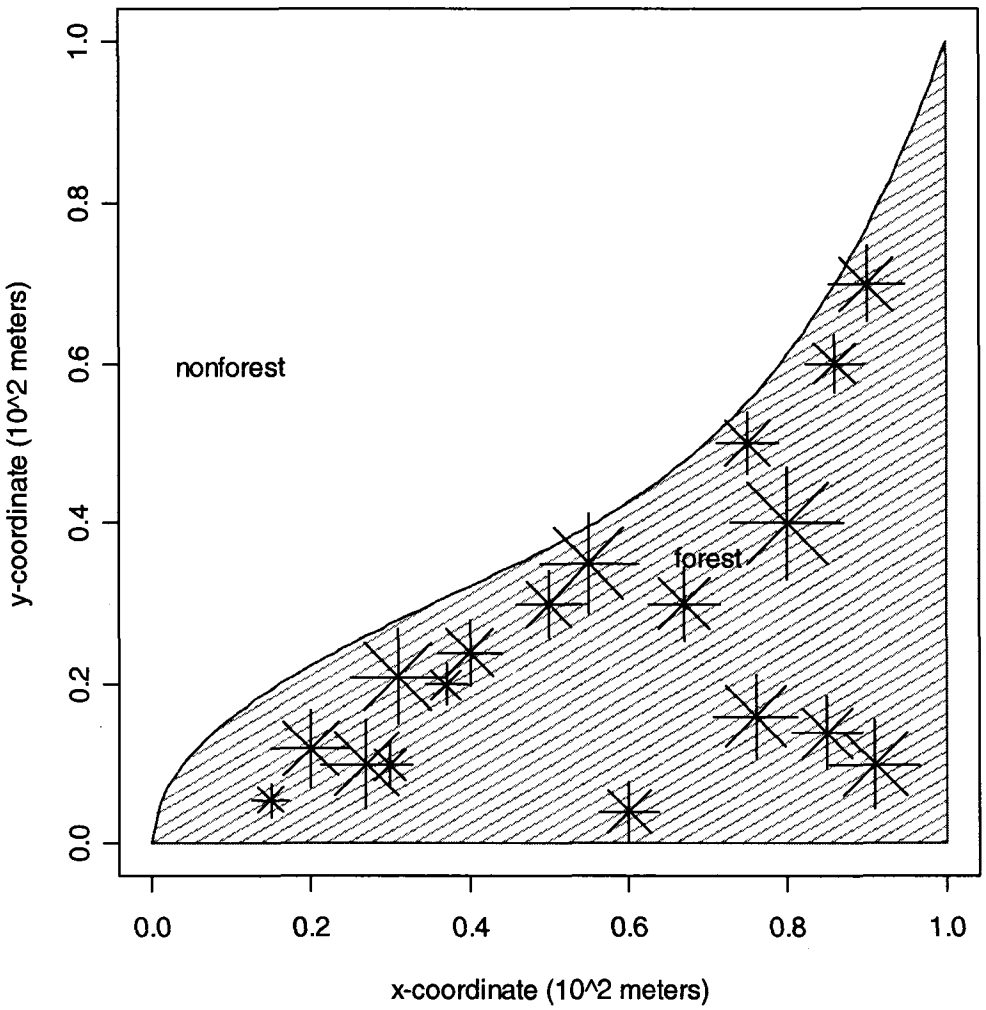
Figure 19. Standard error results for the Canada data set for single cluster plots

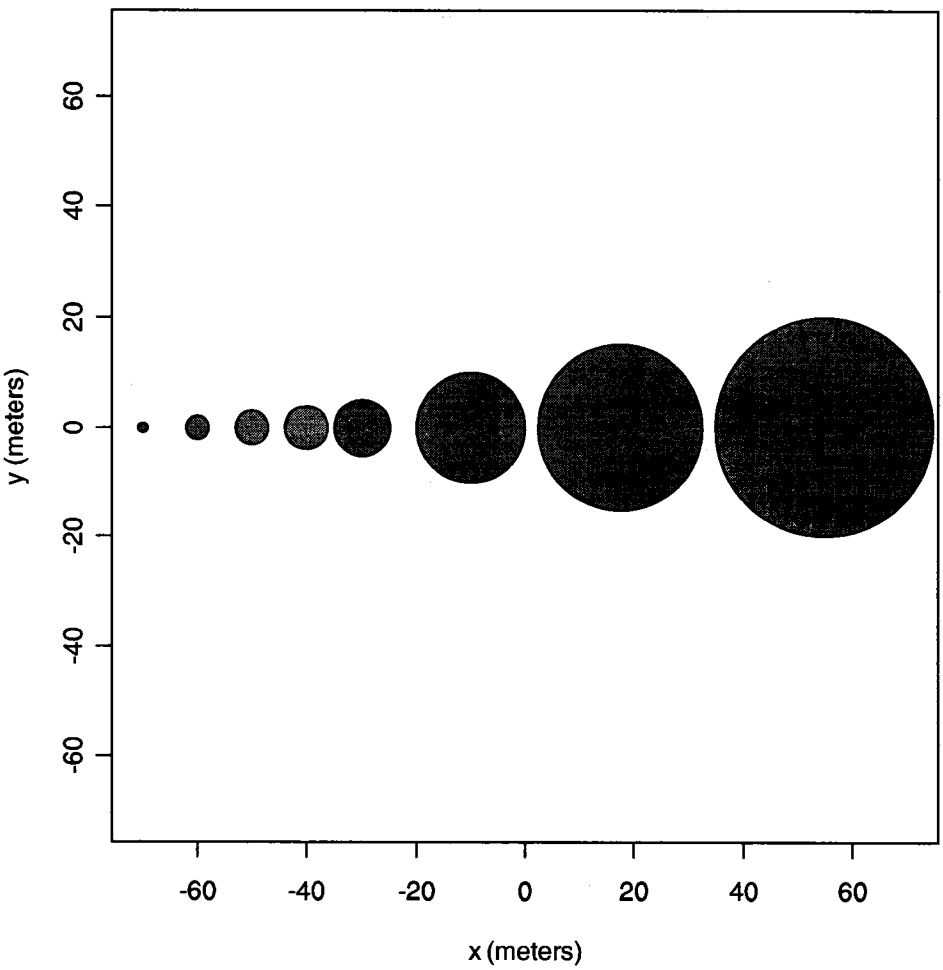
Figure 20. Design effect for the Canada data set using open-cluster plots.

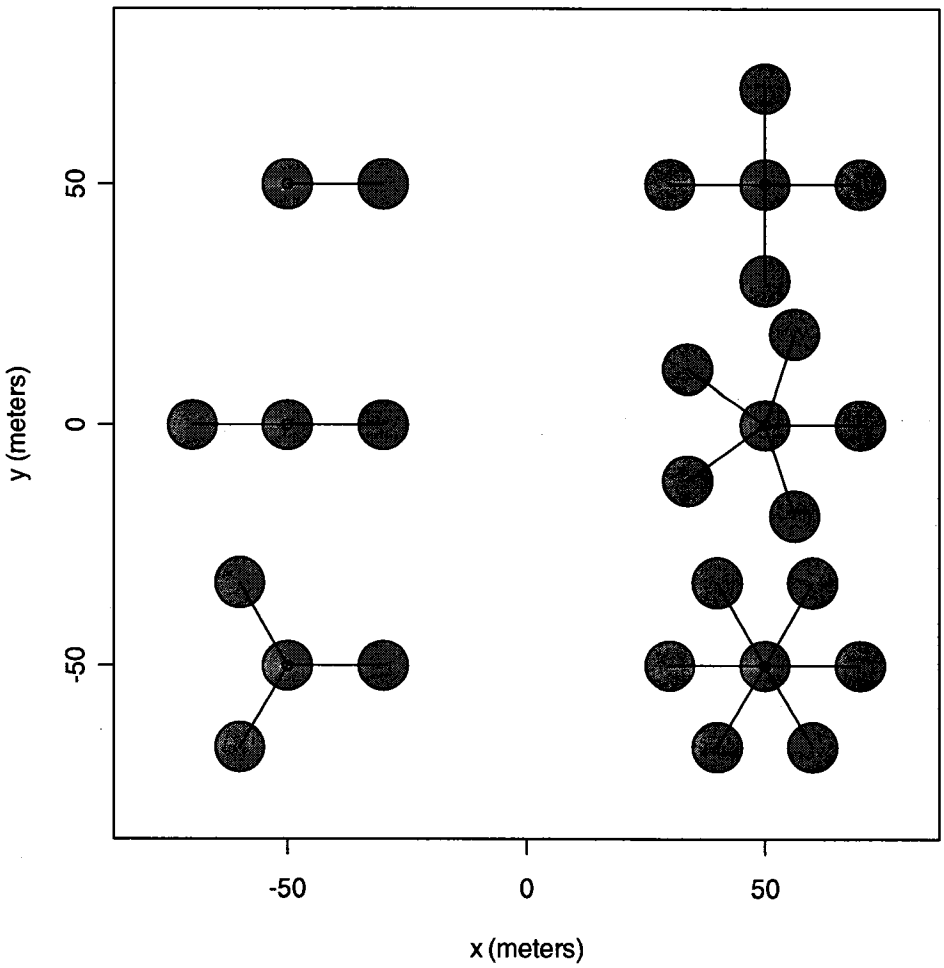
Figure 21. Overall results for all plot configurations for the Canada data set.

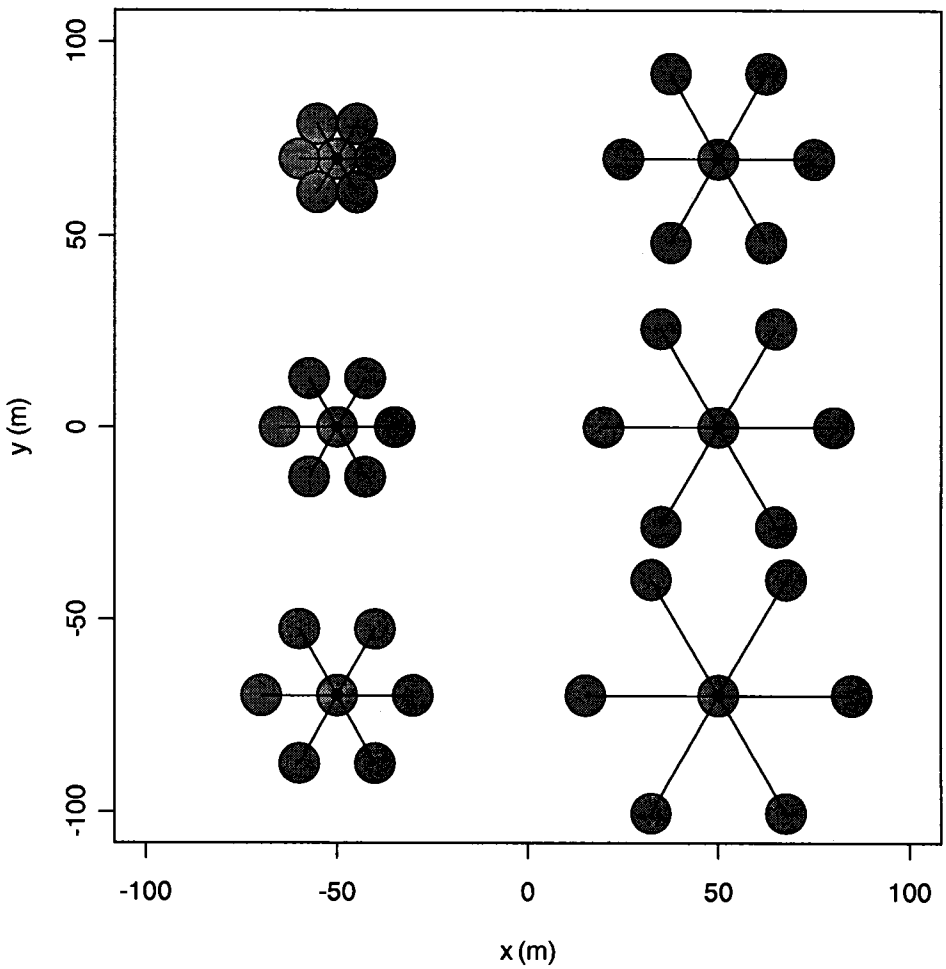
Figure 22. Ratio of the expected value of the estimator to the true total for the three boundary-correction techniques and to an estimator that is uncorrected.

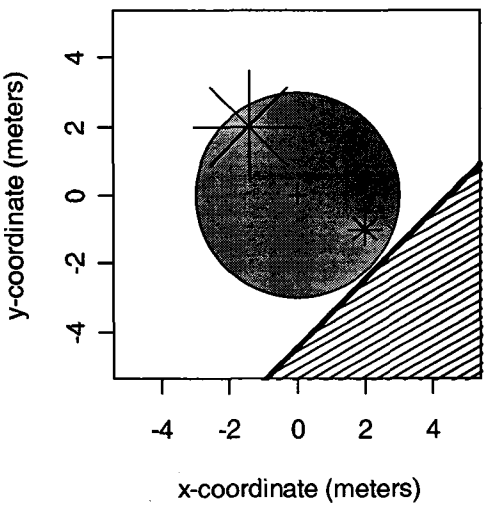


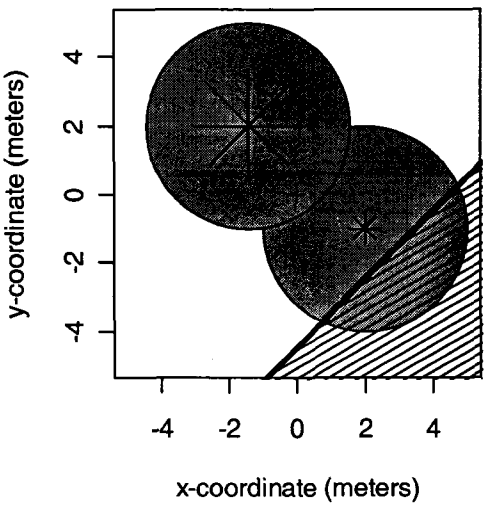


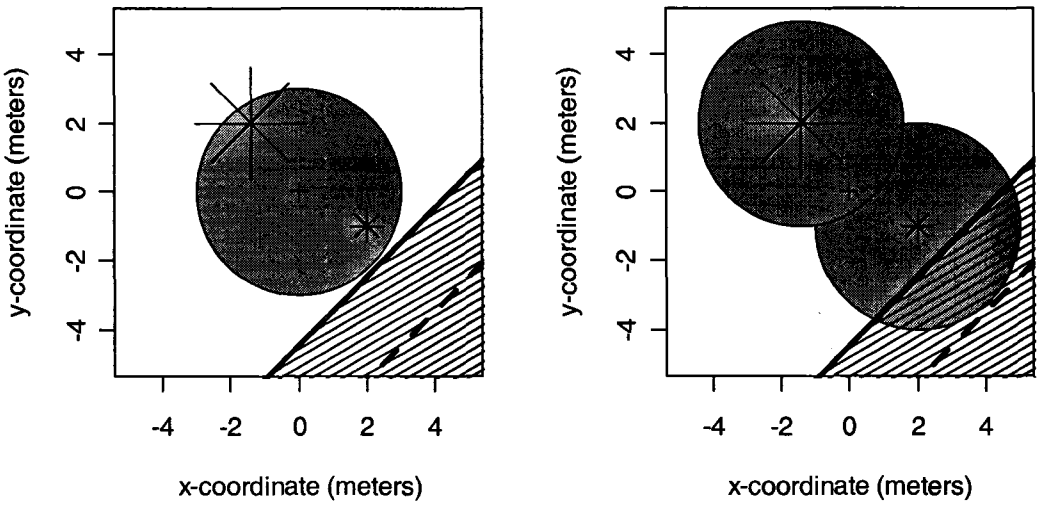


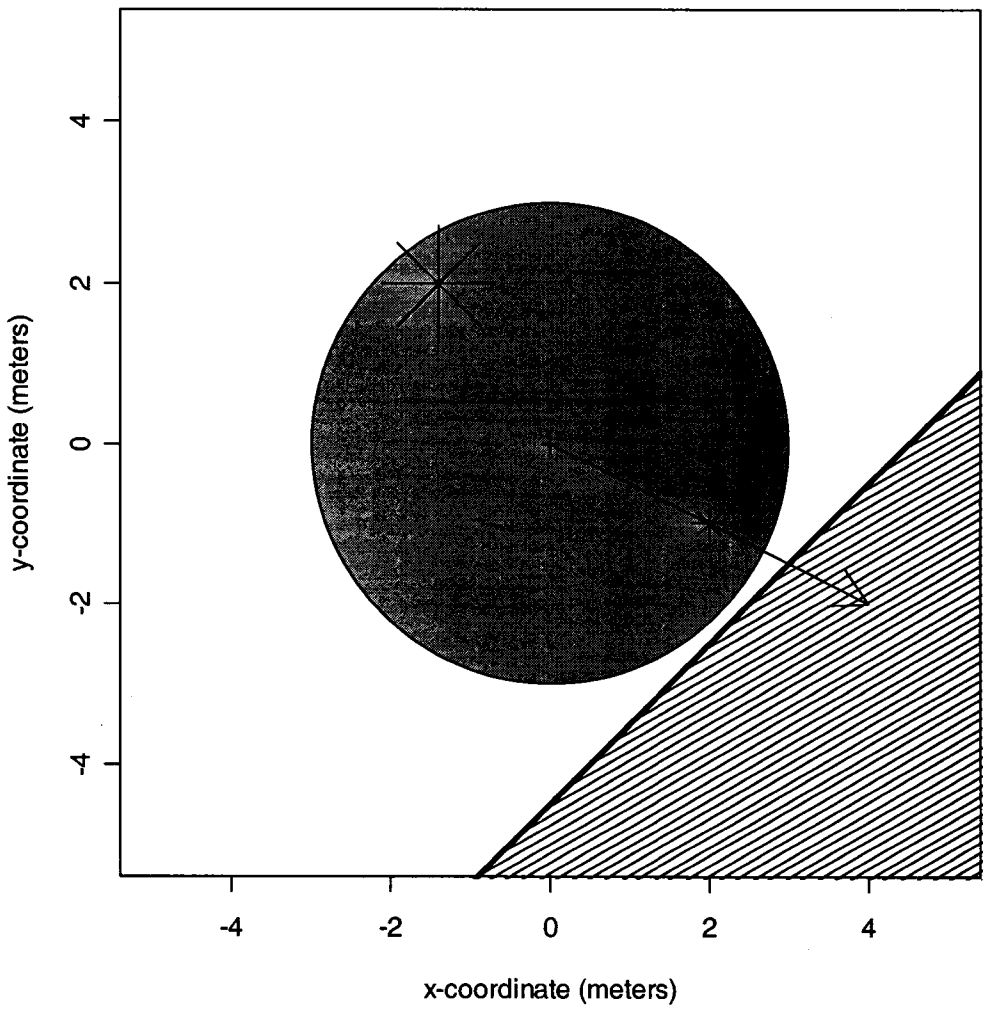


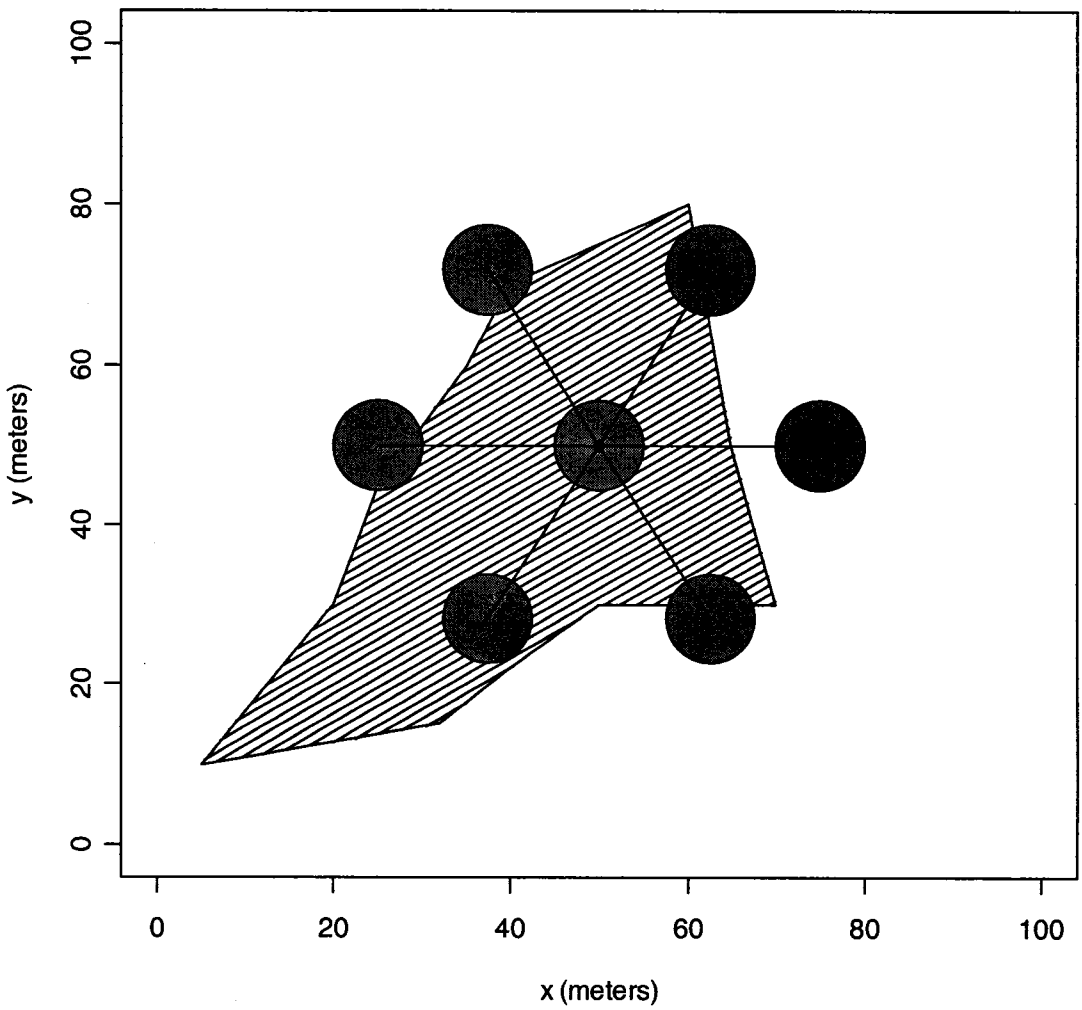


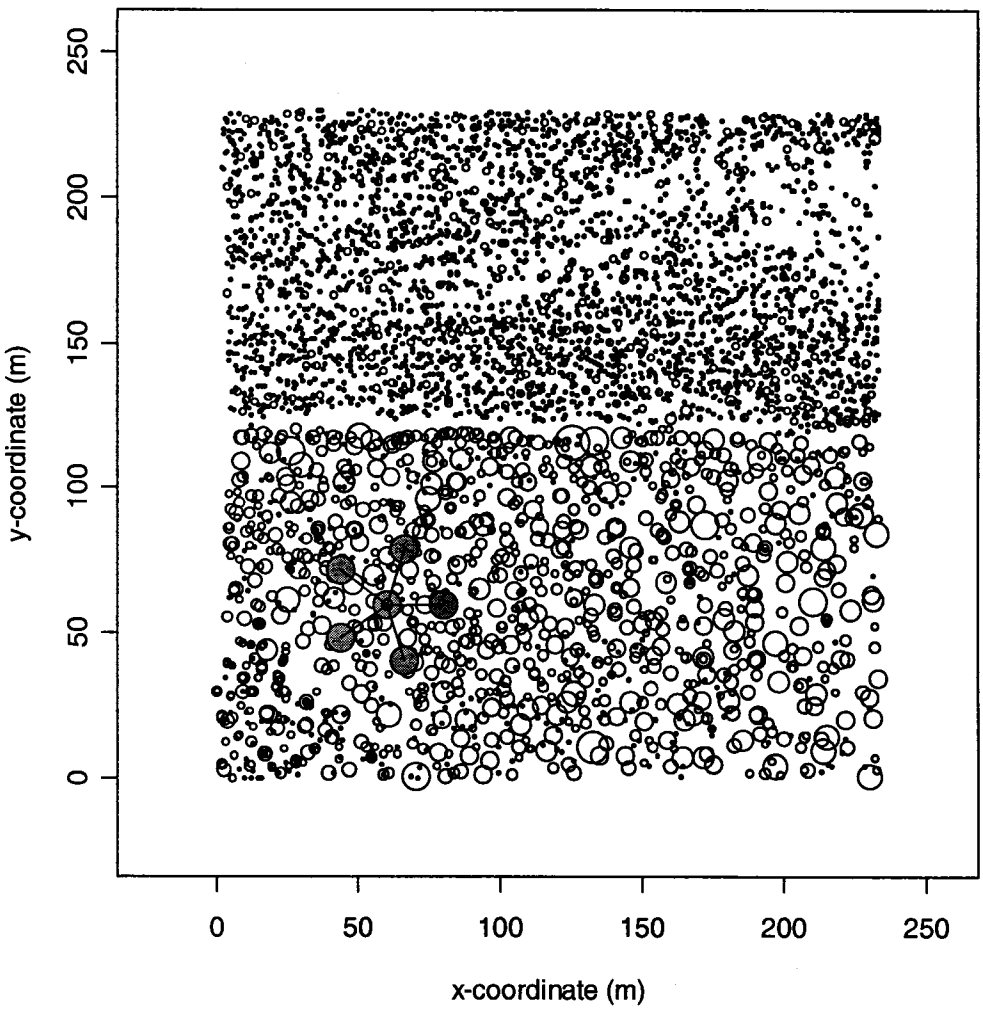


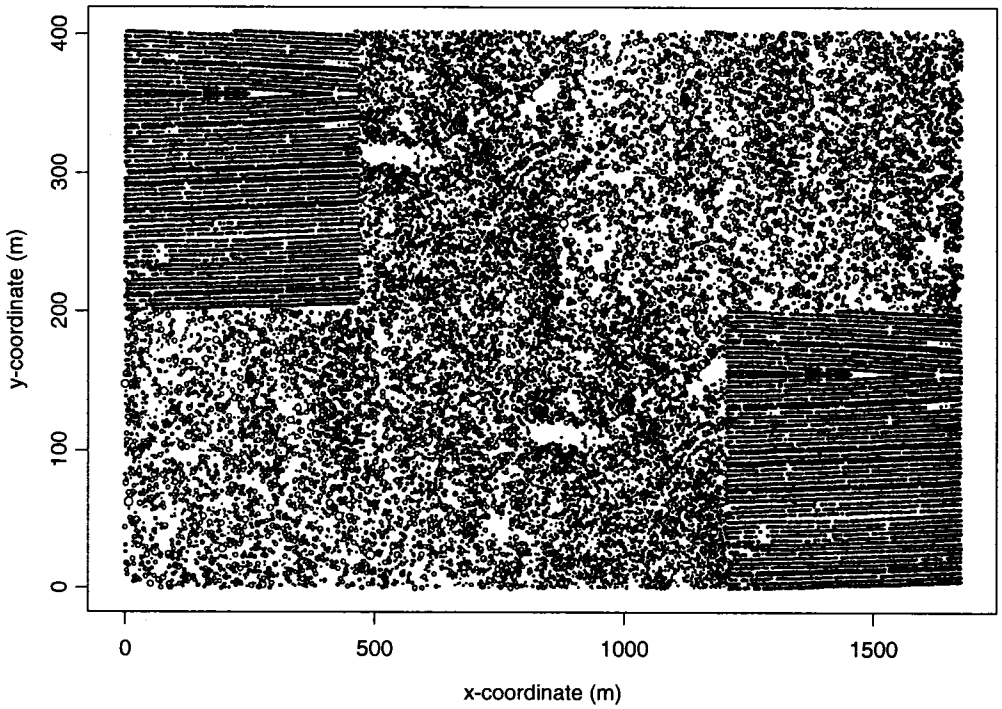


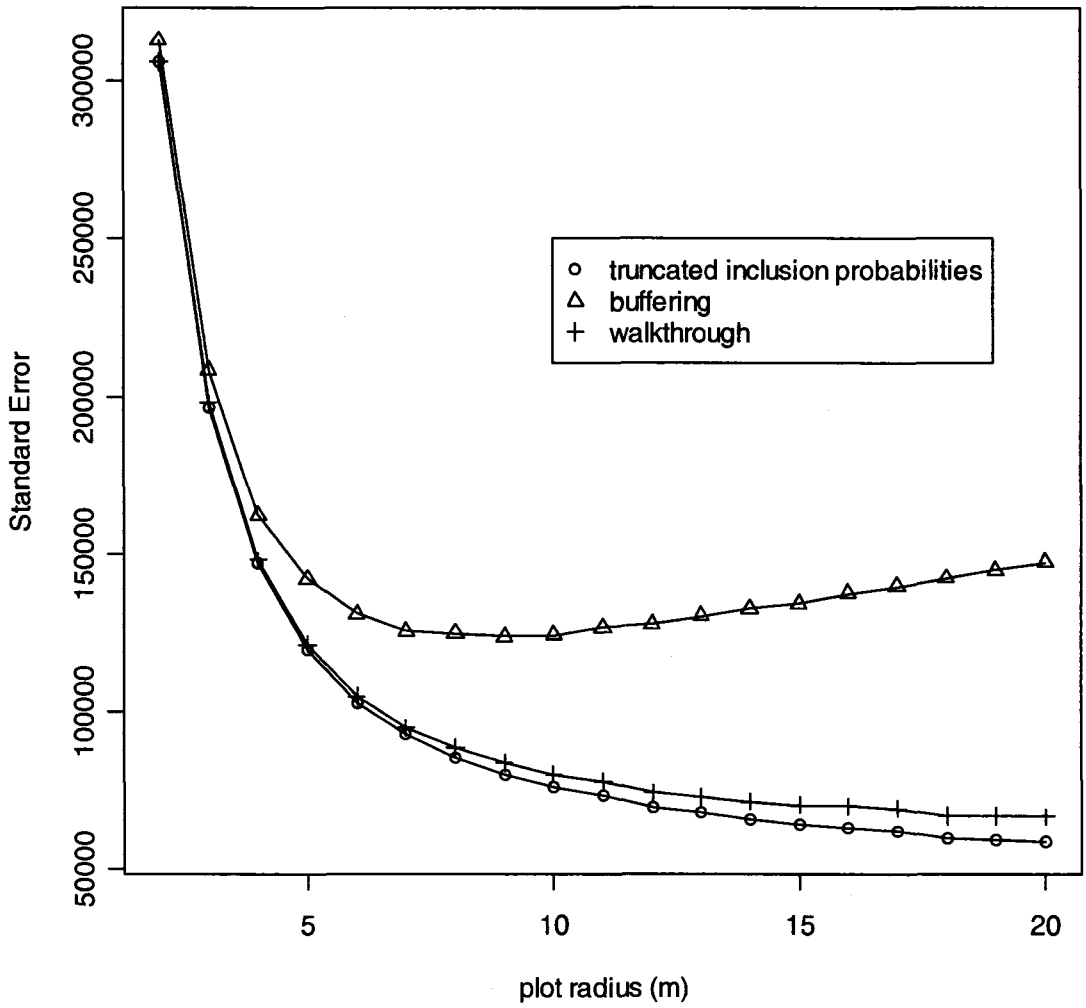


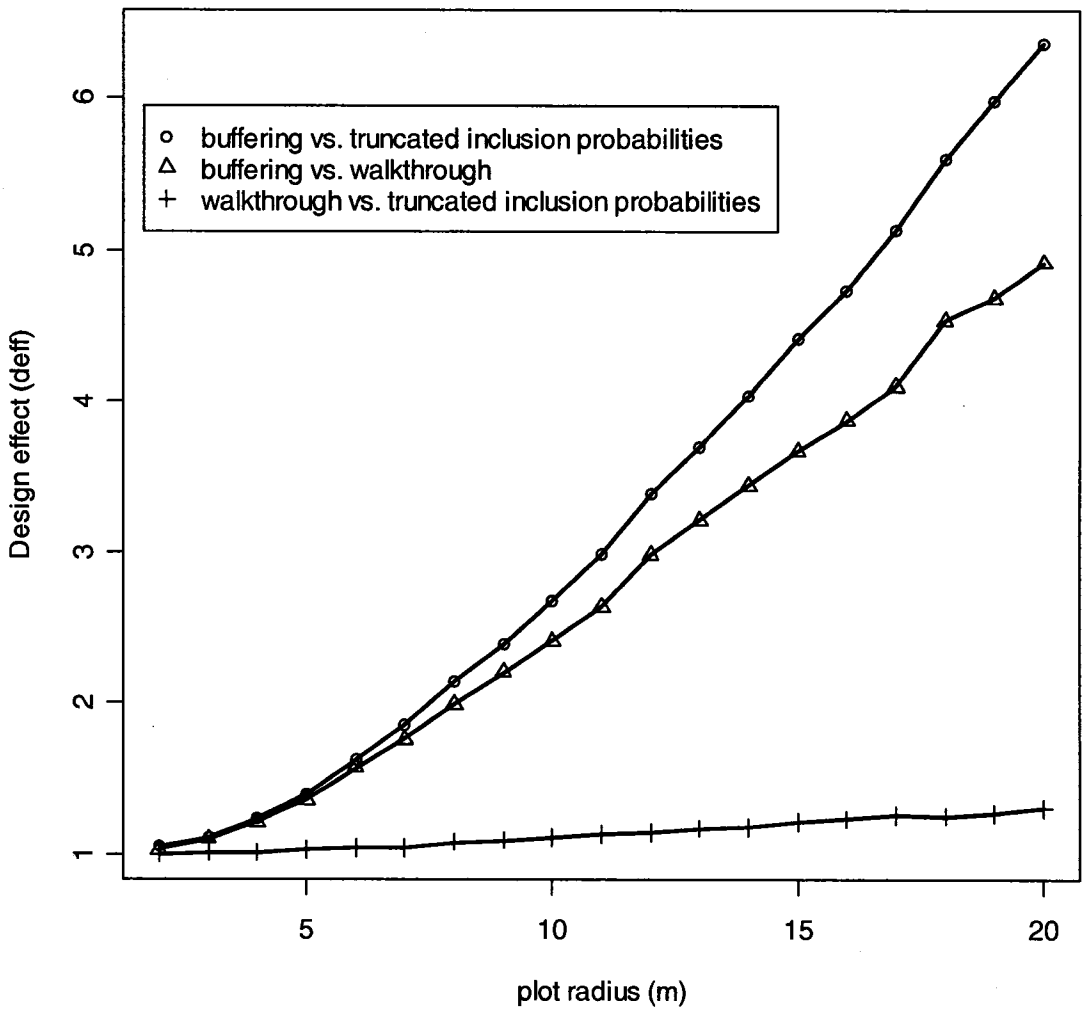


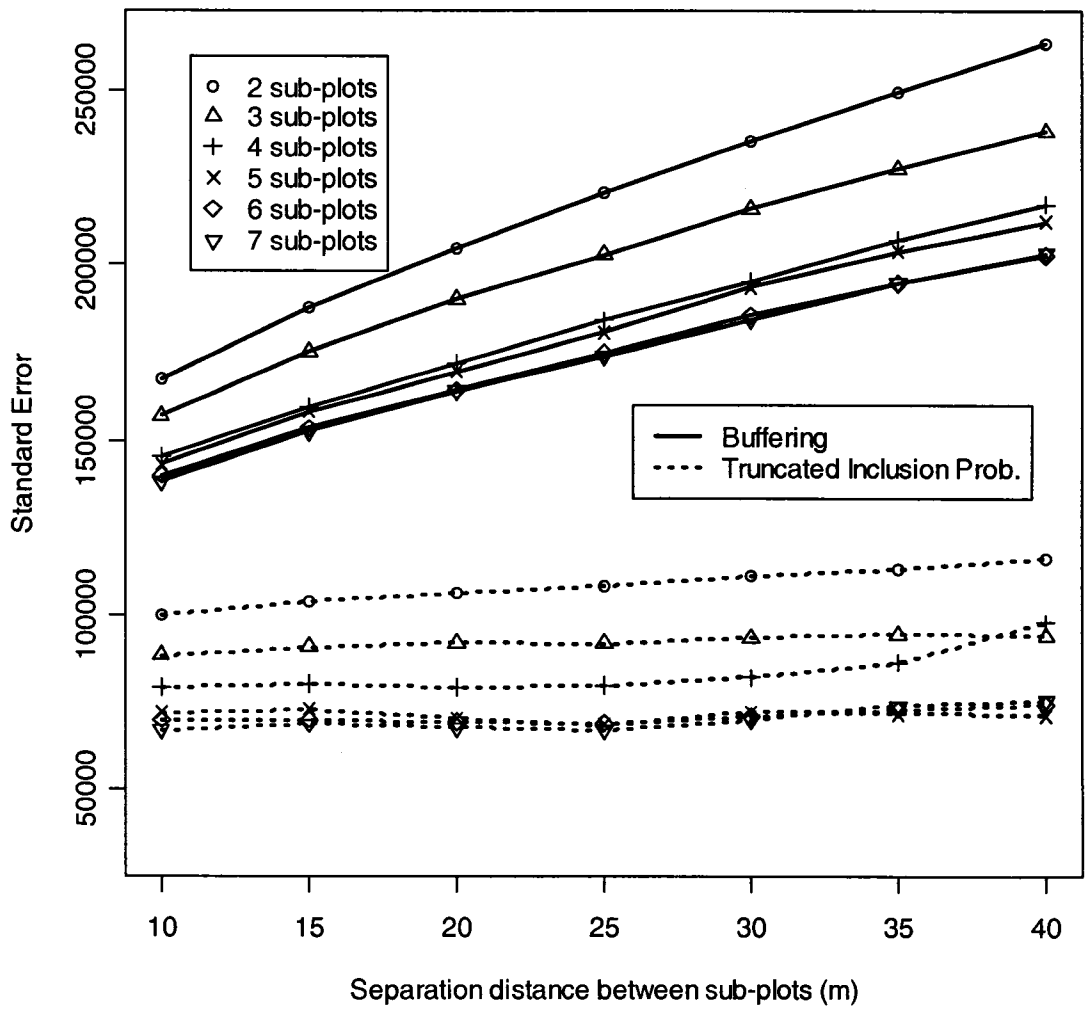


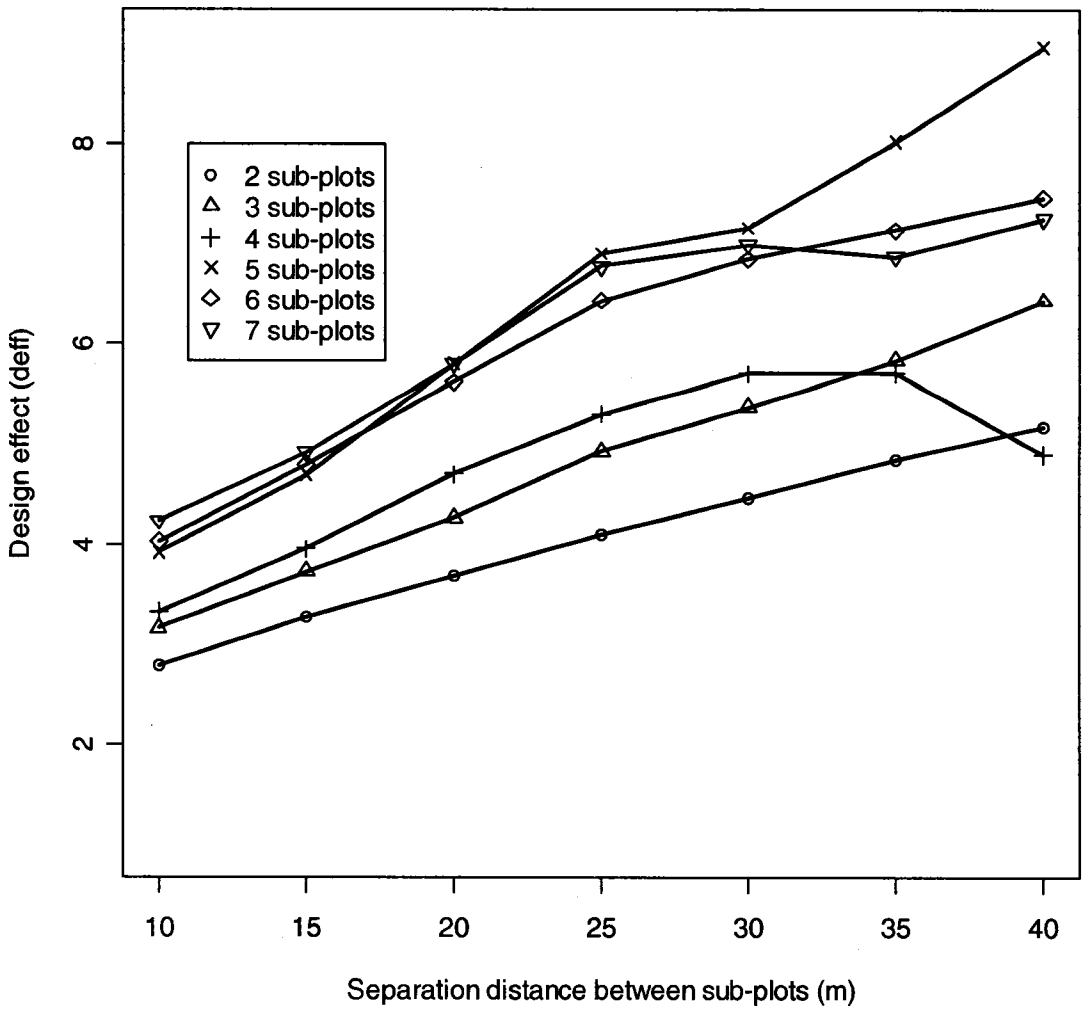


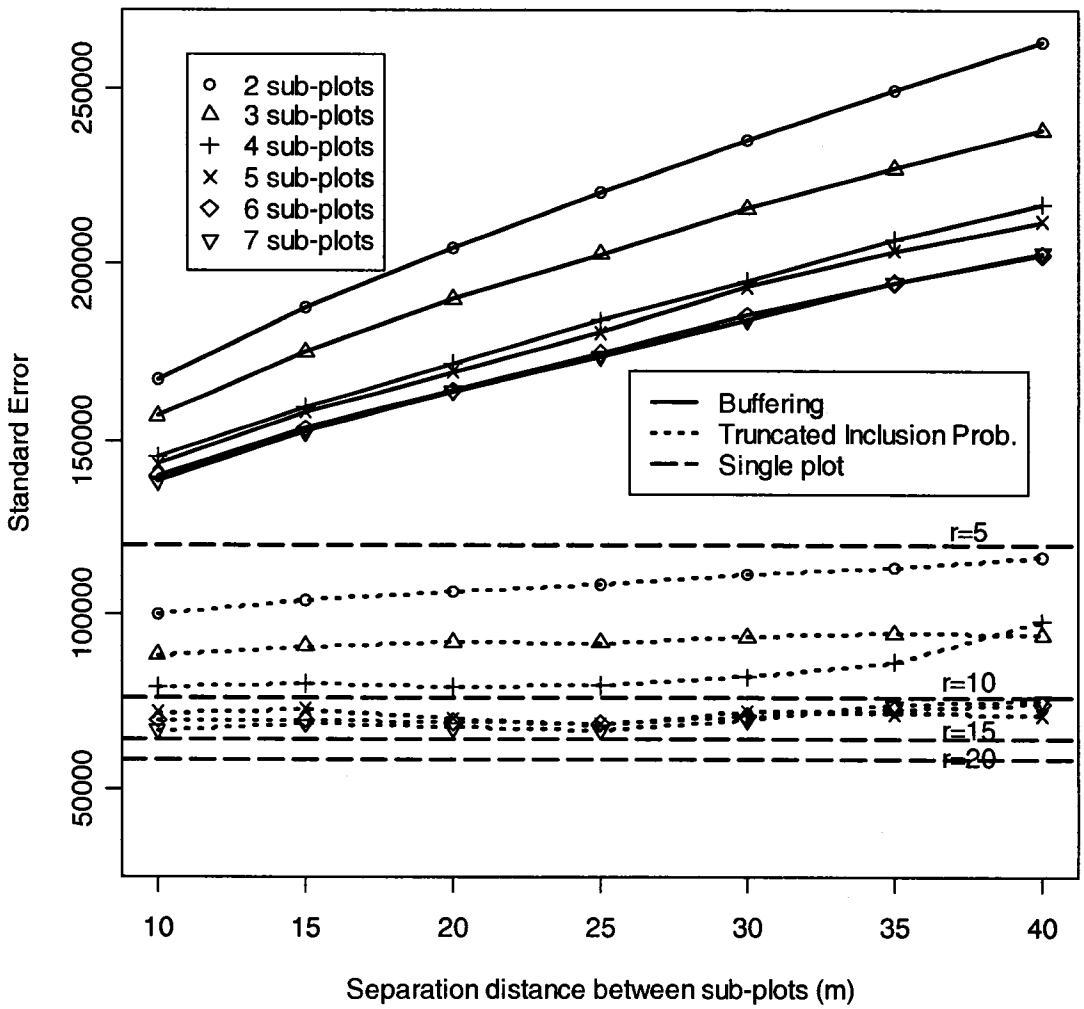


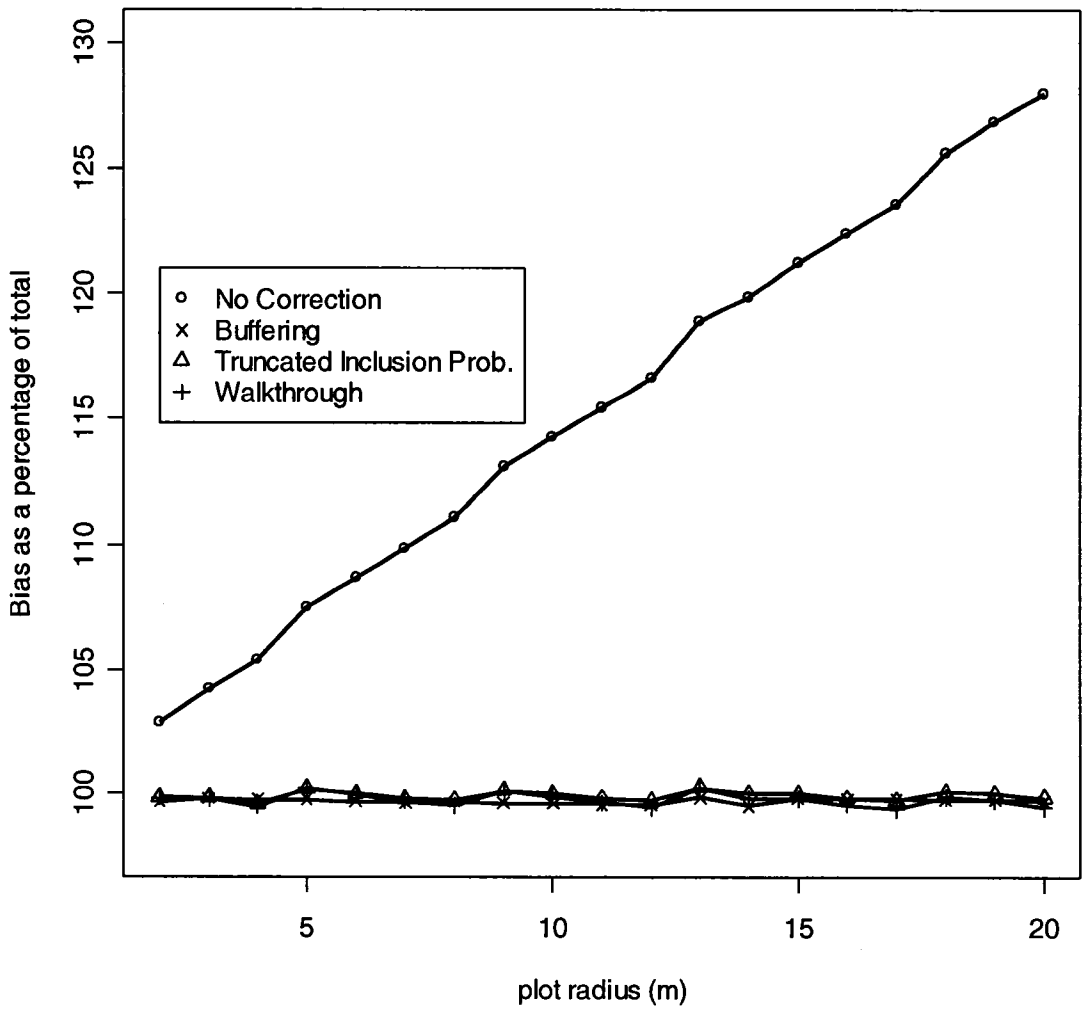


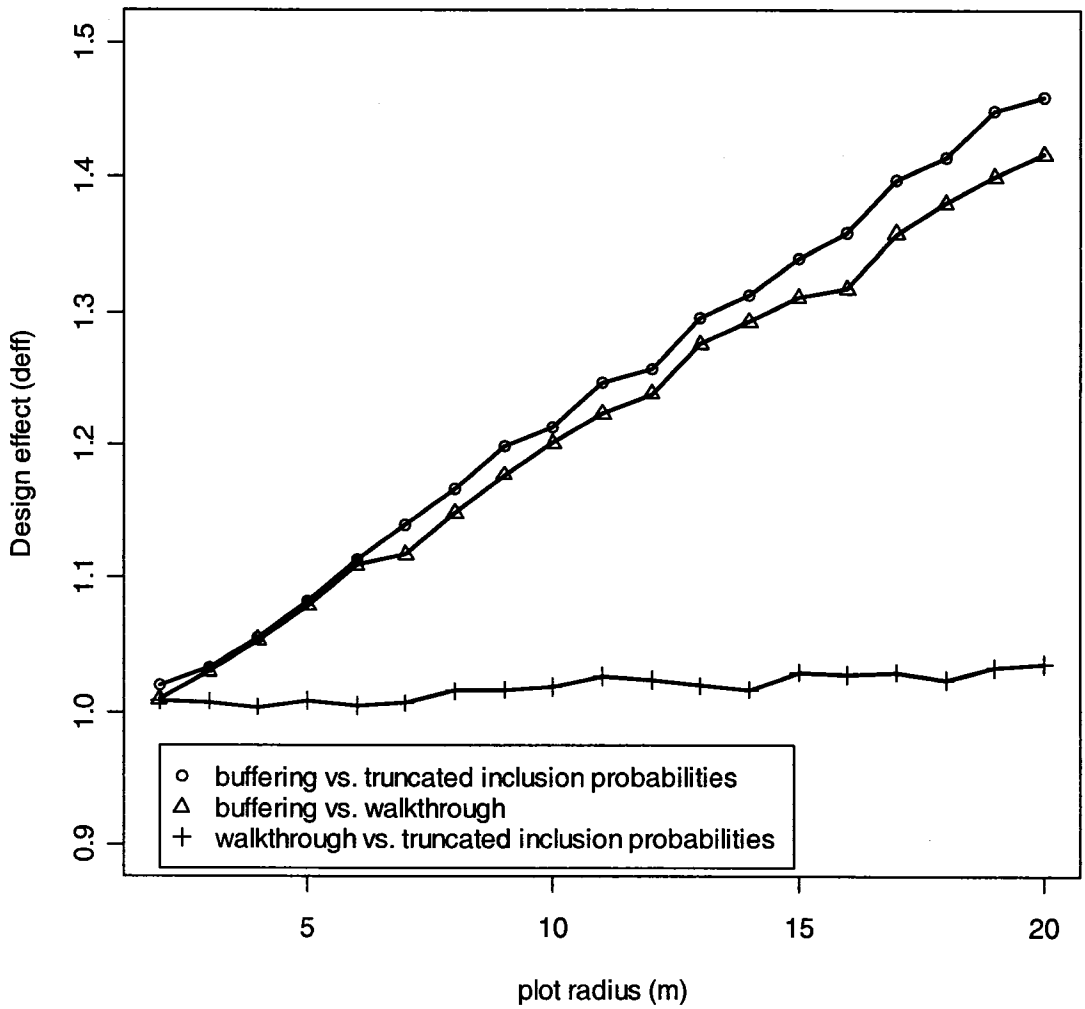


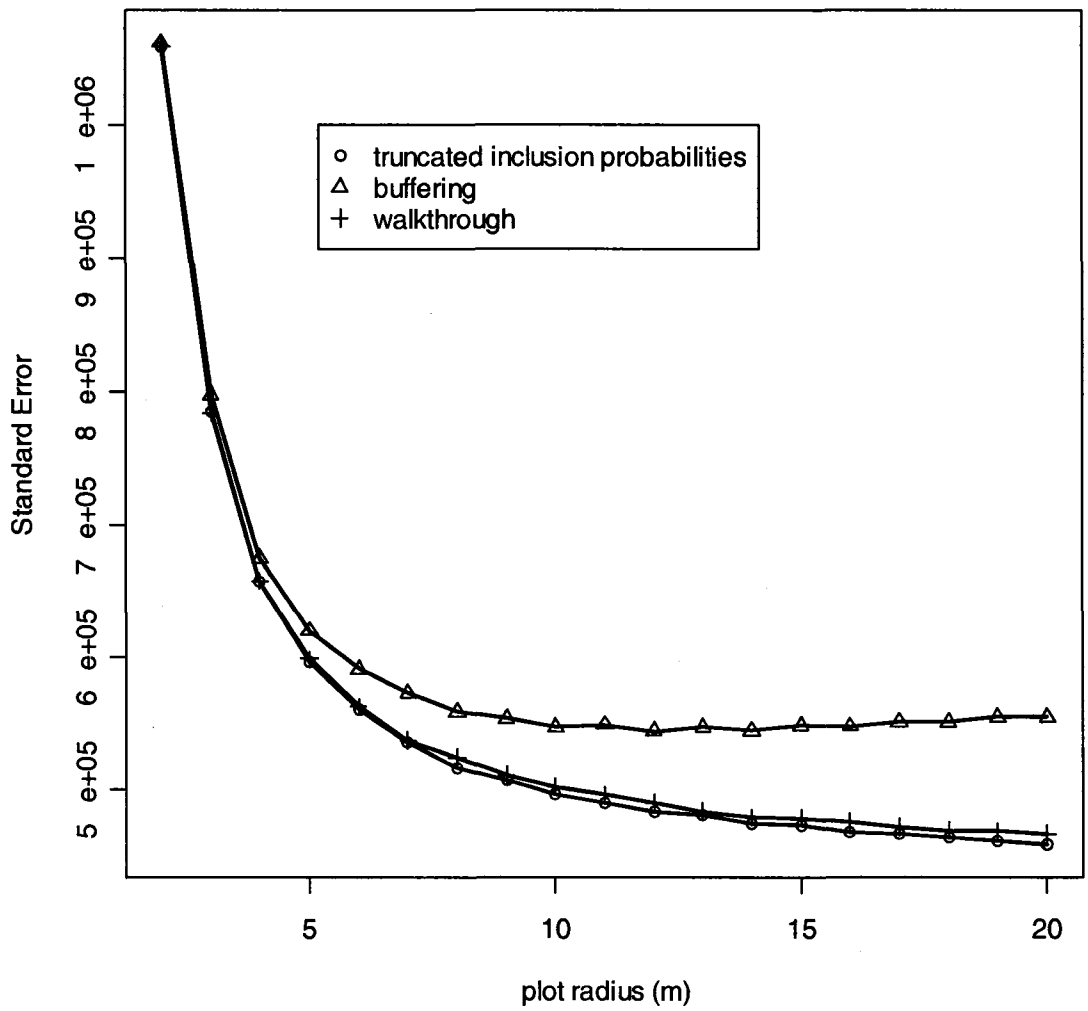


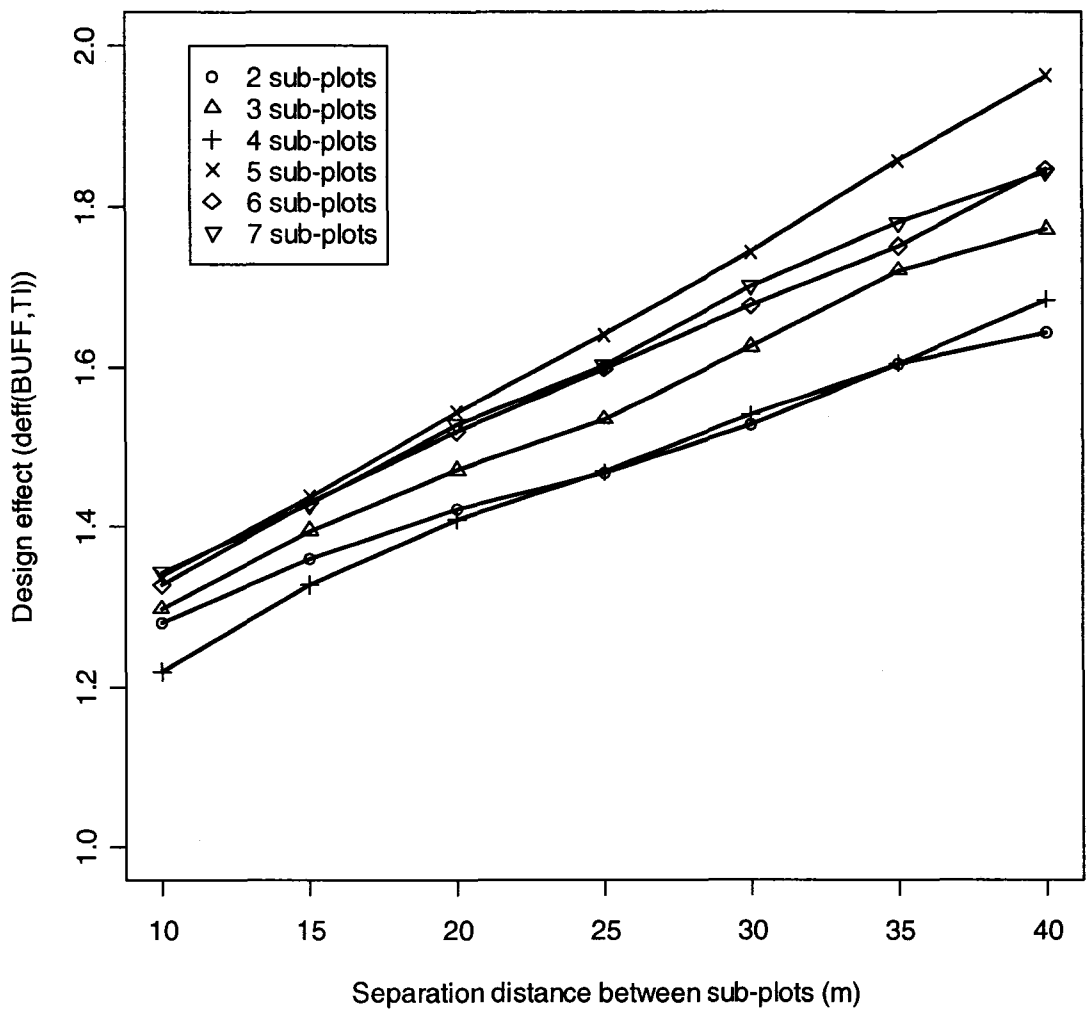


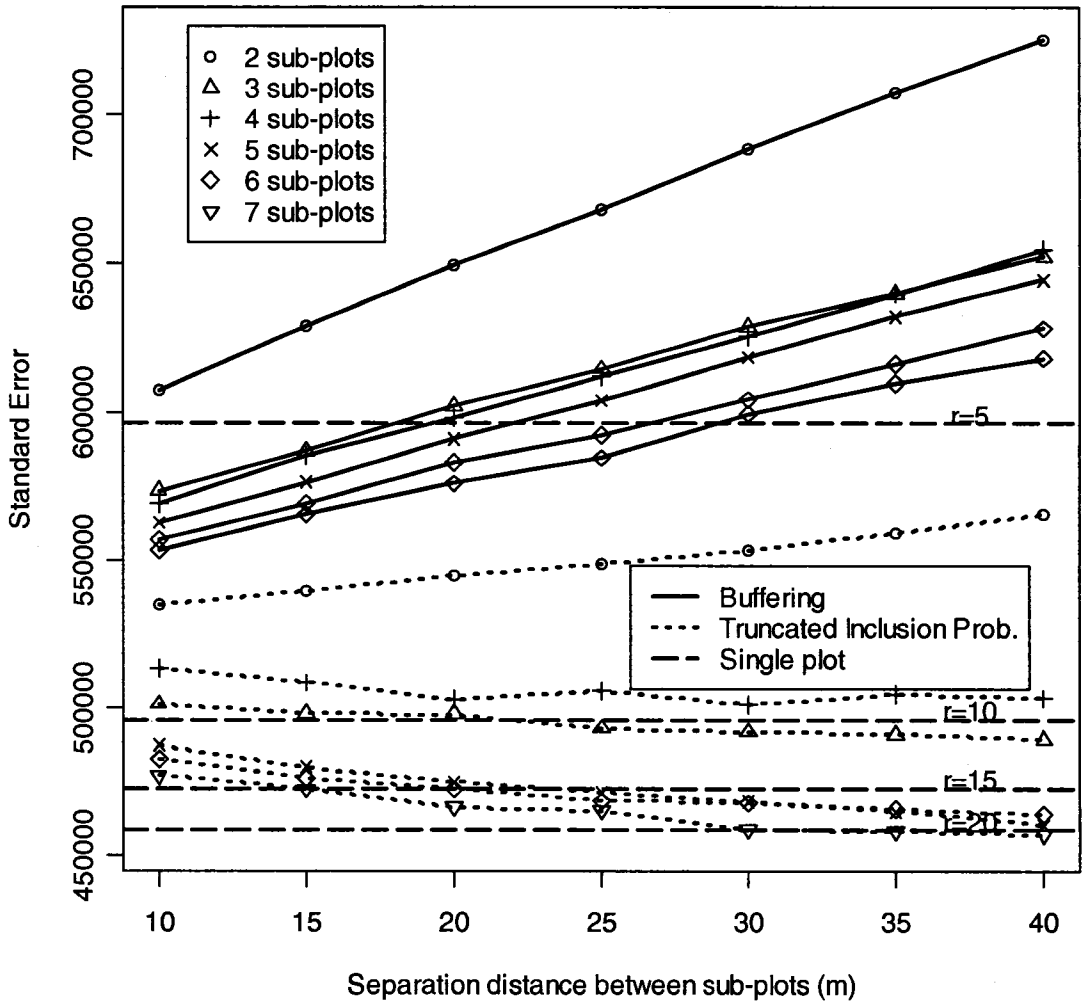


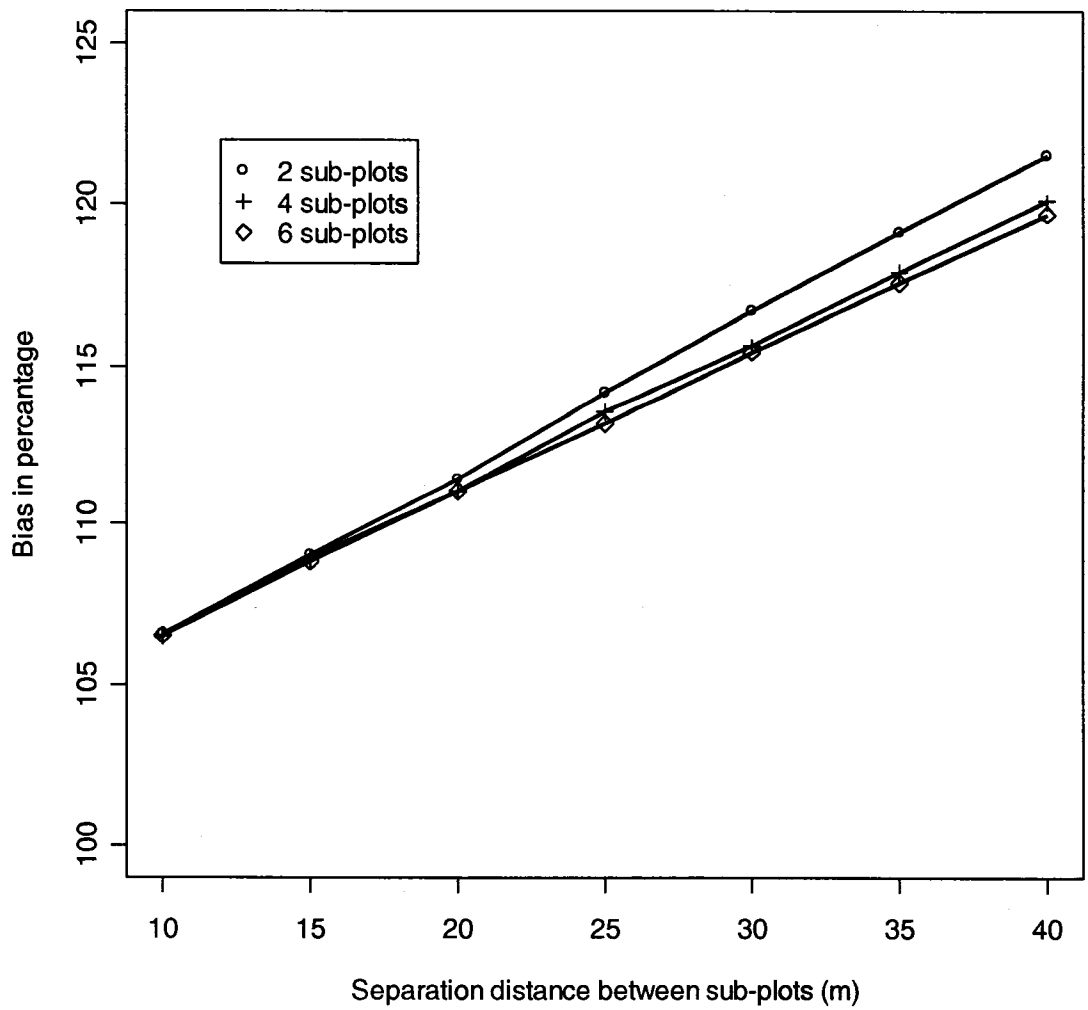












Chapter 2

WOBBLE-PLOT SAMPLING: A TECHNIQUE FOR ESTIMATING SMALL-SCALE SPATIAL VARIABILITY IN ENVIRONMENTAL SURVEYS

Abstract

Geostatistical methods, such as kriging, have become commonly used tools in environmental surveys. In order to perform kriging, a variogram model must be fitted to data collected at various locations throughout the study area. While the production of maps and kriging based estimates is of increasing importance, the design of many environmental surveys is still dictated primarily by more traditional sample survey needs for the data (i.e., the estimation of population totals). A disadvantage of this approach is that a sampling strategy that is appropriate for the sample survey objectives of an environmental survey is often inadequate for modeling spatial correlations because no data are available at "short" lag distances. However, by taking an integrated approach it is possible to meet the data requirements for both the sample survey and spatial modeling objectives. To meet these objectives a new method, referred to as wobble-plot sampling, is introduced. Wobble-plot sampling allows for the accurate estimation of the variogram because it provides an infinite number of observations at very short lag distances. The method is introduced and a simulation study is performed. For the population studied, the variograms derived using the wobble-plot sampling technique were more precise and less biased than the variograms derived from a simple random sample of points that was roughly six times larger. The result of the simulation study also illustrate that a substantial bias can occur in both the nugget and range parameters of the variogram when data at short lag distances are not available.

Key Words: Variogram, lag distance, kriging, spatial modeling

2.1 Introduction

Many environmental surveys collect and analyze data to meet a very broad range of analytical objectives. For example, data from a forest survey are often used primarily to generate descriptive statistics (i.e., the estimation of population means and totals using traditional sample survey techniques such those found in Cochran (1977)). At the same time, data collected at these same locations may be used as part of a monitoring system to detect the spread and severity of a pathogen or the dispersion of pollutants from a point source. Another use of the data is to develop maps illustrating the spatial extent of the resource. Thus, in addition to the usual practical constraints, such as cost, a well designed natural resource survey should simultaneously consider the data collection needs associated with the sample survey, monitoring, and spatial modeling components of the survey.

There are a number of conceptually different ways to view a survey of a population that is spread over an area. However, given the different analytical objectives, it is most convenient to assume that we are trying to infer properties of a continuous surface, $Z(s)$, which we will do by selecting points from the domain of the function $Z()$. The term continuous population sampling will be used to describe sampling strategies that rely on observing a response function $Z(s)$ at point locations indexed by s. There are two different design issues involved in such a survey. The first is to determine how to select each location or point, denoted by s, within the domain. Using the terminology of Stevens and Urquart (2000), this will be referred to as the *sampling design*. Given these sample points, the other design issue is determining what to measure or observe at each point s in order to calculate $Z(s)$. The selection of the appropriate response at each point is referred to as the *response design*. In some situations, the response can be determined by a measurement taken at the point (e.g., soil temperature, wind speed). However, other analyses require the measurement of objects or attributes collected either over an area about the point (e.g., a quarat), or using a size and distance rule to select objects in the vicinity of the point. Examples of the latter include variable radius plot sampling (Husch et al. 1983) and perpendicular distance sampling (Williams and Gove 2003). For these attributes, the response design must consider what is the appropriate plot size or distance rule with which

to collect data about the point. The response, $Z(s)$, is then used to make inferences about the underlying population of interest. This study will only be concerned with the estimation of discrete attributes, such as trees, rocks, and coarse woody debris.

Many of the design aspects for a large-scale survey are still dictated primarily by the sample survey objectives of the survey, with the spatial modeling portion of the design receiving less attention. The concern is that both the sampling and response design that meet the sample survey objectives are not only sub-optimal for some spatial modeling applications, such as kriging, but they may actually prevent valid inference. For example, if one the goals of the survey is to collect data for spatial prediction (i.e., kriging), the data will be used to first estimate a variogram. However, if no two sample locations $s_j, s_{j'}$ are within a distance $\epsilon > 0$ of each other, then it will be nearly impossible to precisely estimate the parameters of the variogram at distances less than ϵ. Given these problems, it would seem that the only viable solution would be to collect additional, or auxiliary data, whose primary use was to estimate the parameters that describe small-scale spatial variability. Examples of this approach are given by Barth and Mason (1984), Atteia et al. (1994), Ritter and LeeCaster (2005). However, by taking an integrated view of the response design, it may often be possible to meet all of the design criteria without the collection of additional data. Unlike other studies related to optimal designs for spatial modeling, such as McBratney et al. (1981), Olea (1984), Russo (1984), Warrick and Myers (1987), and Müller and Zimmerman (1999), the focus of this study will be on the design of the spatial modeling component of an inventory within the constraints imposed by the sample survey requirements of inventory. In this setting, a technique (or trick), referred to as wobble-plot sampling, is introduced that allows for very precise estimation of the parameters associated with small-scale spatial variability. The objectives of this paper are:

- Describe the design constraints imposed by the sample survey design and their impact on spatial modeling objectives .
- Introduce wobble-plot sampling and illustrate the resulting improvement in the variogram.

- Compare the performance of wobble-plot sampling with two other designs using a simulation study on a small forest population.

2.2 Data Description

A small data set is used for the purpose of illustration throughout this study. The data were collected in central New Jersey in 1985 on an area of 5.37 ha, which was extracted from the center of a larger forest stand. Two distinctly different forest types are represented, these being adjacent hardwood stands of roughly equal area containing a total of 4,744 trees. The data set relates to two stands of different structure, one comprised of mature hardwoods of saw-log size and the other of hardwood regeneration. The two stands differed in structure and are separated by a dirt road that divides the population roughly in half and has partially grown over. Conditions within each stand are fairly heterogeneous with areas of visibly different intensity within and between the stands. There is no apparent edge effect along the boundary of the population and none is expected because the data were collected from the interior of a larger stand. Figure 1 depicts the structure of the data set. The diameter of each circle represents the relative size of each tree. Figure 2 shows the response surface for a sampling design where points are uniformly distributed across the area with equal probability and a response design where a circular fixed-area plot of radius $r = 7.5m$ is used to select trees. This surface was created by calculating the response across a very fine grid that covered the population. The grid spacing was $0.7m$, which yielded a set of 108241 responses values $Z(s)$.

2.3 Literature Review

Barnes (1989), Cox et al. (1997), Gregoire (1998), and Gotway and Young (2002) provide limited discussions of concurrent design issues for both sample survey and spatial modeling applications. A rather lengthy review of both issues is provided because it is felt that some widely held results are often misleading when applied to environmental sampling. A summary of the specifics associated with this application is provided at the end of this section.

2.4 Sample survey constraints

The first constraint imposed is that publicly funded surveys should be based on a probability sampling design because such designs ensure objectivity (Smith 1994, Overton and Stehman 1995, Olsen and Schreuder 1997). Design-based inference is also preferable. Probabilistic sampling requires the construction of a frame. However, given the cost of collecting data and the limited information for constructing a sample frame of individual units (e.g., trees, grasses, woody debris), the response design uses some form of cluster sample. The survey design may also be temporally constrained because the survey is required to provide estimates of trend over multiple years, or even decades, as well of current status. A survey is also likely to be spatially constrained in the sense that the design must be sufficiently flexible to adequately categorize both large areas and smaller subpopulations, many of which will not be specified at the design stage. These constraints restrict the use of many of the techniques employed to improve statistical efficiency, such as the extensive use of stratification and unequal probability sampling. These constraints have led numerous authors to suggest that simple designs based on cluster sampling be used for environmental surveys (Olsen and Schreuder 1997). Overton and Stehman (1995) give the following recommendations

- Avoid variable probability sampling when possible.
- Limit stratification to major partitions of the universe, and preferably to classes in which there is little interest in combining strata.
- Select a design for which pairwise inclusion probabilities, in the sense of Cordy (1993), (π_{ij}) are, or can be assumed to be, uniformly close to $\pi_i\pi_j$, indicating near independence.

To address these recommendations, the sampling design portion of the survey often uses plot locations that are regularly distributed so the sample has good spatial *balance*. The most common solution of these is a systematic grid of plot locations. The terms *lattice data* and *centered symmetric sample* are also used. One reason for the use of systematic

grid samples is the efficiency associated with the ease of locating the plots. In the last decade, there has been a substantial interest in sample designs that have fewer drawbacks than those associated with systematic grids. An overview of these alternative designs is provided by Stevens and Olsen (2004). One of the primary reasons for the use of sample designs where the points are regularly distributed across the domain is that these designs tend to be nearly optimal for estimating population means and totals for populations with a variety of spatial patterns (Cochran 1946, Bellhouse 1977, Iachan 1985). The term *systematic sample* will be used to describe the broad class of sampling designs where the points $s_1, s_2, ...s_m$ are spread across A with some degree of spatial regularity.

Many attributes of an environmental survey require measurements taken over some "neighborhood" or plot whose location is determined by the location of the point s. Following the reasoning laid out in traditional finite population sampling references, it is commonly thought that gains in efficiency are achieved through the use of open-cluster plots, which are ground plots that comprise a systematically or randomly arranged groups of small plots or points. De Vries (1986 p. 167), Urban (2002), and Labau et al. (2005) give examples of the open-cluster plots used in various monitoring program in the US and Europe. The premise of sampling with open-cluster plots is that an efficient sampling strategy can be obtained by partitioning the population in such a way that each cluster is as similar as possible to the other clusters in the population. This is equivalent to making the within-cluster variability as large as possible so the between-cluster variability is minimized. It is thought that in most natural resource populations, this condition is often satisfied by using a systematic sample to select primary sampling units, in which the secondary sampling units are systematically arranged within each primary sampling unit as well (Thompson 2002, pp. 134-5). In this setting, a sampling strategy based on widely spaced open-cluster plots will generally have a smaller variance than if the same area were to be sampled by a single plot of equal area because the open-cluster plot will increase the within-cluster variability.

While this result is true in traditional finite population sampling, where samples are drawn from an area frame, Williams and Reich (2005) show that this is generally not

the case in the continuous population setting when a design-unbiased, or nearly design-unbiased estimator is required. They identify four factors that affect the choice of response design, tested a number of different cluster plots designs, and found that a reasonably large single plot is often the most efficient and in many applications the only practical solution for generating descriptive statistics, particularly when the target population is heavily fragmented. The underlying problem of open-cluster plots is that it is impractical to perform the necessary field work to maintain the design-unbiasedness of an estimator when sampling along the edge of the population with an open-cluster plot (Valentine et al. 2005, Williams and Reich 2005).

Two different approaches are used to describe response designs in the continuous population setting. The first is the *point-centered approach*, where the response is a function of the objects measured in some neighborhood about the point (e.g., Campbell 1993, Stevens and Urquhart 2000). The alternative is the *object-centered* approach, which is often used to develop variable radius plot-, line intersect-, and areal sampling in general (see Husch et al. 1982, pp. 224-225). Using circular fixed-area plots as an example, the fundamental idea of the object-centered approach is that visiting a sample point s and selecting elements on a plot of radius r is equivalent to an imaginary circle of radius r being drawn about each object. Each object is selected into the sample any time a sample point falls within the circle, as illustrated in Figure 3. The object-centered approach will be used exclusively to develop the methodology presented in this study.

For each sampling technique, let the population cover an area denoted by A, assuming A is a flat plane with area $|A|$. For each point (s_j), visited on the ground, the estimator considered here is

$$\hat{Z}_j = \sum_{i=1}^{n_j} |A| \frac{z_i \delta_i}{a}, \qquad [1]$$

where z_i is the attribute of interest for object i selected at point s_j, a is the area of the plot, $\delta_i = 1, 2$ is the walkthrough correction factor (Ducey et al. 2004), and n_j is the number of objects tallied at the point. For this study, $z_i = (\pi d_i^2/4)$.

Mandallaz (1991), Eriksson (1995), Williams (2001), and Valentine et al. (2001) discuss methods of deriving the properties of this estimator by treating $\hat{Z}$ as a function of the

sampling design. Thus, the population of interest is the set of all points in A. Using this approach the response surface is a 3-dimensional map of every possible sample outcome over A. The response surface can be constructed across the population using the following algorithm:

- For every object i, place the plot about the object.
- Use the plot boundary to construct a 3-dimensional solid of height $z_i\delta(s)_i|A|/a$, where A is the area of the plot and $\delta(s) = 1, 2$ is the walkthrough correction factor.
- Define the height of the surface at any coordinate (s) as the sum of the heights of the overlapping solids defined for each object that would be selected, i.e. $Z(s) = \sum_{i=1}^{n} z_i\delta_i(s)|A|/a$.

An example of the response surface for New Jersey data set is given in Figure 2.

Thus, in the continuous population setting, the sample design that meets the constraints identified by Overton and Stehman (1995), is an equal probability sample of points that is spread across A in such a way that good spatial balance is achieved. Generally, the only feasible response design is one in which a single point, line, plot, or size and distance rule is used at each point. Finally, the walkthrough technique (Ducey et al. 2004) is the simplest technique for correcting for the differential inclusion probabilities associated with sampling population elements that fall near the boundary.

2.4.1 Literature Review on Spatial Modeling

There are many different applications where inferences are based on some type of spatial model. For example, forestry data are often used for growth and yield studies where individual trees on a plot are "grown" forward in time. The models that predict the rate of growth often require information about the competition between the individual trees, which in turn are related to the size and spacing between an individual tree and its neighbors. Many of these applications can be viewed as marked-point process model (Stoyan et al. (1995), Reich (1997a, 1997b)). A problem with these applications is that trees beyond the boundary of the plot, which contribute to competition, are missing from

the data. While various techniques exist to adjust for this boundary problem (Stoyan et al. (1995), Williams et al. (2001) and references therin), for this type of application it is obvious that a response design based on single circular fixed-area plot would be superior to a number of smaller plots because large circular plots minimize the edge-to-interior area ratio.

On the other hand, many spatial modeling applications use geostatistical methods for the analysis of spatially correlated data. The model in this setting is a random process $Z(.)$. In order to make inferences, some assumptions must be made about the *spatial scale* of the fluctuations in the process (Cressie 1991, p. 112). This choice of scale is partially defined by the level of aggregation needed to collect meaningful data, with two important factors being the plot size and separation between points. A significant challenge is that many ecological processes take place over short distances, but the process of interest is spatially correlated over longer distances. An example, given by Urban (2002), is that seed dispersion takes place over distances in the 10's of meters, but that the resulting distribution of species can occur at 100's of meters. There is no standard definition for what constitutes large-, medium-, or small-scale spatial scale, with the choice of scale being application dependent. The spatial scale of the data is often partially determined by the available data. For example, forest inventory field data are collected across a wide range of plot sizes, with the most recent response design for the Forest Inventory and Analysis (FIA) program being an open-cluster of four fixed-area plots with radii ranging from $r = 7.3 - 17.9m$. Prior to adoption of this design, portions of the northeaster US were sampled by a single circular fixed-area ground plot with a radius of $r = 16m$. The choice of the available remotely sensed data can also determine the spatial scale. For example, Landsat thematic mapper (TM) data have a resolution of approximately 30 m. The problem with many of these data sources is that many ecological phenomena vary at smaller scales, so it is often desirable to model processes at a much smaller scale (Reich et al. 1997a, 1997b, 2004). The spatial scale of the data derived from such remote sensors are often modified to meet specific analytical needs (Joy et al. 2003). For example, each TM pixel can be sub-divided into a 3 by 3 matrix of 10 m pixels and a moving window averaging technique used to derive a new map with a 10 m resolution.

To help define the spatial modeling problem of interest, the basic outline and notation of Müller and Zimmerman (1999) will be used along with examples based on the data set described above. Assume that data are collected at locations within some spatial domain $A \subset R^2$. For a sample of points $s_1, s_2, ...s_m \in A$ (Figure 4), the analysis is carried out using the following six-step process, which is:

1. Propose a model of the large-scale spatial patterns in the data plus an intrinsically stationary error model that describes the small-scale spatial relationships within the data. This model is given by

$$Z(s) = \mu(s) + \epsilon(s), \qquad [1]$$

where $s \in A$ is a sample location within the domain of interest. For this study it is assumed that $E[Z(s)] = \mu(s) = x(s)\beta'$ is a linear function with unknown parameter vector β that describes the large-scale spatial variation, though other methods are available for modeling this component. The assumed model for the small-scale error is such that $E[\epsilon(s)] = 0$ and $var[\epsilon(s+h) - \epsilon(s)]$ depends only on the distance h.

2. Estimate the β parameter for the model describing large-scale variation. This is depicted in Figure 5. The difference between the observed and modeled values at each point s_j is used to determine the residuals for use in estimating the stationary error process.

3. Use the $m(m-1)/2$ paired distances to produce a plot of the error process, which is given by

$$\gamma(h) = \frac{1}{2} Var[\epsilon(s+h) - \epsilon(s)],$$

for every $s, s+h \in A$. The variogram cloud is produced by graphing

$$\hat{\gamma}(h) = \frac{1}{M(h)} \sum_{|s_{j'} - s_j| \in S(h)} [Z(s_{j'}) - Z(s_j)]^2,$$

where $M(h)$is the number of pairs of points separated by a distance h. If the data are irregularly spaced, then $M(h)$ is defined to be some neighborhood about the lag distance h for which there are a reasonable number of lag distances. Note that both the variogram and the lag distance are estimated over an arbitrarily defined neighborhood when the data are irregularly spaced.

4. Choose a valid parametric model $\gamma(h;\theta)$ which is compatible with the plot of the variogram.

5. Fit the chosen model to the estimated variogram to estimate the parameters of the model. Figure 6 illustrates the fit of the Gaussian, exponential and spherical model to the sample variogram data.

6. Use Kriging to predict the values of the spatial process at unobserved locations.

The goal of most studies is the prediction of values at unobserved locations (item 6). However, the focus of this study will be on sampling methods for fitting the variogram (items 4 and 5).

A number of parametric models exist for $\gamma(h;\theta)$, with the Gaussian and exponential models being two of the most common. These two models are of interest because they represent the two extremes of spatial models in terms of the *smoothness*, with the exponential being associated with a random field which is zero times differentiable, and the Gaussian describing an infinitely differentiable random field. The properties of a variogram model are usually described by the nugget, sill and range. The nugget describes both the micro-scale and measurement error component of the small-scale spatial variability. The sill describes variability between sample locations that are sufficiently far apart as to be spatially uncorrelated. The range is the lag distance at which sample locations are no longer spatially correlated, i.e., the distance at which the sill is achieved.

There are two modeling applications in the outline given above, with the first being the estimation of the large-scale trend in the data. Of the possible sampling designs, a systematic sample tends to provide a good balance between efficiency and robustness when estimating the large-scale trend (Cox et al. (1997)). For example, Ylvisaker (1975) and Stein (1995, 1999, Chapter 5) discuss the asymptotic behavior of various best linear predictors (BLPs) used in conjunction with grid data. Results are quite varied with Stein (1999) showing that centered symmetric sampling being only slightly less efficient than any other sampling design when the random field is isotropic or nearly isotropic [1] When the

[1] A random field is isotropic when there is no reason to distinguish one direction from another (i.e., directionally invariant).

assumption of isotropy does not hold, BLPs based on centric symmetric sampling can be badly suboptimal (Ylvisaker (1975)). While these results suggest that the use of grid data may be reasonable for some applications, it should be noted that these results are based on fixed-domain or in-fill asymptotics, where in-fill asymptotics considers the addition of more and more observations within the fixed domain A. Thus, given the small sample size, it seems unreasonable to assume that the asymptotic behavior of a BLP would be valid for most environmental surveys.

A disadvantage of systematic designs is that they can be very poor for estimating the small-scale variability component. The problem is that there are few if any points with "short" between-sample or lag distances. For example, if a square grid is used, where the spacing between grid locations is h units, the set of distances between points is $h, \sqrt{2}h, 2h, ...$ These sampling intervals specify the minimum *grain* in the data that can be used for estimating spatial relationships between locations. The concern is that determining the behavior of $\gamma(h)$ is most critical for small h (Davis and Borgman 1978, 1982, Warrick and Myers 1987). Stein (1988) studied the behaviour of the kriging predictor when the covariance function is incorrectly specified. He shows that as long as the behavior of the any covariance function is "similar" to that of the true covariance function near the origin, the resulting kriging predictor will have the same asymptotic efficiency as the predictor based on the true covariance function. Thus, a major concern in environmental surveys is that a systematic sampling design provides no data for the most critical values of h. This is particularly problematic in many forest and environmental surveys because numerous studies have found that points at distances greater than about $10 - 300m$ are often spatially uncorrelated (Reich et al. 1997, Wallerman 2003), but the spacing between points often ranges from no less than 20-40 m (i.e., 1 to 2 chains, where a chain is defined as 66 ft) to in excess of many kilometers. Thus, a reasonable design for generating descriptive statistics and modeling large-scale spatial variability lacks sufficient information for making valid inferences about the small-scale spatial variability.

The traditional solution to the problem of modeling both large- and small-scale spatial variability is to employ sampling designs where a grid of points is used for the estimation of large-scale variability and clusters of closely spaced points are used to augment the grid.

Examples include Olea (1984), Barth and Mason (1984), West et al. (1993), Atteia et al. (1994), Urban (2002), Ritter and LeeCaster (2005). However, as noted above, such designs are at odds with the sample survey constraints.

2.4.2 Summary of Literature

In order to meet the guidelines of Stevens and Urquhart (2000), the sampling design of an environmental survey should use some form of systematic sample with a response design based on a single, relatively large plot or point, rather than a cluster of smaller plots or points. The reason for such a restrictive response design is the difficulty of sampling along the population boundary. The reasoning for relying on design-based inferences is that while they are more restrictive and elementary, they also tend to be more robust across multiple variables of interest and across time (Cox et al. (1997)). Within this sampling and response design framework, it is reasonable to expect that coefficients of the large-scale spatial variation component, given be $\mu(s) = x(s)\beta'$, can also be reasonably well estimated. However, without the collection of additional data, it is unlikely that the small-scale component can be estimated.

2.5 Wobble-plot sampling: a technique for the estimation of small-scale variability

The advantage of model-based inference is that it offers more latitude in the choice of analytical methods. For example, the sample need not be probabilistic and the scale at which the random process $Z(s)$ is modeled can be varied. The estimation of small-scale variability usually requires the collection of $M(h)$ closely spaced points where h is "small". However, rather than collect an auxiliary sample of points for estimating the small-scale variability, we introduce a new approach, referred to as *wobble-plot sampling*, which exploits a change of scale to extract an infinite set of points that are suitable for modeling small-scale spatial variability.

An example is used to describe the technique. Consider the small forest population of trees (Figure 7a) and assume that the response design used for the sample survey objectives is a fixed-area plot with radius $r = 15m$. Using the plot-centered approach and a scale of

$30m(2r)$, six of the seven trees are included in the sample and the response at $s = (0,0)$ is

$$Z_{30}(s) = \frac{|A|}{\pi * 15^2} \sum_{i=1}^{6} z_i.$$

If the scale is changed to $15m$ only two of the trees would be selected and at $s = (0,0)$

$$Z_{15}(s) = \frac{|A|}{\pi * 7.5^2} \sum_{i=1}^{2} z_i.$$

However, when the object-centered approach is used to view the locations $s + h$ over which each tree would be selected, it can be seen (Figure 7b) that only those trees on the original $r = 15m$ plot would ever be selected by any point that fell within a radius of $r = 7.5m$ of the center point of the plot.

Thus, the set of trees that would be sampled on the $r = 15m$ plot can be used to determine the exact response surface $(Z(s))$ for a plot of $r = 7.5m$ over a circular area with $r = 7.5m$. By changing the scale between the sample survey and spatial modeling portions of the survey it is possible to generate an infinite number of sample locations from which the small-scale spatial variability can be modeled. This is depicted in Figure 8.

Another way to describe wobble-plot sampling is that it is equivalent to placing an infinite number of smaller plots within a larger single larger plot. The only restriction placed on the location of the smaller plots is that they must be completely encompassed by the large plot. For the example given in this study, the spatial modeling plot is one half the radius of the sample survey plot, but there is no particular reason why the ratio between the sample survey and spatial modeling plots must be 2.

Fitting a variogram model to the wobble-plot data poses some difficulties. The first is that there are an infinite number of points available for estimating the variogram, but commonly available software is designed to accomodate a relatively small number of observations. The second is to note that, if $S(h)$ is the set of wobble-plot points separated by a distance h, the area or volume of this set decreases with increasing h. For example, note that $S(h = 0)$ is the set of all points within the wobble-plot and $S(h = 2r)$ is the set of points that fall on the circumference of this plot. The third problem is that the wobble-plot data alone may be insufficient for fitting the variogram because the range parameter is likely to exceed r.

The goal of this study is to introduce the concept of wobble-plot sampling, with theoretical results related to asymptotic properties, optimum fitting approaches, and determining the optimal location of additional auxiliary points covered in later publications. Thus, the methodology used to fit the variogram to the wobble-plot data was fairly crude. The following approach was used;

- For each sample point $s_j, j = 1, ...M$ determine the response $Z(s)$ for every grid point on the response surface that would fall within the wobble-plot distance.
- The square grid produced a set of lag distances $h, \sqrt{2}h, 2h, ...r$, where $h = 0.7m$ in this study. For each lag distance, the sample variogram was calculated using

$$\hat{\gamma}(h) = \frac{1}{2M(h)} \sum_{|s_{j'} - s_j| = h} [Z(s_{j'}) - Z(s_j)]^2,$$

  where $M(h)$is the number of pairs of points separated by a distance h.
- The responses at each point s_j was used to calculate $\hat{\gamma}(h)$ at lag distances greater than $r = 15m$. These data are irregularly spaced and required grouping the data into bins based on their lag distance. The width of each bin was allowed to vary so that the number of points in each bin was equal to the size of the smallest wobble-plot sampling bin. The lag distance used for each of these bins was the mean of the lag distances in the bin.
- Gaussian and exponential variogram model were fitted to the variogram cloud.

An example of the variogram cloud derived from a sample of $M = 36$ wobble-plot plots is given in Figure 9.

2.6 Simulation Study

A simulation study was used to compare the performance of wobble-plot sampling against two more traditional alternatives.

For all three approaches, it was assumed the goal is to model the response at a $15m$ resolution. Thus, the goal of the spatial modeling is to model the response surface given

in Figure 2. For wobble-plot sampling, it is assumed that the sample survey portion of the inventory draws samples from the population using circular fixed-area plots with radius $r = 15m$. Where the methods differed was in the distribution and number of points. For the first approach, that will be referred to as RAND, a sample of $N = 200$ random locations was drawn and the variogram was fitted to the data. By using a large sample with no restrictions on the spacing of points, sufficient information is available to estimate both the small- and large-scale spatial variability. The problem with this approach is that a sample of this size and spatial arrangement would rarely be available for a population of this size, unless the data were collected as part of a special study or specifically for the purpose spatial modeling. The second approach, referred to as SYS, relied on a sample of $N = 36$ points which were drawn using a Strauss distribution. Samples drawn from a Strauss distribution are more regularly distributed than a uniform distribution, with the parameters of the distribution chosen so that no two points were separated by a distance of less than $25m$. While this sample is a small fraction of the size used in the first approach, it is a more realistic representation of the sample sizes that would be used for this sized population. In fact, it would be unusual for many forest surveys to use more than 15 to 20 points for a population of this size and the spacing between points would be rectangular. The third approach, referred to as WP, relied on the same sample of points as the SYS approach, but wobble-plot sampling was used to ensure that reasonable inferences could be made about spatial variation at distances less than $25m$. Given the grid spacing of $h = 0.7m$, roughly 1.8 million paired lags were available for estimating the small-scale spatial variability.

For all three approaches, a sample of points was drawn and the data were fitted using the algorithms outlined above. The large-scale spatial variability was modeled using a third degree polynomial. Once the residuals were calculated, a Gaussian and exponential variogram were used to model the small-scale spatial variability. A spherical variogram was also fitted, but the results were not of sufficient interest to warrant presentation. Fifteen iterations of this process were performed and the resulting variograms were compared. The performance between the three designs was performed by comparing the nugget, sill, and range parameters, as well as a visual comparison of the 15 variograms.

One reason for using such a small number of iterations is that it was not possible to automate the model fitting process for the RAND and SYS approaches. The reason being that the fitted form of the models, and their resulting parameter estimates, can be very sensitive to factors such as the number of bins used for the lag distance, the fitting procedure, and the maximum lag distance at which the sample variogram is calculated. Of the three variogram models considered, the exponential model was by far the most sensitive and it was not uncommon for this model to produce estimates of the range parameter that were nonsensical (Figure 10). These "outliers" were so large that it would have been very difficult to make a meaningful comparison between the different techniques. Thus, many of the samples required that the method of fitting the model be specifically tailored to the sample. While this significantly affected the results of the simulation study, it should be noted that such subjective manipulations of the data are the norm in any practical application. In contrast, the fitting algorithm for wobble-plot sampling was never altered.

2.7 Results

Figures 11-13 summarize the nugget, sill, and range parameter estimates for each of the three methods. The influence of the sample design on the nugget parameter is best illustrated by observing that the nugget parameter was 0 for every sample and model under the SYS sampling design. When comparing the two variogram models, the nugget effect is most often larger for the Gaussian variogram. This is particularly true under the wobble-plot sampling design, where the nugget effect estimated by the Gaussian model is always at least 4 times larger than the exponential model for every iteration. The sill parameter estimates derived from the WP sampling design were the least variable and substantial differences exist between the estimates derived from the SYS sampling design and other two sampling designs, with the mean of the sill estimates always being the smallest. Another interesting observation is that the sill parameters for both models are almost identical for the RAND sampling design and that the sill parameters are quite different for the WP sampling design. The range parameter for the exponential model under the SYS sampling design is significantly different from both the RAND and WP sampling designs.

Figure 14 provides a visual comparison of the variograms across the three sampling designs. The variograms derived from wobble-plot sampling tend to be the least variable and the general shape is very consistent (Figures 14a and b). On the other hand, the variograms derived from the SYS data (Figures 14c and d) clearly illustrate that the range of the variogram is substantially longer. The exponential variogram for the RAND sampling (Figure 14e) is similar to those derived from the WP design, but the tendency is for the RAND variogram to by more variable. The Gaussian variogram for the RAND sampling design (Figure 14f) illustrates substantial variability in the estimated range, with a number of variogram with either very short or long ranges.

One unanticipated benefit of wobble-plot sampling is its utility in model selection. For example, in this application the wobble-plot data can be used to eliminate the Gaussian variogram as a possible model. This conclusion is made by observing the behavior of the Gaussian model as $h \rightarrow 0$ (Figure 15), where it can be seen that the shape of the Gaussian model fails to match the variogram cloud for small h values. This result was observed for every iteration of the simulation and explains why the estimated nugget effect was always significantly larger for the Gaussian variogram under wobble-plot sampling.

2.8 Discussion

The results for the simulation study for the SYS sampling design suggest the variogram can suffer from substantial biases when there is a lack of sufficient data with small lag distances. However, it should be noted that reasonable estimates of the nugget effect can generally not be obtained with any of the open-cluster plot designs that are currently employed. This is because given the very short range of spatial correlation for many environmental applications, the edges of the sub-plots would need to be overlapping and these plots are of questionable value for traditional sample survey objectives, such as estimating population means and totals. Thus, it is possible that the only benefit to augmenting an existing sample survey plot with additional points would be a slight reduction in the bias in the range parameter.

The simulation study also illustrates that not only are the variograms less variable under wobble-plot sampling, but that the objectivity of the model fitting process can es-

sentially be eliminated. While the selection of parameters for fitting a variogram is a necessary process in almost every spatial modeling applications, the impact of this process on the validity of the inferences will depend on how well the resulting variogram models the spatial correlation in the data. For this application, where the range parameter of the variogram is over-estimated, the resulting map of predicted $Z(s_0)$ values will be "too smooth" and the estimated mean square prediction error will under-estimate the true prediction error. This will result in prediction intervals that rarely achieve their nominal coverage.

The practicality of data collection for wobble-plot sampling is a reasonable question because the mapping of objects on a large plot is often a nontrivial task. However, many monitoring programs that are designed to assess changes over time are already collecting these data. For example, one objective of forest inventories is the desire to divide the change in forest resource into specific categories, with these categories being:

- survivor growth, which is the change in size for trees that existed at both the initial and terminal visits to the plot,
- ingrowth, which is the change in the forest due to new trees,
- mortality, which is comprised of the trees that died between the initial and terminal measurement
- removals, which is comprised of those tree removed from the forest.

In order to perform this decomposition, it must be possible to assign each sample tree to one of these categories, so each tree can be identifiable at each visit to the plot. The easiest way to identify each tree, without actually physically marking the trees, is to record the location of each tree in polar coordinates. This task can also be further simplified through the use of laser devices (Kalliovirta et al. 2005).

2.9 Summary

Wobble-plot sampling provides a method for the efficient estimation of small-scale spatial variability within the constraints imposed by the sample survey needs of an inventory.

The key to wobble-plot sampling is the selection of a response design that meets the resolution requirements of the spatial modeling component of the inventory. However, it should be noted that it is quite possible that some spatial processes will still require the collection of auxiliary data for efficient estimation of all of the components of variogram. This situation can occur when the range of the variogram falls somewhere between wobble-plot distance and the spacing of the systematic sample points. In this situation, some form of open-cluster plot or even additional grid points may still be necessary. However, note that these auxiliary points can be collected at a substantially reduced cost because auxiliary spatial modeling plots only need to have a plot radius that is smaller than the sample survey plots, plot mapping is not necessary, and these points do not necessarily need to be a probabilistic sample or maintained over time.

One area that this paper has avoided is the issue of optimal design. There have been a number of studies to determine optimum locations for additional sample locations (e.g., Warrick and Myers 1987). It would seem these results would not be appropriate for data collected under wobble-plot sampling. The reason being that most or all of the common parametric models for the variogram have parameters that explain the nugget, sill, and range. The data derived for the center-point of each can be assumed to be sufficient for estimating the sill parameter and the wobble-plot sampling data provide sufficient data for estimation of the nugget parameter and provides some information on the slope of the variogram for small h values. However, there may still be insufficient data to estimate the range parameter. So, at least in the case of isotropic variograms, it would seem that choosing new point locations that are in some sense "optimal" would require determining auxiliary points whose primary function is to facilitate estimation of the range parameter. However, it is felt that no optimal design will exist for a multi-objective survey. The reasoning being that trying to balance the costs and variance reductions of multiple attributes and multiple objectives across both the sampling and response design is likely to be a futile exercise. A further complicating factor is that is it would be nearly impossible to consider the myriad of ways that auxiliary data, in the form of additional points or remotely sensed data, could be used in the estimation process. Thus, we recommend pursuing a design that is robust

and practical over one that is in some sense "optimal". The following is one possible outline for the design of a multi-objective survey that used circular fixed-area plots.

1. Select the desired resolution for the anticipated spatial modeling applications (e.g., mapping forest cover to a 10 m resolution).

2. Select a plot radius that is as large or larger than the desired mapping resolution to meet the sample survey objectives.

3. Utilize a spatially balanced sampling design to meet the sample survey objectives, such as estimating the total abundance of the resource.

4. Extract the wobble-plot data and use the center-point of each wobble-plot to model the large-scale spatial variability.

5. Use all of the wobble-plot data and the center points of the wobble-plot plots to estimate the variogram.

6. Assess the need for collecting auxiliary point data for either modeling the variogram or for spatial prediction.

7. If auxiliary data are collected, assess the utility of these data for additional efficiency gains in the sample survey objectives of the inventory. Nontrivial improvements may be possible through the use of model-assisted estimation techniques.

2.10 Literature Cited

Atteia, O., Dubois, J.-P., and Webster, R. 1994. Geostatistical analysis of soil concentration in the Swiss Jura. Environ. Polution. 86:315-327.

Barnes, R. J. 1989. A partial history of spatial sampling design. Geostatistics Newsletter. 2:10-12.

Barth, D. S., and Mason, B. J. 1984. Soil sampling quality assurance and the importance of an exploratory study, in Environmental Sampling of Hazardous Wastes, Schweitzer, G. E., and Santolucito, J. A. (eds). American Chemistry Society. Washington, DC, pp. 97-104.

Bellhouse, D. R. 1977. Some optimal designs for sampling in two dimensions. Biometrika 64:605-611.

Campbell. K. 1993. Sampling the continuum: The spatial support of environmental data. In Proceedings of the Statistical Computing Section. 11-19. Alexandrea, VA. American Statistical Association.

Cochran, W. G. 1946. Relative accuracy of systematic and stratified samples from a certain class of populations. Annals of Mathematical Statistics. 17.164-177.

Cochran, W. G., 1977. Sampling Techniques. 3^{rd} ed. J. Wiley & Sons, New York.

Cordy, C. 1993. An extension of the Horvitz-Thompson Theorem to point sampling from a continuous universe. Probability and Statistics Letters. 18:353-362.

Cox, D. D., Cox, L. H. and Ensor, K. B. 1997. Spatial sampling and the environment: some issues and directions. Environmental and Ecological Statistics. 5:219-233.

Cressie, N. 1991. Statistics for Spatial Data. J. Wiley & Sons, New York.

Davis, D., and Borgman, L. 1978. Some exact sampling distributions for variogram estimates. J. Int. Assoc. Math. Geol. 11:643-645.

Davis, D., and Borgman, L. 1982. A note on the asymptotic distribution of the sample variogram. J. Int. Assoc. Math. Geol. 14:189-193.

Ducey, M.J., Gove, J.H., and Valentine, H.T. 2004. A walkthrough solution to the boundary overlap problem. For Sci. 50:427-435.

Eriksson, M., 1995. Design based approaches to horizontal-point-sampling. For. Sci. 41:890-907.

Gotway. C. A., and Young, L. J. 2002. Combining incompatible spatial data. JASA 97:632-648.

Husch, B., Miller, C.I., Beers T.W., 1982. Forest Mensuration 3rd edn. J. Wiley and Sons, New York.

Iachan, R. 1985. Plane sampling. Statistics and Probability Letter. 50:151-159.

Joy, S. M, Reich, R. M., and Reynolds, R. T. 2003. A non-parametric, supervised classification of vegitation types on the Kaibab national forest using decision trees. Int. J. Remote Sensing. 9:1835-1852.

Kalliovirta, J., Laasasenaho, J. and Kangas, A. 2005. Evaluation of the laser-relascope. Forest Ecology and management. 204:181-194.

Labau, V. J., Bones, J. T., Kingsley, N. P., Lund, H. G., and Smith, B. W. 2005. A history of the USFS Forest Survey 1900-2003.

Mandallaz, D. 1991. A unified approach to sampling theory for forest inventory based on infinite population and superpopulation models. Chair of Forest Inventory and Planning, Swiss Federal Institute of Technology, Zurich. 256 p.

McBratney, A. B., and R. Webster 1982. How many observations are needed for regional estimations of soil properties? Soil Sci. 135:177-183.

Müller, W. G. 2001. Collecting spatial data: Optimum design of experiments for random fields. 2nd. edn. Springer-Verlag, New York.

Müller, W. G., and Zimmerman, S.L. 1999. Optimal designs for variogram estimation. Environmetrics. 10:23-37.

Olea, R. A., 1984. Sampling design optimization for spatial functions. Mathematical Geology. 16:369-392.

Olsen, A. R., and Schreuder, H. T. 1997. Perspectives on large-scale natural resource surveys where cause-effect is a potential issue. Environ. Ecol. Stat. 4:167-180.

Overton, W. S. and Stehman S. V. 1995. Design implications of anticipated data uses for comprehensive environmental monitoring programmes. Environ. Ecol. Stat. 2:287-303.

Overton, W. S. and Stehman, S. V. 1995. The Horvitz-Thompson theorem as a unifying perspective for probability sampling: With examples from natural resource sampling. The American Statistician. 49(3).261-268.

Reich, R. M., C.D. Bonham, and K. Metzger. 1997a. Modeling small-scale spatial interaction of shortgrass prairie species. Ecological Modeling. 101:163-174.

Reich, R. M., K. L. Metzger, and C. D. Bonham. 1997b. Application of permutation procedures for comparing multi-species point patterns of grassland plants. Grasslands Science. 43:189-195

Reich, R. M., Joy, S. M., and Reynolds, R. T. 2004. Predicting the location of northern goshawk nests: modeling the spatial dependency between nest locations and forest structure. Ecolog. Mod. 176:109-133.

Ripley, B. D., 1981. Spatial Statistics. J. Wiley & Sons, New York

Ritter, K. J., and Leecatser, M. 2005. Spatial enhancement of grids for the estimation of a variogram in near coastal systems. Submitted to Envir. Ecolo. Stat.

Russo, D. 1984. Design of an optimal sampling network for estimating the variogram. Soil. Sci. Soc. Am. J. 48:708-716.

Smith, T. M. F. 1994. Sample surveys 1975-1990; an age of reconciliation. Int. Stat. Rev. 62:5-34.

Stein, M. L. 1995. Locally lattice sampling designs for isotropic random fields. Ann. Statistics. 23:1991-2012.

Stein, M. L. 1998. Asymptoticly efficient prediction of a random field with a misspecified covariance function. Ann. Statistics. 16:55-63.

Stein, M. L. 1999. Interpolation of spatial data: Some theory for Kriging. Springer, NY

Stevens, D. L. Jr., and Urquhart, N. S. 2000. Response designs and support regions in sampling continuous domains. Environmetrics 11-13-41.

Stevens, D. L. Jr., and Olsen, A. R. 2004. Spatially balanced sampling of natural resources. JASA 99:262-278.

Stoyan, D., Kendall, W. S. and Mecke, J. 1995. Stochastic Geometry and Its Applications, 2^{nd} ed. J. Wiley & Sons, New York.

Thompson, S.K., 2002. Sampling, 2^{nd} ed. J. Wiley & Sons, New York.

Urban, D.L. 2002. Strategic monitoring of landscapes for natural resource management. In J.L. Liu and W.W. Taylor (eds.), Integrating landscape ecology into natural resource management. Cambridge University Press (in press).

Valentine, H. T., Gove, J. H., and Gregoire, T. G. 2001. Monte Carlo approaches to sampling forested tracts with lines or points. Can. J. For. Res. 31:1410-1424.

Valentine, H. T., Ducey, M. J., and Gove, J. H. 2005. A Walkabout correction for cluster-plot slop. For. Sci. (in press).

Vries, P.G. de. 1986. Sampling theory for forest inventory. A teach yourself course. Springer-Verlag, New York. 399 p.

Warrick, A. W. and Myers, D. E. 1987. Optimization of sampling locations for variogram calculations. Water Resources Research. 23:496-500.

West, O. R., Siegrist, R. L., Mitchell, T., Pickering, D. A., Muhr, C. A., Greene, D. W., and Jenkins, R. A. 1993. X231-B technology demonstration for in situ treatment of contaminated soils: contaminant characterization of mean radium-226 concentration in surface soil. Oal Ridge National Laboratory Technical Report ORNL TM-12258.

Williams, M. S. 2001. New approach to areal sampling in ecological surveys. For. Ecol. and Manage. 154:11-22.

Williams, M. S., Williams, M. T., and Mowrer, H. T. 2001. A boundary reconstruction technique for circular fixed-area plots in environmental survey. J. Agric. Biol. Environ. Stat. 6:479-494.

Williams, M. S., and Gove, J. H. 2003. Perpendicular distance sampling: Another method of sampling coarse woody debris. Can. J. of For. Res. 33:1564-1579.

Williams, M. S., and Reich, R. M. 2005. Choice of Ground Plots Configurations for Continuous Population Surveys over a Fragmented Landscape. For. Sci. (in review).

Ylvisaker, D 1975. Designs on random fields. In Survey of statistical design and linear models. Ed. J. N. Srivastava, North-Holland, Amsterdam, 593-608.

2.11 Figure Captions

- Figure 1. Perspective of the data set. The size of each circle represents the relative size of each tree.
- Figure 2. Response surface created by sampling every location with a circular fixed-area plot of radius $r = 7.5m$.
- Figure 3. The set of objects selected with a plot of radius r centered at a point (a) is equivalent to the set of objects that would be selected when the point falls within the object-centered plot (b).
- Figure 4. A set of $N = 36$ points selected from the response surface for the purpose of spatial modeling. The point locations (s) where generated by a Strauss distribution.
- Figure 5. Trend surface that models large-scale spatial variability. The surface was derived from the sample points.
- Figure 6. Exponential, Gaussian, and spherical variograms fitted to the 630 lag data points. Note that there are no lag distances less than roughly $30m$ from which to fit the models and that the resulting variograms differ significantly.
- Figure 7. In (a) two concentric plot measuring $r = 15$ and $r = 7.5m$ are depicted. Six and two of the seven trees are selected on the $r = 15$ and $r = 7.5m$ plots, respectively. In (b) the object-centered approach is taken. Note that only those tree that fell within $15m$ of the point would contribute to the response surface derived from $r = 7.5m$ plots.
- Figure 8. The sample values derived from the $N = 36$ points assuming spatial modeling is performed at 15-meter resolution ($r = 7.5m$), but the data for the sample survey objectives were collected using an ($r = 15m$) plot.
- Figure 9. Variogram cloud derived from the wobble-plot method based on a sample of $N = 36$ data points.

- Figure 10. Three variogram models fitted to one SYS sample. Note that the fit of the exponential model is that the range parameter was $587m$, which is roughly 2 time greater than the maximum lag distance for the data set.

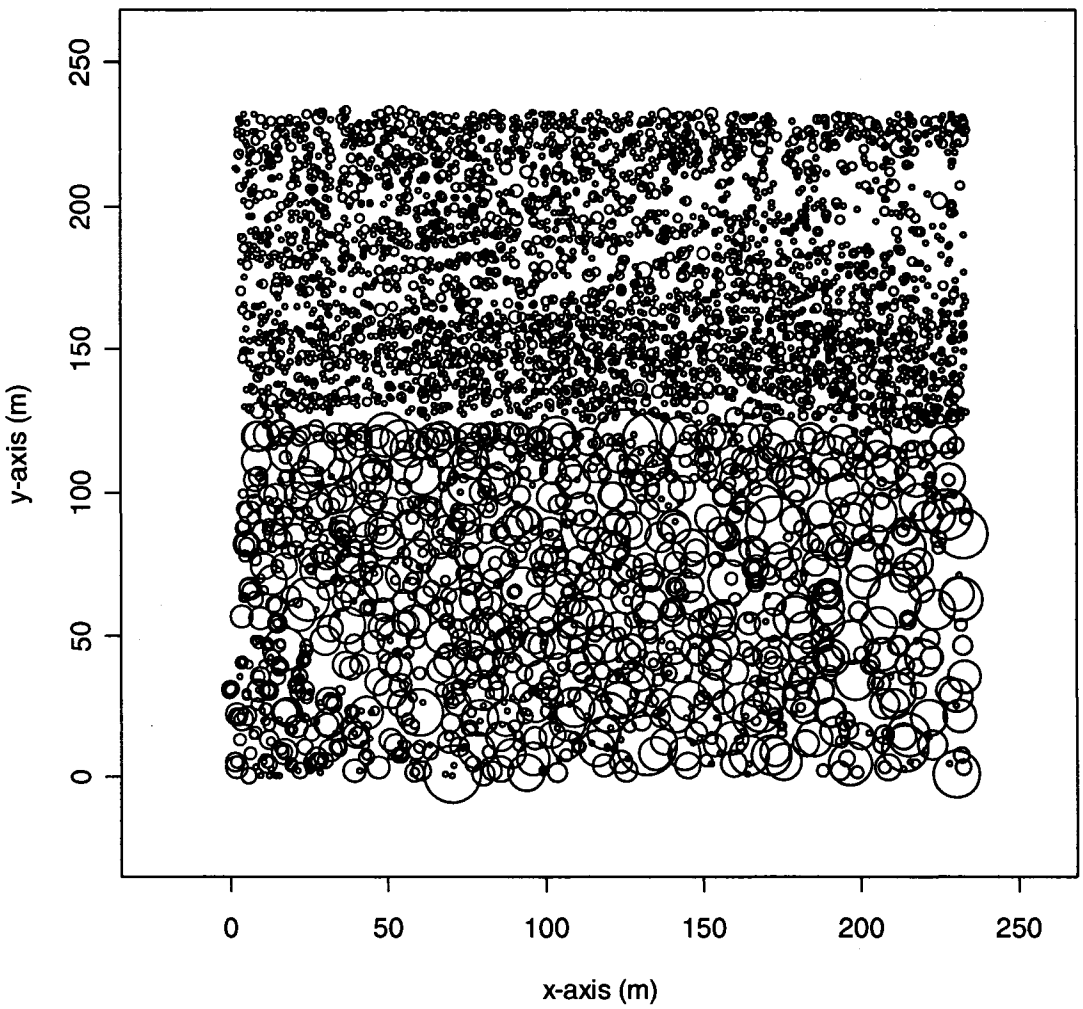
- Figure 11. Summary of the nugget parameter from the simulation study. The sampling design used is indicated by WP, RAND, SYS. Exp and Gau refer to the exponential and Gaussian variogram, respectively.

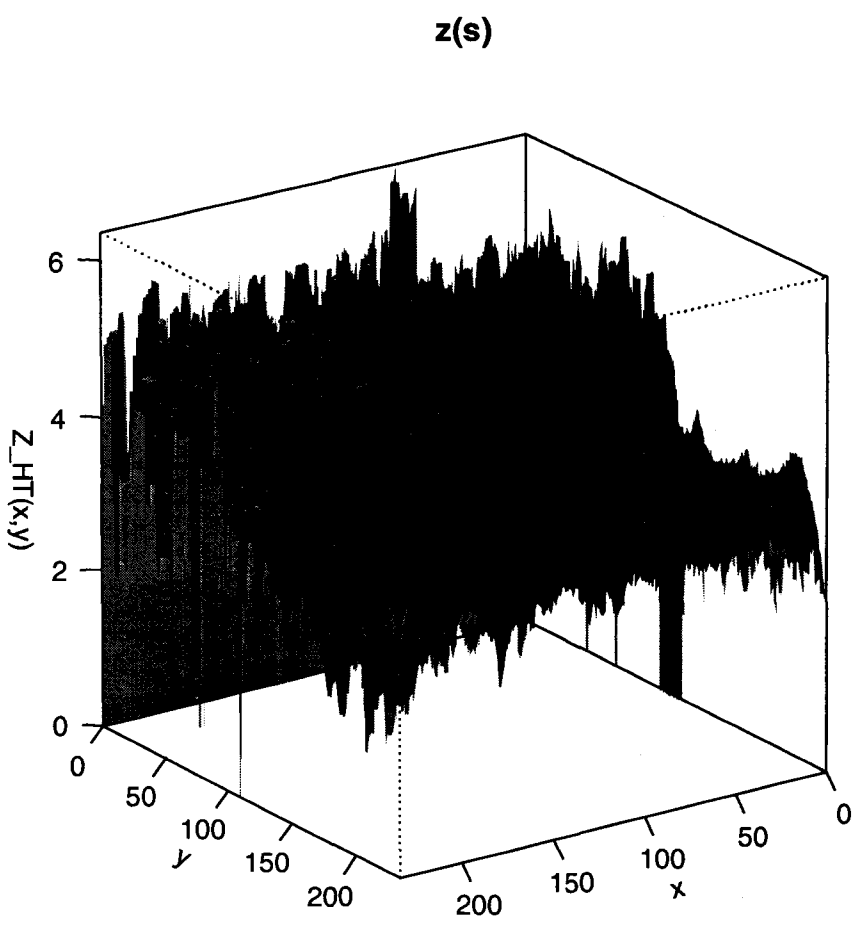
- Figure 12. Summary of the sill parameter from the simulation study. The sampling design used is indicated by WP, RAND, SYS. Exp and Gau refer to the exponential and Gaussian variogram, respectively.

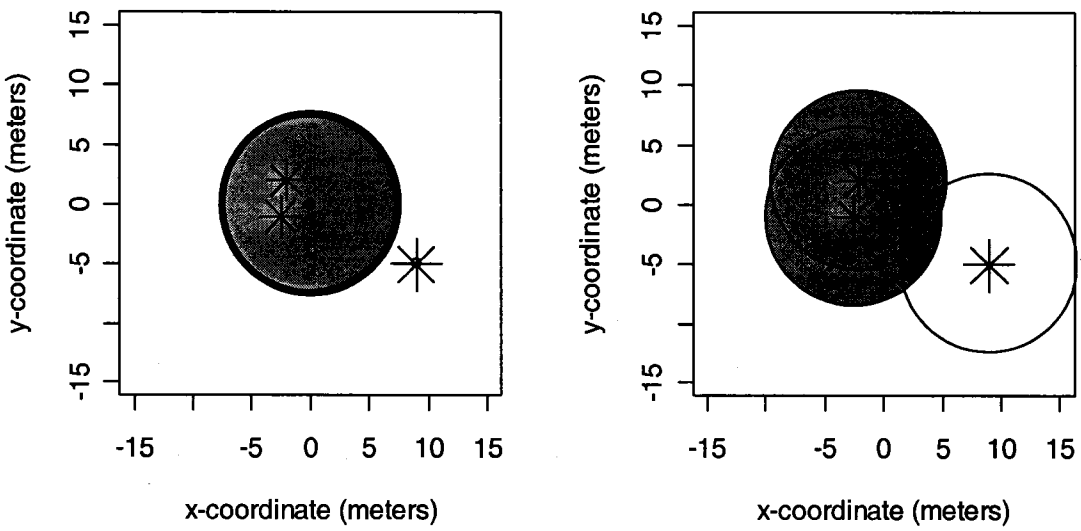
- Figure 13. Summary of the sill parameter from the simulation study. The sampling design used is indicated by WP, RAND, SYS. Exp and Gau refer to the exponential and Gaussian variogram, respectively.

- Figure 14. Variogram models from the simulation study. The exponential variogram results are given on the left-hand side (a,c,e). All graphs on the right-hand side are Gaussian variograms (b,d,f). The top row were fitted using wobble-plot sampling (WP) (a,b). The middle row were fitted using the SYS data (c,d), and the bottom row are the RAND data.

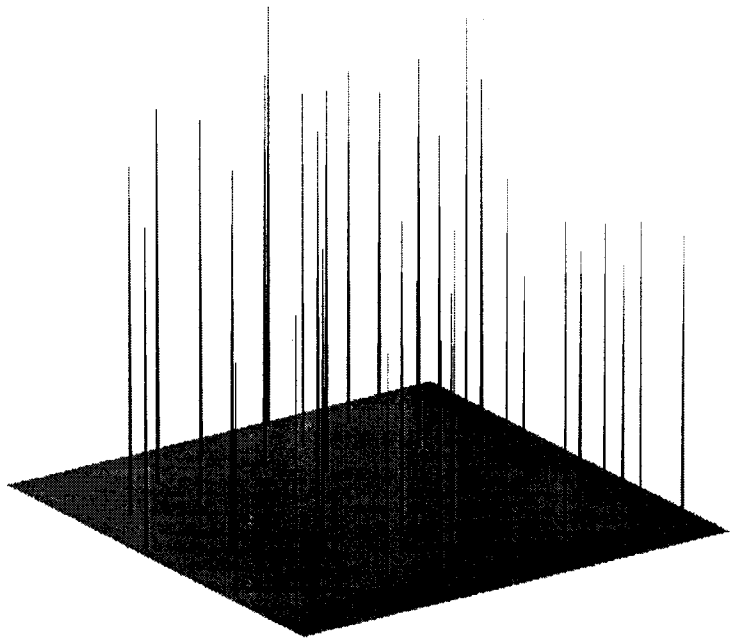
- Figure 15. Gaussian variogram models fitted to the wobble-plot sampling data. The Gaussian model overestimated the nugget effect for all samples.



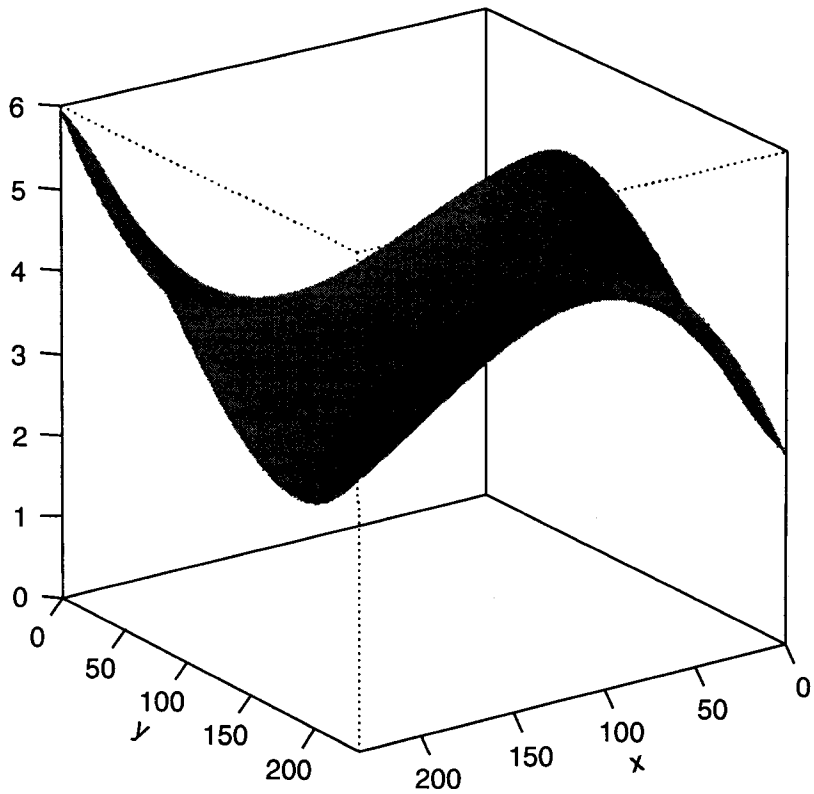


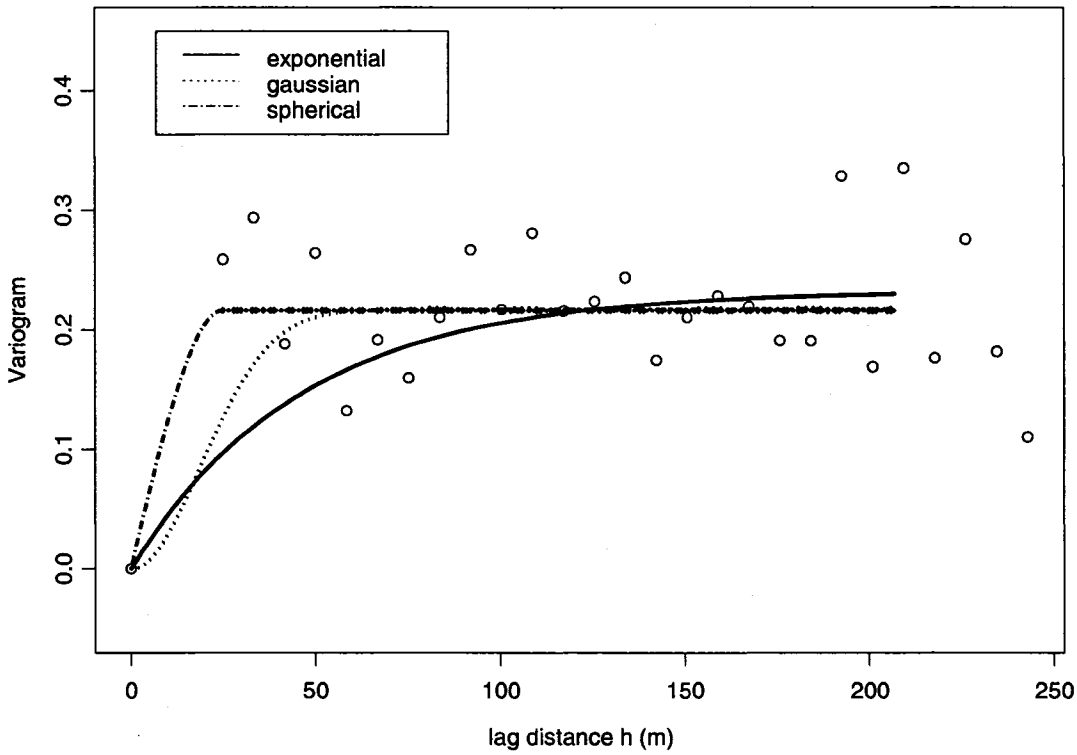


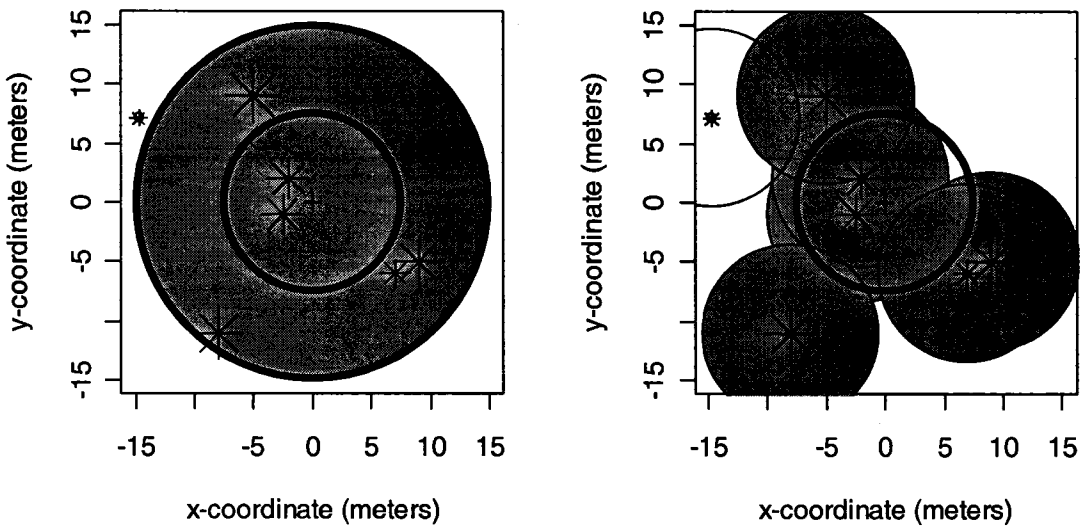
z(s) values for 36 points



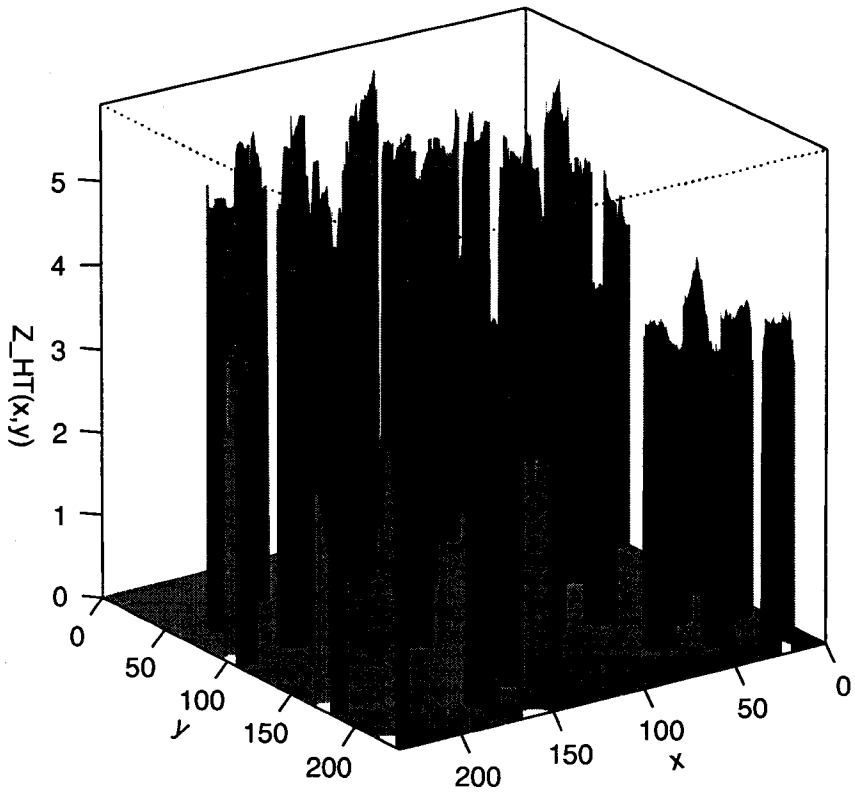
Trend surface model estimated from the 36 points

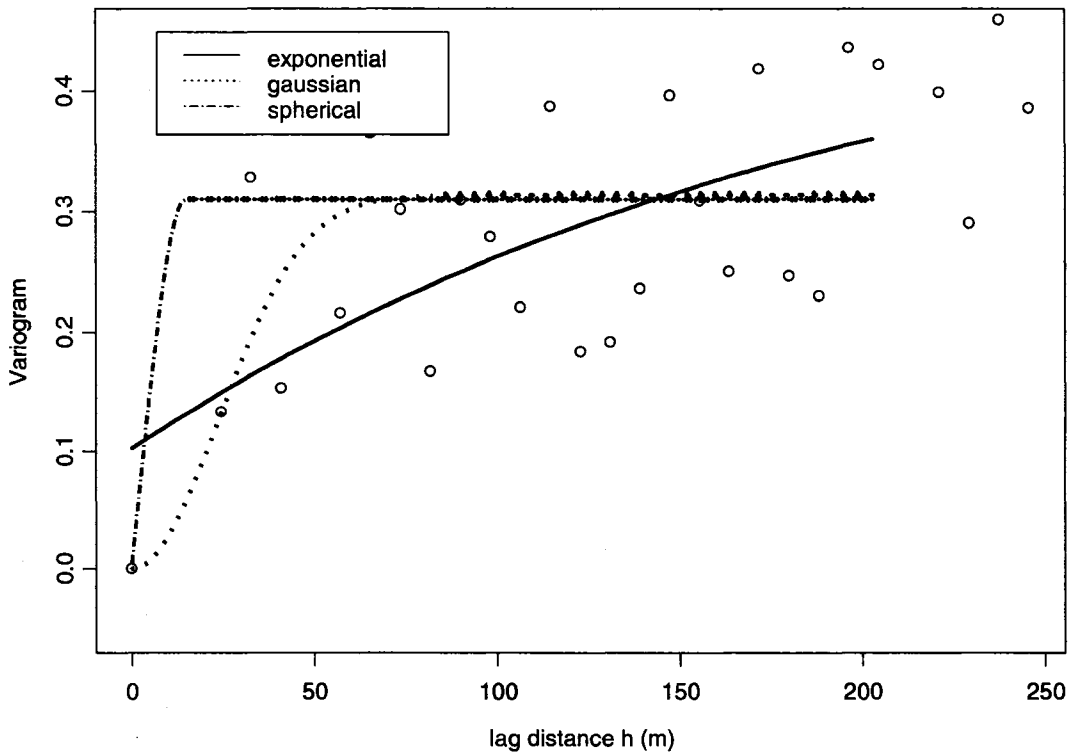


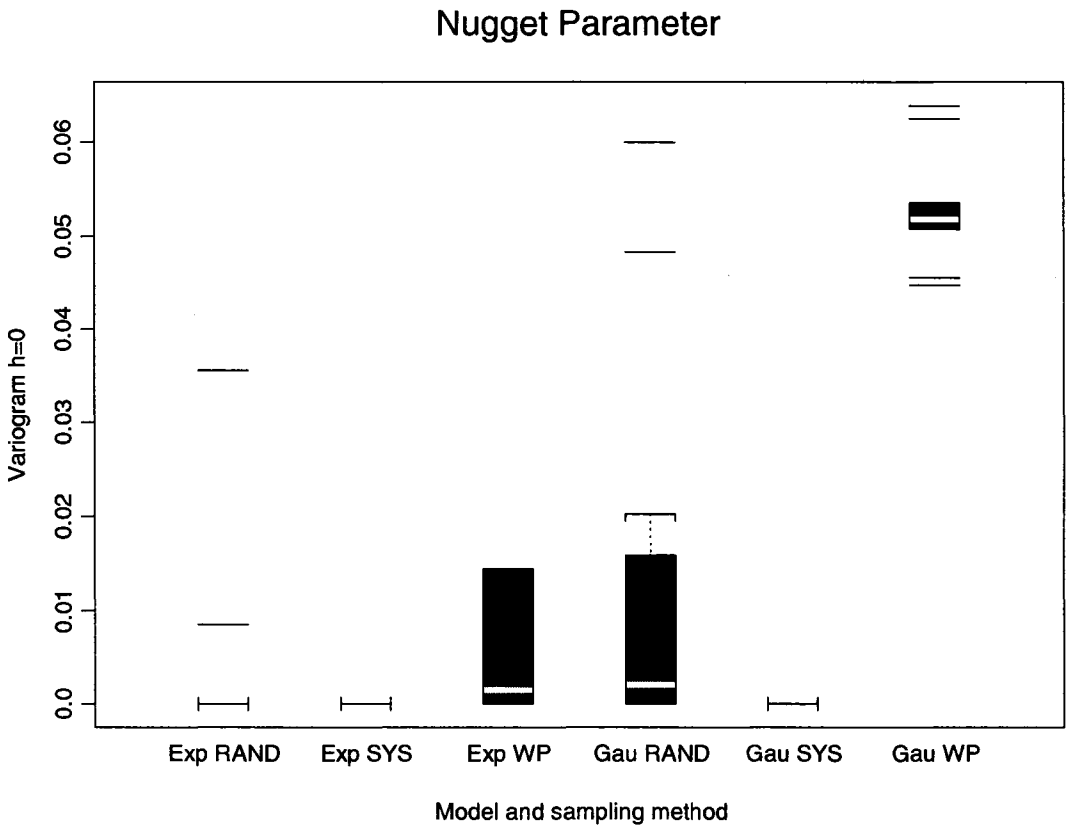


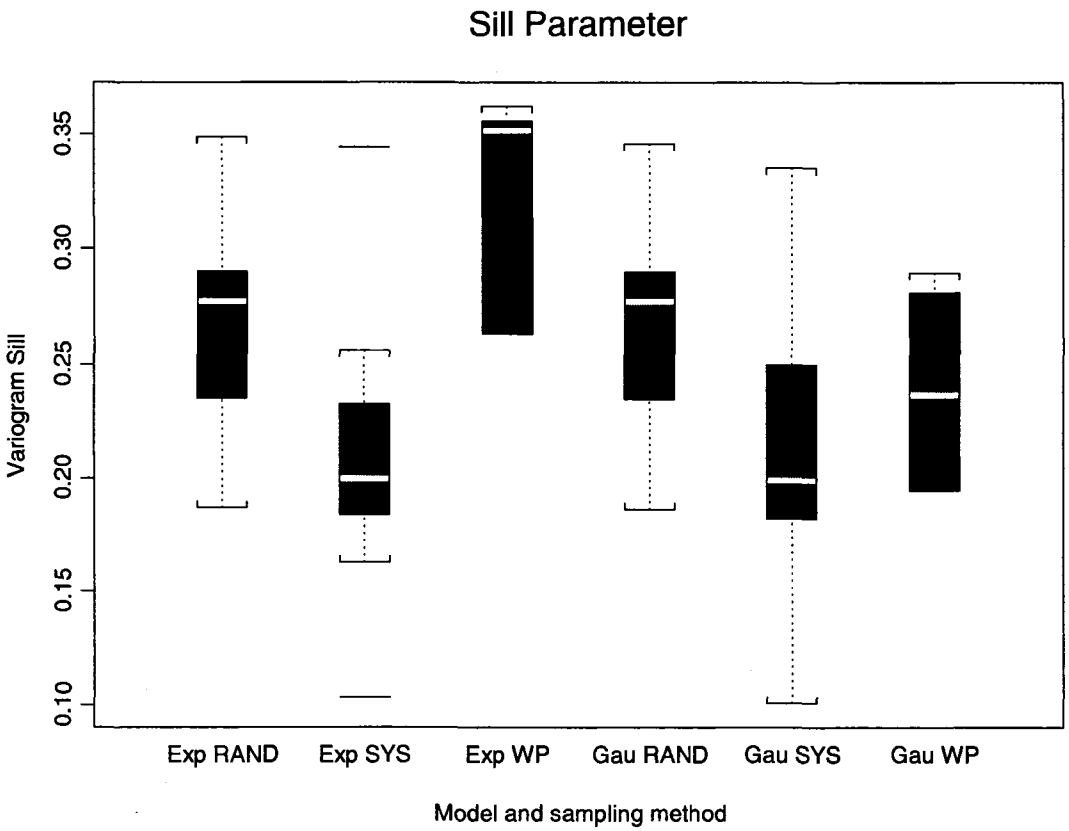


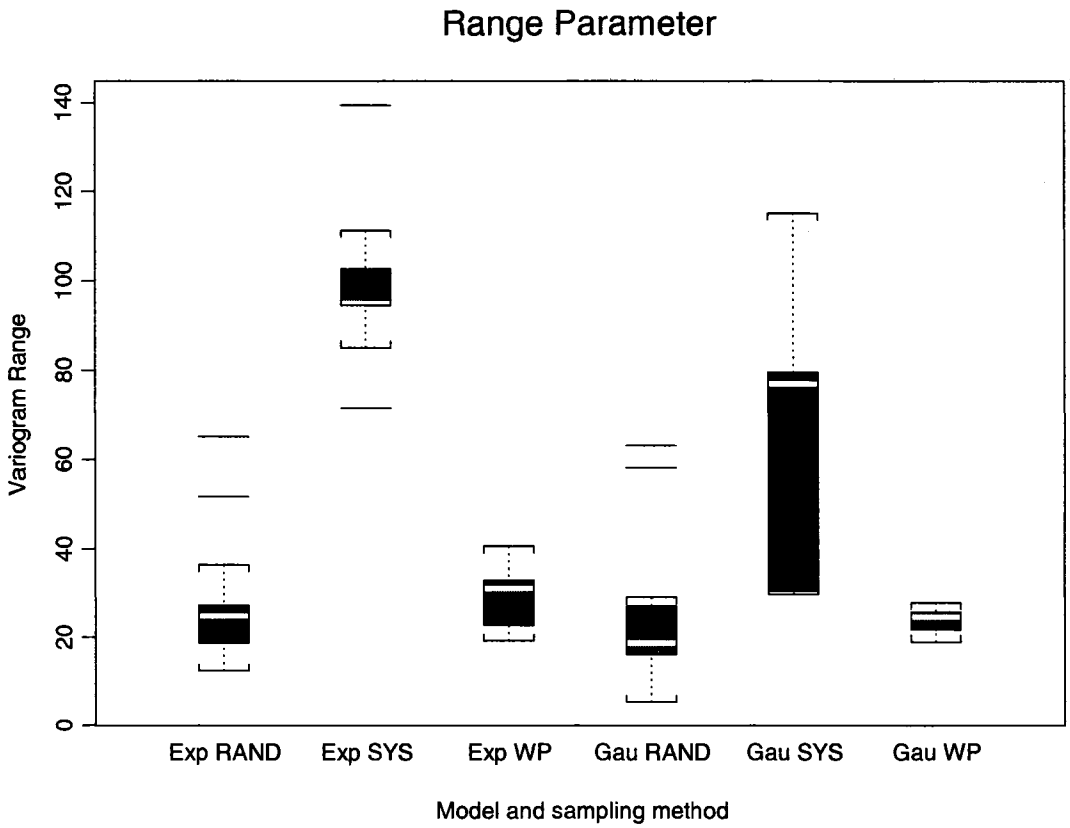
z(s) from the half-sample points

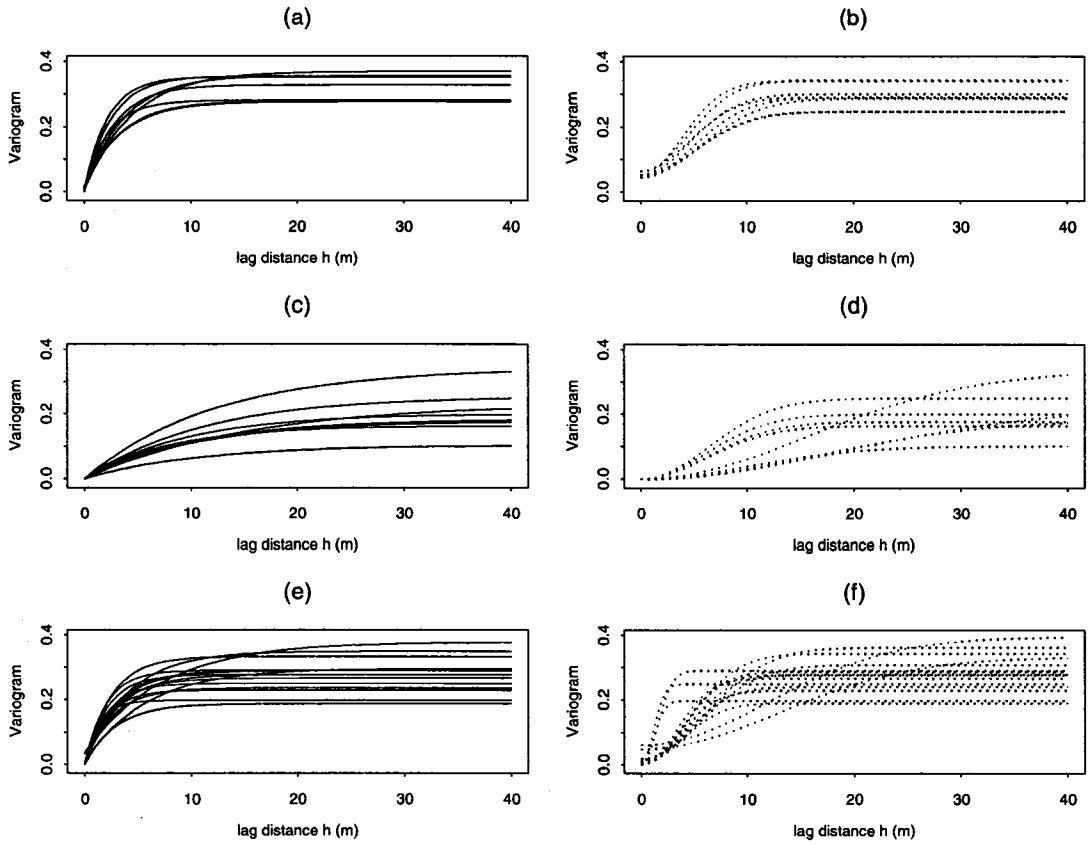


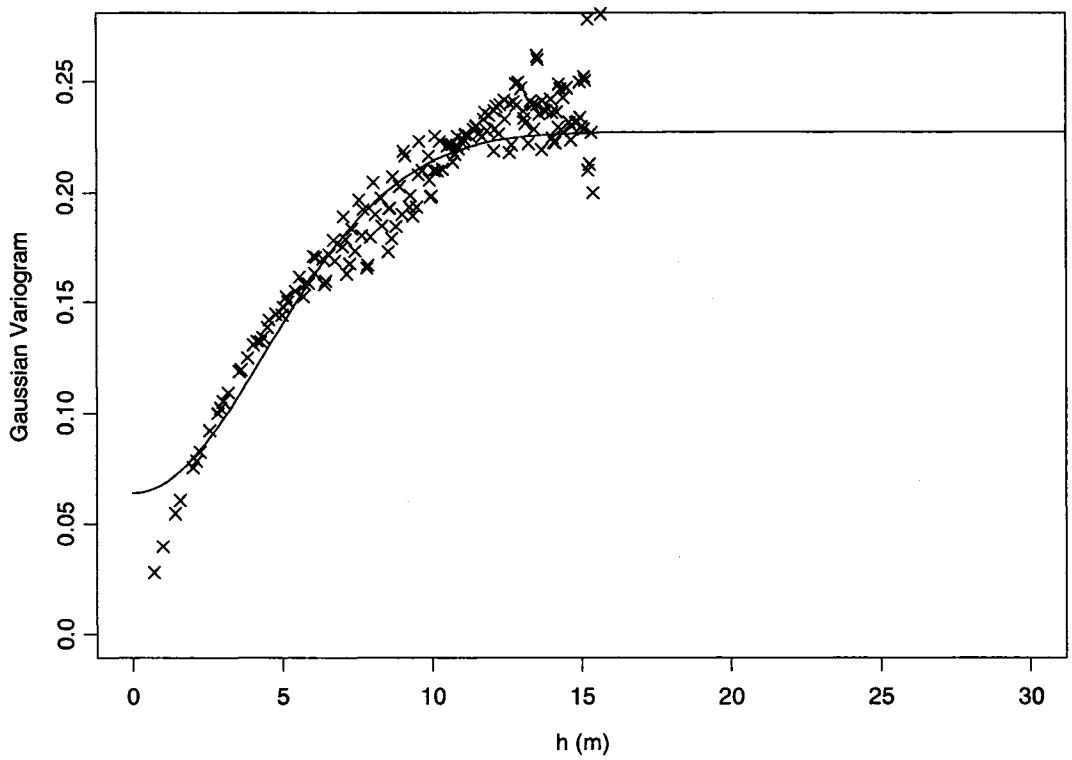












Chapter 3

MODIFICATIONS TO KRIGING FOR WOBBLE-PLOT SAMPLING: AN EMPIRICAL EVALUATION

Abstract

Many environmental surveys collect data for the purpose of estimating the abundance of a discrete attribute, with an example from forestry being the estimation of total biomass or the percentage of damaged trees for a forest stand. Another objective is the production of maps, which requires prediction of the same attributes at unobserved locations. Geostatistical methods, such as kriging, have become commonly used tools in the production of maps in environmental applications. In order to perform kriging, a variogram model must be fitted to data collected at various locations throughout the study area. A concern with using the data derived from a survey that was designed for the estimation of the mean or total is that the sample locations are often separated by distances that are too great to allow accurate estimation of the variogram. One consequence is that the range parameter can be biased, so that the kriging predictor is no longer a best linear unbiased predictor. This study uses a technique, referred to as wobble-plot sampling, to accurately estimate the variogram over short lag distances. It also introduces modifications to kriging that take advantage of the additional information provided by the wobble-plots. The new techniques are compared to a traditional kriging approach that does not used the wobble-plot enhancements using a simulation study. The results show that, on average, using the variogram derived from the wobble-plot data reduced the mean squared prediction error by about 5%. The reduction in mean squared prediction error was 19 % when both the improved variogram and a modified kriging procedure were employed. The performance of 95 % prediction intervals was also assessed, where the coverage rates of the prediction intervals for the standard approach averaged only 74 %, with the achieved coverage rate

for some samples being less than 60%. The improvements in the estimated variogram and a nonparametric prediction interval technique lead to an achieved coverage rates ranging from 91-93 %.

Key Words: Variogram, lag distance, kriging, spatial modeling, prediction interval

3.1 Introduction

Many environmental surveys collect cluster-plot data for the purpose of estimating the abundance of a discrete attribute, with an example from forestry being the estimation of basal area or volume with a fixed-area plot or quadrat. Another objective of these surveys is to perform prediction of the same attributes at unobserved locations. These predictions are often performed using geostatistical methods, such as kriging. The predictions are then used to produce maps or perform other ecological analyses. Examples of kriging applications are the mapping of forest pathogens (Kallas et al. 2003) and forest stem volume (Mandallaz 2000, Wallerman et al. 2002 and Wallerman 2003).

There are many types of environmental inventory systems, but the application considered here mimics the approach commonly used for forest management of public and industrial lands. In this approach, large tracts of forest land are subdivided into smaller stands with similar forest conditions. Sampling is carried out within each of the individual stands to characterize the forest conditions within each stand. The mapping of forest resources can be performed at a rather coarse scale by displaying the mean value of each stand or at finer scales by modeling the spatial variation within each stand. The application considered here is the prediction of forest resources at an unobserved location, s_0, using kriging techniques at the scale of an individual or small group of stands. The attribute of interest for the population of trees is denoted by $z_1, ...z_N$, which is embedded in a domain A. The data are collected by sampling trees, or other discrete population elements, using some form of plot sampling, with circular fixed-area plot sampling being used in this study. The per unit area response, $Z(s)$, at any point location $s \in A$ is

$$Z(s) = \sum_{i=1}^{n(s)} \frac{z_i}{a(s)},$$

where n(s) is the number of trees within the boundary of a circular fixed-area plot, i indexes the trees on the plot, and $a(s)$ is the area of the plot within A.

There are a number of significant issues related to spatial modeling with plot data. The term *support* of an observation is used to describe the finite area or volume that is measured in order to calculate the response $Z(s)$ at a point with coordinates s. The

underlying problem is that inferences made about the underlying population are dependent on the size of the plot. The aggregation of the population elements affects the correlation observed between locations. Generally, greater aggregation leads to an increase in the sample correlation between locations. The dependence of this correlation on plot size has been deemed the *modifiable area unit problem*. This observed correlation between aggregated observations can lead to a conclusion that a significant spatial relationship exists, when in reality there is none. The term *ecological fallacy* is used to describe this situation (Openshaw and Taylor 1979). Gotway and Young (2002) provide an overview of many issues related to this problem. In this study, we assume that inferences that are specific to the individual elements $z_1, ...z_N$, such as estimating the total would be made using traditional survey sampling techniques and that the goal of a spatial modeling application is to produce a map displaying the relative distribution of forest resources across the domain A and a related map of prediction errors.

Many different forms of kriging exist. However, a number of studies have found that, for a given process, the differences in the predictive performance of competing kriging methods have often been limited (Papritz and Moyeed (1999 and references therein), Wallerman (2003)). For example, Moyeed and Papritz (2002) performed an extensive empirical comparison of kriging methods and found no clear differences in the predictive performance of ordinary, disjunctive, indicator, or model-based kriging on populations whose responses were either linear or nonlinear. They also found that the use of the Matérn class of variograms provided no significant improvement in prediction. Moyeed and Papritz (2002) did, however, find that the nonlinear methods did show improvement over linear methods in modeling the conditional distribution of marginally skewed data. This improvement led to prediction interval coverage rates that more closely matched their nominal value. This study takes a different approach, in that rather than attempting to find a better predictor for a specific application or to utilize auxiliary data sources, improvements in the prediction are sought through additional manipulations of the existing data.

Spatial prediction with most forms of kriging relies on the assumption that the variogram is known. However, in practice, the true variogram is never available and must be estimated from the plot data. A common problem when estimating the variogram from

plot sampled data is that few if any closely spaced observations are available. The term *lag* or *lag distance*, denoted by h, is used to describe the spacing between sample locations (i.e., $h = |s_j - s_{j'}|$, where s_j and $s_{j'}$ and the location of sample points). The reason there are rarely any data at short lag distances in many forestry or ecological applications is that it would require that cluster plots be located so that they partially overlap, which is impractical and inefficient when one of the objectives of the survey is to estimate the mean or total of the resource. Thus, the variogram is fitted by extrapolating between $h = 0$ and the closest plot locations. The concern with this extrapolation is that optimal, or nearly optimal, prediction is only achieved when the estimated variogram closely matches the true variogram as $h \to 0$. Stein (1988) studied the behaviour of the kriging predictor when the variogram is incorrectly specified. He shows that as long as the behavior of the misspecified variogram is "similar" to that of the true variogram function near the origin, the resulting kriging predictor will have the same asymptotic efficiency as the predictor based on the true variogram. Thus, a major concern for kriging applications with environmental survey data is that there is often no data for the most critical lag distances.

This issue is adressed in the previous study with the developement of wobble-plot sampling. This simple technique uses a change of spatial scale and mapping of the locations of population elements on the plot to enumerate every possible sample outcome in a neighborhood about each point. In essence, the change of scale and the mapping provides the infinite set of responses that would be observed if one were to allow a plot to *wobble* about the point, hence the name *wobble-plot sampling*. The abbreviation WP will be used in the remainder of the study.

Williams and Reich (2005) compared the method using a small forest stand where 36 systematically arranged WPs were tested against the same sample of 36 points and a sample of 200 points that were randomly located within the stand. They found the following:

- The variogram fitted from the WP data varied very little from sample to sample.
- The exponential variogram was always selected over the Gaussian and spherical models as best model when the WP data were used. The exponential variogram was very unstable and rarely selected for the other two methods.

- Variograms fitted to the systematic data over-estimated the range of spatial dependence by a factor of between 2 and 3.

- Variograms fitted to the sample of 200 points were still substantially more variable than those fitted to the WP data.

Authors, such as Cressie (1993) and Stein (1999), have done much to develop a sound theoretical foundation for kriging. However, as noted by Goovaerts (1997), the outcome of most kriging applications will likely be influenced by the subjective nature of model selection, numerical difficulties with matrix inversions, the bias in the estimation of mean squared prediction error, and other practical issues. WP sampling at least partially addresses and improves some of these issues and Williams and Reich (2005) do not address the possible extensions beyond improved estimation of the variogram. Thus, it seems reasonable to explore whether the application of WP sampling to spatial prediction can lead to additional improvements and to illustrate those improvements through a simulation study on a small forest population.

The goals of this study are:

- Propose practical adaptations to existing methods and software to fit variograms using WP data.

- Propose practical modifications to kriging to improve spatial prediction.

- Test the methods and quantify the improvements in spatial prediction due to WP sampling.

The remainder of paper is organized as follows: Section 2 provides a brief overview of WP sampling. Section 3 outlines issues of spatial prediction where deficiencies and difficulties still exist for spatial prediction. Section 4 provides a brief description of the data set used. Section 5 presents modifications of standard prediction practices when WPs are used. Section 6 tests these modifications on the data to quantify improvements in spatial predictions when WP data are used.

3.2 Review of WP sampling

An example is used to motivate and describe the technique. Consider the following typical application in forest survey, where the edges of a forest stand are delineated on a map and a series of forest survey plots are installed. Most forest inventories are designed with the goal of estimating the mean or total of the forest resources within the stand boundary. Given this goal, cluster plots are located within the boundaries using some form of systematic sampling. The reason for using a systematic sample of points is that the cluster plots are easily located, the sample provides good coverage of the area, and that systematic sampling, while not optimal, is often both efficient and robust when the goal of the survey is to estimate the mean (Cochran 1946, Bellhouse 1977, Iachan 1985). An example of such a survey is given in Figure 1, where $m = 22$ circular fixed-area plots with radius $r = 8.2m$ are installed using a systematic grid. The separation distance between grid points was approximately $h = 50m$. At each location s, all trees within the boundary of the plot are measured to determine the per unit area response

$$Z(s_j) = \sum_{i=1}^{n(s_j)} \frac{z_i}{a(s_j)},$$

where $n(s_j)$ is the number of trees within the plot boundary, $a(s_j)$ is the plot area falling within the stand, and z_i is the attribute of interest for each tree. The mean is estimated by

$$\frac{1}{m} \sum_{j=1}^{m} Z(s_j).$$

While this typical survey design is well suited for the estimation of the mean, describing it as "poor" for most spatial modeling applications is being charitable. The obvious problem is that the design provides no data with short lag distances to facilitate estimation of the variogram. Another less obvious problem is that in order to quantify many environmental phenomena, the population elements must be aggregated and measured over an area or volume, thus some assumptions must be made about the *spatial scale* of the fluctuations in the process (Cressie 1993, p. 112). This choice of scale is partially defined by the level of aggregation needed to collect meaningful data. A significant challenge is that many ecological processes take place over short distances. An example, given by Urban (2005), is

that seed dispersion often takes place over distances of less than 10 meters. The problem with many forest inventory plots is that the large plot size dampens much of the signal. The usual solution to the problem is to either use a completely different design for collecting data for spatial modeling applications or to collect information at an auxiliary sample of points for estimating the small-scale variability. Urban (2005) describes a design based on randomly located auxiliary plots, as depicted in Figure 2, while Olea (1984) and Ritter and Leecaster (2005) use systematically arranged grids of auxiliary plots.

The premise of WP sampling is that much of the necessary data for estimating the variogram are often available. To illustrate, consider a small section of forest and assume that the plot design used is a circular fixed-area plot with radius $r = 15m$ (Figure 3a). Using the plot-centered approach to areal sampling (Husch et al. 1983), six of the seven trees are included in the sample and the response at $s = (0, 0)$ is

$$Z_{r=15}(s) = \frac{|A|}{\pi * 15^2} \sum_{i=1}^{6} z_i.$$

If the size of the plot is reduced to $r = 7.5m$, only two of the trees would be selected and at $s = (0, 0)$ (Figure 3b) and the response is

$$Z_{r=7.5}(s) = \frac{|A|}{\pi * 7.5^2} \sum_{i=1}^{2} z_i.$$

However, note that if the locations of the trees selected on the larger plot were mapped in relation to the location s, the object-centered approach can used to determine $Z_{r=7.5}(s)$ at all locations $s + h$ over the set $B = (s + h|h \leq 7.5m)$, because only those trees on the original $r = 15m$ plot would ever be selected by any point that fell within a radius of $r = 7.5m$ of the center point of the plot. This is equivalent to letting a smaller plot "wobble" within the larger plot.

Another way to describe WP sampling is that it is equivalent to placing an infinite number of smaller plots within a single larger plot. The only restriction placed on the location of the smaller plots is that they must be completely encompassed by the large plot. For the example given in this study, the WP is one half the radius of the original plot, but there is no reason why the ratio must be 2. There is simply the tradeoff between increasing the spatial scale and reducing the range of lag distances for which observations are available. Details regarding the specific implementation of WP sampling are given later.

3.3 Prediction with Kriging

There are many different types of kriging. However, in any application there are a number of practical issues associated with even the simplest forms of kriging. To introduce notation and define the problem, a brief overview of ordinary kriging is provided.

For ordinary kriging, the model assumed is that data $\underline{Z} = Z(s_1), ...Z(s_m)$ are generated by a random process,

$$Z(s) = \mu + \delta(s),$$

where μ is assumed unknown. The goal of ordinary kriging is to derive the best linear unbiased predictor, where *best* is defined by squared error loss. The predictor at an unobserved location s_0 is

$$\hat{Z}(s_0) = \sum_{j=1}^{m} \lambda_j Z(s_j),$$

where $\sum_{j=1}^{m} \lambda_j = 1$. The condition that the λ_j sum to one ensures that the predictor is unbiased. Determining the best linear unbiased predictor (BLUP) at a new point s_0, denoted by $\hat{Z}(s_0)$ is equivalent finding the λ_j that solve

$$\min E[(Z(s_0) - \sum_{j=1}^{m} \lambda_j Z(s_j))^2],$$

subject to the constraint $\sum_{j=1}^{m} \lambda_j - 1 = 0$. The problem can be reformulated using the Lagrange multiplier technique to one of finding the $\lambda_j, j = 1, ...m$ and q that minimize

$$E[(Z(s_0) - \sum_{j=1}^{m} \lambda_j Z(s_j))^2] - 2q \sum_{j=1}^{m} \lambda_j - 1.$$

Taking the partial derivatives with respect to the λ_j and q and equating them to zero yields the system of $m + 1$ equations

$$-\sum_{j'=1}^{m} \lambda_{j'} E[(Z(s_j) - Z(s_{j'}))^2]/2 + E[(Z(s_0) - Z(s_j))^2]/2 - q = 0, \; j = 1, ...m. \qquad [1]$$

$$\sum_{j=1}^{m} \lambda_j = 1.$$

3.3.1 Estimation of the variogram

Equation [1] poses a problem because $E[(Z(s_j) - Z(s_{j'}))^2]/2$ cannot be estimated for the n data points. Thus, prediction is not possible without the key assumption that $E[(Z(s_j) - Z(s_{j'}))^2]/2$ can be described by the known function $\gamma(s_j - s_{j'})$, which is the variogram. In practice this function is not available and it is actually a parameter that must be estimated from the data. As noted by Cressie (1993), stationarity of the variogram is not a necessary condition for the ordinary kriging to be a BLUP, but it must be assumed in order for the variogram to be estimated from the data.

The variogram is estimated by first generating the variogram cloud using the classical variogram estimator

$$\hat{\gamma}(h) = \frac{1}{|M(h)|} \sum_{M(h)} [Z(s_{j'}) - Z(s_j)]^2,$$

where $M(h) = \{(s_j, s_{j'}) : |s_j - s_{j'}| = h; j, j' = 1, 2, ...m\}$ is the set of pairs of points separated by a distance h. If the data are irregularly spaced then $M(h)$ is defined to be some neighborhood about the lag distance h for which there are a reasonable number of points. Note that both the variogram and the lag distance are estimated over a neighborhood chosen by the user when the data are irregularly spaced. Once the variogram cloud is produced, a valid parametric model $\gamma(h; \theta)$ is chosen and fitted to the variogram cloud.

The concern with many kriging applications is that the variogram $\gamma(s_j - s_{j'})$, is often poorly estimated because there is insufficient data to estimate the model parameters θ, which in turn determine the shape of the variogram. The consequence of misspecifying the variogram is that the kriging predictor is no longer *best*, though the constraint that $\sum_{j=1}^{m} \lambda_j = 1$ still ensures unbiasedness with respect to the model.

Stein (1988) studied the behaviour of the kriging predictor when the covariance function is incorrectly specified. He shows that as long as the behavior of any covariance function is "similar" to that of the true covariance function near the origin, the resulting kriging predictor will have the same asymptotic efficiency as the predictor based on the true covariance function.

3.3.2 Estimation of the mean squared prediction error

Under the assumption that the variogram is known, the mean squared prediction error is given by

$$\sigma^2(s_0) = 2\sum_{j=1}^{m} \lambda_j \gamma(s_0 - s_j) - \sum_{j=1}^{m}\sum_{j'=1}^{m} \lambda_j \lambda_{j'} \gamma(s_{j'} - s_j). \quad [2]$$

Given that the true variogram is unknown, the mean squared prediction error is estimated using a "plug in" approach that replaces $\gamma(s_0 - s_j)$ in eq. [2] with the estimated variogram $\hat{\gamma}(s_0 - s_j)$. Then the estimated mean square prediction error in given by

$$\hat{\sigma}^2(s_0) = 2\sum_{j=1}^{m} \lambda_j \hat{\gamma}(s_0 - s_j) - \sum_{j=1}^{m}\sum_{j'=1}^{m} \lambda_j \lambda_{j'} \hat{\gamma}(s_{j'} - s_j). \quad [3]$$

As noted by Cressie (1993), $\hat{\sigma}^2(s_0)$ does not account for the estimation of the variogram. If the notation, $\tilde{p}(\underline{Z}; s_0)$, is used to denote the predictor obtained from an estimated variogram, then one expects that

$$\sigma^2(s_0) \leq E[(\tilde{p}(\underline{Z}; s_0) - Z(s_0))^2].$$

This relationship is expected to hold when both sides are estimated from the data, with the concern being that $\sigma^2(s_0)$ is used to determine prediction intervals when in fact $E[(\tilde{p}(\underline{Z}; s_0) - Z(s_0))^2]$ should be used instead. The notation $\tilde{\sigma}^2(s_0) = E[(\tilde{p}(\underline{Z}; s_0) - Z(s_0))^2]$ will be used to denote the true mean squared prediction error and to emphasize that the λ_j are derived from the estimated variogram.

A number of methods have been proposed to account for the uncertainty associated with the estimation of the variogram. Zimmerman and Cressie (1992) and Harville and Jeske (1992) provide approximate solutions that are valid under strong assumptions about the process $Z()$. Wang and Wall (2003) propose a parametric bootstrap method similar to the methods proposed by Zimmerman and Cressie (1992) and Harville and Jeske (1992). Diggle et al. (1998) developed a method to account for the estimation of model parameters that imbeds a linear kriging predictor within the structure of a generalized linear model. A Bayesian approach is then used for predictive inference. A concern with these methods is that they can be quite computationally intensive and often yield little or no improvement in the actual predictions (Moyeed and Papritz (2002)).

3.3.3 Construction of prediction intervals

The estimated mean square prediction error can be used to construct a "plug-in" prediction interval. Under the assumption that the random process $Z()$ is Gaussian, the interval

$$I(s_0) = \left(\hat{Z}(s_0) - 1.96\sigma(s_0), \hat{Z}(s_0) + 1.96\sigma(s_0)\right),$$

is a nominal 95% prediction interval for $\hat{Z}(s_0)$. Given that $\hat{\sigma}^2$ under-estimates the actual MSPE, the coverage probability of the prediction interval on average will be lower than the nominal probability. Another concern, particularly for this application, is that $Z()$ is not Gaussian.

The advantage of the Bayesian and bootstrap methods described above is that the estimation of the variogram parameter can be accounted for in the construction of the prediction interval. The degree to which the use of $\hat{\sigma}^2$ biases the coverage rates of the prediction interval is often application dependent, with Moyeed and Papritz (2002) finding that "plug-in" methods were no worse than the fully model-based approach proposed by Diggle et al (1998). Wang and Wall (2003) tested the performance of three bootstrap techniques and one Bayesian technique and found that two of the bootstrap methods essentially achieved their 95% nominal coverage. However, the achieved coverage rate of the "plug-in" interval was still 94.12 %, which would probably be acceptable in most applications. In contrast to these results, Sjöstedt-de Luna and Young (2003) found more substantial improvements when comparing the "plug-in" interval to a bootstrap calibrated interval technique. Their extensive simulation study was based on a Gaussian random process and studied both in-fill and increasing domain asymptotic properties.

3.4 Data Description

A small data set is used for the purpose of illustration throughout this study. The data were collected in central New Jersey in 1985 on an area of 5.37 ha, which was extracted from the center of a larger forest stand. Two distinctly different forest types are represented, these being adjacent hardwood stands of roughly equal area containing a total of 4,744 trees. The data set relates to two stands of different structure, one comprised of mature

hardwoods of saw-log size and the other of hardwood regeneration. The two stands differed in structure and are separated by a dirt road that divides the population roughly in half and has partially grown over. Conditions within each stand are fairly heterogeneous with areas of visibly different intensity within and between the stands. Figure 4 depicts the structure of the data set. The diameter of each circle represents the relative size of each tree. This surface was created by calculating the response across a very fine grid that covered the population. The grid spacing was $0.7m$, which yielded a set of 108241 responses $Z(s)$. The data were transformed to improve the assumption of stationarity and detrended using a third degree polynomial.

The distribution of the responses, $Z(s)$, for virtually any plot-sampling application does not follow a Gaussian distribution. This occurs because a tree, and its attribute values (z_i), are only included in the sample whenever the center of the bole falls within the plot boundary. This causes the response over a small neighborhood about s to change only when a population element either enters or leaves the sample. Thus, $Z(s+h)$ is a constant until the plot is moved a sufficient distance and direction so that the term $n(s+h)$ in

$$Z(s+h) = \frac{|A|}{\pi * r^2} \sum_{i=1}^{n(s+h)} z_i,$$

is incremented or decremented. A QQ-plot, derived from the 108241 $Z(s)$ values, is given in Figure 5 to illustrate the substantial departure from Normality.

Another factor that makes this a challenging data set is that there are a number of outliers in the data. These are caused by areas in the data set where there are gaps and small openings such that plots do not sample any trees. Roughly 0.3% of the data values were 0s, with the majority occurring along the upper half of the y-axis, for $x = 0$. No attempt was made to remove these outliers.

Measurement error is also expected to add little to the nugget effect in this application because the inclusion of trees in the sample and their measurement of their basal area are very repeatable measurements. Exceptions to this general statement include situations where tree and vegetation attributes are visually estimated or when models are used to predict various tree attributes, such as volume and biomass.

3.5 Implementation and modifications to kriging with wobble-plot data

The use of WP data poses some unique problems because there are an infinite number of sample points, which are grouped into clusters, with which to fit the variogram and perform spatial prediction. While it may be possible to calculate closed-form expressions to describe the spatial correlation, this would be tedious and it would also require abandoning existing software packages. The goal of this section is to describe one possible implementation of variogram estimation and kriging that utilizes the additional data while still allowing the use of existing software and methodology. Two issues are addressed, the first being changes to the usual approach to estimation of the variogram and the second being modifications to kriging. The data set described above is used for the purpose of illustration, where a sample of $m = 30$ points $s_1, ...s_m$ was selected using a Strauss process with complete inhibition at distances less than $25m$.

To illustrate the proposed modifications it is necessary to first outline a standard approach to modified residual kriging. The outline follows that given by Müller and Zimmerman (1999) where the analysis, given data $Z(s_1), ...Z(s_m)$ is carried using the following six-step process:

1. Propose a model of the large-scale spatial patterns in the data plus an intrinsically stationary error model that describes the small-scale spatial relationships within the data. This model is given by

$$Z(s) = \mu(s) + \delta(s).$$

   For this study, a third degree polynomial with unknown parameter vector β was chosen to describe the large-scale spatial variation.

2. Estimate the β parameter for the model describing large-scale variation. The difference between the observed and modeled values at each point s_j is used to determine the residuals for use in estimating the stationary error process.

3. Use the $m(m-1)/2 = 435$ paired distances to produce a plot of the error process using the classical variogram estimator.

4. Choose a valid parametric model $\gamma(h;\theta)$ which is compatible with the plot of the variogram.

5. Fit the chosen model to the estimated variogram to estimate the parameters of the model.

6. Use kriging in conjunction with the model describing the large-scale trend in the population to predict the values of the spatial process at unobserved locations and the estimated mean square prediction error.

3.5.1 Estimation of the variogram with wobble-plot data

In a standard application, fitting a variogram model is divided into two stages. In the first stage, the distance from each data point to the remaining $m-1$ data points is calculated. This provides a set of $m(m-1)/2$ unique lag distances. These lag distances are broken down into classes based on the lag distance and the classical variogram estimator is used to calculate the variogram cloud. Once the variogram cloud is produced an appropriate model is selected and fitted to the data.

The use of WP data presents a problem because there are an infinite number if points available for fitting the variogram. In order for existing software and approaches to be applicable, only a finite number of points is desired. This number must also be sufficiently small so that the number of paired distances can be calculated and stored using generally available computing resources. To address these problems the following algorithm was employed:

1. Overlay a square grid of points on each WP. The square grid produced a set of lag distances $h, \sqrt{2}h, 2h, ...r$, where $h = 0.7m$ in this study. Provided the plot did not intersect the boundary, the grid created a set of $m' = 368$ points within each WP, which in turn created 67528 lag distances ranging for $h = 0.7 - 15m$. An example is given in Figure 6.

2. The mapped coordinates of each sample element are used to determine the true response $Z(s)$ at every one of the $m' = 368$ points on the grid.

3. The classical variogram estimator term $[Z(s_{j'}) - Z(s_j)]^2$ is calculated for each of the $m'(m'-1)/2 = 67528$ paired observations. These values are stored in an array.

4. Steps 1-3 were repeated for each of the $m = 30$ points.

5. The m individual arrays of $m'(m'-1)/2$ squared differences are concatenated and sorted by h and the number of $[Z(s_{j'}) - Z(s_j)]^2$ terms in each class is calculated to determine $|M(h)|$. The total number of paired distances was nominally $mm'(m'-1)/2 = 2025840$, though this value was reduced to approximately 1.8 million observations by plots that intersected the boundary.

6. The $\hat{\gamma}(h)$ is calculated for each of the individual h classes. An example is given in Figure 7.

This algorithm provides $\hat{\gamma}(h)$ estimates up to a maximum lag distance of twice the WP radius. To provide additional $\hat{\gamma}(h)$ values at greater lag distances, the responses at each of the original m point centers (s_j) were used. For this application, these data were irregularly spaced and required grouping the data into bins based on their lag distance. The width of each bin was allowed to vary so that the number of points in each bin was equal to the size of the smallest bin from the WP data. The lag distance used for each of these bins was the mean of the lag distances in the bin. These $\hat{\gamma}(h)$ values were concatenated onto the array of WP derive $\hat{\gamma}(h)$ values (Figure 8). Finally, an appropriate variogram model was selected and fitted to the variogram cloud.

3.5.2 Modifications to kriging with wobble-plot sampling

Note that there are two forms of prediction that can take place in a standard kriging application, these being interpolation between points and extrapolation beyond the domain of the points. The kriging predictor in both applications is given by $p(\underline{Z}; s_0) = \sum_{j=1}^{m} \hat{\lambda}_j Z(s_j)$. However, estimation of the $\hat{\lambda}$'s requires the inversion on an $m \times m$ matrix, which is generally not possible. The solution is to select only a small subset of size $\tilde{m} << m$ of the nearest points and perform the prediction using only these points.

Prediction with data derived from WPs poses a problem because, given an infinite set of points to use for prediction, it is not clear what constitutes the best subset of "closest"

points. This occurs because, given the infinite number of possible WP points, any set of the $\tilde{m}$ "closest" points all lie on the boundary of the WP that is closest to s_0 and have the same $Z()$ value. If this approach were taken, prediction with WP data would be equivalent to extrapolating the $Z()$ on the boundary of the WP from the edge of wobble plot to the boundary of the Thiessen polygons that tessellated the m points.

The proposed solution is pragmatic and based on the notion that interpolation is preferable to extrapolation. The proposed solution is as follows:

- If the point s_0 falls within the boundary of the WP, use the actual $Z(s_0)$. The MSPE for this point is $\hat{\sigma}^2(s_0) = 0$, unless measurement error is to be accounted for, in which case $\hat{\sigma}^2(s_0) = \hat{\sigma}^2_{nugget}$, which is the estimated nugget effect from the variogram.

- If the point s_0 falls outside of the boundary of all wobble plots, select the new set of m' points, denoted by s'_j that are closest to s_0. These are the points from the boundary of the m' closest wobble plots (Figure 9).

This approach has two advantages, with the first being that the MSPE will likely be reduced because $\gamma(s_0 - s'_j) \leq \gamma(s_0 - s_j)$ for all m' points. The second advantage is that the closest point on the boundary of the wobble plot has not been previously been used in the estimation of large-scale trend or the fitting of the variogram. Thus, concerns regarding correlations in the residuals causing a bias in either the estimation of $\hat{Z}(s_0)$ or $\hat{\sigma}^2(s_0)$ are reduced (Cressie 1993, p. 167.)

3.5.3 Modifications to prediction interval estimation

Under that assumption that the $Z()$ process is stationary, the large number of observations provided by the WPs should greatly reduce the effect of the estimated variogram on the prediction interval. However, as illustrated in Figure 5, basing the upper and lower bounds of a prediction interval on the assumption that $Z()$ is Gaussian is clearly inappropriate.

The solution chosen is to mimic the percentile interval technique used in the construction of simple bootstrap confidence intervals (Efron and Tibshirani 1993, Chapter 13). However, instead of relying on a resampling technique, the method simply uses the large

sample of $Z(s)$ values to estimate the appropriate upper and lower bounds based on the quantiles of the observed values.

Suppose the goal is to construct a $1 - 2\alpha$ prediction interval for a new observation $\hat{Z}(s_0)$. The proposed nonparametric approach first scales the WP observations

$$Z^*(s) = \frac{(Z(s) - \bar{Z})}{\hat{\sigma}_Z}$$

to create a the random variable such that $Z^* \sim (0.1)$. Then the goal is to determine the α and $1 - \alpha$ quantiles the distribution for Z^*. If these quantiles are denoted by q^{α} and $q^{1-\alpha}$, then the proposed prediction interval is given by

$$I_{WP}(s_0) = \left(\hat{Z}(s_0) - q^{1-\alpha}\hat{\sigma}(s_0), \hat{Z}(s_0) - q^{\alpha}\hat{\sigma}(s_0)\right).$$

3.6 Simulation Study

The goal of this study is to test the effect of using data derived from WP sampling on spatial prediction. To quantify the differences in performance between the methods, a simulation study was performed. The simulation study drew samples from the population, fit variogram models and used kriging to predict new observations at unobserved locations. The process described below was repeated 15 times to illustrate patterns in the performance of the different methods. The process for each iteration is as follows: A sample of $m = 30$ points was drawn from the population described above. To mimic an actual forest inventory where plots are systematically located, the point locations where derived from a Strauss distribution, where the parameters of the distribution were chosen so that no two points were separated by a distance of less than $25m$. These data were used in three different estimation strategies to quantify possible improvements in spatial prediction with the incorporation of WP data. For all three methods, kriging was performed using the four nearest points ($m' = 4$).

The first method was used to establish a baseline against which WP methods would be compared. Ordinary kriging was used to preform the spatial prediction. The sample of points provided $m(m-1)/2 = 435$ pairs of unique lag distances to which gaussian, exponential, and spherical variograms were fitted. The "best" variogram model, denoted

by $\hat{\gamma}_{OK}(h)$ was chosen using the AICC index. This variogram was then used in conjunction with ordinary kriging to predict the response $\hat{Z}_{OK}(s_0)$ and the estimated mean square prediction error $\hat{\sigma}^2_{OK}(s_0)$ on a new grid of unobserved locations. The subscript OK denotes that the variogram and kriging predictions were derived from the original data with no additional enhancements to the predictor.

The first area of interest is to quantify the improvement in the accuracy of the predictions when the WP data were used in the fitting of the variogram, denoted by $\hat{\gamma}_{WP}(s_0 - s_j)$. To perform this comparison the same 36 points were used, but the location of each tree was mapped and the WP data were derived from each point. The variogram models were then fitted to the WP data. In every case, the exponential variogram was the best fitting model. This variogram was then used in conjunction with ordinary kriging to predict the response $\hat{Z}_{WPV}(s_0)$ and the estimated mean square prediction error $\hat{\sigma}^2_{WPV}(s_0)$ on the same grid of unobserved locations. The subscript WPV denotes "WP variogram" and indicates that the variogram was fitted using the WP data and the prediction was performed using only the original points $s_1, ...s_m$.

The third variant of simulation used $\hat{\gamma}_{WP}(h)$ in conjunction with the four nearest WP points for performing prediction. The predictor is denoted by $\hat{Z}_{WPP}(s_0)$ and the estimated mean square prediction error is denoted by $\hat{\sigma}^2_{WPP}(s_0)$. The subscript WPP denotes "WP prediction" where both the variogram and points used in prediction were derived from the WP data.

One advantage of using the data set described earlier is that the true value $Z(s_0)$ are known for any location within the population. This allowed the performance to be assessed without relying on techniques such as cross-validation. To perform the comparison, a $M \times M = 20 \times 20$ square grid of point locations was used and the true values $Z(s_0)$ and predicted values $\hat{Z}(s_0)_*$ were determined at each grid location and for each of three predictors. Three metrics were used to compare the performance of the different estimation strategies.

The first metric assessed the true mean square prediction error by calculating the sample-based approximation given by

$$\tilde{\sigma}^2_* \approx \frac{1}{M^2} \sum_{grid} (\hat{Z}_*(s_0) - Z(s_0))^2$$

over the grid, where $* = OK, WPV, WPP$. The difference in the MSPE between the three kriging techniques was assessed using the ratio

$$R(OK, WPV) = \frac{\tilde{\sigma}^2_{OK}}{\tilde{\sigma}^2_{WPV}}, R(OK, WPP) = \frac{\tilde{\sigma}^2_{OK}}{\tilde{\sigma}^2_{WPP}}, R(WPV, WPP) = \frac{\tilde{\sigma}^2_{WPV}}{\tilde{\sigma}^2_{WPP}}.$$

The second metric is the average estimated mean square prediction error, which is estimated by

$$\hat{MSPE}(*) \approx \frac{1}{M^2} \sum_{grid} \hat{\sigma}^2_*(s_0).$$

This metric is used to assess how accurately the MSPE is estimated. Two sources of bias are of interest. The first, as mentioned above, is expected to occur because the range of the variogram is often over-estimated in situations where there are no data at short lag distances, so the variogram is interpolated between $h = 0$ and some relatively large lag distance. The second use for this metric is to compare how well the true MSPE is estimated by $\hat{\sigma}^2$ for each technique. To assess the performance of $\hat{\sigma}^2$ for each technique, the ratio of

$$R = \frac{\hat{\sigma}^2_*(s_0)}{\sigma^2_*(s_0)}$$

is determined for $* = OK, WPV, WPP$. A ratio of 1 suggests unbiasedness. However, R is expected to be less than 1 because $\hat{\sigma}^2_*(s_0)$ is calculated without considering the affect of estimating the variogram.

The final metric assessed the performance of the prediction intervals by comparing the observed and nominal coverage rate over the 400 grid locations. Five different methods were compared. The first three methods compared the performance of the standard plug-in prediction interval, $I(s_0)$ using the data from the OK, WPV, and WPP sampling and modeling approaches. The final two methods used the nonparametric prediction interval, $I_{WP}(s_0)$. only the WPV and WPP methods were compared because the WP data would not be available for OK sampling. The notation WPV, NP and WPP, NP will be used to denote that prediction intervals will be determined using the nonparametric technique described above. All methods were compared by calculating $1 - 2\alpha = 0.95$ prediction interval at each grid location and determining if the true value fell with interval. Then the true coverage rate was assessed by comparing by determining what percentage of the predicted $\hat{Z}(s_0)$ values fell within the prediction intervals.

3.7 Results

The first area of interest is to determine whether the use of the two WP methods leads to significant improvements in prediction. The degree of improvement was quantified by the average mean square prediction error, given by $\tilde{s}^2$, from each iteration of the simulation study. These values are given in Figure 10. The figure shows that WPP kriging always reduced in MSPE when compared to both OK and WPV. The average reduction in MSPE over the 15 iterations was 19 %. The results are not quite as dramatic when comparing OK to WPP, with the predictions generated by WPV exhibiting little or no improvement over the the OK method for 7 out of 15 samples. The average reduction in MSPE using WPV over OK across all 15 iteration was 5%. When comparing the two WP kriging techniques, a reduction in MSPE average reduction in the MSPE was 13.6 %. Thus, for this application the greatest improve in performance was realized by using the closer prediction points on the boundary of the WPs.

The overall performance of the kriging variance estimator was assessed by comparing the achieved coverage rates of the estimated prediction intervals with their nominal coverage rate of 95% (Figure 11). In every case, all methods failed to achieve their nominal coverage rates. However, the performance of the OK intervals was especially poor, with coverage rates as low as 57% and an average coverage rate of only 74.3%. In contrast the average achieved coverage rates for WPV and WPP kriging were 91.6 and 91.2%, respectively. The use of the nonparametric prediction interval increased the coverage rates to 92.6 and 91.8 %, for the WPV,NP and WPP,NP methods, respectively. Some degree of under-coverage is expected for all of the techniques because the estimation of the variogram is not accounted for in the estimated MSPE. Another possible explanation is that roughly 1 % of the validation data set is comprised of outlier values, which were caused by plots that did not sample any trees. This occurs most often along the edge of the population and these outliers were also chosen for prediction points. In contrast, less that 0.3 % of plots in the original data set were outliers.

The very poor performance of the OK prediction intervals is at least partially explained by the range of the variogram being consistently over-estimated. This bias causes the

estimated MSPE to be biased downward. To assess the level of bias, the ratio of the estimated to the achieved MSPE were graphed. (Figure 12). For only 3 of the 15 samples was the OK estimated MSPE within 10 % of the estimated true prediction error. On average, the OK estimated MSPE was only 65 % the true MSPE. In contrast, WPV and WPP estimated prediction errors actually tended to slightly overestimate the true value, with the over-estimates averaging 7.5 and 8 %, respectively.

Another conclusion drawn from the simulation study is that the erratic behavior of the OK predictor suggest that the sample size of $m = 30$ is likely too small for reliable inference in this application without the use of WPs. However, it seems unlikely that higher sampling rates would occur in this application.

3.8 Discussion and Conclusions

The motivation for this study is that literally thousands of forest surveys, similar to the one depicted in Figure 1, are performed every year by both public and private organizations. If these data are used for kriging, the lack of sufficient data at short lag distances can result in a variogram that is poorly estimated. WPs are an efficient alternative to augmenting the data with auxiliary plots (Figure 2). While it is unrealistic to assume that the original plots will always be sufficiently large to effectively apply the WP approach and that tree location will be recorded, there is no reason why an annular plot (i.e., a ring) could not be added to a sub-sample of the plots. For this sub-sample, the tree locations would be recorded and the data for fitting the variogram would be extracted using the WP technique. The only draw-back to this approach is that all spatial analyses must be performed at the scale determined by the remaining points in the sample.

The results of the simulation study raise some additional issues. Stein's (1988) results on mis-specified variograms suggest that for efficient prediction, the behavior of the variogram near the origin must be adequately modeled. Cressie (1993, p. 134) points out that it is difficult to judge the relevance of these asymptotic results in an applied setting because it is not possible to scale Stein's results to a finite-data setting. The results of this study suggest that while both of these points of view are valid, the advantage of having data at short lag distances is that, by providing a good estimate of the slope of the variogram

at short lag distances, the potential for over-estimating the range parameter of the variogram is reduced or eliminated. A bias in the estimated range of spatial autocorrelation not only leads to a predictor that is sub-optimal, but the resulting prediction intervals will on average be far too narrow to achieve their nominal coverage rate.

A number of authors have pointed out the shortcomings of kriging (Gooverats 1997, p. 442) and the data set used in this study is probably best described as "challenging" for any parametric estimation method. However, the one surprising result of this study is how well the simple "plug-in" based methods performed when used in conjunction with the WP methods. Therefore, it would appear that even the least sophisticated kriging methods can be quite robust, provided the variogram is accurately estimated.

3.9 Literature Cited

Bellhouse, D. R. 1977. Some optimal designs for sampling in two dimensions. Biometrika 64:605-611.

Cochran, W. G. 1946. Relative accuracy of systematic and stratified samples from a certain class of populations. Annals of Mathematical Statistics. 17.164-177.

Cressie. N. 1993. Statistics for Spatial Data. J. Wiley & Sons, New York.

Diggle, P. J., Tawn, J. A., and Moyeed, R. A. 1998. Model-based geostatistics (with discussion), Applied Statistics. 47:299-350.

Husch, B., Miller, C.I., Beers T.W., 1982. Forest Mensuration 3rd edn. J. Wiley and Sons, New York.

Goovaerts. P. 1997 Geostatistics for natural resources evaluation. Oxford University Press, New York.

Gotway. C. A., and Young, L. J. 2002. Combining incompatible spatial data. JASA 97:632-648.

Harville, D. A., and Jeske, D. R. 1992. Mean square error of estimation and prediction under a general linear model. JASA 853-862.

Iachan, R. 1985. Plane sampling. Statistics and Probability Letter. 50:151-159.

Kallas, M. A., Reich, R. M., Jacobia, W. R., Lundquist, J. E. 2003. Modeling the probability of observing Armillaria root disease in the Black Hills. Forest Pathology. 33:241252

Mandallaz, D. 2000. Estimation of the spatial covariance in Universal Kriging: Application to forest inventory. Environmental and Ecological Statistics. 7:263-284.

Moyeed, R.A. and Papritz, A. 2002. An empirical comparison of kriging methods for nonlinear spatial point prediction. Mathematical Geology. 34:365-386.

Müller and Zimmerman 1999. Optimal designs for variogram estimation. Environmetrics. 10:23-37.

Olea, R. A., 1984. Sampling design optimization for spatial functions. Mathematical Geology. 16:369-392.

Openshaw, S. U., and Taylor P.J. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In Wrigley, N. (ed): Statistical applications in the spatial science. Pion Limited, London. pp. 127-144.

Papritz, A., and Moyeed, R.A. 1999. Linear and nonlinear kriging kriging methods: Tools for monitoring soil pollution. in:Barnett, V., Turkman, K. and Stein, A. (eds). Statistics for the environment. Vol. 4: Statistics of health and environment: Wiley pp. 303-336.

Ritter, K. J., and Leecaster, M. 2005. Spatial enhancement of grids for the estimation of a variogram in near coastal systems. Submitted to Envir. Ecolo. Stat.

Sjöstedt-de Luna, S and Young, A. 2003. The bootstrap and kriging prediction intervals. Scandanavian J. Stat. 30:175-192.

Stein, M. L. 1988. Asymptotically efficient prediction of a random field with a misspecified covariance function. Ann. Statistics. 16:55-63.

Stein, M. L. 1999. Interpolation of spatial data: Some theory for Kriging. Springer, NY

Urban, D.L. 2005. Strategic monitoring of landscapes for natural resource management. In J.L. Liu and W.W. Taylor (eds.), Integrating landscape ecology into natural resource management. Cambridge University Press (in press).

Wallerman, J. 2003. Remote sensing aided spatial prediction of forest stem volume. Doctorial Thesis. Swedish University of Agricultural Sciences, Umeå

Wallerman, J., Joyce, S., Vencatasawmy, C., and Olsson, H. 2002. Prediction of forest stem volume using kriging adapted to detected edges. Can. J. For. Res. 32: 509-518

Wang, F. and Wall, M. M. 2003. Incorporating parameter uncertainty into prediction intervals for spatial data modeling via a paraetric variogram. Journal of Agricultural, Biological and Environmental Statistics. 8:296-309.

Williams, M. S., and Reich, R. M. 2005. Wobble-Plot Sampling: A Technique for Estimating Small-Scale Spatial Variability in Environmental Surveys. Ecol. Modeling. (in review).

Zimmerman, D. L., and Cressie, N. 1992. Mean square prediction error in the spatial linear model with estimated covariance parameters. Annals of the Institute of Statistical Mathematics. 44:27-43.

3.10 Figure Captions

- Figure 1. Example of a typical stand level forest inventory. A systematic sample of plot locations is chosen and a single plot is installed at each location.

- Figure 2. One possible cluster plot design used specifically for the estimation of the variogram (adapted from Urban 2005).

- Figure 3. In (a) two concentric plot measuring $r = 15$ and $r = 7.5m$ are depicted. Six and two of the seven trees are selected on the $r = 15$ and $r = 7.5m$ plots, respectively. In (b) the object-centered approach is taken. Note that only those trees that fell within $15m$ of the point would contribute to the response surface derived from $r = 7.5m$ plots.

- Figure 4. Overview of the dataset. A number of small openings exist where the plots are not sufficiently large to sample any trees.

- Figure 5. Comparison of the distribution of the response $Z(s)$ against a standard normal distribution.

- Figure 6. The set of data points derived from the $m = 36$ wobble-plots. The spacing used between points within each cluster was $h = 0.7m$ and there were nominally $m' = 368$ points per plot with which to estimate the variogram.

- Figure 7. An example of the classical variogram derived from the WP data. A gaussian variogram model has been fitted to the data. Note that the form the gaussian model does not fit the data for $h \approx 0$.

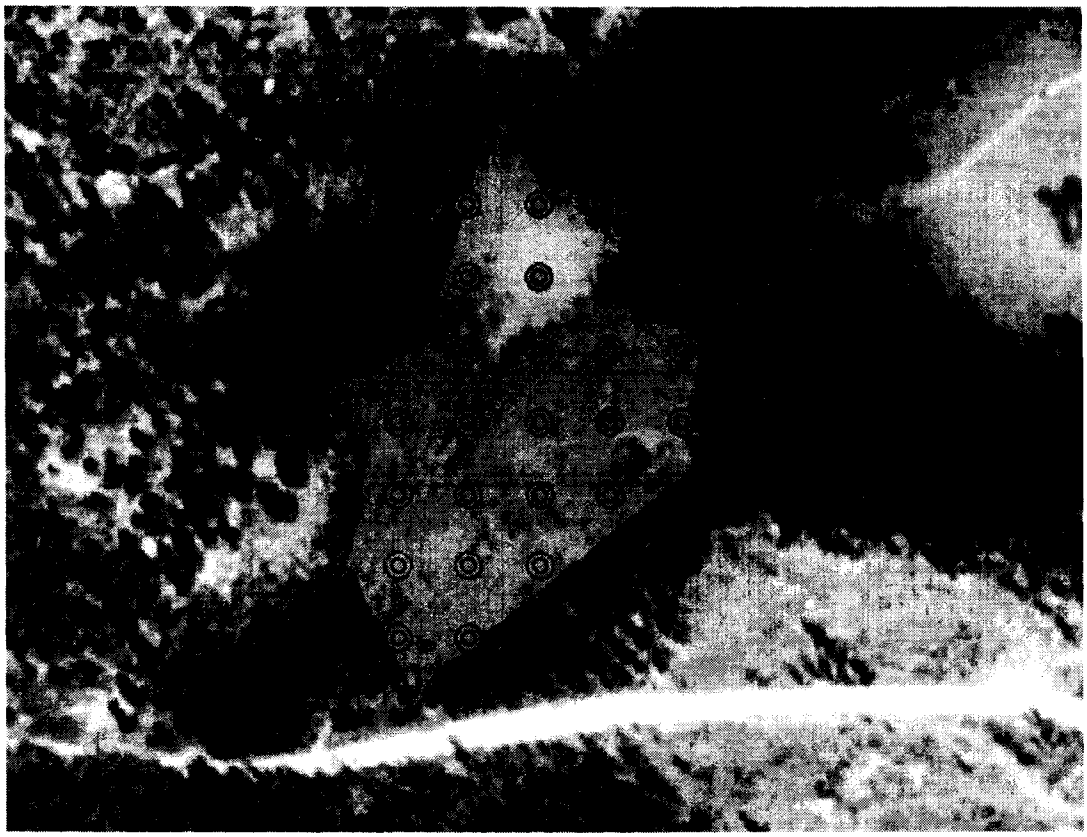
- Figure 8. A different example of the variogram gram cloud with the data from the m original point concatenated to the WP data. Two variograms are displayed. One is the variogram is fitted with all the data WP. The second variogram is the one fitted using just the original $m = 36$ points.

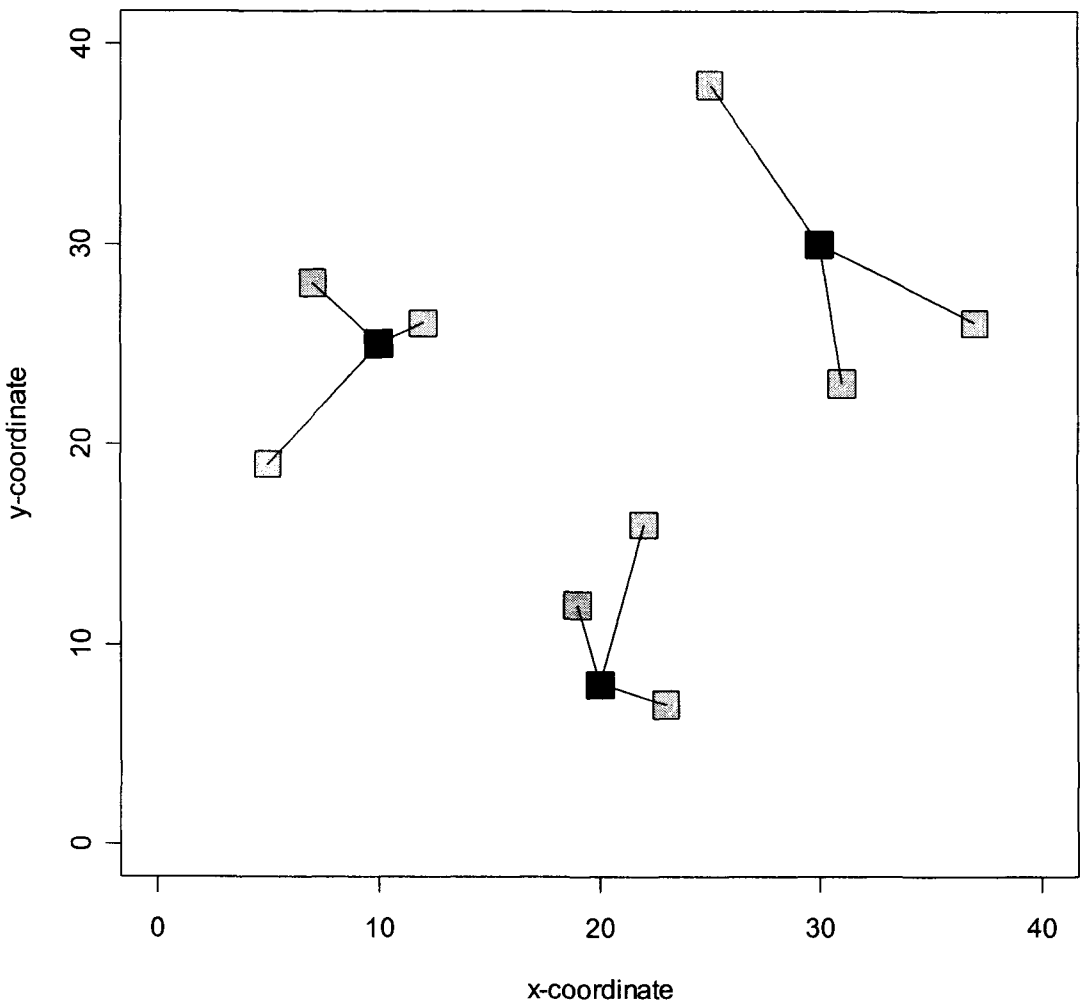
- Figure 9. Prediction is performed by using closest WP data point on the boundary of the wobble plot.

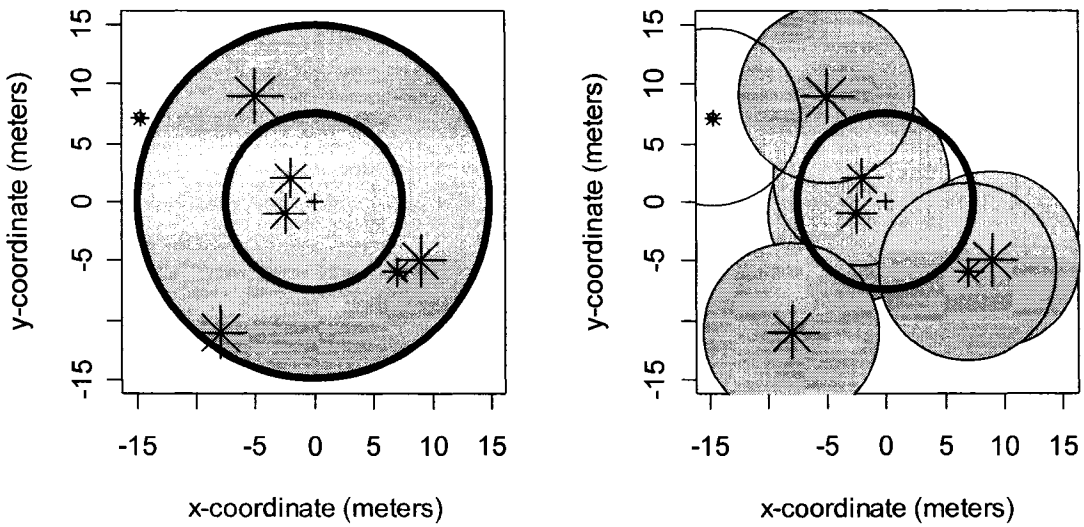
- Figure 10. The average mean squared prediction error from each iteration of the simulation study.

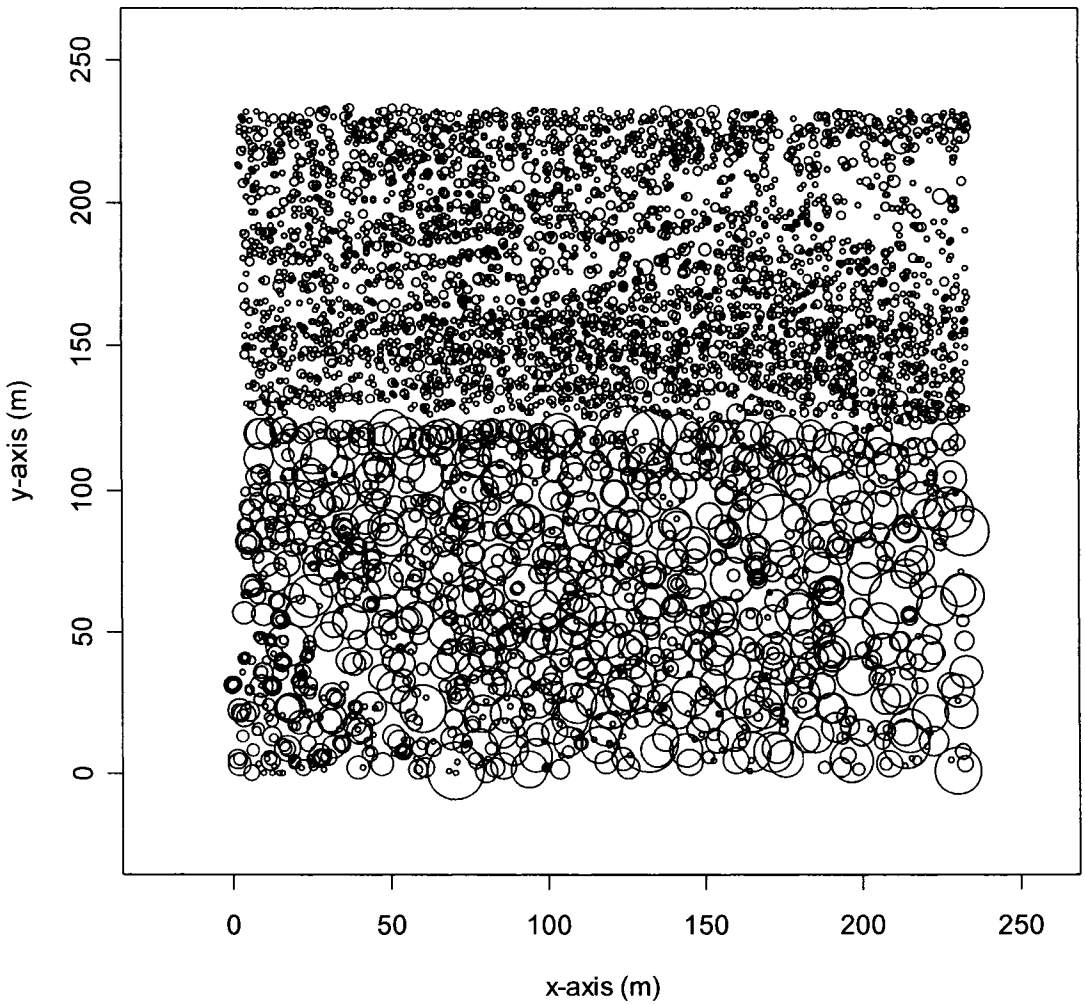
- Figure 11. The achieved coverage rate for each of the three predictors. While none of the methods achieves their nominal coverage rate, the WP based methods are perform markedly better than the standard predictor.

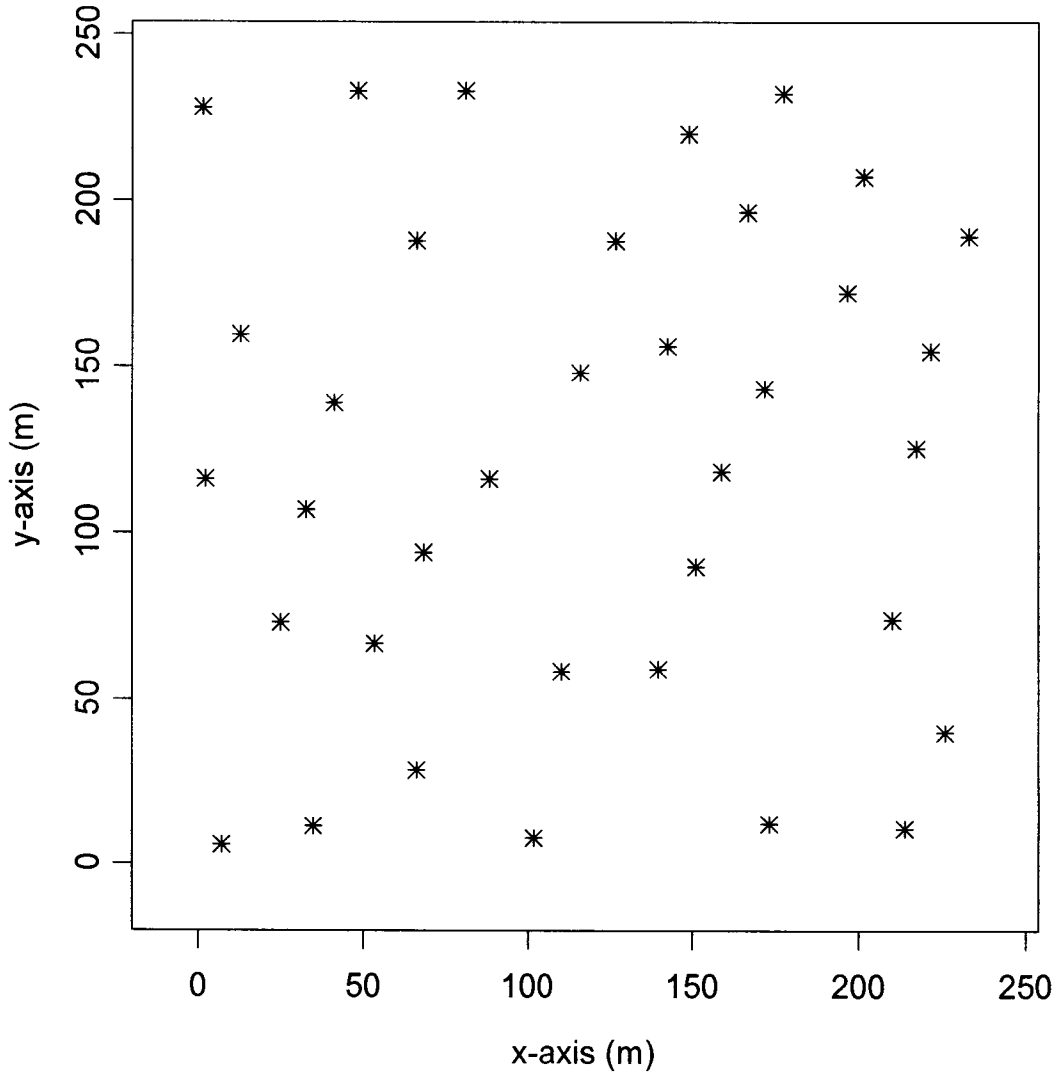
- Figure 12. The ratio of estimated to "true" mean squared prediction error for each of the three methods. The OK method under-estimated the true MSPE for every sample. The WP based methods tended to over-estimate the true MPSE.

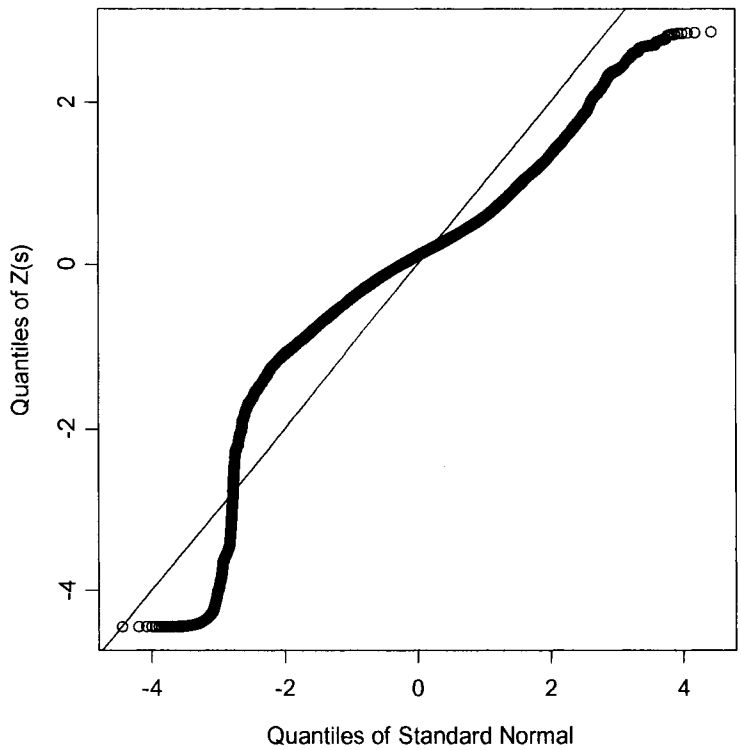


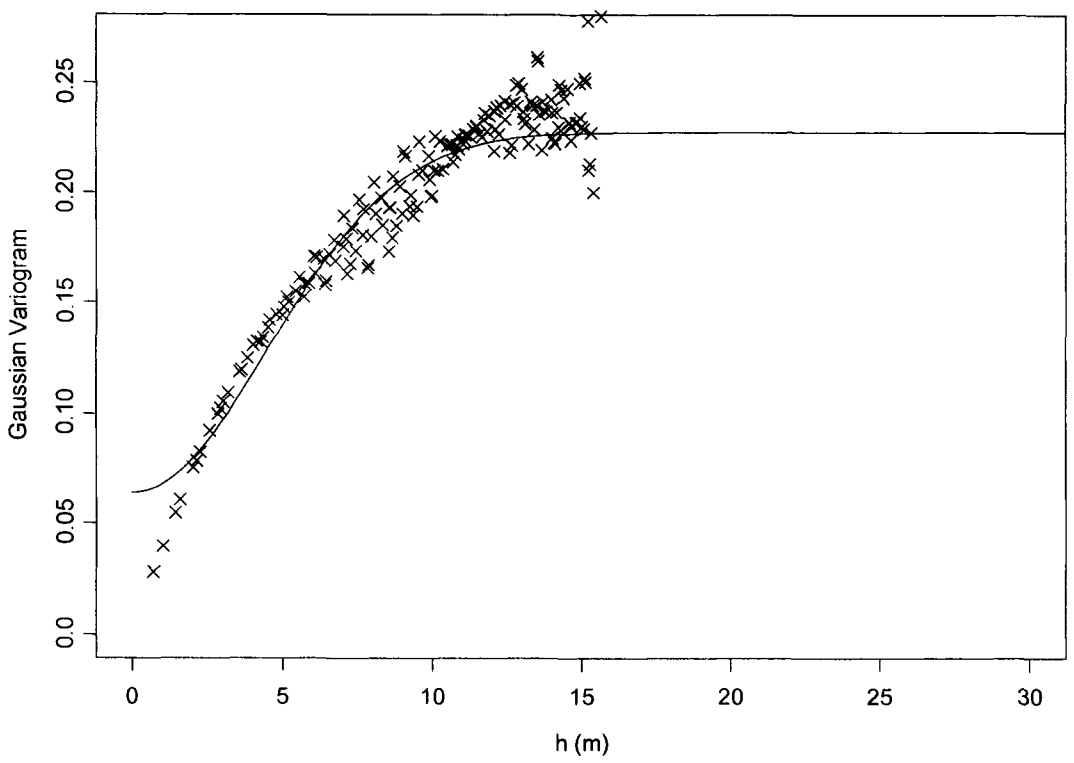


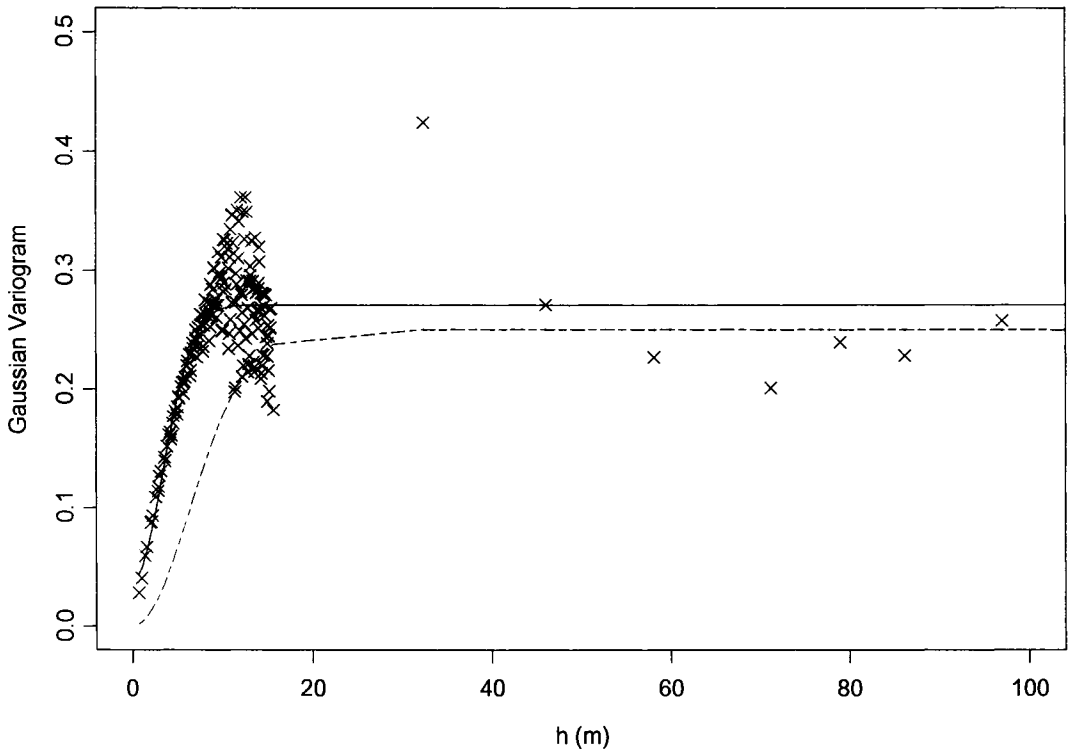


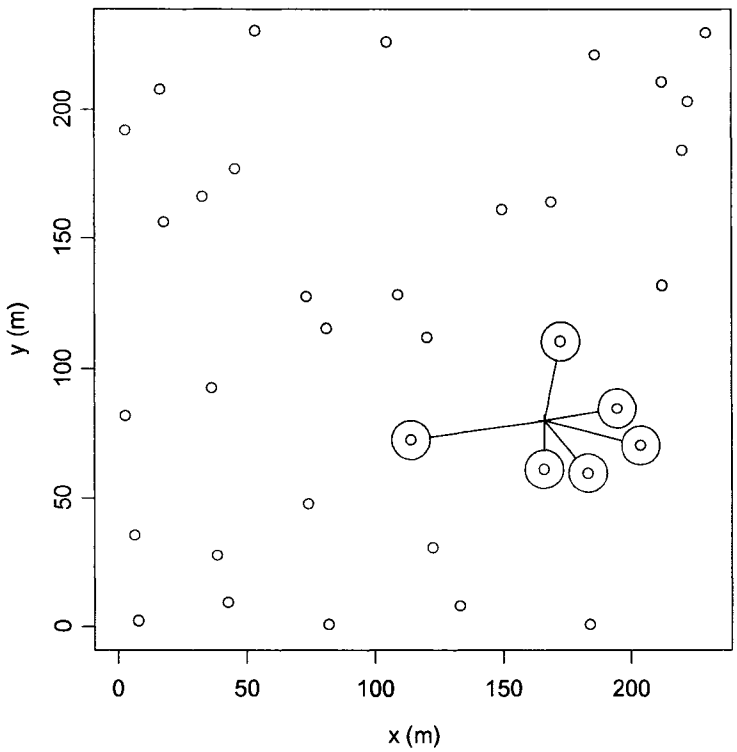


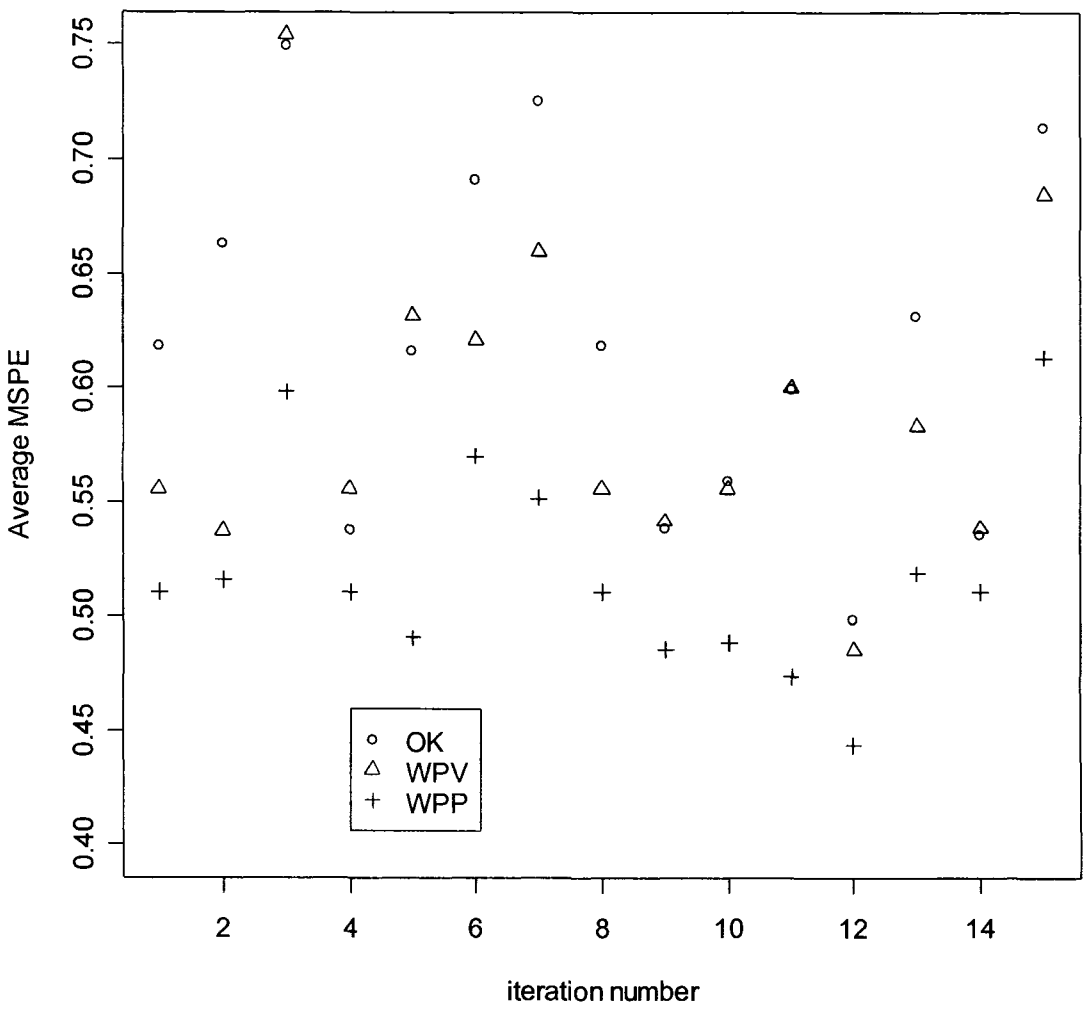


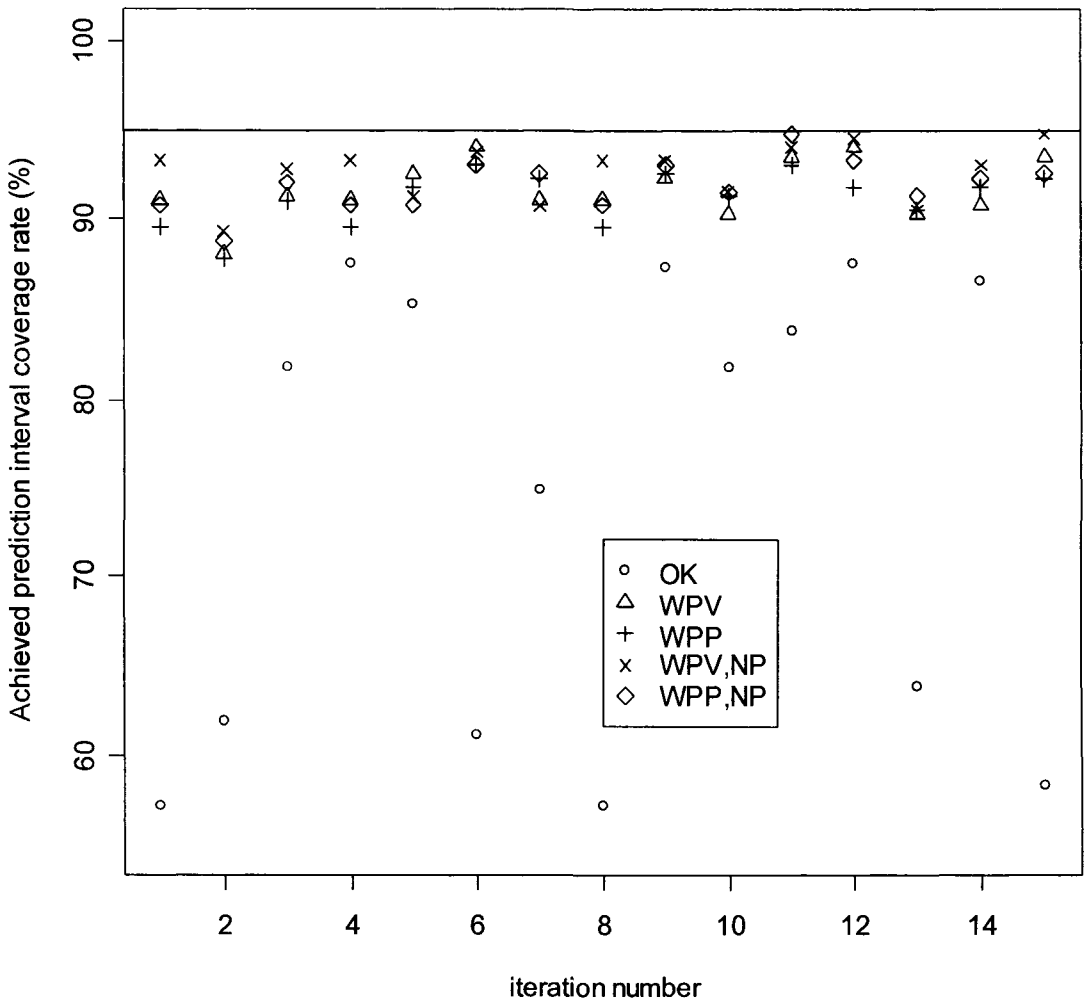


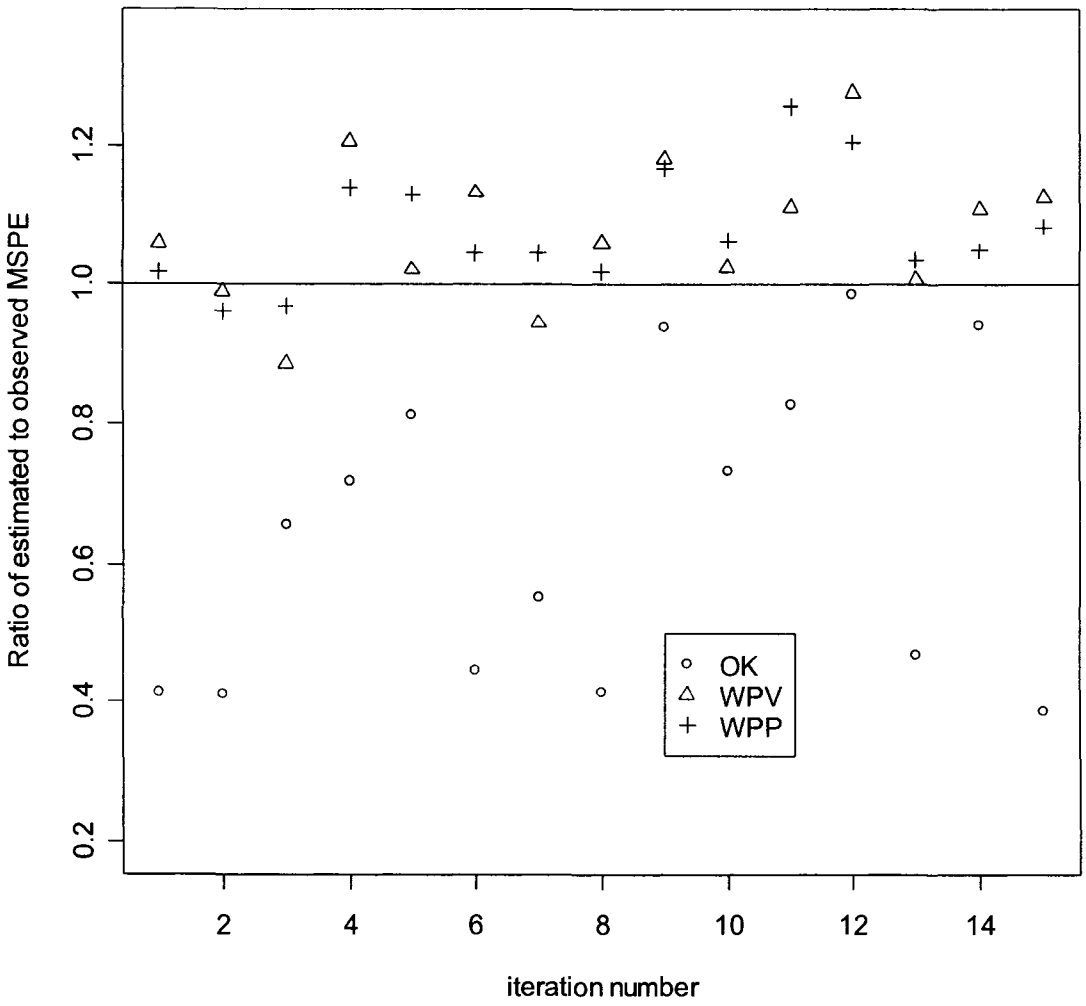












Chapter 4

SUMMARY OF THE STUDIES

The results for each study are briefly summarized.

4.1 Study I:

Many environmental surveys and virtually every large-scale forest inventory in the world utilizes open-cluster plots for the collection of field data, where these are plot configurations that comprise a systematically arranged group of smaller plots or points. While these plots are an efficient way to collect data when sampling from an area frame that tessellated the population into primary and secondary sampling units, their performance in the continuous population setting will almost surely be poor unless the target population is comprised of large patches with a very high interior-to-edge ratio.

The first goal of this study was to identify factors that should be considered in the design of a continuous population survey. The three key factors that were identified are;

1. The method of boundary correction when sampling along the boundary.
2. The trade-off between within- and between-cluster variability.
3. The average area of the plot that covers the population.

The study links these factors by noting that each factor is at least partially related to the level of fragmentation in the forest population.

The level of fragmentation in the US has been studied extensively by landscape ecologists. The results of these studies indicate that there are few area in the lower 48 states where large contiguous patches of forest exist. The study uses these results in the construction of two artificial populations; with the first having an edge-to-interior ratio that

would be typical in heavily fragmented areas and the second having a significantly lower edge-to-interior ratio. A simulation study was preformed to compare the performance of a number response designs based on fixed-area plots and to rank the importance of the three factors identified above. In every case, the boundary-correction method was the single most important factor. While this result is of interest, the more important issue is that boundary correction methods for open-cluster plot either expensive of impossible to implement in the field. From these results it is concluded that response designs based on a single plot or point and used in conjunction with the walkthrough boundary-correction technique are generally the most efficient and practical.

4.2 Study II:

The guidelines for the appropriate sampling and response design established in Study I constitute what is probably the worst possible design for a spatial modeling application. The underlying problem is that the sampling design that meets the sample survey objectives ensures that no data are collected at short lag distances.

While the traditional solution to this problem has been to augment the sample with additional points, this study shows that by using a simple change of scale and re-sampling the plot information, an infinite number of points are available for the modeling the variogram at short lag distances. This technique, for lack of a better term, is referred to as wobble-plot sampling. These data are then combined with the data that are derived from the center-point of each plot to estimate the variogram.

The improvement in the resulting variogram models was assessed using a simulation study. The study compared the fitted variogram models using three different sampling designs. The first was a sample of 200 randomly located plots; the second was a systematic sample of 36 plots; the third was a sample of the same 36 plots, except that roughly 2 million additional points were extracted using the wobble-plot technique. The results showed that the wobble-plot derived variograms were much less variable and than the variograms derived from the other two methods. An additional point of interest was the bias observed in the range parameter when only the 36 points were used to fit the variogram. On average, the

estimated range parameter was more than twice range estimated by the other two methods. The bias is attributed to the lack of observations with short lag distances.

4.3 Study III:

The variogram is generally only the first step in most spatial analyses, though there are exceptions. The overall goal is usually the prediction of the attribute at unobserved locations using one of the many forms of kriging (e.g., block, simple, ordinary, disjunctive, model-based, median polish, and co-kriging to name a few).

The previous study illustrates how data gathered with wobble-plot sampling improved all aspects of the estimated variogram (i.e., reduction of the variance in the parameter estimates and the variogram itself, more consistent model selection, and reduced bias in the range parameter). However, the previous study provided no information regarding how improvements in the variogram affect the predictive ability of kriging. A related area interest is determining the affect of performing spatial prediction when there is insufficient data for valid inference. This occurs in many forestry applications because there is often little or no data collected at small lag distances.

The first goal of this study was to assess the affect the improved variogram on the spatial prediction of attributes such as basal area. This topic was of interest because asymptotic results suggest that the use of wobble-plot sampling would lead to "better" spatial predictions, but the magnitude of the improvement cannot be quantified. The second goal of the study was to determine how some of the limitations of ordinary kriging could be addressed by the data derived from the wobble-plots. Two areas were identified, with the first being improved spatial prediction and the second being improvements in the prediction interval coverage rate.

To improve spatial prediction, the data from the boundaries of the wobble-plots were incorporated into the prediction of the response surface at unsampled locations. The motivation for this approach is, that by utilizing these points, the estimated MSPE for any new point will always be less than or equal to estimated MSPE derived from using the original point locations. Improvements in the prediction interval coverage rates were achieved in two ways. The first is that standard plug-in prediction interval do not account for effect of

estimating the parameters of the variogram. This will cause the achieved coverage rate of the prediction intervals to be less than the nominal coverage rate. It was expected that by reducing the variability in the estimated parameters, this bias would be reduced.

Another weakness of using "plug-in" prediction intervals is that the upper and lower bounds are determined by assuming that the residuals are normally distributed. To address this issue, a nonparametric method for choosing the upper and lower bounds of the intervals was proposed.

The improvements in performance provided by these modifications was assess in the same simulation as used in the previous study. The baseline for comparison was to use a systematic sample of 30 points to develop the variogram and predict the response at 200 unobserved locations using ordinary kriging. The known values at the 200 locations were compared to the predicted values to determine the integrated MSPE and to determine the achieved coverage rates of 95 % prediction intervals.

The MSPE of the baseline method was compared to two methods that utilized the wobble-plot data. The first comparison was between the baseline method and the predictor that used the variogram derived from the wobble-plots. The second comparison was between the baseline method and the predictor that used both the improved variogram and the prediction points on the boundary of the wobble-plots. The simulation results showed that the improved variogram reduced the MSPE by 5 % on average. Using both the improved variogram and the new prediction points reduced the MSPE 20 %.

The achieved coverage rates of the 95 % prediction for the baseline method were compared to three different methods that utilized the wobble-plot data. As in the MSPE comparison, the first two methods used the improved variogram and the improved variogram in conjunction with the new prediction points. The last method used the improved variogram, the additional prediction points, and the upper- and lower- bounds of the intervals were determined using a nonparametric approach.

The simulation results showed that the achieved coverage rate for the baseline approach was on average only 74 %. In contrast the achieved coverage rates for the new methods average between 91 % to 93 %. Thus, not only are the intervals roughly 5 to 20intervals nearly achieved their nominal coverage rate.

4.4 Future Research

The results related to the sample survey objectives (Study I) are fairly clear-cut. These results suggest that for the purpose of estimating population means and totals, there are likely to be very few situations in which the use of open-cluster plots will be beneficial. While this study uses a response design consisting of sampling with fixed-area plots, there is no reason to expect that these same results do not hold for other response designs (e.g., variable radius- and line-intersect sampling). In fact, it is rather obvious that boundary-correction techniques for alternative response designs tend to be far more complicated (e.g., Affleck et al. 2005) if not completely impractical. Thus, areas where additional research is needed are generally only related to practical issues of implementation. A partial list of the issues that determine the appropriate ground plot configuration is:

- The trade-off in variance reduction associated with increasing plot size versus the cost of visiting additional points. This is generally related to the cost of data collection in terms of the amount of time needed to collect the data at a single point versus the travel time between points.

- The compatibility of the response design in relation to the size and type of remotely sensed data that might be used in other analyses.

- The compatibility of the response designs for assessing and relating multiple resources. For example, which sampling strategies for standing trees and soil characteristics have the most highly correlated estimators?

In the context of a large-scale inventory, such as FIA, where round-trip travel times to plot locations can exceed two to three hours, it seems likely that the most appropriate plot design for sampling standing trees is a fixed-area plot whose size is as large as possible while still allowing field crews to collect all the necessary data (i.e., both standing tree and other resource data) in four to six hours.

There are both practical and theoretical areas for additional research related to spatial modeling. The first problem is to note that the observed process is an artifact of the

pragmatic nature of how data are collected. Specifically, note that for an attribute such as basal area or volume, the amount of tree volume within the plot boundary is actually a continuous function over the population. However, in practice the function describing these attributes is a step function because trees are only included in the sample when the center of the tree falls within the plot boundary. Thus, the observed process is not as smooth as the assumed underlying variogram models and that we are doing some aggregation that raises some issues. The general point is that we are estimating $Cov[Z(s), Z(s+h)]$, where $Z(s)$ is the observed response at s and $s+h$. However, in a "counting trees" application, what we are interested in is determining if

$$Cov[Z(s), Z(s+h)] = Cov[\int_{N(s)} \varsigma(t)dt, \int_{N(s+h)} \varsigma(u)du] = \int_{N(s)} \int_{N(s+h)} Cov[\varsigma(t), \varsigma(u)]dtdu,$$

where ς is actually generated from an exponential, Matern, or Gaussian process. While this is a serious concern, it should be remembered that one of the goals of these studies was to build the foundation of the survey on a sampling- and response design under which the proposed estimators would be design-unbiased and the spatial modeling component would be used primarily to describe the relative distribution of the resource. Where this issue of aggregation becomes a more serious problem would be in situations where estimation hinged entirely on the spatial model. Regardless of the application, it would be interesting to see how closely a process that was generated from an exponential variogram would be modeled by wobble plot data.

Another issue for the use of wobble-plots sampling is the methodology used to draw a finite sample from the infinite set of the available data on each plot. In this study, the sample was drawn from a square grid imposed on each plot. The grid spacing was chosen because it appeared to be sufficiently fine and the amount of time and memory required to complete the simulation study was reasonable. However, using a square spacing limits the number of points whose lag distance is near the diameter of the plot. Additional unresolved issues are;

- Are there any tangible advantages to estimating the variogram using a closed-form expression that accounts for the infinite number of available observations?

- What are the advantages and approaches for using a grid with some form of adaptive spacing?
- How does one incorporate points derived from neighboring wobble-plots?
- Adding all the data points from neighboring plots would rapidly overburden any computing environment, so how does one choose a reasonable subset additional points?
- What is the most effective method for allocating new points whose primary purpose is for spatial prediction?

While the second study shows that it is possible to estimate the variogram at short lag distances with single plots, the utility of the wobble-plot method alone is limited to the modeling of spatial processes where the range of the data is "similar" to both the size of the plot and the spacing between existing sample points. Thus, in its current incarnation it could be useful tool for the management of areas that range in size from hundreds to many thousands of hectares, provided the sampling locations are sufficiently "close". This would generally hundreds or thousands of sample points within each stand. Examples of this type of inventory system exist in Sweden and other European countries (Wallerman et al (2003)). and some privately held commercial forest lands in the US, Mexico, and Canada. However, the utility of this approach is limited for large-scale environmental inventories, where the spacing between points is can exceed 10 kilometers. For these types of applications one could argue that it may be beneficial for estimation of the nugget parameter for applications where the lag distance is significantly larger than the plot. However, it seems unlikely that significant improvements in spatial predictions would be realized. The only solution for the spatial prediction of attributes with a "short" range parameter is to use some form of remotely sensed data to "fill in" the significant gaps between ground plots. Thus, an area for future research consists of studying how the variograms derived from the ground data could be incorporated with various forms of remotely sensed data.

The final study attempts to quantify the magnitude of the improvements derived from the use of wobble-plot data. This is done by first assessing the reduction in MSPE and the improvements in prediction interval coverage rates. Then a number of fairly simplistic

modifications are made to ordinary kriging in order to take advantage of the additional data provided by the wobble plots. This leads to a host of additional research areas where these data may be of value. Some possible areas include;

- Much of the theoretical work on kriging (e.g. Stein 1999) assumes in-fill asymptotics. However, there are few applications where the sample size is sufficient to argue that these asymptotic results apply. Wobble-plot data may provide a reasonable setting in which to study the asymptotic behavior of various kriging predictors.

- How to best utilize the wobble-plot data in the prediction process by determining the "best" subset of points to include in the prediction.

- The data set illustrates the potential inadequacy of a Normality assumption for determining the bounds of the predictive interval. It would seem that resampling methods could be adapted for this application.

- The use of wobble-plot data in a Bayesian setting may be difficult. The reasons being that the volume of data is too large for Monte Carlo estimation approaches and that the data set used suggests that the use of a parametric model to describe the distribution of the residuals will be difficult for this application.

4.5 Conclusions

This study has attempted to address many of the issues related to choosing an appropriate ground plot for a multi-objective forest inventory or environmental survey. In terms of the sample survey objectives, the most appropriate plot configuration consists of a single plot or point and these points will be spread across the study area using some form of systematic sample. This response design is thought to be inappropriate for spatial modeling applications because no data are collected at short lag distances. Wobble-plot sampling is developed to show that within the constraints imposed by the sample survey objectives, it is possible to simultaneously collect much, if not all, the data necessary for fitting the variogram. Further extensions and modifications of wobble-plot sampling are shown to improve the performance of ordinary kriging.

An interesting aspect of these studies is that it can be argued that environmental sampling is a mature field with little room for significant improvement. However, the results of this study suggest that many of the commonly used techniques in environmental surveys are inefficient. A rather surprising result is that some of the greatest gains in overall efficiency, as measured by sampling effort and mean squared error, are likely to come from the selection of a ground plot configuration that meet the many analytical objectives of the survey.

4.6 Future Research

The results related to the sample survey objectives (Study I) are fairly clear-cut. These results suggest that for the purpose of estimating population means and totals, there are likely to be very few situations in which the use of open-cluster plots will be beneficial. While this study uses a response design consisting of sampling with fixed-area plots, there is no reason to expect that these same results do not hold for other response designs (e.g., variable radius- and line-intersect sampling). In fact, it is rather obvious that boundary-correction techniques for alternative response designs tend to be far more complicated (e.g., Affleck et al. 2005) if not completely impractical. Thus, areas where additional research is needed are generally only related to practical issues of implementation. A partial list of the issues that determine the appropriate ground plot configuration are;

- The trade-off in variance reduction associated with increasing plot size versus the cost of visiting additional points. This is generally related to the cost of data collection in terms of the amount of time needed to collect the data at a single point versus the travel time between points.

- The compatibility of the response design in relation to the size and type of remotely sensed data that might be used in other analyses.

- The compatibility of the response designs for assessing and relating multiple resources. For example, which sampling strategies for standing trees and soil characteristics have the most highly correlated estimators?

In the context of a large-scale inventory, such as FIA, where round-trip travel times to plot locations can exceed two to three hours, it seems likely that the most appropriate plot design for sampling standing trees is a fixed-area plot whose size is as large as possible while still allowing field crews to collect all the necessary data (i.e., both standing tree and other resource data) in four to six hours.

There are both practical and theoretical areas for additional research related to spatial modeling. The first problem is to note that the observed process is an artifact of the pragmatic nature of how data are collected. Specifically, note that for an attribute such as basal area or volume, the amount of tree volume within the plot boundary is a actually a continuous function over the population. However, in practice the function describing these attributes is a step function because trees are only included in the sample when the center of the tree falls within the plot boundary. Thus, the observed process is not as smooth as the assumed underlying variogram models and that we are doing some aggregation that raises some issues. The general point is that we are estimating $Cov[Z(s), Z(s+h)]$, where $Z(s)$ is the observed response at s and $s+h$. However, in a "counting trees" application, what we are interested in is determining if

$$Cov[Z(s), Z(s+h)] = Cov[\int_{N(s)} \varsigma(t)dt, \int_{N(s+h)} \varsigma(u)du] = \int_{N(s)} \int_{N(s+h)} Cov[\varsigma(t), \varsigma(u)]dtdu,$$

where ς is actually generated from an exponential, Matern, or Gaussian process. While this is a serious concern, it should be remembered that one of the goals of these studies was to build the foundation of the survey on a sampling- and response design under which the proposed estimators would be design-unbiased and the spatial modeling component would be used primarily to describe the relative distribution of the resource. Where this issue of aggregation becomes a more serious problem would be in situations where estimation hinged entirely on the spatial model. Regardless of the application, it would be interesting to see how closely a process that was generated from an exponential variogram would be modeled by wobble plot data.

Another issue for the use of wobble-plots sampling is the methodology used to draw a finite sample from the infinite set of the available data on each plot. In this study, the sample was drawn from a square grid imposed on each plot. The grid spacing was chosen

because it appeared to be sufficiently fine and the amount of time and memory required to complete the simulation study was reasonable. However, using a square spacing limits the number of points whose lag distance is near the diameter of the plot. Additional unresolved issues are;

- Are there any tangible advantages to estimating the variogram using a closed-form expression that accounts for the infinite number of available observations?
- What are the advantages and approaches for using a grid with some form of adaptive spacing?
- How does one incorporate points derived from neighboring wobble-plots?
- Adding all the data points from neighboring plots would rapidly overburden any computing environment, so how does one choose the a reasonable subset additional points?
- What is the most effective method for allocating new points whose primary purpose is for spatial prediction?

While the second study shows that it is possible to estimate the variogram at short lag distances with single plots, the utility of the wobble-plot method alone is limited to the modeling of spatial processes where the range of the data is "similar" to both the size of the plot and the spacing between existing sample points. Thus, in its current incarnation it could be useful tool for the management of areas that range in size from hundreds to many thousands of hectares, provided the sampling locations are sufficiently "close". This would generally hundreds or thousands of sample points within each stand. Examples of this type of inventory system exist in Sweden and other European countries (Wallerman et al (2003)), and some privately held commercial forest lands in the US, Mexico, and Canada. However, the utility of this approach is limited for large-scale environmental inventories, where the spacing between points is can exceed 10 kilometers. For these types of applications one could argue that it may be beneficial for estimation of the nugget parameter for applications where the lag distance is significantly larger than the plot. However, it seems unlikely that

significant improvements in spatial predictions would be realized. The only solution for the spatial prediction of attributes with a "short" range parameter is to use some form of remotely sensed data to "fill in" the significant gaps between ground plots. Thus, an area for future research consists of studying how the variograms derived from the ground data could be incorporated with various forms of remotely sensed data.

The final study attempts to quantify the magnitude of the improvements derived from the use of wobble-plot data. This is done by first assessing the reduction in MSPE and the improvements in prediction interval coverage rates. Then a number of fairly simplistic modifications are made to ordinary kriging in order to take advantage of the additional data provided by the wobble plots. This leads to a host of additional research areas where these data may be of value. Some possible areas include;

- Much of the theoretical work on kriging (e.g. Stein 1999) assumes in-fill asymptotics. However, there are few applications where the sample size is sufficient to argue that these asymptotic results apply. Wobble-plot data may provide a reasonable setting in which to study the asymptotic behavior of various kriging predictors.

- How to best utilize the wobble-plot data in the prediction process by determining the "best" subset of points to include in the prediction.

- The data set illustrates the potential inadequacy of a Normality assumption for determining the bounds of the predictive interval. It would seem that resampling methods could be adapted for this application.

- The use of wobble-plot data in a Bayesian setting may be difficult. The reasons being that the volume of data is too large for monte carlo estimation approaches and that the data set used suggests that the use of a parametric model to describe the distribution of the residuals will be difficult for this application.

4.7 Conclusions

This study has attempted to address many of the issues related to choosing an appropriate ground plot for a multi-objective forest inventory or environmental survey. In

terms of the sample survey objectives, the most appropriate plot configuration consists of a single plot or point and these points will be spread across the study area using some form of systematic sample. This response design is thought to be inappropriate for spatial modeling applications because no data are collected at short lag distances. Wobble-plot sampling is developed to show that within the constraints imposed by the sample survey objectives, it is possible to simultaneously collect much, if not all, the data necessary for fitting the variogram. Further extensions and modifications of wobble-plot sampling are shown to improve the performance of ordinary kriging.

An interesting aspect of these studies is that it can be argued that environmental sampling is a mature field with little room for significant improvement. However, the results of this study suggest that many of the commonly used techniques in environmental surveys are inefficient. A rather surprising result is that some of the greatest gains in overall efficiency, as measured by sampling effort and mean squared error, are likely to come from the selection of a ground plot configuration that meet the many analytical objectives of the survey.

Literature Cited

Affleck, D. L. R., Gregoire, T. G., Valentine, H.T. 2005. Design unbiased estimation in line intersect sampling using segmented transects Environmental and Ecological Statistics. 12:139-154.

Frayer, W. E., and G. M. Furnival. 1999. Forest survey sampling designs: A history. J. For. 97:4-8.

Goebel, J. J., and George, T.A. 1990. Establishing a consistent national base for assessing natural resource issues. *in* proc. 30th International Atlantic Economic Conference. Williams, VA. Oct. 11-14.

Messer, J.J., Linthurst, R. A., and Overton, W. S. 1991. An EPA program for monitoring ecological status and rends. Environ. Mon. Assesst. 20:67-78.

Nusser, S. M., F. J. Breidt, and W. A. Fuller. 1998. Design and estimation for investigating the dynamics of natural resources. Ecol. Applic. 8:234-245.

Olsen, A. R., and Schreuder, H. T. 1997. Perspectives on large-scale natural resource surveys where cause-effect is a potential issue. Environ. Ecol. Stat. 4:167-180.

Overton, W. S. and Stehman S. V. 1995. Design implications of anticipated data uses for comprehensive environmental monitoring programmes. Environ. Ecol. Stat. 2:287-303.

Stein, M. L. 1999. Interpolation of spatial data: Some theory for Kriging. Springer, NY

Wallerman, J., Joyce, S., Vencatasawmy, C., and Olsson, H. 2002. Prediction of forest stem volume using kriging adapted to detected edges. Can. J. For. Res. 32: 509-518