AN ABSTRACT OF THE THESIS OF

Mercedes Berterretche for the degree of Master of Science in Forest Science

presented on May 23, 2001. Title: Comparison of Regression and Geostatistical

Methods to Develop LAI Surfaces for NPP Modeling

Abstract approved:

_____

(Warren Cohen —Julia A. Jones (|

This study aims to compare different methods of obtaining maximum growing

season leaf area index (LAI) maps using remote sensing data, LAI and tree cover

field data in a boreal forest near Thompson, Manitoba, Canada. The comparison

includes aspatial methods such as traditional regression, inverse regression and

reduced major axis, and spatial methods such as kriging, cokriging, kriging with an

external drift, and conditional simulation. The LAI maps will serve as input in

process models to obtain maps of net primary production (NPP).

The present work was done in the context of the BigFoot project

(http://www.fsl.orst.edu/ larse/bigfoot) which focuses on the validation of the

MODIS (Moderate Resolution Imaging Spectrometer) land cover, LAI/fAPAR

(fraction of absorbed photosynthetically active radiation), and NPP products

(http://modarch.gsfc.nasa.gov/MODIS, with the main objective of scaling up from

in situ ground measurements to the moderate spatial resolution of MODIS data

products (250 - 1000 m spatial resolution).

Due to the clumped structure of the boreal forest and the presence of a highly reflective understory, vegetation indices derived from remotely sensed data were not useful in explaining LAI variability. The use of mid-IR bands and tree cover data improved the performance of the models. Kriging with an external drift performed better in the presence of trends and anisotropy. An integrated aspatial (reduced major axis)/spatial (cokriging) method produced a useful compromise between local accuracy and pattern representation. Conditional simulation maintained global accuracy and spatial variability. Conditional simulation also provided a measure of spatial uncertainty useful to assess how LAI variability affects process models, and to evaluate how spatial variability influences the upscaling from Landsat ETM+ (25-30 m) to MODIS (250-1000 m) spatial resolutions.

Our main conclusion is that the selection of the optimal mapping technique depends on user requirements, because not all the desired map characteristics can be achieved simultaneously.

Comparison of Regression and Geostatistical Methods to Develop LAI Surfaces for
NPP Modeling

by

Mercedes Berterretche

A THESIS

submitted to

Oregon State University

In partial fulfillment of

the requirement for the

degree of

Master of Science

Presented May 23, 2001

Commencement June 2002

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

Comparison of Regression and Geostatistical Methods to Develop LAI Surfaces for NPP Modeling

Chapter One: Introduction

Objectives

This study aims to compare different methods of obtaining maximum growing

season leaf area index (LAI) maps using remote sensing data, and LAI and tree

cover field data in a boreal forest near Thompson, Manitoba, Canada. The

comparison includes aspatial methods such as traditional regression, inverse

regression and reduced major axis, and spatial geostatistical techniques such as

kriging, cokriging, kriging with an external drift, and conditional simulation. The

LAI maps will serve as input in process models to obtain net primary production

(NPP).

The present work was done in the context of the BigFoot project

(http://www.fsl.orst.edu/ larse/BigFoot) which focuses on the validation of the

MODIS (Moderate Resolution Imaging Spectrometer) land cover, LAI/fAPAR

(fraction of absorbed photosynthetically active radiation), and NPP products

(http://modarch.gsfc.nasa.gov/MODIS) (Cohen and Justice, 1999). The main

objective of BigFoot is scaling up from *in situ* ground measurements to the

moderate spatial resolution of MODIS data products (250 - 1000 m spatial resolution).

Leaf Area Index

Leaf area index (LAI) is a significant attribute of forest ecosystems that controls, in part, physiological processes such as photosynthesis, transpiration and leaf maintenance respiration, as well as physical processes such as snow melt, canopy water interception and evaporation, and light attenuation (Landsberg and Gower, 1997; Waring and Running, 1998). LAI, defined here as half the total leaf area per unit ground surface area (Chen and Cihlar,1996), is a widely used parameter that drives biogeochemical process models that characterize the primary productivity of extensive terrestrial areas (Running and Gower,1991). Estimates of LAI are often derived from remotely sensed data through empirical relationships with spectral vegetation indices (SVIs). Vegetation indices are calculated from remotely sensed reflectance data, and are often related to field LAI measurements using regression-based relationships. These relationships have been shown to be valid over a wide range of vegetation types and with an array of different sensors (Tucker, 1979; Peterson et al., 1987; Spanner et al., 1990 a and b; Fassnacht et al., 1997; Turner et al., 1999). Regression relationships, however, have limited accuracy in situations where canopy closure varies (Loechel et al., 1997), and where the understory and background materials contribute substantially to the reflectance signal received by

the sensor (Huete et al., 1985; Nemani et al., 1993), characteristics that were found

to be more pronounced for conifer species (Spanner et al., 1990 a).

In closed canopy cover conditions, LAI shows a negative relationship with red

reflectance and a positive one with near infrared (NIR) reflectance, maximizing the

utility of SVIs such as the normalized difference vegetation index (NDVI) and the

simple ratio (SR). But when canopy cover is not closed, the LAI-NIR relationship

may exhibit no relationship at all (Spanner et al., 1990 a; Nemani et al., 1993). To

complicate matters, background and understory reflectances change differently

through the seasons, due to snow cover or understory phenology (Chen and Cihlar,

1996), which sometimes depends on the overstory species composition (Miller et

al., 1997). Many attempts have been made to correct for this understory reflectance;

some of them included the use of mid- infrared (Mid-IR) wavelengths (Loechel et

al.,1997; Nemani et al., 1993). Mid-IR is negatively related to LAI in closed

canopies, and is strongly affected by leaf water content (Lillesand and Kiefer, 1999;

Nemani et al., 1993).

A further difficulty in the estimation of LAI from vegetation indices is a commonly

observed asymptotic relationship between LAI and SVIs, especially at high LAI

values (Spanner et al., 1990 a). Canopy architecture plays an important role in the

interception and reflection of solar radiation, influencing the signal received by the

sensor. Leaf spectral properties (Gates et al., 1965), foliar angular and spatial

distribution (including vertical distribution of foliage, tree height and gap

distribution) (Lefsky et al., 1999; Fournier et al., 1997; Chen and Cihlar, 1996;

Cohen et al., 1990), and foliar hierarchical clumping structure (Gower et al., 1999;

Ni et al., 1997), all contribute to the heterogeneity of the radiation environment

within the forest.

LAI estimation in boreal forests is challenging because these forests have high

variability at all scales of organization, not so much due to the species diversity,

which is low (Landsberg and Gower, 1997), but from structural diversity. The

complex structure of boreal forests is a function of topography, soil parent material,

climate and periodic disturbances (Van Cleve and Vierek, 1981) and stand age. At

the large scale, fire is the most important natural disturbance (followed by tree

damage by insects), influencing species composition, nutrient availability, and

forest age and productivity (Larsen, 1982). Fire frequency may have increased

lately as a result of increased temperatures coupled with drier conditions (Van

Cleve and Vierek, 1981).

Boreal conifers have narrow, columnar crowns, and clumping at the shoot, branch

and crown level, to maximize light interception and reduce damage from snow

loading (Landsberg and Gower, 1997). Black spruce is a good example, with 40 to

50 % of the foliage concentrated in the top of the crown, leading to LAI

underestimation by optical measurement methods (Chen et al., 1997). Another

characteristic of boreal forests is the rich understory and soil cover, composed of

various shrubs, grasses, tree regeneration, and abundant mosses, lichens and

sphagnum species, distinctive of boreal plant communities. Bryophytes exhibit

different spectral characteristics from vascular plants (Bubier et al., 1997; Petzold

and Goward, 1988; Vogelman and Moss, 1993). Depending on the density and

composition of the overstory vegetation and the time of year, these components

affect differently the signal remotely received by overhead sensors (Chen and

Cihlar, 1996; Miller et al., 1997). Chen and Cihlar (1996) found that late spring

images were superior to summer ones to determine overstory LAI because the

effect of the understory is minimized, and that there was no obvious saturation

point in the LAI-NDVI relationship because of the clumpiness of the canopy.

Tree Cover

Cover, like LAI, is an ecological property with great functional significance. It

influences the microenvironment within the forest in terms of light, temperature,

rainfall and snow interception, which play a role in overstory and understory

development. Loechel et al. (1997) observed the positive correlation between LAI

and canopy cover and suggested that cover may be an important variable to add to

LAI models when poor NIR-LAI relationships are present, such as in the open

canopy situations of this study.

Cover was defined by Mueller-Dombois and Ellenberg (1974) as "the vertical

projection of the crown or shoot area of a species to the ground surface expressed

as a fraction or percent of a reference area". Bunnell and Vales (1990) suggested

mean crown completeness (MCC) as a useful measure of canopy cover as observed

from the ground. MCC is a stand or plot measurement denoting the mean of several measurements of the proportion of the sky covered by tree crowns within a specified angle from a single point (Bunnell and Vales, 1990). The area sampled is determined by trigonometric principles and depends on the angle of view, the height to base of live crown and the height from which the angle is projected. Bunnel et al. (1990) and Vales and Bunnel (1988) compared several different cover measurement techniques including ocular, moosehorn, spherical densiometer, regular and hemispherical photographs to evaluate differences among techniques and effects of the observers. They agreed that narrow angles of view and vertical projections are the least biased ways of estimating MCC to include both gaps between and within crowns. They observed that wide angles of view masked small gaps, included objects not directly above the point sampled, and had an angular view of the canopy at the outer edges, overestimating MCC. On the other hand, Bunnel et al. (1990) suggested that wide angles may be more appropriate when examining relationships between overstory and understory radiation in boreal forests, where low solar angles are present.

Digital photography has the advantage of offering a permanent and objective record of the samples that allows flexible laboratory analysis and eliminates observer effects.

Theory of Applied Methods

This section reviews the theory of the different aspatial and spatial LAI estimation

and simulation techniques tested in this study (Fig. 1).
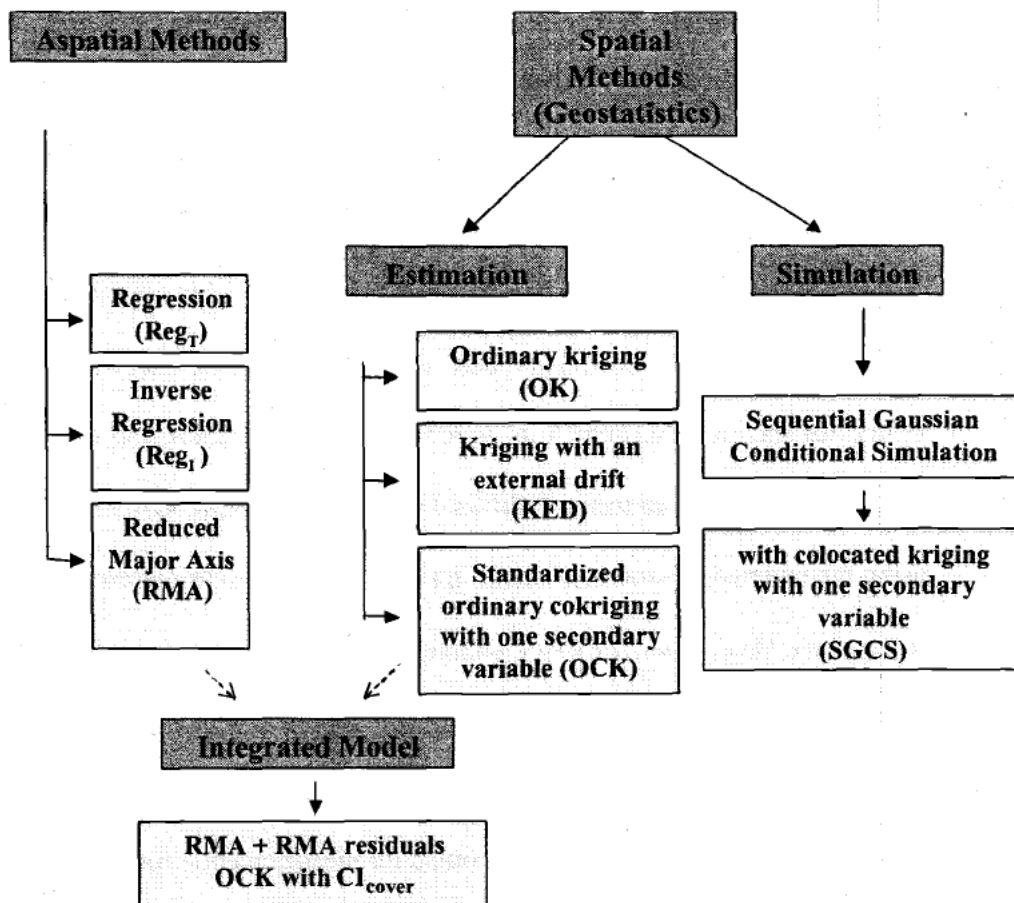
Figure 1. Scheme of methods used in this study.

Aspatial Regression-based Methods

A regression equation can generally be expressed as:

$$Y = \alpha + \beta X + \varepsilon$$

where Y is the variable to be estimated, α is the intercept, β is the slope, X is the independent variable, and ε is the error, assumed to be independent and normally distributed. Regression methods assume that data are spatially independent. In remote sensing, the independent variable is the earth's surface reflectance, and the dependent variable is the signal received by the sensor. Commonly used SVIs are based in traditional regression (Reg$_T$), and are expressed as $Y = \alpha + \beta X + \varepsilon$, where Y is LAI and X is reflectance or one of several SVIs. This type of regression analysis is based on the assumption that there are no measurement errors in the independent variable, and is designed to estimate Y by minimizing the sum of squares errors in Y. Curran and Hay. (1986) described the major measurement errors that should be accounted for in remote sensing variables, and that are generally ignored when using traditional regression. This problem is solved by using the inverted equation $X = \alpha + \beta Y + \varepsilon$, referred to as inverse regression (Curran and Hay., 1986), or Reg$_I$, which minimizes the sum of squared errors in X. However, Chen and Cihlar (1997) pointed out that commonly used allometric methods to estimate LAI can accumulate errors, and it is difficult to keep the total error under 25 %. A method that accounts for measurement errors in both dependent and independent variables is needed in such situations. The reduced major axis method (RMA) improves the traditional and inverse regression procedures in that it minimizes the sum of the cross-products of the differences on both axes, accounting simultaneously for the errors in both dependent and

independent variables (Miller and Kahn, 1962; Davis, 1986) and is given by

$$Y = (\alpha + \beta X + \varepsilon) * (-1).$$

The main difference among these methods is that traditional and inverse regression

equation coefficients (α and β) are determined by least squares, while reduced

major axis intercept and slope are given by:

$$\alpha = \overline{Y} - \frac{\sigma_Y}{\sigma_X} * \overline{X}$$

$$\beta = \frac{\sigma_Y}{\sigma_X}$$

where $\overline{Y}$ and $\overline{X}$ are the means of the dependent and independent variables

respectively, and σ_Y and σ_X are their standard deviations.

Geostatistics

Geostatistics, based on the theory of regionalized variables (Journel and Huijbregts,

1978), is concerned with a variety of techniques aimed at understanding and

modeling spatial variability through estimation and simulation (Deutsch, 2000;

Journel, 1989; Goovaerts, 1997). A regionalized variable equation differs from a

regression equation in that its components are indexed by their position in space or

time:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon''$$

where $Z(x)$ is the variable to be estimated at location x, $m(x)$ is a deterministic

component, $\varepsilon'(x)$ is spatially correlated variability, and ε'' is a spatially

independent random residual, assumed to be normally distributed and interpreted as nugget variance. Geostatistics accounts for the presence of spatial autocorrelation and joint dependence in space and time, which occur in most natural resources variables (Myers, 1997).

In an ecological context, geostatistics has been used to describe scale and pattern of spatial variability (Woodcock et al., 1988; Legendre and Fortin, 1989; Turner et al., 1990; Rossi et al., 1992), to characterize canopy structure (Cohen et al., 1990; St-Onge and Cavayas, 1997; Hudak and Wessman, 1998; Wulder et al., 1998), to estimate continuous and categorical variables (Rossi et al., 1993; Milne and Cohen, 1999), and to assess risk (Myers, 1997; Saito and Goovaerts, 2000).

Biogeochemical models are increasingly adopting an explicitly spatial configuration. Spatial surfaces of meteorological values such as temperature and evapotranspiration or accurate digital elevation models (DEM) are commonly required as model inputs (Running and Nemani, 1987). Geostatistics is currently used to improve such data layers (Goovaerts, 2000; Kyriakidis, 1999). The latest applications of geostatistics have emphasized the use of models of uncertainty that depend on the data values in addition to data configuration (Deutsch and Journel, 1998). Stochastic simulation is an example of a probabilistic approach that provides a distribution of possible values for each cell of the surface, characterizing uncertainty. These uncertainty measurements improve ecological interpretation,

help assess error areas and decrease losses and risks in policy and management decision-making (Rossi et al., 1993).

Spatial structure may be described by structure functions that allow the quantification of spatial dependency. The structure functions used in this study were semivariance and cross-correlation.

Semivariance, $\gamma(h)$, is a measure of spatial variability used to assess the average dissimilarity between data separated by a vector h (Curran, 1988; Goovaerts, 1997). It is computed as follows:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2$$

where h is the lag, or vector separation between the attribute z at locations u_α and $(u_\alpha + h)$, and $N(h)$ is the number of pairs of data locations a vector h apart. The subscript α represents the sample data.

An experimental semivariogram is a graphical representation of the semivariance versus the lag, and summarizes the magnitude, spatial scale, and pattern of the variation for a given set of data (Curran, 1988; Wulder et al., 1998; Milne and Cohen, 1999).

Experimental semivariograms are interpreted in terms of the sill, range, slope, and nugget. The sill represents the maximum semivariance, the point where increments of the lag do not result in increases in semivariance because spatial autocorrelation is no longer present (St. Onge and Cavayas, 1997). The lag distance at which the

sill is reached is called the range and within this distance, sample points are spatially dependent. The slope represents the rate of change of the semivariance with lag dstance. The nugget variance is the apparent discontinuity at the origin of the semivariogram. It is the sum of two components: measurement errors, and variation below the minimum lag distance (Burrough, 1995; Rossi et al., 1992; Wulder 1998; Deutsch 2000). In an image context, the height of the sill is proportional to the global image variance and related to the density of scene objects, and the range is a good indicator of texture coarseness, related to the size of the objects (St-Onge and Cavayas, 1997; Woodcock et al., 1988)

The computation of the experimental semivariogram based on a sample data set allows for the construction of models used in geostatistical estimation and simulation methods. Semivariograms can be computed as either an average over all directions (omnidirectional semivariograms) or in a particular direction (directional semivariograms), allowing isotropic and anisotropic analysis of the data, respectively.

Model functions were fitted to experimental following Isaaks and Srivastava (1989), and Deutsch and Journel (1998). To ensure the positive definiteness condition required by the covariance kriging matrices to guarantee positive variance values, it is advised to use only the functions that are known to be positive definite, such as spherical, exponential, and hole effect (refer to Isaaks and Srivastava, 1989, and Deutsch and Journel, 1998, for a full description of these functions). The present study modeled the semivariograms and cross variograms

using "nested structures", which are linear combinations of the basic functions. The following is an example of a nested isotropic model:

$$\gamma_{Z_{OK}}(h) = a + b * sph_{(dist_a)} + c * hole_{(dist_b)}$$

where a is the nugget effect ; b is the sill contribution to the model sph , a spherical model, up to $dist_a$, the range of the second structure; and c is the sill contribution to the model $hole$, a hole effect model, up to $dist_b$, the range of the third structure.

In this way, the most relevant features of the experimental semivariogram can be represented in the models to be used in the geostatistical procedures. Cross-validation procedures helped assess model fit. Simplicity is an important rule in the construction of the models.

The **cross-correlation** is a unit-free measure of similarity between two different attributes. It is given by:

$$\rho_{ij}(h) = \frac{C_{ij}(h)}{\sqrt{\sigma^2_{i_{-h}} \sigma^2_{j_{+h}}}} \qquad \in [-1,+1]$$

with

$$C_{ij} = \frac{\sum\limits_{n=1}^{N}(z_i - m_i)(z_j - m_j)}{N}$$

where $C_{ij}(h)$ is the covariance between z_i and z_j values separated by vector h,

N is the total number of pairs of sample values (z_i, z_j), m_i and m_j are the means

of the i and j attributes, and with

$$\sigma^2_{i_{-h}} = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} [z_i(u_\alpha) - m_{i_{-h}}]^2$$

$$\sigma^2_{j_{+h}} = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} [z_j(u_\alpha) - m_{j_{+h}}]^2$$

where $\sigma^2_{i_{-h}}$ and $\sigma^2_{j_{+h}}$ are the variances of the two different attributes at the tail and

head of vector h, $m_{i_{-h}}$ and $m_{i_{+h}}$ are the means of the tail and head values, and

$z_i(u_\alpha)$ are the i attributes of variable z at locations u_α (Rossi et al., 1992;

Goovaerts, 1997).

A plot of the cross-correlation values versus distance is called a cross-correlogram

and helps to visualize and interpret how two variables vary jointly in space.

Generally, the cross-correlation between two variables reaches its maximum at lag

0, and decreases as the distance increases.

Geostatistical Methods

Geostatistics is based on the concepts of random variables and random functions

(Isaaks and Srivastava, 1989). A random variable (RV) is a variable whose values

are randomly generated according to some probabilistic mechanism that models the uncertainty about the attribute under study. In predictive statistics, any unsampled (unknown) value is characterized as a random variable. The probability distribution of the random variable is usually location and information-dependent (Deutsch, 1998). The random variable seen as a function of spatial location is called a random function (RF). A random function is a set of random variables that have associated spatial locations and whose dependence on each other is specified by some probabilistic procedure (Isaaks and Srivastava, 1989; Deutsch, 2000).

Kriging is a generalized least-squares linear regression technique that accounts for the spatial dependence among observations. The kriging estimators are exact interpolators, meaning that they honor or reproduce the sample values at their locations (Goovaerts, 1997). Three different kriging variants were applied in this study: ordinary kriging (OK), kriging with an external drift (KED) and ordinary cokriging (OCK).

Ordinary Kriging

Ordinary kriging is a univariate technique that accounts for local fluctuations in the mean value by limiting the domain of stationarity of the mean to a local neighborhood centered on the location being estimated (Goovaerts, 1997). The ordinary kriging estimator is calculated as follows:

$$Z_{OK}^{*}(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{OK}(u) Z(u_{\alpha})$$

under the constraint:

$$\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{OK}(u) = 1$$

where $Z_{OK}^{*}(u)$ is the ordinay kriging estimator at location u, $\lambda_{\alpha}^{OK}(u)$ are the OK

weights corresponding to the n samples at location u, and $Z(u_{\alpha})$ are the nearby

sample values.

Kriging minimizes the error variance and aims to obtain a mean residual error equal

to 0. These constraints and the specification of a pattern of spatial continuity (a

model applied to the experimental semivariogram) lead to a normal system of

linear equations called the ordinary kriging system that provides the OK weights.

For a detailed description of the kriging computation refer to Isaaks and Srivastava,

(1989), Goovaerts, (1977), or Deutsch and Journel, (1998).

The ordinary kriging variance is:

$$\sigma_{OK}^{2}(u) = C(0) - \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{OK}(u) C(u_{\alpha} - u) - \mu_{OK}(u)$$

where $\sigma^2_{OK}(u)$ is the minimum error variance, $C(0)$ is the stationary covariance,

$C(u_\alpha - u)$ is the covariance with the unknown value, and $\mu_{OK}(u)$ is the Lagrange

parameter.

This error variance is dependent on the data configuration but independent of the

data values, so it is not a good measure of error. The variation is always 0 at sample

data locations, increases away from the data and reaches a maximum value at

locations most distant from sample data points (extrapolation situations).

Kriging with an External Drift

Kriging with an external drift (KED) is a variant of kriging that allows for the use

of secondary information known at every location (exhaustive), which is assumed

to reflect the spatial trend of the primary variable (Deutsch and Journel, 1998;

Goovaerts, 1997). At a landscape scale, spatial variation can be decomposed into

two components: large-scale variation and small-scale variation. Large-scale

variability may be influenced by geomorphology, elevation, slope, aspect,

precipitation, or disturbances like fire or disease. Small-scale variability may be

influenced by soil permeability, nutrient availability, or pH. The KED trend

represents the large-scale variability of the primary variable. The residuals from the

trend represent the small-scale variability, and the final KED result combines both.

KED models the trend under the assumption of a linear relationship between

primary and secondary variables, and smooth variation of the secondary variable.

The distinctive feature of KED is that the algorithm considers a non-stationary

random function model, where stationarity is limited within each search neighborhood, yielding more local detail (Deutsch and Journel, 1998). The KED estimator is written:

$$Z^*_{KED}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{KED}_{\alpha}(u) Z(u_\alpha)$$

where $Z^*_{KED}(u)$ is the kriging with an external drift estimator at location u,

$\lambda^{KED}_{\alpha}(u)$ are the KED weights corresponding to the n samples at location u,

and $Z(u_\alpha)$ are the nearby sample values.

The procedure consists of three steps (Goovaerts, 1997); first, the trend coefficients $a^*_0(u)$ and $a^*_1(u)$ of the trend model $m^*_{KED}(u)$ are evaluated within the search neighborhood, from the $n(u)$ data pairs $(z_1(u), z_2(u))$, where $z_1(u)$ are the primary variable sample points, and $z_2(u)$ are the secondary variable sample data. These coefficients are estimated through the kriging system. Then, the trend components $m(u)$ are estimated at all primary sampled locations and at all locations being estimated. Finally, a simple kriging is performed in the residuals of the trend:

$$Z^*_{KED}(u) - m^*_{KED}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{SK}_{\alpha}(u)[Z(u_\alpha) - m^*_{KED}(u_\alpha)]$$

where $m^*_{KED}(u)$ is the trend component, estimated as

$$m^*_{KED}(u) = a^*_0(u) + a^*_1(u) z_2(u) \qquad \vee \alpha = 1,\ldots,n(u)$$

where α are the sample data, and λ_α^{SK} are the simple kriging weights.

The minimized error variance is:

$$\sigma_{KED}^2(u) = C_R(0) - \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{KED(u)} C_R(u_\alpha - u) - \sum_{k=0}^{K} \mu_k^{KED}(u) f_k(u)$$

where $C_R(0)$ is the residuals' stationary covariance, $\mu_k^{KED}(u)$ are the Lagrange

parameters, and $f_k(u)$ are the trend functions. KED is also useful to estimate the

trend itself, not only acting as a local estimation procedure, but also as a global one

(Isaaks and Srivastava, 1989).

Standardized Ordinary Cokriging

Standardized ordinary full cokriging (OCK) is a multivariate extension of kriging

that allows for the use of secondary data, accounting for the spatial cross

correlation between the primary (undersampled) and secondary (exhaustive)

variables. The OCK estimator is written:

$$Z_{OCK}^*(u) = \sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{OCK}(u) Z_1(u_{\alpha_1}) + \sum_{\alpha_2=1}^{n_2(u)} \lambda_{\alpha_2}^{OCK}(u) Z_2(u_{\alpha_2}) * [Z_2(u_{\alpha_2}) - m_2 + m_1]$$

with the constraint

$$\sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{OCK}(u) + \sum_{\alpha_2=1}^{n_2(u)} \lambda_{\alpha_2}^{OCK}(u) = 1$$

where $Z^*_{OCK}(u)$ is the standardized ordinary cokriging estimator at location u,

$\lambda^{OCK}_{\alpha i}(u)$ are the ordinary cokriging weights, and m_1 and m_2 are global means of

the primary and secondary variables, respectively.

The minimum error variance is computed as follows:

$$\sigma^2_{OCK}(u) = C_{11}(0) - \mu^{OCK}_1(u) - \sum_{\alpha_1=1}^{n_1(u)} \lambda^{OCK}_{\alpha 1}(u)C_{11}(u_{\alpha 1} - u) - \sum_{\alpha_2=1}^{n_2(u)} \lambda^{OCK}_{\alpha 2}(u)C_{21}(u_{\alpha 2} - u)$$

where $\sigma^2_{OCK}(u)$ is the minimum error variance, $C_{11}(0)$ is the primary stationary

variance, $C_{11}(u_\alpha - u)$ is the covariance between the primary variable and the

unknown, $C_{211}(u_\alpha - u)$ is the covariance between the secondary variable and the

unknown, and $\mu^{OCK}_1(u)$ is the Lagrange parameter.

Cokriging requires a joint model that includes the modeling of the direct

semivariograms of primary and secondary variables and their respective cross-

semivariograms under the condition that the whole set of semivariograms be

positive definite (Isaaks and Srivastava, 1989; Goovaerts 1997). This need for

positive definiteness imposes many constraints on the models that can be chosen; a

relatively easy way of checking this property is the linear model of

coregionalization (LMC).

The LMC is defined as a set of direct and cross-semivariogram models $\gamma_{ij}(h)$ such that:

$$\gamma_{ij}(h) = \sum_{l=0}^{L} b_{ij}^l g_l(h) \quad \forall i, j$$

where each function $g_l(h)$ is a permissible model and the $(L+1)$ matrices of coefficients b_{ij}^l, corresponding to the sill or slope of the model $g_l(h)$, are all positive semi-definite. Refer to Goovaerts, (1997) for an exhaustive development of the linear model of coregionalization.

The main differences between OCK and KED are first, that the cokriging estimates are directly influenced by the secondary data, while in the KED approach, secondary data only provide information about the trend; and second, that the secondary information influences the KED depending on the slope of the trend, while OCK accounts for the global correlation of the primary and secondary data through the cross-semivariogram.

Integrated Models

In cases where a trend is present, it is better to model the trend using physical knowledge of the phenomenon under study (Isaaks and Srivastava, 1989). Such an approach involves modeling the trend from the observed sample values, the subtraction of the trend to obtain the residuals, and the interpolation estimation on

the residuals (i.e. through kriging), adding up the trend and the estimated residuals at the end.

Metzger, (1997) suggested a combination of trend surface analysis and kriging or cokriging to combine the use of remote sensing data and field data, and she improved the estimation of certain forest variables. The conceptual idea underlying this procedure is to exploit large scale variability using remote sensing data through the trend surface analysis, and to describe small scale variability using the residuals of the trend model cokriged with a secondary exhaustive source of information.

Journel et al., (1989) stated that the choice of a particular trend model acquires more relevance in extrapolation conditions, outside the window data.

Conditional Simulation

Stochastic simulation is a probabilistic approach that provides a distribution of multiple, equally probable realizations of the joint distribution of one or more attribute values in space, generating a model of spatial uncertainty (Goovaerts, 1997; Rossi et al., 1993).

Several different simulation algorithms are used to model spatial uncertainty, depending on the purposes of the final outputs. The approach used in this study was sequential Gaussian conditional simulation (SGCS) with simple collocated

cokriging with one secondary variable. SGCS is based on the Bayes postulate describing the conditional or posterior probability of two events:

$$P(A_1 \mid A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

This expression can be read as the probability that an event A_1 will occur given that event A_2 has occurred, and is equivalent to the ratio of the joint probability between the two events and the individual probability of event A_2 (Rossi et al., 1993). This expression can be expanded to n events, the events being the nodes or cells to simulate on the map.

The general idea of SGCS with collocated cokriging is to produce a set of joint realizations of the spatial distribution of the primary attribute, conditional to both the primary and collocated secondary information.

The collocated simple cokriging estimate of the primary attribute is:

$$Z_{SCK}^{*}(u) = \sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{SCK}(u)[Z_1(u_{\alpha_1}) - m_1] + \lambda_{\alpha_2}^{SCK}(u)[Z_2(u) - m_2] + m_1$$

where $Z_{SCK}^{*}(u)$ is the colocated simple cokriging estimator at location u,

$\lambda_{\alpha i}^{SCK}(u)$ are the collocated simple kriging weights, and m_1 and m_2 are global means of the primary and secondary variables, respectively.

Unlike full cokriging, collocated cokriging just uses the lag correlation function

between the primary and secondary variables, following the Markov-type

approximation:

$$\rho_{12}(h) \approx \rho_{12}(0) * \rho_{11}(h)$$

where $\rho_{12}(h)$ is the lag cross-correlation function of the primary and secondary

variables, $\rho_{12}(0)$ is the correlation coefficient between the primary and secondary

variables, and $\rho_{11}(h)$ is the lag correlation function of the primary variable.

This approximation avoids modeling the linear model of coregionalization; only the

semivariogram of the primary variable needs to be modeled. The cross-variogram

model is derived as a linear rescaling of the primary variable semivariogram model

by the correlation coefficient $\rho_{12}(0)$. In collocated cokriging, the dependence of

the secondary variable on the primary one is limited to the collocated datum to

avoid matrix instability problems caused by highly redundant secondary

information, and to speed up the process (Goovaerts, 1997).

The smoothing effect of kriging and cokriging is due to a missing variance

component (Deutsch, 1998). The conditional simulation realizations have added

back the missing variance estimated by cokriging, representing in a more realistic

way the spatial distribution of the features observed in the original data. Because

spatial heterogeneity is a natural feature of ecosystems, conditional simulation

realizations have the advantage of representing high spatial entropy, producing more realistic output displays of the variable in terms of spatial variability.

The variance of the cokriged estimate is

$$Var\{Z_{SCK}^*(u)\} = \sigma^2 - \sigma_{SCK}^2(u)$$

where $Var\{Z_{SCK}^*(u)\}$ is the variance of the simple collocated cokriged estimate,

σ^2 is the variance of the stationary random function or the variance at lag $= 0$, and

$\sigma_{SCK}^2(u)$ is the simple collocated cokriging variance.

So, the cokriging variance of the estimate is smaller than the total variance and that explains why cokriging maps are smooth. A way of incorporating this missing variance is to add an independent random component $R(u)$, while reproducing the covariance properties of cokriging. The simulated value $Z_S(u)$ would then be:

$$Z_S(u) = Z_{SCK}^*(u) + R(u)$$

The SGCS algorithm used here (Deutsch, 1998) proceeds as explained in the following steps (Deutsch, 2000; Goovaerts, 1997). First, the original data are transformed to a standard normal distribution. A multivariate multi-point Gaussian random function model is adopted if the joint normal score random functions are normally distributed. Subsequently, a random path to visit once all the nodes of the grid is defined. Then, a location u is visited randomly and collocated cokriging is performed to obtain the cokriged estimate and its corresponding variance, using the

conditional information provided by neighboring primary and collocated secondary normal score data. Next, the algorithm draws a random residual (through Monte Carlo simulation) that follows a normal distribution with mean equal to 0 and variance equal to the cokriging variance. The cokriged estimate and the random residual are added to obtain the simulated value. The simulated value is included in the data set and the algorithm proceeds to the next node, looping until all nodes are simulated. Finally, the simulated normal scores are backtransformed to their original values. Another realization can be obtained by repeating the process with a different random visiting path.

Simulations were developed initially to provide measures of spatial uncertainty, but they have increasingly been used as maps of the variable of interest in cases where the reproduction of the spatial variability is more important than local accuracy and where sample and exhaustive data are available (Deutsch, 1998). Unlike kriging, conditional simulation emphasizes the global accuracy of the simulated variable, reproducing its sample statistics (mean, histogram, covariance) and its pattern of spatial continuity (Deutsch, 1998; Dungan, 1999). Another difference with interpolation algorithms is that stochastic simulation provides measures of local and global joint accuracy. Table 1 summarizes the relevant properties of the methods used. It also provides information for use in comparing the methods.

Table 1. Summary of relevant properties of the methods. This summary is based on methods as applied in this study.

Method	Type of Method	Number of variables, how are they used?
Traditional Regression Reg$_T$	Parametric, single estimation	Multivariate, coefficients calculated over entire area
Inverse regression Reg$_I$	Parametric, single estimation	Multivariate (condensed to do the inversion), coefficients calculated over entire area
Reduced major axis RMA	Parametric, single estimation	Multivariate (condensed to do coef. calculation), coefficients calculated over entire area
Ordinary kriging OK	Best linear unbiased estimator, exactness property, single estimation	Univariate, uses local search neighborhoods
Kriging with an external drift KED	Best linear unbiased estimator, exactness property, single estimation	Bivariate, secondary variable only characterizes the trend of the primary variable. Trend coefficients calculated within local search neigborhoods
Ordinary cokriging OCK	Best linear unbiased estimator, exactness property, single estimation	Multivariate, secondary variables directly influence the estimate through cokriging system, uses local search neighborhoods
Integrated methods	Single estimation, exactness property	Multivariate, secondary variables directly influence the estimate through cokriging system, uses local search neighborhoods
Conditional simulation SCGS	Multiple equally probable realizations, exactness property	Uni-Multivariate, secondary variable used based on Markov approximation

Table 1 (continued).

Method	Spatial autocorrelation (SAC) is exploited? How?	Limitations / Assumptions
Traditional Regression Reg$_T$	Aspatial	Assumes: residuals' normality, independence, and homoscedasticity. Does not account for measurement error in explanatory variable
Inverse regression Reg$_I$	Aspatial	Assumes: residuals' normality, independence, and homoscedasticity. Does not account for measurement error in response variable
Reduced major axis RMA	Aspatial	Assumes that measurement errors for explanatory and response variables are similar in magnitude
Ordinary kriging OK	Yes, through modeling the variogram, managing the neigborhood size, shape, and the number of samples used for estimation	Assumes stationarity within the search neighborhood. Useful only for interpolation cases
Kriging with an external drift KED	Yes, through modeling of variogram of primary variable, managing neighborhood size, shape, and the number of samples used for estimation	Assumes gradual trend over area of interest. Assumes linear relationship between primary and secondary variable. Secondary variable has to be known at all locations
Ordinary cokriging OCK	Yes, through modeling the direct and cross-semivariograms of primary and secondary variables, managing neighborhood size, shape, and the number of samples used for estimation	Assumes stationarity within the search neighborhood. Useful only for interpolation cases. The linear model of corregionalization has several constraints when modeling.
Integrated methods	Yes, through modeling the residuals and the secondary variable direct and cross-semivariograms, managing neighborhood size, shape, and the number of samples used for estimation	Combines RMA limitations with cokriging ones, but extrapolation situations are improved
Conditional simulation SCGS	Yes, through modeling the primary variable variogram, managing neighborhood size, shape, the number of samples and simulated values, and the searching strategy	Assumes multiGaussian distribution. Markov-type approximation assumes semivariogram of primary variable proportional to cross semivariogram

Chapter Two

Comparison of Regression and Geostatistical Methods to Develop LAI Surfaces for NPP Modeling

Mercedes Berterretche[1], Warren B. Cohen[2], Andrew T. Hudak[2], Thomas K. Maiersperger[2], Michael A. Lefsky[2],Karin S. Fassnacht[2], and Stith T. Gower[3]

[1]Department of Forest Science, Oregon State University, Corvallis, Oregon

[2]Pacific Northwest Research Station, USDA Forest Service, Corvallis, Oregon

[3]Department of Forest Ecology and Management, University of Wisconsin, Madison

Abstract

This study aims to compare different methods of obtaining maximum growing

season leaf area index (LAI) maps using remote sensing data, LAI and tree cover

field data in a boreal forest near Thompson, Manitoba, Canada. The comparison

includes aspatial methods such as traditional regression, inverse regression and

reduced major axis, and spatial methods such as kriging, cokriging, kriging with an

external drift, and conditional simulation. The LAI maps will serve as input in

process models to obtain maps of net primary production (NPP).

The present work was done in the context of the BigFoot project

(http://www.fsl.orst.edu/ larse/bigfoot) which focuses on the validation of the

MODIS (Moderate Resolution Imaging Spectrometer) land cover, LAI/fAPAR

(fraction of absorbed photosynthetically active radiation), and NPP products

(http://modarch.gsfc.nasa.gov/MODIS, with the main objective of scaling up from

in situ ground measurements to the moderate spatial resolution of MODIS data

products (250 - 1000 m spatial resolution).

Due to the clumped structure of the boreal forest and the presence of a highly

reflective understory, vegetation indices derived from remotely sensed data were

not useful in explaining LAI variability. The use of mid-IR bands and tree cover

data improved the performance of the models. Kriging with an external drift

performed better in the presence of trends and anisotropy. An integrated aspatial

(reduced major axis)/spatial (cokriging) method produced a useful compromise

between local accuracy and pattern representation. Conditional simulation maintained global accuracy and spatial variability. Conditional simulation also provided a measure of spatial uncertainty useful to assess how LAI variability affects process models, and to evaluate how spatial variability influences the upscaling from Landsat ETM+ (25-30 m) to MODIS (250-1000 m) spatial resolutions.

Our main conclusion is that the selection of the optimal mapping technique depends on user requirements, because not all the desired map characteristics can be achieved simultaneously.

Keywords: LAI maps, reduced major axis, kriging, cokriging, kriging with an external drift, conditional simulation , boreal forest.

Introduction

Leaf Area Index

Leaf area index (LAI) is a significant attribute of forest ecosystems that controls, in part, physiological processes such as photosynthesis, transpiration and leaf maintenance respiration, as well as physical processes such as snow melt, canopy water interception and evaporation, and light attenuation (Landsberg and Gower, 1997; Waring and Running, 1998). LAI, defined here as half the total leaf area per unit ground surface area (Chen and Cihlar,1996), is a widely used parameter that drives biogeochemical process models that characterize the primary productivity of extensive terrestrial areas (Running and Gower,1991). Estimates of LAI are often derived from remotely sensed data through empirical relationships with spectral vegetation indices (SVIs). Vegetation indices are calculated from remotely sensed reflectance data, and are often related to field LAI measurements using regression-based relationships. These relationships have been shown to be valid over a wide range of vegetation types and with an array of different sensors (Tucker, 1979; Peterson et al., 1987; Spanner et al., 1990 a and b; Fassnacht et al., 1997; Turner et al., 1999). Regression relationships, however, have limited accuracy in situations where canopy closure varies (Loechel et al., 1997), and where the understory and background materials contribute substantially to the reflectance signal received by the sensor (Huete et al., 1985; Nemani et al., 1993), characteristics more pronounced for conifer species (Spanner et al., 1990 a).

In closed canopy cover conditions, LAI shows a negative relationship with red reflectance and a positive one with near infrared (NIR) reflectance, maximizing the utility of SVIs such as the normalized difference vegetation index (NDVI) and the simple ratio (SR). But when canopy cover is not closed, the LAI-NIR relationship may exhibit no relationship at all (Spanner et al., 1990 a; Nemani et al., 1993). To complicate matters, background and understory reflectances change differently through the seasons, due to snow cover or understory phenology (Chen and Cihlar, 1996), which sometimes depends on the overstory species composition (Miller et al., 1997). Many attempts have been made to correct for this understory reflectance; some of them included the use of mid- infrared (Mid-IR) wavelengths (Loechel et al.,1997; Nemani et al., 1993). Mid-IR is negatively related to LAI in closed canopies, and is strongly affected by leaf water content (Lillesand and Kiefer, 1999; Nemani et al., 1993).

Canopy architecture plays an important role in the interception and reflection of solar radiation, influencing the signal received by the sensor. Leaf spectral properties (Gates et al., 1965), foliar angular and spatial distribution (including vertical distribution of foliage, tree height and gap distribution) (Lefsky et al., 1999; Fournier et al., 1997; Chen and Cihlar, 1996; Cohen et al., 1990), and foliar hierarchical clumping structure (Gower et al., 1999; Ni et al., 1997), all contribute to the heterogeneity of the radiation environment within the forest.

LAI estimation in boreal forests is challenging because these forests have high

variability at all scales of organization, not so much due to the species diversity,

which is low (Landsberg and Gower, 1997), but from structural diversity. The

complex structure of boreal forests is a function of topography, soil parent material,

climate and periodic disturbances (Van Cleve and Vierek, 1981) and stand age. At

the large scale, fire is the most important natural disturbance, influencing species

composition, nutrient availability, and forest age and productivity (Larsen, 1982).

Boreal conifers have narrow, columnar crowns, and clumping at the shoot, branch

and crown level, to maximize light interception and reduce damage from snow

loading (Landsberg and Gower, 1997). Black spruce is a good example, with 40 to

50 % of the foliage concentrated in the top of the crown, leading to LAI

underestimation by optical measurement methods (Chen et al., 1997). Another

characteristic of boreal forests is the rich understory and soil cover, composed of

various shrubs, grasses, tree regeneration, and abundant mosses, lichens and

sphagnum species, distinctive of boreal plant communities. Bryophytes exhibit

different spectral characteristics from vascular plants (Bubier et al., 1997; Petzold

and Goward, 1988; Vogelman and Moss, 1993). Depending on the density and

composition of the overstory vegetation and the time of year, these components

affect differently the signal remotely received by overhead sensors (Chen and

Cihlar, 1996; Miller et al., 1997). Chen and Cihlar (1996) found that late spring

images were superior to summer ones to determine overstory LAI because the

effect of the understory is minimized, and that there was no obvious saturation point in the LAI-NDVI relationship because of the clumpiness of the canopy.

Tree Cover

Cover, like LAI, is an ecological property with great functional significance. It influences the microenvironment within the forest in terms of light, temperature, rainfall and snow interception, which play a role in overstory and understory development. Loechel et al. (1997) observed the positive correlation between LAI and canopy cover and suggested that cover may be an important variable to add to LAI models when poor NIR-LAI relationships are present, such as in the open canopy situations of this study.

Cover was defined by Mueller-Dombois and Ellenberg (1974) as "the vertical projection of the crown or shoot area of a species to the ground surface expressed as a fraction or percent of a reference area". Bunnell and Vales (1990) suggested mean crown completeness (MCC) as a useful measure of canopy cover as observed from the ground. MCC is a stand or plot measurement denoting the mean of several measurements of the proportion of the sky covered by tree crowns within a specified angle from a single point (Bunnell and Vales, 1990). The area sampled is determined by trigonometric principles and depends on the angle of view, the height to base of live crown and the height from which the angle is projected. Bunnel et al. (1990) and Vales and Bunnel (1988) compared several different cover measurement techniques including ocular, moosehorn, spherical densiometer,

regular and hemispherical photographs to evaluate differences among techniques and effects of the observers. They agreed that narrow angles of view and vertical projections are the least biased ways of estimating MCC to include both gaps between and within crowns. They observed that wide angles of view masked small gaps, included objects not directly above the point sampled, and had an angular view of the canopy at the outer edges, overestimating MCC. On the other hand, Bunnel et al. (1990) suggested that wide angles may be more appropriate when examining relationships between overstory and understory radiation in boreal forests, where low solar angles are present.

Digital photography has the advantage of offering a permanent and objective record of the samples that allows flexible laboratory analysis and eliminates observer effects.

This study aims to compare different methods of obtaining maximum growing season leaf area index maps using remote sensing data, and LAI and tree cover field data in a boreal forest near Thompson, Manitoba, Canada. The comparison includes aspatial methods such as traditional regression, inverse regression and reduced major axis, and geostatistical techniques such as kriging, cokriging, kriging with an external drift, and conditional simulation. The LAI maps will serve as input in process models to obtain net primary production.

The present work was done in the context of the BigFoot project

(http://www.fsl.orst.edu/ larse/BigFoot) which focuses on the validation of the

MODIS (Moderate Resolution Imaging Spectrometer) land cover, LAI/fAPAR

(fraction of absorbed photosynthetically active radiation), and NPP products

(http://modarch.gsfc.nasa.gov/MODIS) (Cohen and Justice, 1999). The main

objective of BigFoot is scaling up from *in situ* ground measurements to the

moderate spatial resolution of MODIS data products (250 - 1000 m spatial

resolution).

Mapping LAI with Aspatial and Geostatistical Methods

This section reviews briefly the theory of the different aspatial and spatial LAI

estimation and simulation techniques tested in this study (Fig. 1).

Aspatial Regression-based Methods

Commonly used SVIs are based in traditional regression (Reg_T), and are expressed

as: $Y = \alpha + \beta X + \varepsilon$,

where α is the intercept, β is the slope, ε is the error, assumed to be independent

and normally distributed, Y is LAI and X is reflectance or one of several SVIs.

Regression methods assume that data are spatially independent and that there are

no measurement errors in the independent variable; it is designed to estimate Y by

minimizing the sum of squares errors in Y.
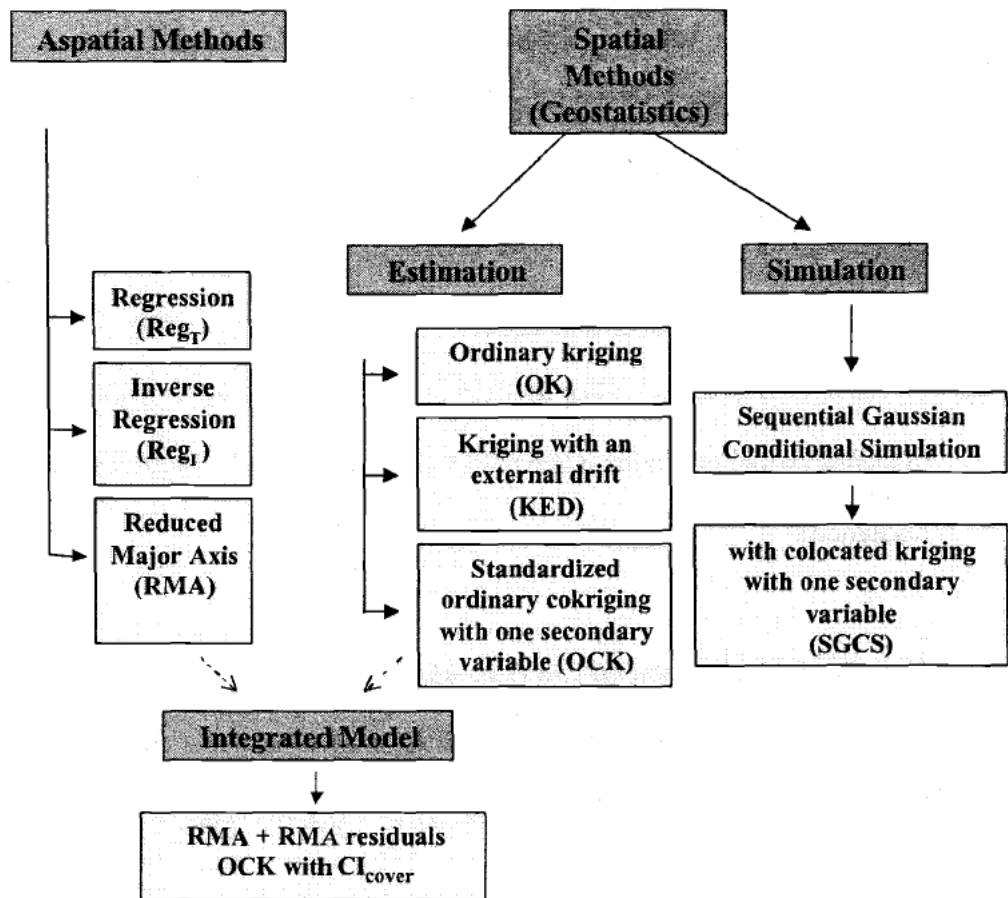
Figure 1. Scheme of methods used in this study.

Curran and Hay (1986) described the major measurement errors that should be accounted for in remote sensing variables, and that are generally ignored when using traditional regression. This problem is solved by using the inverted equation $X = \alpha + \beta Y + \varepsilon$, referred to as inverse regression (Curran and Hay, 1986), or Reg_I, which minimizes the sum of squared errors in X. However, Chen and Cihlar (1997) pointed out that commonly used allometric methods to estimate LAI can accumulate errors, and it is difficult to keep the total error under 25 %. A method

that accounts for measurement errors in both dependent and independent variables is needed in such situations. The reduced major axis method (RMA) improves the traditional and inverse regression procedures in that it minimizes the sum of the cross-products of the differences on both axes, accounting simultaneously for the errors in both dependent and independent variables (Miller and Kahn, 1962; Davis, 1986) and is given by $Y = (\alpha + \beta X + \varepsilon) * (-1)$.

The main difference among these methods is that traditional and inverse regression equation coefficients (α and β) are determined by least squares, while reduced major axis intercept and slope are given by:

$$\alpha = \overline{Y} - \frac{\sigma_Y}{\sigma_X} * \overline{X}$$

$$\beta = \frac{\sigma_Y}{\sigma_X}$$

where $\overline{Y}$ and $\overline{X}$ are the means of the dependent and independent variables respectively, and σ_Y and σ_X are their standard deviations.

The Geostatistical Framework

Geostatistics, based on the theory of regionalized variables (Journel and Huijbregts, 1978), is concerned with a variety of techniques aimed at understanding and modeling spatial variability through estimation and simulation (Deutsch, 2000; Journel, 1989; Goovaerts, 1997). A regionalized variable equation differs from a

regression equation in that its components are indexed by their position in space or time:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon''$$

where $Z(x)$ is the variable to be estimated at location x, $m(x)$ is a deterministic component, $\varepsilon'(x)$ is spatially correlated variability, and ε'' is a spatially independent random residual, assumed to be normally distributed and interpreted as nugget variance. Geostatistics accounts for the presence of spatial autocorrelation and joint dependence in space and time, which occur in most natural resources variables (Myers, 1997).

In an ecological context, geostatistics has been used to describe scale and pattern of spatial variability (Woodcock et al., 1988; Legendre and Fortin, 1989; Turner et al., 1990; Rossi et al., 1992), to characterize canopy structure (Cohen et al., 1990; St-Onge and Cavayas, 1997; Hudak and Wessman, 1998; Wulder et al., 1998), to estimate continuous and categorical variables (Rossi et al., 1993; Milne and Cohen, 1999), and to assess risk (Myers, 1997; Saito and Goovaerts, 2000).

Biogeochemical models are increasingly adopting an explicitly spatial configuration. Spatial surfaces of meteorological values such as temperature and evapotranspiration or accurate digital elevation models (DEM) are commonly required as model inputs (Running and Nemani, 1987). Geostatistics is currently used to improve such data layers (Goovaerts, 2000; Kyriakidis, 1999). The latest

applications of geostatistics have emphasized the use of models of uncertainty that depend on the data values in addition to data configuration (Deutsch and Journel, 1998). Stochastic simulation is an example of a probabilistic approach that provides a distribution of possible values for each cell of the surface, characterizing uncertainty. These uncertainty measurements improve ecological interpretation, help assess error areas and decrease losses and risks in policy and management decision-making (Rossi et al., 1993).

Geostatistical Estimation Procedures

Geostatistics is based on the concepts of random variables and random functions (Isaaks and Srivastava, 1989). A random variable (RV) is a variable whose values are randomly generated according to some probabilistic mechanism that models the uncertainty about the attribute under study. In predictive statistics, any unsampled (unknown) value is characterized as a random variable. The probability distribution of the random variable is usually location and information-dependent (Deutsch, 1998). The random variable seen as a function of spatial location is called a random function (RF). A random function is a set of random variables that have associated spatial locations and whose dependence on each other is specified by some probabilistic procedure (Isaaks and Srivastava, 1989; Deutsch, 2000).

Kriging is a generalized least-squares linear regression technique that accounts for the spatial dependence among observations. The kriging estimators are exact

interpolators, meaning that they honor or reproduce the sample values at their locations (Goovaerts, 1997). Three different kriging variants were applied in this study: ordinary kriging (OK), kriging with an external drift (KED) and ordinary cokriging (OCK).

Ordinary Kriging

Ordinary kriging is a univariate technique that accounts for local fluctuations in the mean value by limiting the domain of stationarity of the mean to a local neighborhood centered on the location being estimated (Goovaerts, 1997). The ordinary kriging estimator is calculated as follows:

$$Z^*_{OK}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{OK}_{\alpha}(u) Z(u_{\alpha})$$

under the constraint:

$$\sum_{\alpha=1}^{n(u)} \lambda^{OK}_{\alpha}(u) = 1$$

where $Z^*_{OK}(u)$ is the ordinay kriging estimator at location u, $\lambda^{OK}_{\alpha}(u)$ are the OK weights corresponding to the n samples at location u, and $Z(u_{\alpha})$ are the nearby sample values.

Kriging minimizes the error variance and aims to obtain a mean residual error equal to 0. These constraints and the specification of a pattern of spatial continuity (a

model applied to the experimental semivariogram) lead to a normal system of linear equations called the ordinary kriging system that provides the OK weights. For a detailed description of the kriging computation refer to Isaaks and Srivastava, (1989), Goovaerts, (1977), or Deutsch and Journel, (1998).

The kriging error variance is dependent on the data configuration but independent of the data values, so it is not a good measure of error. The variation is always 0 at sample data locations, increases away from the data and reaches a maximum value at locations most distant from sample data points.

Kriging with an External Drift

Kriging with an external drift (KED) is a variant of kriging that allows for the use of secondary information known at every location (exhaustive), which is assumed to reflect the spatial trend of the primary variable (Deutsch and Journel, 1998; Goovaerts, 1997). At a landscape scale, spatial variation can be decomposed into two components: large-scale variation and small-scale variation. Large-scale variability may be influenced by geomorphology, elevation, slope, aspect, precipitation, or disturbances like fire or disease. Small-scale variability may be influenced by soil permeability, nutrient availability, or pH. The KED trend represents the large-scale variability of the primary variable. The residuals from the trend represent the small-scale variability, and the final KED result combines both. KED models the trend under the assumption of a linear relationship between

primary and secondary variables, and smooth variation of the secondary variable. The distinctive feature of KED is that the algorithm considers a non-stationary random function model, where stationarity is limited within each search neighborhood, yielding more local detail (Deutsch and Journel, 1998). The KED estimator is written:

$$Z^*_{KED}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{KED}_\alpha(u) Z(u_\alpha)$$

where $Z^*_{KED}(u)$ is the kriging with an external drift estimator at location u,

λ^{KED}_α (u) are the KED weights corresponding to the n samples at location u,

and $Z(u_\alpha)$ are the nearby sample values.

The procedure consists of three steps (Goovaerts, 1997); first, the trend coefficients $a^*_0(u)$ and $a^*_1(u)$ of the trend model $m^*_{KED}(u)$ are evaluated within the search neighborhood, from the $n(u)$ data pairs $(z_1(u), z_2(u))$, where $z_1(u)$ are the primary variable sample points, and $z_2(u)$ are the secondary variable sample data. These coefficients are estimated through the kriging system. Then, the trend components $m(u)$ are estimated at all primary sampled locations and at all locations being estimated. Finally, a simple kriging is performed in the residuals of the trend:

$$Z^*_{KED}(u) - m^*_{KED}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{SK}_\alpha(u)[Z(u_\alpha) - m^*_{KED}(u_\alpha)]$$

where $m^*_{KED}(u)$ is the trend component, estimated as

$$m_{KED}^*(u) = a_0^*(u) + a_1^*(u)z_2(u) \qquad \lor \alpha = 1,\ldots\ldots, n(u)$$

where α are the sample data, and λ_α^{SK} are the simple kriging weights.

The minimized error variance is:

$$\sigma_{KED}^2(u) = C_R(0) - \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{KED(u)} C_R(u_\alpha - u) - \sum_{k=0}^{K} \mu_k^{KED}(u) f_k(u)$$

where $C_R(0)$ is the residuals' stationary covariance, $\mu_k^{KED}(u)$ are the Lagrange

parameters, and $f_k(u)$ are the trend functions. KED is also useful to estimate the

trend itself, not only acting as a local estimation procedure, but also as a global one

(Isaaks and Srivastava, 1989).

Standardized Ordinary Cokriging

Standardized ordinary full cokriging (OCK) is a multivariate extension of kriging

that allows for the use of secondary data, accounting for the spatial cross

correlation between the primary (undersampled) and secondary (exhaustive)

variables. The OCK estimator is written:

$$Z_{OCK}^*(u) = \sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{OCK}(u)Z_1(u_{\alpha_1}) + \sum_{\alpha_2=1}^{n_2(u)} \lambda_{\alpha_2}^{OCK}(u)Z_2(u_{\alpha_2})*[Z_2(u_{\alpha_2}) - m_2 + m_1]$$

with the constraint

$$\sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{OCK}(u) + \sum_{\alpha_2=1}^{n_2(u)} \lambda_{\alpha_2}^{OCK}(u) = 1$$

where $Z^*_{OCK}(u)$ is the standardized ordinary cokriging estimator at location u,

$\lambda^{OCK}_{\alpha i}(u)$ are the ordinary cokriging weights, and m_1 and m_2 are global means of

the primary and secondary variables, respectively.

Cokriging requires a joint model that includes the modeling of the direct

semivariograms of primary and secondary variables and their respective cross-

semivariograms under the condition that the whole set of semivariograms be

positive definite (Isaaks and Srivastava, 1989; Goovaerts 1997). This need for

positive definiteness imposes many constraints on the models that can be chosen; a

relatively easy way of checking this property is the linear model of

coregionalization (LMC).

The LMC is defined as a set of direct and cross-semivariogram models $\gamma_{ij}(h)$ such

that:

$$\gamma_{ij}(h) = \sum_{l=0}^{L} b^l_{ij} g_l(h) \quad \forall i, j$$

where each function $g_l(h)$ is a permissible model and the $(L+1)$ matrices of

coefficients b^l_{ij}, corresponding to the sill or slope of the model $g_l(h)$, are all

positive semi-definite. Refer to Goovaerts, (1997) for an exhaustive development

of the linear model of corregionalization.

The main differences between OCK and KED are first, that the cokriging estimates are directly influenced by the secondary data, while in the KED approach, secondary data only provide information about the trend; and second, that the secondary information influences the KED depending on the slope of the trend, while OCK accounts for the global correlation of the primary and secondary data through the cross-semivariogram.

Integrated Models

In cases where a trend is present, it is better to model the trend using physical knowledge of the phenomenon under study (Isaaks and Srivastava, 1989). Such an approach involves modeling the trend from the observed sample values (i.e. through regression), the subtraction of the trend from the sample values to obtain the residuals, and the interpolation estimation on the residuals (i.e. through cokriging), adding up the trend and the estimated residuals surfaces at the end.

Metzger, (1997) suggested a combination of trend surface analysis and kriging or cokriging to combine the use of remote sensing data and field data, and she improved the estimation of certain forest variables. The conceptual idea underlying this procedure is to exploit large-scale variability using remote sensing data through the trend surface analysis, and to describe small-scale variability using the residuals of the trend model cokriged with a secondary exhaustive source of information.

Geostatistical Simulation Procedures

Stochastic simulation is a probabilistic approach that provides a distribution of
multiple, equally probable realizations of the joint distribution of one or more
attribute values in space, generating a model of spatial uncertainty (Goovaerts,
1997; Rossi et al., 1993).

Several different simulation algorithms are used to model spatial uncertainty,
depending on the purposes of the final outputs. The approach used in this study was
sequential Gaussian conditional simulation (SGCS) with simple collocated
cokriging with one secondary variable.

The general idea of SGCS with collocated cokriging is to produce a set of joint
realizations of the spatial distribution of the primary attribute, conditional to both
the primary and collocated secondary information.

The collocated simple cokriging estimate of the primary attribute is:

$$Z_{SCK}^*(u) = \sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{SCK}(u)[Z_1(u_{\alpha_1}) - m_1] + \lambda_{\alpha_2}^{SCK}(u)[Z_2(u) - m_2] + m_1$$

where $Z_{SCK}^*(u)$ is the colocated simple cokriging estimator at location u,

$\lambda_{\alpha i}^{SCK}(u)$ are the collocated simple kriging weights, and m_1 and m_2 are global

means of the primary and secondary variables, respectively.

Unlike full cokriging, collocated cokriging just uses the lag correlation function

between the primary and secondary variables, following the Markov-type

approximation:

$$\rho_{12}(h) \approx \rho_{12}(0) * \rho_{11}(h)$$

where $\rho_{12}(h)$ is the lag cross-correlation function of the primary and secondary

variables, $\rho_{12}(0)$ is the correlation coefficient between the primary and secondary

variables, and $\rho_{11}(h)$ is the lag correlation function of the primary variable.

This approximation avoids modeling the linear model of coregionalization; only the

semivariogram of the primary variable needs to be modeled. The cross-variogram

model is derived as a linear rescaling of the primary variable semivariogram model

by the correlation coefficient $\rho_{12}(0)$. In collocated cokriging, the dependence of

the secondary variable on the primary one is limited to the collocated datum to

avoid matrix instability problems caused by highly redundant secondary

information, and to speed up the process (Goovaerts, 1997).

The smoothing effect of kriging and cokriging is due to a missing variance

component (Deutsch, 1998). The conditional simulation realizations have added

back the missing variance estimated by cokriging, representing in a more realistic

way the spatial distribution of the features observed in the original data. Because

spatial heterogeneity is a natural feature of ecosystems, conditional simulation

realizations have the advantage of representing high spatial entropy, producing more realistic output displays of the variable in terms of spatial variability.

The variance of the cokriged estimate is

$$Var\{Z^*_{SCK}(u)\} = \sigma^2 - \sigma^2_{SCK}(u)$$

where $Var\{Z^*_{SCK}(u)\}$ is the variance of the simple collocated cokriged estimate, σ^2 is the variance of the stationary random function or the variance at lag = 0, and $\sigma^2_{SCK}(u)$ is the simple collocated cokriging variance.

So, the cokriging variance of the estimate is smaller than the total variance and that explains why cokriging maps are smooth. A way of incorporating this missing variance is to add an independent random component $R(u)$, while reproducing the covariance properties of cokriging. The simulated value $Z_S(u)$ would then be:

$$Z_S(u) = Z^*_{SCK}(u) + R(u)$$

The SGCS algorithm used here (Deutsch, 1998) proceeds as explained in the following steps (Deutsch, 2000; Goovaerts, 1997). First, the original data are transformed to a standard normal distribution. A multivariate multi-point Gaussian random function model is adopted if the joint normal score random functions are normally distributed. Subsequently, a random path to visit once all the nodes of the grid is defined. Then, a location u is visited randomly and collocated cokriging is performed to obtain the cokriged estimate and its corresponding variance, using the

conditional information provided by neighboring primary and collocated secondary

normal score data. Next, the algorithm draws a random residual (through Monte

Carlo simulation) that follows a normal distribution with mean equal to 0 and

variance equal to the cokriging variance. The cokriged estimate and the random

residual are added to obtain the simulated value. The simulated value is included in

the data set and the algorithm proceeds to the next node, looping until all nodes are

simulated. Finally, the simulated normal scores are backtransformed to their

original values. Another realization can be obtained by repeating the process with a

different random visiting path.

Simulations were developed initially to provide measures of spatial uncertainty, but

they have increasingly been used as maps of the variable of interest in cases where

the reproduction of the spatial variability is more important than local accuracy and

where sample and exhaustive data are available (Deutsch, 1998). Unlike kriging,

conditional simulation emphasizes the global accuracy of the simulated variable,

reproducing its sample statistics (mean, histogram, covariance) and its pattern of

spatial continuity (Deutsch, 1998; Dungan, 1999). Another difference with

interpolation algorithms is that stochastic simulation provides measures of local

and global joint accuracy. Table 1 summarizes the relevant properties of the

methods used. It also provides information for use in comparing the methods.

52

Methods and Materials

Study Site Description

The study site (Fig. 2) is within the Boreal Ecosystem-Atmosphere Study

(BOREAS) Northern Study Area (NSA), near Thompson, Manitoba, Canada, at

latitude 56°N and longitude 98°W (http://www.eosdis.ornl.gov/BOREAS/bhs/bhs.html).

It lies within the Canadian Shield Province, and has a gentle topography that

reflects the glacially smoothed Pre-Cambrian bedrock surface modified by glacial

drift. Most of its soils are derived from Glacial Lake Agassiz sediments and consist

of clays, organics, and some

sand deposits. Bogs and fens are present in low areas, and permafrost can often be

found a few feet below the surface of bogs or at greater depths in clay soils. Being

flat and having abundant wetland areas, the drainage of much of the surface is poor.

There are several tributaries of the Sapochi and Odei Rivers and a few small lakes.

There are five major overstory cover types in the study site. The predominant

species is black spruce (*Picea mariana*), of variable density, with stand ages up to

80 years. Black spruce may be present in muskeg, an open canopy cover type,

associated with tamarack (*Larix laricina*), or in closed canopies, mixed with

tamarack and a low occurrence of balsam poplar (*Populus balsamifera*) and jack

pine (*Pinus banksiana*).

UTM X: 535350
UTM Y: 6196025

N

a

UTM X: 530125
UTM Y: 6190850
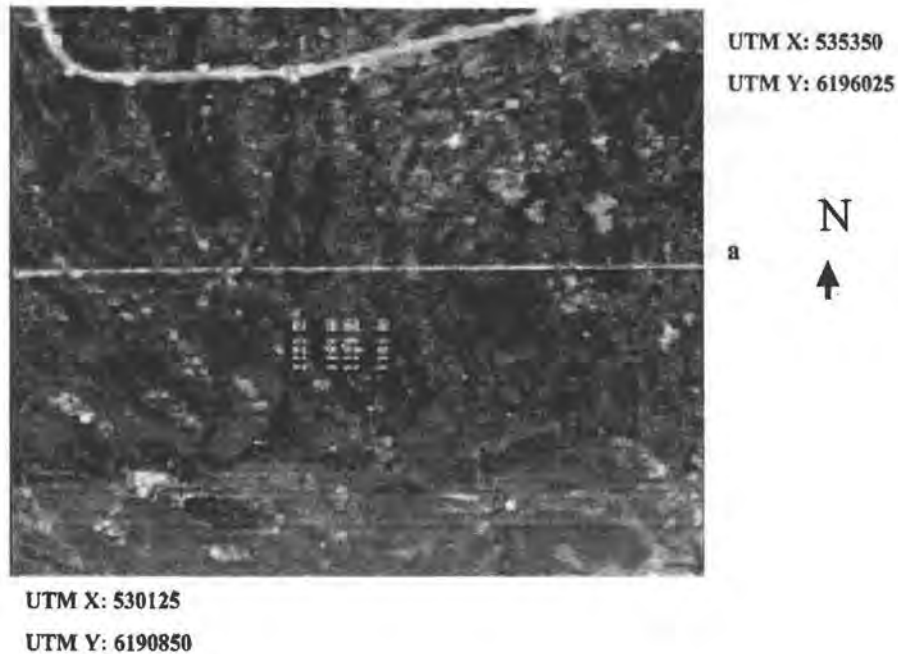
Figure 2. Landsat ETM+ 432 (RGB) composite image of the 5 by 5 km BigFoot study area. Yellow dots represent actual sample plot locations.

The aspen cover type is dominated by trembling aspen (*Populus tremuloides*), distributed in small patches, and mixed with white spruce (*Picea glauca)*, balsam poplar, black spruce and jack pine. Wetlands are covered by scattered bog birch (*Betula spp.*), and tamarack, and are located in flooded peatlands, creek borders, and beaver ponds. Jack pine stands, associated with white spruce, balsam poplar, black spruce and aspen, are located preferentially on sandy upland soils.

A distinctive feature of the flora of boreal forests is the abundance and diversity of sphagnum, feathermosses and lichens. Many vascular plants are also present such

as *Vaccinium spp.*, *Cornus spp.*, *Ledum spp.* (labrador tea), *Rose spp.* (wild roses), *Salix spp.* (willow), Alnus spp. (alder), and *Bromus spp.*, *Poa spp.*, *Calamagrostis spp.* (grasses). Regeneration of the overstory species at different stages of development was also observed in the area.

The boreal forest is well adapted to development following natural disturbances, which include fire, floods, snow breakage, and insect outbreaks. *Pinus banksiana* has serotinous cones, and *Picea mariana* has semiserotinous cones. A large burn occurred in 1981 in the southern section of the 5 by 5 km area (Fig. 2). No logging has occurred in the study area.

Data Collection

Sampling Design

This study used the BigFoot field sampling design for BOREAS NSA, which is a systematic spatial cluster (Fig. 3 b), based on Clinger and Vann Ness, (1976). This design is a spatial application of unequally, but cyclically, spaced time points of a discrete time series that distributes pairs of plots at different distances (lags), a major advantage for computing geostatistics. The design is efficient in that it allows the retention of the sampling density needed, while decreasing redundant measurements at constant lag increments (Fortin et al., 1989). Burrows et al., (2001) used this sampling scheme and found that it maximized information about the spatial variability of vegetation in heterogeneous landscapes at the landscape

level, and was more efficient than other sampling designs such as random sampling. This nested sampling allowed the coverage of a range of spatial scales between 25 and 500 m.

LAI and cover were measured at each of the 106 plots (86 within the BigFoot flux tower footprint, and 20 outside the flux tower footprint, (Fig. 2, 3 and 4)). The present analysis considered only the 86 plots within the BigFoot flux tower footprint (Fig. 4). Each plot was 25 x 25 m, to match the resolution of resampled Landsat ETM+ data. The area covered by the 86 plots within the flux tower footprint was 925 m east to west and 550 m north to south, although the modeled LAI maps were developed over a total extent of 1.82 km^2 (Fig. 4).

Each plot had multiple fixed area sub-plots arranged in such a way that they were evenly spread and did not overlap each other (Fig. 3 c). These sub-plot measurements were then averaged to produce a unique value per plot.

To assess the representativeness of the data captured with the sampling design, an evaluation was performed in the context of a land cover map developed by Cohen et al., (in preparation) (Fig.5).

56
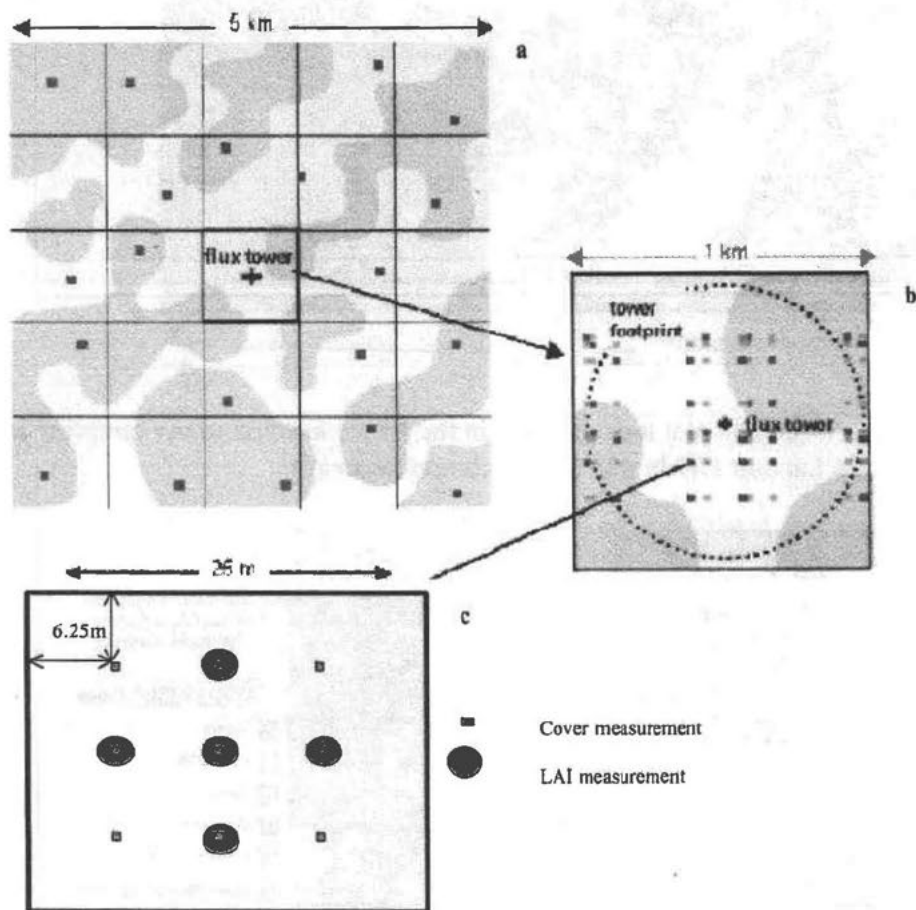
Figure 3. Idealized sampling design at a) 5 by 5 km, b) 1 by 1 km, and c) single plot 25 by 25 m.

57
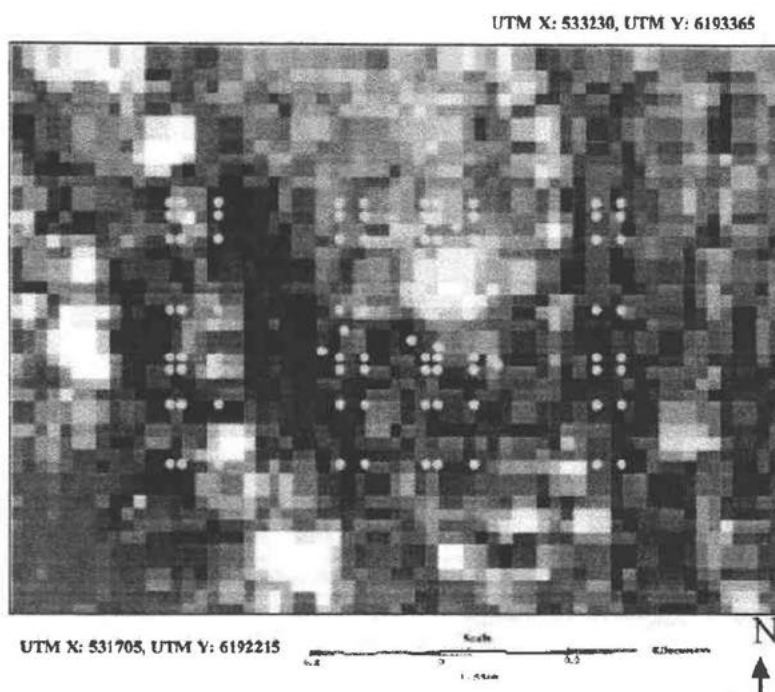
Figure 4. Sample plot locations within the 1 by 1 km flux tower footprint cell on top of a Landsat ETM+ 432 (RGB) composite image.
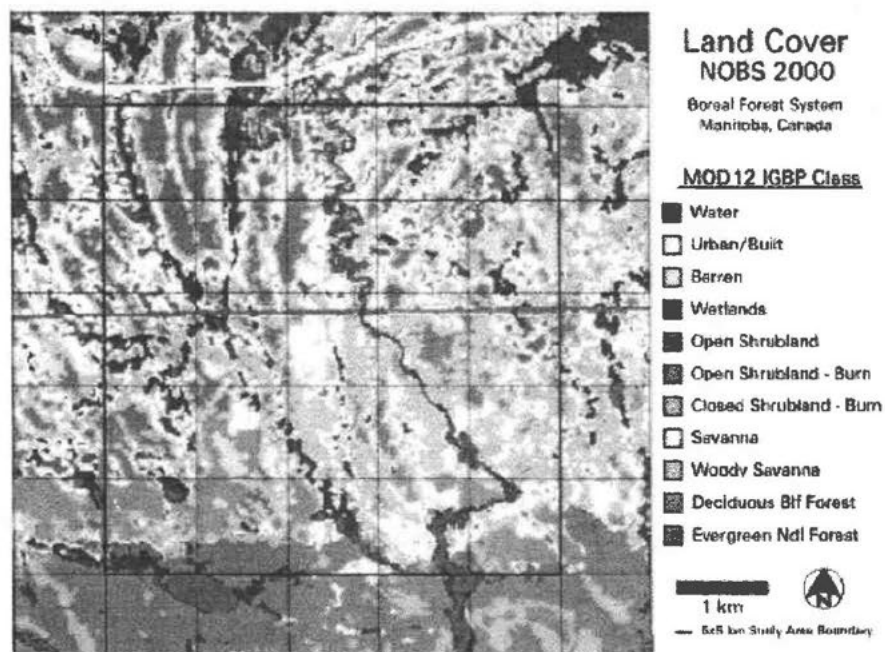
Figure 5. IGBP land cover classification for NOBS's 5 by 5 km study area.

LAI Measurements

In this study, LAI is expressed as one-half total overstory leaf surface area per unit

ground surface area. LAI was measured by the BigFoot team via harvesting and

allometric equations following the methods summarized by Gower et al., (1999)

and Campbell et al., (1999).

Cover Measurements

Canopy cover was quantified with a digital, true-color camera (Sony Mavica FD91)

oriented vertically on a monopod at a height of 1.70 m above ground. Photographs

were taken using an angle of view of 30°. Initially, four sub-plot measurements

were suggested to measure cover within each plot, but a pilot study during the

summer of 1999 determined that a better characterization of cover, with an

acceptable standard error for all components, required nine sub-plots located on a

regular grid (Fig. 3 c) (Cohen et al, 1999). Previous the analysis, the digital photos

were enhanced with adjustments for color balance, contrast, and saturation using

Adobe Photoshop v5.0.2, (1998). Several grid densities were analyzed and finally a

reticular grid with 96 intersection points and a reliability of 11.5 % was selected

(Thompson, 1992) for cover determination. The proportions of conifer, hardwood

and snag were then quantified on the computer screen at the intersection points of

the grid, transformed to percents, and averaged for the nine photos covering each

plot to obtain a unique value per plot. The total cover values used in the

calculations included only conifer and hardwood values and excluded snags.

Data Analysis

Satellite Image Processing

This study used a Landsat ETM+ image (path 33/ row 21), acquired on July 10th,

1999, and an IKONOS image, acquired on May 20th, 2000. The July 1999 image

digital numbers were converted to percent reflectance following the COST model

as developed by Chavez (1996).

The IKONOS panchromatic band (1m x 1m) was georeferenced to UTM, WGS84

using ground control points provided by BigFoot personnel, measured using a GPS

with real-time correction and accurate and accurate to the nearest 0.1 m (Cohen et

al., 1999). The July 1999 image was then coregistered to match the panchromatic

IKONOS and resampled to 25 m using cubic convolution in ERDAS Imagine v8.3

(http://www.erdas.com). Misregistration was in the order of one pixel or less.

Reflectance values of Landsat ETM+ bands were extracted for the 86 plots within

the flux tower footprint, and NDVI ([NIR – Red] / [NIR + Red]) and SR (NIR /

Red) spectral vegetation indices were calculated. Two canonical correlation

analyses (CCA) were performed, the first having LAI on one side of the equation,

and the ETM+ bands on the other side, and the second constructed in the same way,

this time using total cover instead of LAI. The CCA was preferred over of a

multivariate regression to obtain a unique set of values (the canonical index)

representing the linear combinations with the largest possible correlation of the two

60

sets of variables (Johnson, 1998; Ramsey and Shafer, 1997) that allowed the

posterior application of the inverse regression and reduced major axis procedures.

Using the modeler tool in ERDAS Imagine, the two canonical index images (CI_{LAI} and CI_{cover}) were created based in the following equation:

$$CI_v = \sum SCC_{b_x} * \frac{b_x - mean_{b_x}}{std_{b_x}}$$

where:

CI_v is the canonical index, where v is LAI or cover; SCC_{bx} is the standardized canonical coefficient for band x, $mean_{bx}$ is the mean of the band x values used in the CCA calculation, and std_{bx} is the standard deviation of the band x values used in the CCA calculation.

Regression Methods

Traditional regression (Reg$_T$), inverse regression (Reg$_I$) and a reduced major axis regression (RMA) were performed as follows:

Traditional Regression

$$LAI = 4.19 - 1.23\ CI_{LAI}$$

Inverse Regression

$$CI_{LAI} = 1.83 - 0.44 * LAI \quad, \text{ solved for LAI as } \quad LAI = \frac{CI_{LAI} - 1.83}{0.44}$$

Reduced Major Axis

$$LAI = (4.19 - 1.68 * CI_{LAI})(-1)$$

where the intercept is given by:

$$\alpha_2 = 4.19 - \frac{1.68}{1.00} * CI_{LAI}$$

and the slope is defined as the ratio of the standard deviation of the two variables:

$$\beta_2 = \frac{1.68}{1.00}$$

These three regression methods are aspatial in that the LAI estimations are derived only from the explanatory variables at each single sampling point, regardless of surrounding sample point information. These approaches assume that the residual values are spatially uncorrelated. Inspection of the residuals' spatial autocorrelation for the three models was done using the Moran's I and Geary's C coefficients calculated at a 0.05 significance level (Griffith, 1987; Reich, 1999), and by visual inspection of their respective semivariograms.

Moran's I coefficient:

$$I = \frac{N \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij} Z_i - Z_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{ij} \sum_{i=1}^{N} Z_i^2}$$

Geary's C coefficient:

$$C = \frac{(N-1)}{2(\sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{ij})} * \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij} (Z_i - Z_j)^2}{\sum_{i=1}^{n} Z_i^2}$$

where: δ_{ij} are the weights of a distance weight matrix. These coefficients are interpreted as follows: as Moran's I values approach 1, there is positive spatial

correlation; values near 0 indicate that data are spatially independent; as Geary's C values approach 0, there is positive spatial correlation; values near 1 denote independence, while values near 2 point out negative spatial correlation.

Geostatistical Methods

Geostatistical procedures were perfomed with GSLIB (Geostatistical software library, Deutsch et al., 1998). For all geostatistical analyses, the data were first transformed using a normal score transformation (Goovaerts, 1997; Deutsch and Journel, 1998). The normal scores transform is a non-linear, rank-preserving transformation that matches the original data to a standard normal distribution (zero mean, unit variance) (Fig. 6). After geostatistics procedures were performed, and before mapping the results, the predictions were backtransformed to their original data distributions.

An exploratory analysis of the data consisted of constructing location maps for LAI and cover, histograms, scatterplots, semivariograms and cross-correlograms. Pearson's correlations were calculated for all of the variables. Omnidirectional semivariograms for the normal score LAI and cover sample values were developed. Because a periodic behavior was observed, omnidirectional semivariograms at different spacings (lags = 25, 50, 75, and 100m) were computed for the normal score LAI values, to see if this behavior was present at other scales.

Cross-correlograms were also calculated between the primary (LAI) and auxiliary variables (CI_{LAI} and CI_{cover}). These data descriptors helped to reveal important spatial and aspatial aspects of the data.

Declustering techniques were not applied because data were not spatially clustered due to the pattern of the sampling design used; declustering-weight magnitudes were very close to one another.
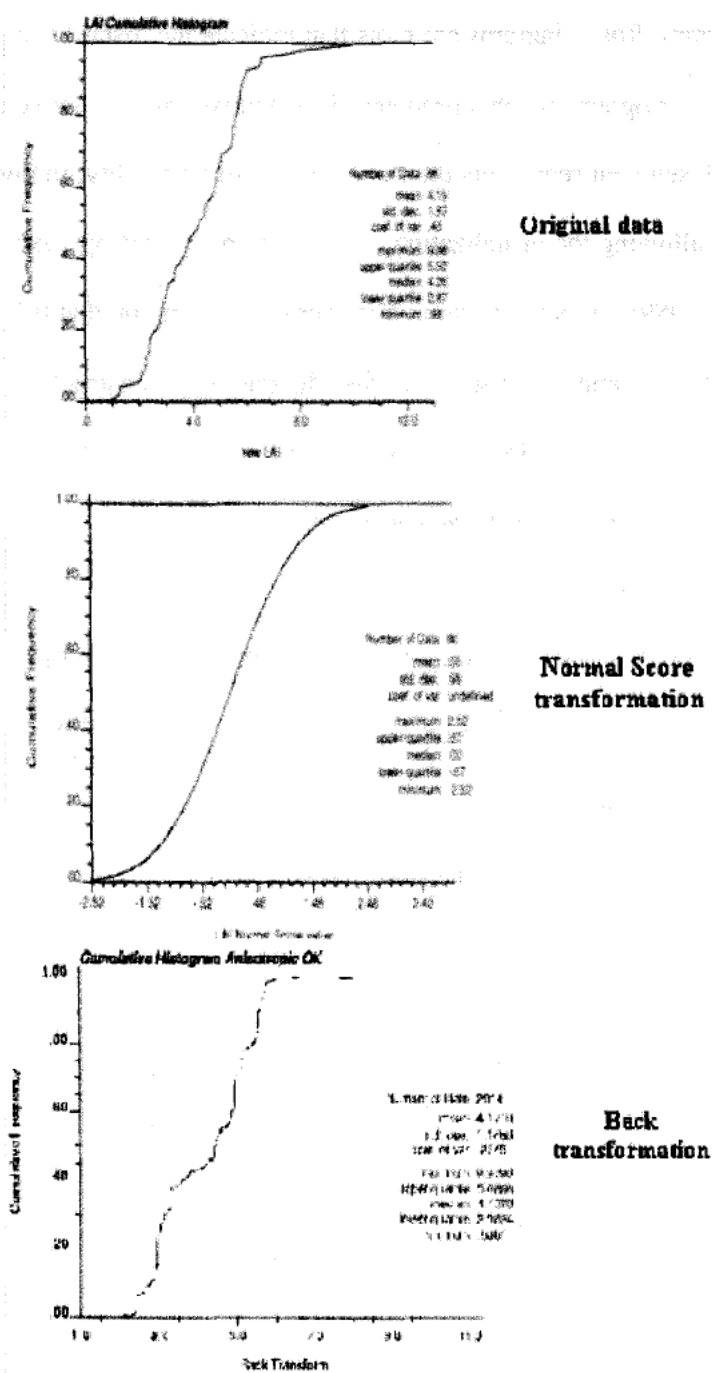
Figure 6. Example of normal scores transformation and backtransformation for LAI.

An anisotropy analysis was performed for LAI with directional semivariograms and a rose diagram. Rose diagrams are plots that indicate the distances at which directional semivariograms reach a predetermined semivariance in those different directions. Each segment represents the same amount of variability, though the distances vary, allowing the visualization of the principal anisotropy axes (Isaaks and Srivastava, 1989). Directional semivariograms were computed at 0, 10, 30, 50, 70, 90, 100, 120, 140, and 160 degrees. A rose diagram was constructed for the LAI normal scores. The directions of maximum and minimum continuity were identified and used to construct the anisotropic models.

Experimental semivariograms were modeled with different positive linear combinations of three permissible models (Deutsch and Journel, 1998; Goovaerts, 1997): nugget effect, spherical, and hole effect. In all situations, the simpler the model, the lower the cross validation errors shown.

1. Nugget effect model

$\gamma(h) = 0$ if $h=0$, or c otherwise

2. Spherical model

$$\gamma(h) = c * Sph\left(\frac{h}{a}\right) = \left\{ c*\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right] \right. , \text{ if } h \le a \text{ or } c \text{ otherwise}$$

3. Hole effect model

$$\gamma(h) = c * \left[1.0 - \cos\left(\frac{h}{a} * \pi \right) \right]$$

where

c is a positive variance contribution or sill value and a is the actual range.

The geostatistics algorithms applied consider a limited nearby conditioning data closest to the location being estimated, called the search neighborhood. These local search neighborhoods limit the stationarity assumption to small areas, allow the calculation of local trends and means required for some of the isotropic and anisotropic methods, and decrease computational time. GSLIB allows the user to manage search neighborhoods by changing their size, shape, and the amount of closest data to be used in the calculation, with closeness measured by the Euclidean distance. Cross validation was used to evaluate the influence of different search parameters (Deutsch and Journel, 1998).

Ordinary Kriging

Isotropic and anisotropic ordinary kriging methods were performed. The isotropic model for LAI was a combination of a spherical plus a hole effect model, to account for the periodic behavior of the experimental semivariogram (Fig. 7):

$$\gamma_{LAI_{OK}}(h) = 0.05 + 0.76 * sph_{(130m)} + 0.17 * hole_{(115m)}$$

The anisotropic model for LAI was given by the combination of one model at azimuth 70 degrees (minimum direction of continuity) and the other at azimuth 160 degrees (maximum direction of continuity) as follows:

Azimuth 70: $\gamma_{LAI}(h) = 0.05 + 0.95 * sph_{(90m)}$

Azimuth 160: $\gamma_{LAI}(h) = 0.05 + 0.95 * sph_{(135m)}$

The assembled model (Fig. 8) was:

$$\gamma_{LAI_{OK}}(h) = 0.05 + 0.95 * sph_{(h\min=90, h\max=135)}$$

This is a case of geometric anisotropy, where the range changes with direction while the sill remains constant.
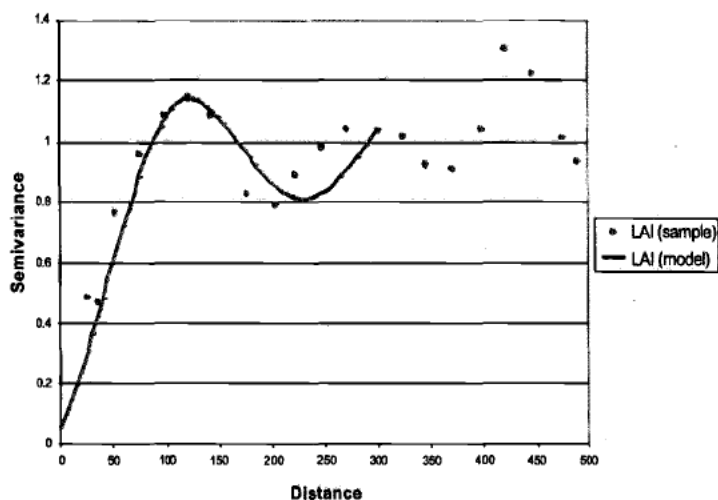
Figure 7. Experimental omnidirectional LAI semivariogram and isotropic model.
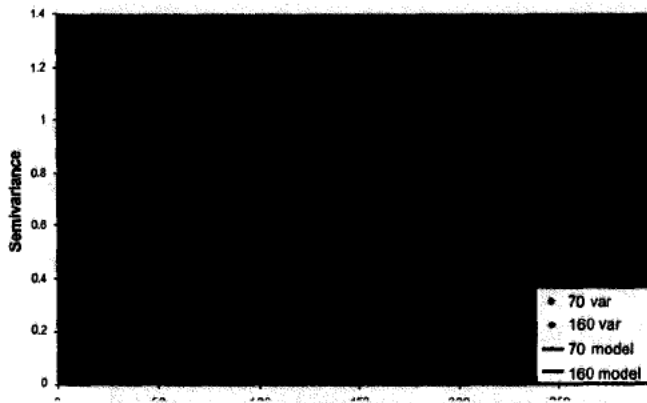
Figure 8. Experimental directional LAI semivariograms in 70 and 160 degrees and anisotropic model.

Kriging with an External Drift

KED allows for the use of exhaustive secondary information. In this case, CI_{cover} was chosen to characterize the trend of the primary attribute (LAI), then a simple kriging was performed on the corresponding residuals. KED assumes that the primary and secondary variables are linearly related. This assumption was met because LAI was linearly related to CI_{cover}. This relationship makes physical sense, since LAI was related to both cover and the reflectance captured by the sensor.

The anisotropic model used to perform KED (Fig. 8) was given by:

$$\gamma LAI_{KED}(h) = 0.05 + 0.95 * sph_{(h\min=90, h\max=135)}$$

The trend was kriged with ordinary kriging.

Standardized Ordinary Cokriging

Another approach for incorporating secondary information is cokriging. The application of cokriging requires modeling the semivariograms for the primary variable (LAI) and the secondary variable (CI$_{cover}$), and the cross-semivariogram for both variables (LAI-CI$_{cover}$). The linear model of corregionalization allows for the modeling of these direct and cross-semivariograms so that the variance of the variables is always positive. The linear model of corregionalization for LAI and CI$_{cover}$ (Fig. 9) was given by:

$$\gamma LAI_{OCK}\ (h) = 0.05 + 1.1 * sph_{(130m)}$$

$$\gamma CI\,cover_{OCK}\ (h) = 0.05 + 0.95 * sph_{(130\,m)}$$

$$\gamma CI\,cover-LAI_{OCK}\ (h) = 0.05 - 0.8 * sph_{(130\,m)}$$
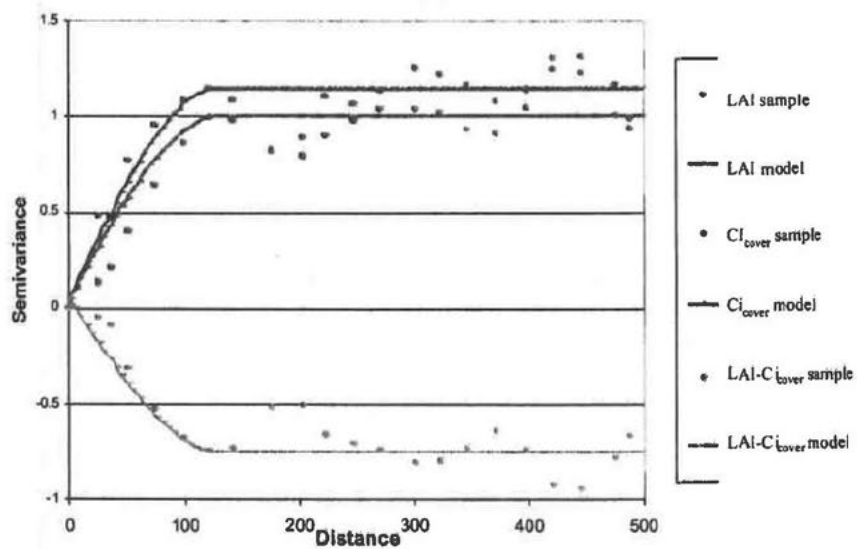
Figure 9. Experimental LAI and CIcover semivariograms, their cross variogram, and respective models.

Integrated Models

Cross-correlograms of the residuals of each regression method and CI_{cover} were calculated. The integrated model consisted of the RMA residuals cokriged with CIcover, and then added back to the RMA LAI regression predictions. The linear model of corregionalization for the RMA residuals and CI_{cover} (Fig. 10) was given by:

$$\gamma_{RMAresiduds_{OCK}}(h) = 0.65 + 0.1 * sph_{(126m)} + 0.35 * sph_{(280m)}$$

$$\gamma_{CI\,cov\,er_{OCK}}(h) = 0.01 + 0.92 * sph_{(126m)} + 0.09 * sph_{(280m)}$$

$$\gamma_{RMAresiduds-CI\,cov\,er_{OCK}}(h) = 0.01 + 0.18 * sph_{(126m)} + 0.175 * sph_{(280m)}$$
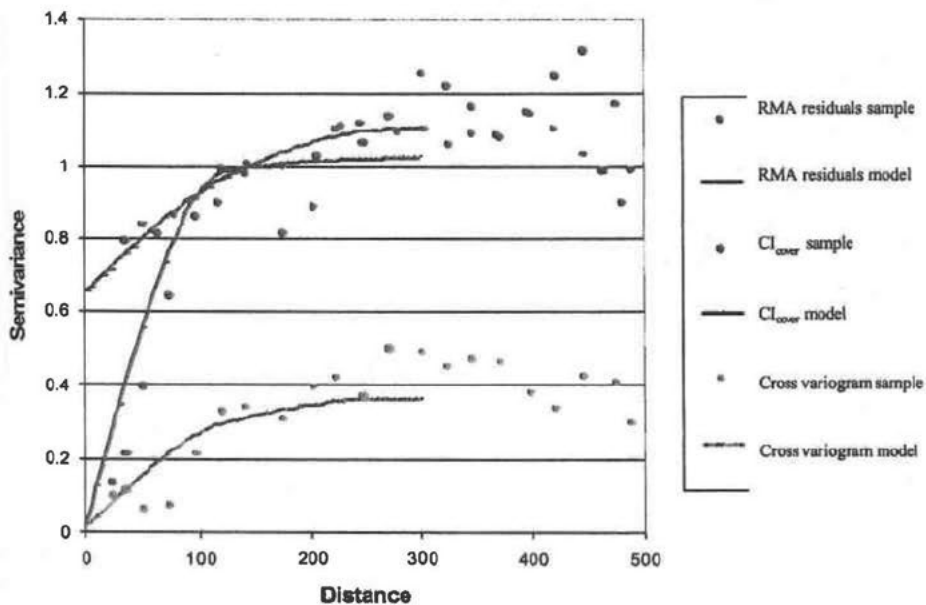
Figure 10. Experimental RMA residuals and CI_{cover} semivariograms, their cross variogram, and respective models.

72

Conditional Simulation

Sequential Gaussian conditional simulation (SGCS) with simple collocated cokriging with one secondary variable was performed using CI_{cover} as the secondary variable. The same anisotropic model utilized in KED was provided to the algorithm to model LAI, and 51 realizations were produced. A variance reduction factor of 0.85 was applied to ensure that the LAI normal score variance was close to one. The 51 realizations were subsequently processed.

Post-processing of the Realizations

The set of 51 realizations provided a measure of uncertainty for the spatial distribution of the primary variable. A visual assessment of uncertainty could be done by inspecting several probability and spread maps displays.

Probability maps were derived using local uncertainty, which can be described by the probability of exceeding a fixed threshold z:

$$F(u; z/(n)) = \Pr ob\{Z(u) \le z/(n)\}$$

where:

$F(u; z/(n))$ is the conditional distribution defined over location u.

Maps with the probability of LAI being greater than 2, 4 and 6 m^2/m^2 were obtained. Areas with probability values close to 0.5 are highly uncertain.

The conditional variance $\sigma^2(u_i)$ measures the spread of the conditional probability distribution around its mean $z_E^*(u_i)$.

$$\sigma^2(u_i) = \sum \frac{(u*u)}{n} - (\sum \frac{u}{n})^2$$

A map of the conditional variance of the 51 realizations was generated.

The uncertainty of a probabilistic model was defined by Deutsch, (2000) as the average conditional variance of all locations in the area of interest:

$$U = \frac{1}{N} \sum_{i=1}^{N} \sigma^2(u_i)$$

Uncertainty increases as the spread of the probability distribution increases. The average conditional variance (uncertainty) value lies between 0 and 1 in normal space.

Uncertainty measures were calculated for different conditional simulation runs, which differed in the model of the primary variable used, and in the covariate variable.

The optimal estimate of the unknown can be given by the expected value of the conditional cumulative distribution function at location u_i. The E-type estimate is the mean of the simulated LAI values at location u_i. Its computation is:

$$z_E^*(u_i) = \frac{1}{L} \sum_{l=1}^{L} z^{(l)}(u_i)$$

where:

$z_E^*(u_i)$ is the expected value of the unknown LAI value at location u

$l = 1,...., L$ are the simulations, and $z^{(l)}$ are the simulated values.

Cross Validation

The validation of the different techniques applied posed two problems. The first one derived from the exactness property of the geostatistical methods. The sample data are honored at their locations, so summary statistics of the estimations could not be used to directly compare the different methods. Second, the number of samples was limited, so methods such as jacknife (removal of a percent of the samples, and re-estimation from a non-overlapping data set) were not helpful. Cross validation is a procedure in which each sample value is removed one at a time from the data set, and that location is re-estimated from the remaining data (Deutsch and Journel, 1998;Goovaerts, 1997; Isaaks and Srivastava,1989; Wackernagel, 1998). This technique served two purposes in this study. First, it helped to identify the most appropriate geostatistical models in terms of type of variogram model, variogram parameters, and search strategy (size and geometry of the neighborhood). Second, it helped to compare the performances of the different methods used.

The cross validation root mean square error (RMSE) measures the average square difference between the sample $z(u_\alpha)$ and the predicted $z^*(u_\alpha)$ data (Wackernagel, 1998):

$$RMSE = \sqrt{\frac{1}{n}\sum_{\alpha=1}^{n}[z(u_\alpha)-z^*(u_\alpha)]^2}$$

The cross validation residuals were assessed with scatterplots of the observed versus estimated values, scatterplots of residuals versus estimated values to check

for unbiasedness (over or underestimation), and homoscedasticity (error variance independent of the magnitude of the estimated value). Histograms and semivariograms of cross validation errors were also made and inspected for their distribution and spatial independence, respectively (Deutsch and Journel, 1998). Cross validation can also be assessed qualitatively by analyzing the spatial arrangement of the residuals (Isaaks and Srivastava, 1989). The predictions and the errors were mapped to identify possible areas of consistent LAI over or underestimation.

The cross validation procedures were performed with different tools depending on the method: for traditional regression, cross validation was performed with the SAS REG procedure (SAS v.7); for inverse regression, with a program implemented in IDL (IDL); for RMA in Microsoft Excel 2000, and for the geostatistical methods with GSLIB (Deutsch and Journel, 1998 and Goovaerts, personal communication).

Results

LAI values ranged from 0.98 m^2/m^2 in a fen type location, where minimal willow

canopy was present, to 9.98 m^2/m^2, in an area where there was a mixture of conifer

and hardwood species (black spruce, balsam poplar and aspen) (Table 2a). The LAI

mean value was 4.19, relative to a median of 4.26, indicating a slight positive skew

in the data. The LAI histogram (Fig. 11) showed a bimodal distribution and

confirmed some degree of positive skew. The distribution of LAI values in space

revealed a slight N-S trend (Fig. 12).

Table 2 a. Univariate Statistics for ground measurements (n=86).

Variable	Min	Max	Mean	STD
LAI (m^2 m^{-2})	0.98	9.98	4.19	1.68
Cover (conifer + hardwood) (%)	0.00	63.89	38.55	16.10
Conifer (%)	0.00	63.77	36.75	15.48
Hardwood (%)	0.00	18.06	1.79	3.12

Table 2 b. Univariate Statistics for Landsat ETM+ bands (n=86). Numbers are in terms of percent reflectance.

Variable	Min	Max	Mean	STD
Band 1	2.58	3.54	3.19	0.22
Band 2	3.50	4.78	4.01	0.32
Band 3	3.19	5.17	3.85	0.40
Band 4	15.45	23.64	18.06	1.76
Band 5	10.45	18.36	13.65	2.02
Band 7	5.71	10.47	7.45	1.05
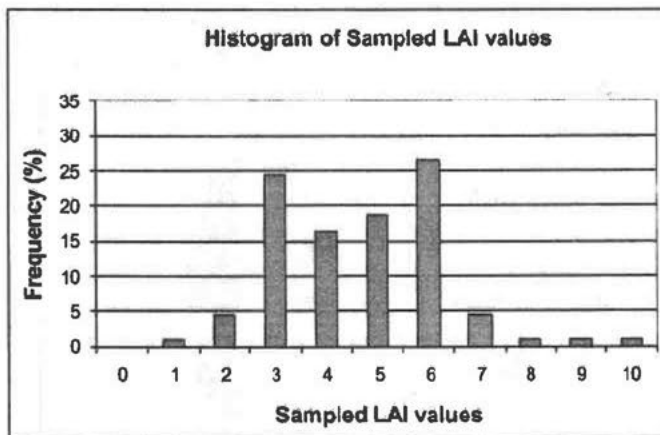
a.)
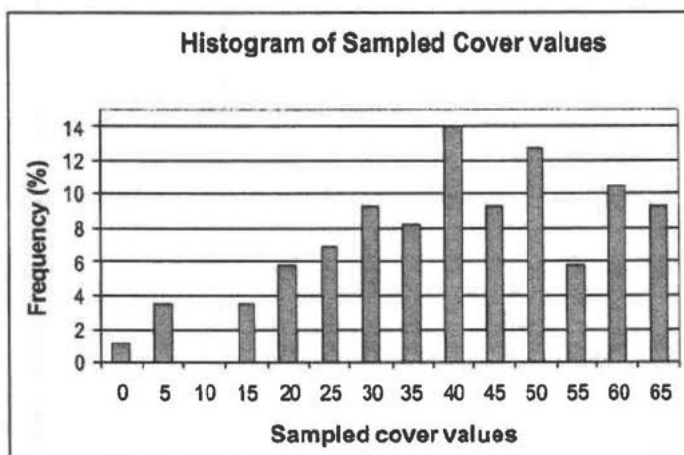
b.)

Figure 11. a.) Histogram of LAI sample data. b.) Histogram of total cover sample data. The numbers in the X axis correspond to the upper-limit of each class.
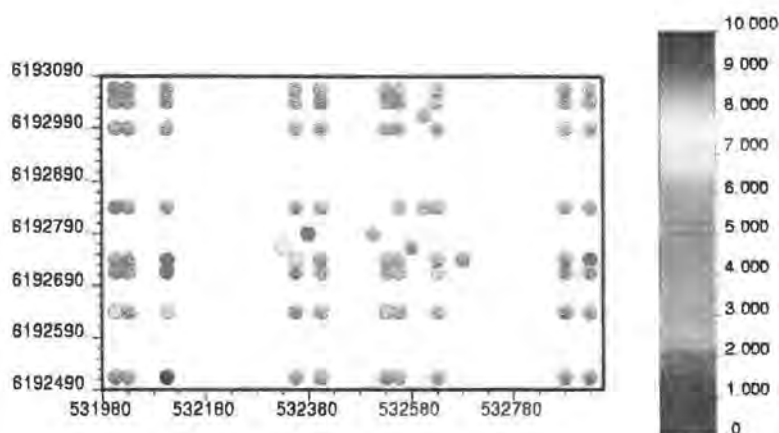
Figure 12. Location and magnitude of LAI for each plot.

Cover ranged from 0 to 64 % with a mean of 39 %. Conifer cover prevailed with a mean of 37 % compared to a hardwood cover mean of 1.79 % (Table 2 a). These low canopy closures explain the high influence of understory and soil cover in the reflectance signal. The spatial distribution of cover reveals, to a greater degree, the presence of a N-S trend (Fig. 13).

Scatterplots of LAI versus ETM+ reflective bands and cover can be seen in fig. 14, and univariate statistics in Table 2 b. Cover was divided into cover classes according to the land cover classification in fig. 5. The cover classes present in this area were wetlands, open shrubland, savanna, woody savanna, and evergreen needleleaf forest (see class definitions in Table 3).
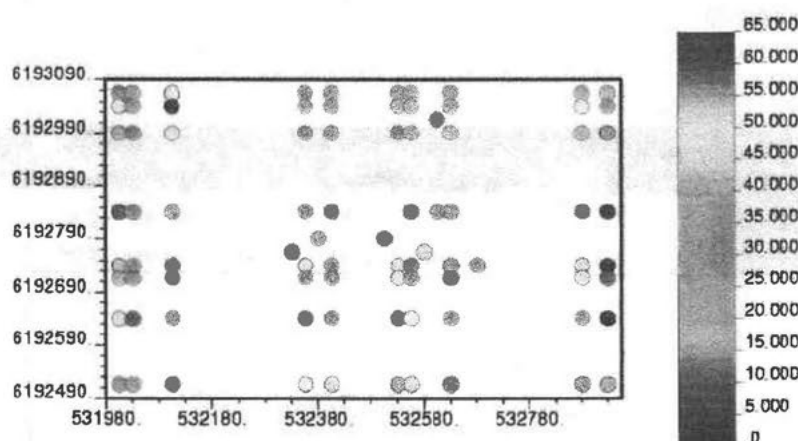
Figure 13. Location and magnitude of total cover (conifer + hardwood) for each plot.

Table 3. Description of cover classes.

Cover Class	Definitions
Wetland	Permanent water/herbaceous/woody <10% tree cover, presence of water
Open shrubland	<10% tree cover, 10-60% shrub
Savanna	10-30% tree cover
Woodland	30-60% tree cover
Evergreen needleleaf forest (ENF)	>60% ENF tree cover

Negative linear relationships were observed for all ETM+ reflective bands and LAI (Table 4), band 5 and 7 being the strongest (r^2 = -0.73 and -0.69, respectively). Cover classes such as savanna, woody savanna, and evergreen needle leaf forest presented a broad range of LAI values and were not separable by their signal

Table 4. Pearson's correlation coefficients for field and Landsat ETM+ data (n=86).

Variable	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7	CILAI	Clcover
LAI	-0.29	-0.60	-0.58	-0.63	-0.73	-0.69	-0.73	-0.70
Total cover	-0.33	-0.73	-0.70	-0.74	-0.78	-0.78	-0.77	-0.80

Stands with predominance of hardwood species (> 60% of deciduous broadleaf forest) were not sampled, as can be seen in the S-W corner of the image ETM+ image (Fig. 4) and in the land cover map (Fig. 5). Open shrubland received only one sample (Fig.14).

Table 5 shows the standardized canonical coefficients obtained using LAI and cover. Scatterplots between LAI and the canonical indices for LAI (CI_{LAI}) and cover (CI_{cover}), and with the spectral vegetation indices NDVI and SR are shown in fig. 15. The canonical indices improved the ability of the spectral data to account for the variability observed in LAI; R^2 changed from < 0.01 for LAI and NDVI to 0.54 for LAI and CI_{LAI}. These low coefficients of determination for NDVI and SR may be caused by high reflectance from the developed summer understory and the clumped structure of the canopy (Turner et al., 1999).

In closed canopies, SVIs such as NDVI and SR maximize the information content of the LAI traditional trends (negative correlation with red and positive with NIR).

When canopy closure is low, LAI and NIR show virtually no relationship (Chen
and Cihlar, 1996; Loechel et al., 1997; Turner et al., 1999) suggesting that more
complicated models may be needed in open canopy situations. In this study, bands
3 and 4 both showed a negative relationship with LAI, limiting the utility of NDVI
and SR. This negative relationship for band 4 becomes flatter if the three wetland
points are eliminated from the data set (Fig. 14 f).

Table 5. Canonical Correlation Coefficients (n=86).

ETM+BAND	Mean *	STD	SCC LAI**	SCC Cover**
1	3.19	0.22	-0.05	-0.07
2	4.01	0.32	-0.08	0.17
3	3.84	0.40	-0.12	0.06
4	18.06	1.76	-0.05	0.25
5	13.65	2.02	1.15	0.20
7	7.45	1.05	0.09	0.41

* Percent reflectance ,
** Standardized canonical coefficients.

82

Figure 14 a.) Scatterplot of cover versus LAI, b.) scatterplot of cover class versus
LAI, c.) band 1 versus LAI, d.) band 2 versus LAI, e.) band 3 versus LAI, f.) band
4 versus LAI, g.) of band 5 versus LAI, and h.) band 7 versus LAI.

14a.)

14b.)

14c.)

84
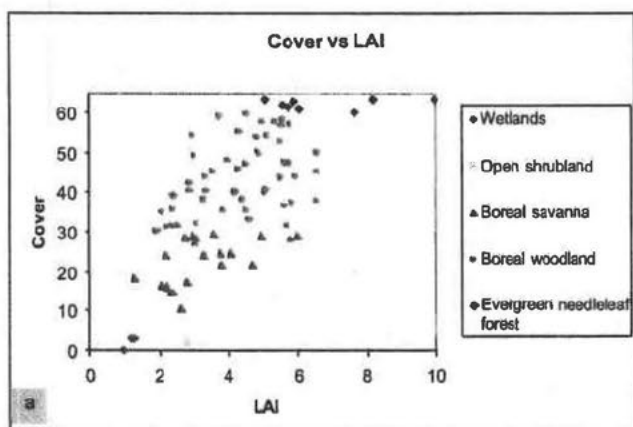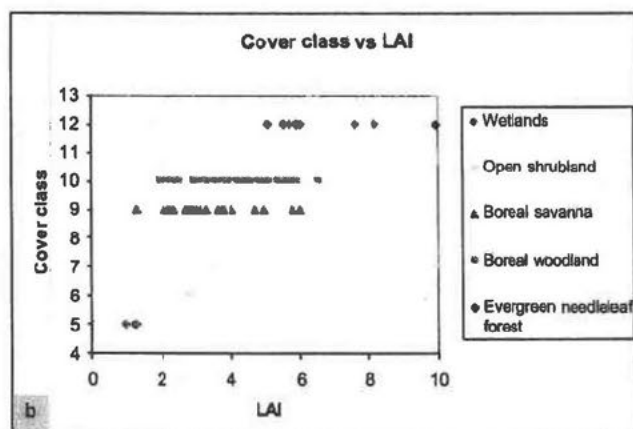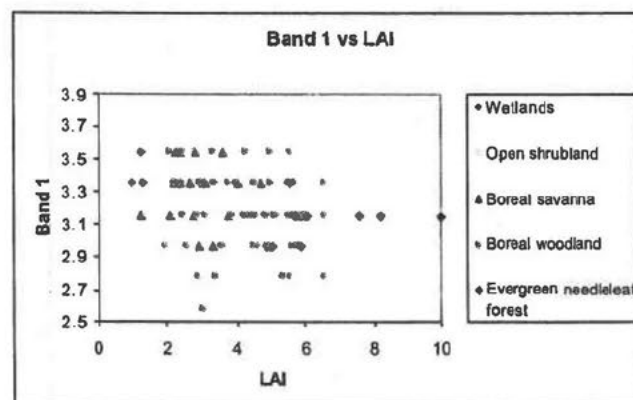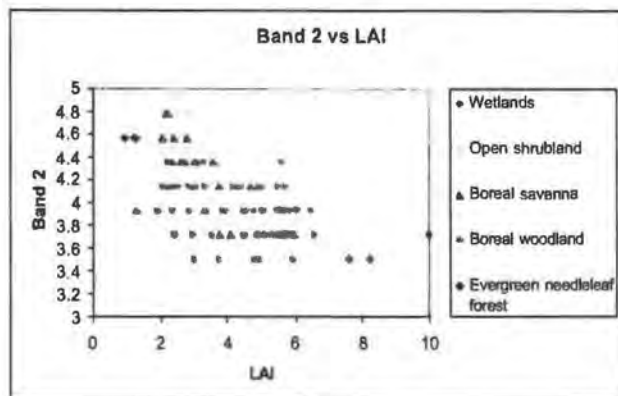
14d.)
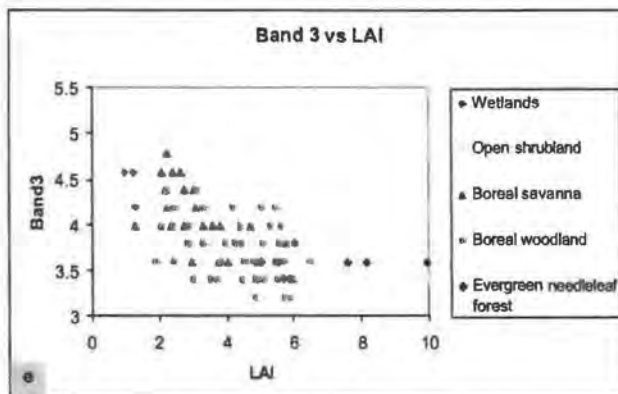
14e.)
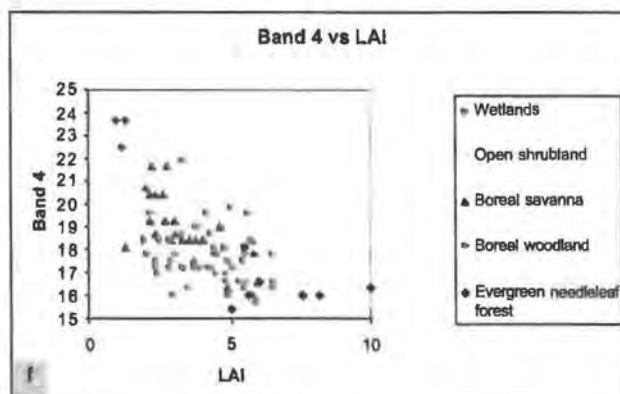
14f.)
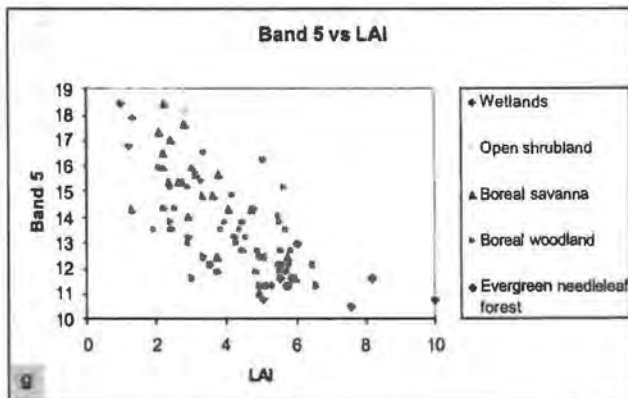
14g.)
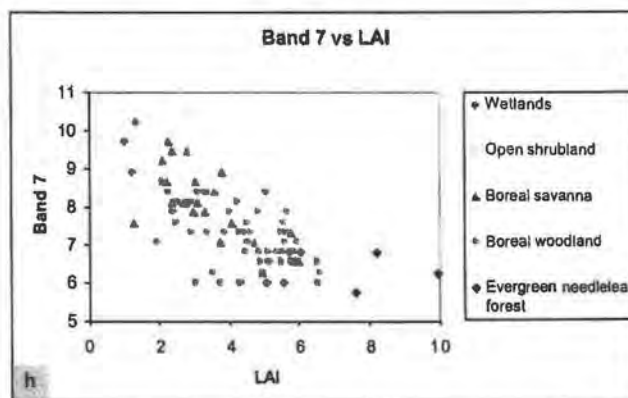
14h.)

Figure 15 a.) Scatterplot of NDVI versus LAI, b.) scatterplot of SR versus LAI, c.) scatterplot of CI_{LAI} versus LAI, and d.) scatterplot of CI_{cover} versus LAI.

NDVI vs LAI

$R^2=0.0002$

15a.)

SR vs LAI

$R^2=0.0005$
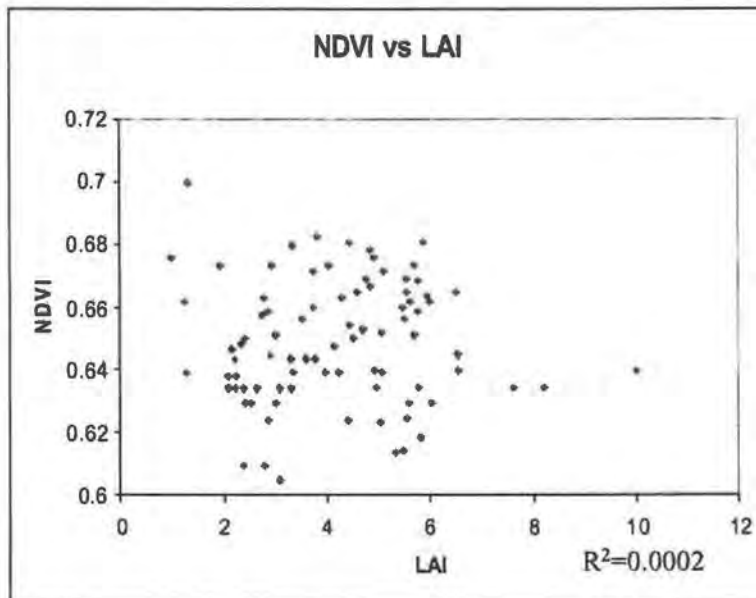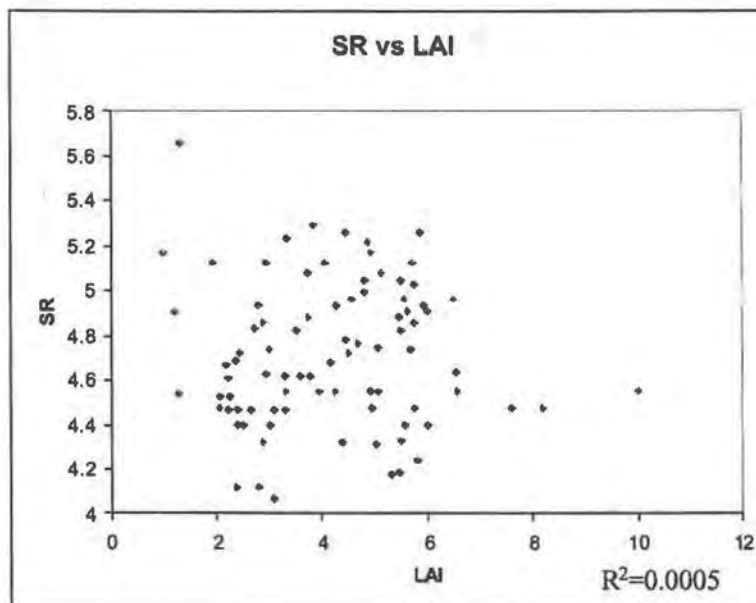
15b.)
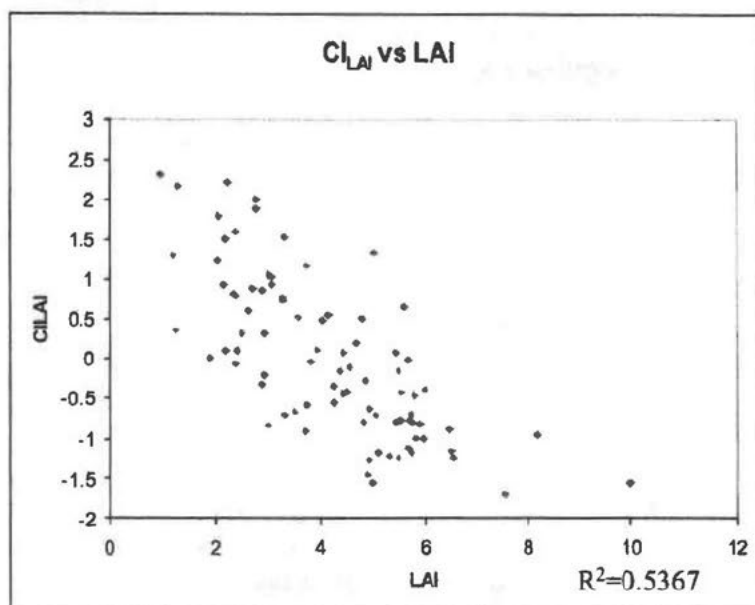
15c.)
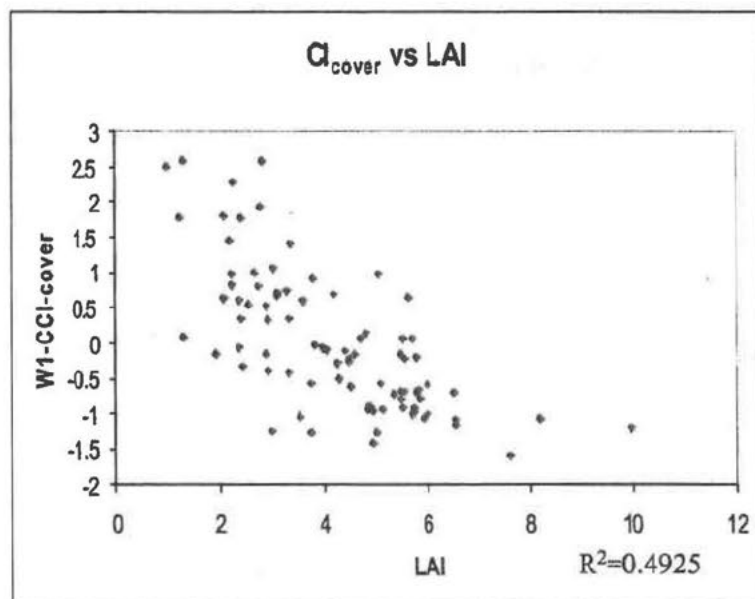
15d.)

Regression

Regression equations, predictions, and residuals for the sample locations can be seen in Tables 6, 7, and 8, respectively. Plotted as predicted versus observed, Reg_T resulted in the least pronounced slope, with overestimation for lower LAI values and underestimation for high LAI values (Figs.16 and 17). Reg_I had the steepest slope, while the RMA slope was intermediate (Fig. 16). The mean was well preserved by all methods, the actual LAI range was reduced by all techniques, and the standard deviations were too low for Reg_T and excessively high for Reg_I (Table 7).

Table 6. Regression models and coefficients of determination.

Method	Regression Models	R^2
RegT	$LAI = 4.1922 - 1.2286 \cdot CI_{LAI}$	0.537
RegI	$CI_{LAI} = 1.8314 - 0.4369 \cdot LAI$	0.537
RMA	$LAI = 4.1922 - 1.6769 \cdot CI_{LAI}$	0.465

Table 7. Summary statistics of predictions for the sample locations.

Method	Min	Max	Mean	Median	STD	Coef. Var.
Field LAI	0.98	9.98	4.19	4.26	1.68	0.40
RegT	1.36	6.27	4.19	4.36	1.23	0.29
RegI	-1.09	8.07	4.19	4.51	2.28	0.54
RMA	0.32	7.03	4.19	4.43	1.68	0.40

Table 8. Summary statistics of prediction residuals for the sample locations.

Method	Min	Max	Mean	Median	STD	R^2 (residuals vs pred.)
RegT	-2.48	3.88	0.00	0.03	1.14	0.00
RegI	-3.12	3.92	0.00	-0.1	1.56	0.46
RMA	-2.61	3.18	0.00	0.09	1.23	0.13
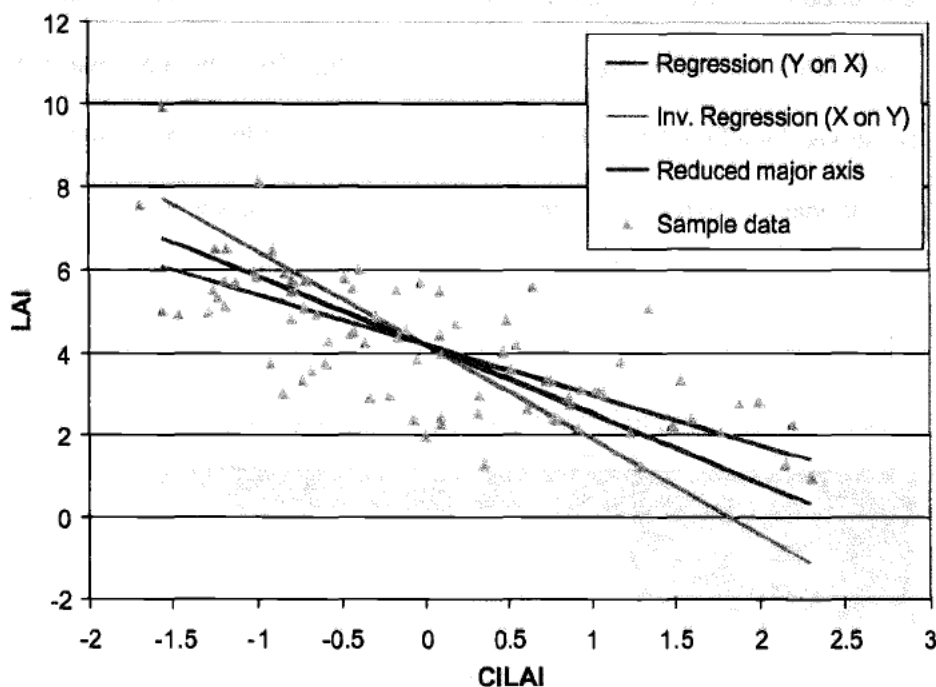
Figure 16. Traditional regression, inverse regression, and reduced major axis models for LAI.

Reg$_T$ and Reg$_I$ residuals showed a homoscedastic distribution, with no apparent trend; however, RMA residuals showed a slight trend, with overestimations for low LAI values and underestimation for high LAI values (Fig. 18). There was no bias exhibited for any model (ie. mean = 0.00 in Table 8). However, the R^2 between predictions and residuals and the range of residuals were highest for Reg$_I$.

Figure 17 a.) Scatterplot of predicted versus observed LAI values for traditional regression, b.) scatterplot of predicted versus observed LAI values for inverse regression, and c.) scatterplot of predicted versus observed LAI values for reduced major axis.

a) Traditional Regression
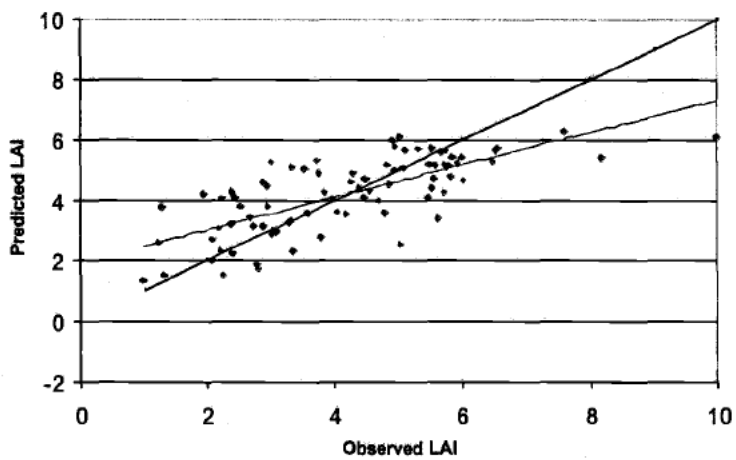
17a.)

b) Inverse Regression

17b.)

c) Reduced Major Axis

17c.)

94
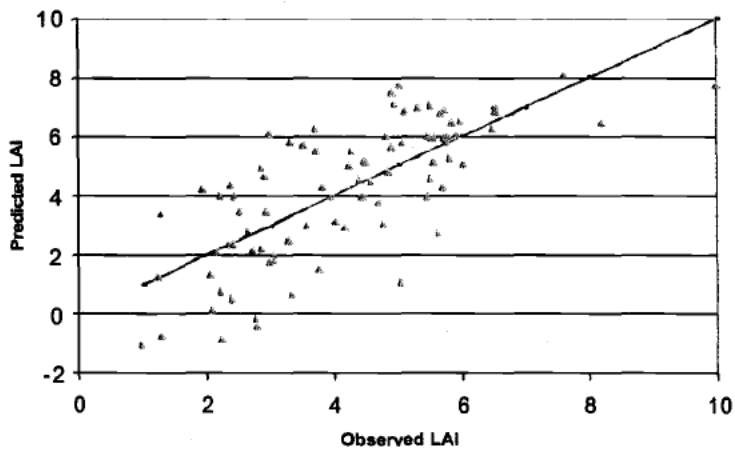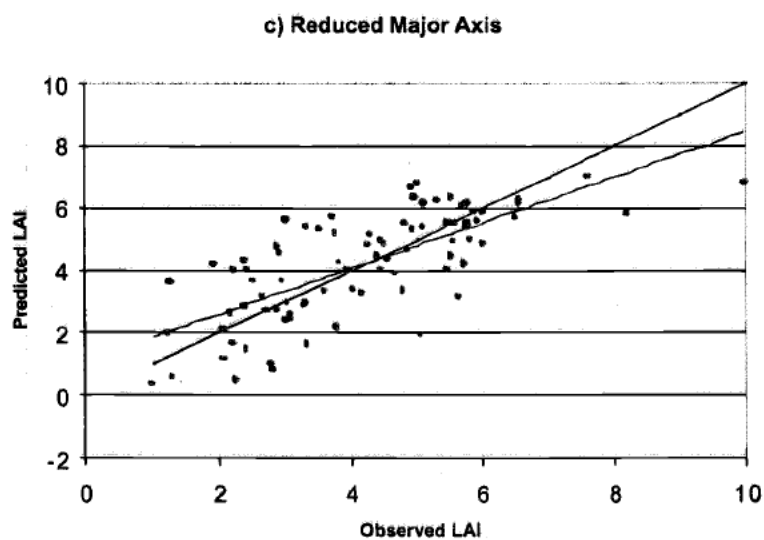
Figure 18 a.) Scatterplot of residuals versus predicted LAI values for traditional regression, b.) scatterplot of residuals versus predicted LAI values for inverse regression, and c.) scatterplot of residuals versus predicted LAI values for reduced major axis.

a) Traditional Regression
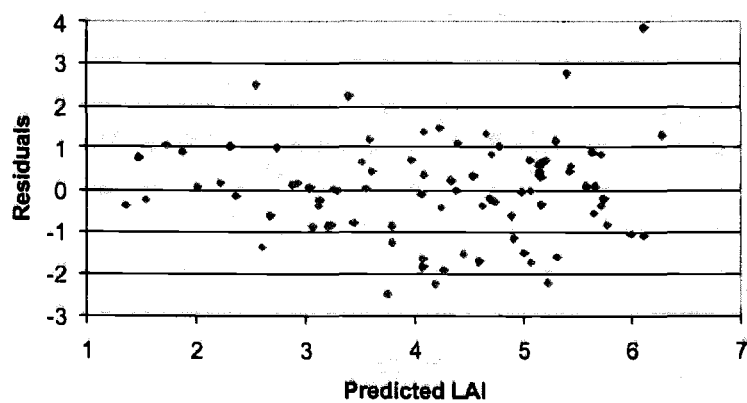
18a.)

b) Inverse Regression

18b.)

c) Reduced Major Axis
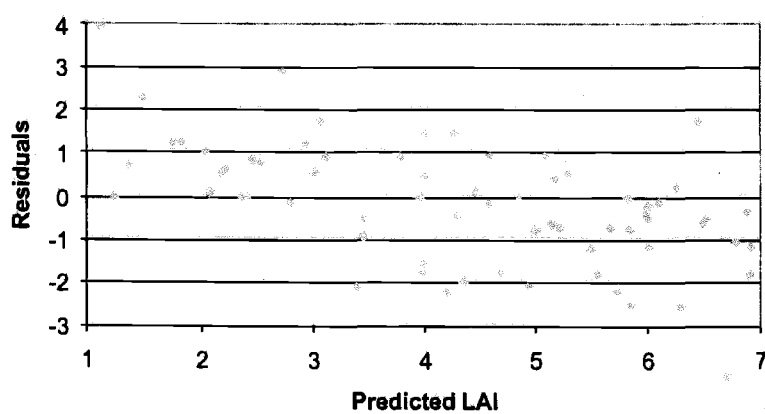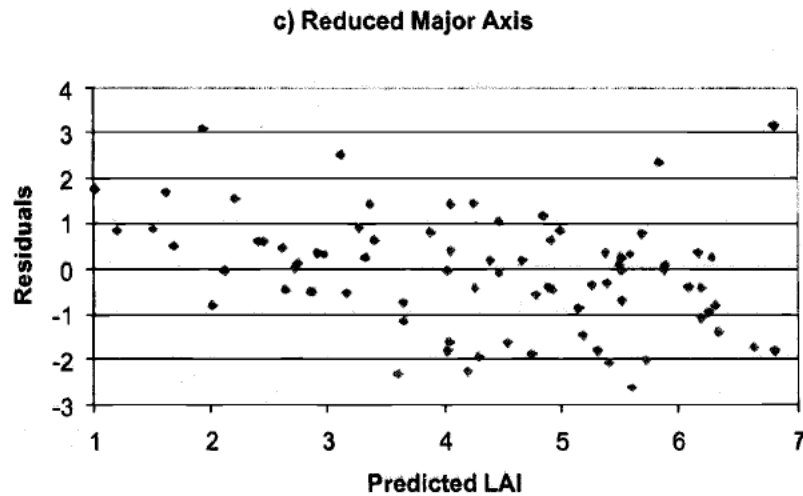
18c.)

For each method, the spatial autocorrelation of the residuals was inspected using Moran's I and Geary's C indices and semivariograms. Such measures indicated low spatial autocorrelation, especially for Reg$_r$ residuals, which were not significantly autocorrelated at a significance level of 0.05 (Table 9). The high nugget effects obtained by the three methods (Fig. 19) resulted from an initial low spatial autocorrelation of LAI (Table 10), and because the models did a good job accounting for the variation of their respective response variables.

Table 9. Moran's I and Geary's C autocorrelation coefficients for residuals of regression methods.

Variable	Moran's I	p-value	Geary's C	p-value
RegT residuals	0.003	0.58	0.99	0.76
RegI residuals	0.077	0.00	0.91	0.02
RMA residuals	0.028	0.04	0.96	0.14

Table 10 Moran's I and Geary's C coefficients for LAI and cover data.

Variable	Moran's I	p-value	Geary's C	p-value
LAI	0.068	0.00	0.93	0.00
Cover	0.080	0.00	0.91	0.00
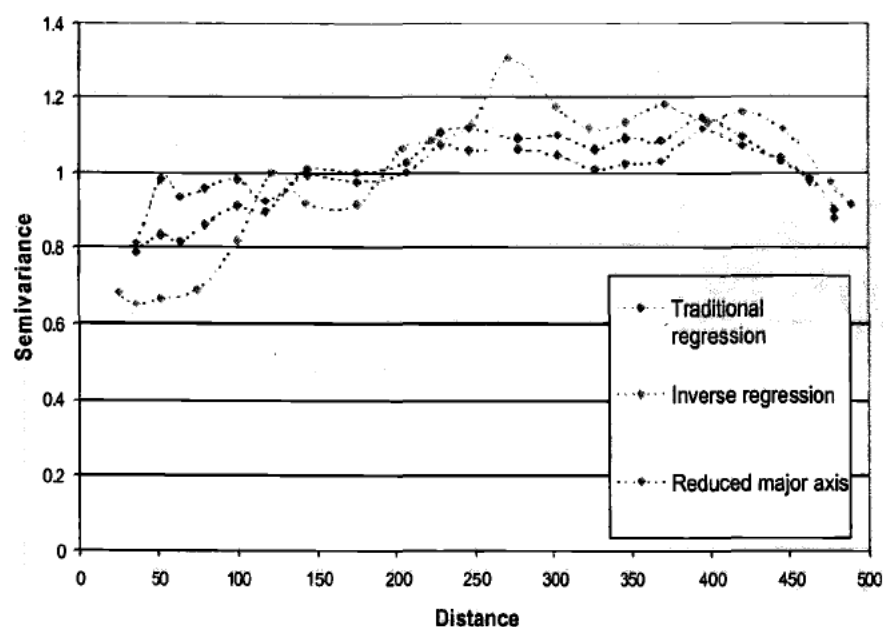
Figure 19. Semivariograms of residuals of the traditional regression, inverse regression and reduced major axis models.

Cross validation results revealed that Reg_T had the lowest RMSE, Reg_I the highest and RMA was again intermediate (Table 11). Compared to the field data, all methods reproduced closely the mean (Table 12), although the median showed a slight positive bias. RegT exhibited a high degree of attenuation as shown in the range of predictions in Table 12. Although the range of predictions was also attenuated for RMA at the high end, it was close to the observed low values and it was the only model to nearly match the variability of observed values. This can also be observed in the histograms of cross-validation predictions (Fig.20). Graphs of predicted versus observed values showed the attenuation effect of the Reg_T model, with underpredictions at the high end and overpredictions at the low end when compared to the sample LAI range (Fig. 21 a, b, and c).

Table 11. Cross validation RMSE and coefficient of determination of predicted versus observed LAI values for regression methods.

Method	RMSE	R^2
RegT	1.16	0.515
RegI	1.59	0.526
RMA	1.23	0.521

Table 12. Summary statistics of cross validation predictions for regression methods.

Method	Min	Max	Mean	Median	Std
Field LAI	0.98	9.98	4.19	4.26	1.68
RegT	1.39	6.21	4.19	4.36	1.23
RegI	-1.27	8.10	4.18	4.51	2.32
RMA	0.37	6.94	4.19	4.42	1.65

99
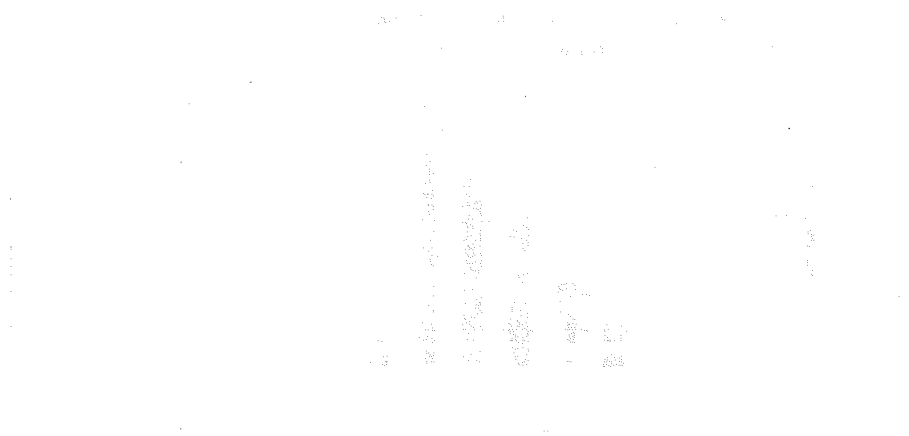
Figure 20 a.) Histogram of cross validation predictions for traditional regression, b.) histogram of cross validation predictions for inverse regression, c.) histogram of cross validation predictions for reduced major axis.
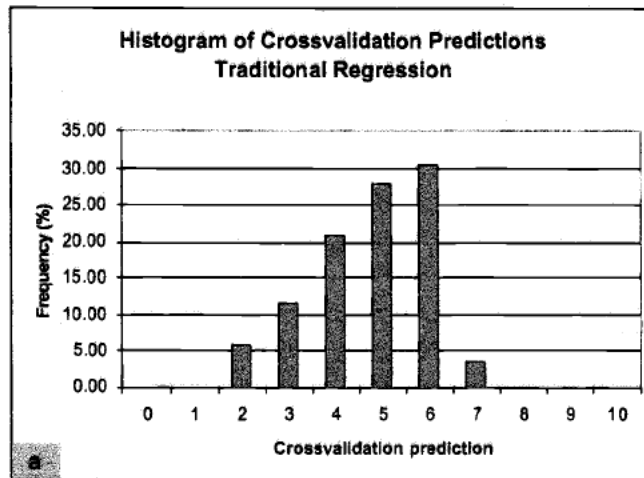
20a.)
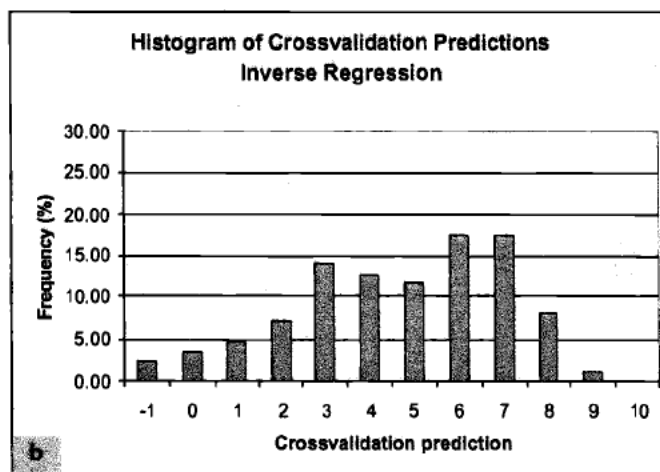
20b.)
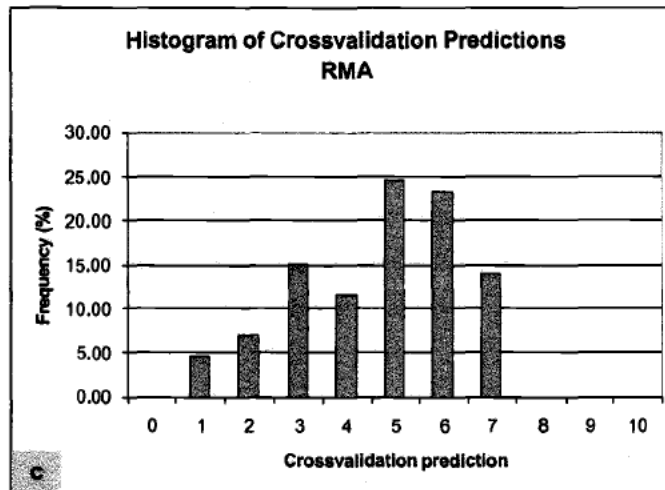
20c.)

Scatterplots of the cross validation residuals vs. estimated LAI showed no trend for

Reg_T, high trend for Reg and a slight trend for RMA (Fig. 21). Table 13 showed a

slight positive skew for Reg and RMA, confirmed by the cross validation error

histograms (Fig. 22). This bias may result from the narrower estimated LAI ranges.

Moran's I coefficients for the residuals of cross validation indicated that only the

Reg residuals had significant spatial autocorrelation (Table 13). This was

confirmed by the Reg_T and RMA semivariograms in fig. 22, which showed a pure

nugget effect.

Among the regression methods tested, RMA was judged to be the best model. This

is because RMA accounted for the errors in both variables, and preserved the mean

and the variance of the observed values in both, the estimation and cross validation

procedures.

Figure 21 a.) Scatterplot of cross validation estimated LAI versus observed LAI for traditional regression, b.) scatterplot of cross validation estimated LAI versus observed LAI for inverse regression, c.) scatterplot of cross validation estimated LAI versus observed LAI for reduced major axis, d.) scatterplot of cross validation errors versus observed LAI for traditional regression, e.) scatterplot of cross validation errors versus observed LAI for inverse regression, f.) scatterplot of cross validation errors versus observed LAI for reduced major axis.
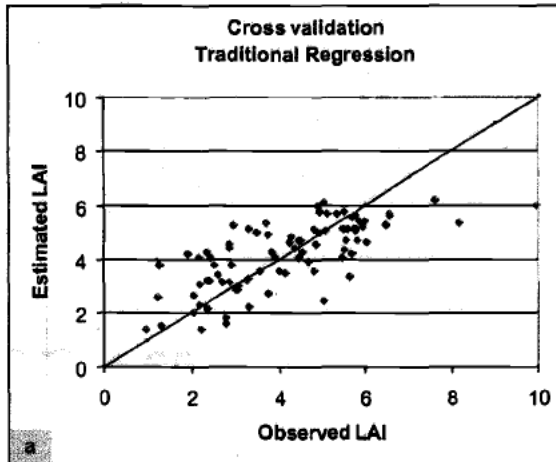
21a.)
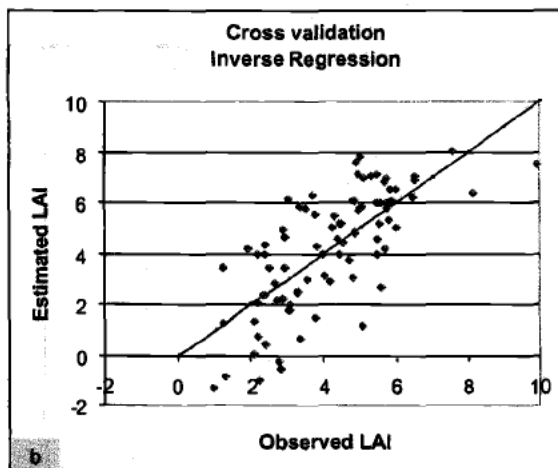
21b.)
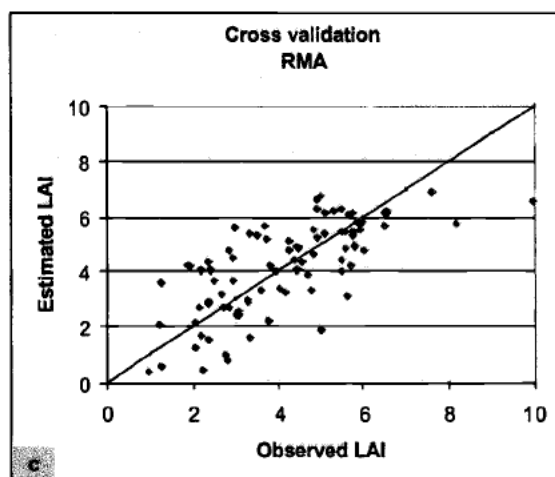
21c.)
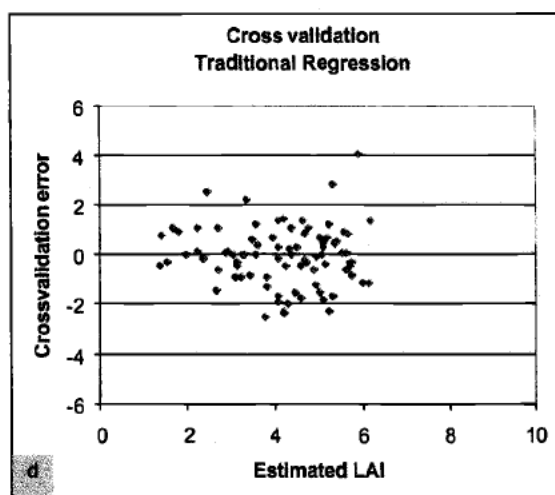
21d.)
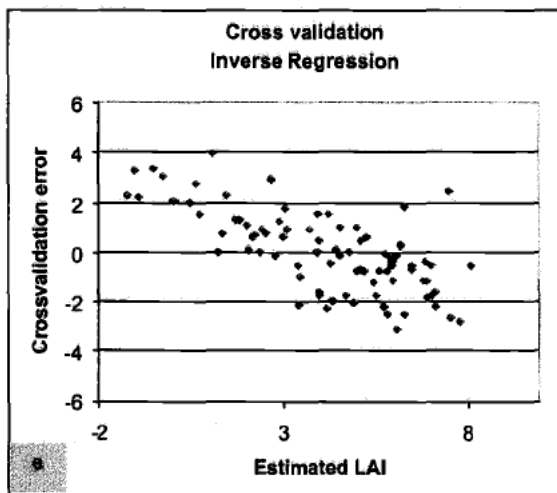
21e.)
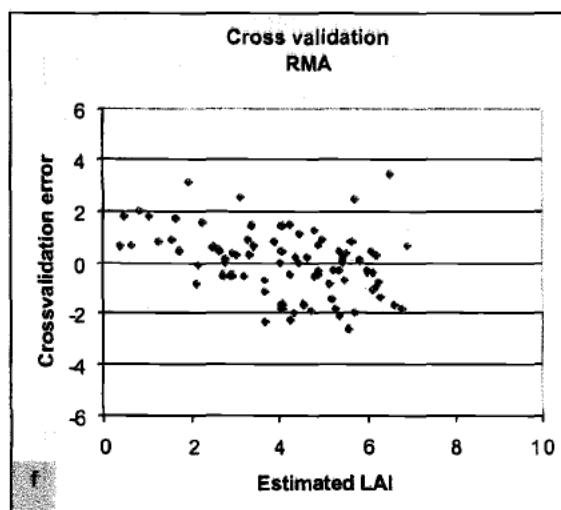
21f.)

Table 13. Summary statistics of cross validation errors for regression methods.

Method	Min	Max	Mean	Median	Std	Moran'sI	p-value
RegT	-2.51	4.04	0.00	0.03	1.17	0.02	0.34
RegI	-3.13	3.92	0.01	-0.10	1.60	0.08	0.02
RMA	-2.60	3.43	0.01	0.10	1.24	0.02	0.14

Figure 22 a.) Histogram of cross validation errors for traditional regression, b.) histogram of cross validation errors for inverse regression, c.) histogram of cross validation errors for reduced major axis, d.) experimental semivariogram of cross validation errors for traditional regression, e.) experimental semivariogram of cross validation errors for inverse regression, f.) experimental semivariogram of cross validation errors for reduced major axis.

22a.)

22b.)

22c.)

22d.)

109
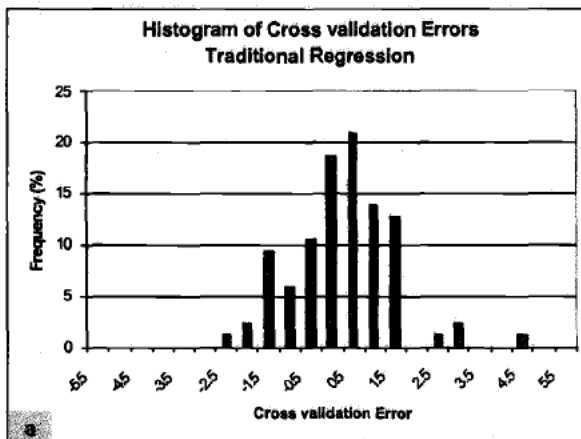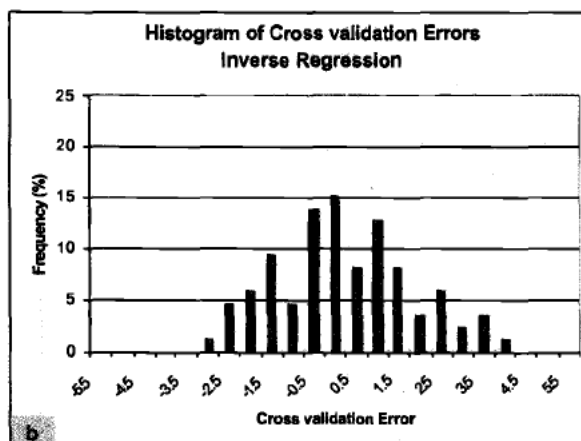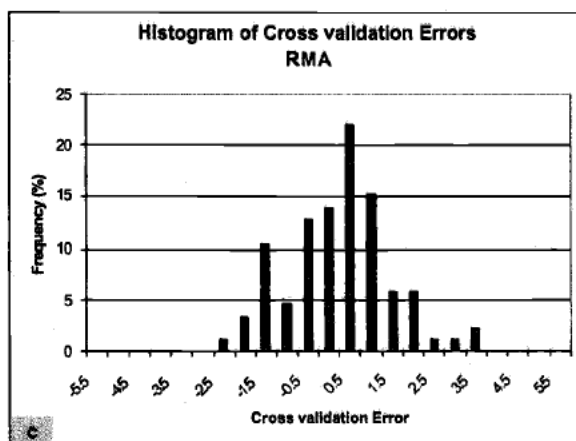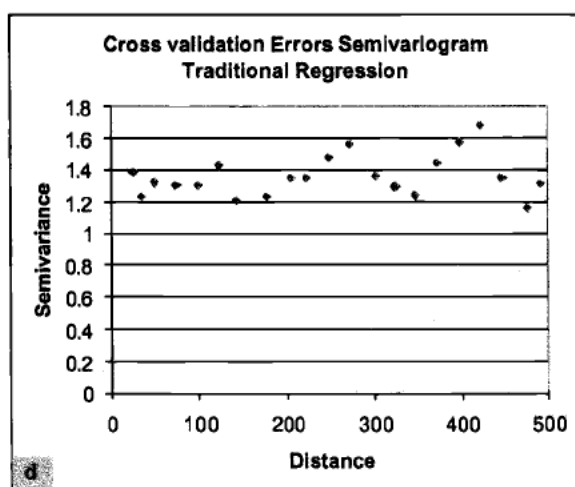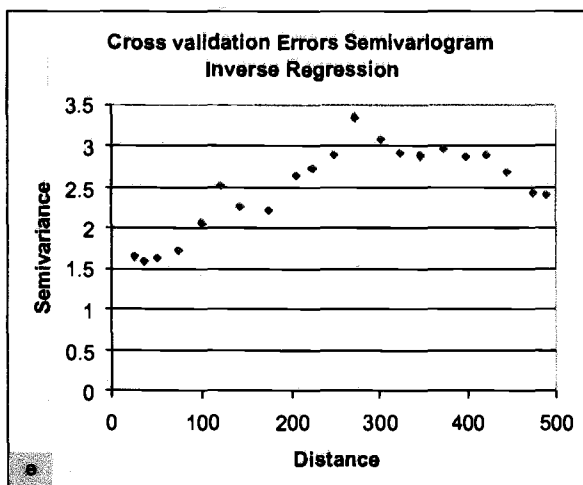
22e.)
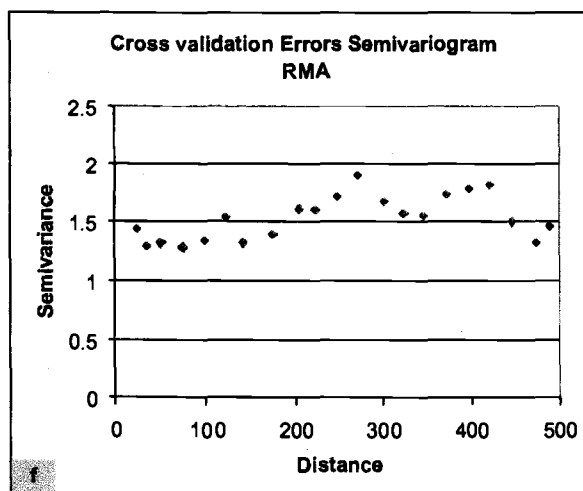
22f.)

Geostatistics: Estimation Methods

Moran's I and Geary's C coefficients were calculated for LAI and cover, showing that both variables had low but significant spatial autocorrelation (Table 10). LAI and cover omnidirectional semivariograms looked very similar (Fig. 23), having the same range of spatial autocorrelation of 122 meters, and differing slightly in their normalized sills (1.15 for LAI and 1.13 for cover).
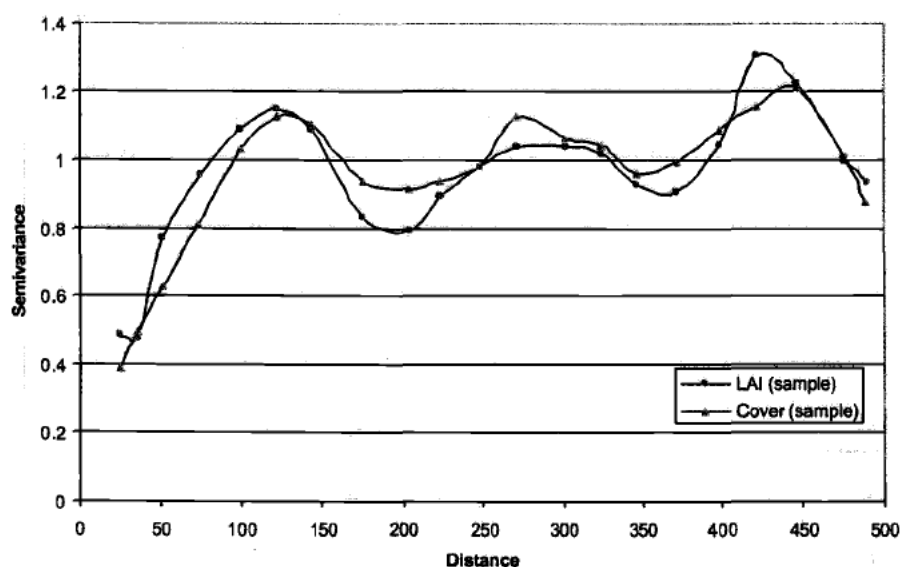
Figure 23. Experimental semivariograms for the normal scores of LAI and cover.

LAI semivariograms calculated at lags of 25, 50, 75 and 100 meters showed periodicity, no matter what the scale of the calculation (Fig. 24).

Directional LAI semivariograms indicated that the direction of major continuity was 160 degrees, as also can be seen in the rose diagram (Fig. 25 c). Features with

elongated shapes in this direction can also be observed in the original TM image

(Figs. 2 and 4).
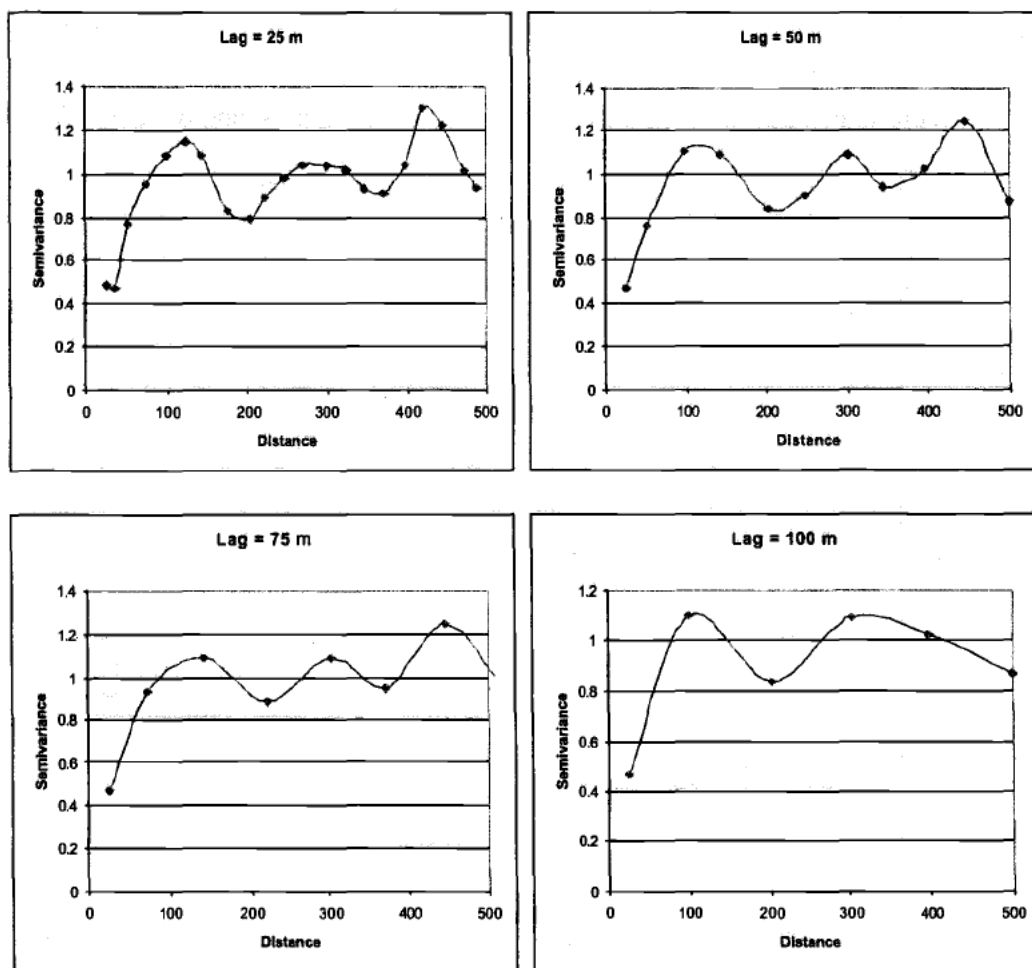
Figure 24. Experimental semivariograms of the normal scores of LAI at lags of 25, 50, 75, and 100 m.
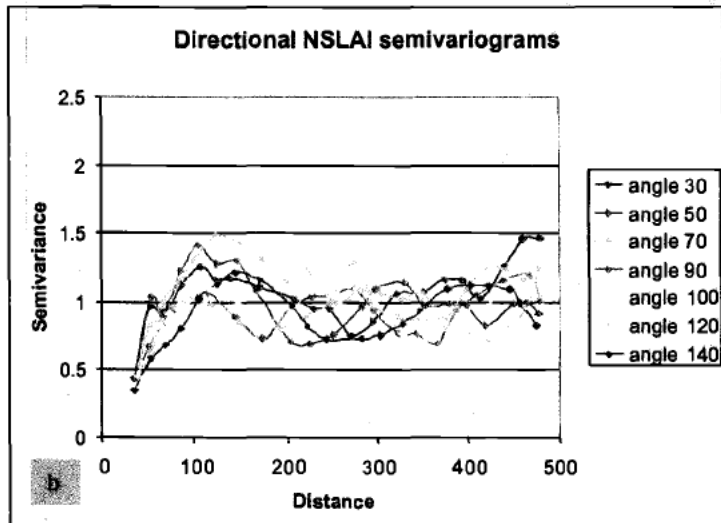
Cross correlograms of LAI with CI_{LAI} and CI_{cover} showed a similar behavior, with a

negative correlation decreasing rapidly up to 100 m, and then slowly up to 325 m

where the correlation becomes positive (Fig. 26). Because of their similar spatial

behavior associated to LAI, they could both be used as secondary variables for the

geostatistics bivariate procedures. LAI and cover were positively cross-correlated

up to 300 m but only highly cross-correlated up to 100 m (Fig. 26).

Figure 25. Experimental directional semivariograms of the normal scores of LAI grouped in a.) directions of maximum continuity and b.) directions of minimum continuity c.) Rose diagram. Each segment represents the same semivariance magnitude, allowing the visualization of the principal anisotropy axes.

a.)

b.)

(Continued)

0°   10°
30°
50°
70°
90°
100°
120°
140°
c
160°
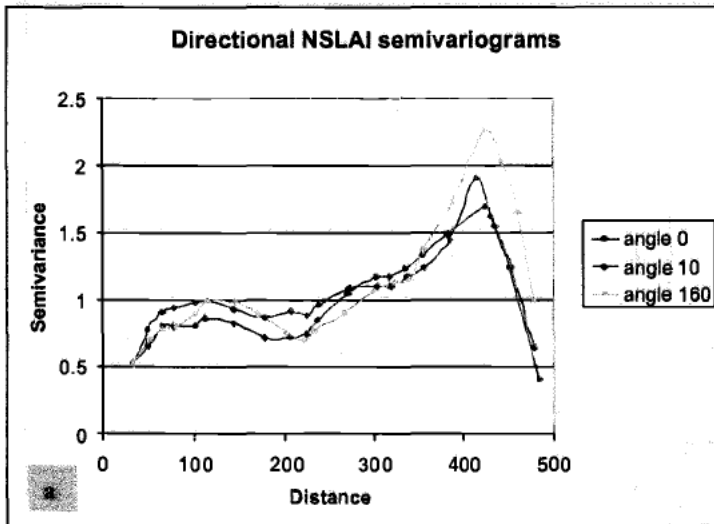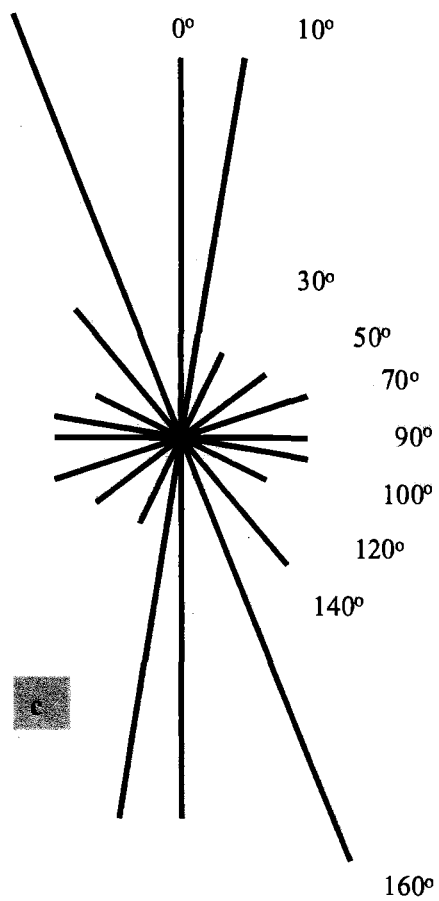
c.)
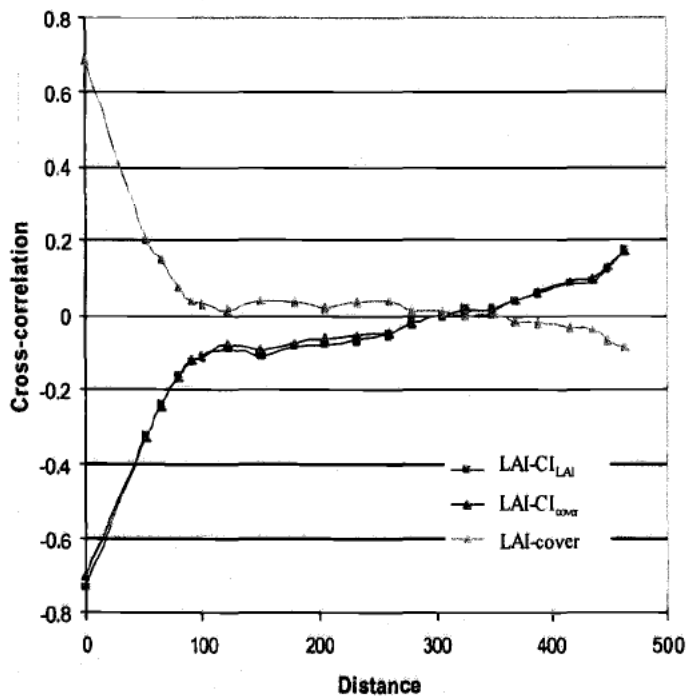
Figure 26. Cross correlograms of the nscores of LAI with the nscores of CI_{LAI}, CI_{cover} and cover.

Cross correlograms of the residuals of each regression method with CI_{LAI} and CI_{cover} were very similar (Fig. 27). All of them showed a positive correlation up to approximately 250 m, but after this distance the correlation was negative. The Regr residuals showed the lowest cross correlation of all, with certain particular features: a 0 correlation at lag 0, and a possible periodicity from positive to negative correlation changing approximately each 200 m. This 0 correlation at lag 0 suggested that the residuals of the regression model were independent at this lag, but their spatial autocorrelation increased with increasing lags, although not by much. Reg and RMA residuals were not spatially dependent (Table 9).

117

Figure 27 a.) Cross correlogram of the normal scores of the residuals from traditional regression with the normal scores of CI_{LAI} and CI_{cover}. b.) cross correlogram of the normal scores of the residuals from inverse regression with the normal scores of CI_{LAI} and CI_{cover}. c.) cross correlogram of the normal scores of the residuals from reduced major axis with the normal scores of CI_{LAI} and CI_{cover}.
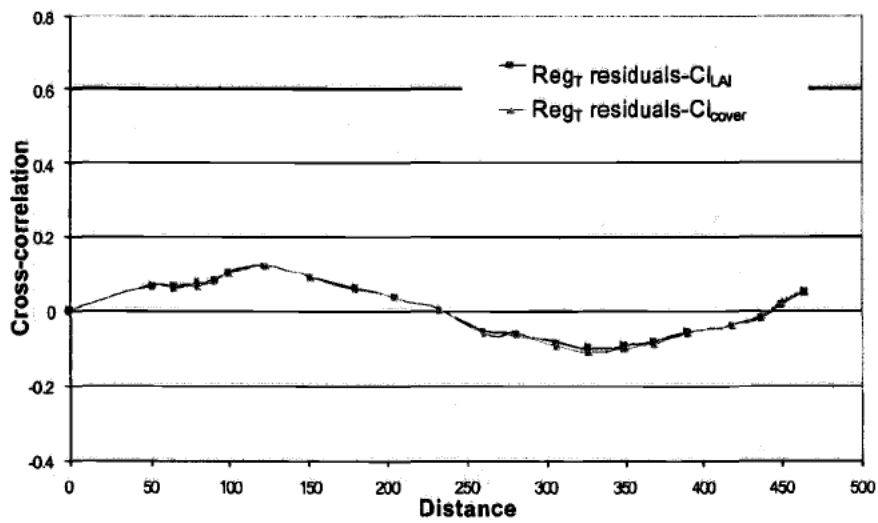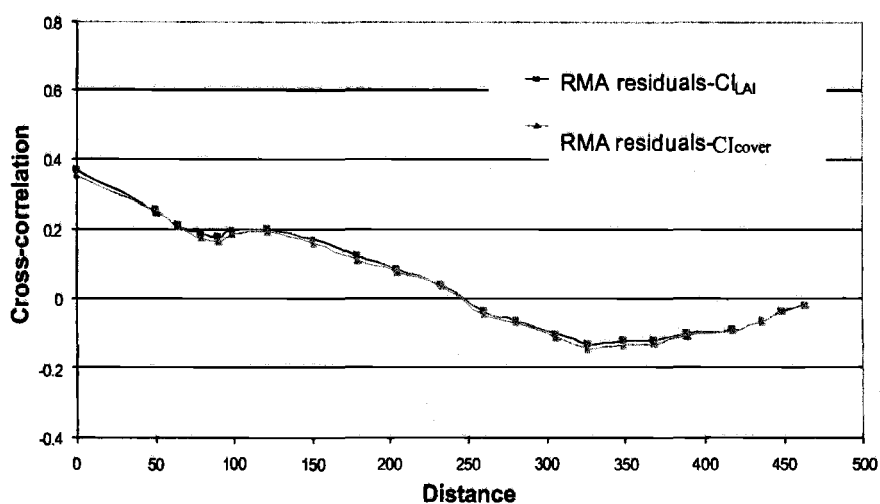
a.)
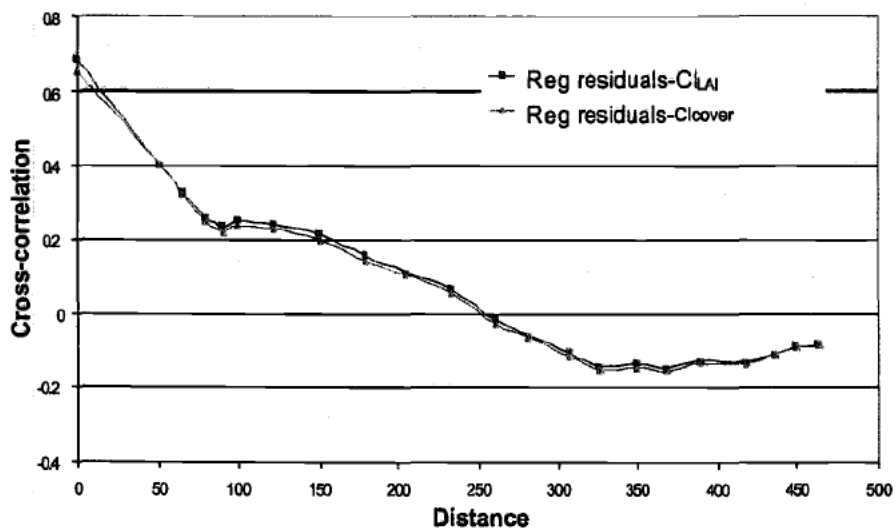
b.)

(Continued)

c.)

Cross validation RMSE of geostatistical methods indicated that the integrated

model performed best (Table 14). Geostatistical methods closely reproduced the

mean, showed some degree of positive skew, and OK and OCK had the lowest

standard deviations because of their smoothing effect (Table 15).

Table 14. Cross validation RMSE and coefficient of determination of predicted versus observed LAI values

Method	RMSE	R^2
Isotropic OK	1.68	0.159
Anisotropic OK	1.58	0.201
OCK CI_{cover}	1.51	0.234
Anisotropic KED CI_{cover}	1.44	0.397
Integrated Model CI_{cover}	1.23	0.495

Table 15. Summary statistics of cross validation predictions for geostatistical methods.

Method	Min	Max	Mean	Median	Std
Field LAI	0.98	9.98	4.19	4.26	1.68
Isotropic OK	1.16	7.93	4.14	4.25	1.36
Anisotropic OK	1.27	7.75	4.19	4.45	1.27
OCK CI_{cover}	2.09	6.87	4.22	4.26	1.19
Anisotropic KED CI_{cover}	0.33	9.98	4.15	4.36	1.69
Integrated Model CI_{cover}	0.62	6.66	4.17	4.20	1.51

The KED histogram of cross validation predictions was the closest to the sample LAI data (Fig. 28), as well as its standard deviation (Table 15). Scatterplots of the cross validation residuals vs. estimated LAI showed a consistent overestimation for low and underestimation for high LAI values, respectively (Fig. 29). Although KED cross validation predictions covered the whole range of the LAI sample values, the integrated method showed a better predictive ability (Fig. 29 d and e), smaller bias (Fig. 29 i and j), and cross validation error distribution closer to normal (Fig. 30 d).
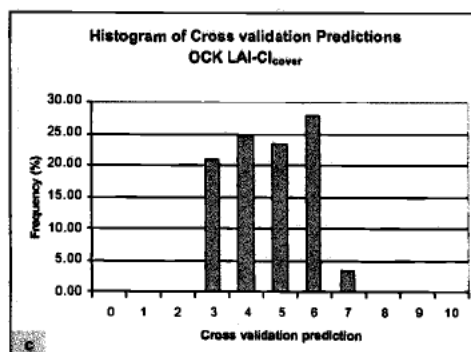
Figure 28. Histogram of cross validation predictions: a.) isotropic kriging, b.) anisotropic kriging, c.) cokriging, d.) kriging with an external drift, e.) integrated model.
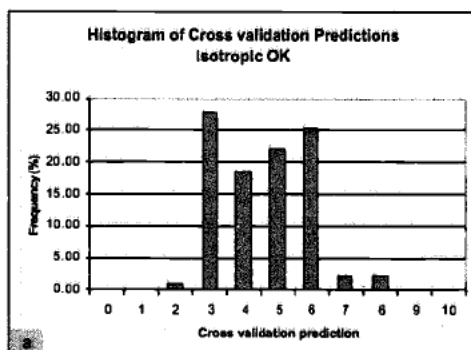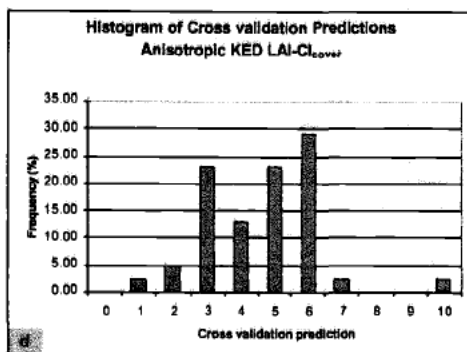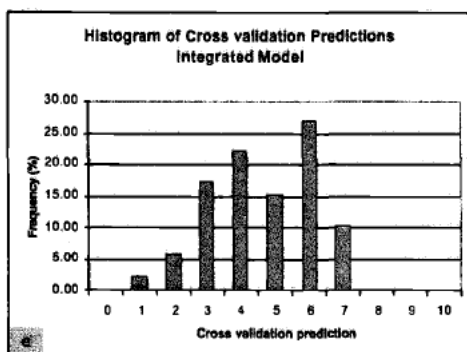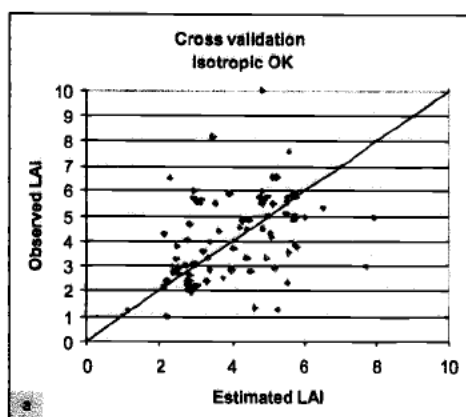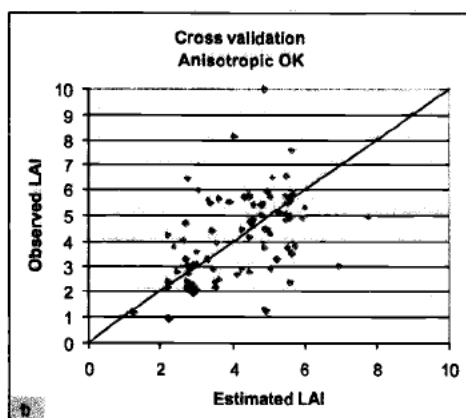
a.)
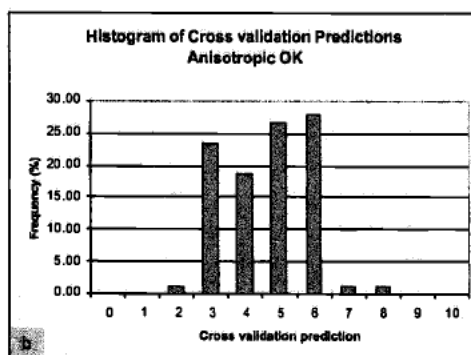
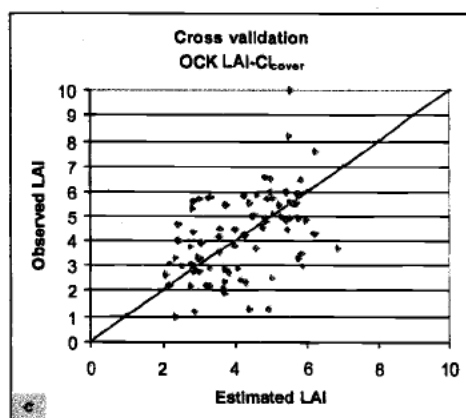b.)

c.)

(Continued)

d.)

e.)

Figure 29. Scatterplots of observed LAI versus cross validation estimated LAI: a.) isotropic kriging, b.) anisotropic kriging, c.) cokriging, d) kriging with an external drift, e.) integrated model, f.) isotropic kriging, g.) anisotropic kriging, h.) cokriging, i.) kriging with an external drift, j.) integrated model.

a.)

b.)

c.)

(Continued)

d.)

e.)

f.)

127

(Continued)

g.)

h.)

i.)

128

(Continued)

Cross validation
Integrated Model

j.)

129

Figure 30. Histograms of cross validation errors: a.) isotropic kriging,
b.) anisotropic kriging, c.) cokriging, d.) kriging with an external drift,
e.) integrated model, and experimental semivariograms of cross validation errors
f.) isotropic kriging, g.) anisotropic kriging, h.) cokriging,
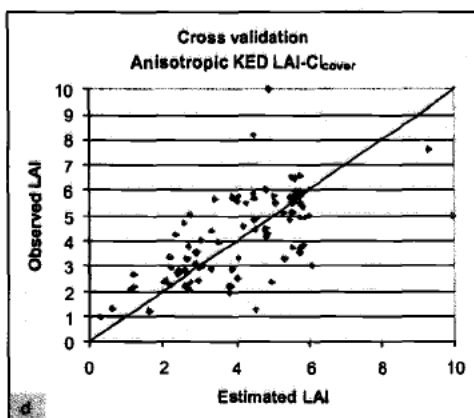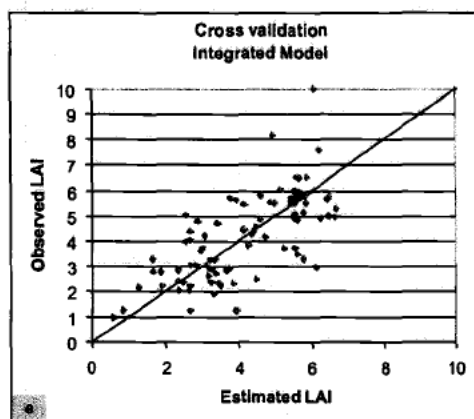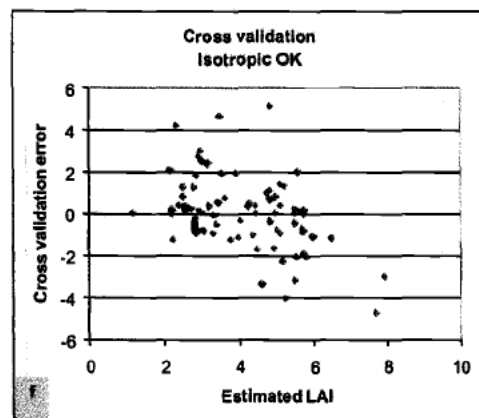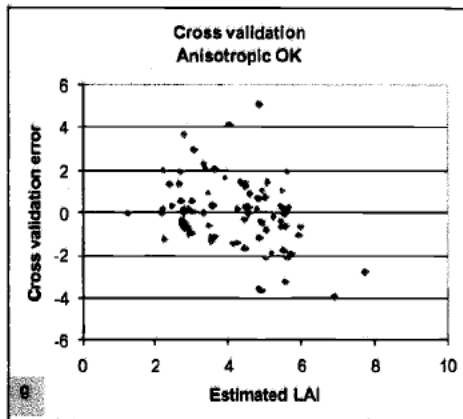i.) kriging with an external drift, j.) integrated model.
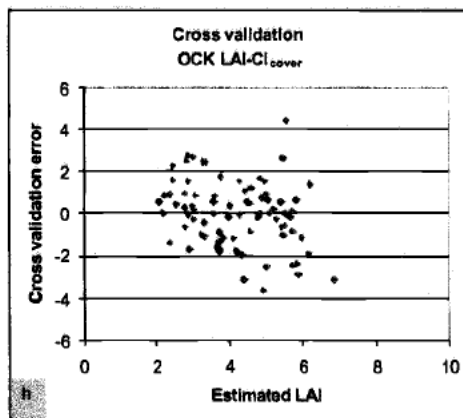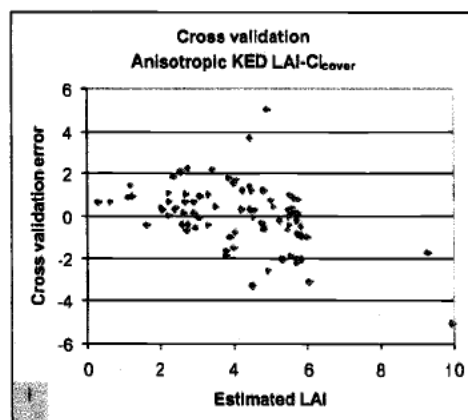
a.)

b.)

c.)

(Continued)

d.)

e.)

f.)

(Continued)

g.)
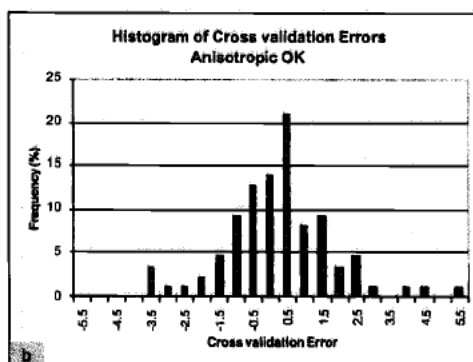

h.)

i.)

(Continued)

Semivariogram of Cross validation Errors
Integrated Model
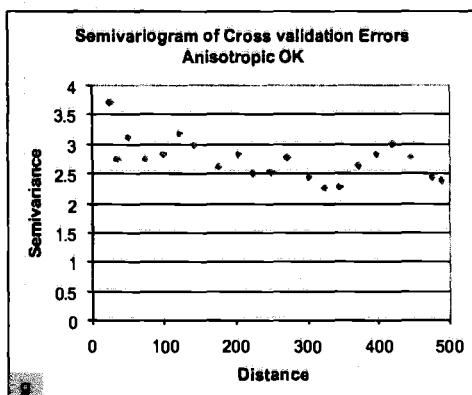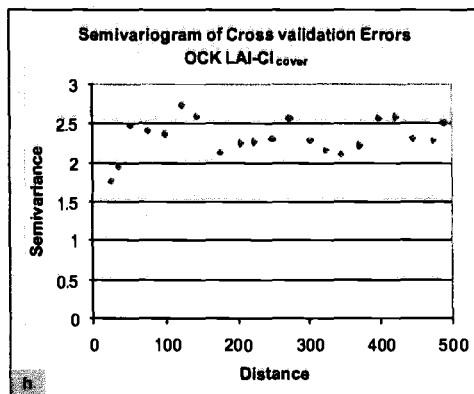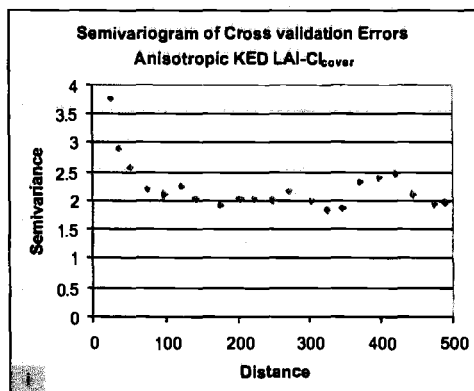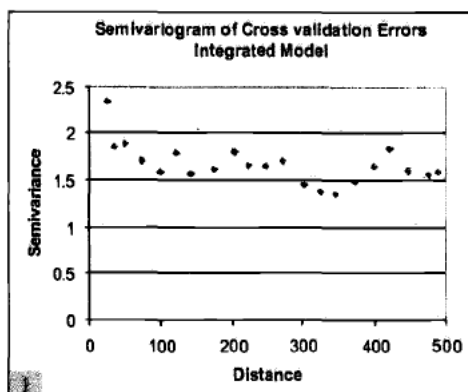
j.)

Maps of the predicted cross validated values (Fig. 31) were compared with the

sample LAI location map (Fig. 12). The N-S trend was reproduced by all methods.

The OK cross validation LAI location maps were the ones that most resembled the

sample LAI map. Geostatistical methods showed higher local accuracy than

regression models.

Transitions between different areas were smooth for regression and integrated

methods, while geostatistical methods showed a better capability of reproducing

heterogeneous contiguous areas. Although geostatistical techniques reproduced

well homogeneous areas (i.e. the northern 3 first sampling rows), in the

heterogeneous ones, high and low values were sometimes shifted (i.e. the eastern

two sampling columns).

Figure 31. Location maps of LAI cross validation estimations: a.) traditional regression, b.) inverse regression, c.) reduced major axis, d.) integrated model, e.) isotropic kriging, f.) anisotropic kriging, g.) cokriging, h.) kriging with an external drift.

135

a.)

b.)

c.)

d.)

136

(Continued)

e) Isotropic Kriging

e.)

f) Anisotropic Kriging

f.)

g) Cokriging

g.)

h) Kriging with an external drift
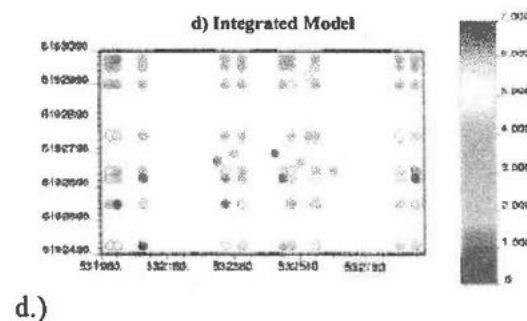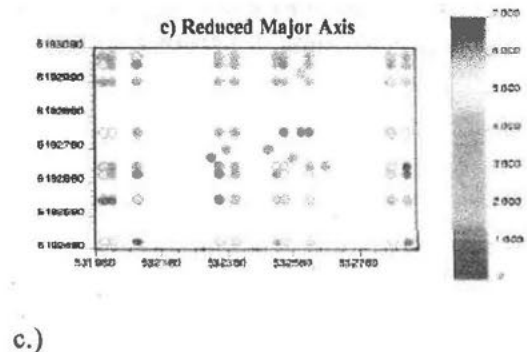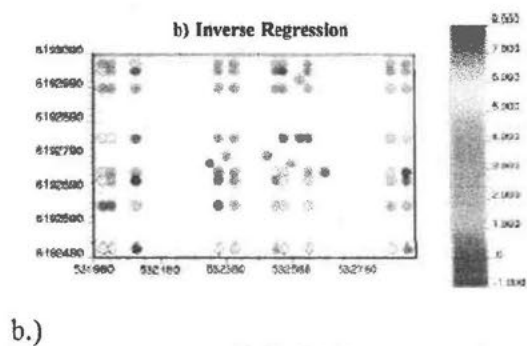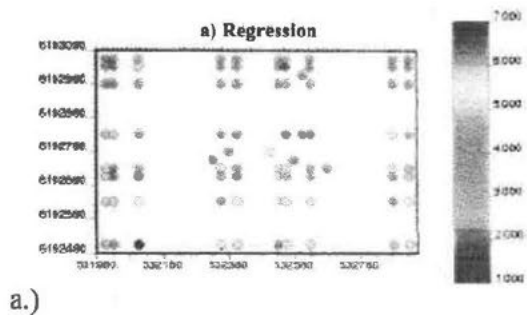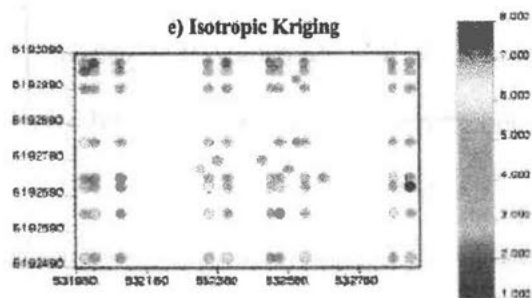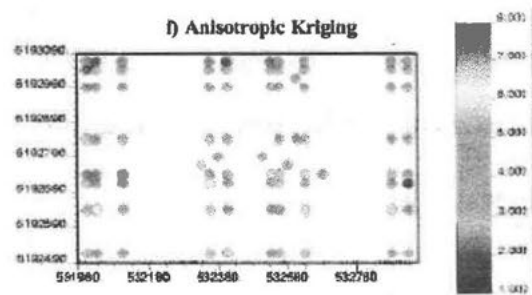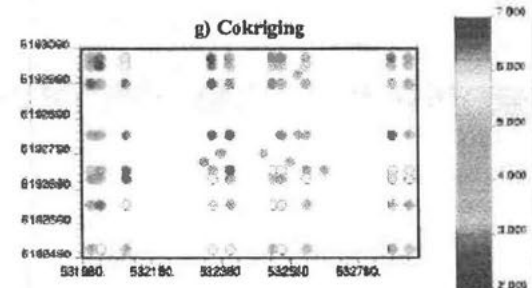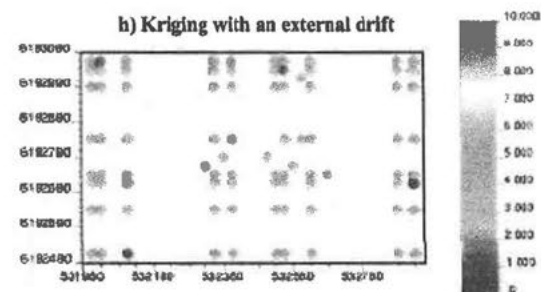
h.)

In an attempt to spatially analyze the cross validation errors, their Moran's I coefficients were calculated and maps of the errors were created for each surface. No clusters of errors were observed in the location maps of cross validation residuals (Fig. 32). OK and CK underpredicted high LAI values and overpredicted low LAI values, confirming their smoothing effect.

The error semivariograms and Moran's I coefficients showed that the residuals had very low spatial autocorrelation, although most where significant at a 0.05 level, suggesting that some LAI variability was left unexplained (Tables 13 and 16and Fig. 30). SAC of cross validation residuals of regression methods was positive while the geostatistical ones were was negative.

Table 16. Summary statistics of cross validation errors for geostatistical methods.

Method	Min	Max	Mean	Median	Std	Moran's I	p-value
Isotropic OK	-4.71	5.15	0.05	0.06	1.69	-0.10	0.00
Anisotropic OK	-3.94	5.13	0.00	0.01	1.59	-0.09	0.00
OCK CI$_{cover}$	-3.63	4.41	-0.02	0.01	1.51	-0.03	0.44
Anisotropic KED CI$_{cover}$	-5.03	5.06	0.05	0.16	1.45	-0.10	0.00
Integrated Model CI$_{cover}$	-3.12	3.90	0.02	0.06	1.24	-0.08	0.00

138

Figure 32. Location maps of LAI cross validation errors: a.) traditional regression , b.) inverse regression, c.) reduced major axis, d.) integrated model, e.) isotropic kriging, f.) anisotropic kriging, g.) cokriging, h.) kriging with an external drift.

a) Regression

a.)

b) Inverse Regression

b.)

c) Reduced Major Axis

c.)

d) Integrated Model

d.)

(Continued)

e) Isotropic Kriging

e.)

f) Anisotropic Kriging

f.)

g) Cokriging

g.)

h) Kriging with an external drift
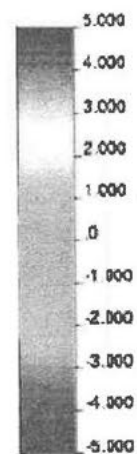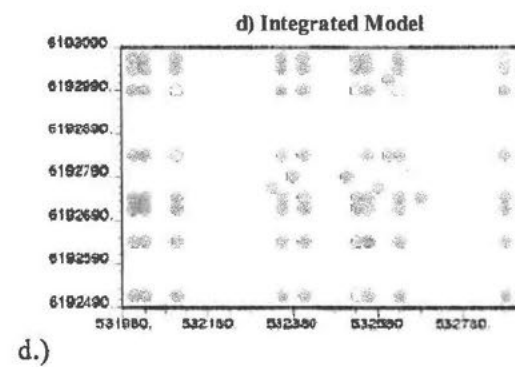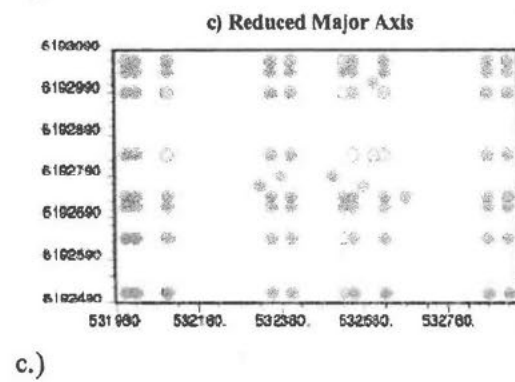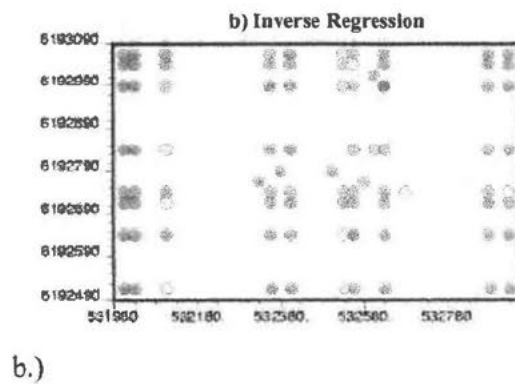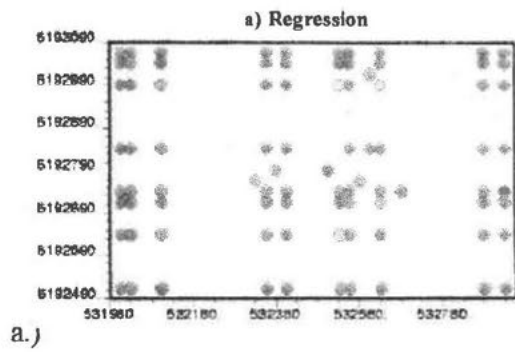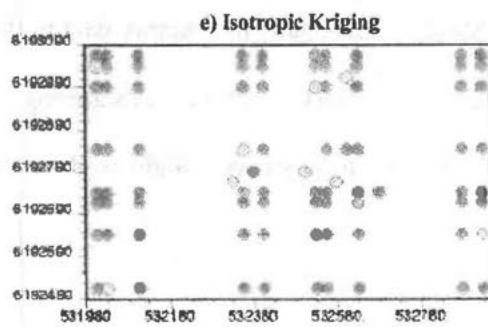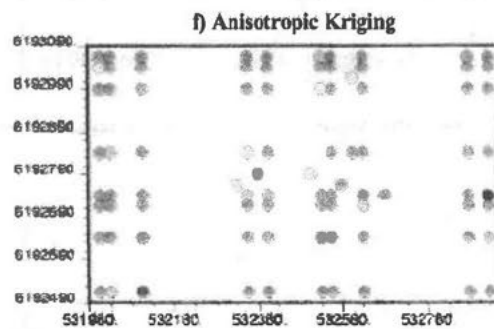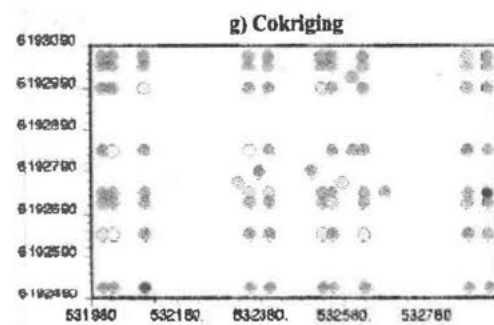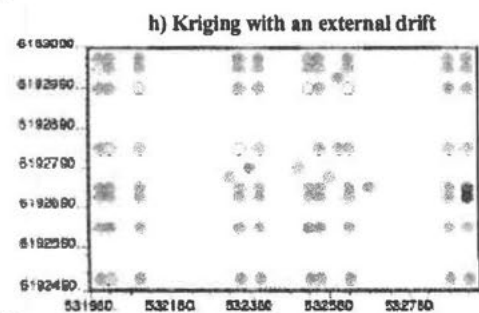
h.)
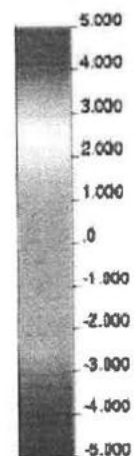
Comparison of LAI Estimated Surfaces

Because all the geostatistical methods used reproduce the sample data at their locations (exactness property), comparisons of estimated surfaces among geostatistical techniques and with the regression ones were addressed only in terms of the whole study area.

Regression methods covered the smallest LAI ranges (if negative estimations were set to 0), and produced some negative LAI values (Table17). The means were lower for regression methods, KED, and the integrated model; those methods showed also a slight negative skew.

In contrast, OK and OCK produced higher means with a slight positive skew. The highest standard deviation for Reg probably was due to the large negative predictions. The small variances of OK and OCK responded to the smoothing effect of these methods. Histograms of predictions resulted in unimodal (regression methods, OCK, integrated method) and bimodal (OK, KED) distributions (Fig. 33).

Table 17. Summary statistics of predictions for the whole area (see fig. 4) for regression and geostatistical methods.

Method Estimations (Whole area)	Min	Max	Mean	Median	STD	Coef. Var.
RegT	-3.70	6.78	3.85	4.03	1.44	0.37
RegI	-10.51	9.02	3.54	3.88	2.68	0.76
RMA	-6.58	7.73	3.72	3.96	1.96	0.53
Isotropic OK	0.98	9.98	4.21	4.40	1.22	0.29
Anisotropic OK	0.98	9.98	4.17	4.44	1.15	0.28
OCK CI$_{cover}$	0.89	10.00	4.24	4.33	1.24	0.29
Anisotropic KED CI$_{cover}$	0.00	9.98	3.77	3.33	1.49	0.40
Integrated Model CI$_{cover}$	-4.94	9.98	3.49	3.36	1.65	0.47

2-D maps (Fig. 34) showed increasing variability in LAI pattern going from Reg$_T$ to RMA to Reg$_I$. This can also be seen in the semivariograms of these surfaces (Fig. 35), where the semivariance magnitudes increased in the same order. If we compare these three maps with the CI$_{LAI}$ map, we observe that the distribution of the patches is very similar (Fig. 34), as is the anisotropy represented by the semivariograms (Fig. 35). Omnidirectional and 160 degrees semivariograms of the whole area showed increasing semivariance, with no apparent sill, suggesting that maybe a larger scale would better represent LAI variability.

143

Figure 33. Histograms of whole predicted surfaces. a.) traditional regression,
b.) inverse regression, c.) reduced major axis, d.) isotropic kriging, e.) anisotropic
kriging, f.) cokriging, g.) kriging with an external drift, h.) integrated model. The
numbers in the X axis correspond to the upper-limit of each class.

144

a.)

b.)

c.)

d.)

(Continued)

e.)

f.)

g.)

h.)
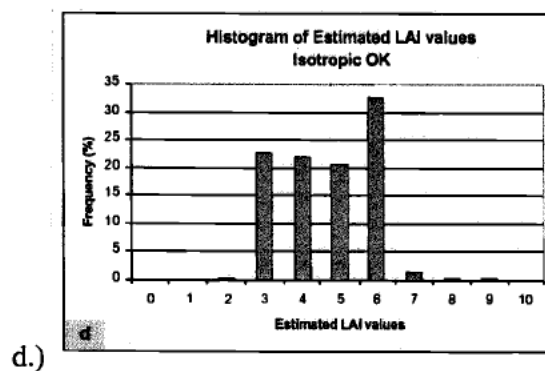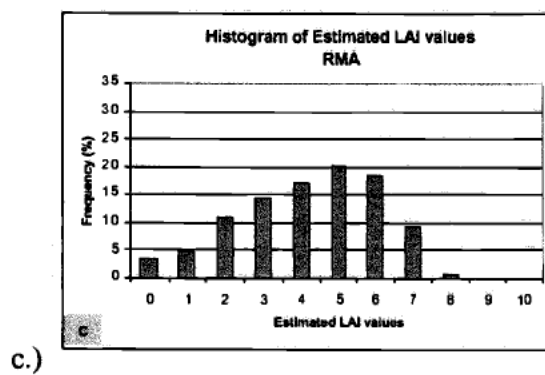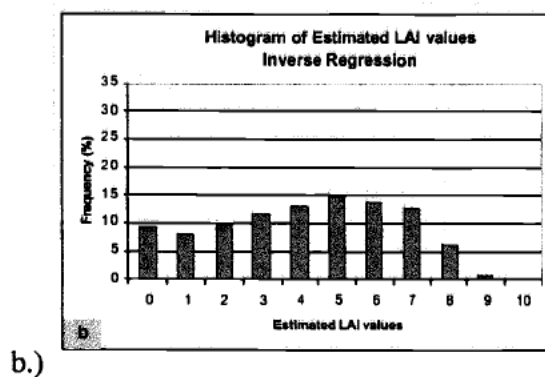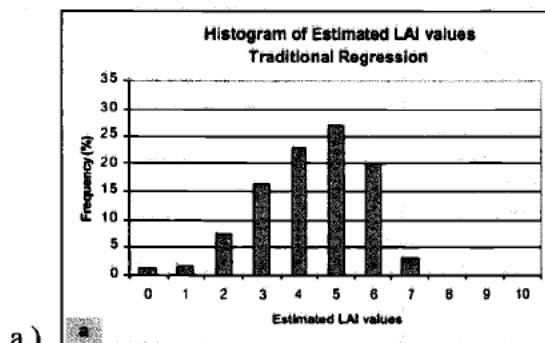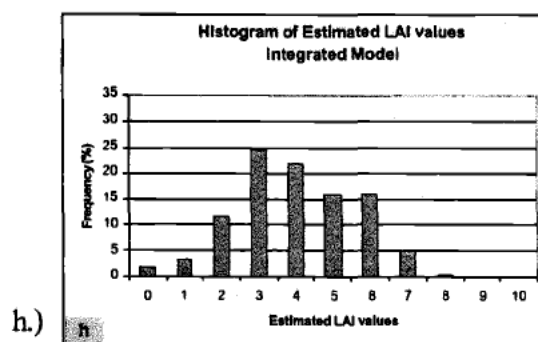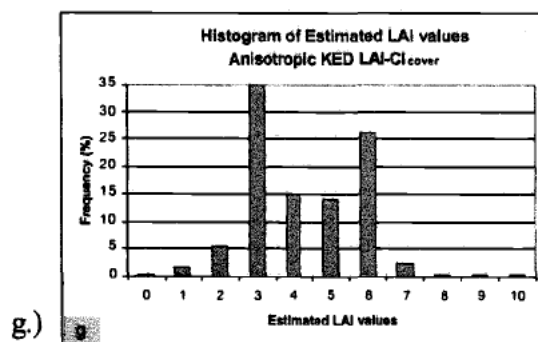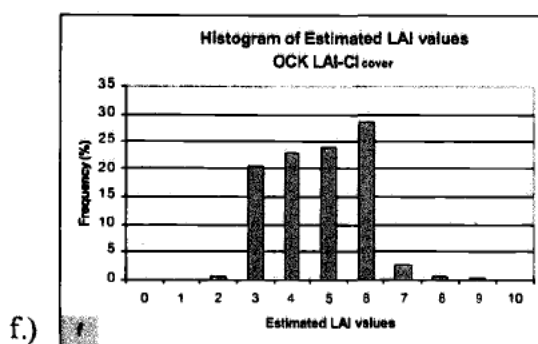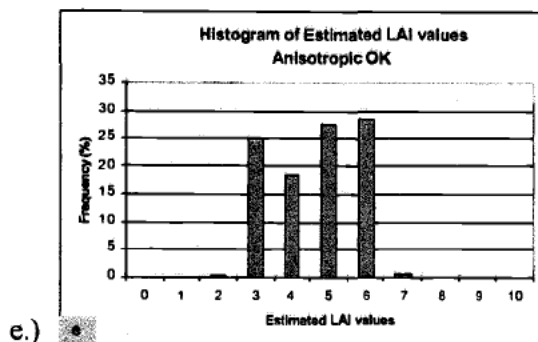
OK maps resulted in smooth surfaces, but the N-S trend can be seen in the 2-D maps (Fig. 36) as well as in the mesh plots (Fig. 39). Semivariances for isotropic kriging were a little higher than for anisotropic kriging. A possible explanation is that the isotropic model includes a hole effect model that better represents the periodicity observed in the experimental LAI semivariograms. Hole effect models cannot be applied to different directions so they cannot be used in the anisotropic model.

The variability of the maps and mesh plots increased in OCK and KED approaches with respect to ordinary kriging (Figs. 36 and 39) because of the use of extensive secondary data. OK and OCK semivariograms of the whole resultant LAI surfaces did not show the anisotropy observed in the canonical index CI_{LAI} semivariograms. They all showed the same semivariance behavior up to 100 m. KED did a better job of reproducing anisotropy features compared to OCK (Fig. 37)

Figure 34. 2-D map of a.) CI_{LAI} (used as reference to observe pattern distribution). 2-D maps of LAI predicted surfaces: b.) traditional regression, c.) inverse regression, d.) reduced major axis.
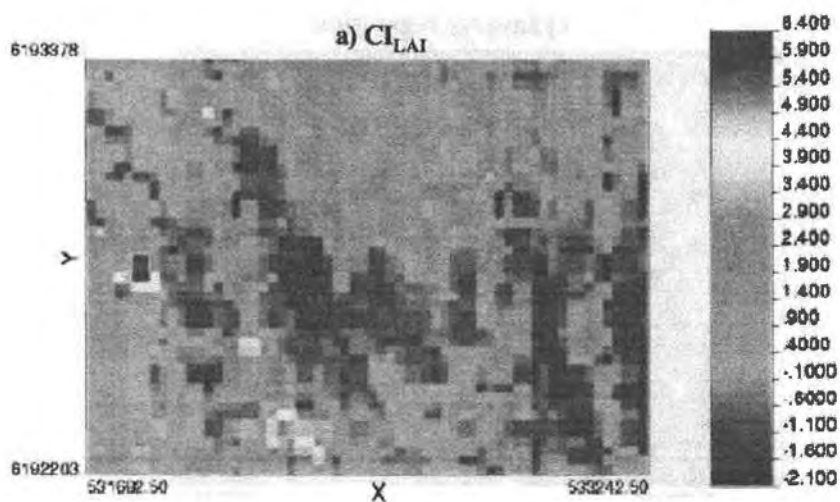
a)
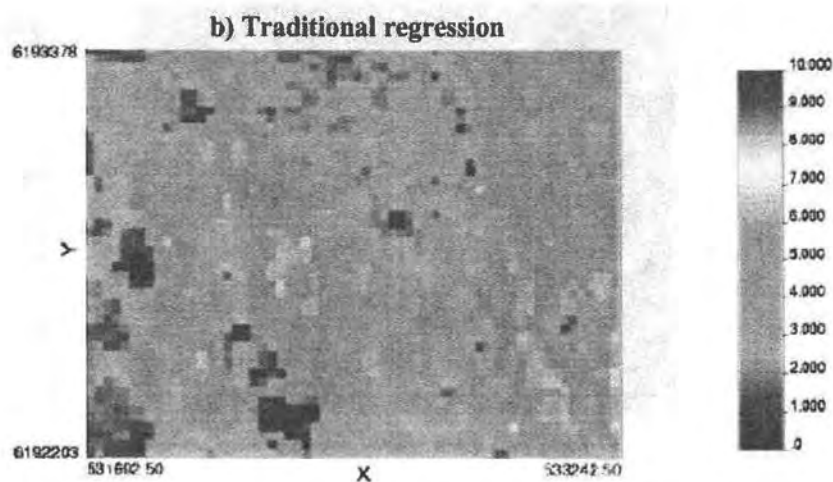
b)

(Continued)

c) Inverse regression
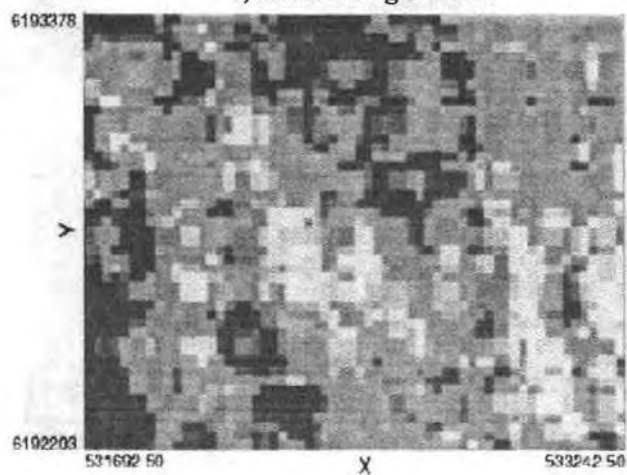
c.)

d) Reduced major Axis
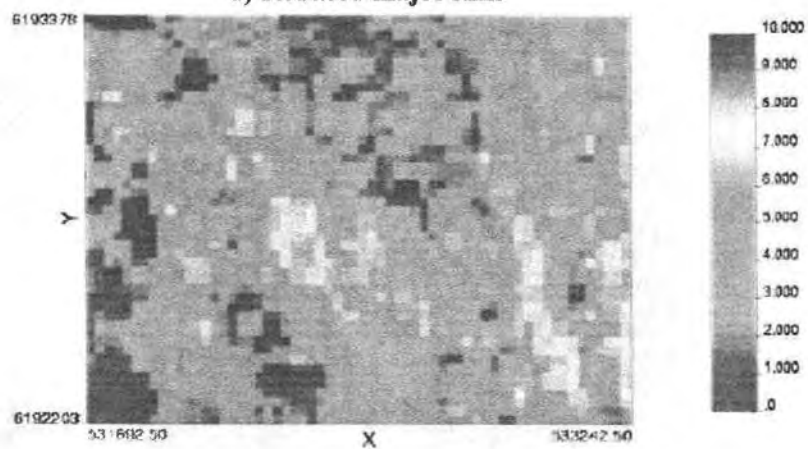
d.)

Figure 35. Omnidirectional and directional (70 and 160 degrees) semivariograms of: a.) CILAI (used as reference to observe anisotropy), and of LAI predicted surfaces: b.) traditional regression, c.) inverse regression, d.) reduced major axis. Note different scale in the Y axis.

Semivariograms of whole CI$_{LAI}$ estimated surface

a.)
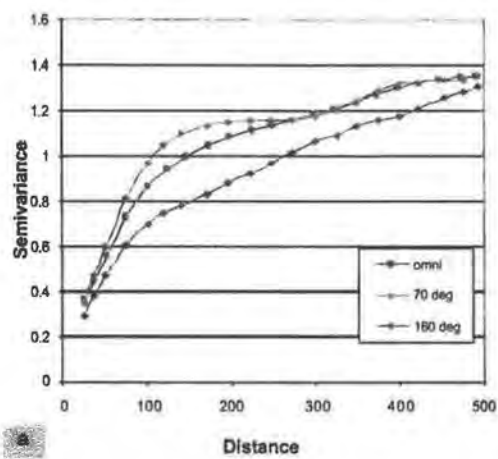
Semivariograms of whole LAI estimated surfaces
Traditional Regression

b.)

152
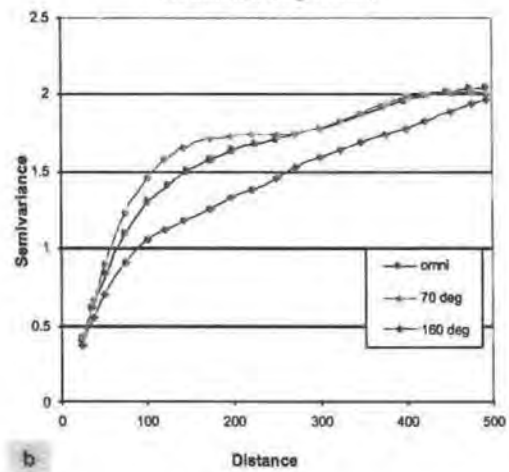
**Semivariograms of whole LAI estimated surfaces
Inverse Regression**
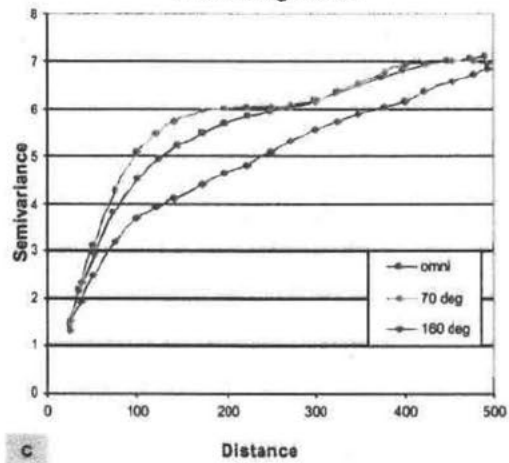
c.)

**Semivariograms of whole LAI estimated surfaces
Reduced Major Axis**
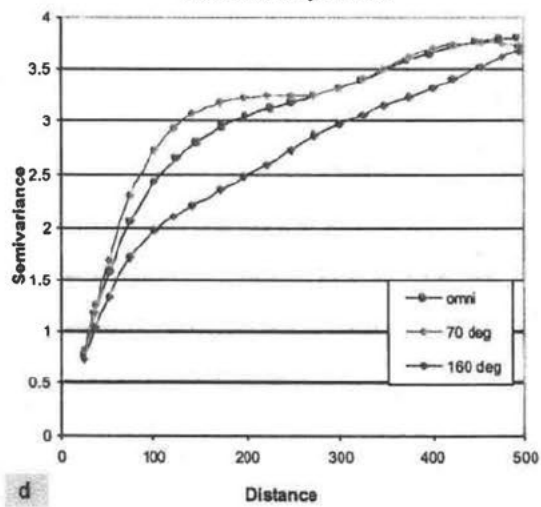
d.)

Figure 36.   2-D maps of LAI predicted surfaces: a) isotropic kriging.
2-D maps of LAI predicted surfaces: b) anisotropic kriging, c) cokriging, d) kriging
with an external drift.

154

a) Isotropic kriging
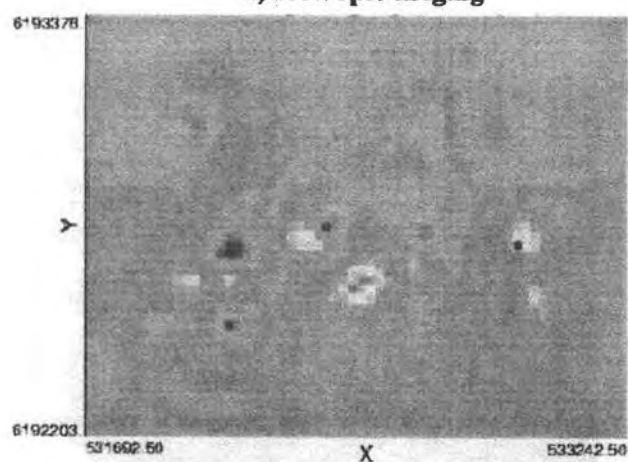
a.)

b) Anisotropic kriging

b.)

155
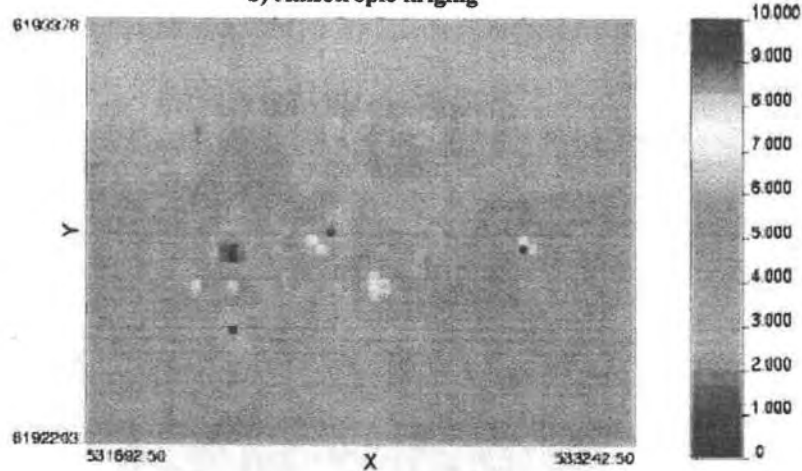
(Continued)

c) Ordinary Cokriging
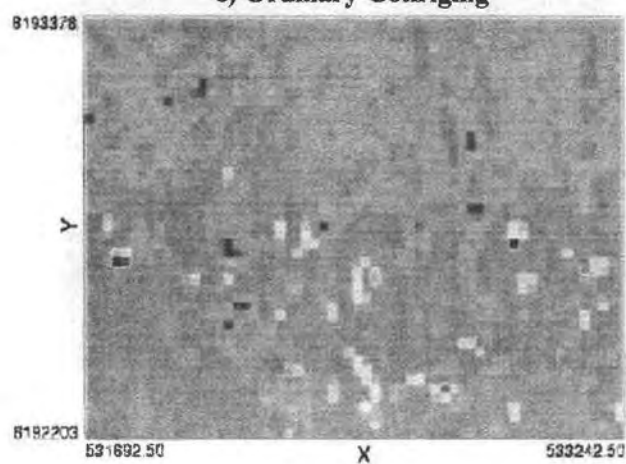
c.)

d) Kriging with external drift
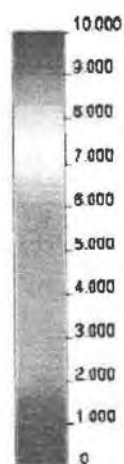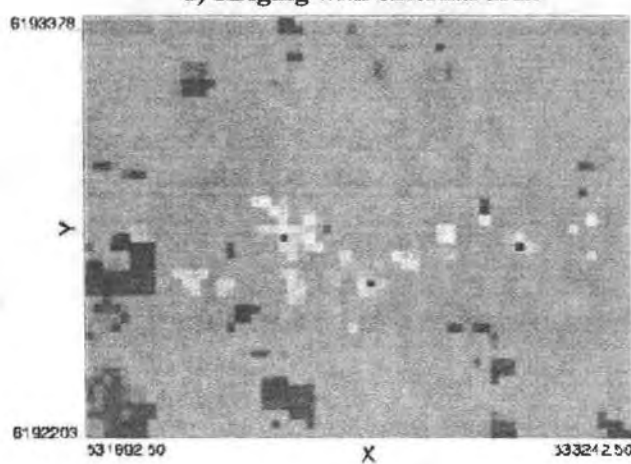
d.)

Figure 37. Omnidirectional and directional (70 and 160 degrees) semivariograms of
a.) CILAI (used as reference to observe anisotropy), and of LAI predicted surfaces:
b.) isotropic kriging, c.) anisotropic kriging, d.) cokriging, e.) kriging with an
external drift, f.) integrated model. Note different scale in the Y axis.

Semivariograms of whole CILAI estimated surface

a.)
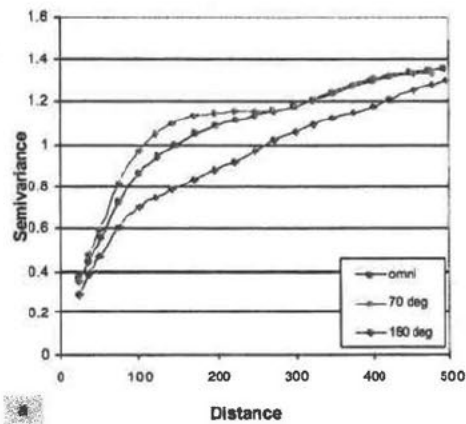
Semivariograms of whole LAI estimated surfaces Isotropic OK

b.)
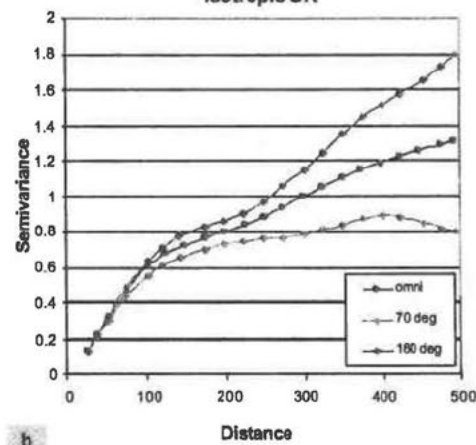
Semivariograms of whole LAI estimated surfaces Anisotropic OK

c.)

158

(Continued)

Semivariograms of whole LAI estimated surfaces
OCK CIcover

d.) d

Semivariograms of whole LAI estimated surfaces
KED CIcover

e.) e

Semivariograms of whole LAI estimated surfaces
Integrated Model
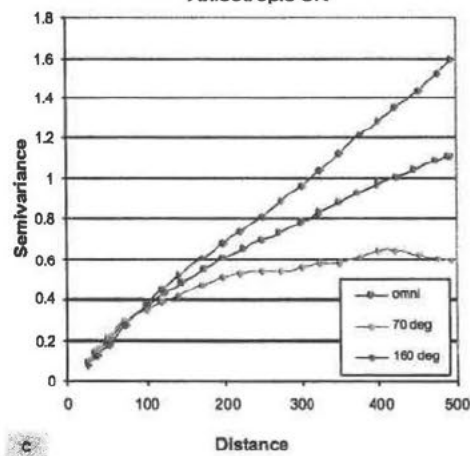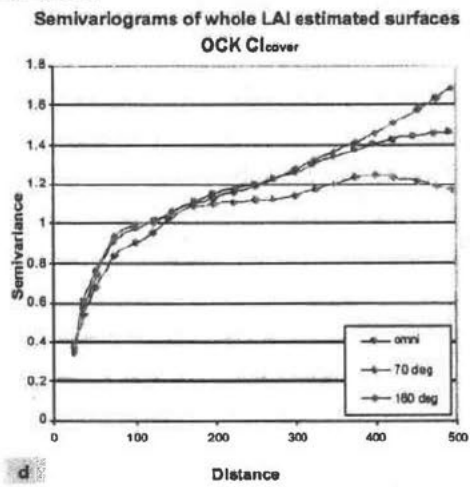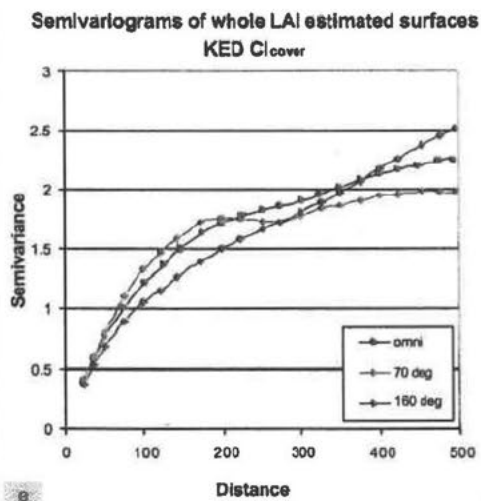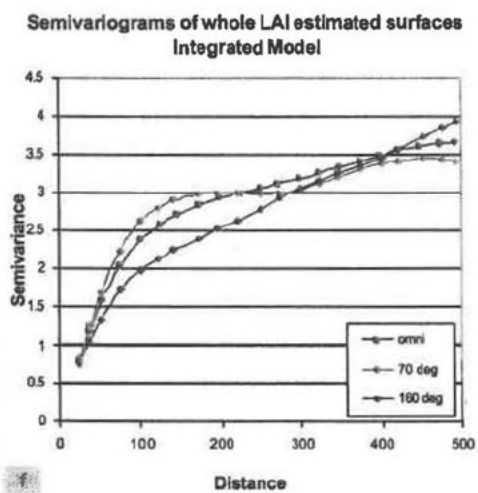
f.) f

The integrated model showed reduced variability with respect to the RMA surface (Fig.38), but the semivariograms in figs. 35 and 37 showed that this reduction in variability was more pronounced in the direction of higher spatial discontinuity (70 degrees). The omnidirectional semivariogram remained very similar to the omnidirectional RMA one, while at large distances, the variability increased for the direction of major continuity (160 degrees).

Figure 38. 2-D maps of LAI predicted surfaces: a.) reduced major axis LAI predictions, b.) reduced major axis residuals cokriged with CI_{cover}, c.) integrated model, addition of a + b.

a.)

b)

c)

162
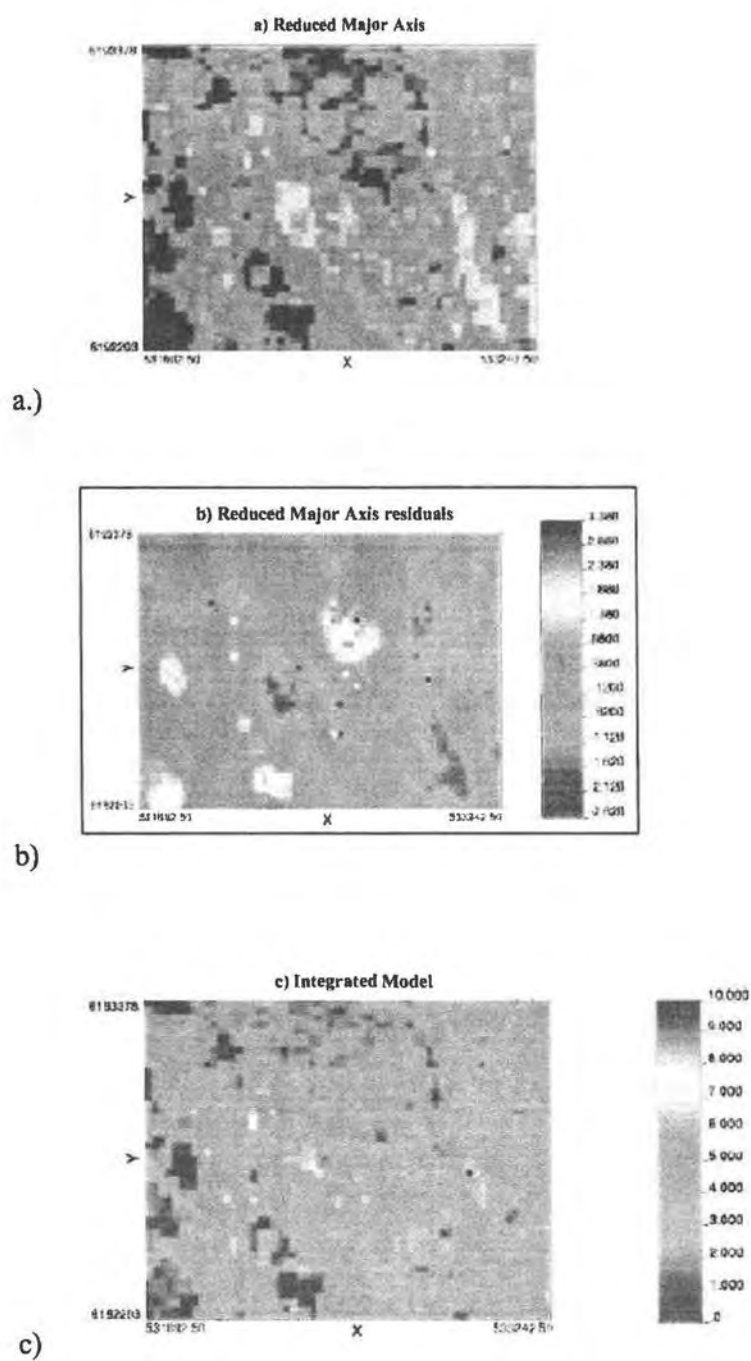
Figure 39. Mesh plots of LAI predicted surfaces: a.) CILAI(used as reference to observe patterns and trends), b.) traditional regression, c.) inverse regression, d.) reduced major axis. E.) isotropic kriging, f.) isotropic kriging, g.) cokriging, h.) kriging with an external drift, i.) integrated model.

163

a) Canonical Index CI_LAI
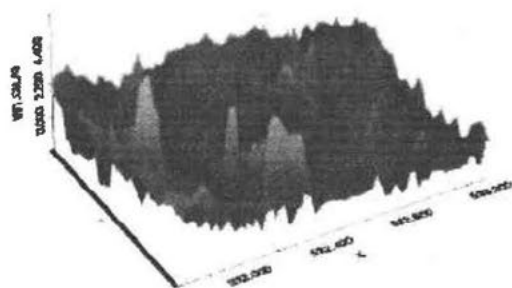
a.)

b) Traditional Regression
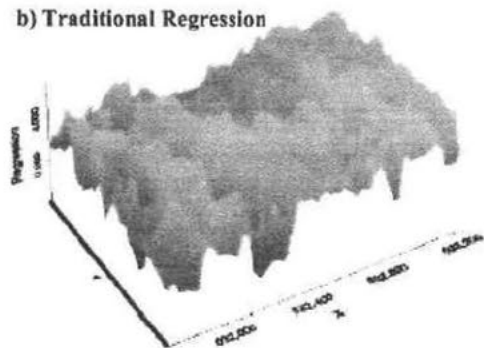
b.)

c) Inverse Regression
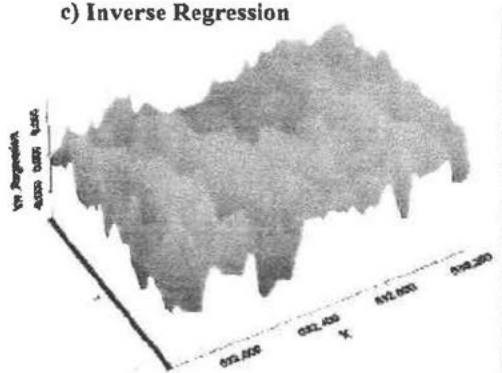
c.)

(Continued)

d) Reduced Major Axis
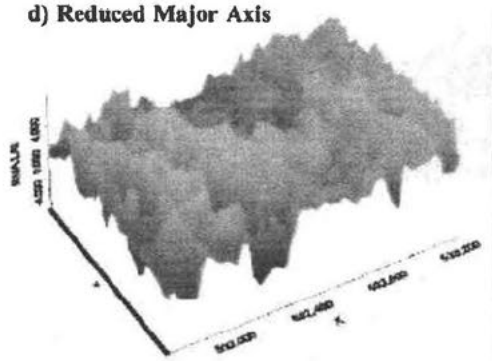
d.)

e) Isotropic Ordinary Kriging
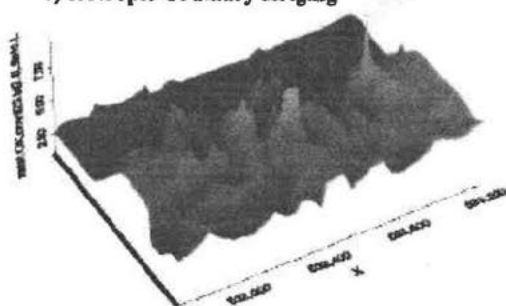
e.)

f) Anisotropic Ordinary Kriging
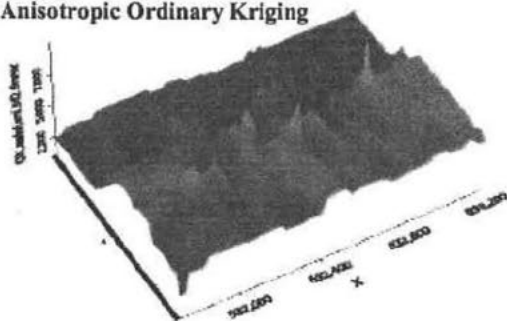
f.)

(Continued)

g) Ordinary Cokriging

g.)

h) Kriging with an External Drift
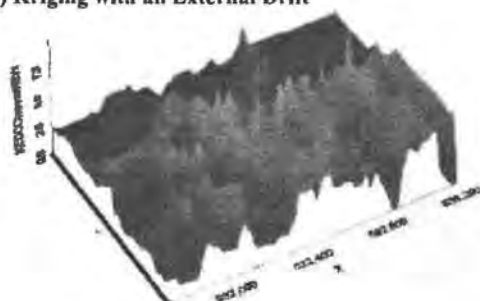
h.)

i) Integrated model
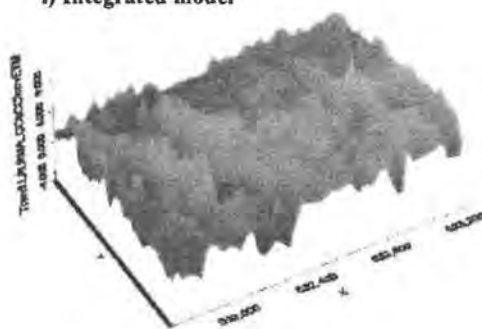
i.)

Considering all non-simulation methods, the integrated model was the technique that had best cross validation results, and the most realistic pattern distribution with an improved local accuracy over RMA.

Geostatistics: Simulation Methods

Conditional simulation realizations resulted in bimodal distributions (Fig. 40).

Means and medians were very close, showing less bias that non-simulation

methods (Table 18). Standard deviations were higher than those of the non-

simulation methods, and similar to the integrated model variance.

Table 18. Summary statistics of sequential Gaussian conditional simulations for the whole area (see fig. 4).

Simulations (Whole area)	Min	Max	Mean	Median	STD	Coef. Var.
SGCS realization 1	0.05	9.99	4.13	4.18	1.61	0.39
SGCS realization 2	0.07	10.00	4.19	4.25	1.68	0.40
SGCS realization 15	0.03	10.00	4.10	4.10	1.66	0.40
SGCS E-type	0.74	9.98	4.16	4.09	1.27	0.31

Conditional simulation maps showed increased spatial variability with respect to

the other geostatistical maps and more closely reflected the pattern distribution

displayed by the CI_{LAI} map, denoting a higher global accuracy (Figs. 41 and 42).

The N-S trend was visible and the semivariograms reproduced the observed

anisotropy (Fig. 43).

Figure 40. Histograms of simulated surfaces: a.) sequential Gaussian conditional simulation, realization 1, b.) realization 2, c.) realization 15, d.) E-type. The numbers in the X axis correspond to the upper-limit of each class.

Histogram of SGCS LAI values
LAI-Cl cover Realization 1

a.)

Histogram of SGCS LAI values
LAI-Cl cover Realization 2

b.)

Histogram of SGCS LAI values
LAI-Cl cover Realization 15

c.)

Histogram of SGCS LAI values
LAI-Cl cover E-type

d.)

170
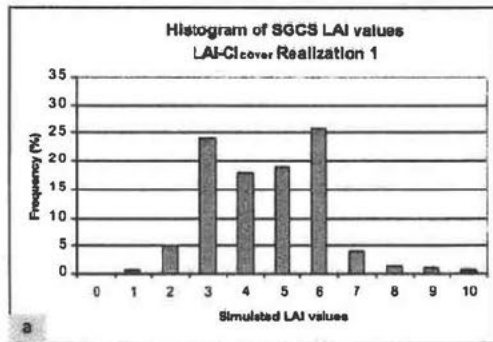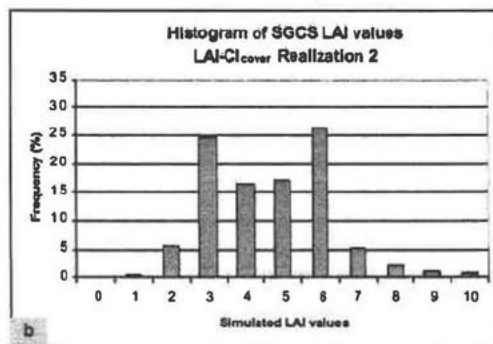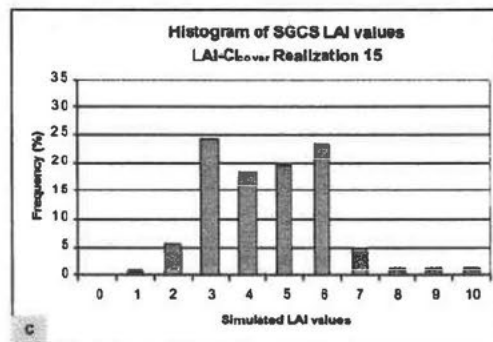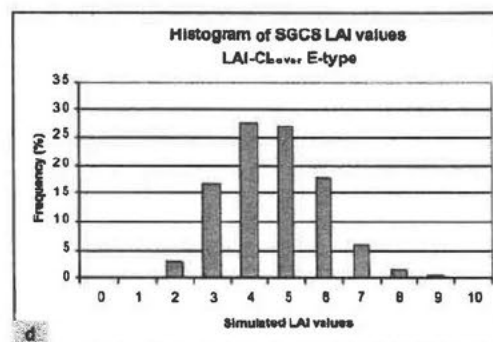
Figure 41. Examples of 2-D LAI simulated surfaces: a.) sequential Gaussian conditional simulation, realization 1, b.) realization 2, c.) realization 15. d.) SGCS E-type of 51 simulations.

a) SGCS Realization1

a.)

b) SGCS Realization 2

b.)
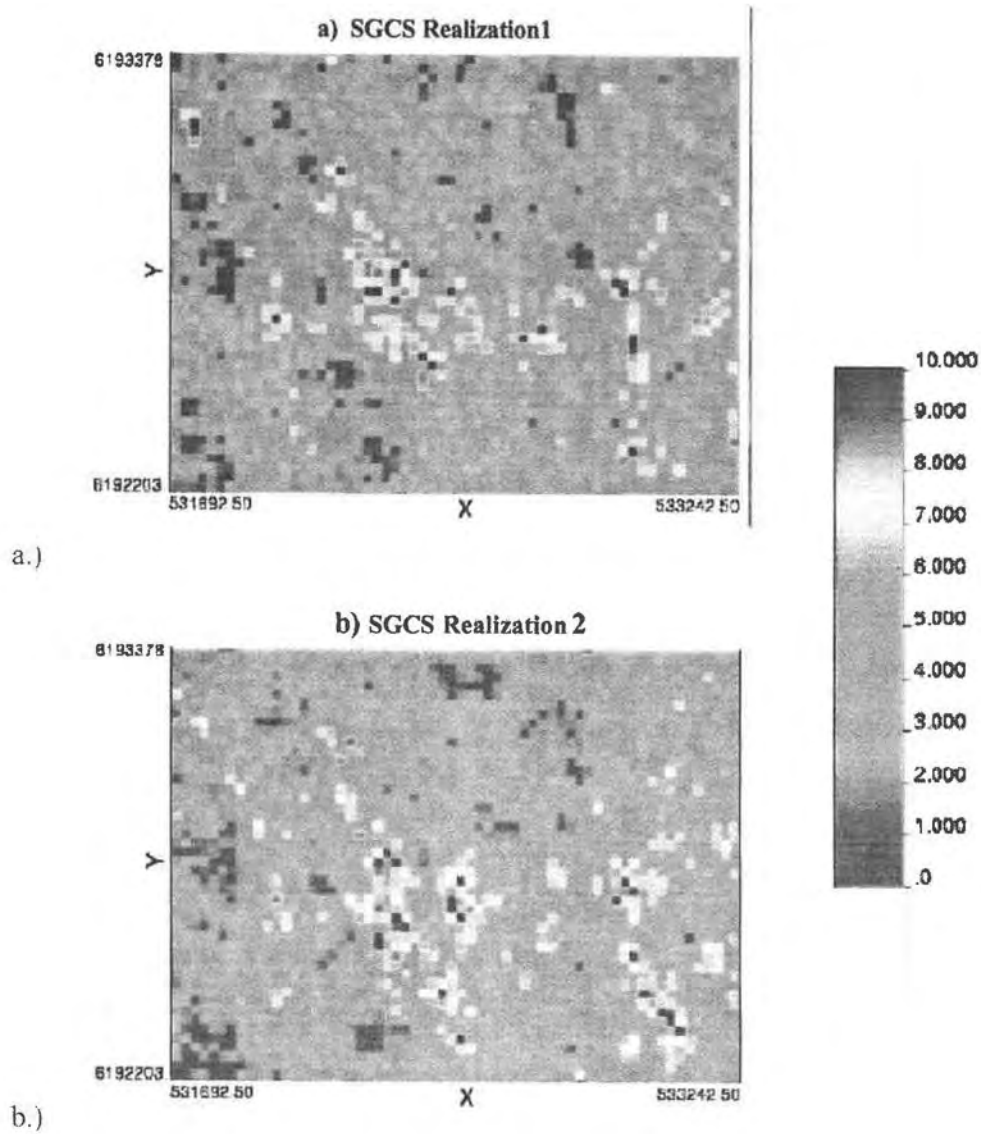
(Continued)

c) SGCS Realization 15
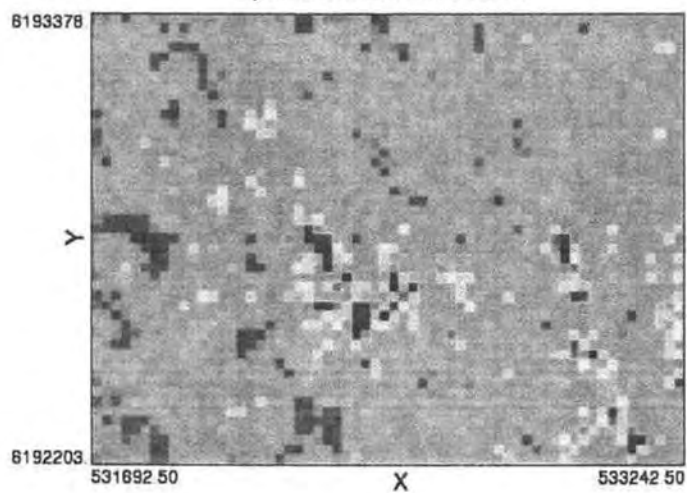
c.)

d) SGCS E-type
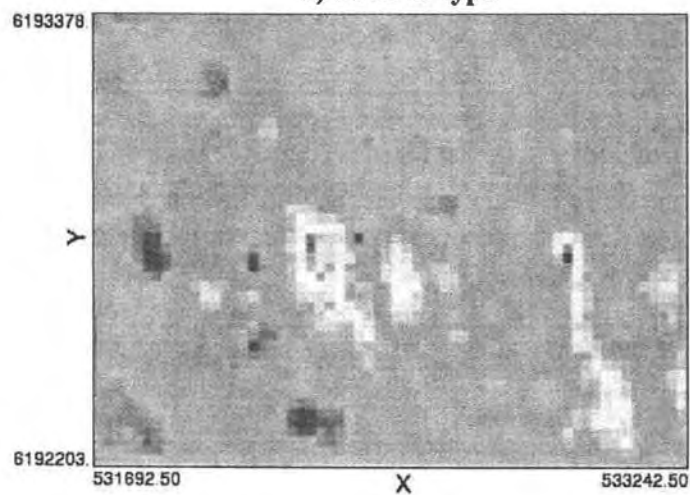
d.)
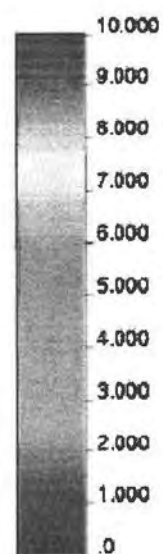
Figure 42. Mesh plots of LAI simulated surfaces: , a.) SGCS realization 1, b.) realization 2, c.) E-type of 51 simulations.

a) Realization 1
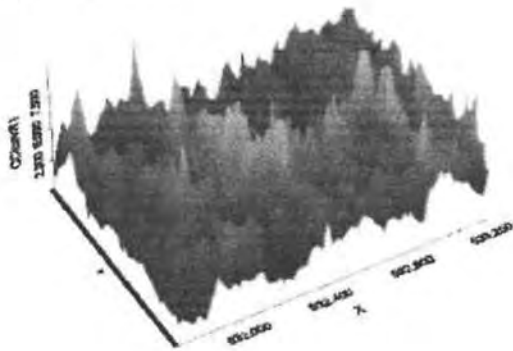
a.)

b) Realization 2
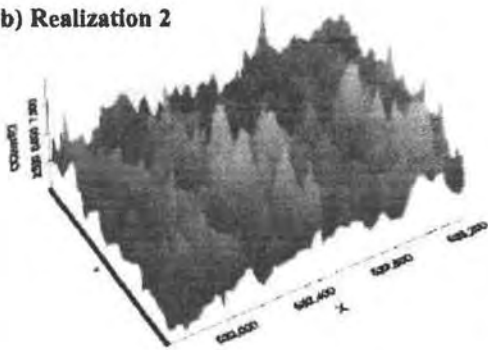
b.)

c) E-type
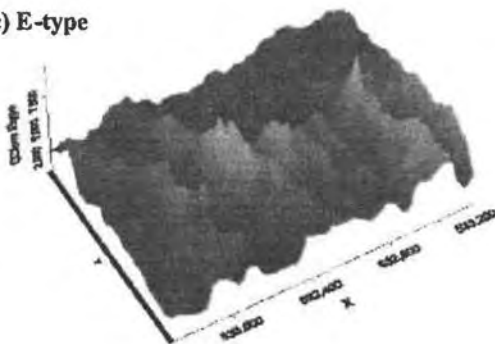
c.)

Figure 43. Omnidirectional and directional (70 and 160 degrees) semivariograms of LAI simulated surfaces: a.) SGCS realization 1, b.) realization 2, c.) realization 15, d.) E-type. Note different scale in the Y axis.

Semivariograms of whole LAI simulated surfaces
SGCS Realization 1

a.)

Semivariograms of whole LAI simulated surfaces
SGCS Realization 2

b.)

Semivariograms of whole LAI simulated surfaces
SGCS Realization 15

c.)
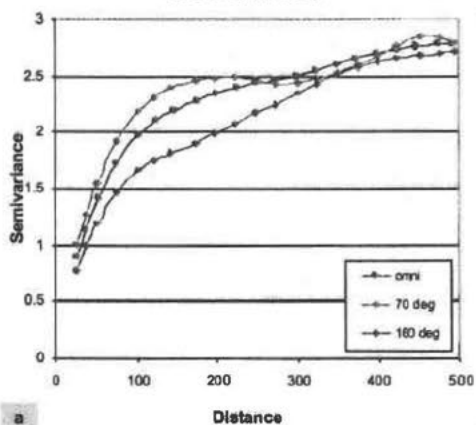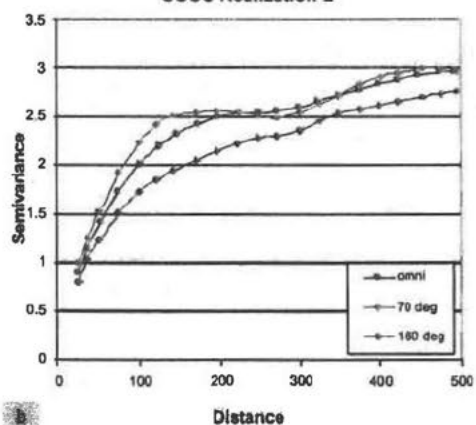
(Continued)

Semivariograms of whole LAI simulated surfaces
SGCS E-type
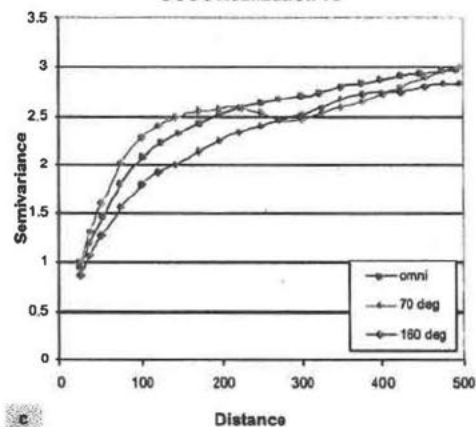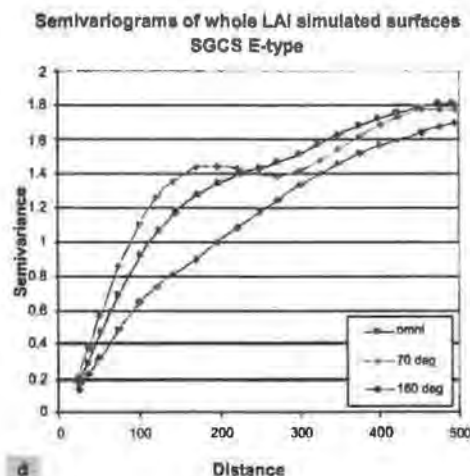
Two more conditional simulation runs were performed varying the model of the primary variable in the first case, and the secondary variable used in the second place, to evaluate how the uncertainty of the maps changed. The run described above modeled LAI with an anisotropic model and used CI_{cover} as a secondary variable, resulting in an uncertainty of 0.39. The second run modeled LAI with the isotropic model used in kriging, and resulted in a slight reduction of the uncertainty to 0.35. The third run used the isotropic model for LAI and CI_{LAI} as the secondary variable, decreasing the uncertainty to 0.31.

Discussion and Conclusions

The results obtained in this study suggest that when mapping natural resource attributes such as LAI, the optimal technique to select depends on the final uses of the map, because not all the desired map characteristics can be achieved in a unique model.

The location of the ground measurements has two important aspects: accuracy in the match between ground and image data (georectification), and representativeness of the ground data for the area to be estimated. Representativeness, defined by Myers (1997) as "the degree to which sample data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, or an environmental condition", is of crucial importance in some geostatistical methods because attributes not sampled will not be present in the final outputs. A good example of this is the hardwood area in the SW corner of the image (Fig. 4), which was not sampled. This area was reproduced in the output maps by regression methods, although it constitutes an extrapolation situation.

The same area was ignored by kriging and cokriging maps, but kriging with an external drift and conditional simulation accounted for it because of their use of the auxiliary variables. An unsupervised classification and/or a feature space analysis are recommended as first steps in evaluating representativeness of the sampling design.

The sampling design plays a crucial role in the process of producing environmentally-derived maps. The spatial distribution of the ground measurements may have great influence in the final results. The systematic spatial cluster sampling design used was appropriate for this study because of its efficiency in diminishing redundant data collection, while distributing pairs of points at different lags. However, orientation of the sampling units was not optimal for our purposes because it did not correspond to the maximum direction of continuity (Fig. 19). A basic a-priori anisotropy analysis of the image could have avoided this inconvenience. Nevertheless, the available samples and the anisotropy analysis allowed the identification of different spatial ranges that may be a result of the glacial drift occurring in the area.

Measurement errors are intrinsic in both ground and image measurements. Traditional regression and inverse regression take into account measurement errors only in the response or the explanatory variables respectively, assuming that the other variable is free of error. RMA is a useful improvement for many applications, including this one, because it takes into account error measurements in both the response and explanatory variables. The results of this study agreed with Curran and Hay. (1986), who found the same effect when the regression slope was underestimated under the assumption of measurement free of errors for the explanatory variable.

Measurement support, defined as the shape, size and orientation of the area or volume being measured (Curran, 1999), is significant in determining at which scale the analysis is valid. This study tried to match the ground-based measurements to Landsat ETM+ pixel size. In both cases, averages over the support surface are used, in which case the variograms are called "regularized". The effects of regularization are the same as those for aggregation: the variance is reduced, the range increases, and the mean remains unchanged (Isaaks and Srivastava, 1989; Woodcock et al., 1988).

The variance reduction may be more important if there is a large nugget effect, as was observed when going from the RMA to the integrated model. One possible explanation for this reduction is that originally, LAI had low spatial autocorrelation, and the RMA model accounted for most of its variability. The residuals of the RMA model had a high nugget effect and very low spatial autocorrelation, and although Moran's I coefficient was significant (at a 0.05 significant level), there was little spatial autocorrelation to be further explained. The high nugget effect suggested that there was unexplained variability at scales smaller than the support.

Statistical properties of spatial attributes strongly depend on the support of the observations (Heuvelink, 1998). From the results of this study, processes concerning LAI at this particular forest are relevant at scales smaller than 150 m, so the scaling up to MODIS resolutions (250 to 1000 m) and the use of layers of

information at these scales to inform process models would imply certain amount of unknown error that should be quantified to assess the relevance of the final NPP maps.

Spatial autocorrelation is present in natural environments at all spatial scales (Legendre, 1993). This is reflected in the patterns observed in most ecological phenomena. If spatial autocorrelation is present, one can decide to remove it so that traditional statistical procedures can be used. A preferable alternative is to take it into account by using other statistical methods such as geostatistics. Although in this study LAI exhibited low spatial autocorrelation, the geostatistical techniques exploited it and improved some features of the maps, thus showing that in some cases it is worth the extra modeling effort.

Table 19 shows a summary of the spatial variability features obtained in the LAI maps identified as trend, pattern and anisotropy, and explains how each method accounted for them. In general, regression methods reproduced pattern and anisotropy well but were not good at reproducing trends. A possible explanation for this is that the regression coefficients were calculated over the entire area. When search neighborhoods were applied in the geostatistical methods, the trends became more apparent. Kriging and cokriging resulted in smoothed patterns, and a poor anisotropy reproduction that may be a consequence of the minimization of the variance. Kriging with an external drift improved the pattern reproduction but had

less predictive ability than the integrated method. Although the integrated model had a reduced semivariance over RMA, it performed best among the regression and geostatistical estimation methods if a compromise between global and local accuracy were the objective.

Conditional simulation reproduced trend, pattern and anisotropy well. Conditional simulation used the secondary data in a different way than cokriging and kriging with an external drift. While cokriging uses the full secondary data through the cokriging system, kriging with an external drift uses them only to characterize the trend of the primary variable, and conditional simulation uses only the colocated data scaled by the correlation coefficient between the primary and the secondary variables. Though KED maps reproduced a better pattern inside as well as outside the sample grid (Fig. 36), not all the features were reproduced in the extrapolation situation. This is more evident in the northern area. Although the anisotropic LAI model used in KED and in OK was the same, the semivariograms of the whole area showed a better anisotropy reproduction for KED, suggesting that the secondary information provided by CI_{cover} was useful in accounting for LAI spatial variability.

Besides, the use of this secondary information seemed to be more appropriate in the KED case, compared to OCK. This may be explained by the way OCK incorporated the full secondary data, some of which may be redundant. This is confirmed by the colocated use of the secondary data by conditional simulation, which yielded the best pattern representation of all.

The methods used in this study suggest that high levels of local and global accuracy cannot be achieved simultaneously. Kriging techniques aim to provide estimators as close as possible to the true unknown values, with the criteria of unbiasedness and minimal estimation variance. The minimization of the estimation variance involves a smoothing of the true distribution of the original variable, resulting in poor reproduction of spatial variability. On the other hand, conditional simulation values are not the best possible estimators on a point by point basis, but the variance of estimation is greater than the kriging variance, resulting in a better reproduction in the pattern of the modeled variable (Journel, 1978; Dungan, 1999).

If local accuracy were the objective, and RMA the desired method, it would be preferable to stratify first by land cover and apply different models for the different cover types, as suggested by Turner et al. (1999).

Table 19. Summary of the relevant characteristics of the resultant maps.

Method	Reproduction of summary statistics	Level of accuracy	Error, uncertainty measures?
Traditional Regression Reg$_T$	*Reduced range of values, low STD, poor histogram reproduction	Low local	Standard error (aspatial)
Inverse regression Reg$_I$	*Reduced range of values (negative values not considered), high STD, poor histogram reproduction, negative estimated values	Low local	Standard error (aspatial)
Reduced major axis RMA	*Reduced range of values, similar STD, poor histogram reproduction	Low local	Standard error (aspatial)
Ordinary kriging OK	*Reduced range of values, low STD, low mean, improved histogram reproduction	High local Low global	The error variance only depends on the data configuration, not on the data values (not useful as an error measurement)
Kriging with an external drift KED	*Full range of values, similar STD, improved histogram reproduction, low mean	High local Low global	The error variance only depends on the data configuration, not on the data values (not useful as an error measurement)
Ordinary cokriging OCK	*Reduced range of values, low STD, poor histogram reproduction	High local Low global	The error variance only depends on the data configuration, not on the data values (not useful as an error measurement)
Integrated methods	*Reduced range of values, low STD, poor histogram reproduction	Improved local over RMA	The error variance only depends on the data configuration, not on the data values (not useful as an error measurement)
Conditional simulation SGCS	** Full range of values, similar STD, good histogram reproduction, similar mean	Low local High global	Quantitative and qualitative assessment of global uncertainty is possible.

* Based on statistics of cross validation predictions.
** Based on statistics of whole simulated surfaces.

Table 19 (continued).

Method	Spatial variability features		
	Trend reproduction	Pattern reproduction	Anisotropy reproduction
Traditional Regression Reg$_T$	N-S trend not very evident	Original pattern reproduced	Accounted for by the sampling and the image information
Inverse regression Reg	N-S trend not very evident	Original pattern reproduced	Accounted for by the sampling and the image information
Reduced major axis RMA	N-S trend not very evident	Original pattern reproduced	Accounted for by the sampling and the image information
Ordinary kriging OK	Accounts for trends because of local estimation of the mean within each search neighborhood.	Original pattern smoothed. Smoothing not uniform, less smooth where more samples are available	Accounted for by the variogram model and search neighborhood but is not shown well because of minimized variance
Kriging with an external drift KED	Accounts for local trends within search neighborhoods, calculated as smooth varying functions of the secondary variable	Original pattern smoothed, but improved over kriging and cokriging. More realistic in extrapolated areas	Improved over OK and OCK, accounted for by the variogram model and the search neighborhood
Ordinary cokriging OCK	Accounts for trends because of local estimation of the mean within each search neighborhood.	Original pattern smoothed, but improved over kriging. Smoothing not uniform, less smooth where more samples are available	Accounted for by the variogram model and search neighborhood but is not shown well because of minimized variance
Integrated methods	N-S trend improved over RMA	Original pattern reproduced	Accounted for by the sampling, the image information, the variogram model, and the search neighborhood
Conditional simulation SGCS	Accounts for trends because of local estimation of the mean within each search neighborhood.	Emphasizes spatial pattern through addition of missing kriging variance	Accounted for by the semivariogram model of the primary variable, the search strategy, and multiple grid simulation

Estimation methods that take a deterministic approach provide a single estimate map with aspatial (i.e. standard error in regression techniques) or poor (i.e. kriging variance) assessment of errors. In contrast, simulation methods take a probabilistic approach, where simulated values are conditional to the samples, honor their statistics, and provide a spatially quantitative measure of uncertainty.

A good spatial error map should depend on the values of the samples, their distances and geometry. Fig. 44 shows an E-W transect that goes through 10 samples. The kriging error variance shows the variance is dependent on the data configuration only; the variance decreasing when approaching the sample locations, and increasing when going away from them. In contrast, the conditional variance depends on the data configuration and the data values. The variance was greater when two consecutive samples had dissimilar LAI values, and smaller when consecutive values were more alike. The conditional variance of the whole set of realizations provided a visual and quantitative spatial uncertainty measure (Fig32 a)

Figure 44 a.) LAI sample values (red dots) and the corresponding error variances of the kriging variance. b.) LAI sample values (red dots) and the corresponding error variances for the conditional variance. Kriging variance only depends on data configuration, while conditional variance depends on both, data configuration and data values.

OK local uncertainty
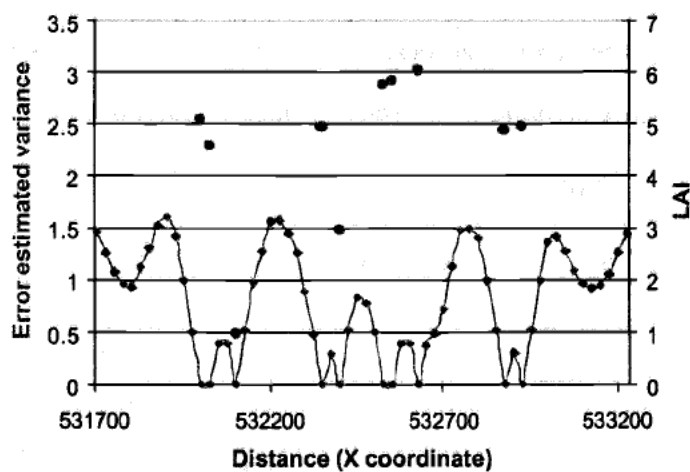
(a)

SGCS local uncertainty

(b)

Because cross validation RMSEs were so similar, they were difficult to evaluate. A disadvantage of the cross validation method is that the RMSE is a global measure, not specific to any particular location, and cannot indicate areas that may be more likely in error (Kyriakidis et al., 1999). A better assessment of the methods could be done with a further analysis of the cross validation predictions and cross validation errors in terms of their summary statistics, distribution, and spatial autocorrelation.

Besides the satisfactory statistical reproduction and improved pattern representation provided by conditional simulation, this technique offers other advantages: a spatial uncertainty measure allows one to identify areas with potentially higher errors, to improve future sampling efforts where uncertainty is high, and to assess the quality of the map. The areas that showed highest uncertainty were not sampled at the field, (i.e. near the center of the map on Fig. 45 a). An example of the usefulness of the uncertainty measure was shown by changing the secondary variable CI_{cover} to CI_{LAI} in the conditional simulation run. The use of CI_{LAI} as a covariate in the conditional simulation process resulted in a decline of the uncertainty of the map. A possible explanation for this decline in the uncertainty when using CI_{LAI} is that the Markov approximation rescales the primary variable with the correlation coefficient of the primary and secondary variables. Because CI_{LAI} has a slight stronger correlation with LAI than CI_{cover} (-0.73 vs. -0.70, see Table 4), this effect

may have had a positive effect on the estimates. Therefore, this uncertainty measure would also help to identify the best covariate in the conditional simulation process. Some applications may need to identify areas of interest, and conditional simulated maps may be post-processed to obtain probability maps of the classes of interest (Fig. 45 b, c, and d).

The expected LAI map resulting from the average of the 51 realizations resulted in very similar pattern compared to the CI_{LAI} map (Fig. 41 d) and reproduced the anisotropy, although at lower semivariance values, because of the averaging (Fig. 43 d).

Figure 45. 2-D maps of selected post-processing outputs: a.) conditional variance of 51 realizations, b.) probability map of LAI being grater than 2, c.) grater than 4, and d.) greater than 6.

a) Conditional Variance

b) Probability map LAI > 2

c) Probability map LAI > 4

d) Probability map LAI > 6

The set of simulations may also serve as input for sensitivity analysis in two ways: to assess how LAI variability affects process models, and to assess how the variability influences the upscaling from Landsat ETM+ (25-30 m) to MODIS (250-1000 m) spatial resolutions.

Conditional simulation may be promising for upscaling for other reasons as well. Aggregation usually decreases the variance; if smoothed maps were used for upscaling, the final variability would not be representative of the actual patterns. The variance and semivariance of the variable of interest should be determined at the different scales studied, to preserve the ecological significance and validity of the maps, because the variability observed in the system is conditional to the scale of description. If maps were used at MODIS scales, conditional simulation would be suggested. MODIS-scale grids are being developed to capture global trends and will be used in highly generalized models. Conditional simulation provides global accuracy and a better representation of the landscape patterns, essential properties for the mentioned applications. Another advantage of conditional simulation is that it allows for the generation of maps with associated uncertainty.

Spectral vegetation indices such as NDVI and SR may not be useful for characterizing complex canopy structure such as in boreal forests, which has variable canopy closure, clumped canopy and highly reflective understory. The use of mid-IR bands helped to explain more of the LAI variability. When analyzed

individually, bands 5 and 7 were the best-correlated with LAI, and they contributed substantially to the canonical indices (Table 5). Nemani et al., (1993) used a mid-IR corrected NDVI that responded to canopy closure for correcting for understory and other background materials (mid-IR decreased as canopy closure increased and as water content of the understory increased). The bryophyte species covering the soil in our boreal forest have water holding capacities of 16 to 26 times their own dry weight (Vogelman and Moss, 1993), which may have lowered the mid-IR signal.

Cover was a significant variable in this study because of the particular canopy structure of boreal forests, its variable canopy closure, and the poor NIR-LAI relationship in open canopies. It was useful to help explain LAI spatial and aspatial variability and served as a good covariate in the multivariate geostatistical models. The boreal forest canopy and understory, especially the soil cover components, have dissimilar spectral properties. For these reasons, image spectral mixture models (Peddle et al., 1999) for estimating subpixel fractions of overstory and understory cover components might be explored to derive more accurate LAI maps.

Acknowledgements

References

Adobe Photoshop Version 5.0.2, 1998, Adobe Systems Incorporated.

Bonan, G. 1993, Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. Remote Sensing of Environment, 43, 303-314.

Bonnor, G. 1967, Estimation of ground canopy density from ground measurements. J. For. 65, 544-547.

Bubier, J., Barret, N., and Crill, P. 1997, Spectral reflectance measurements of boreal wetland and forest mosses. Journal of Geophysical Research, Vol. 102, No. D24, 29,483-29,494.

Bunnell, F., and Vales, D. 1990, Comparison of methods for estimating forest overstory cover: differences among techniques. Can. J. For. Res. 20, 101-107.

Burrows, S., Gower, S., Clayton, M., Mackay, D., Ahl, D., Norman, J., and Diak, G. 2001, Application of Geostatistics to characterize LAI from flux tower to landscape scales using a cyclic sampling design. Submitted.

Campbell, J., Burrows, S., Gower, S., and Cohen, W. 1999, BigFoot: Characterizing land cover, LAI, and NPP at the Landscape Scale for EOS/MODIS. http://www.fsl.orst.edu/larse/bigfoot/pubs_rpts.html

Chavez, P. 1996, Image-based atmospheric corrections – Revisited and improved. Photogrammetric Engineering & Remote sensing, Vol. 62, No. 9, 1025-1036.

Chen, J., and Cihlar, J. 1996, Retrieving LAI of boreal conifer forests using Landsat TM images. Remote Sensing of Environment, 55, 153-162.

Clinger, W., and Van Ness, J. 1976, On equally spaced time points in time series. The Annals of Statistics, Vol. 4, 736-745.

Cohen, W., Spies, T., and Bradshaw, G. 1990, Semivariograms of digital imagery for analysis of conifer canopy structure. Remote Sensing of Environment, 35, 167-178.

Cohen, W., and Justice, C. 1999, Validating MODIS terrestrial ecology products: Linking in situ and satellite measurements. Remote Sensing of Environment, 70, 1-3.

Cohen, W.B., Thornton, P., and Maiersperger, T. 1999, BigFoot Pilot Study: Test of the Vegetation Cover Component Characterization System (3CS). http://www.fsl.orst.edu/larse/bigfoot/pubs_rpts.html

Curran, P., and Hay, A. 1986, The importance of measurement error for certain procedures in remote sensing at optical wavelengths. Photogrammetric Engineering and Remote Sensing, Vol. 52, No.2, 229-241.

Davis, J. 1986, Statistics and data analysis in geology. John Wiley & Sons, Inc., New York, 2nd ed.

Deutsch, C. 2000, Geostatistical reservoir modeling. Draft to be published by Oxford University Press.

Deutsch, C., and Journel, A. 1998, GSLIB: Geostatistical software library and user's guide. Oxford University Press, New York, 2nd ed.

Dungan, J. 1999, Conditional simulation: An alternative to estimation for achieving mapping objectives. In: Spatial statistics for remote sensing, edited by A. Stein and F. Van der Meer, Kluwer Academic Publishers.

Fassnacht, K., Gower, S., MacKenzie, M., Nordheim, E., and Lillesand, T. 1997, Estimating the LAI of North Central Wisconsin forests using the Landsat Thematic Mapper. Remote Sensing of Environment, 61, 229-245.

Fortin, M., Drapeau, P., and Legendre, P. 1989, Spatial autocorrelation and sampling design in plant ecology.

Fournier, R., Rich, P., Landry, R. 1997, Hierarchical characterization of canopy architecture for boreal forest. Journal of Geophysical Research, Vol. 102, No. D24, 29,445-29,454.

Gates, D., Keegan, H., Schleter, J., and Weidner, V. 1965, Spectral properties of plants. Applied Optics, Vol.4, No.1, 11-19.

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

Goovaerts, P. 2000, Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228, 113-129.

Gower, S., Kucharik, C., and Norman, J. 1999, Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote Sensing of Environment, 70, 29-51.

Grifith, D. 1987, Spatial Autocorrelation. Resource publications in geography.

Helms, J. 1998, The dictionary of Forestry. Society of American Foresters, John Helms, editor.

Heuvelink, G. 1998, Error propagation in environmental modeling with GIS. Taylor & Francis. London.

Hudak, A., and Wessman, C. 1998, Textural analysis of historical aerial photography to characterize woody plant encroachment in south african savanna. Remote Sensing of Environment, 66, 317-330.

Huete, A., Jackson, R., and Post, D. 1985, Spectral response of a plant canopy with different soil backgrounds. Remote Sensing of Environment, 17, 37-53.

IDL Version 5.4. 2000, Research Systems, Inc.

Isaaks, E., and Srivastava, R. 1989, An introduction to applied geostatistics. Oxford University Press, New York

Johnson, D. 1998, Applied multivariate methods for data analysis. , 1st ed., Duxbury Press.

Journel, A., and Huijbregts, C. 1978, Mining geostatistics. Academic Press, New York.

Kyriakidis, P., Shortridge, A., and Goodchild, M. 1999, Geostatistics for conflation and accuracy assessment of digital elevation models. Int. J. Geographical Information Science, Vol.13, No.7, 677-707.

Landsberg, J., and Gower, S. 1997, Applications of physiological ecology to forest management. Academic Press, Inc., San Diego, California.

Larsen, J. 1982, The northern forest border in Canada and Alaska: biotic communities and ecological relationships, ecological studies 70. Springer-Verlag, New York.Larsen, J. 1982, The northern forest border in Canada and Alaska: biotic communities and ecological relationships, ecological studies 70. Springer-Verlag, New York.

Lefsky, M., Cohen, W., Acker, S., Parker, G., Spies, T., and Harding, D. 1999, Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environment, 70, 339-361.

Legendre, P., and Fortin, M. 1989, Spatial pattern and ecological analysis. Vegetatio, 80, 107-138.

Legendre, P. 1993, Spatial autocorrelation: Trouble or new paradigm? Ecology, 74 (6), 1659-1673.

Lillesand, T., and Kiefer, R. 1999, Remote Sensing and image interpretation. 4th ed. John Wiley & Sons, Inc. New York.

Loechel, S., Walthall, C., Brown de Colstoun, E., Chen, J., Markham, B., and Miller, J. 1997, Variability of boreal forest reflectances as measured from a helicopter platform. Journal of Geophysical Research, Vol. 102, No.D24, 29,495-29,503.

Metzger, K. 1997, Modeling forest and structure to a ten meter resolution using Landsat TM data. MsSc. Dissertation, Colorado State University, Fort Collins, Colorado.

Miller, R., and Kahn, J. 1962, Statistical analysis in the geological sciences. John Wiley & Sons, Inc. New York.

Miller, J., White, H.,Chen, J.,Peddle, D., McDermid, G, Fournier, R.,Shepherd, P.,Rubinstein, I.,Freemantle, J.,Soffer, R., and LeDrew, E. 1997, Seasonal change in understory reflectance of boreal forests and influence on canopy vegetation indices. Journal of Geophysical Research, Vol.102, No.D24, 29,475-29,482.

Milne, B., and Cohen, W. 1999, Multiscale assessment of binary and continuous landcover variables for MODIS validation, mapping, and modeling applications. Remote Sensing of Environment, 70, 82-98.

Mueller-Dombois, D., and Ellenberg, H. 1974, Aims and methods of vegetation ecology. John Wiley & Sons, Inc. New York.

Myers, J. 1997, Geostatistical Error Management: Quantifying uncertainty for environmental sampling and mapping. Van Nostrand Reinhold.

Nemani, R., Pierce, L., Running, S., and Band, L. 1993, Forest Ecosystem processes at the watershed scale: sensitivity to remotely sensed Leaf Area Index estimates. Int. J. Remote Sensing, Vol. 14, No. 13, 2519-2534.

Ni, W., Li, X., Woodcook, C., Roujean, J., Davis, R. 1997, Transmission of solar radiation in boreal conifer forests: Measurements and models. Journal of Geophysical Research, Vol. 102, No. D24, 29,555-29,566.

Peddle, D., Hall, F., and LeDrew, E. 1999, Spectral mixture analysis and geometric-optical reflectance modeling of boreal forest biophysical structure. Remote Sensing of Environment, 67, 288-297.

Peterson, D., Spanner, M., Running, S., and Teuber, K. 1987, Relationship of Thematic Mapper Simulator data to leaf area index of temperate coniferous forests. Remote Sensing of Environment, 22, 323-341.

Petzold, D., and Goward, S. 1988, Reflectance spectra of subartic lichens. Remoste Sensing of Environment, 24, 481-492.

Ramsey, F., and Shafer, D. 1997, The statistical sleuth: a course in methods of data analysis. 1st ed., Duxbury Press.

Reich, M. 1999, Spatial statistical modeling of ecosystem respources and the environment. Department of Forest Sciences, Colorado State University.

Rossi, R.,Mulla, D., Journel, A., and Franz, E. 1992, Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 62(2), 277-314.

Rossi, R., Borth, P., and Tollefson, J. 1993, Stochastic simulation for characterizing ecological spatial patterns and appraising risk. Ecological Applications, 3(4), 719-735.

Running, S., and Nemani, R. 1987, Extrapolation of synoptic meteorological data in mountainous terrain and its use in simulating forest evapotranspiration rate and photosynthesis. Canadian Journal of Forest Research, Vol.17, 472-483.

Running, S., and Gower, S. 1991, FOREST-BGC, a general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets. Tree Physiology 9, 147-160.

Saito, H., and Goovaerts, P. 2000, Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site. Environmental Science Technology, 34, 4224-4235.

SAS/STAT®, Version 6, SAS Institute INC.

Spanner, M., Pierce, L., Peterson, D., and Running, S. 1990 (a), Remote sensing of temperate coniferous forest leaf area index: The influence of canopy closure, understory vegetation and background reflectance. Int. J. of Remote Sensing, 11, 95-111.

Spanner, M., Pierce, L., Running, S., and Peterson, D. 1990 (b), The seasonality of AVHRR data of temperate coniferous forests: relationships with LAI. Remote Sensing of Environment, 33, 97-112.

S-PLUS Version 2000 MathSoft, Inc.

St-Onge, B., and Cavayas, F. 1997, Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates. Remote Sensing of Environment, 61, 82-95.

Thompson, S. 1992, Sampling. John Wiley & Sons, INC.

Tucker, C. 1979, Red and infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8, 127-150.

Turner, D., Cohen, W., Kennedy, R., Fassnacht, K., anf Briggs, J. 1999, Relationships between LAI and Landsat TM spectral vegetation indices across three temperate zone sites. Remote Sensing of Environment, 70, 52-68.

Turner, S., O'Neill, R. Conley, W., Conley, M., and Humphries, H. 1990, Pattern and scale: Statistics for landscape ecology. In: Turner, M. and Garchner, R. 1990, Quantitative methods in landscape ecology. Springer Verlag, New York.

Vales, D., and Bunnell, F. 1988, Comparison of methods for estimating forest overstory cover.I. Observer effects. Can. J. For. Res. 18, 606-609.

Van Cleve, K., and Viereck, L. 1981, Forest succession in relation to nutrient cycling in the boreal forest of Alaska, in Forest Successions: Concepts and application (D.C. West, H. Shugart, and D. Botkin, Eds.), Springer-Verlag, New York.

Vogelmann, J. and Moss, D. 1993, Spectral reflectance measurements in the genus Sphagnum. Remote Sensing of Environment, 45,273-279.

Wackernagel, H. 1998, Multivariate geostatistics: An introduction with applications. Springer-Verlag Berlin Heidelberg New York, 2nd ed.

Waring, R., and Running, S. 1998, Forest Ecosystems: Analysis at multiple scales. Academic Press, San Diego, 2nd ed.

WINGSLIB, Version 1.3, Deutsch, C. and Journel, A. (http://www.statios.com)

201

Woodcock, C., Strahler, A., and Jupp, D. 1988, The use of variograms in remote sensing: I. Scene models and simulated images. Remote Sensing of Environment, 25, 323-343.

Chapter Three: Discussion and Conclusions

The results obtained in this study suggest that when mapping natural resource attributes such as LAI, the optimal technique to select depends on the final uses of the map, because not all the desired map characteristics can be achieved simultaneously.

The location of the ground measurements has two important aspects: accuracy in the match between ground and image data (georectification), and representativeness of the ground data for the area to be estimated. Representativeness, defined by Myers (1997) as "the degree to which sample data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, or an environmental condition", is of crucial importance in some geostatistical methods because attributes not sampled will not be present in the final outputs. A good example of this is the hardwood area in the SW corner of the image, which was not sampled. This area was reproduced in the output maps by regression methods, although it constitute an extrapolation situation. The same area was ignored by kriging and cokriging maps, but kriging with an external drift and conditional simulation accounted for it because of their use of the auxiliary variables. An unsupervised classification and/or a feature space analysis are recommended as first steps in evaluating representativeness of the sampling design.

The sampling design plays a crucial role in the process of producing environmentally-derived maps. The spatial distribution of the ground measurements may have great influence in the final results. The systematic spatial cluster sampling design used was appropriate for this study because of its efficiency in diminishing redundant data collection, while distributing pairs of points at different lags. However, orientation of the sampling units was not optimal for our purposes because it did not correspond to the maximum direction of continuity (Figs. 4 and 25). A basic a-priori anisotropy analysis of the image could have avoided this inconvenience. Nevertheless, the available samples and the anisotropy analysis allowed the identification of different spatial ranges that may be a result of the glacial drift occurring in the area.

The conditional bias observed in the scatterplots of cross validation errors versus estimated LAI may be a result of the smoothing in the case of the geostatistical methods, but this may also be a consequence of a smaller number of samples of lower and higher LAI values.

Measurement errors are intrinsic in both ground and image measurements. Traditional regression and inverse regression take into account measurement errors only in the response or the explanatory variables respectively, assuming that the other variable is free of error. RMA is a useful improvement for many applications, including this one, because it takes into account error measurements in both the

response and explanatory variables. The results of this study agreed with Curran and Hay. (1986), who found the same effect when the regression slope was underestimated under the assumption of measurement free of errors for the explanatory variable. The total error of the maps was not thoroughly calculated, but includes the errors generated in the processing of the images, the misregistration error, the measurement errors of the ground variables, the interpretation, estimation, and simulation errors. Heuvelink, (1998) stated that there is not still a single, accepted method for handling these kinds of errors and their propagation in GIS.

Measurement support, defined as the shape, size and orientation of the area or volume being measured (Curran, 1999), is significant in determining at which scale the analysis is valid. This study tried to match the ground-based measurements to Landsat ETM+ pixel size. In both cases, averages over the support surface are used, in which case the variograms are called "regularized". The effects of regularization are the same as those for aggregation: the variance is reduced, the range increases, and the mean remains unchanged (Isaaks and Srivastava, 1989; Woodcock et al., 1988).

The variance reduction may be more important if there is a large nugget effect, as was observed when going from the RMA to the integrated model. One possible explanation for this reduction is that originally, LAI had low spatial

autocorrelation, and the RMA model accounted for most of its variability. The residuals of the RMA model had a high nugget effect and very low spatial autocorrelation, and although Moran's I coefficient was significant (at a 0.05 significant level), there was little spatial autocorrelation to be further explained. The high nugget effect suggested that there was unexplained variability at scales smaller than the support.

Statistical properties of spatial attributes strongly depend on the support of the observations (Heuvelink, 1998). The description of pattern implies the description of variation and variation quantification requires the determination of scales (Levin, 1992), and supports.

From the results of this study, processes concerning LAI at this particular forest are relevant at scales smaller than 150 m, so the scaling up to MODIS resolutions (250 to 1000 m) and the use of layers of information at these scales to inform process models would imply certain amount of unknown error that should be quantified to assess the relevance of the final NPP maps.

Spatial autocorrelation is present in natural environments at all spatial scales (Legendre, 1993). This is reflected in the patterns observed in most ecological phenomena. If spatial autocorrelation is present, one can decide to remove it so that traditional statistical procedures can be used. A preferable alternative is to take it into account by using other statistical methods such as geostatistics. Although in this study LAI exhibited low spatial autocorrelation, the geostatistical techniques

exploited it and improved some features of the maps, thus showing that in some cases it is worth the extra modeling effort.

Spatial variability is generated by the complexity of the processes that reflect ecological functions, which we intend to interpret. LAI has important ecological meaning because it influences the microclimate of the forest, and regulates ecophysiological processes that determine the exchange of heat, water, nutrients, carbon, and other elements.

Table 19 shows a summary of the spatial variability features obtained in the LAI maps identified as trend, pattern and anisotropy, and explains how each method accounted for them. In general, regression methods reproduced pattern and anisotropy well but were not good at reproducing trends. A possible explanation for this is that the regression coefficients were calculated over the entire area. When search neighborhoods were applied in the geostatistical methods, the trends became more apparent. Kriging and cokriging resulted in smoothed patterns, and a poor anisotropy reproduction that may be a consequence of the minimization of the variance. Kriging with an external drift improved the pattern reproduction but had less predictive ability than the integrated method. Although the integrated model had a reduced semivariance over RMA, it performed best among the regression and geostatistical estimation methods if a compromise between global and local accuracy were the objective.

Conditional simulation reproduced trend, pattern and anisotropy well. Conditional

simulation used the secondary data in a different way than cokriging and kriging

with an external drift. While cokriging uses the full secondary data through the

cokriging system, kriging with an external drift uses them only to characterize the

trend of the primary variable, and conditional simulation uses only the colocated

data scaled by the correlation coefficient between the primary and the secondary

variables. KED maps reproduced a better pattern inside as well as outside the

sample grid (Figs. 36 and 39), not all the features were reproduced in the

extrapolation situation. This is more evident in the northern area. Although the

anisotropic LAI model used in KED and in OK was the same, the semivariograms

of the whole area showed a better anisotropy reproduction for KED, suggesting that

the secondary information provided by CI_{cover} was useful in accounting for LAI

spatial variability. Besides, the use of this secondary information seemed to be

more appropriate in the KED case, compared to OCK. This may be explained by

the way OCK incorporated the full secondary data, some of which may be

redundant. This is confirmed by the colocated use of the secondary data by

conditional simulation, which yielded the best pattern representation of all. Ideally,

LAI maps used as inputs in process models should represent actual LAI spatial

variability to attain more realistic results.

The methods used in this study suggest that high levels of local and global accuracy

cannot be achieved simultaneously. Kriging techniques aim to provide estimators

as close as possible to the true unknown values, with the criteria of unbiasedness and minimal estimation variance. The minimization of the estimation variance involves a smoothing of the true distribution of the original variable, resulting in poor reproduction of spatial variability. On the other hand, conditional simulation values are not the best possible estimators on a point by point basis, but the variance of estimation is greater than the kriging variance, resulting in a better reproduction in the pattern of the modeled variable (Journel, 1978; Dungan, 1999). If local accuracy were the objective, and RMA the desired method, it would be preferable to stratify first by land cover and apply different models for the different cover types, as suggested by Turner et al. (1999).

Estimation methods that take a deterministic approach provide a single estimate map with aspatial (i.e. standard error in regression techniques) or poor (i.e. kriging variance) assessment of errors. In contrast, simulation methods take a probabilistic approach, where simulated values are conditional to the samples, honor their statistics, and provide a spatially quantitative measure of uncertainty.

A good spatial error map should depend on the values of the samples, their distances and geometry. Fig. 44 shows a E-W transect that goes through 10 samples. The kriging error variance shows the variance is dependent on the data configuration only; the variance decreasing when approaching the sample locations, and increasing when going away from them. In contrast, the conditional variance depends on the data configuration and the data values. The variance was greater

when two consecutive samples had dissimilar LAI values, and smaller when consecutive values were more alike.

The conditional variance of the whole set of realizations provided a visual and quantitative spatial uncertainty measure (Fig.45 a).

Because cross validation RMSEs were so similar, they were difficult to evaluate. A disadvantage of the cross validation method is that the RMSE is a global measure, not specific to any particular location, and cannot indicate areas that may be more likely in error (Kyriakidis et al., 1999). A better assessment of the methods could be done with a further analysis of the cross validation predictions and cross validation errors in terms of their summary statistics, distribution, and spatial autocorrelation.

Besides the satisfactory statistical reproduction and improved pattern representation provided by conditional simulation, this technique offers other advantages: a spatial uncertainty measure allows one to identify areas with potentially higher errors, to improve future sampling efforts where uncertainty is high, and to assess the quality of the map. The areas that showed highest uncertainty were not sampled at the field, (i.e. near the center of the map on Fig. 45 a). An example of the usefulness of the uncertainty measure was shown by changing the secondary variable CI_{cover} to CI_{LAI} in the conditional simulation run. The use of CI_{LAI} as a covariate in the conditional simulation process resulted in a decline of the uncertainty of the map. A

possible explanation for this decline in the uncertainty when using CI_{LAI} is that the Markov approximation rescales the primary variable with the correlation coefficient of the primary and secondary variables. Because CI_{LAI} has a slight stronger correlation with LAI than CI_{cover} (-0.73 vs. -0.70, see Table 4), this effect may have had a positive effect on the estimates. Therefore, this uncertainty measure would also help to identify the best covariate in the conditional simulation process. Some applications may need to identify areas of interest, and conditional simulated maps may be post-processed to obtain probability maps of the classes of interest (Fig. 45 b, c, and d).

The expected LAI map resulting from the average of the 51 realizations resulted in very similar pattern compared to the CI_{LAI} map (Fig. 41 d) and reproduced the anisotropy, although at lower semivariance values, because of the averaging (Fig. 43).

The set of simulations may also serve as input for sensitivity analysis in two ways: to assess how LAI variability affects process models, and to assess how the variability influences the upscaling from Landsat ETM+ (25-30 m) to MODIS (250-1000 m) spatial resolutions.

Conditional simulation may be promising for upscaling for other reasons as well. Aggregation usually decreases the variance; if smoothed maps were used for

upscaling, the final variability would not be representative of the actual patterns. The variance and semivariance of the variable of interest should be determined at the different scales studied, to preserve the ecological significance and validity of the maps, because the variability observed in the system is conditional to the scale of description. If maps were used at MODIS scales, conditional simulation would be suggested. MODIS-scale grids are being developed to capture global trends and will be used in highly generalized models. Conditional simulation provides global accuracy and a better representation of the landscape patterns, essential properties for the mentioned applications. Another advantage of conditional simulation is that it allows for the generation of maps with associated uncertainty.

Spectral vegetation indices such as NDVI and SR may not be useful for characterizing complex canopy structure such as in boreal forests, which has variable canopy closure, clumped canopy and highly reflective understory. The use of mid-IR bands helped to explain more of the LAI variability. When analyzed individually, bands 5 and 7 were the best-correlated with LAI, and they contributed substantially to the canonical indices (Table 5). Nemani et al., (1993) used a mid-IR corrected NDVI that responded to canopy closure for correcting for understory and other background materials (mid-IR decreased as canopy closure increased and as water content of the understory increased). The bryophyte species covering the soil in our boreal forest have water holding capacities of 16 to 26 times their own

dry weight (Vogelman and Moss, 1993), which may have lowered the mid-IR signal.

Cover was a significant variable in this study because of the particular canopy structure of boreal forests, its variable canopy closure, and the poor NIR-LAI relationship in open canopies. It was useful to help explain LAI spatial and aspatial variability and served as a good covariate in the multivariate geostatistical models. The boreal forest canopy and understory, especially the soil cover components, have dissimilar spectral properties. For this reason, image spectral mixture models (Peddle et al., 1999) for estimating subpixel fractions of overstory and understory cover components might be explored to derive more accurate LAI maps.

Bibliography

Adobe Photoshop Version 5.0.2, 1998, Adobe Systems Incorporated.

Bonan, G. 1993, Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. Remote Sensing of Environment, 43, 303-314.

Bonnor, G. 1967, Estimation of ground canopy density from ground measurements. J. For. 65, 544-547.

Bubier, J., Barret, N., and Crill, P. 1997, Spectral reflectance measurements of boreal wetland and forest mosses. Journal of Geophysical Research, Vol. 102, No. D24, 29,483-29,494.

Bunnell, F., and Vales, D. 1990, Comparison of methods for estimating forest overstory cover: differences among techniques. Can. J. For. Res. 20, 101-107.

Burrough, P. 1995, Spatial aspects of ecological data. In Data analysis in community and landscape ecology, edited by Jongman, R., Ter Braak, C., and Van Tongeren, O. New ed. Cambridge University Press 1995.

Burrows, S., Gower, S., Clayton, M., Mackay, D., Ahl, D., Norman, J., and Diak, G. 2001, Application of Geostatistics to characterize LAI from flux tower to landscape scales using a cyclic sampling design. Submitted.

Campbell, J., Burrows, S., Gower, S., and Cohen, W. 1999, BigFoot: Characterizing land cover, LAI, and NPP at the Landscape Scale for EOS/MODIS. http://www.fsl.orst.edu/larse/bigfoot/pubs_rpts.html

Chavez, P. 1996, Image-based atmospheric corrections – Revisited and improved. Photogrammetric Engineering & Remote sensing, Vol. 62, No. 9, 1025-1036.

Chen, J., and Cihlar, J. 1996, Retrieving LAI of boreal conifer forests using Landsat TM images. Remote Sensing of Environment, 55, 153-162.

Clinger, W., and Van Ness, J. 1976, On equally spaced time points in time series. The Annals of Statistics, Vol. 4, 736-745.

Cohen, W., Spies, T., and Bradshaw, G. 1990, Semivariograms of digital imagery for analysis of conifer canopy structure. Remote Sensing of Environment, 35, 167-178.

Cohen, W., and Justice, C. 1999, Validating MODIS terrestrial ecology products: Linking in situ and satellite measurements. Remote Sensing of Environment, 70, 1-3.

Cohen, W.B., Thornton, P., and Maiersperger, T. 1999, BigFoot Pilot Study: Test of the Vegetation Cover Component Characterization System (3CS). http://www.fsl.orst.edu/larse/bigfoot/pubs_rpts.html

Curran, P., and Hay, A. 1986, The importance of measurement error for certain procedures in remote sensing at optical wavelengths. Photogrammetric Engineering and Remote Sensing, Vol. 52, No.2, 229-241.

Curran, P. 1988, The semivariogram in remote sensing: An introduction. Remote Sensing of Environment, 24, 493-507.

Davis, J. 1986, Statistics and data analysis in geology. John Wiley & Sons, Inc., New York, 2nd ed.

Deutsch, C. 2000, Geostatistical reservoir modeling. Draft to be published by Oxford University Press.

Deutsch, C., and Journel, A. 1998, GSLIB: Geostatistical software library and user's guide. Oxford University Press, New York, 2nd ed.

Dungan, J. 1999, Conditional simulation: An alternative to estimation for achieving mapping objectives. In: Spatial statistics for remote sensing, edited by A. Stein and F. Van der Meer, Kluwer Academic Publishers.

Fassnacht, K., Gower, S., MacKenzie, M., Nordheim, E., and Lillesand, T. 1997, Estimating the LAI of North Central Wisconsin forests using the Landsat Thematic Mapper. Remote Sensing of Environment, 61, 229-245.

Fortin, M., Drapeau, P., and Legendre, P. 1989, Spatial autocorrelation and sampling design in plant ecology.

Fournier, R., Rich, P., Landry, R. 1997, Hierarchical characterization of canopy architecture for boreal forest. Journal of Geophysical Research, Vol. 102, No. D24, 29,445-29,454.

Gates, D., Keegan, H., Schleter, J., and Weidner, V. 1965, Spectral properties of plants. Applied Optics, Vol.4, No.1, 11-19.

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

Goovaerts, P. 2000, Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228, 113-129.

Gower, S., Kucharik, C., and Norman, J. 1999, Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote Sensing of Environment, 70, 29-51.

Grifith, D. 1987, Spatial Autocorrelation. Resource publications in geography.

Helms, J. 1998, The dictionary of Forestry. Society of American Foresters, John Helms, editor.

Heuvelink, G. 1998, Error propagation in environmental modeling with GIS. Taylor & Francis. London.

Hudak, A., and Wessman, C. 1998, Textural analysis of historical aerial photography to characterize woody plant encroachment in south african savanna. Remote Sensing of Environment, 66, 317-330.

Huete, A., Jackson, R., and Post, D. 1985, Spectral response of a plant canopy with different soil backgrounds. Remote Sensing of Environment, 17, 37-53.

IDL Version 5.4. 2000, Research Systems, Inc.

Isaaks, E., and Srivastava, R. 1989, An introduction to applied geostatistics. Oxford University Press, New York

Johnson, D. 1998, Applied multivariate methods for data analysis. , 1st ed., Duxbury Press.

Journel, A., and Huijbregts, C. 1978, Mining geostatistics. Academic Press, New York.

Journel, A. 1989, Fundamentals of Geostatistics in five lessons. American Geophysical Union Washington, D.C.

Kyriakidis, P., Shortridge, A., and Goodchild, M. 1999, Geostatistics for conflation and accuracy assessment of digital elevation models. Int. J. Geographical Information Science, Vol.13, No.7, 677-707.

Landsberg, J., and Gower, S. 1997, Applications of physiological ecology to forest management. Academic Press, Inc., San Diego, California.

Larsen, J. 1982, The northern forest border in Canada and Alaska: biotic communities and ecological relationships, ecological studies 70. Springer-Verlag, New York.Larsen, J. 1982, The northern forest border in Canada and Alaska: biotic communities and ecological relationships, ecological studies 70. Springer-Verlag, New York.

Lefsky, M., Cohen, W., Acker, S., Parker, G., Spies, T., and Harding, D. 1999, Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environment, 70, 339-361.

Legendre, P., and Fortin, M. 1989, Spatial pattern and ecological analysis. Vegetatio, 80, 107-138.

Legendre, P. 1993, Spatial autocorrelation: Trouble or new paradigm? Ecology, 74 (6), 1659-1673.

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

Lillesand, T., and Kiefer, R. 1999, Remote Sensing and image interpretation. 4th ed. John Wiley & Sons, Inc. New York.

Loechel, S., Walthall, C., Brown de Colstoun, E., Chen, J., Markham, B., and Miller, J. 1997, Variability of boreal forest reflectances as measured from a helicopter platform. Journal of Geophysical Research, Vol. 102, No.D24, 29,495-29,503.

Metzger, K. 1997, Modeling forest and structure to a ten meter resolution using Landsat TM data. MsSc. Dissertation, Colorado State University, Fort Collins, Colorado.

Miller, R., and Kahn, J. 1962, Statistical analysis in the geological sciences. John Wiley & Sons, Inc. New York.

Miller, J., White, H.,Chen, J.,Peddle, D., McDermid, G, Fournier, R.,Shepherd, P.,Rubinstein, I.,Freemantle, J.,Soffer, R., and LeDrew, E. 1997, Seasonal

change in understory reflectance of boreal forests and influence on canopy vegetation indices. Journal of Geophysical Research, Vol.102, No.D24, 29,475-29,482.

Milne, B., and Cohen, W. 1999, Multiscale assessment of binary and continuous landcover variables for MODIS validation, mapping, and modeling applications. Remote Sensing of Environment, 70, 82-98.

Mueller-Dombois, D., and Ellenberg, H. 1974, Aims and methods of vegetation ecology. John Wiley & Sons, Inc. New York.

Myers, J. 1997, Geostatistical Error Management: Quantifying uncertainty for environmental sampling and mapping. Van Nostrand Reinhold.

Nemani, R., Pierce, L., Running, S., and Band, L. 1993, Forest Ecosystem processes at the watershed scale: sensitivity to remotely sensed Leaf Area Index estimates. Int. J. Remote Sensing, Vol. 14, No. 13, 2519-2534.

Ni, W., Li, X., Woodcook, C., Roujean, J., Davis, R. 1997, Transmission of solar radiation in boreal conifer forests: Measurements and models. Journal of Geophysical Research, Vol. 102, No. D24, 29,555-29,566.

Peddle, D., Hall, F., and LeDrew, E. 1999, Spectral mixture analysis and geometric-optical reflectance modeling of boreal forest biophysical structure. Remote Sensing of Environment, 67, 288-297.

Peterson, D., Spanner, M., Running, S., and Teuber, K. 1987, Relationship of Thematic Mapper Simulator data to leaf area index of temperate coniferous forests. Remote Sensing of Environment, 22, 323-341.

Petzold, D., and Goward, S. 1988, Reflectance spectra of subartic lichens. Remoste Sensing of Environment, 24, 481-492.

Ramsey, F., and Shafer, D. 1997, The statistical sleuth: a course in methods of data analysis. 1st ed., Duxbury Press.

Reich, M. 1999, Spatial statistical modeling of ecosystem respources and the environment. Department of Forest Sciences, Colorado State University.

Rossi, R.,Mulla, D., Journel, A., and Franz, E. 1992, Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 62(2), 277-314.

Rossi, R., Borth, P., and Tollefson, J. 1993, Stochastic simulation for characterizing ecological spatial patterns and appraising risk. Ecological Applications, 3(4), 719-735.

Running, S., and Nemani, R. 1987, Extrapolation of synoptic meteorological data in mountainous terrain and its use in simulating forest evapotranspiration rate and photosynthesis. Canadian Journal of Forest Research, Vol.17, 472-483.

Running, S., and Gower, S. 1991, FOREST-BGC, a general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets. Tree Physiology 9, 147-160.

Saito, H., and Goovaerts, P. 2000, Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site. Environmental Science Technology, 34, 4224-4235.

SAS/STAT®, Version 6, SAS Institute INC.

Spanner, M., Pierce, L., Peterson, D., and Running, S. 1990 (a), Remote sensing of temperate coniferous forest leaf area index: The influence of canopy closure, understory vegetation and background reflectance. Int. J. of Remote Sensing, 11, 95-111.

Spanner, M., Pierce, L., Running, S., and Peterson, D. 1990 (b), The seasonality of AVHRR data of temperate coniferous forests: relationships with LAI. Remote Sensing of Environment, 33, 97-112.

S-PLUS Version 2000 MathSoft, Inc.

St-Onge, B., and Cavayas, F. 1997, Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates. Remote Sensing of Environment, 61, 82-95.

Thompson, S. 1992, Sampling. John Wiley & Sons, INC.

Tucker, C. 1979, Red and infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8, 127-150.

Turner, D., Cohen, W., Kennedy, R., Fassnacht, K., anf Briggs, J. 1999, Relationships between LAI and Landsat TM spectral vegetation indices across three temperate zone sites. Remote Sensing of Environment, 70, 52-68.

Turner, S., O'Neill, R. Conley, W., Conley, M., and Humphries, H. 1990, Pattern and scale: Statistics for landscape ecology. In: Turner, M. and Garchner, R. 1990, Quantitative methods in landscape ecology. Springer Verlag, New York.

Vales, D., and Bunnell, F. 1988, Comparison of methods for estimating forest overstory cover.I. Observer effects. Can. J. For. Res. 18, 606-609.

Van Cleve, K., and Viereck, L. 1981, Forest succession in relation to nutrient cycling in the boreal forest of Alaska, in Forest Successions: Concepts and application (D.C. West, H. Shugart, and D. Botkin, Eds.), Springer-Verlag, New York.

Vogelmann, J. and Moss, D. 1993, Spectral reflectance measurements in the genus Sphagnum. Remote Sensing of Environment, 45,273-279.

Wackernagel, H. 1998, Multivariate geostatistics: An introduction with applications. Springer-Verlag Berlin Heidelberg New York, 2nd ed.

Waring, R., and Running, S. 1998, Forest Ecosystems: Analysis at multiple scales. Academic Press, San Diego, 2nd ed.

WINGSLIB, Version 1.3, Deutsch, C. and Journel, A. (http://www.statios.com)

Woodcock, C., Strahler, A., and Jupp, D. 1988, The use of variograms in remote sensing: I. Scene models and simulated images. Remote Sensing of Environment, 25, 323-343.

Wulder, M., LeDrew, E.,Lavigne, M. 1998, Aerial image texture informationin the estimation of northern deciduous and mixed wood forest LAI. Remote Sensing of Environment, 64(1), 64-76.