University of Southampton

GEOSTATISTICAL PREDICTION OF VEGETATION AMOUNT USING GROUND AND REMOTELY SENSED DATA

by

Jennifer Lee Dungan.

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the University of Southampton

Department of Geography

November 1999.

DECLARATION

This work has not been previously accepted in substance for any degree and is not being concurrently submitted in candidature for any degree. This thesis is the result of my own investigations, expect where otherwise stated. Other sources are acknowledged by giving explicit references. A bibiliography is appended. I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed _____

Date _____

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE

GEOGRAPHY

Doctor of Philosophy

GEOSTATISTICAL PREDICTION OF VEGETATION AMOUNT USING GROUND AND REMOTELY SENSED DATA

by Jennifer Lee Dungan

Maps of vegetation amount are needed at many scales, from the field scale for the purpose of managing crop production to the global scale to understand biogeochemical cycles. Current methods of predicting vegetation amount use remote sensing to provide a spatially exhaustive data source.

Aspatial regression was recognized as the most commonly used statistical prediction method in a detailed survey of 51 studies from the peer-reviewed literature. Regression used in these studies was aspatial in that it did not incorporate spatial support, sample location or geometry, spatial variability or spatial dependence. Therefore, geostatistical methods, which assume spatial dependence, have an untapped potential to map vegetation amount using ancillary data from remote sensing images. Two geostatistical methods, cokriging and conditional simulation, were contrasted for the first time with aspatial regression in terms of their accuracy and uncertainty description for vegetation amount prediction.

For a synthetic data set constructed from imaging spectrometer data, aspatial regression was most accurate when ground and spectral variables were very closely related (r between the data exceeding .89). Cokriging was more accurate in all other situations. Conditional simulation, though not as accurate, was superior to the other two methods in reproducing the univariate and spatial characteristics of vegetation amount. The sample size was 300 and the sampling fraction was .3%. For a real data set from western Montana, USA, over 300 ground measurements of conifer canopy cover made in each of two years by the US Forest Service and collocated NDVI values from Landsat TM were used to predict canopy cover in a 97 square km^2 subarea where the sampling fraction was .03%. The nonlinear aspatial regression model between canopy cover and NDVI had statistically identical parameters in both years, but prediction intervals were very wide and accuracy was low at test points. Cokriged maps had much higher accuracy but were affected by the small sampling fraction and clumped distribution of ground measurements. Conditionally simulated realizations using collocated cokriging displayed the desirable aspects of cokriging at the same time as presenting plausible global and spatial distributions of canopy cover and were therefore considered preferable to the cokriged maps. AVHRR data from the same region confirmed the difficulties of relating coarse spatial resolution (large spatial support) data to ground measurements. A simple change-of-support model between TM and AVHRR explained the behavior of near infrared reflectance but did not explain NDVI, a nonlinear transform of reflectance.

Results of both the synthetic and real data sets showed that uncertainty described by aspatial regression was completely data-value dependent whereas from cokriging it was completely data-location dependent. The uncertainty description from conditional simulation was dependent on both value and location and therefore should be more useful in describing the geographical distribution of prediction uncertainty.

Contents

Abstract iii

Acknowledgements xiii

Acronyms and symbols xiv

1 Introduction 1

1.1 Understanding the Earth system 1

1.2 Measurements of vegetation amount 2

1.3 Predictions of vegetation amount 3

1.4 Remote sensing for spatial prediction 5

1.4.1 Theory-based methods 6

1.4.2 Empirical methods 6

1.5 Geostatistics for spatial prediction 7

1.6 Research questions in geostatistical prediction of vegetation amount using remotely sensed data 10

2 Vegetation amount prediction using optical remote sensing 12

2.1 Aspatial regression with vegetation and spectral variables 12

2.1.1 Purposes of vegetation-spectral relationship studies 22

2.1.2 Variables used 22

2.1.3 Image preprocessing steps 24

2.1.4 Sampling considerations 26

2.1.5 Methods of fitting regression models to the data 28

2.1.6 Form of the models 29

2.1.7 Evaluating the results of aspatial regression models 30

2.1.8 Summary 31

2.2 Geostatistical analyses of remotely sensed data 32

2.2.1 Recognition of spatial dependence 33

2.2.2 Uses of spatial dependence in geostatistical prediction 34

2.3 Conclusion 37

3 Models and methods for prediction 39

3.1 The random function model 39

3.1.1 Forms of random functions 40

3.1.2 Stationarity 42

3.1.3 Ergodicity 42

3.1.4 Second-order moments of random functions 43

3.2 Prediction methods 48

3.2.1 Aspatial Regression 52

3.2.2 Cokriging 56

3.2.3 Conditional Simulation 61

Sequential simulation 62

Probability field simulation 65

Adding ancillary information to a simulation 66

3.3 Relevant properties of optical remote sensing data 67

3.4 Conclusions 69

4 A synthetic analog to the remote sensing of vegetation amount 72

4.1 The synthetic data set 73

4.2 Methods 75

4.2.1 Sampling from the data set 75

4.2.2 Aspatial regression 78

4.2.3 Cokriging 80

4.2.4 Conditional simulation 81

4.3 Results 82

4.4 Discussion 90

4.5 Conclusions 91

5 Predicting canopy cover in western Montana using ground and image data 93

5.1 Methods 93

5.1.1 Study region 93

5.1.2 Ground measurements 96

5.1.3 Remotely sensed data 99

Landsat Thematic Mapper data 99
NOAA AVHRR data 106
5.1.4 Collocation of ground and remotely sensed data 106
5.1.5 Choice of subarea 108
5.1.6 Aspatial regression 110
5.1.7 Cokriging 114
5.1.8 Conditional simulation 114
5.2 Results and Discussion 115
5.2.1 Reflectance of conifer forests 115
5.2.2 Spatial statistics of primary and ancillary data 128
5.2.3 Prediction of canopy cover 134
Aspatial regression 134
Cokriging 137
Conditional simulation 138
5.3 Conclusions 142

6 Conclusions **145**

6.1 Summary 145
6.1.1 The implications of choosing geostatistical models 145
6.1.2 The error and uncertainty components of prediction methods . 146
6.1.3 Aspatial regression versus cokriging 147
6.1.4 Statistical relationships observed 147
6.1.5 The effect of spatial support 148
6.2 Recommendations 149
6.2.1 Theoretical considerations 149
6.2.2 Practical considerations 151
6.2.3 Future research 152

A Effect of image resampling on statistics **155**

References **157**

List of Tables

2.1 Fifty-one published studies relating ground measurements of vegetation amount variables to spectral variables from remotely sensed images. Acronyms and symbols can be found starting on page xiv. 15

2.1 (continued) 16

2.1 (continued) 17

2.1 (continued) 18

2.1 (continued) 19

2.1 (continued) 20

2.1 (continued) 21

3.1 Summary of the general properties of and characteristics taken into account by three prediction methods. A '√' indicates the method has the property or preserves the property in the spatial predictions. 69

4.1 Parameters of the regression models (β_0 and β_1) and correlation coefficients (r) between the collocated primary data and the ancillary data from bands 7, 9, 10, 12, 13, 15 and 22. 78

4.2 Parameters of the semivariogram models used for cokriging, where C_0 is the nugget variance, C_1 is the coefficient on the exponential model and C_2 is the coefficient on the spherical model. $\gamma_a(\mathbf{h})$ is the semivariogram model for the ancillary variable from the band listed in the left column and $\gamma_{pa}(\mathbf{h})$ is the cross-semivariogram for the ancillary variable with the primary variable. 81

5.1 Codes used for ranges of canopy cover values in ECODATA. 99

5.2 TM calibration coefficients in units of $W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$ applied in equation 5.1. 102

5.3 Implications from 6S for dark object characteristics. 105

5.4 Linear correlation coefficients between reflectances in TM wavebands for pixels at 1988 ECODATA plot locations (n=335). 112

5.5 Linear correlation coefficients between TM spectral variables and canopy cover at ECODATA plot locations. 121

5.6 Parameter estimates for aspatial regression model 5.9. 134

5.7 Global statistics of canopy cover (%) maps predicted by aspatial regression (Reg), cokriging (Cok) and simulation (Sim) alongside the statistics from the 335 1988 ECODATA plot measurements and the 344 1994 plot measurements. 137

List of Figures

3.1 Typical form of a semivariogram, showing the nugget, sill and range. . . . 47

3.2 Illustration of how the quantile algorithm uses values of u, which are uniformly distributed, to draw values of w, which are distributed according to the cdf plotted here. 66

4.1 (a) AVIRIS band 21 used as the true map (unitless) and (b) the locations of the 300 sample pixels from the true map (symbols do not represent the actual size of pixels, but are enlarged for visibility). 74

4.2 Statistics from random samples of different sizes from the true map: (a) mean, (b) standard deviation and (c) root-mean-square error between the semivariogram calculated from each sample and the true map's exhaustive semivariogram. 76

4.3 Omnidirectional semivariograms from random samples of different sizes (50, 100, 200, 300, 400 and 500) from the true map (–•–) and the exhaustive omnidirectional semivariogram from the true map (—). 77

4.4 Scatterplots of primary data versus collocated values from ancillary data from (a) band 7, (b) band 9, (c) band 10, (d) band 12, (e) band 13, (f) band 15 and (g) band 22. 79

4.5 Three of the 30 probability fields. The grey scale ranges from black (representing 0) to white (representing 1). 82

4.6 Semivariogram values from the uniform scores of the primary data (–•–) and exhaustive omnidirectional semivariograms from seven probability fields (—). 83

4.7 Maps resulting from the regression method applied using (a) band 7, (b) 12 and (c) 22; maps resulting from the cokriging method applied using (d) band 7, (e) 12 and (f) 22 and single realizations from probability field simulation using (g) band 7, (h) 12 and (i) 22. The color scale is unitless. . 84

4.8 The trend in root-mean-square-error (RMSE) among maps predicted with the three methods as a function of the correlation between the primary and ancillary information. The prediction resulting from ordinary kriging with 300 primary data only is plotted for reference. 86

4.9 Q-Q plots representing the correspondence between the true distribution of the primary variable and that of predicted maps using ancillary images from (a) band 7, (b) 12 and (c) 22. 87
4.10 Exhaustive omnidirectional semivariogram calculated from the true map (—) compared with those calculated with the predicted maps (- - -) using the seven ancillary images (bands 7, 9, 10, 12, 13, 15 and 22) and (a) regression, (b) cokriging and (c) a single probability field realization. . . . 88
4.11 Square root of prediction variance maps modeled from (a) regression (b) cokriging and (c) probability field simulation. The grey scale is unitless. . 89
5.1 The study region location within the state of Montana; inset shows major park and national forest boundaries and major lakes. The subarea outlined with dashed line is described in section 5.1.5. Source: Digital coverages from the Montana State Library, Helena, MT. 94
5.2 View of a typical watershed in the study region, showing variation in canopy cover governed by land use, aspect and elevation. 96
5.3 The number of ECODATA plots surveyed between 1986 and 1996. 97
5.4 (a) Histogram of the 1988 canopy cover data, (b) histogram of the 1994 canopy cover data and (c) quantile-quantile plot of the 1994 vs 1988 canopy cover data. 100
5.5 Locations of 1988 and 1994 ECODATA plots on a relief-shaded representation of a digital elevation model for the study region. 107
5.6 (a) Scatterplot of elevations for ECODATA plot locations versus those for corresponding DEM pixels, (b) Histogram of differences between plot elevations and DEM elevations and (c) scatterplot of aspects for ECODATA plot locations versus those for corresponding DEM pixels. The 1:1 line is shown in (a) and (c). 109
5.7 Subarea from TM false color composite (Bands 4, 5 and 3 in red, green and blue, respectively) from (a) 1988 with 30, 1988 ECODATA plot locations in open circles and b) 1993 with 30, 1994 ECODATA plot locations in open circles. 110

5.8 1993 TM image in figure 5.7 'draped' onto elevation model of the subarea. (a) Harrison Creek, (b) USFS landing strip, (c) Meadow Creek cabin, (d) Bunker Creek and (e) South Fork Flathead River. 111

5.9 Values of canopy cover (%) at ECODATA plot locations in (a) 1988 and (b) 1994. 111

5.10 Histogram of (a) TM band 3 reflectance at 1988 ECODATA plot locations, (b) TM band 3 reflectance at 1993 locations, (c) AVHRR band 1 reflectance at 1988 locations and (d) AVHRR band 1 reflectance at 1993 locations. . . 117

5.11 Histogram of (a) TM band 4 reflectance at 1988 ECODATA plot locations, (b) TM band 4 reflectance at 1993 locations, (c) AVHRR band 2 reflectance at 1988 locations and (d) AVHRR band 2 reflectance at 1993 locations. . . 118

5.12 Histogram of (a) TM NDVI at 1988 ECODATA plot locations, (b) TM NDVI at 1993 locations, (c) AVHRR NDVI at 1988 locations and (d) AVHRR NDVI at 1993 locations. 119

5.13 Relationship between canopy cover at the ECODATA plot locations and TM (a) reflectance in band 3 from 1988, (b) reflectance in band 4 from 1988, (c) reflectance in band 5 from 1988 and d) NDVI from 1988. 122

5.14 Relationship between canopy cover at the ECODATA plot locations and (a) reflectance in TM band 3 from 1993, (b) reflectance in TM band 4 from 1993, (c) band 5 from 1993 and (d) NDVI from 1993. 123

5.15 TM NDVI versus canopy cover from (a) 1988 and (b) 1993, where distribution of NDVI values for each value of canopy cover is summarized using a boxplot. The width of each box is proportional to the number of plots with the value of canopy cover on the abscissa. 124

5.16 Relationship between canopy cover at the ECODATA plot locations and AVHRR (a) reflectance in band 1 from 1988, (b) reflectance in band 2 AVHRR from 1988, (c) NDVI from 1988, (d) reflectance in band 1 from 1993, (e) reflectance in band 2 AVHRR from 1993 and (f) NDVI from 1993.125

5.17 (a) Reflectance in TM band 4 versus aspect and b) TM NDVI versus aspect.126

5.18 Canopy cover measurements versus NDVI at ECODATA plot locations in a range of aspect classes. The dark bar within the shaded panel above each scatterplot indicates the range of plot aspects, from low values (near 0°) in the lower left corner to high values (near 360°) in the upper right. . 127
5.19 Histogram of distances between plot locations from (a) 1988 and (b) 1994. 129
5.20 Omnidirectional semivariograms of canopy cover (%) from (a) 1988 and (b) 1994. Symbol size is proportional to the number of pairs at that distance. 130
5.21 Omnidirectional semivariograms of TM NDVI from (a) 1988 and (b) 1993. Symbol size is proportional to the number of pairs at that distance. 132
5.22 Omnidirectional semivariograms and models for (a) 1988 canopy cover (%), (b) 1994 canopy cover (%) and (c) 1988 NDVI and d) 1993 NDVI. Omnidirectional cross-semivariograms and models for (e) 1988 canopy cover and NDVI and (f) 1993 canopy cover and NDVI. 133
5.23 Nonlinear regression models from (a) 1988 and (b) 1993 data. 135
5.24 Predicted canopy cover (%) using regression from (a) 1988 data and (b) from 1993 data. ECODATA plot locations are shown as black squares. . . 136
5.25 Cokriged maps of canopy cover (%) from (a) 1988 data, (b) 1993 data and cokriging variance maps from (c) 1988 data and (d) 1993 data. ECODATA plot locations are shown as black squares on (a) and (b). 139
5.26 Six conditionally simulated realizations of canopy cover (%) from 1988 data. 140
5.27 (a) Mean of 50 conditionally simulated realizations of canopy cover (%) from 1988 data, (b) mean map from 1993 data, (c) variance map from 50 conditionally simulated realizations of canopy cover (%) from 1988 data and (d) variance map from 1993 data. ECODATA plot locations are shown as black squares on (a) and (b). 141
A.1 (a) Histogram of differences (unitless) between 300 pixel values from an image resampled by cubic convolution and the corresponding pixels from an image resampled by nearest-neighbor method and (b) Difference between the semivariogram of the nearest-neighbor resampled image and the semivariogram of the cubicly convoluted image, expressed as a proportion of the nearest-neighbor resampled image's semivariogram. 156

Acknowledgments

This thesis was begun during the time I worked as a contractor at NASA Ames Research Center; my thanks go to my employer at that time, Johnson Controls, for their contribution toward my tuition. The California State University Monterey Bay Foundation, my current employer, generously allowed leave-of-absences during the final stages of thesis writing. The Earth Science Division and the Ecosystem Science and Technology Branch at NASA Ames provided the resources and facilities that made this research possible. The Montana State Library provided some of the digital geographical data.

In addition, I want to acknowledge several people who were instrumental during my Ph.D. process. My supervisor, Paul Curran, taught and encouraged me throughout our collaboration. Geoff Smith, John Kupiec and Terry Dawson gave friendship and help as we overlapped in our postgraduate days. Dave Peterson at NASA Ames, from whom I first heard the word 'geostatistics', remained supportive throughout my time at Southampton. André Journel at Stanford has an immense dedication to his field which has been an inspiration to witness. I benefitted greatly from his excellent tutelage. Maria Mantas, of the Forest Service in Kalispell, Montana gave generously of her time and expert knowledge about the Flathead National Forest. Rick Rossi, with his curiosity about geostatistics and admirable assiduity in teasing apart its complexity, shared its philosophical underpinnings with me. The North American Council on Geostatistics and the people who attended over the years provided a convivial environment for learning how geostatistics is applied. Brian O'Connor helped me overcome difficulties with scientific software on numerous occasions. Finally, Brian and the rest of my beloved family and friends helped me stay the course.

Acronyms and symbols

α range parameter of semivariogram

β regression parameter

γ semivariance

$\gamma(\mathbf{h})$ semivariogram

Γ general cross-product statistic

ρ reflectance

ρ_{NIR} reflectance in a near infrared waveband

ρ_R reflectance in a red waveband

ρ_∞ reflectance of an infinitely deep vegetation canopy

ρ_0 reflectance of background

$\rho(\mathbf{h})$ correlogram

a ancillary variable

ADAR Airborne Data Acquisition and Registration

aMSS Airborne Multispectral Scanner

APAR Absorbed Photosynthetically Active Radiation

ASD Analytical Spectral Devices, Inc.

ATM Airborne Thematic Mapper

AVHRR Advanced Very High Resolution Radiometer

AVIRIS Airborne Visible/Infrared Imaging Spectrometer

B blue

BA basal area

C spatial covariance

CASI Compact Airborne Spectrographic Imager

cdf cumulative distribution function

ccdf conditional cumulative distribution function

CV coefficient of variation

DBH diameter at breast height

DEM digital elevation model

DN digital number

ECODATA Ecosystem Inventory and Analysis

EDIPS EROS Digital Image Processing System

EPA Environmental Protection Agency

EROS Earth Resources Observation Systems

f fraction of vegetation cover

fAPAR fraction of Absorbed Photosynthetically Active Radiation

FIFE First ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment

G green

GCP ground control point

GIS Geographical Information System

GLAI Green Leaf Area Index

GPS Global Positioning System

GVI Global Vegetation Index

h lag or pair separation distance

HAPEX Hydrologic and Atmospheric Pilot Experiment

iNDVI Time-integrated Normalized Difference Vegetation Index

L radiance, usually at-sensor radiance

LAI Leaf Area Index

LMC Linear model of coregionalization

MAUP Modifiable Areal Unit Problem, defined by Openshaw (1984)

MERIS Medium Resolution Imaging Spectroradiometer

MIR middle infrared

MODIS Moderate Resolution Imaging Spectroradiometer

MRLC Multi-Resolution Land Characteristics

MSS Landsat Multispectral Scanner

MTF Modulation Transfer Function

NAD27 1927 North American Datum

NAD83 1983 North American Datum

NASA National Aeronautics and Space Administration

NDVI Normalized Difference Vegetation Index

$NDVI_\infty$ Normalized Difference Vegetation Index of an infinitely deep vegation canopy

$NDVI_0$ Normalized Difference Vegetation Index of the background

NDVIc MIR corrected Normalized Difference Vegetation Index, derived by Nemani *et al.* (1993)

NIR near infrared

NOAA National Oceanic and Atmospheric Administration

OLS ordinary least squares

p primary variable

pdf probability density function

r Pearson correlation coefficient

R^2 regression coefficient of determination

R red

RMSE root-mean-square error

RT radiative transfer

SAVI Soil Adjusted Vegetation Index, derived by Huete (1988)

SEE standard error of the estimate

SEP standard error of the prediction

SPOT-HRV Système Pour L'Observation de la Terre -High Resolution Visible

SR simple ratio, described by (Tucker 1977)

SRF spatial random field

TC Kauth-Thomas tasseled cap transformation. This includes brightness (TCb), greenness (TCg) and wetness (TCw)

TM Landsat Thematic Mapper

TMS Thematic Mapper Simulator

TVI Transformed Vegetation Index, defined by Rouse *et al.* (1975)

USFS US Forest Service

USGS US Geological Survey

UTM Universal Transverse Mercator (UTM)

VIS visible

WDVI Weighted Difference Vegetation Index, derived by Clevers (1988)

1 Introduction

1.1 Understanding the Earth system

Vegetation covers a large portion of the Earth's terrestrial surface. Describing the spatial distribution of terrestrial vegetation is important for understanding ecological, climatological and hydrological systems. For example, models of the carbon cycle (Sellers *et al.* 1997, Pan *et al.* 1998), volatile organic carbon emission (Simpson *et al.* 1995) and hydrologic fluxes such as evapotranspiration and runoff (Mackay and Band 1997) all contain variables that quantify the amount of vegetation as a function of location. At regional to global scales, mapping vegetation amount may be an intermediate step on the way to understanding the biogeochemical cycles, the hydrologic cycle, or climate change (Dickinson *et al.* 1991, Xue *et al.* 1991, Bonan 1993). At local or regional scales, the maps may provide more direct information for monitoring crop productivity (Easterling *et al.* 1998) or forest health (Olsen and Schreuder 1997, Laurance *et al.* 1997).

The amount of vegetation at a given location is influenced by climate and soil factors at that location as well as human and animal action. Climate, which controls light, heat and water supply and soil factors relating to water and nutrient supply join human management practices, prior floristic composition, decomposition, disease and herbivory by mammals and insects as the major determinants of vegetation amount (Schultz 1995). Vegetation amount, in turn, affects the surface energy exchange and therefore climate (Henderson-Sellers 1993, Nemani *et al.* 1996), the interception of precipitation (Teklehaimanot and Jarvis 1991) and the erosion of soil by wind and water (Zhang *et al.* 1996, Stohlgren *et al.* 1998). To further understand the role of vegetation and its feedbacks in the Earth system, vegetation amount mapping is required at a variety of scales.

Understanding the spatial distribution of vegetation amount is an inherently geographical problem. Beyond such limited descriptions as 'clumped, 'patchy', or 'random', what is needed is a map. If dynamic processes are involved, many maps representing a landscape over time may be required. The scientific purposes to which such maps are to be used require cartometry (Maling 1989), rather than cartography. Cartometry, or measurement from maps, is behind the application of

maps in computer models of Earth systems via Geographical Information Systems (GIS). The maps (often called layers in a GIS) can be analyzed in a number of ways. They can be used to obtain estimates of the total effect of a process by integrating over arbitrarily defined zones (Gross *et al.* 1987), combined with other maps to infer relationships with other environmental quantities (Davidson and Lefebvre 1993) or input to ecological models (Running 1990, Coughlan and Dungan 1996). The quantitative use of maps in GIS put a special burden on prediction models: they must be *everywhere* as accurate as possible.

1.2 Measurements of vegetation amount

There are several possible ways to quantify vegetation amount required by Earth system studies, depending on the question being asked. Quantities that are typically used include leaf area index (Watson 1947); fraction of canopy cover (Stewart 1988); biomass, specifically total, aboveground, live, dead, or foliar (Gholz 1982); and volume (Eid and Naesset 1998). Leaf area index (LAI) and canopy cover are dimensionless values, comprised of area per area expressions. Biomass is expressed either as a mass or mass per area and volume is expressed as either a volume or volume per area. Although LAI, canopy cover, biomass and volume all quantify some aspect of the amount of vegetation at a particular place at a particular time, each is unique and may or may not be highly correlated with each other. For example, a one hectare plot located in a woodland with grass understory might have very high foliar biomass (including foliage from both grass and woody vegetation), but have a very low stem volume. Therefore, studies must clearly identify the vegetation amount quantity or quantities of interest, which will herein be called 'primary' quantities. Other quantities that may be closely related to vegetation amount, such as primary production and absorbed and intercepted photosynthetically active radiation, are also separate and unique quantities.

Data on a vegetation amount quantity may come from direct measurement, such as weighing the plant components harvested from a plot or determining the total area of foliage removed from the canopy (Daughtry 1990) or non-destructive, indirect measurements. When direct measurements involve plant removal, their

number is naturally limited by the desire not to reduce the availability of experimental material. To avoid destroying the experimental subject and to limit the labor involved in collecting measurements, allometric models are often used to predict one vegetation quantity from another (Kittredge 1944, Gower *et al.* 1987) using non-destructive methods. Other indirect methods include light transmittance measurements made under the canopy that are then used to infer the amount of light-obstructing vegetation (Parton *et al.* 1991, Vose *et al.* 1995, Gower *et al.* 1999). In practice, most measurements of vegetation canopies are indirect (Dufrene and Breda 1995, Gower *et al.* 1999).

Ground-based measurements of vegetation amount are necessarily limited to small areas, on plots of various shapes ranging in size from less than a square meter to not more than about a hectare. If the area to be mapped is large, for example a geomorphological province or level 1 watershed (Seaber *et al.* 1987), the total area actually measured is a minuscule fraction of the area to be mapped. Mapping therefore requires inference at unmeasured locations, or spatial prediction.

1.3 Predictions of vegetation amount

Models must be used to predict values at locations where no measurements are available (Journel and Alabert 1989). Models are either statistical or deterministic – the former use mathematics to achieve results with certain statistical properties, the latter use mathematics to represent processes presumed to be important in the physical environment. Language describing how measurement data are used for prediction is specific to the type of model used and can be inconsistent. This thesis deals with statistical models, therefore the term 'variable' is used to represent quantities that vary spatially or temporally, for example the primary vegetation amount quantities. The term 'parameter' is reserved here for (constant) coefficients of mathematical equations within a model.

Across the literature relevant to this topic, there is no clear convention on the use of the words 'prediction' and 'estimation'. For example, in the remote sensing literature, variables are 'estimated', and increasingly the word 'retrieval' is used as a synonym for the prediction of variables (i.e., Beaudoin *et al.* 1994, Rollin and Milton

1998, Bicheron and Leroy 1999). The classical statistical sense (Draper and Smith 1998) of 'prediction' will be used in this thesis unless otherwise stated. Accordingly, the value of a variable is predicted. The word 'estimation' is therefore reserved for obtaining values of model parameters.

Whether spatial or not, any prediction from a model should be accurate, meaning that it should be unbiased (the expected value of the error should be zero) and precise (the errors should be small). Additional factors in spatial prediction may include the need for reproduction of the mean, variance and frequency distribution over the spatial domain, its spatial pattern at several scales and its co-occurrence with other spatial variables. Unavoidably, predictions will be in error, so the model should include a term or terms for error.

Whereas error is a factual, objectively definable quantity, that is the difference between the predicted value and a true value, to have accurate knowledge of the error is to obviate the prediction problem. In other words, perfect knowledge of model error would entail perfect knowledge of the true values. Therefore, the best that can be done to understand the quality of a prediction is to have a model for uncertainty. While there are many aspects to uncertainty (Kundzewicz 1995, Brincombe 1997, Atkinson 1999), it is used here in the specific sense of 'the range of predicted values'. It is a more subjective quantity than error since it is a function of a model or models. For spatial prediction problems, uncertainty should also be mappable, because the range of predicted values will be smaller in some areas than in others. It is this *geographical* description of uncertainty that can identify areas for further sampling, model refinement or comparison.

Because ground measurement plots are exceedingly small relative to the area to be mapped, they are often considered quasi-points. Models for mapping may simply rely on these quasi-point data of the primary variable. Or the models may include data from other, ancillary sources that are related to vegetation amount. They are ancillary in the sense that they supplement the information provided by the measurements of the primary variable. Remote sensing is an important source of ancillary information for vegetation amount prediction.

1.4 Remote sensing for spatial prediction

The reflectance of vegetation canopies in near infrared and visible wavelengths is physically related to vegetation amount (Myers and Allen 1968, Jordan 1969, Colwell 1974, Curran 1980). Therefore, reflectance data in these optical wavelengths, collected by satellite sensors such as the Landsat Thematic Mapper (TM), the Système Pour L'Observation de la Terre (SPOT) High Resolution Visible (HRV) instrument and the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR) should be useful in the prediction and mapping of vegetation amount (Rouse *et al.* 1975, Jensen 1983, Running *et al.* 1986, Wulder 1998). Images of the Earth's surface acquired with optical and radar sensors have been used as sources for information on the spatial distribution of vegetation since aerial photography became a tool for forest inventory in the 1920s (Seeley 1929, Heller and Ulliman 1983). Currently it is accepted that remote sensing is the only practical means of obtaining spatially extensive and exhaustive data for mapping over large regions (Ustin *et al.* 1991). But it provides only indirect information on vegetation amount (Curran *et al.* 1998), since these sensors record measurements of reflected radiation not of a vegetation quantity.

The use of remote sensing for predicting vegetation amount spatially is similar to other prediction problems in remote sensing, such as mapping evapotranspiration (Moran and Jackson 1991) and snow cover (Lillesand *et al.* 1982, Rosenthal and Dozier 1996) on land or suspended sediment (Curran 1988) and chlorophyll concentration (Johnson 1978, Allee and Johnson 1999) in water bodies. Two general types of methods to solve these problems were recognized early in the development of satellite remote sensing applications. McCluney (1976) categorized the transformation of remotely sensed images of water bodies to maps of water color as theory-based and empirically based methods. By theoretical methods, McCluney meant the inversion of analytical, deterministic models of light/water interaction to predict color quantities. Empirical methods included the development of statistical models built from image and ground data collected at the same locations. This categorization is equally applicable to mapping of vegetation quantities (Hall *et al.* 1995).

1.4.1 Theory-based methods

Many have reported on the theoretical, also known as the physically-based method for predicting vegetation quantities (Clevers 1988, Pinty *et al.* 1990, Rosema *et al.* 1992, Fischer 1994, Woodcock *et al.* 1997, Kuusk 1998, Gemmell and Varjo 1999). Theoretical methods require a comprehensive understanding of all the significant physical phenomena contributing to the remotely sensed signal and an invertible mathematical model describing those phenomena. The difficulties in applying these methods at their present state of development are numerous (Myneni *et al.* 1995, Gemmell and Varjo 1999). Difficulties, called 'caveats' in Myneni *et al.* (1995), are factors *besides* vegetation amount that affect spectral reflectance, including view and illumination direction, atmosphere, canopy structure, background or soil, nonlinearity of scattering, spatial heterogeneity, adjacency, topography and mixing effects. Each factor has different manifestations at different scales and spatial resolutions. Therefore, parameterizing such deterministic models becomes another prediction problem, since these factors also vary spatially. Scatter in relationships between vegetation amount variables and spectral variables can be caused by all of these factors. Scatter is not due solely to error in the observations which might be caused by instrument noise or misregistration.

1.4.2 Empirical methods

Investigations using empirical methods are more numerous than those using theory-based methods. The objective of an empirically developed statistical model is to convert the value at each pixel or spatial region to the physical units of the primary variable, be it biomass (Tucker *et al.* 1983, Asrar *et al.* 1985, Box *et al.* 1989, Ardo 1992, Merrill *et al.* 1993, Friedl *et al.* 1994, Shippert *et al.* 1995, Todd *et al.* 1998), leaf area index (Heilman *et al.* 1977, Pollock and Kanemasu 1979, Curran and Williamson 1987, Peterson *et al.* 1987, Running *et al.* 1989, Curran *et al.* 1992, Nemani *et al.* 1993, Friedl *et al.* 1994, Coops *et al.* 1997, Gower *et al.* 1999, Turner *et al.* 1999), or canopy cover (Butera 1986, Larsson 1993, Ripple 1994). The statistical model employed by the vast majority of these empirical studies is regression. Despite this fact, regression is not generally covered by textbooks on remote sensing.

Curran and Hay (1986) and Curran and Williamson (1986) are among the very few papers that deal explicitly with methodological issues of regression in remote sensing. Recent reviews of the remote sensing of terrestrial surface properties, especially vegetation quantities (Goel and Norman 1992, Myneni *et al.* 1995, Verstraete *et al.* 1996), state that empirical use of spectral ratios are critical tools without actually stating the methodology for their use. Although the literature is rife with examples of relationships between remotely sensed data and vegetation amount, maps constructed using regression are surprisingly rare, despite the fact that regression is the statistical method of choice for developing prediction models. This curious vacuum in the remote sensing literature should be filled with a careful examination of the assumptions of the regression model, its usefulness in past studies and alternatives.

Although advances continue to be made in theory-based methods, at present a practical spatial prediction method must be based on an empirical approach. An empirical model may have a functional form drawn from theory (Price 1992, Leblon *et al.* 1993, Myneni *et al.* 1995) but utilize a statistical model rather than a deterministic model of the remote sensing process to obtain values for the primary vegetation amount variable.

1.5 Geostatistics for spatial prediction

Applications of regression in empirical studies of remotely sensed quantities historically have been developed without regard to spatial considerations, since neither absolute location coordinates nor relative variables (distance, direction) feature in the models. Alternative statistical models for spatial prediction, such as kriging (Journel and Huijbregts 1978), splines (Dubrule 1983) or Kalman filtering (Anderson and Moore 1979), all involve spatial attributes in their formulation and can therefore be considered part of that special field of geostatistics. Geostatistics can be defined as the field[1] that utilizes spatial information to make predictions.

As in other statistical approaches, there are descriptive functions and inferential functions in geostatistics (Isaaks and Srivastava 1989). There are tools to

[1]Though geostatistics is sometimes equated to the 'theory of regionalized variables' (Matheron 1971), the definition is usually taken to be broader by its diverse practitioners. Journel (1986) says it is a branch of statistics dealing with spatial phenomena. Cressie (1990) cites Hart (1954) as coining the term to refer to geographical statistics emphasizing location with areal distributions.

describe spatial structure and predict values at unsampled locations (Rossi *et al.* 1992). The description of spatial structure as a function of distance is constitutive to all geostatistical prediction algorithms. Geostatistical prediction is a form of regression (Cressie 1990), but to distinguish geostatistical regression from regression without spatial variables, the latter is called aspatial regression here. Aspatial regression has been the paradigm in remote sensing of vegetation amount and spatial regression is as yet relatively untried.

Another feature of geostatistical approaches is the recognition of regionalized nature of all environmental variables. That is, everything that can be measured in the environment must be defined over a distance, area or volume. In geostatistics, this distance, area or volume, including its size and shape, is called the spatial *support* (Olea 1990). For example, the area of a quadrat would be the support of destructive biomass and leaf area measurements made within that quadrat. The support of radiance measurements made by a Landsat sensor would be the area described by the point spread function (Forshaw *et al.* 1983) though precise knowledge of this function is often missing (Fisher 1997).

There are two main types of geostatistical prediction. The first type is called kriging, a spatial regression method that provides optimal, in the sense of unbiased and minimum-error, predictions at each unsampled location (Cressie 1990). The second type is conditional simulation, used to provide multiple, 'equally probable' maps of a variable which has been sampled. The original development of conditional simulation is attributed to Georges Matheron by Journel and Huijbregts (1978) and Cressie (1991). It focuses on the variability in spatial fields. Conditional simulation is the generation of synthetic realizations of a random function that equal the observations at sample locations and possess the spatial statistics of the sample data as measured by statistical moments up to the second order. Therefore, it is especially appropriate for some objectives of mapping, because it emphasizes the global attributes of a spatial field, rather than emphasizing local unbiasedness and precision as does kriging (Journel 1996). Conditional simulation algorithms yield not one, but several maps, each of which is equally likely to result from the algorithm. At any location it does not yield one predicted value, but a predicted probability

distribution.[2]

Both kriging and conditional simulation can use ancillary data along with primary data to increase the accuracy and/or decrease uncertainty about predictions. The multivariate form of kriging is called cokriging (Myers 1983). Conditional simulation can also incorporate multiple sets of information and is sometimes referred to as co-conditional simulation (Carr and Myers 1985).

Disciplinary and scientific field boundaries have evidently created impediments to the adoption of geostatistical methods by researchers in physical geography and the natural sciences who ask similar spatial questions. Kriging was originally applied to locating ore in gold-bearing regions in the late 1950s, though some of the statistical theory was developed earlier (Cressie 1990). The seminal geostatistics texts of the 1970s (David 1977, Journel and Huijbregts 1978, Clark 1979) were aimed primarily at resource exploitation and the journal *Mathematical Geology* was the principal outlet for the development of geostatistical models. In the early 1980s, Richard Webster and his colleagues promoted the use of such models for soil science objectives by providing a series of examples of soil sampling and mapping (McBratney *et al.* 1981, McBratney and Webster 1983, Webster and Burgess 1984). The 1980s also saw geostatistical applications in hydrology and meteorology (i.e., Hughes and Lettenmaier 1981, Bilonick 1988). Then ecologists had their awareness of geostatistics increased by Robertson (1987) and Rossi *et al.* (1992). Geostatistical techniques have also gained popularity in geography since the late 1980s (Oliver *et al.* 1989, Oliver 1989, Curran and Atkinson 1998). The application of geostatistics to remote sensing problems entered the refereed literature in the late 1980s (Curran 1988, Jupp *et al.* 1988, 1989, Woodcock *et al.* 1988*a*,*b*).

The slow diffusion of geostatistical applications beyond geological resource extraction probably has many causes. One is a peculiar jargon that many found difficult to penetrate (Journel 1986). Delay in the extension of geostatistics to remote sensing may be explained also by an apparent dissimilarity between remote sensing problems and reservoir characterization problems. Petroleum or mineral

[2]It should be noted that conditionally simulated realizations by convention are not considered 'predictions' because they are not optimal. Simulations are put into a prediction context in this thesis, in that they yield values for variables, which allows explicit comparison with optimal predictors.

reservoirs are usually sampled irregularly and sparsely, whereas in remote sensing an entire two-dimensional field is sampled on a regular grid. However, the instruments used to obtain information on this regular grid record electromagnetic energy reflected from the surface and reflectance is usually not the primary variable. Since more direct measurements of the primary variable may only be available on a sparse irregular basis, the problem then becomes how to use the remotely sensed information as ancillary data to interpolate between more direct measurements of the primary quantity. For this problem, geostatistical methods have potential to provide accurate spatial predictions and associated spatial descriptions of uncertainty.

1.6 Research questions in geostatistical prediction of vegetation amount using remotely sensed data

To make progress towards spatial prediction of vegetation amount using remote sensing it is reasonable to test geostatistical prediction methods, trying their usefulness for accurate mapping and for mapping uncertainty. In particular, this thesis addresses five research questions:

1. What are the implications of choosing geostatistical models for spatial prediction of vegetation amount using remotely sensed ancillary data? The implications are both theoretical and practical and will naturally be different for cokriging and conditional simulation. Since conditional simulation is as yet rarely applied in remote sensing applications, particular consideration will be given to the rationale, assumptions, methods and results necessary to build a case for conditional simulation using remotely sensed data. Both cokriging and simulation will be considered in relation to traditional aspatial regression methods.

2. How do models from aspatial regression, cokriging and conditional simulation relate to error and/or uncertainty about predictions made? New concepts (Journel 1990, Goovaerts 1997) of local uncertainty and spatial uncertainty are explored for maps of vegetation variables created using these three methods.

3. How closely do remote sensing data and vegetation amount have to be related

before aspatial regression becomes a more accurate method than cokriging? Rarely have cokriging and aspatial regression been considered side-by-side (an exception is Lesch *et al.* 1995). Practitioners wishing to make a choice of method should be aware of these alternatives and some heuristic recommendations.

4. How do spectral reflectances and vegetation indices from multispectral optical sensors relate statistically to quantitative regional measurements of vegetation amount from forest plots? Past studies on this question have been preponderantly limited to a small number of collocated ground plots and image pixels. This thesis contains the novel aspect of order-of-magnitude larger sample sizes than have been commonly available.

5. What is the effect of spatial support on these relationships? The examination of sample support is only beginning to be addressed as a fundamental aspect of the prediction problem in remote sensing (Atkinson 1997*b*). Here, sample support is addressed using two sensors with different spatial resolutions, Landsat TM and NOAA AVHRR.

Chapter 2 reviews the literature on the relationships observed between spectral variables from the optical region and vegetation amount variables. The use of geostatistics for remote sensing problems is also reviewed. In chapter 3, the theory of statistical predictors is described with particular reference to a comparison of aspatial regression and geostatistical methods. This theory is then applied to vegetation amount prediction using ground and image data from a synthetic data set (chapter 4) and a real data set (chapter 5). Maps are produced using these three methods for each data set and their accuracies are compared. Chapter 6 provides a summary of the results of this work and recommendations for further research.

2 Vegetation amount prediction using optical remote sensing

The remote sensing literature is replete with reports in which spectral measures from selected image pixels are compared with ground measurements, from coinciding locations, of a vegetation amount variable, such as biomass, leaf area index or stand volume. Many of these studies can be considered exploratory in nature and look for statistical associations between two or more variables of interest. Others go a step beyond the exploratory and propose regression models relating vegetation and spectral variables. Finally, a few apply regression models to produce a predicted spatial field, or map, describing a vegetation variable. As this chapter will show, for terrestrial applications of remotely sensed data, regression has been a prominent model in the quest for a means of deriving maps of vegetation amount.

In the way that regression has been used for these purposes, it is not explicitly spatial in that it does not include variables of location, distance or area. Alternative geostatistical models, that by definition do incorporate spatial attributes, are beginning to be recognized as applicable to remote sensing problems. This chapter reviews the literature on aspatial and geostatistical methods and examines it with respect to how closely vegetation variables and spectral measurements relate, the regression models that have been constructed from the relationships and how spatial information is introduced using alternative geostatistical models.

2.1 Aspatial regression with vegetation and spectral variables

The literature on prediction of vegetation amount using remote sensing is over thirty years old. Early work (Thomas *et al.* 1967, Thomas and Gerbermann 1977) utilized aerial photographs that were digitized using densitometers, thereby making it possible to quantitatively analyze the spectral response of sizable areas of vegetation canopies. The work accelerated soon after the launch of the first Landsat satellites in 1972 made spectral image data available over large regions of the Earth's surface. Studies shared a general methodology first published by Thomas *et al.* (1967). Their

methodology used digitized aerial photographs of a cotton crop. The resulting optical density values were transformed to transmittance and related to the percentage cover (in %) and yield (in kg ha^{-1}) of the cotton. Subsequent studies followed this methodology including: the selection of a small number of vegetation and spectral variables, the collection of a number of measurements of these variables made on the ground and from an image, and the comparison of statistical relationships between ground and image variables with null hypotheses. The strongest relationships (as quantified by the Pearson correlation coefficient, r, or the regression coefficient of determination, R^2) were emphasized as the main results. The relationships resulting from this research methodology were usually expressed as regression models. Well-known physical mechanisms, such as absorption of red light by chlorophyll in green vegetation and radiation scattering by leaves in near infrared wavelengths (Gates *et al.* 1965, Knipling 1970) were usually mentioned as assumptions for the selection of specific wavebands. Data values that did not fit the models were sometimes explained as erroneous or as influenced by other mechanisms. In some cases, the available data were reduced to a single model that was then applied to image data to generate a map of vegetation amount. Alternative forms of models for the same data were not usually considered, unless it was to consider multivariate models in comparison to univariate models. Regression diagnostics such as those recommended by Chatterjee and Hadi (1988) were rarely presented. The temporal and spatial limits of the regression model were not usually addressed, except by inference that the available data fit within those limits. Models were evaluated by standard errors of the regression and occasionally by checking the accuracy of predicted values for locations that have been measured separately. This general approach also has been extensively used for quantities besides vegetation amount, such as water temperature (Lathrop and Lillesand 1987), water sediment load (Curran 1988), plant canopy foliar biochemical content (Wessman *et al.* 1989, Treitz and Howarth 1999), mineral quantities in surface deposits (White *et al.* 1997*b*) and for vegetation studies using remotely sensed data outside the optical range (Dobson *et al.* 1992, Franklin *et al.* 1994).

To understand the characteristics of this general methodology, it is useful to look more closely at past studies. Table 2.1 is a summary of 51 peer-reviewed pa-

pers[1] describing studies that used this approach for comparing vegetation amount variables measured using ground-based methods with spectral image data measured from airborne or satellite platforms. While the table does not contain an exhaustive list, it is representative of the work reported in remote sensing and other journals from the late 1960s when the methodology emerged until the present. The table entries include the citation, where the study took place, when the data were collected, the type of plant canopy measured and the type of imaging sensor providing the remotely sensed measurements. Next are listed the response variable(s) and explanatory variable(s), the spatial support (area) of sample measurements representing these variables, the number of observations in the sample, whether the strongest univariate relationship was found to have positive or negative slope and the value of the largest correlation coefficient reported. In addition, if the significance of the relationship was tested, a regression model proposed and/or a map of predicted values was generated, a check mark ($\surd$) appears in the corresponding column. In the 'error' column, values of the regression standard error, in the units of the response variable, are listed if they were reported. In other cases, it is noted that confidence limits or a summary of the predicted values versus observed values were presented. If a table entry is blank, this indicates that information in the cited article was missing or insufficient. All of this information allows a consistent comparison of the studies, their findings and limitations. What follows is a discussion of the purposes, variables, image preprocessing, sampling design and regression models used in these studies. In most cases, more than one model was discussed. For space reasons, only the model with the highest r or R^2 is listed.

[1]Papers appearing in conference proceedings have been omitted from the table.

Table 2.1: Fifty-one published studies relating ground measurements of vegetation amount variables to spectral variables from remotely sensed images. Acronyms and symbols can be found starting on page xiv.

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Thomas *et al.* 1967	Weslaco, Texas	6/65	Cotton	Camera	Yield (% of maximum) 0-100		Green-red transmittance			+	r=.87	√	√	2.74	
Thomas & Gerberman 1977	Weslaco, Texas	1967	Cabbage	K-17 camera	Cover (%) 61-75, Transmittance (%) 7-12	15.2 × 3.9 m	Optical density .88-1.12, Yield (% of maximum) 35-61	15.2 × 3.9 m	15	- +	r=-.90, r=.91	√	√	Pred vs obs.	
Pollack & Kanemasu 1979	3 counties in Kansas	5-7/73	Winter wheat	MSS	LAI 0-2	.5 m^2	DN, DN ratios	80 × 80 m	115		R^2= .68 (multivariate)		√		
Wiegand *et al.* 1979	3 counties in Kansas	1974 - 1976	Wheat	MSS	LAI 0-4	Average of 3 ~1m^2 subsamples	TCg, PVI, TVI (from DN)	Avg from 40 ha area	25, 40	+	r=.953 (v. PVI) r=.866 (v PVI)	√	√	.08 .51	
Tucker *et al.* 1983	Ferlo region, Senegal	8/81	Grass/ forb	Radiometer, AVHRR	iNDVI 0-.75	1m^2, 1.1 × 1.1km	Total biomass (kg ha^{-1}) 0-8000	1m^2	150, 18	+	R^2= .75,		√		√
Aase *et al.* 1984	Sidney, Montana	5/81	Spring wheat	Airborne MSS	Grain yield (T ha^{-1}) 0-.9	Average of 6 1m^2 subsamples	SR	400 × 263 m	9	+	R^2= .89	√	√		
Musick 1984	Jornada range, New Mexico	9/80, 7/81	Shrub/ grass	MSS	NIR DN	80 × 80 m	Total cover (%) 0-35	Various	22	-	r=-.84	√	√	Conf. limits	

Table 2.1: (continued)

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Wardley & Curran 1984	Derbyshire, UK	9/82	Grasses	Milton radio-meter, aMSS	PVI (from ρ) .2–.45		GLAI .5-3		9	+	r=.72	√	√	Pred vs obs.	√
Jensen & Hodgson 1985	Aiken, South Carolina	8/82	Loblolly pine	Airborne scanner	Total biomass (kg) 0-4	Indivi-dual trees	31 trans-formations of DN	1.4 × 1.4 m	64	+	r=.76 (v SR)	√	√		√
Butera 1986	San Juan National Forest, Colorado	9/81	Pine/ aspen	TMS	Canopy closure (%) 0-100	10 ha	MIR DN	~10 ha	32		r=.8	√	√	Pred vs obs.	√
Everitt *et al.* 1986	Mercedes, Texas	4, 5/85	Grass	Sony video camera	Phytomass (kg ha^{-1}) 900-6500	4 .25 m^2 subsam-ples/ plot	DN, SR (from DN)	.9 ft^2	19	+	R^2= .71 (v SR)	√	√		
Franklin 1986	Mendocino, California	5/84	Conifer	TMS	Foliar biomass, basal area (m^2 ha^{-1}) 1-195	100, 250 or 314 m^2	DN, principal compo-nent, SR	105 × 105m, pooled	74, 19	-	R^2=.29 (R v log biomass) R^2=.67 (R v pooled log(bio-mass))	√			
Peterson *et al.* 1986	Sequoia National Park, California	9/83	Mixed conifer/ broad-leaved	ATM	Canopy closure (%) , 8-100, BA (m^2 ha^{-1}) 0-800	.1 ha	DN, TC, SR from DN	~24 × 24m	123	-	r=-.69 (MIR v closure)	√			
Curran & William-son 1987	Derbyshire, UK	6/84	Grasses	Milton radio-meter, aMSS	NDVI from ρ	5 × 5 m	GLAI 0-5	400 cm^2	24 , 10	+	r=.83, r=.72		√	Pred vs obs.	√

Table 2.1: (continued)

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Danson 1987	Nottinghamshire, UK	6/86	Corsican pine	SPOT-HRV	Height (m) 7-24, age (yrs), density, DBH, cover (%) 63-87, etc.	Various (10× 10m, other subsamples)	G, R NIR DN	20 × 20 m	28	-	r=-.83 (height v R)	√			
Gross *et al.* 1987	Lewes, Delaware	6/85	*Spartina alterniflora*	NASA Mark II radiometer, TM	Live biomass ($g\ m^{-2}$) 0-500	.25m^2	NDVI from ρ	.25m^2	53	+	r=.84	√	√		√
Ormsby *et al.* 1987	Washington, D.C	7/81	Woodlands, crop, urban	MSS, simulated AVHRR	Fractional vegetation cover (%) 0-100	10 × 10 km	NDVI, SR from L	10km × 10km	82	+	r=.95 (v SR)	√	√	7.22	
Peterson *et al.* 1987	Oregon	8/83	Conifer	ATM	SR from L		LAI .6-16.1	4 30 × 30m subsamples	18	+	R^2=.91	√	Non-linear	.77	
Vujakovic 1987	Okavango, Botswana	5/83	Savanna	MSS	ρ	154 × 134 m	Woody cover (%) 0-100	200 × 200m	49	-	r=-94 (v NIR)	√	√	7.12	√
Box *et al.* 1989	Earth land surface	1964 - 1984	Vegetation	AVHRR	iNDVI	225 km^2-900 km^2	NPP (g dry matter m^{-2} yr^{-1}) 0-4000	Various	95	+	r=.713		Non-linear		√
Clevers 1989	Wageningen, The Netherlands	1982	Barley	Camera	LAI 0-8	.13 m^2	R, NIR, PVI,NDVI, WDVI from ρ		~50	+			Non-linear	CV of residuals	
Cook *et al.* 1989	Illinois, North Carolina, New York	84, 85	Broadleaved forest	TM	Productivity kg ha^{-1} yr^{-1} 400-4500	.1 ha	DN, ratios (from DN)	8100m^2	32, 111, 44	-	R^2= .39, R^2= .27, R^2= .42	√		Conf. limits	√

Table 2.1: (continued)

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Nemani & Running 1989	Western Montana	7/84, 9/85	Conifer forest	TM, AVHRR	TM NDVI .35-.55, AVHRR NDVI .30-.65	9 TM pixel avg, 1.1 × 1.1 km	LAI (projected) 1.2-4.5		17, 53	+	R^2= .58 R^2= .88		Non-linear	.04, .03	
De Wulf *et al.* 1990	NE Belgium	2, 6/87	Pine	SPOT-HRV	L in VIS and NIR	Various	Density, age, DBH, height, vol	Various	Various					Class % correct	
Herwitz *et al.* 1990	Central Massachusetts	9/82, 9/87	Pine	TM	NDVI from L	8100 m^2	LAI (projected) 4-7	300 m^2 or 6750 m^2	15	+	r=.31	√			
Ahern *et al.* 1991	NW New Brunswick, Canada	7/85	Spruce /fir forest	TM	Net annual volume change (m^3 ha^{-1} yr^{-1}) -9 – 6	Avgd from .01-.04 ha subplots	ρ and ratios	1200 × 1200m	72, 8	-	R^2= .88 R^2= .91 (multivariate)	√	√	1.36 1.26	
Diallo *et al.* 1991	Senegal	1987 - 1988	Grassland, savanna	AVHRR	Biomass production (kg ha^{-1}) 300-6000	1000 m^2	iNDVI (from DN)		17	+	R^2= .82		√		√
Ripple *et al.* 1991	Corvallis, Oregon	7/88	Conifer	TM, SPOT-HRV	Volume (m^3 ha^{-1}) 28-840		TM NIR DN, HRV NIR DN	Various (2-74 ha)	46	-	r=-.83 r=-.89	√	√		
Ardo 1992	Oscarshamn, Sweden	10/88	Pine	TM	Stand volume (m^3 ha^{-1}) 0-300	Various	B, G, R, NIR, MIR5, MIR7 L	Various	198	-	r=-.74, -.78, -.71, .-48, -.79, and -.74	√	Non-linear	46.5	
Curran *et al.* 1992	Gainesville, Florida	2/88, 8/88, 3/89	Slash pine	TM	LAI 2.5-8	50 × 50 m	NDVI	30 × 30 m	16	+	R^2= .86, R^2= .82, R^2= .83		√	.33 .82 .52	

Table 2.1: (continued)

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Oza *et al.* 1992	Karnataka, India	3/85, 3/88	Teak	MSS	Canopy diameter (m) 1.9-4	30 × 30 m	L, NDVI, other ratios	240 × 240 m or 400 × 400 m	23	+	r=.73	√	√	.4	
Chong *et al.* 1993	African continent	1986 - 1989	Vegetation	AVHRR	NPP (tons dry matter ha^{-1} yr^{-1}) 0-32	Various	INDVI	48 × 48 km	78	+	r=.84		Non-linear		√
Duncan *et al.* 1993	Jornada, New Mexico	3, 6, and 9/89	Shrub	SPOT-HRV	TCg, TCb, ratios from ρ	100 × 100 m or 60 × 60 m	Canopy cover (%) 17-45	100 × 100m	70	+	R^2= .41 (v NDVI)	√	√	.5	
Larsson 1993	Kassala Province, The Sudan	2/90	Acacia	MSS, TM, SPOT-HRV	Canopy cover (%) 0-55	50 × 50 m	NDVI (from ρ)	6400m^2 900m^2 400m^2	15-19	+	r=.70 r=.55 r=.72	√	√	12-15	
Merrill *et al.* 1993	Yellowstone, Wyoming	8/87	Grass and shrub	MSS	Green phytomass (kg ha^{-1}) 480-3100	50 .18m^2 subplots	DN and ratios	2.7 ha	25		R^2= .63 (multivariate)	√	√	350	
Nemani *et al.* 1993	Seeley Swan Valley, Montana	1985	Conifer forest	TM	NDVI, NDVIc from L	90 × 90 m	LAI (projected) 1-6	Various	29	+	R^2= .64		Non-linear		√
Friedl *et al.* 1994	Konza Prairie, Kansas	6-8/87	Grass	TM	LAI .3-2.3, biomass (g m^{-2}) 200-700	1 m^2 quadrats, agg. to stations	TCb, SAVI	30 × 30 m	27	+	Adj R^2= .6 (v LAI) R^2= .49 (v biomass)	√		Conf. limits	
Ripple 1994	Oregon	7/88	Conifer forest	MSS, AVHRR	DN, NDVI	1 × 1 km	Closed conifer canopy cover (%) 0-80	1 × 1 km	89	-	R^2= .46 (multivariate)	√	√	.02	√

Table 2.1: (continued)

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Spanner *et al.* 1994	Oregon	7/90, 8/90 10/90	Conifer forest	TMS, AVIRIS CASI	SR from ρ	Various	LAI (projected) .5-10.5	Various	26, 11, 8	+	R^2=.97, R^2=.82, R^2=.92,		√	.47, 2.47, 2.15	
Baulies & Pons 1995	Catalonia, Spain	5/91	Pine forest	CASI	Biomass (MT ha^{-1}) 5.5-44, LAI .1-1.7	20 m diameter	DN	15 × 15 m	29		r=.85 (v log(biomass)), r=.77 (v log(LAI))	√	Non-linear	Pred vs obs.	√
Gong *et al.* 1995	Oregon	5/91	Conifer forest	CASI	LAI 1-11	Various	ρ, log(ρ), NDVI, other ratios	Various	30		R^2= .96 (multivariate)	√		RMSE	√
Shippert *et al.* 1995	Kuparuk River watershed, Alaska	7/89, 8/93	Tundra	ASD spectrometer, SPOT-HRV	Biomass (g m^{-2}) 200-900, LAI (0-2)	Various	NDVI from L	Various	4	+	R^2= .96, R^2= .97	√			√
Chen & Cihlar 1996	Candle Lake, Saskatchewan; Thompson, Manitoba, Canada	94	Black spruce/ Jack pine	TM	NDVI, SR	7-9 pixel avg	LAI, Effective LAI	Various	22	+	R^2= .66				
Jakubauskas 1996	Yellowstone, Wyoming	7/91	Lodgepole pine	TM	L	28.5 × 28.5 m	LAI 0-15, biomass (MT ha^{-1}), 0-17, others	20 × 25 m	69	-	r=-.6 (R v LAI) r=-.53 (R v biomass)	√	Non-linear		
Coops *et al.* 1997	New South Wales, Australia	13 dates 77-89	Eucalyptus	MSS	LAI .8-4.95	18 × 18 m	SR, NDVI from L	80 × 80 m or 240 × 240 m	13	+	r=.84 (v NDVI)		√	.57	

Table 2.1: (continued)

Reference	Location	Date	Type of plant canopy	Sensor	Response variable(s) *y*	Support of *y*	Explanatory variable(s) *x*	Support of *x*	# of obs.	Dir.	Max r or R^2	Sig.	Model	Error	Map
Fassnacht *et al.* 1997	Northern Wisconsin	8/93	Forest	TM	LAI (projected) 1.4-8.4	25 × 25 m	NDVI, SR, TC from L	90 × 90m	24	+	r=.74 (v conifer TCg)	√	Non-linear	RMSE	
Franklin *et al.* 1997	New Brunswick, Canada	8/92	Forest	TM	LAI 4-7	20 × 20m	NDVI	30 × 30m	17		R^2= .15	√			
White *et al.* 1997	Glacier National Park, Montana	8/94	Forest/ grass	TM	LAI 3-14		Ratio transformations with ρ	90 × 90m	109	+	R^2= .90 (v NDVI)		Non-linear		
Rasmussen 1998	Peanut Basin, Senegal	1990 -91	Crop	AVHRR	Millet yield (kg ha^{-1}) 0-1450	Plots agg. to 4 km^2 area	iNDVI from ρ		11, 15	+	r=.78 (in 1990), r=.83 (in 1991)	√	√	254 186	√
Todd *et al.* 1998	North-central Colorado	7/91	Grass	TM	Standing crop ($g\ m^{-2}$) 70-160	90 × 90 m	R DN	90 × 90 m	12	-	R^2=.35	√	√	1606	
Phinn *et al.* 1999	San Diego, California	6/95, 7/96	Salt marsh	ADAR	Canopy cover (%) 0-100	.75 × .75 m	DN, NDVI	.75 × .75 m	72	+	R^2= .72 (v NDVI)	√		RMSE	√

2.1.1 Purposes of vegetation-spectral relationship studies

Though the general procedures had much in common among the 51 studies in the table, their purposes differed. Six had the purpose of only testing association between vegetation amount and spectral variables. The null hypothesis, while not explicitly stated in these papers, was that the association does not differ from what would occur with randomly generated data. All four rejected this null hypothesis. Of these, the smallest maximum correlation was found by Herwitz *et al.* (1990), who reported an r of .31 between NDVI and leaf area index (LAI) from pine plantations. The largest maximum correlation was found by Danson (1987), who reported an inverse correlation (r=-.83) between average stand height in a Corsican pine plantation and DNs in the red waveband of SPOT-HRV data.

The rest of the studies went beyond reporting of the linear correlation coefficient to develop regression models. The relevance of the regression equation(s) was often left unstated, but it can be inferred that the immediate or eventual goal was prediction of the vegetation amount variable. Twenty-seven of the studies stopped short of applying the regression equation(s) for this purpose, while the other 18 produced maps with them. The maps were presented in choropleth form, with a color or grey level to represent each interval class.

Reliance on the regression equation for producing a map varied; for example Wardley and Curran (1984) developed a regression equation from ground-based measurements and field radiometer data but cited radiometric problems in using this model with airborne multispectral scanner imagery to predict green leaf area index (GLAI). Instead, they classified the image using density slicing to spatially predict GLAI and evaluated the accuracy of the result based on the ground measurement information. Most others producing maps utilized the regression model by applying it to each pixel of the image(s) of the explanatory spectral variable(s).

2.1.2 Variables used

Many different spectral variables were used in these studies. Some used digital numbers (DNs), others transformed DNs to the physical units of radiance (L in units of power/solid angle/waveband/area) and some calculated a reflectance variable (ρ

in units of %). Many studies used spectral ratios based on DNs, L or ρ in different wavebands. The most common ratio was the Normalized Difference Vegetation Index (NDVI, Rouse *et al.* (1975)); others were the Simple Ratio (SR, Tucker (1977)), Soil Adjusted Vegetation Index (SAVI, Huete (1988)), Perpendicular Vegetation Index (PVI, Richardson and Wiegand (1977)) and the Transformed Vegetation Index (TVI, Rouse *et al.* (1975)). 'Tasseled Cap' (TC) components (Crist and Kauth 1986), such as component 2 (greenness) and 6, have also been used. Verstraete and Pinty (1996) showed analytically that a single spectral transform can never be optimal for all purposes. Guyot and Gu (1994) showed that image preprocessing, including correction for a sensor's modulation transfer function (MTF), transformation to reflectance and correction for atmospheric path radiance, has a large effect on the magnitude of spectral ratios. Teillet *et al.* (1997) rigorously define the multiple combinations of spectral variables that can be used to calculate ratios and further emphasize that even when data are absolutely calibrated to physical units, inter-sensor comparison is imperfect because each sensor has different waveband ranges. Therefore, both the variety of sensors and the lack of consistent methods for obtaining ratios from them makes it impossible to compare ratios in absolute units.

A variety of vegetation amount variables have been studied as well, including percentage vegetation cover, mean canopy diameter, leaf area index, biomass and timber volume. In theory, any of these variables can be compared in absolute units, though myriad ground measurement techniques create discrepancies that reduce the potential for comparison. For example, leaf area index can be expressed as total LAI, including all sides of the leaves, projected LAI, the surface area intersected by downward radiation only or even effective LAI, the product of one half the total LAI with a factor describing the nonrandomness of the foliage (Chen and Cihlar 1995). For some leaf geometries, projected LAI is slightly different to single-sided LAI (Chen and Black 1992). Authors do not always state explicitly which form of LAI was used in a particular study.

While LAI, biomass, vegetation fraction etc. all relate to the 'state' of a vegetation canopy, some have argued that spectral variables are more functionally related to 'rate' variables, such as Absorbed Photosynthetically Active Radiation or APAR (Sellers *et al.* 1986, Myneni and Williams 1994). APAR is the amount of radiation

in the blue through red spectral region (commonly defined (Nobel 1983) as 400-700 nm) absorbed by a vegetation canopy per unit time. APAR can be normalized by irradiance to express the fraction of radiation absorbed by the canopy (fAPAR). Radiative transfer theory can be interpreted to show that, because reflectance in the near infrared (ρ_{NIR}) is proportional to $e^{-2h_{NIR}LAI/f}$, where f is the vegetation cover fraction, and fAPAR is proportional to $e^{-kLAI/f}$ and k is approximately equal to $2h$, then ρ_{NIR} is linearly related to fAPAR (Sellers 1987). This relationship may be complicated by different definitions of fAPAR based on the time period over which it is integrated (Goward and Huemmrich 1992). Empirical studies have also been accomplished with measurements of APAR versus remotely sensed image data (Demetriades-Shah *et al.* 1992, Li *et al.* 1997, Gower *et al.* 1999). They share a similar approach to those that use state variables, but are not included in table 2.1.

Early in the Landsat era, agricultural canopies were of paramount interest to those studying vegetation amount/reflectance relationships. By the late 1980s forest canopies had become the predominant subject of study, a trend that continues to the present. In theory, the larger biomass and leaf area index values of forests can put these values in the region of saturation in the near infrared and red regions, making forests a more difficult target for vegetation amount prediction. Interestingly though, correlations with forest canopy data do not seem to be any weaker than those from canopies with low values of vegetation amount (table 2.1).

2.1.3 Image preprocessing steps

A variety of geometric and radiometric preprocessing algorithms were used on the spectral data in these studies. Geometric preprocessing was a requirement for collocating ground and image pixels. Distortion caused by topography and the absence of high-contrast, recognizable features in natural or semi-natural areas often made it challenging to select ground control points. Most of the studies occurred before Global Positioning System (GPS) technology (Lance 1993) increased the accuracy possible for locating plots. Lack of correspondence between plot locations and image pixel locations in the correlation and regression studies is a source of error in the spectral variables, but it is not possible to evaluate its magnitude in past studies

because not enough information is available. Descriptions of the geometric preprocessing steps were often brief or non-existent in these reports. Some investigators cited the lack of geometric control as a reason to average several pixels centered on the location estimated to correspond to the ground measurements.

Resampling algorithms differed among the studies, with some reporting using nearest neighbor resampling, some reporting using cubic convolution while the majority did not report a choice of algorithm. Any resampling algorithm adds error to the image data; nearest neighbor preserves the radiometric values but puts them in the wrong location (Lillesand and Kiefer 1994) whereas cubic convolution can add new radiometric values to the population of spectral data (Schowengerdt 1997). Some studies did not resample the image data at all, but simply transformed the geographical coordinates of the plots into image coordinates, a way to avoid this resampling error.

Radiometric preprocessing was also diverse in the cited studies. Satellite-borne sensors measure radiance at the top of the atmosphere and airborne sensors measure radiance at aircraft altitude. But quantitative remote sensing of vegetation depends on measurements of radiance or reflectance at the top of the canopy. Therefore, image data from satellite or airborne sensors should be transformed to minimize the path radiance added by the atmosphere. In the papers summarized here, correction for atmospheric path radiance was spotty, though it was a concern even in early studies (Pollock and Kanemasu 1979). The most frequently used was a dark-object subtraction technique developed by Chavez (1988). The use of ratios of red and near infrared values by some authors was cited as a way to reduce the influence of atmospheric path radiance. Peterson *et al.* (1987) specifically showed that the sensitivity of the SR to LAI increased when path radiance was subtracted from the signals. Because the contribution of path radiance to the signal varies with location and date, it is not possible to determine retrospectively how its removal or lack thereof might have affected regression results in other studies. Path radiance has a larger contribution in the red than the near infrared (Schowengerdt 1997), so those models relying on red wavebands without path radiance reduction may be at a disadvantage.

2.1.4 Sampling considerations

Considerations of sample size, extent, arrangement and spatial support are crucial to spatial prediction. Only some of these aspects were consistently reported on in the literature summarized here. The smallest sample sizes were used by Aase *et al.* (1984) with an n of 9 and Shippert *et al.* (1995) who used pooled data to arrive at an n of 4. The largest size was 198 (Ardo 1992). Most sample sizes fell in the range of 20 to 100. Investigators strive for large sample sizes to achieve statistical power but are usually limited by the logistical constraints in collecting ground-based measurements. The study with the largest ground-based sample size did not result in greater significance than did any of the smaller studies, despite the increased power.

Extents of the studies also varied, from localities representing less than a few hectares (Thomas and Gerbermann 1977) to global extents (Box *et al.* 1989). In some cases, the extent of the study region required images to be collected during different overflights or overpasses of the aircraft or satellite, making scene normalization particularly important.

If the regression equation is to be applied to every pixel in the image to create a map of vegetation amount, the ground sample should be representative of the population of vegetation amount values that will be encountered in the region to be mapped (Steven 1987). A few studies cited this consideration for the collection of ground data. Another aspect of representativeness is that the remotely sensed data should be obtained so as to represent the conditions on the ground – they should therefore be obtained as closely in time to the ground data. The simultaneous collection of ground and image data is the ideal, rarely obtained because of cloud cover, sensor problems, the length of time required to collect ground observation and numerous other logistical reasons. The 'date' column in table 2.1 is only approximate and does not reflect the discrepancies that actually occurred in many of these studies.

Though several papers have described spatial sampling for accuracy assessment of image classification results (Hay 1979, Labovitz *et al.* 1982, Congalton 1988*a*, Stehman 1992, Stehman and Czaplewski 1998, Stehman 1999), the design of ground measurements of vegetation amount is relatively neglected. Few of the reports in the

table discussed the spatial arrangement of measurement plots (an exception is Friedl *et al.* 1994). The well-established principles of designing a ground-based spatial sample to achieve representativeness (Maling 1989, deGruijter and terBraak 1990) are not mentioned by most of these reports. In general, the stated consideration for sampling was to locate homogeneous plots or their surroundings while assuring their accessibility.

One of the inherent difficulties in designing correlation and regression studies of ground and airborne or satellite sensor measurements is obtaining equivalent spatial supports. The cited studies represented a large variety of supports for both types of variables. In most cases, there is a support mismatch between the response and explanatory variables. The mismatch has been increasingly understood as a problem (Atkinson 1997*b*), for both physical and statistical reasons. The physical reason is that covariance of two or more variables can only lead to insight into functional relationships when the variables describe entities that occupy the same physical area. For example, if biomass measurements made on a 1 m^2 quadrat are compared to reflectance measured from a 900 m^2 area, their association will only be meaningful if the 1 m^2 biomass value happens to be nearly equal to the biomass value from the larger 900 m^2 area. This is why investigators have tried to choose ground plots that are spatially 'homogeneous'. Curran and Williamson (1986) state the requirement for compatible spatial support and further emphasize that that multiple ground measurements (subsamples) must be acquired for every image pixel, particularly when such pixels are large. They point out that subsamples of adequate number are rarely acquired and in some cases are logistically unfeasible.

The statistical reason for the requirement of matching supports is the fact that the statistical characteristics of a spatial variable change as the support changes (Matheron 1985). In general, as support gets larger, the variance and range decrease, the minimum increases, the maximum decreases and the shape of the distribution becomes more symmetrical. These changes in the univariate characteristics have been called the support effect (Olea 1990) in the geostatistical literature. It follows that if the supports of two or more variables are changing independently, their covariance and correlation statistics also change. This effect in multivariate data has been identified in the geography literature as the modifiable areal unit problem,

or MAUP (Openshaw 1984). The MAUP is not a recent discovery (Gehlke and Biehl 1934), but has only become an active area of research since the 1980s (Fotheringham and Wong 1991). It states that changing the shape and/or size of the units on which data are mapped can change the resulting correlations or statistical models generated from the data. Openshaw (1984) distinguished the 'zoning' (changing shape) from the 'aggregation' (increasing size) components of the problem. The MAUP could be called a corollary of the support effect. Given that the supports of the ground data and remotely sensed data are not the same in most of the 51 studies, the MAUP alone makes the extension of these studies beyond a single data set problematic.

In cases where image registration error was perceived to be potentially large, averages of 3×3 or larger 'pixel windows' were used (i.e. Nemani and Running 1989, Ahern *et al.* 1991). This averaging effectively increases the support of the sensor measurement while leaving the ground measurement unchanged, increasing the support mismatch. An inherent mismatch also existed in studies that attempted to develop a calibration at the support of ground measurements using field radiometer measurements and apply the calibration equation to image data (i.e. Tucker *et al.* 1983, Wardley and Curran 1984, Curran and Williamson 1987, Gross *et al.* 1987, Shippert *et al.* 1995). The MAUP also makes this an untenable approach without a correction for the larger support of image data recorded by satellite or aircraft sensors.

2.1.5 Methods of fitting regression models to the data

The designation of response (also called 'dependent' or 'regressand') and explanatory[2] (also called 'independent' or 'regressor') variables is another choice in the regression methodology that relates to the purpose of the model and knowledge of relative errors in each. Of the 45 studies reporting regression models, 14 use the vegetation amount variable(s) as explanatory and the spectral variable(s) as response, 30 use the the converse and one looks at both combinations.

Though papers prior to 1986 do not describe the methods used to develop

[2]The term 'explanatory' is a misnomer, since in regression one variable cannot explain another (Johnston 1978). The term is used here in preference to 'independent' (which might be confused with spatial independence) and 'regressor' (which is so similar to 'regressand', it can be confusing).

the regression model, it can reasonably be assumed that an ordinary least squares (OLS) fit of the response variable on the explanatory variable was done in each case. After the publication of Curran and Hay (1986), which pointed out that OLS methods are not appropriate in cases where both response and explanatory variables are measured with significant error, several studies chose non-OLS methods for developing equations (Gross *et al.* 1987, Dewulf *et al.* 1990, Ardo 1992, Larsson 1993).

The use of the vegetation amount quantity as the response variable is contrary to the physical explanation that reflectance depends on the characteristics of vegetation (Curran and Hay 1986), but is in accordance with the statistical objective of predicting vegetation amount from spectral reflectance of the canopy. If spectral variables are used as explanatory variables and vegetation amount the response variable it is an appropriate form to predict vegetation amount values. If a spectral variable is used as the response and the objective is to predict vegetation amount values, the regression equation must be inverted. This inverse form is commonly used for calibration (Shukla 1972). However, as Webster *et al.* (1989) points out, this calibration approach is only germane when error is small in the response, in this case the spectral, variable. As a rule, an analysis of relative error is not carried out, but measurement errors achieved with current technology (Gu *et al.* 1992, Moran *et al.* 1995) appear to preclude the calibration approach.

2.1.6 Form of the models

Radiative transfer theory says that there is an asymptotic relationship between green vegetation amount (in the form of LAI) and reflected radiation (Price 1992). The asymptote, or the region of 'saturation', occurs at values of one-sided LAI of approximately 3 to 4 (Gobron *et al.* 1997). This is true in the simplest radiative transfer model, Beer's Law (Gates 1980, Price 1992), to models using less approximate radiative transfer equations, such as the Simple Biosphere model (Sellers *et al.* 1986). Yet few of the regression models developed from empirical data propose an asymptotic function. Instead, most are first order polynomials, linear in the variables. Part of the reason for this may be that the ranges of vegetation amount are

low, below the region of saturation (Pollock and Kanemasu 1979). However, others (i.e. Peterson *et al.* 1987, Ripple *et al.* 1991) used closed canopies with high values of green biomass, but did not observe asymptotic behavior in the relationship. The many error sources enumerated here may prevent the recognition of a saturation zone because of scatter in the relationships.

Many of the studies compared univariate models to multivariate ones, but rarely offered interpretation of the increased precision by added explanatory (but often collinear) variables. Interestingly, no two multivariate models in table 2.1 used the same combination of variables.

2.1.7 Evaluating the results of aspatial regression models

Though correlation coefficients are the most consistently reported statistic (only two of the 51 papers did not report values of r or R^2), they cannot be easily compared because sample sizes and variable ranges are not consistent among the studies. What should be more comparable are the standard errors of estimate (SEE) reported from the regression model. Values of SEE (listed in the error column in table 2.1) are in the units of the response variable and are reported in approximately one third of the studies. For Gaussian-distributed errors, observed mean values for a given explanatory value would be expected to fall within the range described by the SEE about 67% of the time. Confidence intervals, giving a 95% or 99% range of possible values for the mean predicted value, are based on the SEE and are a more realistic representation of the error predicted by a given model. Only a few studies report confidence intervals for the estimate (Cook *et al.* 1989, Friedl *et al.* 1994) or for newly predicted values (Musick 1984, Cook *et al.* 1989). The latter is constructed from the standard error of prediction, or SEP (equation 3.1.5 in Draper and Smith 1998).

Error analysis of the results of vegetation amount prediction from remote sensing, when it is done, has been confined to a check on predictions at a separate set of sample locations that were withheld from the model development stage (a 'validation' or 'test' set). In the table, 'Pred. vs. obs.' is listed in the error column, indicating that the study reported a statistic from the results from such a test set.

Three studies specifically reported a root mean square error on the predicted vs. observed data, though only two of these were on test sets.

Another method for seeking useful models is to perform cross-validation, a leave-n-out procedure in which the test set is composed of n values and the model developed with the remaining data (Efron 1982). This is a common practice in geostatistics (Davis 1987), with n often set to 1. It effectively includes spatial considerations by repeating comparisons between training and test data at different locations. This procedure does not truly validate the model itself, but only lends support to its usefulness for the region being studied (Davis 1987, Solow 1990). There is no example of cross-validation in table 2.1. Salvador and Pons (1998) recently showed that aspatial regression models for forest vegetation may not stand up to this kind of scrutiny.

2.1.8 Summary

Certainly, the studies represented in table 2.1 have improved the understanding of how vegetation reflects radiation in specific wavebands. They have been designed to answer such questions as, 'are vegetation amount variables and remotely sensed variables related?' and 'which variables?' Questions about what mechanisms cause the relationships under various conditions have also been explored by these studies. But fundamental understanding would be further strengthened if studies were developed to address the questions, 'can robust prediction models be developed?' 'what are the temporal and spatial limits of the models?' and 'what is the precision of prediction and is it unbiased?' Very few of the studies to date have been devised to answer these questions. While the studies have many aspects in common, there is enough variety in the sensors, spatial supports, canopy types, spectral variables, vegetation amount variables and methods as to prevent direct, quantitative comparison of even two regression model coefficients or error estimates. Another factor that makes comparison difficult is the heterogeneous preprocessing methods used with the image data. Authors often compare their studies to others, but only in general terms. The methods they use to evaluate their results are almost as variable as the experimental methods. Despite the optimism expressed in the early 1980s

that 'remote sensing of biomass is close to becoming an operational reality', (Jensen 1983, page 128), prediction of biomass and all other vegetation amount variables is still in a research phase.

2.2 Geostatistical analyses of remotely sensed data

The classical assumptions of aspatial regression are the lack of measurement error and independence of the explanatory variable(s) and homoscedasticity and absence of serial autocorrelation in the residuals (Poole and O'Farrell 1971, Shaw and Wheeler 1994). The previous section has pointed out the weaknesses in the first two assumptions for remotely sensed data. Homscedasticity, the stipulation that the variance of residuals be similar for all values of the explanatory variable can be tested with regression diagnostics (Chatterjee and Hadi 1988), mostly absent in the studies cited in table 2.1. The last assumption, the lack of serial autocorrelation of residuals, is likely to be violated when the data themselves are spatially autocorrelated unless the regression model itself accounts for all causes of the spatial autocorrelation (Anselin and Griffith 1988, McGwire *et al.* 1993).

Anselin and Griffith (1988) show that spatial dependence in geographical data can have a large impact in typical applications of aspatial regression. Haining (1990) states that spatial autocorrelation that is not accounted for in a regression can cause underestimation of the true sampling variance and overestimation of R^2. Borgman (1988) and Webster and Oliver (1990) describe the risks of using classical statistical predictors for autocorrelated data. deGruijter and terBraak (1990) counter these arguments and suggest that by designing spatial samples appropriately, spatial autocorrelation in the experimental data can be avoided and robust prediction models can be developed. Lesch *et al.* (1995) also discount the effect of autocorrelation on spatial prediction, if only for explanatory and response variables that are highly correlated. This lack of a consensus points to a need for further comparison of the aspatial and spatial prediction models.

Geostatistics includes alternative prediction models that do not require a model of spatial independence. The paramount characteristic of geostatistical methods is that they exploit spatial dependence in the data instead of assuming it does not exist.

Though the recognition of spatial dependence in remotely sensed data occurred very early (Leachtenauer 1977), it is only recently that geostatistical models have been employed to analyze these data (Curran and Atkinson 1998, Stein *et al.* 1998).

2.2.1 Recognition of spatial dependence

Spatial structure in remotely sensed images has been quantified in many ways, such as with the correlation function (Craig 1979, Labovitz *et al.* 1982, Schachter 1980), spectral density (Weszka *et al.* 1976, Leachtenauer 1977), texture (Haralick *et al.* 1973, Irons and Peterson 1981), structure function (Lesieur and Katsaros 1981) and fractal dimension (Ramstein and Raffy 1989, Lam 1990, Vignesadler *et al.* 1991, Jaggi *et al.* 1993, Emerson *et al.* 1999, Maitre and Pinciroli 1999). These measures of spatial dependence are also related to the now classical statistics of spatial dependence in geography – Geary's c and Moran's I (Cliff and Ord 1981). Legendre and Fortin (1989) show that Geary's c is closely related to the correlogram. Moran's I and Geary's c were designed to be used to distinguish between autocorrelation or independence in observed spatial data. For example, Congalton (1988*b*) used Moran's I to test for spatial autocorrelation in results from image classification. In contrast, spatial dependence is *assumed* to exist, therefore its existence is not tested for in geostatistical models. It is described using spatial covariance statistics, especially the semivariogram. The semivariogram was first introduced in the remote sensing literature by Curran (1988) and in a suite of papers by Woodcock, Jupp and Strahler (Jupp *et al.* 1988, 1989, Woodcock *et al.* 1988*a*,*b*).

Semivariograms of image data presented by numerous authors (Ramstein and Raffy 1989, Cohen *et al.* 1990, Lacaze *et al.* 1994, Phinn *et al.* 1996, Bruniquel-Pinel and Gastellu 1998, Abarca-Hernandez and Chica-Olmo 1999, Collins and Woodcock 1999, St-Onge 1999) have shown a variety of shapes and ranges. Methods of semivariogram calculation differed in these studies, with some calculating semivariance versus interpixel distance from transects of pixels, others from random samples of pixels and a few from all image data. All semivariograms were modeled as smooth curves, using spherical, exponential, Gaussian, or unbounded (fractal) functions. With few exceptions, spatial dependence had a range of at least many times the

pixel width.

Of the papers in table 2.1, only a few acknowledge that spatial autocorrelation may affect results of aspatial regression models (Box *et al.* 1989, Friedl *et al.* 1994). It is not possible to evaluate, in retrospect, whether it did affect results for any of the 51 studies, since information generally is not given on the spatial configuration of samples or analysis of regression residuals.

2.2.2 Uses of spatial dependence in geostatistical prediction

All of the spatial dependence statistics, fractal dimension, spatial covariance, correlation and the semivariogram, are related yet unique expressions. It is only when assumptions are made (forms of stationarity and ergodicity) that they are useful for inference (see chapter 3). In the past ten years, several types of inference from semivariograms and associated geostatistical models have been done with remotely sensed data (Curran and Atkinson 1998). They can be categorized into noise estimation, sample design, classification and prediction of continuous variables.

Noise estimation – Curran and Dungan (1989) suggested that the extrapolation of the experimental semivariogram to a lag of 0, called the nugget variance (see chapter 3), can be interpreted as an estimate of sensor noise. They applied this method to imaging spectrometer data from a United States National Aeronautics and Space Administration (NASA) experimental sensor. Wald (1989) made the same interpretation from semivariograms calculated from NOAA AVHRR data. The method was subsequently used by Smith and Curran (1996). There remain issues on the application of this method (Gao 1993, Atkinson 1997*a*), for example, how to select appropriate regions of images to calculate the semivariogram and how to extrapolate the function to the vertical axis.

Sampling design – The geostatistical literature contains many examples of how knowledge of spatial dependence, obtained from pilot sampling campaigns or other sources, can be used to minimize the redundancy of samples and help to specify where and how many places to collect measurements (McBratney *et al.* 1981, Russo 1984, Morris 1991). Webster *et al.* (1989) and Atkinson (1991) use this approach

for planning field radiometric measurements. Sample design also has to do with the question of the appropriate scale or support at which to collect measurements. Just as with all variance statistics, the semivariogram is affected by the size and shape of these measurements, or in the case of images, the effective spatial resolution element. By synthesizing digital images from simple disks of various shapes and sizes or size distributions, Woodcock *et al.* (1988*a*) and Jupp *et al.* (1989) showed how these shapes and sizes are reflected in the semivariogram's sill, shape near the origin and range. They also described the concept of 'regularization', the effect on spatial statistics (including the semivariogram) of a change in the pixel size. This effect had been noted earlier by Labovitz *et al.* (1982). 'Scale variance analysis', the effect of changing spatial resolution on variance among resolution elements, was formally developed by Moellering and Tobler (1972). It has been used with remotely sensed images by Townshend and Justice (1990) and Justice *et al.* (1991) among others to find the spatial resolution at which variance is maximum; this is considered the 'optimal resolution' for measurement. Similar reasoning has been made with the variances obtained from the semivariogram (Atkinson and Curran 1995, 1997). They are joined by Marceau *et al.* (1994) and Hyppanen (1996) in showing that because the shapes and sizes of land surface features are a heterogeneous mixture, a single optimum for all purposes is not possible.

Classification – Oliver and Webster (1989) proposed increasing the spatial contiguity of classes determined from image data using semivariogram parameters to modify multivariate distances calculated during the classification process. Other ways of using the semivariogram in classification included using the semivariogram range (Ramstein and Raffy 1989), the semivariance at selected lags (Miranda *et al.* 1994, van der Meer 1996, Carr 1996) and the semivariance at a single lag (Lloyd *et al.* 2000, Berberoglu *et al.* 2000) to act as an additional variable for the classifier. Each reported increased classification accuracy using the spatial information.

Prediction – Models of the semivariogram are used in interpolation functions to weight nearby sample measurements to make predictions spatially, in techniques called kriging (with one variable), cokriging (with more than one variable) or conditional simulation (Journel and Huijbregts 1978, Goovaerts 1997). Geostatistical pre-

diction models were first applied to remotely sensed data in the mid-1980s (Switzer 1983, Carr and Myers 1984). Despite these earlier papers focusing on prediction, descriptive geostatistics still comprise well over half the related remote sensing literature (Curran and Atkinson 1998). Carr and Myers (1984) suggested that kriging could be applied to spatial filtering and for resampling during geometric registration. They proposed that the resolution of an image could be increased using the cross correlation between it and higher spatial resolution images of the same area. Ramstein and Raffy (1989) discussed resampling applications and demonstrated them on undersampled Landsat Thematic Mapper scenes. Atkinson *et al.* (1990) also suggested that images could be stored in this undersampled state and reconstructed using the semivariogram information. Kriging techniques have also been used to fill gaps in images caused by clouds or sensor characteristics (Haining *et al.* 1989, Rossi *et al.* 1994, Inggs and Lord 1996, Addink and Stein 1999).

Cokriging is an alternative approach to kriging for predicting variables at unsampled locations. With cokriging, ground data can be treated as the primary variable and the remotely sensed data can be treated as an ancillary variable to increase the accuracy of prediction over what can be achieved by kriging alone. To use this method assumes that the spatial pattern of the variable to be predicted and that of the ancillary variable or variables are related (they co-vary). Cokriging is being increasingly used with remote sensing data. Bhatti *et al.* (1991) cokriged 172 soil measurements and a ratio of Landsat TM wavebands 4 and 5 to create maps of soil organic matter content. Atkinson *et al.* (1992) and Atkinson *et al.* (1994) demonstrated the combination of field or airborne radiometer data with ground-sampled biomass to cokrige maps of biomass. Gohin and Langlois (1993) used ship-based temperature measurements and thermal images from the NOAA AVHRR to create sea-surface temperature maps. Lohani and Mason (1999) constructed a Digital Elevation Model using estimates of shoreline height from Airborne Thematic Mapper imagery as ancillary data and ground-surveyed measurements of height as primary data.

Conditional simulation, a stochastic interpolation technique whose objective is to reproduce the univariate statistics and spatial pattern of the entire field of a variable of interest, is beginning to appear in the remote sensing literature. Geo-

statistical methods of simulation have gone under several names. Early papers on this topic refer to it as conditional simulation (Journel 1974, Journel and Huijbregts 1978), to distinguish it from other Monte Carlo methods that do not reproduce the data at sample locations. Seo *et al.* (1990) and others have used the term stochastic interpolation while Journel and colleagues are more recently using the terms stochastic simulation and stochastic imaging (Journel 1993, Kyriakidis and Journel 1999). Christakos (1992), in his efforts to put geostatistical methods of simulation in a more general mathematical context, describes 'spatial random field (SRF) simulation'. Simulation methods that are unconditional to data were introduced in geography by Goodchild (1980) and have been used by Friedl *et al.* (1995). Conditional simulation with a primary variable alone has been tested with synthetic data by Chainey and Stuart (1998) and with forest data by Mowrer (1997). Geostatistical simulation that is conditional to ground measurements of a variable while exploiting ancillary measurements from remote sensing have the potential to yield useful uncertainty models.

2.3 Conclusion

Nearly all published studies of vegetation amount variables find statistically significant relationships between the amount of radiation reflected by the Earth's surface and the amount of green vegetation on that surface.[3] The challenge that faces those using remote sensing as a tool is how to turn this statistical significance into a robust and accurate methodology for predicting vegetation amount spatially. Advancing this research would benefit by reproducible methods, explicit descriptions of spatial support, diagnostic tests of regression models and increased comparison among alternative models. The absolute calibration of spectral variables to physical units, traceable to standards, would facilitate the comparison of data and models resulting from them. The field of remote sensing is currently seeing the development of geostatistical methods, which have advantageous attributes in their exploitation of spatial dependence, explicit recognition of measurement support and models of

[3]There are exceptions to this rule (i.e. Badwhar *et al.* 1986), and it is difficult to estimate the 'file drawer effect' – the number of studies that were not published because they found no statistical significance.

uncertainty. However, and of relevance to this thesis, their potential for replacing aspatial regression for prediction of vegetation amount has yet to be tested.

The lack of attention to the regression methodology in the remote sensing literature may be caused by a perception that physically based models address the *causes* of reflectance and therefore are naturally more robust. A prevailing premise is that once radiative transfer in vegetation canopies is fully understood, these models can be inverted to generate maps of any relevant vegetation variable. A challenge with this approach is that the input variables vary spatially, making the inversion problem ill-posed. The literature on physical models is at least the size of that on correlation and regression, yet even fewer of these studies deal with spatial issues or result in maps (Rosema *et al.* 1992, Kuusk 1998). None of these models include spatial location as variables, therefore maps can be generated by treating location as an exogenous variable, and applying the model at each location. However, spatial support is not explicitly considered in these models, nor is the redundancy inherent in using spatially autocorrelated data for their parameterization. The next chapter will describe the basis for statistical models that do account for spatial support and autocorrelation.

3 Models and methods for prediction

This chapter describes a general statistical model, the random function model, for spatially predicting vegetation amount variables using ground and remotely sensed measurements. Depending on how the model is formulated, it yields several alternative methods for prediction. The three methods considered here are those that can be used with sparse ground measurements of vegetation amount and exhaustive indirect measurements from remote sensing instruments. The methods are aspatial regression, cokriging and conditional simulation. The following discussion is drawn from statistical, geographical and geostatistical texts in an attempt to unify typical aspatial regression treatment in the remote sensing literature (see chapter 2), which does not explicitly discuss a random function basis for the method, and geostatistical treatment in the natural sciences (i.e. Goovaerts 1997), which generally does not consider random functions without spatial dependence.

3.1 The random function model

A random function is composed of random variables. A random variable $Y\{s : s \in \Omega\}$ is a function for which one and only one value is defined for each s in the set of all possible outcomes (Ω) and each value has a probability (Cox and Miller 1965). Therefore, a probability distribution exists over the range of these outcomes. The distribution can be described as a cumulative distribution function (cdf), the probability of $(Y(s) \leq s_i)$ for all s_i. The cdf increases monotonically from 0 to 1 as s_i increases from $-\infty$ to $+\infty$. The probability density function (pdf) is an alternative way of characterizing this distribution; it is the first derivative of the cdf.

A collection of random variables can be defined for a single dimension, $\{Z(x, s) : x \in D, s \in \Omega\}$ where D is a (one-dimensional) index set[1]. Such a collection is called a random or stochastic process (Cox and Miller 1965) and this model is often used for time series data where time defines the index set. A collection of random variables with D defined in n dimensions is called a random function or random field (VanMarcke 1983). When D is two or three dimensions (i.e. a plane or volume) the random function lends itself to modeling natural phenomena existing in space.

[1] s is usually left out for efficient notation

Christakos (1992) calls these spatial random fields (SRFs). For some purposes in modeling the terrestrial surface, the third dimension (depth or altitude), can be ignored, making x a two element vector, $\mathbf{x}$, defining a location in the region being studied.

Sets of values that arise from random functions are called realizations (Olea 1990). The simple example often given is to regard throws of a die as a one-dimensional random process, the index set being defined by the number of throws being made (Isaaks and Srivastava 1989). For one throw, the number that appears on the die can be regarded as a realization. Generalizing this to a two dimensional random function, a single realization could be described as a vector $[z(\mathbf{x}_1), z(\mathbf{x}_2), z(\mathbf{x}_n)]$, where $\mathbf{x}_i$ describes a coordinate (x_i, y_i) in a two dimensional plane and the $z(\mathbf{x}_i)$ are values at those coordinates. Any variable distributed in space and therefore having at least two dimensions is called a regionalized variable (Olea 1990) and so regionalized variables can be modeled as realizations from a random function. Herein, upper case letters will be used to symbolize random variables or functions (i.e. $Z(\mathbf{x})$) while lower case will denote realizations (i.e. $z(\mathbf{x})$).

3.1.1 Forms of random functions

Random functions (RFs) can be defined to have different properties. The properties of a particular RF form lead to certain prediction methods because of the simplifications and identities the properties allow. Two major forms of RFs consist of independent and dependent random variables, respectively.

Independent random variables If the random variables (RVs) comprising a random function are independent, their correlation is zero. This property makes the spatial dimension essentially irrelevant. Further, the RVs may be considered to be identically distributed (their cdfs are the same) and if so are sometimes called purely random functions (VanMarcke 1983). A still more strict assumption is that they all come from a particular distribution function with an analytical description, such as the Gaussian (normal) distribution.

Dependent random variables If the random variables comprising a random function are not independent, this means that the random variable $Z(\mathbf{x})$ is related to random variable a distance $\mathbf{h}$ away, $Z(\mathbf{x}+\mathbf{h})$ and is therefore spatially dependent. The distance $\mathbf{h}$ is a vector, conformable with $\mathbf{x}$ and, for 2D, equal to (h_1, h_2) where h_1 is the distance in the x direction and h_2 is the distance in the y direction. Geostatistics is founded on random functions consisting of RVs that are spatially dependent in specific ways (Matheron 1971). This quality makes geostatistics appealing to geography, whose 'first law' has been given as, 'Everything is related to everything else, but near things are more related than distant things' (Tobler 1970, page 236).

An important aspect of real measurements of spatial phenomena is that they are always defined over some area or volume. This aspect of measurements is incorporated into random function models with the concept of spatial support. A random function $Z_v(\mathbf{x})$ is defined for the support v as

$$Z_v(\mathbf{x}) = \frac{1}{v}\int_{v(\mathbf{x})} Z(\mathbf{y})d\mathbf{y} \tag{3.1}$$

(Rendu 1981), the mean value of the RF $Z(\mathbf{y})$ for all points within the area (or volume) v. $Z_v(\mathbf{x})$ is said to be regularized from the RF defined by $Z(\mathbf{y})$. A complete specification of the support entails a description of the geometrical shape, size and orientation of the area or volume (Olea 1990). In theory, the support can be as small as a point (in practice, nearly a point) or as large as the extent of the entire field (in practice, the entire study region).

For RFs defined with dependent random variables, statistics, especially above first order, are affected by support (Matheron 1985). In contrast, purely random fields defined at one support will also be purely random fields when regularized (Journel and Huijbregts 1978), showing no spatial dependence. Again, this makes geographical phenomena, which change their statistics with support, more analogous to dependent random functions than independent random functions.

Since only a small portion of a random function is represented by data on the Earth's surface, that is, one 'realization', important assumptions must be made to apply random functions to real data to construct statistical prediction meth-

ods (Matheron 1989, Myers 1989). Some of these assumptions are various forms of stationarity (also called homogeneity, VanMarcke 1983), which declare certain characteristics of random functions to be invariant, and ergodicity.

3.1.2 Stationarity

Strong (also called strict) stationarity means that any two (or more) random vectors at k locations, $[Z(\mathbf{x}_1), Z(\mathbf{x}_2), \cdots, Z(\mathbf{x}_k)]$ and $[Z(\mathbf{x}_1+\mathbf{h}), Z(\mathbf{x}_2+\mathbf{h}), \cdots, Z(\mathbf{x}_k+\mathbf{h})]$, have the same joint cdf, no matter what the $\mathbf{h}$ (Cox and Miller 1965). If the mean and spatial covariance exist, they are invariant with location. A purely random function would, by definition, exhibit strong stationarity. Weak stationarity implies the expected value of a random field is a constant with a spatial covariance depending only on $\mathbf{h}$. No assumptions are made about joint distributions with weak stationarity.

Second-order stationarity (a form of weak stationarity) is the stationarity of second moments, that is, the variance of the RVs at every location is unchanging (Myers 1989). Also, the expectation or mean of $Z(\mathbf{x})$ does not vary with $\mathbf{x}$ and the spatial covariance and therefore the semivariogram and correlogram, depend only on $\mathbf{h}$ not $\mathbf{x}$. Second-order stationarity of the increment [$Z(\mathbf{x})$ and $Z(\mathbf{x}+\mathbf{h})$] is called 'intrinsic' stationarity (Journel and Huijbregts 1978). Again, the mean does not vary with $\mathbf{x}$. The spatial covariance need not exist, only the semivariogram $\gamma(\mathbf{h})$ must exist and is dependent on $\mathbf{h}$. If the spatial covariance does exist, it is related to the semivariogram by

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h}) \tag{3.2}$$

The semivariogram is equivalent to a constant (the variance) minus the spatial covariance. Intrinsic random functions are the basis for much of geostatistics. The use of intrinsic random functions for this purpose is called 'regionalized variable theory' (Matheron 1971).

3.1.3 Ergodicity

Ergodicity is the property that the statistics of a realization of infinite extent are equal to those of the random function (Cox and Miller 1965). As the extent of the

domain being studied increases, estimated values of these statistics are supposed to converge to their random function values. Ergodicity, like stationarity, can never be tested, it can only be deemed reasonable in making the model relevant to real-world situations (Myers 1989). And it can be seen that some kind of decision is necessary for modeling regionalized variables as random functions, since, unlike in the case of dice, the landscape represents a single reality, not multiple realizations (Matheron 1989). To use an ergodic random function means that the statistics of real data can be used as estimates of the parameters of that random function. Stationarity of some kind is critical to allow data from some locations or time periods to be used to make inferences about other locations or time periods. To date, the second-order moments form the foundation for inference including prediction using random functions.

3.1.4 Second-order moments of random functions

There is a variety of two-point statistics defined for random functions. Of course, these are particularly relevant for random functions consisting of spatially dependent RVs. Some of the most important are described here[2]:

Variance A fundamental second-order moment of an RF, $Z(\mathbf{x})$, is its variance[3].

$$\begin{aligned} Var(Z(\mathbf{x})) &= E\{[Z(\mathbf{x}) - E\{Z(\mathbf{x})\}]^2\} \\ &= E\{Z(\mathbf{x})^2\} - E\{Z(\mathbf{x})\}^2 \end{aligned} \tag{3.3}$$

usually estimated by

$$\hat{Var}(z(\mathbf{x})) = \frac{1}{n}\sum_{i=1}^{n} z(\mathbf{x}_i)^2 - m[z(\mathbf{x}_i)]^2 \tag{3.4}$$

where $m[z(\mathbf{x})]$ is the mean of $z(\mathbf{x})$ and n is the number of sample values.

[2]Except where noted, equations in this section are based on Olea (1990).
[3]In regionalized variable theory, this variance is called the 'a priori' variance (Olea 1990).

Spatial covariance The spatial covariance C is defined for the two RVs $Z(\mathbf{x})$ and $Z(\mathbf{x}+\mathbf{h})$ where $\mathbf{h}$ is the distance and direction of the vector between the RVs as,

$$\begin{aligned} C(Z(\mathbf{x}), Z(\mathbf{x}+\mathbf{h})) &= E\{[Z(\mathbf{x}) - E\{Z(\mathbf{x})\}][Z(\mathbf{x}+\mathbf{h}) - E\{Z(\mathbf{x}+\mathbf{h})\}]\} \\ &= E\{Z(\mathbf{x})Z(\mathbf{x}+\mathbf{h}) - E\{Z(\mathbf{x})\}E\{Z(\mathbf{x}+\mathbf{h})\}\} \end{aligned} \tag{3.5}$$

At $\mathbf{h} = \mathbf{0}$, C becomes equal to the a priori variance of $Z(\mathbf{x})$. Generally, as $\mathbf{h}$ increases, C decreases and, though it can fall below zero, it rarely does in practice. Models exists for which C does not exist, such as for the Brownian motion (Wiener-Levy) process (Journel and Huijbregts 1978) or other cases where variance continues to increase without bound as $\mathbf{h}$ increases.

The traditional estimator of this statistic is

$$\hat{C}(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{i=1}^{N(h)} [z(\mathbf{x}_i)z(\mathbf{x}_i+\mathbf{h}) - m[z(\mathbf{x})]^2] \tag{3.6}$$

where $N(\mathbf{h})$ is the number of pairs of data that are $\mathbf{h}$ apart from one another. This estimator is based on conditions of stationarity and ergodicity, which leads to the expression of C as a function of $\mathbf{h}$ alone. Weakening the requirement for these conditions yields the estimator,

$$\hat{C}(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i)z(\mathbf{x}_i+\mathbf{h}) - m[z(\mathbf{x}_i)]m[z(\mathbf{x}_i+\mathbf{h})]] \tag{3.7}$$

(Isaaks and Srivastava 1988), referred to as the 'non-ergodic spatial covariance'.

Correlogram The correlogram normalizes the spatial covariance to the variance of the random function and is defined as

$$\rho(\mathbf{h}) = \frac{C(\mathbf{h})}{C(\mathbf{0})} = \frac{C(\mathbf{h})}{Var(Z(\mathbf{x}))} \tag{3.8}$$

(Isaaks and Srivastava 1989). It has similar characteristics as the spatial covariance except that it has a value of 1 when $\mathbf{h} = 0$ and values between -1 and 1 elsewhere.

General cross-product statistic Haining (1990) describes the general cross-product statistic,

$$\Gamma = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} G_{ij} S_{ij} \tag{3.9}$$

where G_{ij} is a matrix describing the proximity of $\mathbf{x}_i$ and $\mathbf{x}_j$, S_{ij} is an expression describing the similarity of $Z(\mathbf{x}_i)$ and $Z(\mathbf{x}_j)$ and n_i and n_j are the number of $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. Moran's index (Moran 1948) and Geary's coefficient (Geary 1954) are examples of cross-product statistics that are regularly used components of geographical analysis (Griffith 1987). Moran's I is defined with $S_{ij} = \{z(\mathbf{x}) - m[z(\mathbf{x})]\}\{z(\mathbf{x}+\mathbf{h}) - m[z(\mathbf{x}+\mathbf{h})]\}$; Geary's c with $S_{ij} = \{z(\mathbf{x}) - z(\mathbf{x}+\mathbf{h})\}^2$. G_{ij} has been defined in a variety of ways, from the simplest definition using 1's when $\mathbf{x}$ and $\mathbf{x}+\mathbf{h}$ are neighbors and 0's when they are not, to a complex weighting scheme with weights proportional to $\mathbf{h}$. With a large amount of spatial dependence, Geary's c takes on small values (close to 0) while Moran's I takes on values close to 1. When neighboring values are less similar than those farther away, Geary's c becomes large and Morans' I becomes negative. This is a situation sometimes called 'negative spatial correlation' (Legendre 1993).

Cross-product statistics are used for hypothesis testing, where the null hypothesis is that no spatial autocorrelation exists. Expected values of the statistics under the null hypothesis are derived from, in essence, a purely random function with a normal distribution (Tiefelsdorf and Boots 1995).

Semivariogram Like Geary's c, the semivariogram is based on the increment $\{Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})\}$. The variance of this increment, or the expected value of the squared difference between every pair of values at the lag $\mathbf{h}_i$, is by convention twice the semivariance $\gamma(\mathbf{h}_i)$. The semivariogram (abbreviated to variogram in much of the literature), a function of $\mathbf{h}$, is therefore

$$\gamma(\mathbf{h}) = \frac{1}{2} E[Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})]^2 \tag{3.10}$$

The ordinary estimator of this statistic (Cressie 1991) is

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i) - z(\mathbf{x}_i+\mathbf{h})]^2 \tag{3.11}$$

where $N(\mathbf{h})$ is, as above, the number of pairs of data that are $\mathbf{h}$ apart from one another. While semivariograms calculated from data may take a variety of forms, very often they have asymptotic behavior, reaching a relatively constant plateau at some positive value (figure 3.1). This plateau is called the sill. The value of $\mathbf{h}$ at which this sill is reached is called the range.

Though the definition of the semivariogram implies that $\gamma(\mathbf{0}) = 0$, observed values at $\mathbf{h}$ close to $\mathbf{0}$ often indicate a positive intercept of the semivariogram. This intercept is called the 'nugget' variance, explained as including both measurement error (the variance that would be obtained by making repeated measurements in the same location) and variance at supports smaller than the measurement support (Isaaks and Srivastava 1989). By definition, a purely random function would have a 'flat' semivariogram, equal to the nugget variance for its entire range.

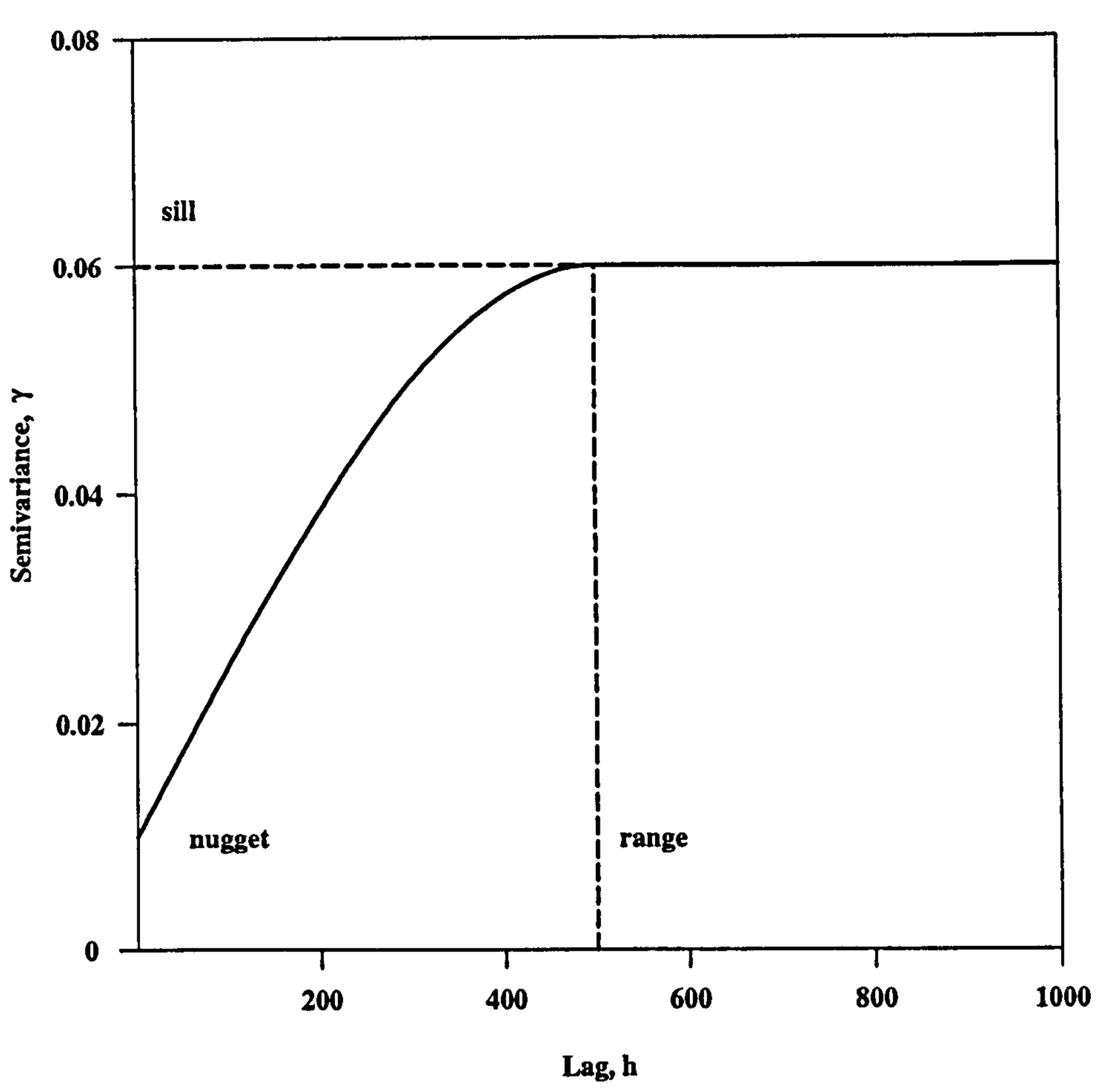
Dispersion variance A second-order statistic that is not expressed as a function of separation distance is called the dispersion variance,

$$D^2(v/V) = E\{\frac{1}{N}\sum_{i=1}^{N}[Z_V(\mathbf{x}) - Z_v(\mathbf{x}_i)]^2\} \tag{3.12}$$

(Journel and Huijbregts 1978), where Z_V is defined on the large support V, Z_v is defined on the smaller support v, v is within V and N is the number of units of support v (Journel and Huijbregts 1978). Dispersion variance is the same statistic advanced by Moellering and Tobler (1972) who first described 'scale variance analysis' for geographers. They recommended the investigation of $D^2(v/V)$ for a sequence of supports and implied the supports at which this variance was high would be the most useful to study. Scale variance analysis has subsequently been used in many instances with remotely sensed data (Townshend and Justice 1990, Justice *et al.* 1991, Belward 1992, Townshend and Justice 1995), though not recognized as the geostatistical dispersion variance.

Cross-semivariogram A realization of the joint distribution of two random functions, say $Z(\mathbf{x})$ and $W(\mathbf{x})$, is called a coregionalization. All of the above statistics have corresponding definitions for two random variables; for example the cross-covariance, cross-semivariogram and cross-correlogram. The cross-semivariogram is

Figure 3.1: Typical form of a semivariogram, showing the nugget, sill and range.



defined for two variables $Z(\mathbf{x})$ and $W(\mathbf{x})$ as

$$\gamma_{ZW}(\mathbf{h}) = \frac{1}{2}E\{[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})][W(\mathbf{x}+\mathbf{h}) - W(\mathbf{x})]^2\} \tag{3.13}$$

Unlike the semivariogram, which is strictly positive, the cross-semivariogram can take negative values indicating an inverse relationship between $Z(\mathbf{x})$ and $W(\mathbf{x})$. The normalized value of the cross-semivariogram, the cross-correlogram, has a value at $\mathbf{h} = \mathbf{0}$ equal to the Pearson correlation coefficient between $Z(\mathbf{x})$ and $W(\mathbf{x})$.

Though all of the above statistics are two-point statistics and as such their estimates all say something about the spatial dependence within the observed variable, each is unique and may yield different information. It is only when stationary, ergodic random fields are used as models of spatial phenomena that some of these statistics can be equated and used in methods of prediction.

3.2 Prediction methods

If the objective is to predict a biophysical variable, such as vegetation amount, across a region, the variable will be called the *primary variable* and can be modeled as a random function. Other variables used in the prediction of the primary variable, say spectral variables from remotely sensed measurements, will be called *ancillary variables*. These too can be considered random functions. The RFs are then used to construct prediction methods that have the highest possible accuracy. The commonly accepted definition of accuracy is a lack of error (Taylor and Kuyatt 1994), and in geographical applications (Johnston 1978, Olea 1990) it is taken to mean more specifically that

1. the error averages out to be insignificantly small – the prediction is unbiased and
2. the spread of those errors is also small – the prediction is precise.

Inaccurate predictions may therefore occur if they are biased though precise, unbiased and imprecise or, finally, biased and imprecise[4].

[4]Note that others (Maling 1989, Goovaerts 1997) equate accuracy directly with unbiasedness so that imprecise predictions, as long as they are unbiased, are declared accurate.

One way of achieving maximum precision is to minimize the variance of the error. Accuracy is further assured with a lack of bias; the expected value of the error is 0. Together, these criteria have come to define 'optimal' predictors (Cressie 1990). Though the term optimal seems to imply there can be no better predictors, there are in fact other criteria that can be used to define worthwhile methods (Srivastava 1987, Goovaerts and Journel 1995). Estimates of model parameters achieved through the optimality criteria are called Best Linear Unbiased Estimators (B.L.U.E.) and predictions of specific values are called Best Linear Unbiased Predictors (B.L.U.P.) (Goldberger 1962). Depending on the form of the random function model used as a premise for the prediction, different optimal predictors can be derived.

Regression methods constitute a large proportion of the B.L.U.P. class. The response variable is to be predicted using a linear or nonlinear combination of explanatory variables. In a spatial context, the response and explanatory may both be primary variables, but at different locations. This kind of regression is called kriging (Cressie 1990). The responses are the values at unsampled locations; the explanatory are the values at measured locations. Ancillary variables can be added to kriging, including explanatory variables at both sampled and unsampled locations, called cokriging. Regression without regard to location, aspatial regression, uses the primary variable as the response, ancillary as the explanatory.

Optimal predictors, while having maximum accuracy at each prediction location, lead to fields of predicted values that have no guarantee of representing the spatial variability of the actual or true field. A different category of methods called conditional simulation methods (Journel 1974, 1996) are designed on the same random function foundations but do not use the optimality criteria in the same way. Instead they aim to generate spatial fields of the primary variable that reproduce the global (or aggregate) statistics, both univariate and bivariate, including the second order moments described above. Aggregate characteristics can be particularly important when remotely sensed vegetation variables are used in physically-based process models (Band and Moore 1995).

Given the sparse data available for the primary variable and the many ways the aggregate statistics can be represented in a spatial field, there is no single solution for a conditional simulation. In fact, simulations are multiple alternative spatial fields,

each chosen with equal probability. Though B.L.U.P. and simulation methods have different objectives, both are considered and compared in this thesis. They can be considered to be prediction of values at individual locations and predictions of fields of values, respectively.

Since the prediction method follows directly from the choice of a random function model, how can this choice be evaluated? Christakos (1992) suggests that spatial prediction should be judged on its ability to incorporate six physical realities. The first is *measurement error*, which is inherent in any measurement and therefore should be taken into account in prediction. The second is *measurement support*, the shape and size of the area of each measurement. A difficult aspect of using remotely sensed data in combination with ground measurements is that the support of the latter is almost always smaller, sometimes substantially, than the support of image data (see chapter 2). Since statistics vary with support, this support mismatch should be considered in the prediction method. Ground measurement *locations* and their *geometry* are two other aspects of spatial data that should bear on results. To illustrate, imagine two sample sets with the same number of observations, one in which measurements are located in one small zone and another composed of a systematic arrangement throughout the region to be mapped. One would expect the first arrangement to result in a predicted field with much larger error, at least outside the sampled zone, than the second, despite the equivalent sample size. Lastly, *spatial variation* and *spatial correlation* are closely related properties of most physical attributes of the Earth's surface. Capturing these two properties of spatial variables is the defining characteristic of geostatistical predictors.

With every model for predicting values spatially there should be a model for the error arising from the prediction. In practice, prediction methods only give information on the variance of the error, or the precision of the prediction. Precision can be used to build models, both local and global, of uncertainty (Myers 1997, Goovaerts 1997). As discussed in chapter 1, a map of error is not feasible, so a map of uncertainty is a practical substitute. A map of predictions should be accompanied by a map of uncertainty to aid decisions about where additional measurements would be most useful or to estimate where risk is highest because of wrong answers. The uncertainty model for the chosen spatial predictor should reflect the six realities listed

above, because if predictions vary across space so do their accuracy. Ideally, a spatial uncertainty map should depend on both the values of nearby sample measurements *and* their distances and geometry (Switzer 1993).

As seen in chapter 2, aspatial regression has been the method most often used by investigators interested in predicting vegetation amount using remote sensed data. *Linear* regression models fitted with OLS, for example, assume (Johnston 1978):

1. a linear trend between response and explanatory variables,
2. normality of the errors,
3. unbiasedness of the errors for each value of the explanatory variable, (conditional unbiasedness),
4. homoscedasticity of the errors,
5. absence of autocorrelation in the errors and
6. lack of measurement error, at least in the explanatory variables.

Failure of the data to fit any of these six ideals has consequences for the usefulness of the model. Since spatial data is almost always autocorrelated, autocorrelation in the errors will typically be present unless the sources of spatial autocorrelation have been completely accounted for in the regression model. On theoretical grounds alone then, this reduces the appropriateness of aspatial regression models for the spatial prediction of vegetation quantities.

Unlike regression, geostatistical methods are designed for spatial data and take advantage of spatial autocorrelation. These methods include kriging (Cressie 1990) and conditional simulation (Journel and Alabert 1989). Cokriging, a kriging method in which more than one variable is used for prediction, has been shown to be particularly useful when measurements of the variable to be predicted are sparse and measurements of a second, related variable are plentiful or exhaustive (Leenaers *et al.* 1990, Goovaerts 1998*b*). For this reason, it seems ideally suited to remote sensing problems. Conditional simulation has a different goal to that of kriging. Simulation emphasizes the reproduction of the frequency distribution of the data

and their spatial pattern. Instead of a single resulting image with a value at each location that is in some sense 'optimal', simulation generates many images, each a possible representation of the variable conditional to the data. The resulting set of maps can be used as a model of uncertainty about the spatial distribution of the variable of interest.

Simulation has not yet been applied to the remote sensing of vegetation amount. Though cokriging and conditional simulation have been applied to many other environmental prediction problems, i.e. for soil and hydrologic quantities, contaminant concentrations and meteorological variables (Ahmed and de Marsily 1987, Fedorov 1989, Webster and Oliver 1990, Bierkens and Burrough 1993, Cassiani and Medina 1997, Goovaerts 1999), cokriging has been used rarely (Bhatti *et al.* 1991, Atkinson *et al.* 1994, Carroll *et al.* 1995, Lohani and Mason 1999, Ishida and Ando 1999) and conditional simulation even less (Dungan *et al.* 1994, van der Meer 1994, Dungan 1998) in remote sensing applications (Curran and Atkinson 1998).

In the next three sections, the derivation of each method from forms of the random function model is given.

3.2.1 Aspatial Regression

The simplest and most often used form of linear regression as applied to spatial problems has the following form:

$$\mathbf{p} = \mathbf{a}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{3.14}$$

where $\mathbf{p}$ is a vector $[p(\mathbf{x}_1), p(\mathbf{x}_2), p(\mathbf{x}_3), \cdots, p(\mathbf{x}_{N_1})]'$ of N_1 primary variable observations, $\mathbf{a}$ is a matrix of N_1 observations on m ancillary variables,

$$\begin{bmatrix} 1 & a_1(\mathbf{x}_1) & a_2(\mathbf{x}_1) & \cdots & a_m(\mathbf{x}_1) \\ 1 & a_1(\mathbf{x}_2) & a_2(\mathbf{x}_2) & \cdots & a_m(\mathbf{x}_2) \\ \cdots & \cdots & \cdots & \cdots & \\ 1 & a_1(\mathbf{x}_{N_1}) & a_2(\mathbf{x}_{N_1}) & \cdots & a_m(\mathbf{x}_{N_1}) \end{bmatrix} \tag{3.15}$$

$\boldsymbol{\beta}$ is a vector of coefficients $[\beta_0, \beta_1, \ldots, \beta_m]$ and $\boldsymbol{\epsilon}$ is a vector of errors (Johnson and Wichern 1982). Note that the $\boldsymbol{\beta}$ coefficients are not functions of $\mathbf{x}$. To apply

this model to vegetation amount prediction, p must be known at some locations $\mathbf{x}_s$ (from a ground measurement sample) and the a must have been measured at the coincident (collocated) locations. This method is usually explained with p and a as random variables (Johnson and Wichern 1982). When applied to spatial data, $\mathbf{p}$ and $\mathbf{a}$ can be considered random functions.

The optimality of regression is defined by the unbiasedness and minimum error variance criteria. The first criterion is

$$E(\epsilon) = \mathbf{0} \tag{3.16}$$

and the second is to minimize

$$Var(\epsilon) = E(\epsilon'\epsilon) = \mathbf{\Sigma} \tag{3.17}$$

Ordinary least squares (OLS) regression assumes independent identically distributed errors in the response variable. (Johnston 1978). This means that

$$\mathbf{\Sigma} = \sigma^2\mathbf{I} \tag{3.18}$$

where $\mathbf{I}$ is the identity matrix and σ^2 the variance of the distribution of errors (Johnson and Wichern 1982). In a spatial context, ϵ can be considered to be a function of $\mathbf{x}$, $[\epsilon(\mathbf{x}_1), \epsilon(\mathbf{x}_2), \ldots, \epsilon(\mathbf{x}_{N_1})]$. This implies that $\epsilon(\mathbf{x})$ must be a purely random function.

To solve for the coefficients β when this variance of the errors is at a minimum,

$$\begin{aligned} \mathbf{\Sigma} &= Var(\mathbf{p} - \mathbf{a}\beta) \\ &= (\mathbf{p} - \mathbf{a}\beta)'(\mathbf{p} - \mathbf{a}\beta) \\ &= (\mathbf{p}' - \beta'\mathbf{a}')(\mathbf{p} - \mathbf{a}\beta) \\ &= \mathbf{p}'\mathbf{p} - 2\mathbf{a}'\mathbf{p}\beta + \mathbf{a}'\mathbf{a}\beta^2 \end{aligned} \tag{3.19}$$

the derivative of Σ with respect to β is set to 0:

$$\begin{aligned} \frac{\delta\Sigma}{\delta\beta} &= 0 = 2\mathbf{a}'\mathbf{p} - 2\mathbf{a}'\mathbf{a}\beta \\ 2\mathbf{a}'\mathbf{p} &= 2\mathbf{a}'\mathbf{a}\beta \\ \beta &= (\mathbf{a}'\mathbf{a})^{-1}\mathbf{a}'\mathbf{p} \end{aligned} \tag{3.20}$$

This predictor is unbiased (expected value of the error is 0), since

$$\begin{aligned} \hat{\epsilon} &= \mathbf{p} - \hat{\mathbf{p}} \\ &= \mathbf{p} - \mathbf{a}\beta \\ &= (\mathbf{I} - \mathbf{a}(\mathbf{a}'\mathbf{a})^{-1}\mathbf{a}')\mathbf{p} \\ &= (\mathbf{I} - \mathbf{a}(\mathbf{a})^{-1}(\mathbf{a}')^{-1}\mathbf{a}')\mathbf{p} \\ &= (\mathbf{I} - \mathbf{I})\mathbf{p} = 0 \end{aligned} \tag{3.21}$$

To put this model into practice, a sample of ground measurements must be collected and associated to the corresponding pixels from an image or images. The β values are solved for and multiplied with a at every pixel in the image, resulting in an image representing $\mathbf{p}$. This amounts to a simple rescaling of spectral values to vegetation amount values, a 'transformation from feature space [the relation between a and p] to geographical space' (page 1035, Steven 1987). This points to the major advantages of this method, its simplicity in concept and implementation.

One property of spatial predictions made with aspatial regression is that at measurement locations $p(\mathbf{x}_s)$, the predictions will not in general be equal to the measured values. Predictors that replicate measured values at those locations are called 'exact' predictors (Journel 1990) – therefore aspatial regression is inexact. It implies that the error at measurement locations is not treated any differently than error at unsampled locations. The variance of the error depends on the magnitude of the prediction in relation to the mean only and is not dependent on spatial location. In the case of $m = 1$, the standard model of the variance of a local prediction at $\mathbf{x}_i$ is given by:

$$Var(\hat{p}(\mathbf{x}_i)) = \sigma_p^2 + \frac{\sigma_p^2}{n} + \frac{[a(\mathbf{x}_i) - m[a(\mathbf{x}_s)]]^2\sigma_p^2}{\sum[a(\mathbf{x}_s) - m[a(\mathbf{x}_s)]]^2} \tag{3.22}$$

(Draper and Smith 1998) where σ_p^2 is estimated using the variance of the ground measurements and n is the number of sample measurements. $Var(\hat{p}(\mathbf{x}_i))$ is some-

times referred to as the standard error of prediction (SEP), in contrast to the standard error of the estimate (SEE). If the purely random function of the ϵ is further assumed to be normally distributed, confidence limits can be constructed as functions of equation 3.22. These confidence limits are the standard quantification of uncertainty from this method (Draper and Smith 1998).

As seen in chapter 2, different interpretations are available for whether spectral data should be considered the response or explanatory variable. If the OLS system described above is used, prediction takes into account only error in the response (vegetation amount) measurements but not that of the remotely sensed measurements. In Curran and Hay's (1986) criticism of this aspect of the OLS approach, other solutions to the system, such as the 'reduced major axis method' (Jones 1937) which takes into account measurement error in the a values, are recommended. However, such alternatives have not become routine (Todd *et al.* 1998, Salvador and Pons 1998).

The aspatial regression method inherently does not include spatial support in its formulation. That is, the purely random function underlying the method does not require that the support of the ground measurements be similar to the support of the remotely sensed measurements. Therefore, the method does not preclude a regression equation developed from a set of ground measurements and spectral measurements made from one sensor to be directly applied to spectral measurements from another sensor (or the same sensor at a different flying height). Yet support mismatches are almost universally present with these types of measurements and support effects do occur in remotely sensed data as shown by scale variance analysis (Townshend and Justice 1995), examination of the semivariogram (Atkinson 1993, McGwire *et al.* 1993, Collins and Woodcock 1999) and other image statistics (Wickham and Riitters 1995). So the aspatial regression method does not obviously offer an associated model for 'scaling up' the ground measurements to adjust for the support of the remotely sensed measurements (Atkinson 1997*b*, Raffy and Gregoire 1998).

Except for the requirement that each $a(\mathbf{x}_i)$ and $p(\mathbf{x}_i)$ pair refers to the same location, this application of regression is completely aspatial. Sample values near to one another are assumed to be equally informative as distant sample values; sample

geometry and relative measurement locations are ignored. Since spatial location does not enter into the parameters of the regression model, measurement location and geometry cannot feature directly in regression predictions. This means that predictions (and uncertainty about predictions) do not change in areas of sparse or no measurement locations relative to areas of plentiful measurements.

If the predicted map is a linear transform of the remotely sensed data, its spatial pattern is usually similar to that of the remotely sensed data though spatial pattern information is not used directly in the model. A nonlinear transform would likely generate a different spatial pattern. Information about the actual spatial pattern of the primary variable is not included in the model.

3.2.2 Cokriging

Rather than using purely random functions, using intrinsic or other random functions that include spatial dependence leads to a group of geostatistical predictors called kriging predictors (Journel and Huijbregts 1978). The most elementary forms of kriging use only primary variables, using values at measurement locations to predict those at unsampled locations. Multivariable kriging, called cokriging, adds ancillary variable or variables to the prediction equations.

By adding an ancillary variable obtained using remote sensing to a kriging analysis, there exists the potential to increase the precision of that interpolation with cokriging (Myers 1983). The availability of a complete grid of these ancillary data allows for a thorough description of their spatial autocorrelation. In cokriging, it is assumed that the spatial pattern of the primary and ancillary variable are related (they co-vary). The ordinary[5] cokriging predictor at an unsampled location is

$$\hat{p}(\mathbf{x}_o) = \sum_{j=1}^{N_1} \lambda_j p(\mathbf{x}_j) + \sum_{k=1}^{N_2} \omega_k a(\mathbf{x}_k) \tag{3.23}$$

where $\hat{p}(\mathbf{x}_o)$ is the prediction at location $\mathbf{x}_o$, $p(\mathbf{x}_j)$ are N_1 nearby sample values at locations $\mathbf{x}_j$ weighted with factors λ_j and $a(\mathbf{x}_k)$ are N_2 nearby ancillary image values at locations $\mathbf{x}_k$ weighted with factors ω_k (Myers 1982).

[5]Only the 'ordinary' forms of kriging, rather than 'simple' forms are discussed here.

The error ϵ therefore is

$$\epsilon = p(\mathbf{x}_o) - \{\sum_{j=1}^{N_1} \lambda_j p(\mathbf{x}_j) + \sum_{k=1}^{N_2} \omega_k a(\mathbf{x}_k)\} \tag{3.24}$$

and the variance of the error (also called kriging variance or precision) is

$$\begin{aligned} Var(\epsilon) = \Sigma &= Var\{p(\mathbf{x}_o) - \textstyle\sum_{j=1}^{N_1} \lambda_j p(\mathbf{x}_j) + \sum_{k=1}^{N_2} \omega_k a(\mathbf{x}_k)\} \\ &= Var\{p(\mathbf{x}_o)\} - 2\sum \lambda_j C\{p(\mathbf{x}_j)p(\mathbf{x}_o)\} - 2\sum \omega_k C\{a(\mathbf{x}_j)p(\mathbf{x}_o)\} + \\ &\quad 2\sum\sum \lambda_j \omega_k C\{a(\mathbf{x}_k)p(\mathbf{x}_j)\} + \sum\sum \lambda_j \lambda_k C\{p(\mathbf{x}_j)p(\mathbf{x}_k)\} + \\ &\quad \sum\sum \omega_j \omega_k C\{a(\mathbf{x}_j)a(\mathbf{x}_k)\} \end{aligned} \tag{3.25}$$

Since cokriging is an optimal predictor, the unbiasedness and minimum error variance constraints are used. For unbiasedness, look at the expected value of the prediction:

$$\begin{aligned} E\{\hat{p}(\mathbf{x}_o)\} &= E\{\textstyle\sum_{j=1}^{N_1} \lambda_j p(\mathbf{x}_j) + \sum_{k=1}^{N_2} \omega_k a(\mathbf{x}_k)\} \\ &= \textstyle\sum_{j=1}^{N_1} \lambda_j E\{p(\mathbf{x}_j)\} + \sum_{k=1}^{N_2} \omega_k E\{a(\mathbf{x}_k)\} \end{aligned} \tag{3.26}$$

This constraint implies any number of possibilities for the weights. The most obvious is

$$\begin{aligned} \textstyle\sum_{j=1}^{n} \lambda_j &= 1 \\ \textstyle\sum_{k=1}^{m} \omega_k &= 0 \end{aligned} \tag{3.27}$$

When the mean values of the primary and ancillary data are known, or can be estimated accurately, a slight modification of the cokriging equations yields the constraints:

$$\sum_{j=1}^{n} \lambda_j + \sum_{k=1}^{m} \omega_k = 1 \tag{3.28}$$

(Goovaerts 1998*b*).

For maximum precision, the first derivatives of Σ are set to 0. To solve these equations for the weights, the method of Lagrange multipliers is used (Kitanidis 1997). This is a mathematical procedure for constrained minimization, involving

two parameters, μ_o and μ_1. The equations become, for the constraints 3.27,

$$\begin{aligned}
\frac{\delta\Sigma}{\delta\lambda} = 0 &= 2\sum\lambda_j C\{p(\mathbf{x}_j)p(\mathbf{x}_k\} + 2\sum\omega_k C\{a(\mathbf{x}_k)p(\mathbf{x}_j)\} - \\
&\quad 2C\{p(\mathbf{x}_o)p(\mathbf{x}_k)\} + 2\mu_0 \\
\frac{\delta\Sigma}{\delta\omega} = 0 &= 2\sum\lambda_j C\{p(\mathbf{x}_j)a(\mathbf{x}_k\} + 2\sum\omega_k C\{a(\mathbf{x}_j)a(\mathbf{x}_k)\} - \\
&\quad 2C\{p(\mathbf{x}_o)a(\mathbf{x}_k)\} + 2\mu_1 \\
\frac{\delta\Sigma}{\delta\mu_0} = 0 &= 2\sum\lambda_j - 1 \\
\frac{\delta\Sigma}{\delta\mu_1} = 0 &= 2\sum\omega_k
\end{aligned} \tag{3.29}$$

The resulting normal equations in matrix form (Myers 1983) are:

$$\begin{bmatrix} \mathbf{C}_{pp} & \mathbf{C}_{pa} & \mathbf{I} \\ \mathbf{C}_{ap} & \mathbf{C}_{aa} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Lambda \\ \Omega \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{px_o} \\ \mathbf{C}_{ax_o} \\ \mathbf{I} \end{bmatrix} \tag{3.30}$$

$$\mathbf{Cw} = \mathbf{D}$$

where $\mathbf{C}_{pp}$ is the spatial covariance of the vegetation variable, $\mathbf{C}_{aa}$ is the spatial covariance of the remotely sensed variable and $\mathbf{C}_{pa}$ is their cross-covariance. $\mathbf{C}_{px_o}$ is the spatial covariance for the primary variable at a measurement location and at an unsampled location. $\mathbf{C}_{ax_o}$ is the spatial cross-covariance for the ancillary variable at a measurement location and the primary variable at an unsampled location. The vector of weights on the vegetation variable is Λ or $[\lambda_1, \lambda_2, \ldots, \lambda_j]'$; Ω is the vector of weights $[\omega_1, \omega_2, \ldots, \omega_j]'$ on the ancillary variable, μ is the vector of Lagrange multipliers and $\mathbf{I}$ is the identity vector or matrix. The solution is obtained by inverting the $\mathbf{C}$ matrix and multiplying it by $\mathbf{D}$ to get the weights. The weights can be seen as solely a function of spatial covariances, allowing the use of expressions of spatial dependence in the solution.

To use cokriging, an adequate number of collocated primary and ancillary measurements are needed to obtain estimates of the spatial covariances and cross-covariances (via equation 3.6 or other estimators of covariance (Isaaks and Srivastava 1988)). Alternatively, semivariograms and cross-semivariograms are estimated and equation 3.2 is used to obtain the equivalent covariance under the intrinsic random function model. Estimates of covariance from the data are not sufficient however,

since these only give $\mathbf{C}_{pp}$, $\mathbf{C}_{aa}$, $\mathbf{C}_{ap}$ and $\mathbf{C}_{pa}$. To obtain $\mathbf{C}_{px_o}$ and $\mathbf{C}_{ax_o}$, values of covariance at distances besides those defined by pairs of data are needed. This is done with a so-called 'model of coregionalization' obtained by fitting certain types of functions through the experimentally derived estimates of covariance. Strict limitations are placed on this coregionalization model (Isaaks and Srivastava 1989, Goovaerts 1997) because mathematically it must ensure a positive cokriging variance.

To construct the coregionalization model, the usual practice (Isaaks and Srivastava 1989, Goovaerts 1997) is to use a linear combination of covariance models of L independent random variables. For a single primary variable and one ancillary variable, the set of spatial covariances ($C_{11}(\mathbf{h})$ and $C_{22}(\mathbf{h})$) and cross-covariances ($C_{12}(\mathbf{h})$ and $C_{21}(\mathbf{h})$) to be modeled can be represented as $C_{ij}(\mathbf{h})$:

$$C_{ij}(\mathbf{h}) = \sum_{l=1}^{L} b_{ij}^{l} c^{l}(\mathbf{h}) \; \forall \; i, j \tag{3.31}$$

where $c^{l}(\mathbf{h})$ is an individual basic covariance model and b_{ij}^{l} is the sill of this model. For for every individual model l, the matrix of sills

$$\begin{bmatrix} b_{11}^{l} & b_{21}^{l} \\ b_{12}^{l} & b_{22}^{l} \end{bmatrix} \tag{3.32}$$

must be positive semi-definite, which implies that $b_{11}^{l} \geq 0$, $b_{22}^{l} \geq 0$ and (by Schwarz's inequality)

$$|b_{12}^{l}| = |b_{21}^{l}| \leq \sqrt{b_{11}^{l} b_{22}^{l}} \tag{3.33}$$

This way of achieving positive semi-definiteness entails that every basic model l that appears in a cross-covariance model $C_{12}(\mathbf{h})$ must also appear in the two covariance models $C_{11}(\mathbf{h})$ and $C_{22}(\mathbf{h})$. Conversely, all covariance models comprising $C_{11}(\mathbf{h})$ and/or $C_{22}(\mathbf{h})$ need not be in $C_{12}(\mathbf{h})$. The b_{11}^{l} and b_{22}^{l} coefficients must be non-negative and the rest must be checked for condition 3.33. The modeling of the coregionalization is done in a similar way with semivariograms, in which case linear combinations of basic models such as spherical, exponential and Gaussian expressions (Deutsch and Journel 1998) will satisfy the positive semi-definiteness criterion. For this linear model of coregionalization (LMC), the exhaustive grid of remote mea-

surements can be used to model the spatial covariance of the ancillary variable, but only collocated measurements can be used for the cross-covariance modeling.

Two caveats have been generally recognized in cokriging methods (Goovaerts 1997): 1) it is difficult to find an LMC that is a good fit to experimental semivariograms and cross-semivariograms and 2) a much denser grid of data for the ancillary variable than that for the primary variable can lead to the instability of the cokriging matrix system. 'Collocated' cokriging is an approximation of the cokriging method that addresses these two problems (Almeida and Journel 1994, Goovaerts 1997, Deutsch and Journel 1998). At each grid cell to be predicted, collocated cokriging uses only the ancillary datum at that cell rather than all neighboring ancillary data (i.e. only 'collocated' ancillary data).

As in the case of aspatial regression, it is only the variance of the error (equation 3.25), not the error itself, that can be calculated from the model. Values of kriging variance (frequently called 'estimation variance') are often displayed in map form. Switzer (1993) and Webster and Oliver (1990) have stated that this error model is the main contribution of geostatistics. But, as acknowledged by Switzer (1993) and demonstrated by Journel and Rossi (1989) and Rossi *et al.* (1994) a kriging variance map does not provide a useful *spatial* description of error and is a useful description of local precision only in ideal, multi-Gaussian circumstances. Kriging variance does not depend on the magnitude of variable values, only their location and geometry. Therefore, the precision of the prediction can be quantified from the model as soon as a sample network is designed (Journel 1990).

Measurement error in the primary and ancillary variables is captured within the 'nugget variance' terms of the model of coregionalization. This is the variance that is spatially independent, but is not separated in this model from the variation occurring at smaller supports than those measured (Atkinson *et al.* 1996). This nugget variance can be filtered (Carr 1990, Bourgault 1994) in an attempt to reduce the influence of noise on the predictions. Unless filtered, kriging predictors guarantee that at measurement locations they return the measured value of the vegetation variable. So they are 'exact' unlike aspatial regression.

Models are available in kriging theory to address a difference in support between primary data and the support of predictions. This change of support can be

accomplished by block kriging, whereby values at one support can be used along with estimates of point-to-block semivariograms or spatial covariances to predict values at a larger support. The case of block cokriging also relies on point-to-block cross-semivariograms, which are estimated from point-to- point cross-semivariograms. This procedure implies that the supports of direct and ancillary data should be equivalent. When they are not, as in most applications of remote sensing data, scaling up may be precluded or alternatives to block cokriging must be found.

B.L.U.P interpolators, like cokriging, all have the property of being 'smoothers' (Journel 1990). This means that the global distribution of all values in a cokriged map will have a smaller variance than the actual variable and the map itself will have less overall variability. Paradoxically, semivariograms or spatial covariances carefully modeled from the data and supposed to represent the spatial continuity of the actual phenomenon will not be similar to the semivariograms calculated from the cokriged map. Journel (1990, page 31) goes as far as to say, 'In all rigor, estimates based on a local accuracy criterion such as kriging should only be tabulated; to map them is to entice the user to read on the map patterns of spatial continuity that may be artifacts of the data configuration and the kriging smoothing effect.' Another paradox is that cokriged maps will show the most continuity where data values are sparsest – leading to the often false impression that these regions are less variable than regions where sampling was more dense. Simulation methods, described in the next section, are in contrast designed to describe spatial variability of the variable being studied.

3.2.3 Conditional Simulation

Aspatial regression and cokriging provide a *single* spatial field of predictions, one prediction for each unsampled location or grid cell. In contrast, conditional simulation methods provide *multiple* alternative spatial fields. Conditional simulation is the generation of synthetic realizations of a random function that preserve the sample values at their measured locations and possess the same spatial statistics (moments up to the second order) as the data that have been collected about a variable (Olea 1990). The preservation of the sample values in the realizations is

what the modifier 'conditional' refers to, so the realizations are exact. Conditional simulation is an appropriate method for mapping, because it emphasizes the global attributes of a spatial field, its histogram and spatial pattern (Journel 1974, 1996).

There are many algorithms devised to produce conditional simulations, including sequential methods (Journel and Alabert 1989), methods in the frequency domain (Friedl *et al.* 1995), through matrix decompositions (Alabert 1987), transforming two or three dimensions to a single dimension (Christakos and Panagopoulos 1992), simulated annealing (Deutsch and Cockerham 1994) and probability field simulation (Srivastava 1992). Each is based on one of three general strategies (Dungan 1999). The first strategy is to add the missing spatial variability back in to a prediction made by kriging. The second strategy is to construct the simulation from Bayesian first principles. The third is to use optimization techniques to converge on solutions that meet statistical criteria. The number of algorithms is still growing and at present, no single text or software package covers all published algorithms. Conditional simulation is discussed in Journel and Huijbregts (1978), Cressie (1991), Christakos and Panagopoulos (1992), Goovaerts (1997), Kitanidis (1997) and Deutsch and Journel (1998) and is implemented in the software packages GSLIB (Stanford University, Stanford, CA), Gstat (University of Amsterdam) and Isatis (Geovariances, Avon, France), among others.

The original method for conditional simulation was one based on the strategy of adding variability 'back in' to smooth predicted maps called the turning bands algorithm (Journel 1974). Simulations were done in one dimension for vectors (the bands) that were then moved (turned) through the 2D or 3D field. Because of its inflexibility (Dowd 1992), turning bands has been largely discarded in favor of other algorithms. Sequential simulation, based on a Bayesian strategy, is described herein.

Sequential simulation

Many of the commonly used algorithms developed after turning bands are variants of sequential simulation (Journel and Alabert 1989). The sequential approach exploits Bayes Theorem, which describes how the prior (marginal) probabilities for a random variable can be revised (or updated) to reflect additional information contributed from nearby (and therefore dependent) data to create a posterior, joint

probability. According to the theorem, the joint probability of A_1 and A_2 is equal to the probability that event A_1 will occur given A_2 has occurred multiplied by the probability of A_2:

$$Pr(A_1 \bigcap A_2) = Pr(A_2)Pr(A_1|A_2) \tag{3.34}$$

This can be extended to m events as follows:

$$\begin{aligned} Pr(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_m) &= Pr(A_m|A_1 \cap A_2 \cap A_3 \ldots \cap A_{m-1}) \\ & \quad Pr(A_1 \cap A_2 \cap \ldots \cap A_{m-1}) \\ &= Pr(A_m|A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_{m-1}) \\ & \quad Pr(A_{m-1}|A_1 \cap A_2 \cap A_3 \ldots \cap A_{m-2}) \\ & \quad Pr(A_1 \cap A_2 \cap A_3 \ldots \cap A_{m-2}) \\ & \quad \vdots \\ &= Pr(A_m|A_1 \cap A_2 \cap A_3 \ldots \cap A_{m-1}) \\ & \quad Pr(A_{m-1}|A_1 \cap A_2 \cap A_3 \ldots \cap A_{m-2}) \\ & \quad Pr(A_1 \cap A_2 \cap A_3 \ldots \cap A_{m-2}) \\ & \quad \cdots Pr(A_2|A_1) \ldots Pr(A_1) \end{aligned} \tag{3.35}$$

(Isaaks 1990). For conditional simulation, the objective is to determine the joint ccdf of all random variables within the random function given all the n data, that is,

$$Pr(Z(\mathbf{x}_{n+1}) \leq z, Z(\mathbf{x}_{n+2}) \leq z, \ldots, Z(\mathbf{x}_{n+m}) \leq z|Z(\mathbf{x}_1), Z(\mathbf{x}_2), Z(\mathbf{x}_3), \ldots, Z(\mathbf{x}_n)) \tag{3.36}$$

To draw from this conditional joint distribution to create realizations, sequential algorithms do a step-by-step updating of the ccdfs (Journel 1990). The first ccdf, $Pr(Z(\mathbf{x}_{n+1}) \leq z)$, is obtained through kriging at a location, with the conditional mean defined by the kriged value and the conditional variance defined by the kriging variance. It is drawn from to obtain a simulated value for that location. This simulated value is added to the pool of nearby data to provide data to the kriging system at the next location, $Pr(Z(\mathbf{x}_{n+1}) \leq z|Z(x) = z)$. At each step, the simulated value is added to the pool of data and previously simulated values and steps continue until all cells in the grid have a value. The locations in the grid are visited in a random order. Each new realization is obtained by going through the

procedure beginning from a different point in the simulation grid and drawing from the ccdfs randomly. The number of simulated cells, s, is much larger than the number of data, n. Because very close simulated cells screen data farther away (Isaaks and Srivastava 1989), the conditioning data and ccdf are taken from a neighborhood surrounding the node to be simulated (referred to as the search neighborhood). The use of a search neighborhood (instead of all the data) to define the conditional cdfs has two advantages. It prevents the need to solve very large $(n+s)(n+s)$ kriging systems and it reduces the impact that the lack of second-order stationarity will have on the results (Deutsch and Journel 1998).

When the kriging algorithm used to build the ccdfs is simple kriging, each ccdf is defined by the simple kriging mean and variance (Journel 1990). Ordinary kriging is an approximation to simple kriging and may be used in its place. Using simple or ordinary kriging approaches result in what has been called sequential Gaussian simulation, a parametric approach (Isaaks 1990). Data that do not obviously fit a multinormal model are transformed to normal scores before applying the simulation algorithm. Realizations are then transformed back to the data distribution, though this does not ensure that a multinormal model will be successful. Conditional cdfs may also be obtained through indicator kriging, IK (Journel 1990). This is a direct modeling of the ccdf without resort to any kriging variance or Gaussian hypothesis. Since the conditional expectation of the indicator variable identifies the ccdf value at threshold z_k,

$$I(x;z_k) = Pr(Z(x) \leq z_k|(n)) \qquad (3.37)$$

where $I(x;z_k)$ is an indicator random variable at location x, an estimate of that indicator can be used to build a model for the z ccdf. Continuous ccdfs are effectively discretized by dividing the values of the primary variable into classes and transforming each class (k) to an indicator variable. Using the IK approach has been called sequential indicator simulation and can provide non-Gaussian ccdfs (Journel and Isaaks 1984). Journel and Alabert (1989) describe the latter approach using an example in which the exhaustive 'true' image is known, so that simulated realizations can be compared objectively.

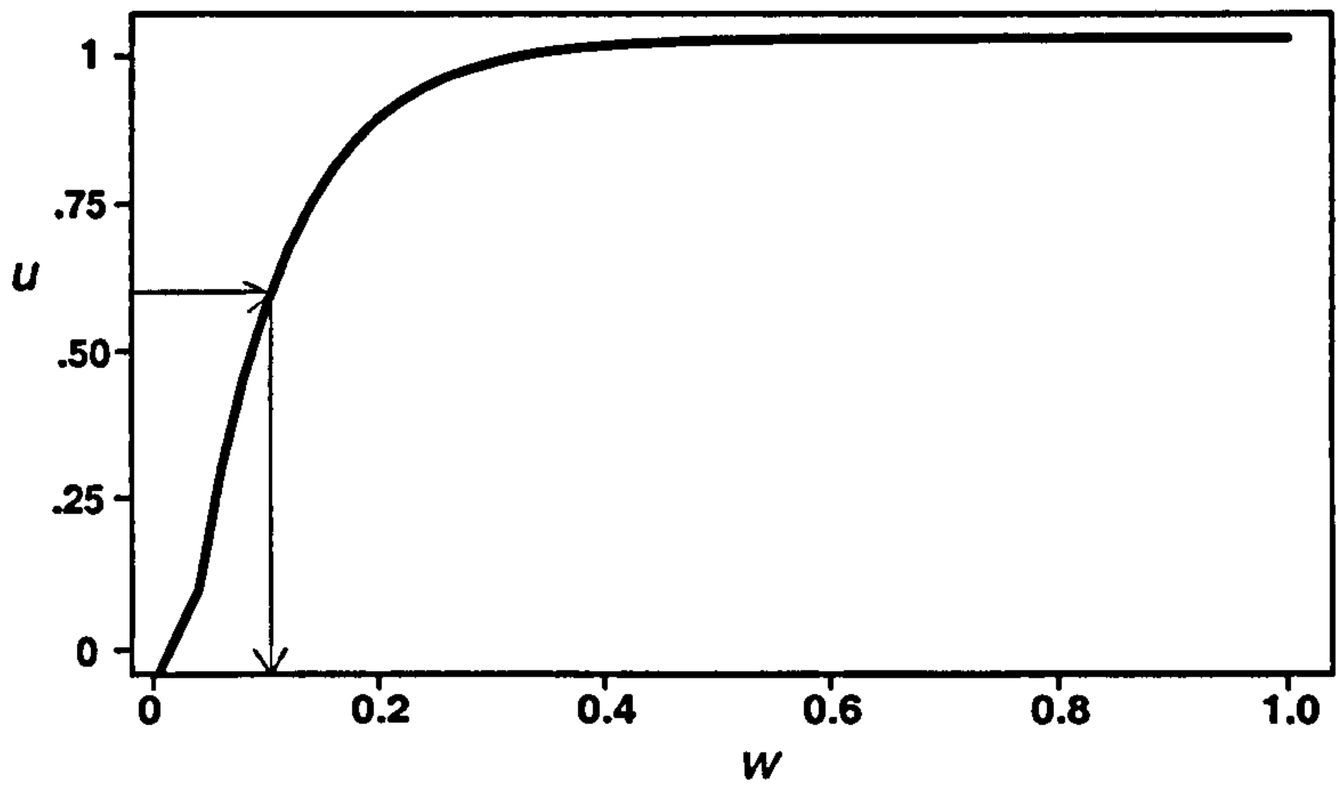
Probability field simulation

An algorithm closely related to sequential simulation is called probability field simulation, first conceived by Srivastava (1992) and Froidevaux (1993). It is conceptually simpler and less computationally taxing than sequential methods. It represents a shortcut to these more rigorously defined methods; work remains to be done on the relationship of probability field simulation to the random function model, $Z(x)$.

Probability field simulation separates the steps of creating the local ccdfs and drawing from them. The ccdfs at every grid cell are defined in the first step using any algorithm, be it simple kriging, ordinary kriging, indicator kriging etc. In fact, distributions generated from regression using equation 3.22 or similar error model could also be used (Srivastava 1992), though these are not conditioned to nearby data. The second step is the selection of a value from every local ccdf to generate a realization. A probability field is a set (image) of values uniformly distributed between 0 and 1 used to draw from the field of ccdfs. A constant, flat probability field where all values are equal to 0.5 would result in the median value from each ccdf being selected at each location. The resulting realization would be smooth, like a kriged map, and not reflect the spatial pattern of the variable. To obtain realizations that represent the spatial pattern, probability fields must have a spatial pattern of their own. This pattern is characterized by the semivariogram of the uniform scores of the data (which have values between 0 and 1). Any simulation algorithm that can generate unconditional simulations can be used to create probability fields that reproduce this semivariogram (Deutsch and Journel 1998). Any number of unconditional simulations can be produced in order to generate the probability fields; each is used to draw from the z ccdfs and yield a realization.

Drawing from ccdfs within the sequential or probability field algorithms employs the so-called quantile algorithm (Hogg and Tanis 1983). The quantile algorithm exploits the theorem that if U has a uniform distribution between 0 and 1, and $F(W)$ is a permissible cdf, strictly increasing between 0 and 1, then the random variable W defined by $W = F^{-1}(U)$ is a continuous type random variable with distribution function $F(W)$. Therefore, U can be used in the inverse of the local ccdf to draw from that ccdf (figure 3.2).

Figure 3.2: Illustration of how the quantile algorithm uses values of u, which are uniformly distributed, to draw values of w, which are distributed according to the cdf plotted here.



Adding ancillary information to a simulation

To this point, only algorithms for univariate conditional simulation have been discussed. To add information from ancillary variables, in particular from dense grids of spectral data, realizations must somehow be further conditioned to take into account this ancillary information. If the ancillary data are informative, they should act to reduce the variance of ccdfs modeled at each grid cell.

With two of the simulation strategies, adding back variability and the Bayesian approach, multivariable estimation (cokriging) lends itself to conditioning the ccdfs. So, for example, either sequential Gaussian simulation or probability field simulation could be implemented using ccdfs defined using cokriging estimation variances. If cokriging variances are used, simulations should show less variation from realization to realization than those using just the primary variable approach given the same data, since as Myers (1983) showed the cokriging variances are always less than kriging variances.

There are several advantages of conditional simulation methods, like probability field simulation, as spatial field predictors. If local ccdfs are generated geostatistically, such as by cokriging, the method incorporates measurement error, sample geometry and relative locations of the measurements in a similar way to that algorithm. It recognizes spatial correlation as an important characteristic of the predictions. Simulations are generally more spatially variable than are the results

of other prediction methods – the occurrence of high and low values is more likely. Its main disadvantage is its lower local accuracy relative to B.L.U.P. methods; it is non-optimal.

The mean of all realizations generated by a particular conditional simulation algorithm, calculated by estimating the mean at each grid cell across realizations, is expected to be more locally accurate than any individual realization. This mean realization, one of a class of 'E-type estimates' (with 'E' for expected value, Journel 1990), is similar to a kriged map and has the same smoothing properties. A variance map of realizations, or another statistic of spread such as the inter-quartile range (Hoaglin *et al.* 1983), is one summary of uncertainty described by the conditional simulation method.

3.3 Relevant properties of optical remote sensing data

Images collected by satellite or airborne sensors generate data with special characteristics that bear on the theoretical and practical reasons for selecting a random function model and an associated method for predicting the values of a primary variable. For example, the sizes of remotely sensed images are substantial, with most images typically containing hundreds of thousands or millions of pixels. Linear or nonlinear rescaling of large grids does not at present pose any computational challenge. However, the much more computationally intensive geostatistical methods are less practical to apply to such large grids. These methods have been developed on small grids (100-300 × 100-300) and their use on large grids is limited. Of some compensating benefit, the relatively consistent spacing of image pixels allows one-time calculation of distances for spatial dependence calculations, reducing the computational burden.

With the existence of large grids from remote sensing comes an inevitable diminishment in the number of ground measurements *relative* to the number of image measurements. So, although the absolute number of ground measurements within a sample has been discussed in the literature (Curran and Williamson 1986), the density of measurements, what Maling (1989) calls the sampling fraction, is as important, at least for geostatistical methods. These methods rely on a sufficient

number of nearby sample measurements where each prediction will be made. Some illustrative sampling fractions from the geostatistical literature are 0.6% (Isaaks and Srivastava 1989), 1% (Journel and Alabert 1989), 7.5% (Olea and Pawlowsky 1996) and 1–12% (Chainey and Stuart 1998). However, in remote sensing studies the fraction may be an order or more of magnitude smaller. A study of sampling fraction could provide information to create heuristics for the choice of models and methods.

The effective spatial resolution (Forshaw *et al.* 1983) brought about by the sensor and platform characteristics, point spread function and atmosphere is the support of image data. Though this value may not be known very well (Forshaw *et al.* 1983, Fisher 1997), it can at least be approximated. It is a relatively consistent value across any given scene for many sensors, except for those with widely varying view angles (such as the NOAA Advanced Very High Resolution Radiometer) where spatial resolution elements increase greatly in size outside the central swath. This consistency of support makes geostatistical models, which rely on fixed support, tenable. The flip side of this advantage is that for spaceborne sensors, supports are usually much larger than those of ground measurements, particularly measurements of vegetation amount. This support mismatch between image and ground measurements creates problems for the robustness of prediction methods– they may only be applicable to the exact combination of supports for which they are developed unless support effects are explicitly accounted for. One of the longstanding grails in remote sensing is to apply methods developed with one sensor to data from a second sensor (Wardley and Curran 1984, Raffy 1994, Iverson *et al.* 1994). Change of support models are needed here (Atkinson 1997*b*, Dubayah *et al.* 1997, Raffy and Gregoire 1998, Chen 1999), but they are complicated by the differences in radiometric characteristics among sensors (Atkinson and Emery 1999). Careful choice of model and methods to fully exploit the data is therefore crucial.

Certainly, a major advantage of remotely sensed image data is their exhaustive spatial sampling of a region. Unfortunately, the actual measurement made by remote sensors is of a variable that is only related indirectly to many of the biophysical, ecological or other environmental processes of interest to researchers (Curran *et al.* 1998).

Table 3.1: Summary of the general properties of and characteristics taken into account by three prediction methods. A '√' indicates the method has the property or preserves the property in the spatial predictions.

	Aspatial regression	Cokriging	Conditional simulation
optimality	√	√	
exactness		√	√
measurement error	√	√	√
measurement support		√	√
measurement location		√	√
sample geometry		√	√
spatial variability			√
spatial correlation	[a]		√

[a]Except as a side-effect

3.4 Conclusions

In this chapter, the groundwork has been laid for a comparison of aspatial regression, cokriging and conditional simulation. The three methods deal differently with measurement error, measurement support, sample geometry, relative locations of the measurements, spatial variability and spatial correlation and have, or fail to have, the properties of exactness and optimality. Table 3.1 summarizes these properties. While all methods have foundations in a random function model, the form of this model, in particular whether it is composed of independent or dependent random variables, creates major differences in the possible outcomes of prediction. Because the forms of the random function model used to premise the methods are different, it is not possible to compare theoretically the potential accuracy of aspatial vs geostatistical methods.

As Lesch *et al.* (1995) showed, if cokriging is applied to a purely random function model, the result is an aspatial regression prediction. So, a critical issue in prediction is which form of the random function to rely on as a useful model. The difference between geostatistical approaches and aspatial ones is the willingness to take spatial dependence for granted and exploit it, or to ignore it and hope that it does not reduce the quality of the resulting prediction. In fact, another term for the aspatial regression could be nongeographical regression, since it uses a model that

is antithetical to geography's 'first law' (see page 41).

For many geographical phenomena, there exist obvious spatial processes that give rise to spatial dependence. In the case of vegetation amount, indirect environmental factors often cause such dependence. For example, there exists spatial continuity of climate, topography and soil attributes that provide growing conditions and a substrate favorable to the development of certain ranges of leaf area. Transport of seed is non-random, making nearby seed sources (and their concomitant species mix and eventual development into canopies) more plentiful than those farther away. Spatial continuity of human land use and ownership patterns also affects vegetation, with the behavior of the tractor and the plow generating characteristic spatial patterns. Natural disturbances such as fire and wind also affect large areas. The choice of spatially dependent models would appear to be logical under these conditions. Aspatial regression has been applied far more often in this context than has any spatially dependent model however.

Theoretical differences among the methods may have less to do with their relative frequency in the research literature than practical considerations. It is likely that few workers in remote sensing have been familiarized with geostatistical methods of prediction. Other practical differences exist, with increasing complexity in procedures, decreasing availability of software for implementation and decreasing speed of computation in the progression from aspatial regression to cokriging to conditional simulation.

Kriging variance maps are common in the geostatistical literature, in contrast to aspatial regression studies where the standard error of estimate is most often reported as a single number. In fact, spatial uncertainty description for continuous variables is not yet routine outside of geostatistics and is currently an important area of research (Lunetta *et al.* 1991, Openshaw 1992, Unwin 1995, 1996). Though regression theory includes a fully specified suite of uncertainty models, these have not been used to describe error spatially in predictions of vegetation amount.

In the geostatistical approach, ground measurements are used directly in prediction. Aspatial regression uses the ground sample values only to develop the model; these values are ignored in the predicted map. The potential exists for bias in the regression coefficients, especially without taking into account measurement

error in the ancillary variable. This kind of bias can cause large errors, especially where values of the ancillary variable are far away from their mean. Geostatistical methods have a different disadvantage; in the absence of local data it may not even be possible to make predictions at all. The tradeoffs involved in applying these methods to synthetic and real data sets will be explored in the following chapters.

4 A synthetic analog to the remote sensing of vegetation amount

In chapter 3, the theoretical elements of geostatistical prediction were contrasted with those of aspatial regression. The purpose of this chapter is to begin to explore how these elements can be implemented in practice for the remote sensing of vegetation amount. It compares maps resulting from aspatial regression with those produced using cokriging and one type of conditional simulation called probability field simulation. The features and tradeoffs involved in choosing a particular spatial prediction method are discussed, especially the potential to generate accurate predictions and descriptions of uncertainty. A remote sensing analogy is presented to illustrate the application of each method and its results.

Statistical models of vegetation amount and predictions from them developed from previous remote sensing studies cannot be compared quantitatively for several reasons. As seen in chapter 2, specific models relating spectral data and vegetation amount tend to be reported but not applied; if they are applied, they are not applied in the same study area more than once. The unique characteristics of these studies, their diverse combinations of variables, spatial supports and preprocessing steps, preclude comparison of accuracy. Further, exhaustive description of a vegetation amount variable is never available for a real data set. Information about true values is always limited to a small subset of ground observations and even these contain measurement error. To achieve an objective comparison of the actual accuracy of maps predicted using alternative methods and to act as a benchmark data set with which future developments in prediction might be tested, a synthetic data set was constructed. Such data sets, with properties that are known completely, are commonplace in image processing (Gregotski and Jensen 1993, Foody *et al.* 1995) and geostatistics (Journel and Alabert 1989, Isaaks and Srivastava 1989, Zimmerman *et al.* 1998).

The synthetic data set presented here is useful as an analog to data collected for remote sensing studies. The data came from actual sensor data rather than being generated from some ideal statistical model and therefore shared some of the qualities of real data. But unlike a real data set, the synthetic data set allows objective

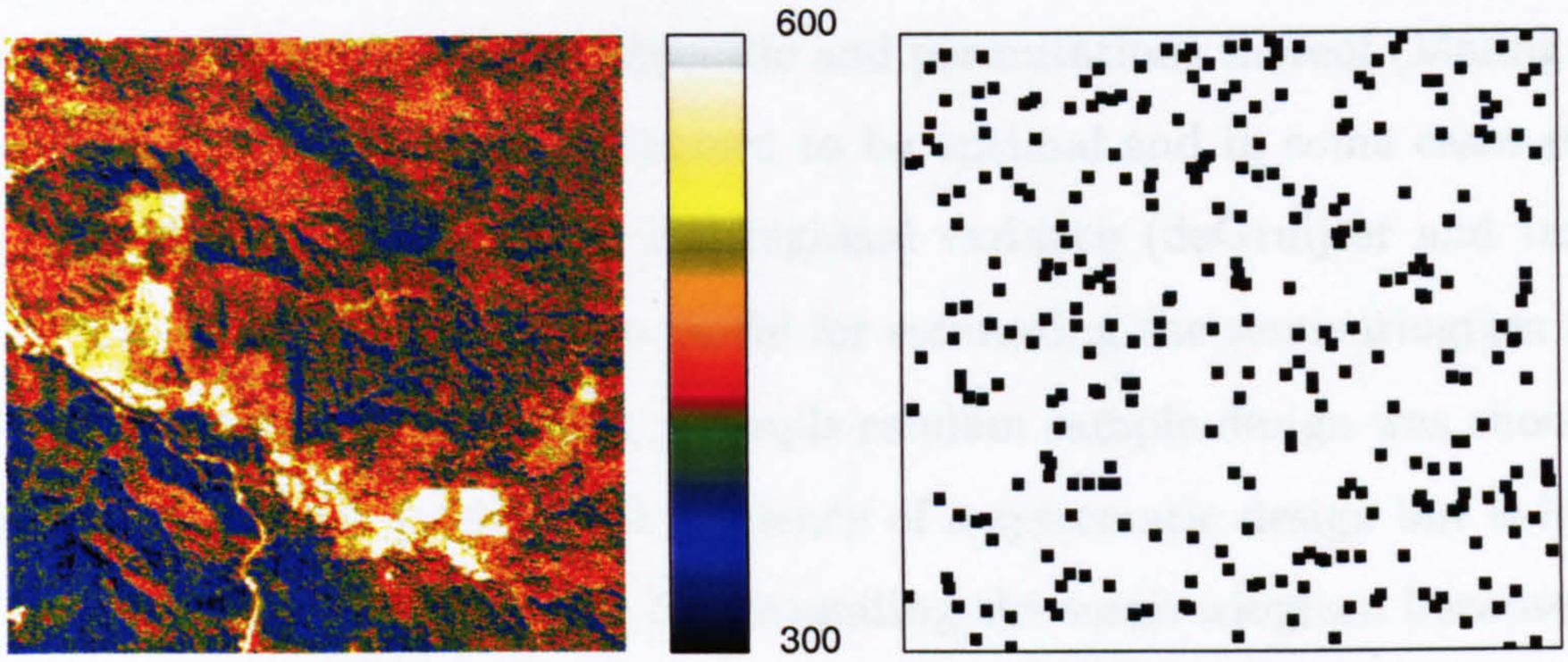
comparison of the consequences of using specific algorithms. It allows 'controlled experiments', in the sense that it provides consistency in some circumstances that are likely to vary in real situations. For example, this analog does not address the different spatial supports of remotely sensed data and ground data. By controlling for this factor, the effect of different supports on the results is not included and the comparison, although limited in scope, can be evaluated clearly. Results from the synthetic data set are illustrative only – relative accuracies of spatial prediction methods will differ in every specific case and with sample size.

A major finding of the studies reported in table 2.1 was that spectral variables and vegetation amount are related. A next logical question is, how close does the relationship have to be before prediction using spectral data as an ancillary variable becomes more accurate than using vegetation amount data (as the primary variable) alone? Another is, how close does the relationship between primary and ancillary variables have to be for cokriging to be more accurate than aspatial regression? To address these questions, the synthetic data set contains several alternative ancillary images representing remote sensing data, each with a different correlation with primary (representing vegetation amount) data. A similar synthetic data set was used by Dungan *et al.* (1994), using a single ancillary image, to illustrate the possibility of conditional simulation for mapping vegetation amount and an earlier version of the results presented herein were reported by Dungan (1998).

4.1 The synthetic data set

The synthetic data were from an image obtained by NASA's Airborne Visible/Infrared Imaging Spectrometer, AVIRIS (Vane *et al.* 1993), collected near the coast of Oregon, USA on 14 August 1990. This sensor collects data in 224 contiguous visible and near infrared bands that are, on average, approximately 10 nm wide. The close spacing of the spectral bands leads to high correlation between adjacent bands; within a spectral region, this correlation tends to decrease as band separation increases. A 300 by 300 pixel subset from a single band (band 21, centered at 596.7 nm) was selected from the AVIRIS image cube to represent a 'true map' (figure 4.1*a*), or an exhaustive and error-free description of a vegetation amount variable. This ex-

Figure 4.1: (a) AVIRIS band 21 used as the true map (unitless) and (b) the locations of the 300 sample pixels from the true map (symbols do not represent the actual size of pixels, but are enlarged for visibility).



a b

haustive description of the primary variable is based on a support equal to the area of one pixel. For AVIRS data, this support is approximately 20×20 m, or 400 m^2. Though the physical units of these data were those of radiance ($\mu Wcm^{-2}nm^{-1}sr^{-1}$), for the purposes of this study the data are assumed to be unitless.

Seven other bands from the AVIRIS cube in the visible region were selected to represent remotely sensed images of the area, termed 'ancillary images' herein. These bands; 7, 9, 10, 12, 13, 15 and 22; represented a range of linear relationships with the true map. The linear correlation coefficients (r) were .45, .55, .61, .75, .79, .88 and .94 respectively.

4.2 Methods

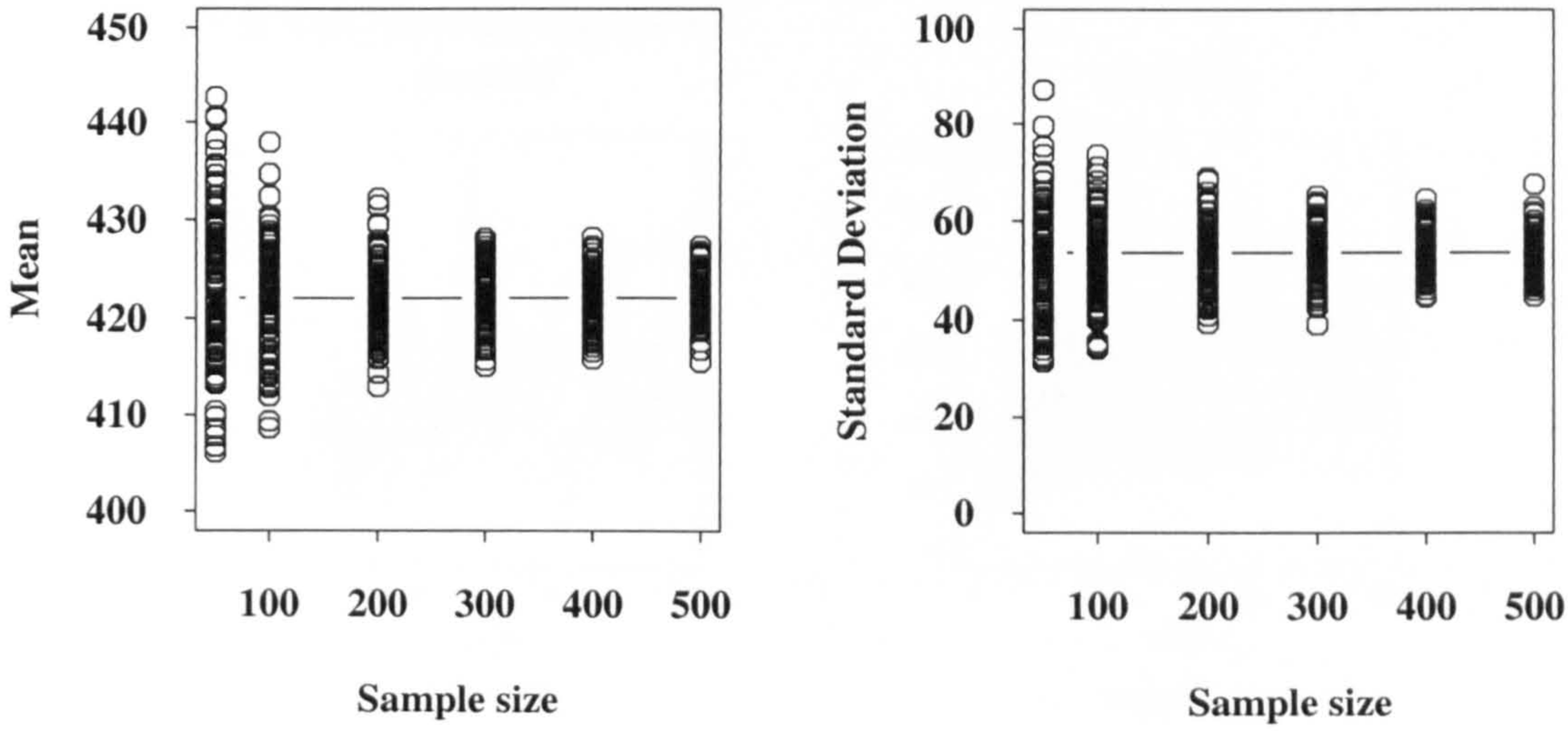
4.2.1 Sampling from the data set

The use of synthetic data allows complete freedom in choosing samples to represent ground measurements of the primary variable. Spatial sampling includes simple random, stratified random, systematic and permutations thereof (Maling 1989). Sampling designs of this type are known to be optimal and in some cases efficient for estimating the regional mean and regional variance (deGruijter and terBraak 1990). Other designs may be more useful for estimating the semivariogram (Russo 1984, Morris 1991). For this study, a simple random sample design was chosen as a compromise that does not have the efficiency of a systematic design but is likely to be superior to a systematic design for estimating the semivariogram because of the increased number of pairs at smaller lags (Brus and deGruijter 1994). It also has the advantage of a straightforward implementation.

As has been established in chapter 3, the specific objectives of spatial prediction may dictate the choice of a prediction method. The optimal *number* of sampled values also varies depending on estimation or prediction objectives (Cressie 1991). Sample sizes of 50, 100, 200, 300, 400 and 500 from the true map (figure 4.1*a*) were used to show the variation in estimates of the regional mean, variance and semivariogram. One hundred sample sets of each size were taken from the true map and these three statistics were calculated from each set (figure 4.2). The regional mean and variance calculated from these sets were unbiased (figure 4.2*a*, *b*), with increasing sample size showing diminishing returns after a sample size of about 200. Figure 4.3 shows the semivariograms resulting from ten sample sets of each size. The root-mean-square-error between the semivariogram estimated from each sample set and the exhaustive semivariogram (figure 4.2*c*) indicated an eventual convergence toward 0, again with the variance decreasing slowly after a sample size of 200.

These considerations led to the choice of a simple random sample of 300 pixels (or .3% of the true map) to represent ground-based measurements of the primary variable (figure 4.1*b*); these were the primary data. The ancillary images were utilized one at a time in combination with the primary data to predict the values in the true map using regression, cokriging and probability field simulation. The

Figure 4.2: Statistics from random samples of different sizes from the true map: (a) mean, (b) standard deviation and (c) root-mean-square error between the semivariogram calculated from each sample and the true map's exhaustive semivariogram.



c

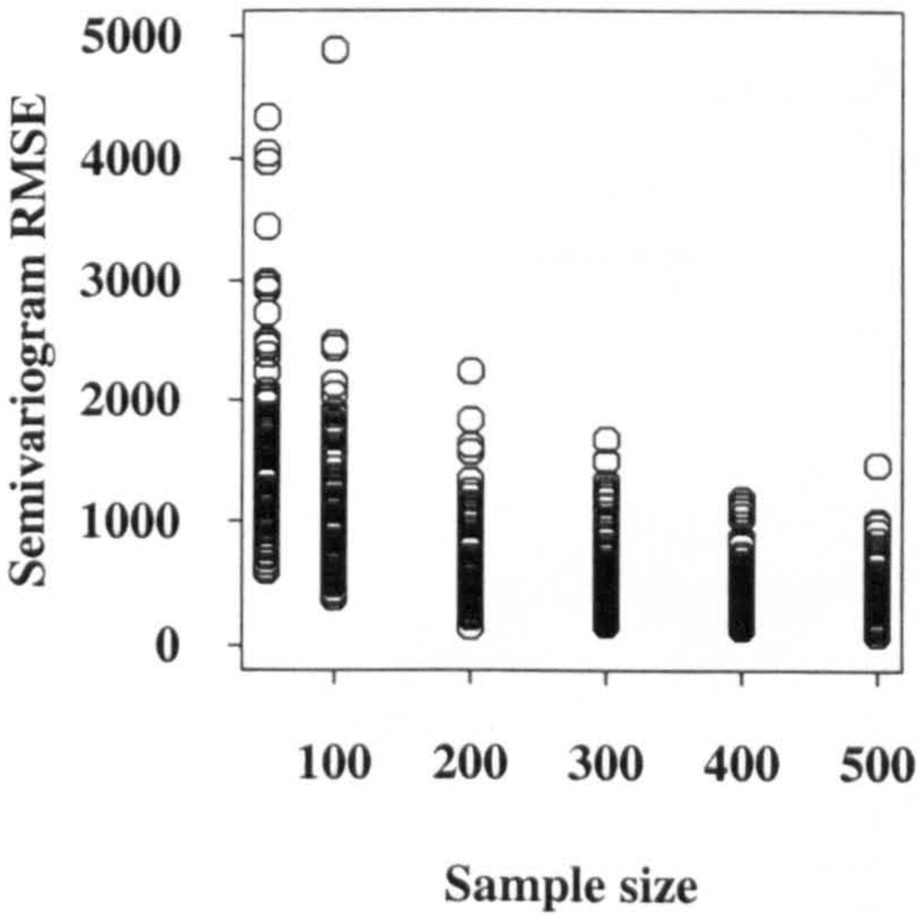
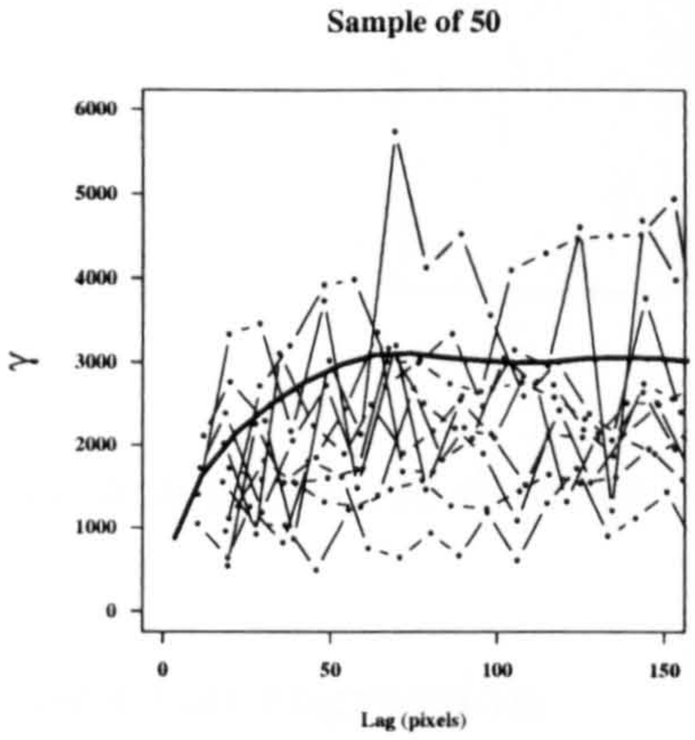
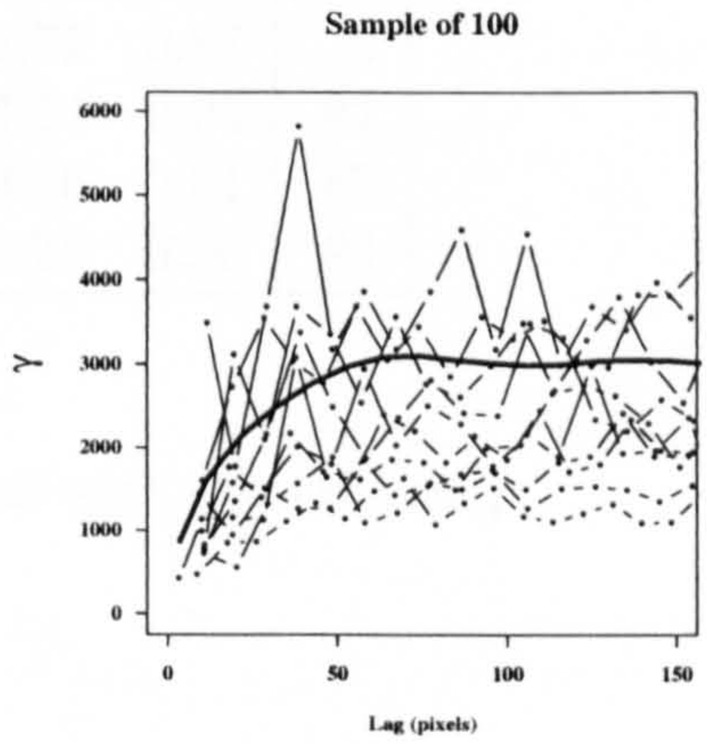
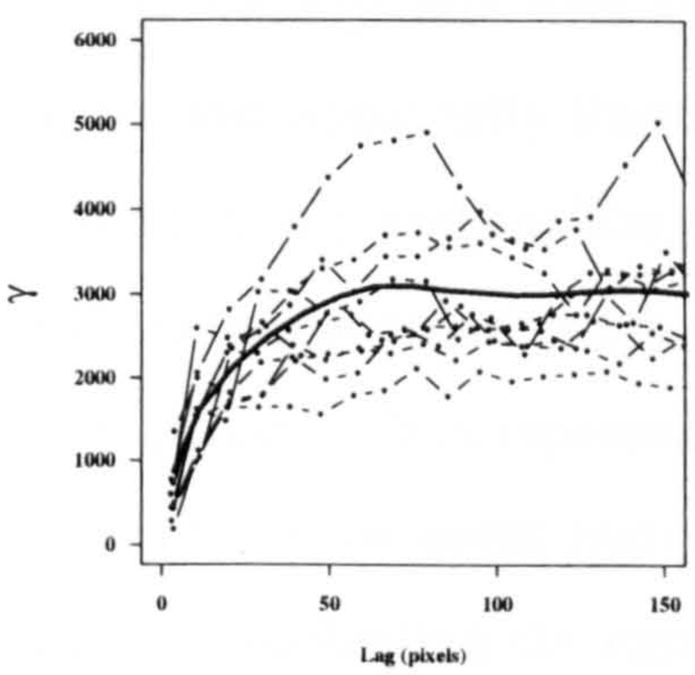
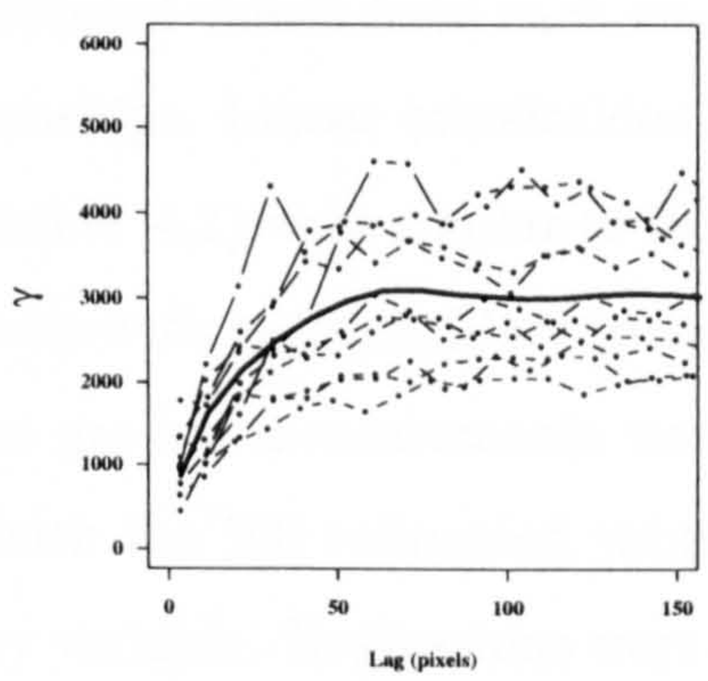


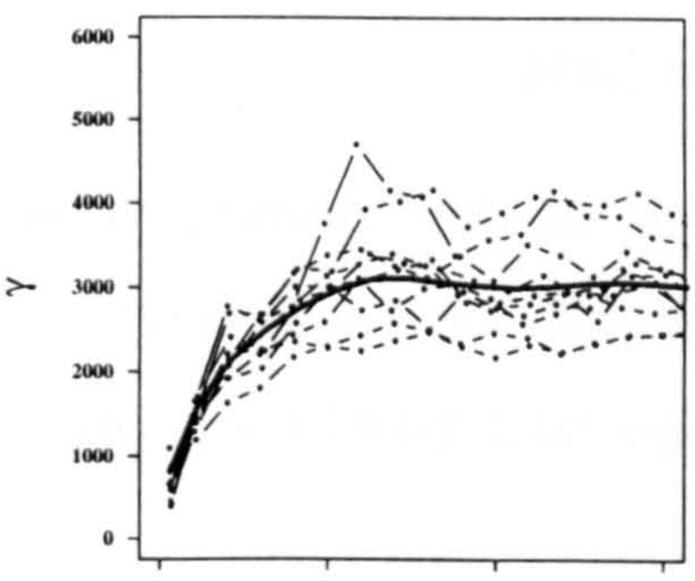
Figure 4.3: Omnidirectional semivariograms from random samples of different sizes (50, 100, 200, 300, 400 and 500) from the true map (–•–) and the exhaustive omnidirectional semivariogram from the true map (—).











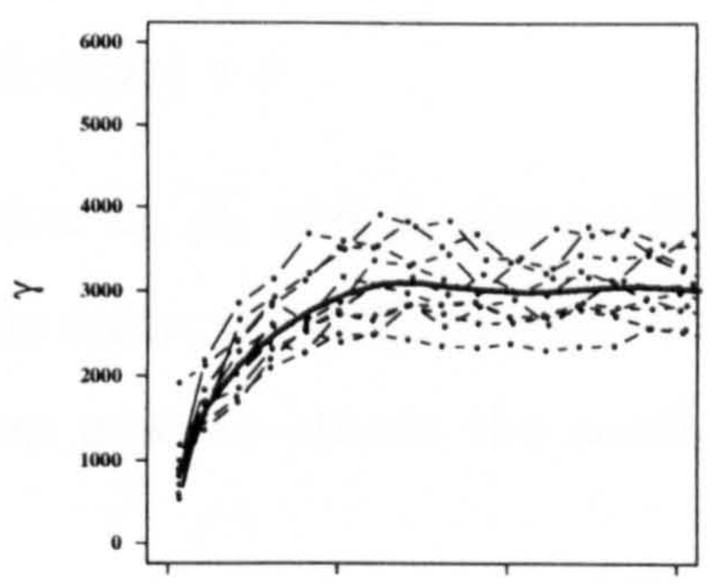


Table 4.1: Parameters of the regression models (β_0 and β_1) and correlation coefficients (r) between the collocated primary data and the ancillary data from bands 7, 9, 10, 12, 13, 15 and 22.

Ancillary image band	β_0	β_1	r
7	66.81	0.42	.41
9	3.32	0.56	.53
10	-8.01	0.64	.54
12	-102.71	0.87	.72
13	-134.54	0.95	.82
15	-99.76	0.81	.87
22	70.28	0.90	.94

predicted maps were then directly compared with the true map.

4.2.2 Aspatial regression

Scatterplots of primary data with their collocated values from each ancillary image (figure 4.4) showed apparently linear relationships. Linear correlations between the 300 collocated primary and ancillary data (table 4.1) were similar to the coefficients from the exhaustive population (the true map with each ancillary image).

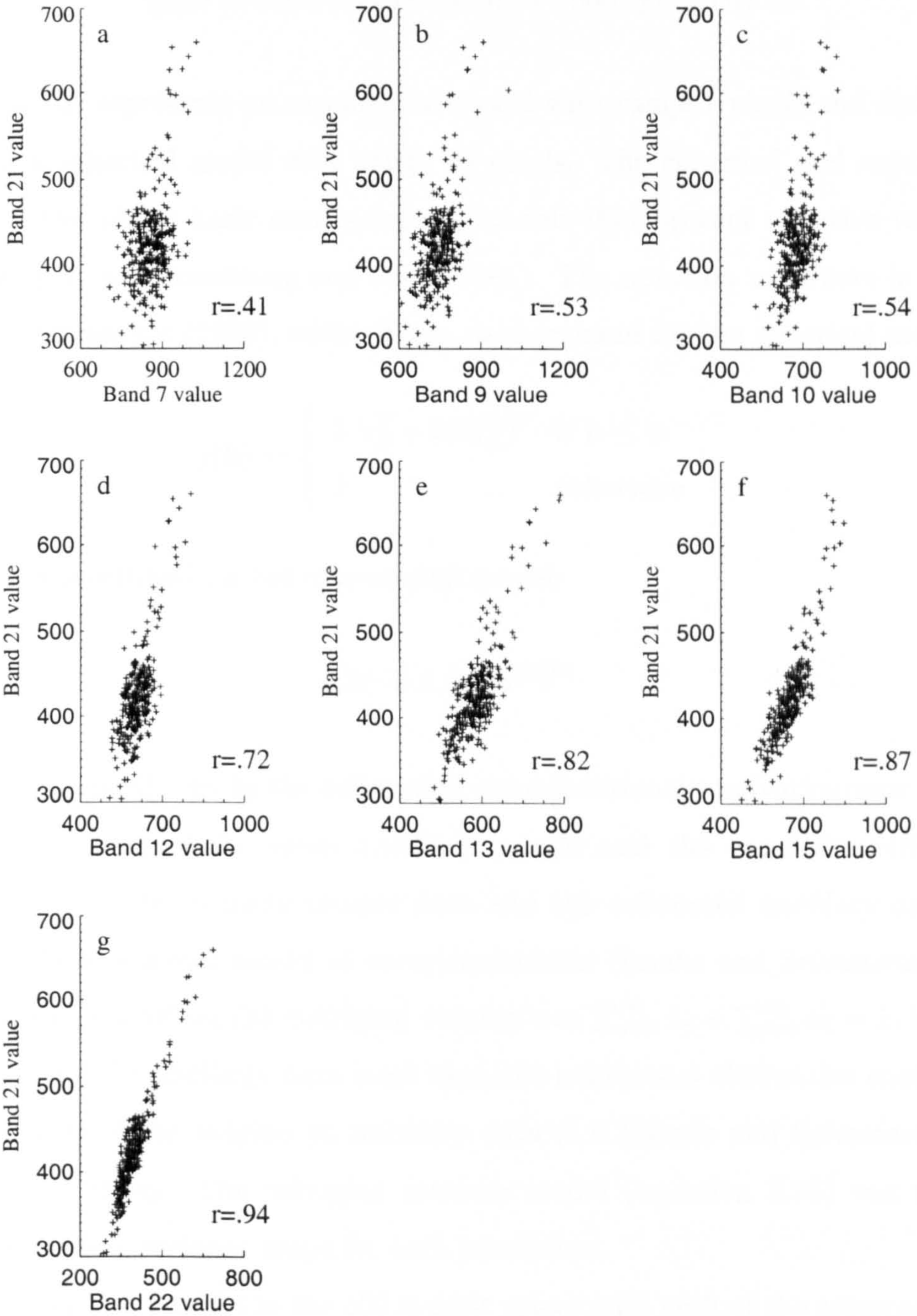
The 300 primary data representing the ground measurements were treated as the response variable in linear regressions with the 300 collocated values from each ancillary image representing the explanatory variable. Regressions were of the form:

$$p(\mathbf{x}_i) = \beta_0 + \beta_1 a_1(\mathbf{x}_i) + \epsilon \tag{4.1}$$

where p is the primary variable at a location $\mathbf{x}_i$, β_0 and β_1 are coefficients of the model, a_1 is the ancillary variable and ϵ is the error, the difference between predicted and true values. Ordinary least squares was used to obtain the coefficients (table 4.1).

The regression equation for each ancillary image was applied to all 90000 pixels in it to predict the primary variable. The variance of the prediction at each grid cell was also generated using equation 3.22.

Figure 4.4: Scatterplots of primary data versus collocated values from ancillary data from (a) band 7, (b) band 9, (c) band 10, (d) band 12, (e) band 13, (f) band 15 and (g) band 22.



4.2.3 Cokriging

To implement cokriging, a first step is the description of spatial dependence of the primary and ancillary variables. A model was fit to the omnidirectional semivariogram ($\gamma_p(\mathbf{h})$) calculated from the 300 sample values from the true map,

$$\gamma_p(\mathbf{h}) = 160 + 1410 Exp_7(\mathbf{h}) + 1600 Sphr_{80}(\mathbf{h}) \tag{4.2}$$

where $Exp_7(\mathbf{h})$ represents an exponential model with range 7 pixels and $Sphr_{80}(\mathbf{h})$ represents a spherical model with range 80 pixels. The spherical and exponential models are two of the basic semivariogram models that prevent negative variances in kriging systems (Armstrong and Jabin 1981). The notation used here is that of Isaaks and Srivastava (1989), where $Sphr_\alpha$ is shorthand for the spherical model:

$$\gamma(\mathbf{h}) = \begin{cases} 1.5\frac{\mathbf{h}}{\alpha} - 0.5(\frac{\mathbf{h}}{\alpha})^3 & \text{if } \mathbf{h} \leq \alpha \\ 1 & \text{otherwise} \end{cases} \tag{4.3}$$

and Exp_α is shorthand for the exponential model:

$$\gamma(\mathbf{h}) = 1.0 - e^{(\frac{-3\mathbf{h}}{\alpha})} \tag{4.4}$$

Models were also fit to the exhaustive omnidirectional semivariograms ($\gamma_a(\mathbf{h})$) calculated for each of the seven ancillary images and the cross-semivariograms ($\gamma_{pa}(\mathbf{h})$) between the primary sample data and the collocated ancillary data (table 4.2) to form a linear model of coregionalization (Isaaks and Srivastava 1989). The constraint placed on the cokriging weights was $\sum_{j=1}^{N_1} \lambda_j + \sum_{l=1}^{N_2} \omega_l = 1$, because this emphasizes the ancillary data more than the traditional alternative constraints where the sum of the weights on ancillary data is 0 (Isaaks and Srivastava 1989, Dungan *et al.* 1994). The cokriging variance model (equation 3.25) was used to generate cokriging variance maps for each prediction.

Cokriging was applied to the 300 sample values with each of the seven ancillary images using the code *cokb3d* (Deutsch and Journel 1992). A search neighborhood of 90 pixels, just beyond the range of the semivariogram model, was used for both the primary and ancillary data. To indicate the potential for prediction with no

Table 4.2: Parameters of the semivariogram models used for cokriging, where C_0 is the nugget variance, C_1 is the coefficient on the exponential model and C_2 is the coefficient on the spherical model. $\gamma_a(\mathbf{h})$ is the semivariogram model for the ancillary variable from the band listed in the left column and $\gamma_{pa}(\mathbf{h})$ is the cross-semivariogram for the ancillary variable with the primary variable.

Ancillary band	$\gamma_a(\mathbf{h})$			$\gamma_{pa}(\mathbf{h})$		
	C_0	C_1	C_2	C_0	C_1	C_2
7	1980.	430.	550.	1.	400.	1000.
9	1250.	490.	570.	1.	810.	780.
10	1050.	600.	450.	1.	740.	760.
12	560.	850.	650.	1.	900.	1000.
13	480.	900.	850.	1.	900.	1100.
15	350.	1460.	1400.	1.	1400.	1450.
22	160.	1410.	1870.	1.	1400.	1700.

ancillary (remote sensing) data, a prediction was also made using ordinary kriging with just the 300 primary data. Ordinary kriging was accomplished with *okb3d* (Deutsch and Journel 1992).

4.2.4 Conditional simulation

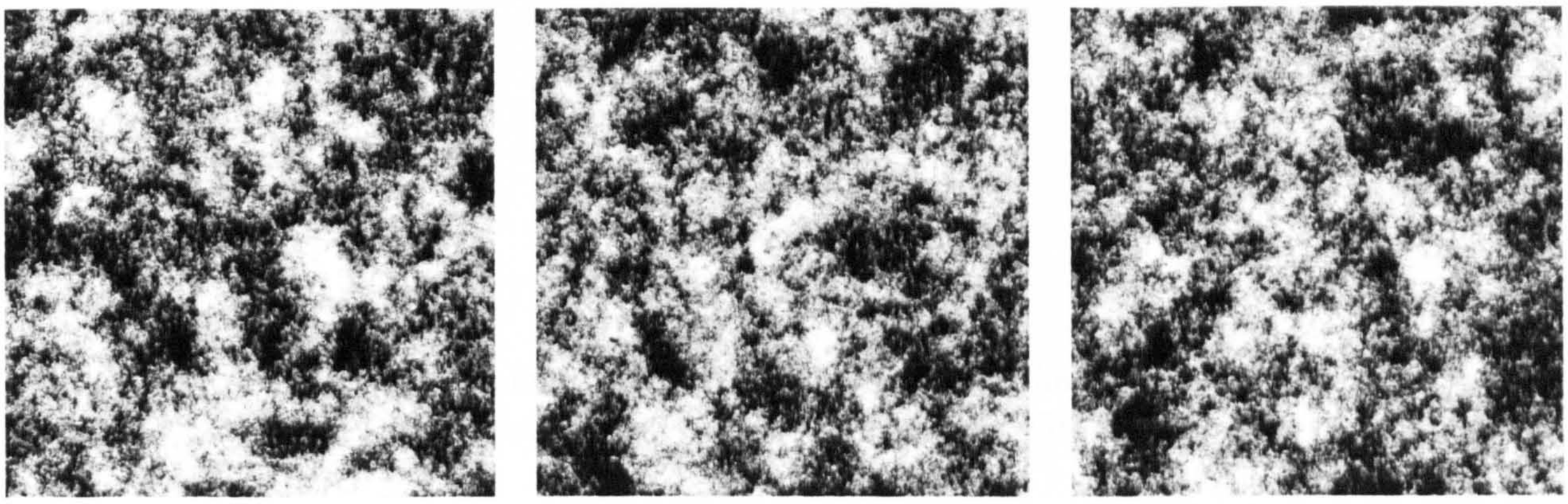
To generate probability field simulations (Srivastava 1992), the primary data were first transformed to a uniform distribution. The omnidirectional semivariogram of the transformed data was then modeled by

$$\gamma_u(\mathbf{h}) = .02 + .032Exp_7(\mathbf{h}) + .03Sphr_{80}(\mathbf{h}) \tag{4.5}$$

Thirty unconditional simulations (three are shown in figure 4.5) were generated to reflect this uniform-score semivariogram using the sequential Gaussian simulation code *sgsim* (Deutsch and Journel 1992). In general the exhaustive omnidirectional semivariograms of the probability fields satisfactorily fit the experimental variogram (figure 4.6). Departures from a single model can be ascribed to the small sample size as seen in figure 4.3*c*.

The probability fields were then used to draw from the local conditional distributions defined by the cokriging variance (equation 3.25) to create thirty realizations for each of the ancillary images. The code used to accomplish this was written in

Figure 4.5: Three of the 30 probability fields. The grey scale ranges from black (representing 0) to white (representing 1).



the Interactive Data Language (IDL, Research Systems Inc., Boulder, CO). The variance among the thirty realizations at each pixel was computed as a description of uncertainty.

4.3 Results

For all three methods, the accuracy of predictions increased with increasing correlation between primary and ancillary data, as expected. This progression is seen in figures 4.7*a*-*c*, which shows maps predicted with regression using bands 7, 12 and 22 respectively. The map predicted with band 7 (figure 4.7*a*), was noisy and shows little continuity and few areas of extreme values. The prediction with band 12 (figure 4.7*b*), was more accurate and that predicted with band 22 (figure 4.7*c*) was difficult to distinguish from the true map (figure 4.1*a*).

Maps predicted with cokriging (figures 4.7*d*-*f*) showed a more consistent resemblance to the true map and had a 'smooth' character not seen in the regression predictions. Single realizations from probability field simulation are shown in fig-

Figure 4.6: Semivariogram values from the uniform scores of the primary data (-•-) and exhaustive omnidirectional semivariograms from seven probability fields (—).

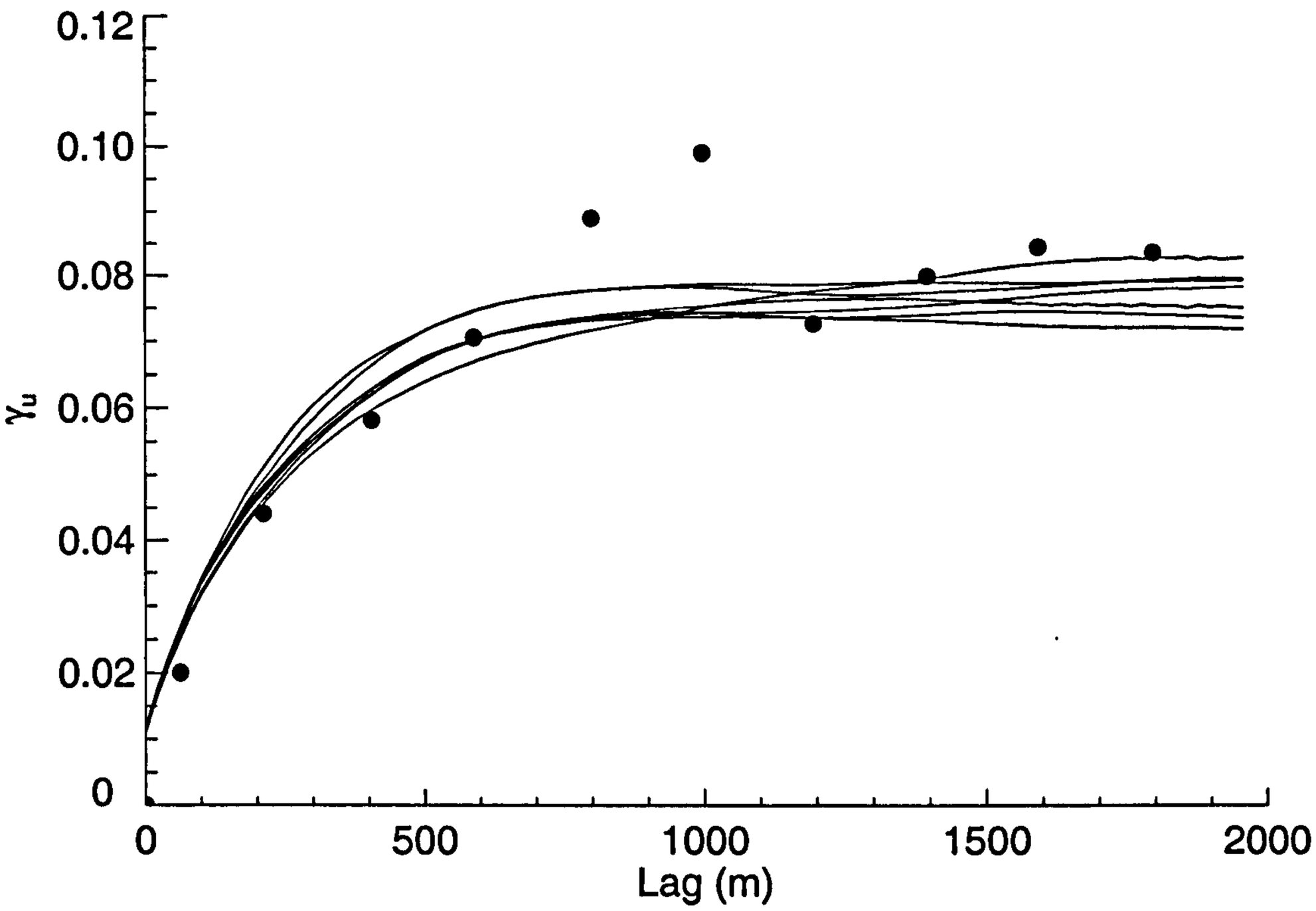
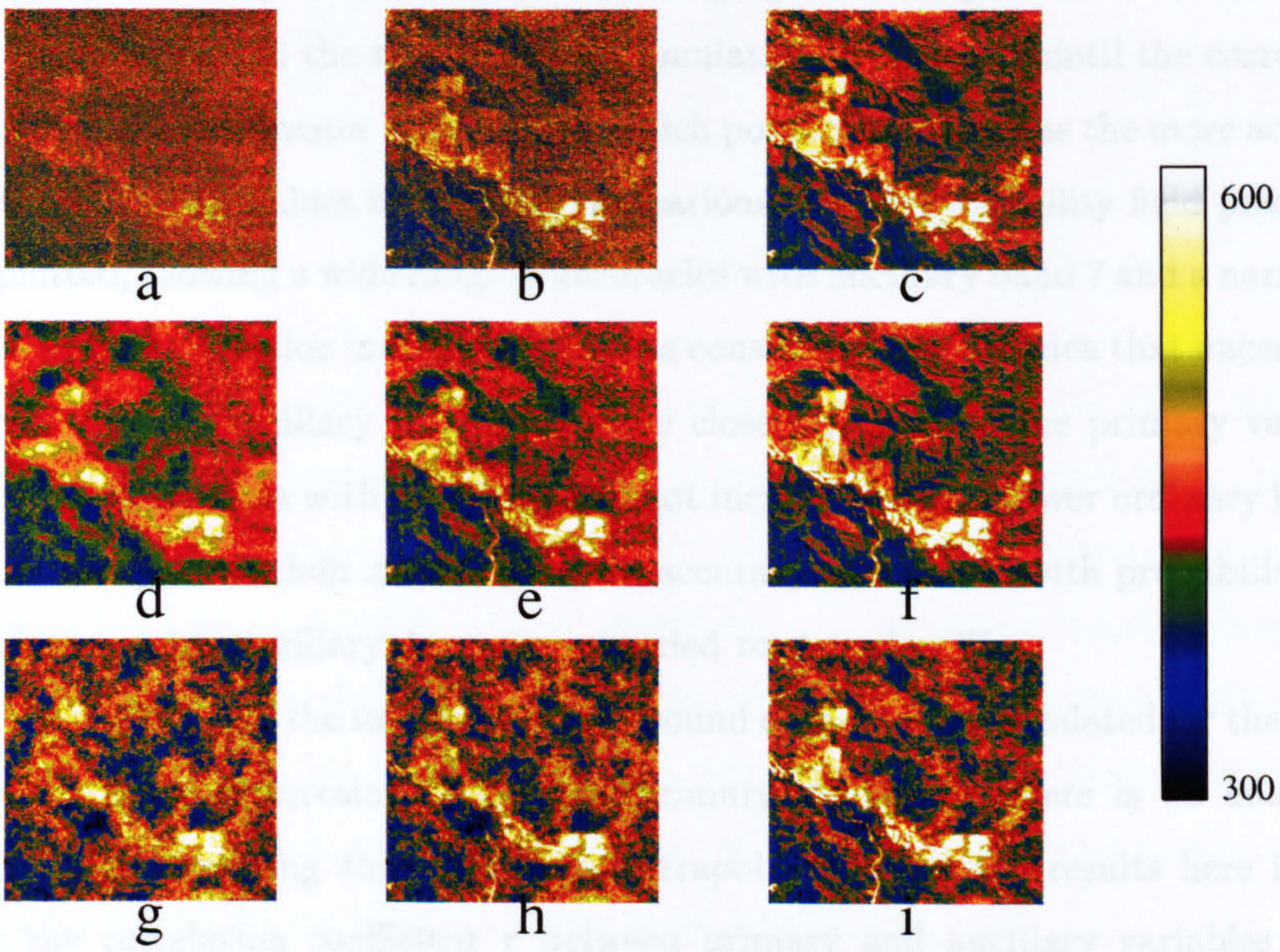


Figure 4.7: Maps resulting from the regression method applied using (a) band 7, (b) 12 and (c) 22; maps resulting from the cokriging method applied using (d) band 7, (e) 12 and (f) 22 and single realizations from probability field simulation using (g) band 7, (h) 12 and (i) 22. The color scale is unitless.



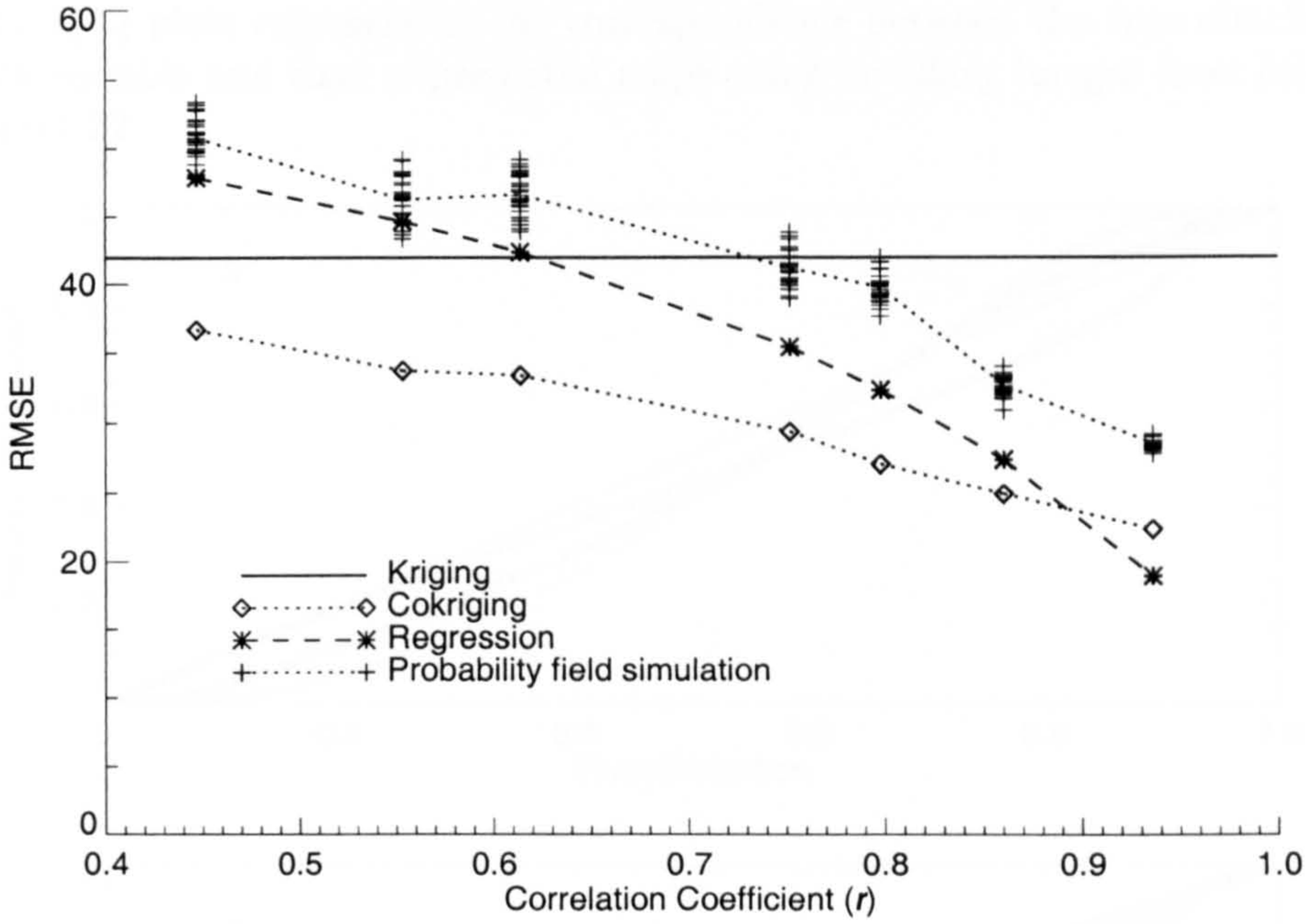
ures 4.7*g*-*i*. Simulations using poorly correlated ancillary data, such as figure 4.7*g*, appeared to do almost as well as cokriging in predicting areas of low and high values. But realizations using the most highly correlated ancillary data (figure 4.7*i*) were less similar to the true map than were the corresponding cokriging and regression predictions.

The overall accuracy of each predicted map can be measured using the root-mean-square-error (RMSE) between it and the true map. The RMSE from ordinary kriging with just the 300 primary data was 43; ancillary data that do not increase accuracy beyond this value are not likely to be considered useful. The trends of RMSE with the correlation coefficient of the ancillary data (figure 4.8) showed that cokriging was always more accurate than kriging for these particular ancillary data and that it exceeded the regression and simulation in accuracy until the correlation coefficient (r) was greater than $\sim$.89 at which point regression was the more accurate predictor. RMSE values from all 30 realizations of each probability field prediction are plotted, showing a wide range of accuracies with ancillary band 7 and a narrowing range as the correlation increases. This was consistent with the idea that uncertainty decreases as an ancillary variable is more closely related to the primary variable. Use of ancillary data with regression did not increase accuracy over ordinary kriging until r was greater than .6. Likewise, for accuracy to increase with probability field simulation using ancillary data, the r needed to exceed $\sim$.75.

How closely do the image data and ground data have to be related for the image data to reduce uncertainty about the primary variable? There is no analytical method for answering this question. Extrapolating from the results here implies that the correlation coefficient r between primary and ancillary variables would have to exceed around .3 for the root-mean-square error of cokriging to be less than that from ordinary kriging. Kupfersberger and Bloschl (1995) tested data sets with correlations of .4, .6 and .8 and found ancillary data with r values as low as .6 were significantly informative in estimation via cokriging. Asli and Marcotte (1995)'s results were similar, with .4 found to be the threshold. Yates and Warrick (1987) suggested that .5 is a threshold.

Predicted maps whose univariate distributions were closest to that of the true map appeared to be those from probability field simulation. Q-Q plots (figure 4.9)

Figure 4.8: The trend in root-mean-square-error (RMSE) among maps predicted with the three methods as a function of the correlation between the primary and ancillary information. The prediction resulting from ordinary kriging with 300 primary data only is plotted for reference.



of single probability field realizations were closer to the 1:1 line than are those from cokriging or regression. Both geostatistical methods also had better pattern-preserving qualities than did regression, at least when ancillary data were not very closely related to the primary. This is expressed in a comparison between exhaustive semivariograms of predicted and true maps (figure 4.10). The shapes of semivariograms from probability field simulation were always similar to that of the true map, although the sill was underestimated, while regression-predicted maps' semivariograms were flatter from six of the seven regressions.

Maps resulting from the prediction variance models of the three methods showed different magnitudes and patterns (figure 4.11). The variance of prediction from regression was spatially more homogeneous than the others; what small variation there was followed the pattern of the ancillary data values. The cokriging variance pattern was virtually a map of sample data location, with variance increasing isotropically with distance from sample locations. Variance among probability field simulation realizations was a mixture of both patterns.

Figure 4.9: Q-Q plots representing the correspondence between the true distribution of the primary variable and that of predicted maps using ancillary images from (a) band 7, (b) 12 and (c) 22.

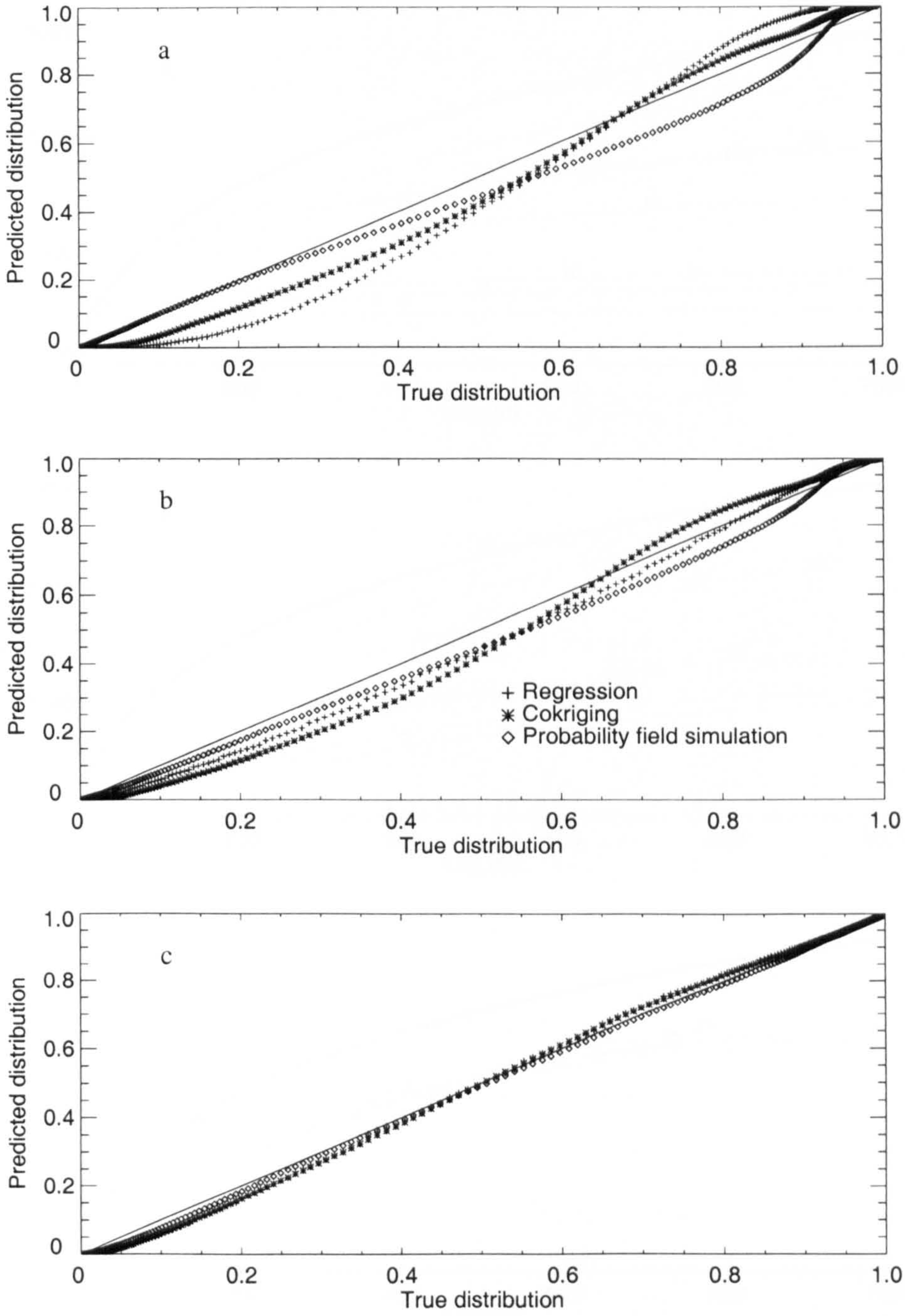


Figure 4.10: Exhaustive omnidirectional semivariogram calculated from the true map (—) compared with those calculated with the predicted maps (- - -) using the seven ancillary images (bands 7, 9, 10, 12, 13, 15 and 22) and (a) regression, (b) cokriging and (c) a single probability field realization.

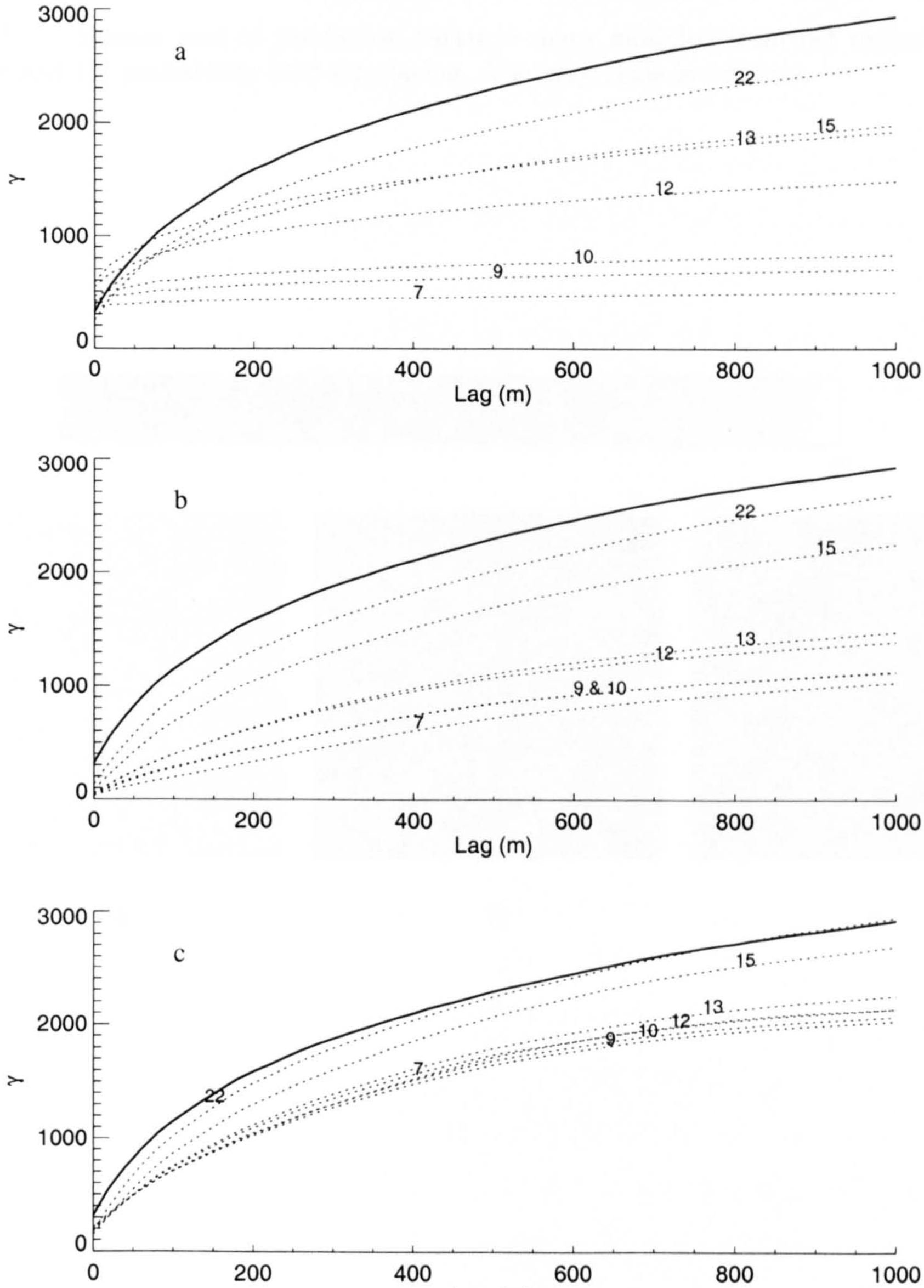
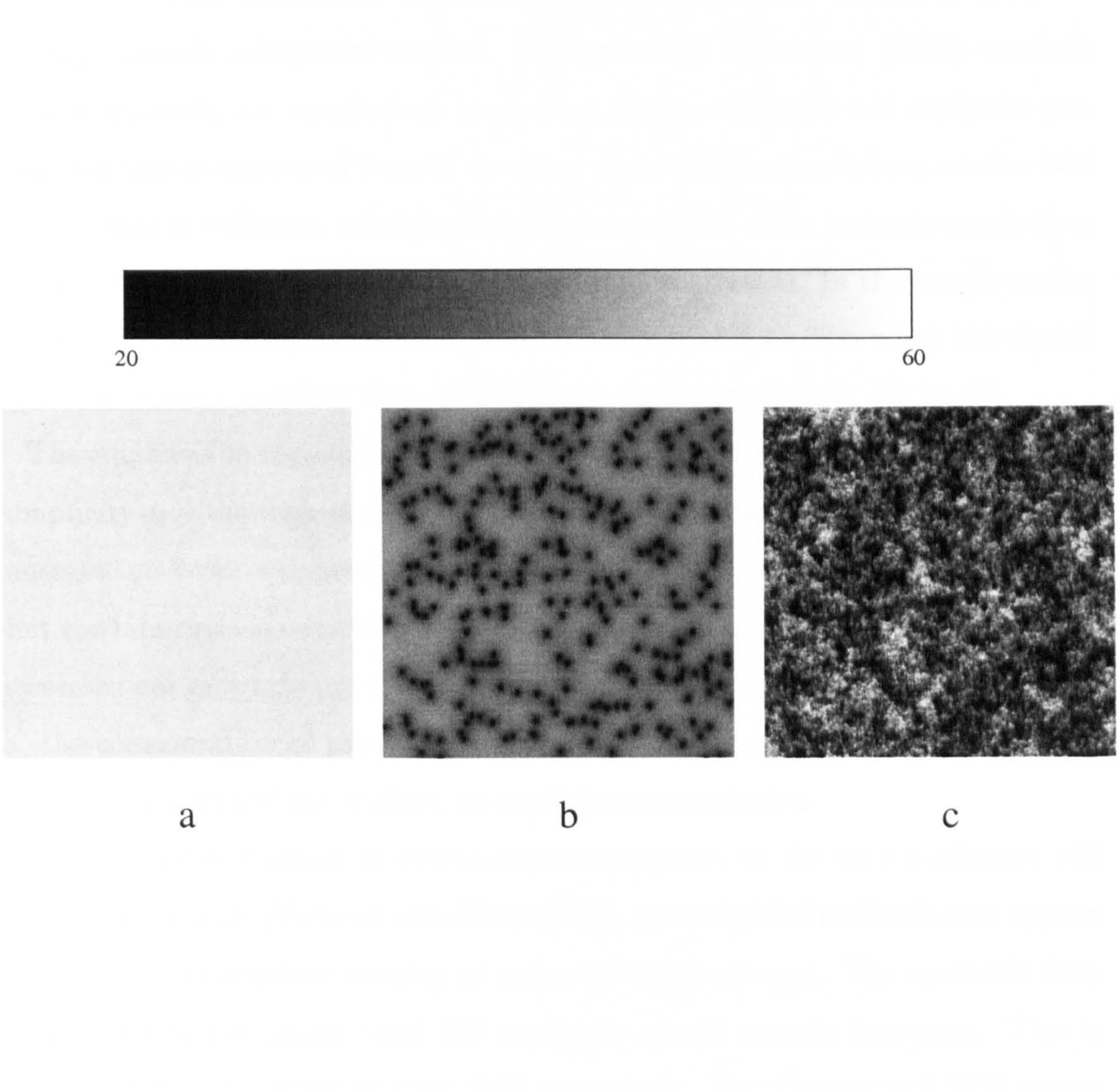


Figure 4.11: Square root of prediction variance maps modeled from (a) regression (b) cokriging and (c) probability field simulation. The grey scale is unitless.



4.4 Discussion

Relationships between remotely sensed variables and ground measurements of vegetation amount are often statistically significant but not strong. For example, Friedl *et al.* (1994) reported significant correlation coefficients ranging between .57 and .68 for grassland LAI and Kauth-Thomas greenness derived from coincident TM images. In a review of relationships between Landsat Thematic Mapper data and conifer forest stand variables (Lambert *et al.* 1995), values ranged from .17 to .96, some of which were significant. Table 2.1 includes many other examples where relationships contain substantial scatter. Holmgren and Thuresson (1998) conclude that these generally low correlations imply that remote sensing is not useful for predicting vegetation amount of forests. In cases where linear correlations are low and sample number is sufficient, cokriging has the potential to make accurate predictions and should be considered an alternative to aspatial regression. In the simple analog described here, cokriging increased overall accuracy by 8% to 23% when correlation coefficients between ancillary data and primary data ranged from .88 to .45.

The emphasis on regression in the remote sensing literature can be explained by its simplicity and the wide availability of software to implement it. These practical advantages can be outweighed by the need for spatial prediction techniques that fully exploit the information content of remotely sensed ancillary data. Other advantages of geostatistical methods are the inclusion of measurement error for the ancillary data, the consideration of measurement location and sample arrangement and the exploitation (instead of the neglect) of spatial autocorrelation.

Because the estimation of semivariograms requires, at the very minimum, 100 sample measurements (Webster and Oliver 1992), geostatistical methods may appear impractical to investigators wishing to map vegetation amount. The synthetic data set illustrated in this paper used 300 randomly placed sample locations. This is a large number, uncommon in most field campaigns. But Curran and Williamson (1986) point out that sample sizes as large as this are usually not sufficient, even for an aspatial regression prediction of the regional mean. If investigators want accurate spatial predictions, increasing sample size will be an important factor regardless of the method used. Sampling design is an important factor as well, since purposive

sampling has been used almost to the exclusion of design-based spatial sampling schemes in field studies of vegetation amount (see chapter 2).

Though geostatistical methods are considered multivariate (it is theoretically possible to include more than one source of ancillary data into the predictions), practical considerations in modeling a valid linear model of coregionalization effectively discourage their use with more than one co-variate data source. In cases where two or more such sources are deemed relevant, they could be combined into a single ancillary variable for cokriging or conditional simulation. Multivariate aspatial regression suffers from no such limitations. The only barriers to the inclusion of a large number of co-variates in regression is the need to avoid over-fitting and collinearity. Though additional ancillary data will usually increase correlations among the sample data, this does not ensure increasing prediction accuracy.

Probability field simulation does not increase overall accuracy over regression methods in the analog reported here, nor would it in real applications. But this method does have promise in providing spatial predictions that are more accurate in their aggregate characteristics, their univariate and spatial statistics, which can be particularly important when the predicted maps are used to locate extreme values or when they are used in combination with other variables in deterministic process models (Abdulla and Lettenmaier 1997). In addition, the variation among and between realizations has the potential to become a useful model of prediction uncertainty.

4.5 Conclusions

This chapter has highlighted some of the considerations in the selection of a method which uses remote sensing as supporting information for quantitative, spatial prediction. Though traditional, aspatial regression is likely to be less accurate than cokriging in situations where radiance or reflectance are not highly correlated with the vegetation variable of interest. Neither regression nor cokriging have the advantages of conditional simulation for the reproduction of univariate distribution and spatial pattern in predicted maps.

The uncertainty model from the regression method is completely data-value

dependent, whereas that from kriging is completely data-location dependent. The uncertainty model from conditional simulation methods is dependent on both location and value and therefore should be more useful in describing the geographical distribution of prediction uncertainty. But since geostatistical approaches to the remote sensing of vegetation are very new, only practical experience with real data will bear out this potential.

Given the many alternatives, how can a choice of a spatial prediction method be made? Criteria upon which to base such a choice include the strength of the relationship between the vegetation variable and remote sensing data, the number of ground measurements collected, the need for an uncertainty description, the scale required for the result, and the character of autocorrelation measured or expected at a given support. No one method is best suited for all situations and goals.

5 Predicting canopy cover in western Montana using ground and image data

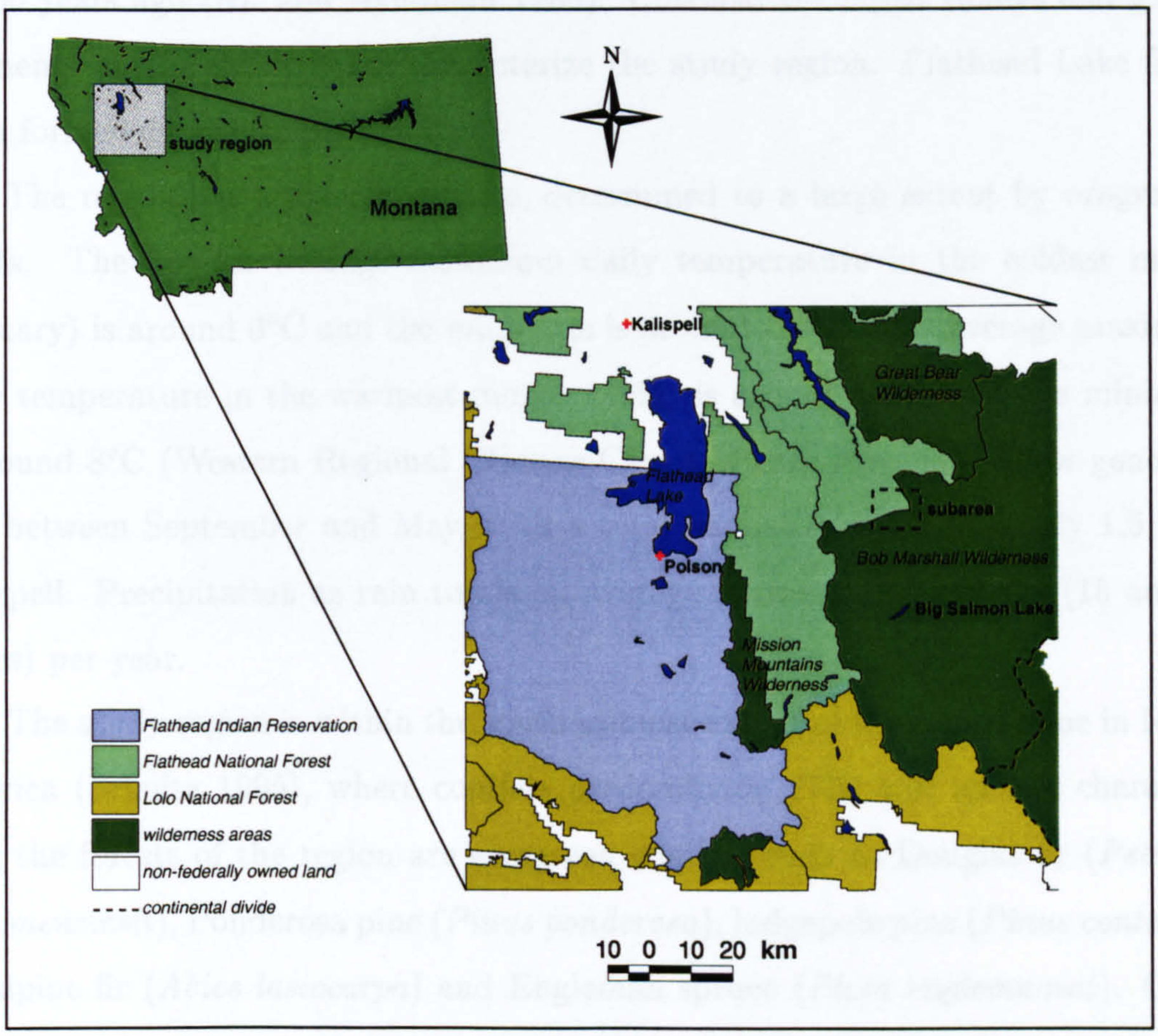
In chapter 4, aspatial regression was contrasted with two geostatistical methods, cokriging and conditional simulation in a study using synthetic data as an analog to the remote sensing of vegetation amount. These analog data allowed direct control over issues of spatial support, measurement error and sampling design so that prediction methods could be applied and compared easily. With real data, these factors often present major challenges to analysis. It is therefore important to investigate the three prediction methods in the context of real data to see how the outcomes are affected by these factors. This chapter is an analysis of actual vegetation amount and remotely sensed data from a region surrounding Flathead Lake, Montana. The primary vegetation amount variable was canopy cover (%) measured at an unusually large number of locations by the US Forest Service (USFS) during two seasons. The ancillary variable was calculated using spectral reflectance data from two different satellite sensors (measuring at two different spatial supports), Landsat TM and NOAA AVHRR.

5.1 Methods

5.1.1 Study region

The study region is located within western Montana, USA, just west of the Continental Divide and surrounding Flathead Lake, the largest natural freshwater lake in the western United States (figure 5.1). The region includes federally designated wilderness, where commercial developments, roads, motorized transportation, buildings or other permanent structures and commercial logging are prohibited. Wilderness areas include the Bob Marshall Wilderness, approximately 385,000 hectares (950,000 acres) north of Lolo National Forest and the Mission Mountain Wilderness, 31,000 hectares (76,000 acres) comprising a portion of the Swan Mountain range and located east of the Flathead Indian Reservation (Graetz 1985). The national forests including the Flathead and Lolo (see key in figure 5.1), where logging occurs through leases to private companies, comprise more than 1.6 million hectares (four million

Figure 5.1: The study region location within the state of Montana; inset shows major park and national forest boundaries and major lakes. The subarea outlined with dashed line is described in section 5.1.5. Source: Digital coverages from the Montana State Library, Helena, MT.



acres). The two largest settlements in the region are Kalispell (population around 12,000) just north of Flathead Lake and Polson (population around 3,500) just south of it (US Census Bureau 1990 estimate).

The oldest recognized geological formation in this region is the Precambrian Belt formation, comprised of primarily sedimentary rock which extends beyond western Montana to northern Idaho and into British Columbia (Fields 1971). This formation, along with sedimentary rock formed during Paleozoic and Mesozoic eras, was reshaped during the collision of the North American continent with the Pacific plate that began 175 million years ago. Particularly rapid tectonic movement 70 to 90 million years ago is presumed to have created the Rocky Mountain range and the overthrust belt, where slabs of older Belt rock lie over younger formations at a multitude of angles. These processes generated the modern-day relief of the

area, ranging from around 800 m to around 3,000 m above sea level. Other major geological features of the region were formed in the late Pleistocene, with two major glaciations called the Bull Lake and Pinedale, the latter ending approximately 10,000 years ago (Alt and Hyndman 1986). Classical U-shaped valleys and glacial sediments on the valley floors characterize the study region. Flathead Lake fills a basin formed during glaciation.

The region has a diverse climate, determined to a large extent by orographic effects. The 30-year average maximum daily temperature in the coldest month (January) is around 0°C and the minimum is around -10°C. The average maximum daily temperature in the warmest month (July) is around 26°C and the minimum is around 8°C (Western Regional Climate Center, Reno, Nevada). Snow generally falls between September and May, with a total annual snowfall of nearly 1.5 m in Kalispell. Precipitation as rain totals on average between 38 to 76 cm (15 and 30 inches) per year.

The study region is within the southernmost extent of the boreal zone in North America (Schultz 1995), where conifers predominate. The tree species characterizing the forests of the region are evergreen conifers such as Douglas fir (*Pseudotsuga menziesii*), Ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), Subalpine fir (*Abies lasiocarpa*) and Engleman spruce (*Picea englemannii*). Other prevalent species include the deciduous conifer larch (*Larix occidentalis*) and deciduous broadleaves such as black cottonwood (*Populus trichocarpa*) and aspen (*Populus tremuloides*). Patterns of forest growth are controlled primarily by aspect, altitude, fire history, land use practice and hydrology (figure 5.2). South-facing slopes tend to have more open canopies with less cover than north-facing slopes (Lathrop and Pierce 1991). Clearcuts, both recent and decades old, often have much lower cover values than surrounding tracts and are often the most visible features on the landscape. Riparian areas are often zones of high canopy cover because of increased water availability. The understory typically is composed of grass or shrub species; common examples are beargrass *Xerophyllum tenax* or snowberry *Symphoricarpos albus* (Jensen *et al.* 1992).

Figure 5.2: View of a typical watershed in the study region, showing variation in canopy cover governed by land use, aspect and elevation.

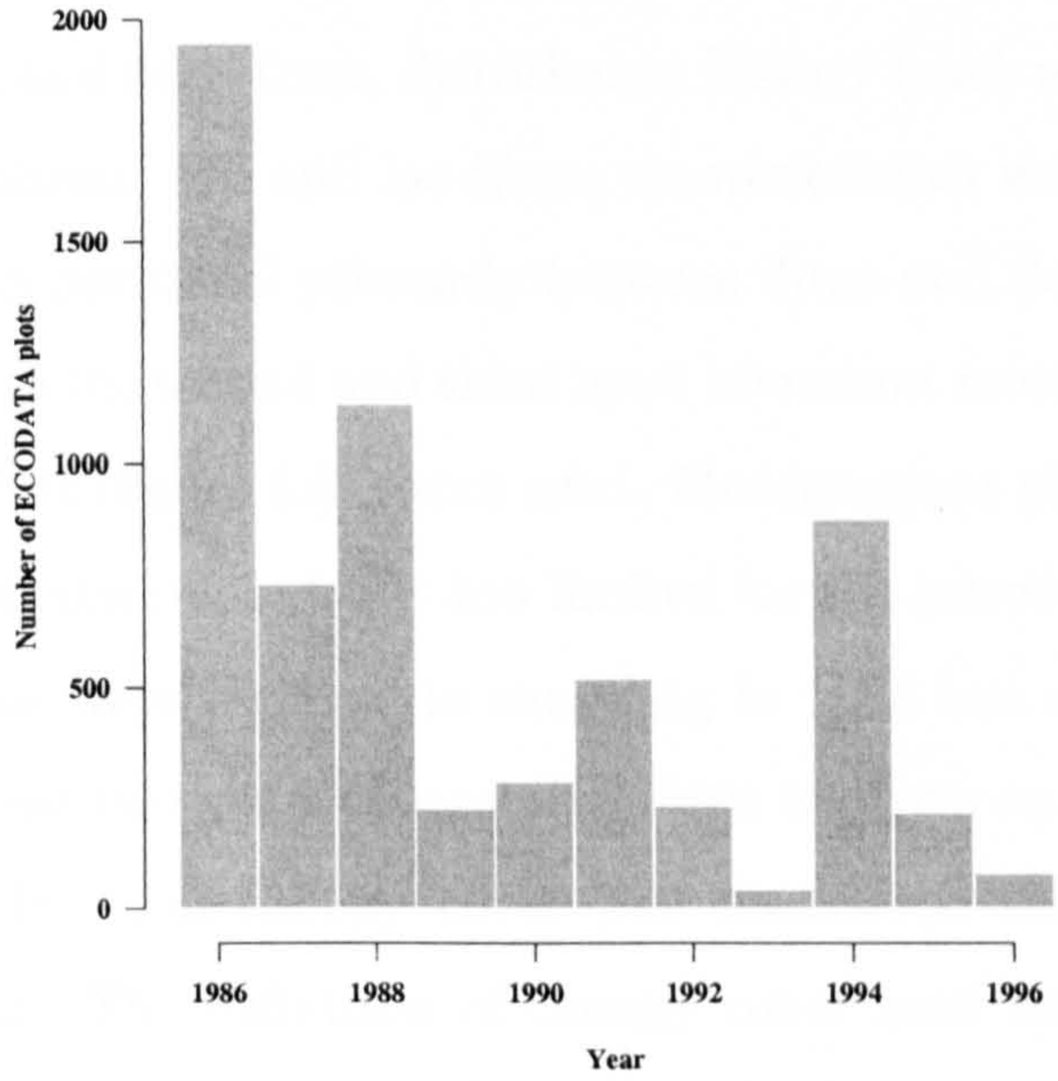


5.1.2 Ground measurements

Investigations of relationships between vegetation amount and remotely sensed spectral data have rarely had access to large numbers of ground measurements, especially in forest environments. Forest measurements are particularly time-consuming and labor intensive both because of the logistics of travel to roadless areas and because of the inherent difficulties in characterizing forest stands. Because governments can deploy greater resources to achieve larger sample sizes, government-sponsored sampling campaigns can be more suited to regional analysis with remotely sensed data than the limited field campaigns most commonly reported in chapter 2. In this study, the USFS's Ecosystem Inventory and Analysis ('ECODATA') database (Jensen *et al.* 1992) was used as a source of plentiful ground measurements. The USFS has been developing standard protocols for forest survey and has built up a database on over 6,000 plots (figure 5.3) in and beyond the study region. The database includes observations of vegetation amount and therefore provides a valuable opportunity for testing relationships and developing predictive models.

The ECODATA database includes information from surveys that took place between 1986 and 1996. Surveys were undertaken for different reasons, such as

Figure 5.3: The number of ECODATA plots surveyed between 1986 and 1996.



vegetation, soil or riparian inventories, maps of existing vegetation or site potential. The database was constructed from the surveys using a standardized coding scheme. Depending on the purpose for a given survey, different quantities were measured at each plot location, though all shared a basic set (Jensen *et al.* 1992). As far as a spatial sampling design most of these surveys can be considered purposive, since plot locations were not chosen as a probability sample. Locations were chosen subjectively, but without any preconceived bias (Mueller-Dombois and Ellenberg 1974). The criterion of a 'relatively uniform plant community and environment' was used for specific selection of plots (page 2–1, Jensen *et al.* 1992). The basic spatial sampling unit for ECODATA ground measurements used in this study is the 'macroplot', a circular area with an 11.3 m (37 foot) radius. The support of these measurements is therefore about 400 m^2, or about one-tenth of an acre.

There were over 200 fields of information for each plot stored in the database. For example, location information was held in twelve fields, with fields for latitude and longitude degrees, minutes, seconds and hundredths of seconds, northings and eastings, Universal Transverse Mercator (UTM) zone and datum year. In the first years of the survey, plot locations were obtained by reading 1:24,00 scale maps and

recording UTM coordinates with the 1927 North American Datum (NAD27). In more recent surveys, locations were obtained exclusively using Global Positioning System (GPS) units and recorded in latitude/longitude. At each location, data on dominant and co-dominant overstory and understory species, size and height structure of living and dead trees, disturbance history (such as fire and management practices), and habitat, soil and landform characteristics were recorded. Sampling was done by USFS personnel primarily between June and September in each year.

The years with the second and third most abundant measurement plots, 1083 in 1988 and 876 in 1994 (figure 5.3), were used. Measurement plots were most plentiful in 1986 but their spatial extent was too limited for the purposes of this study. More than twenty people were involved in sampling in 1988 but only two individuals in 1994. Treating these two years separately allows analysis on each to be compared.

For this study, total canopy cover was considered the variable representing vegetation amount. The definition of canopy cover used for the ECODATA measurements was 'The percentage of the ground included in a vertical projection of imaginary polygons drawn about the total natural spread of foliage of the individuals of a species' (page 43, Daubenmire 1968). Separate estimates of tree canopy cover for saplings, poles and trees were made in each stand, as well as the total. The estimates were made by eye by trained personnel. Different individuals were involved in sampling in each year, with certain experienced individuals providing continuity across several years. The measurement scale for canopy cover (Bonham 1989) has an interval between recorded values around 10% throughout most of the range with a greater precision in the sparsest canopies (table 5.1). The stated accuracy standard is $\pm$ one class (Jensen *et al.* 1992) though no information was available to verify whether this was achieved.

The distributions of the 1988 and 1994 canopy cover data were quite similar in shape (figure 5.4*a* and *b*) and clearly nonnormal. There appeared to be a larger proportion of small values in 1994 compared to 1988 (figure 5.4*c*). The Kolmogorov-Smirnov test rejects, at the 99% confidence level, that the two data sets come from the same distribution. The lack of randomness in the spatial sampling design makes it difficult to say with confidence that these plots were representative of the population of canopy cover values. The larger 1988 survey included more plots that were

Table 5.1: Codes used for ranges of canopy cover values in ECODATA.

Class code	Range of cover (%)
00	0
01	0– <1
03	1– <5
10	5– <15
20	15– <25
30	25– <35
40	35– <45
50	45– <55
60	55– <65
70	65– <75
80	75– <85
90	85– <95
98	95– 100

farther afield than the 1994 survey. Possible explanations for the observed difference in the canopy cover distribution are an actual decrease in canopy cover over these six years, sampling bias because of the different individuals making the measurements or simply because the geographical locations of the ECODATA plots were different in the two years.

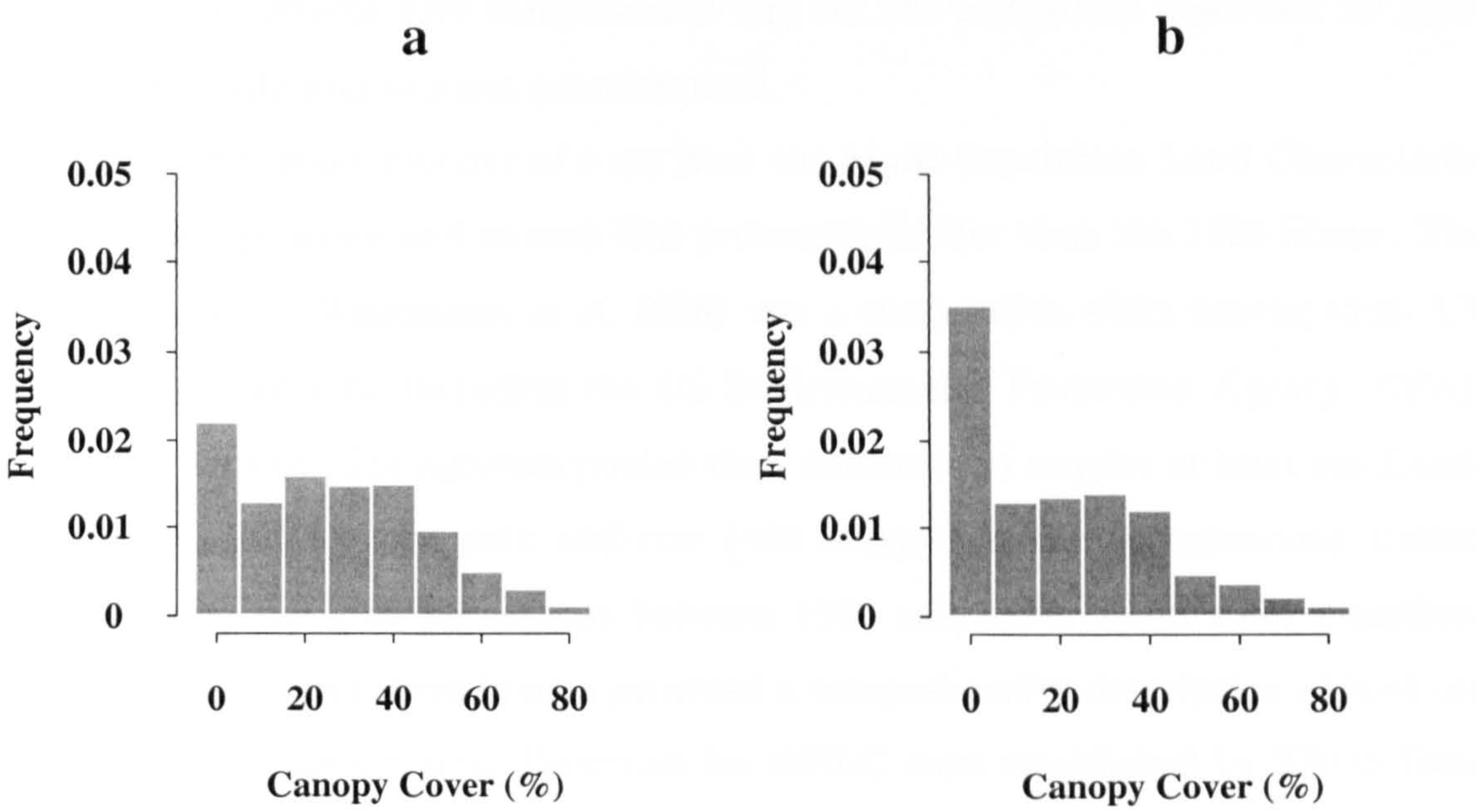
5.1.3 Remotely sensed data

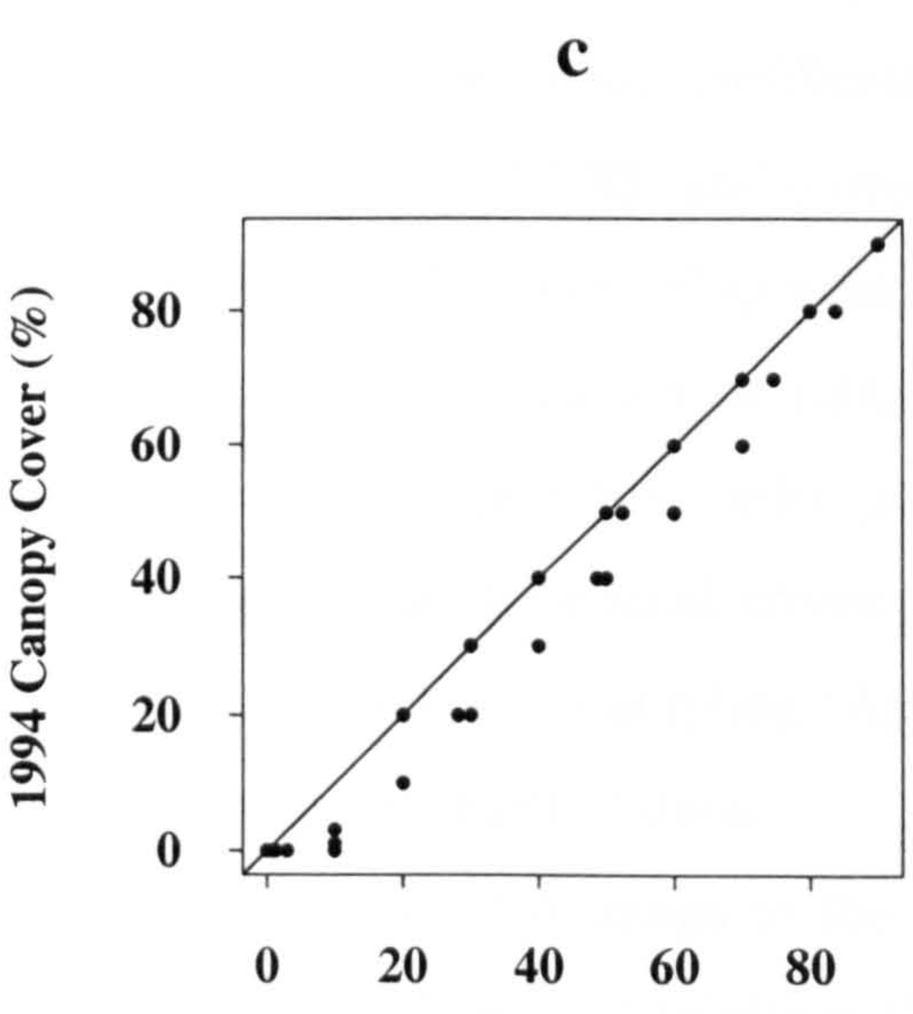
Landsat Thematic Mapper data

Two Landsat TM images from Path 41 Row 27 were acquired, one from August 28, 1988 and the other from September 27, 1993. The area represented by these images extends from 115°30'W to 113°W and 46°30'N to 48°15'N and covers the study region. The data were supplied by the US Geological Survey's (USGS) Earth Resources Observation Systems (EROS) Data Center. The thermal waveband (band 6) was not used. Because the number of ECODATA measurements collected in 1993 was insignificant, the 1993 scene was considered to be coincident to the 1994 ECODATA measurements. When these data are discussed together herein, they are sometimes referred to as '1993' data for convenience.

Preprocessing steps used to obtain reflectance values from these images at the ECODATA plot locations include geometric and radiometric processing. The

Figure 5.4: (a) Histogram of the 1988 canopy cover data, (b) histogram of the 1994 canopy cover data and (c) quantile-quantile plot of the 1994 vs 1988 canopy cover data.





two images supplied by EROS represented different levels of processing along this path. The 1988 image was produced by the EROS Digital Image Processing System (EDIPS, US Geological Survey 1977). Effects of nonlinear scanning by the sensor's mirror, sequential detector sampling, detector offsets, Earth curvature and rotation and other effects were compensated for, but the image was described by pixel coordinates only and was not georeferenced.

The 1993 image was one of a set from the Multi-Resolution Land Characteristics (MRLC) program and as such was processed further than the 1988 image. The MRLC program (Vogelmann *et al.* 1998) was a cooperative effort among three US Government agencies, including the US Environmental Protection Agency (EPA), USGS and NOAA. The agencies pooled their resources to acquire at least one Landsat TM image for every path and row (430 images) in the conterminous United States during the growing seasons between 1990 and 1994. Land cover classifications generated from these images provided a comprehensive description of land use patterns across the country. Protocols for MRLC data established by EROS Data Center guaranteed a more uniform set of geometric and radiometric processing steps than would be possible under other circumstances. These steps included systematic correction as in the 1988 image, debanding to reduce the variation across the 16 detectors (Jensen 1996), rectification to a UTM projection with the 1983 North American Datum (NAD83) and correction for relief displacement (Vogelmann *et al.* 1998). EROS used at least 20 ground control points (GCPs) collected from 1:100,000 Digital Line Graph data and/or 1:24,000 USGS topographic maps and accomplished the rectification with a first-order polynomial model. Twelve or more GCPs were used to check that locational errors were on average less than 1 pixel. Cubic convolution was used for resampling (Appendix A). A digital elevation model was also provided with the MRLC data.

To bring the 1988 image to the same projection as the 1993 image, the latter was used as a reference, exploiting the fact that MRLC processing achieved rectification with small errors. Twenty-five spatially well-distributed GCPs in the 1988 image were chosen that corresponded with pixels in the MRLC image. These points were used in a collinearity solution that includes the effects of relief displacement (ERDAS 1997). According to Pala and Pons (1995), relief displacement could cause

Table 5.2: TM calibration coefficients in units of $W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$ applied in equation 5.1.

Scene Date	Coefficient	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
1988	L_{min}	0	0	0	0	0	0
	L_{max}	153.6	299.6	205.5	207.7	27.56	14.53
1993	L_{min}	-1.50	-2.80	-1.20	-1.50	-0.37	-0.15
	L_{max}	152.1	296.8	204.3	206.8	27.19	14.38

location errors of more than 10 pixels in this scene, given the high elevations present (maximum elevation = 2,890 m). Nearest-neighbor resampling was used to minimize the support change that comes with smoothing. An average error of 0.41 pixels (12.3 m) occurred with the 25 control points and 0.49 pixels (14.7 m) with six independently chosen test points.

A deterministic approach was chosen to remove atmospheric path radiance and estimate surface reflectance (%) from the digital numbers (DNs) in each TM band for both years. The first step in this procedure was to convert the DNs to at-sensor radiance. Several studies have attempted to quantify the radiometric degradation in TM sensitivity over time (Moran *et al.* 1995, Olsson 1995, Thome *et al.* 1997). The calibration coefficients of Olsson (1995) were used to obtain radiance (L) in $W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$ (table 5.2) for each band:

$$L = L_{min} + (\frac{L_{max} - L_{min}}{255})DN \tag{5.1}$$

Olsson's study encompassed 1988 to 1991, so there was less confidence in extending his results beyond 1991 to the 1993 scene. In fact, the coefficients when extrapolated to 1993 imply negative at-sensor radiances in TM bands 5 and 7 for dark objects. These were set to 0 in subsequent analysis.

From at-sensor radiance, assuming a Lambertian surface under cloudless conditions, surface reflectance can be calculated using the equation:

$$\rho = 100\frac{\pi(L - L_p)/E}{1 + s\pi(L - L_p)/E} \tag{5.2}$$

where L_p is the upwelling radiance of the atmosphere, also called path radiance; s is the proportion of radiation backscattered by the atmosphere and E is the

downwelling irradiance (Chandrasekhar 1960). Since L_p, E and s all depend on atmospheric conditions that vary, especially aerosol content, many efforts to derive surface reflectance from remotely sensed data rely on coincident measurements of these conditions (Holm *et al.* 1989, Spanner *et al.* 1990, Moran *et al.* 1992). In this study, coincident measurements were not available. In addition, there was evident spatial variation of atmospheric conditions in the the images, so the measurement of atmospheric optical properties at a single location was not sufficient. For these reasons, the algorithm of Liang *et al.* (1997) seems promising. This whole-scene correction algorithm considers aerosol optical depth as spatially variable and estimates it from dark objects found within small windows. No independent estimates of atmospheric conditions are needed for this algorithm. However, problems with the software implementation of the algorithm (Liang, personal communication) prevented its use for this study.

An alternative whole-scene correction algorithm was found in Wrigley *et al.* (1992). This algorithm contained several simplifications to allow correction on millions of pixels within a reasonable amount of time. It assumed spatially invariant aerosol optical depth and irradiance, neglected diffuse irradiance (skylight) and described Rayleigh and aerosol scattering separately (the 'single-scattering' approximation). The algorithm was applied by Wrigley *et al.* (1992) on data from FIFE[1] and HAPEX[2], achieving accuracy of within 10% of measured values of surface reflectance. It was applied to the 1988 and 1993 TM scenes to estimate surface reflectance.

A visibility value of 10 km was reported at the Kalispell airport (located just north of Flathead Lake) on September 27 1993 (National Climate Data Center, Asheville, North Carolina). It is a common practice at airports to use 10 km as the maximum visibility value and in fact none of the values in the meteorological database show values exceeding 10. A safe inference from this information was that the visibility on the 1993 date was equal to or greater than 10 km. No historical visibility data were available from the 1988 data.

To obtain estimates for aerosol optical depth, the atmospheric radiative transfer code 6S (Second Simulation of the Satellite Signal in the Solar Spectrum, Vermote

[1]First ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment, (Sellers *et al.* 1992)

[2]Hydrologic and Atmospheric Pilot Experiment, (Prince *et al.* 1995)

et al. 1997) was used to characterize dark objects in the scenes. Big Salmon Lake, a glacial lake within the study region (figure 5.1) at an elevation of 1,313 m, was used as a dark object. In an approach similar to that of Olsson (1995), 6S was run iteratively, changing the visibility parameter until surface reflectance in TM band 1 fell between 2 and 3%. The calibration coefficients from table 5.2 yielded at-sensor radiances for the darkest pixel in this lake that were remarkably similar for the two scene dates (table 5.3). Accordingly, a visibility value of 15 km was used for both the 1988 and 1993 scenes. Aerosol optical depths used in the Wrigley *et al.* (1992) algorithm were therefore the same for the two dates and the irradiance differences were a function of solar position (table 5.3).

The atmospheric correction program was run using the parameters obtained from 6S. Water vapor was neglected. Output (o) from the program in two-byte format was transformed to reflectance (%) using

$$\rho = 100o/(1 + so) \tag{5.3}$$

The Normalized Difference Vegetation Index (NDVI) was calculated for the TM images as

$$NDVI = (\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R) \tag{5.4}$$

where ρ_R is reflectance in band 3 (0.63-0.69 μm) and ρ_{NIR} is reflectance in band 4 (0.76-0.9 μm).

Table 5.3: Implications from 6S for dark object characteristics.

Scene Date	Parameter	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
1988	At-sensor radiance ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$)	36.70	23.57	12.97	9.34	.22	0..11
	Reflectance (%)	2.6	3.2	1.4	4.0	0.10	0.7
	Aerosol optical depth (dimensionless)	0.36	0.30	0.26	0.19	0.08	0.05
	Total irradiance ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$)	707.7	678.5	618.8	445.1	96.67	36.54
	Backscatter ratio (dimensionless)	0.17	0.13	0.10	.07	.02	.01
1993	At-sensor radiance ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$)	35.16	20.97	10.89	7.39	0	0
	Reflectance (%)	2.8	2.4	.4	2.8	0	0
	Aerosol optical depth (dimensionless)	0.36	0.30	0.26	0.19	0.08	0.05
	Total irradiance ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$)	602.68	577.68	531.83	386.43	84.92	32.15
	Backscatter ratio (dimensionless)	0.17	0.13	0.10	.07	.02	.01

NOAA AVHRR data

NOAA AVHRR data were acquired as near to the dates of the TM images as possible. Images judged to be of good quality from August 27, 1988 from the NOAA-9 satellite and September 29, 1993 from the NOAA-12 satellite were in the EROS archive. Geometric and radiometric processing was done by EROS according to Eidenshink and Faundeen (1994), including calibration according to Teillet and Holben (1993). No information was given in Eidenshink and Faundeen (1994) about the resampling method used. The resulting images represented reflectance in each of the five AVHRR bands. The map projection used by EROS was UTM with NAD27 datum.

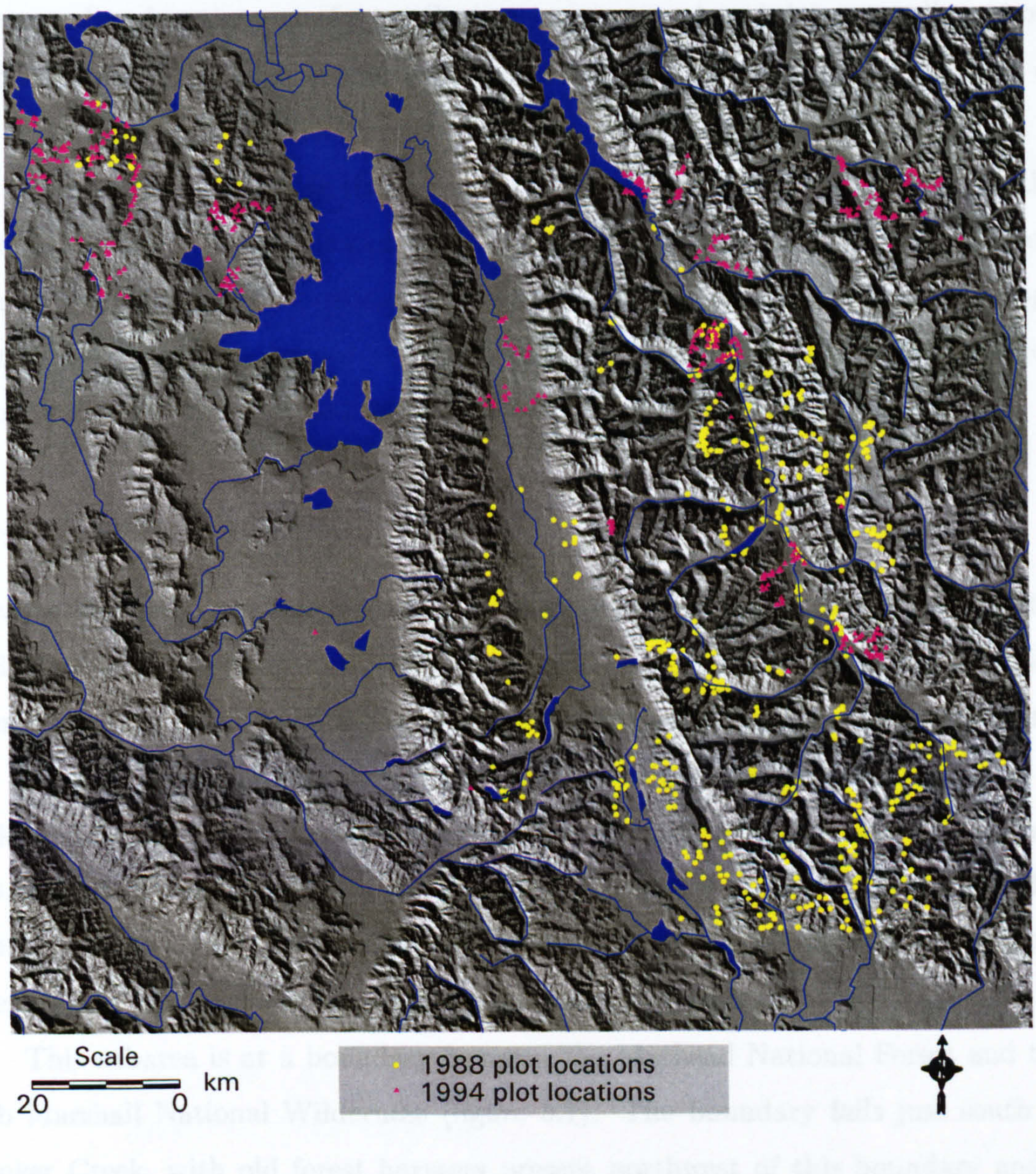
The NDVI was also calculated for the AVHRR images using equation 5.4 where ρ_R is reflectance in band 1 (0.58-0.68 μm) and ρ_{NIR} is reflectance in band 2 (0.725-1.0 μm).

5.1.4 Collocation of ground and remotely sensed data

Coordinate transformations were needed to find corresponding ECODATA plots and remotely sensed measurements at pixels. The 1988 ground locations were originally recorded in UTM with NAD27 datum. These points were transformed to the NAD83 datum using the High Performance GPS Network conversion method in Arc/Info (ESRI, Redlands, CA) to match the TM image data. The 1994 ground data were recorded in latitude and longitude to the nearest hundredth of a second. These were projected to UTM with NAD83 for comparison with TM and to UTM with NAD27 for comparison with AVHRR data. In this region of the country the effective distance between coordinates expressed in the NAD83 datum and those expressed in the NAD27 datum is 220 to 230 m. The number of ECODATA plots occurring within the TM image area was 335 for 1988 and 344 for 1993.

Plot locations used in this study on a representation of a Digital Elevation Model (figure 5.5) show that the majority of the plots were in the Swan Valley and South Fork of the Flathead River drainage. Others were in the upper reaches of the Middle Fork of the Flathead River drainage a few miles west of the Continental Divide, with another cluster east of Flathead Lake in the Salish range.

Figure 5.5: Locations of 1988 and 1994 ECODATA plots on a relief-shaded representation of a digital elevation model for the study region.



To increase confidence that the image data and ground data were in the same coordinate system, elevation values recorded in ECODATA were compared with elevations in corresponding pixels from the DEM. The mean elevation discrepancy was less than 10 m, less than the contour interval for the 1:24,000 scale maps from which DEM's are generated; this was taken as an indication the collocation was successful. The root-mean-square-error (RMSE) was 60 m; the correlation between the two measures was 0.93. (figure 5.6*a* and *b*). Differences between the values at corresponding locations can be ascribed to measurement and data-recording errors by field survey personnel, errors in the DEM and the different spatial supports of the DEM data and ECODATA observations as well as misregistration. Though it was not possible to separate these sources of error in this study, they indicate the problems inherent in ground and image data comparison even with a precisely estimable variable such as elevation. Discrepancies with aspect (figure 5.6*c*) and slope (not shown), variables derived from elevation, were much larger than for elevation alone, indicating the expected increase in error inherent in taking derivatives.

5.1.5 Choice of subarea

Because the study region is large (approximately 12 million TM pixels) and geostatistical methods are computationally intensive, a small subarea was chosen on which to make canopy cover predictions. It was chosen to coincide with the largest number of ground measurements per area to allow enough neighboring measurements to make geostatistical interpolation feasible. This area was 9,700 ha, represented by a 348×310 pixel square in a TM image (figure 5.7). Every pixel is considered in this approach, even those that do not include forest cover. Pixels that do not contain forest should be output as 0% cover by prediction methods.

This subarea is at a boundary between the Flathead National Forest and the Bob Marshall National Wilderness (figure 5.1). The boundary falls just south of Bunker Creek, with old forest harvests present northwest of this boundary and a lack of such harvests southeast of the boundary (figure 5.8). There were 30 plots surveyed in this area in both 1988 and 1994. The density of plots may be explained by its accessibility from an aircraft landing strip south of the Harrison Creek/South

Figure 5.6: (a) Scatterplot of elevations for ECODATA plot locations versus those for corresponding DEM pixels, (b) Histogram of differences between plot elevations and DEM elevations and (c) scatterplot of aspects for ECODATA plot locations versus those for corresponding DEM pixels. The 1:1 line is shown in (a) and (c).

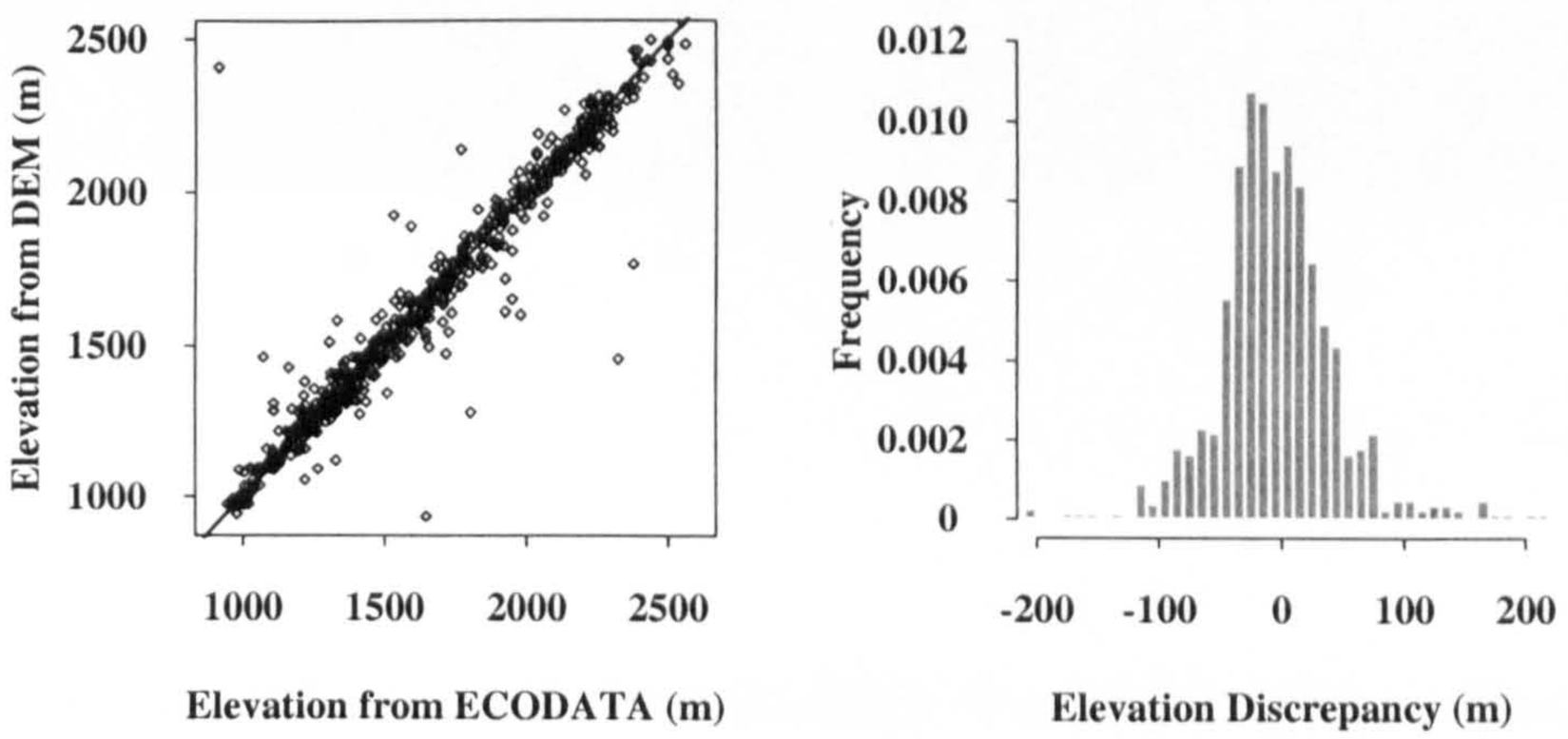
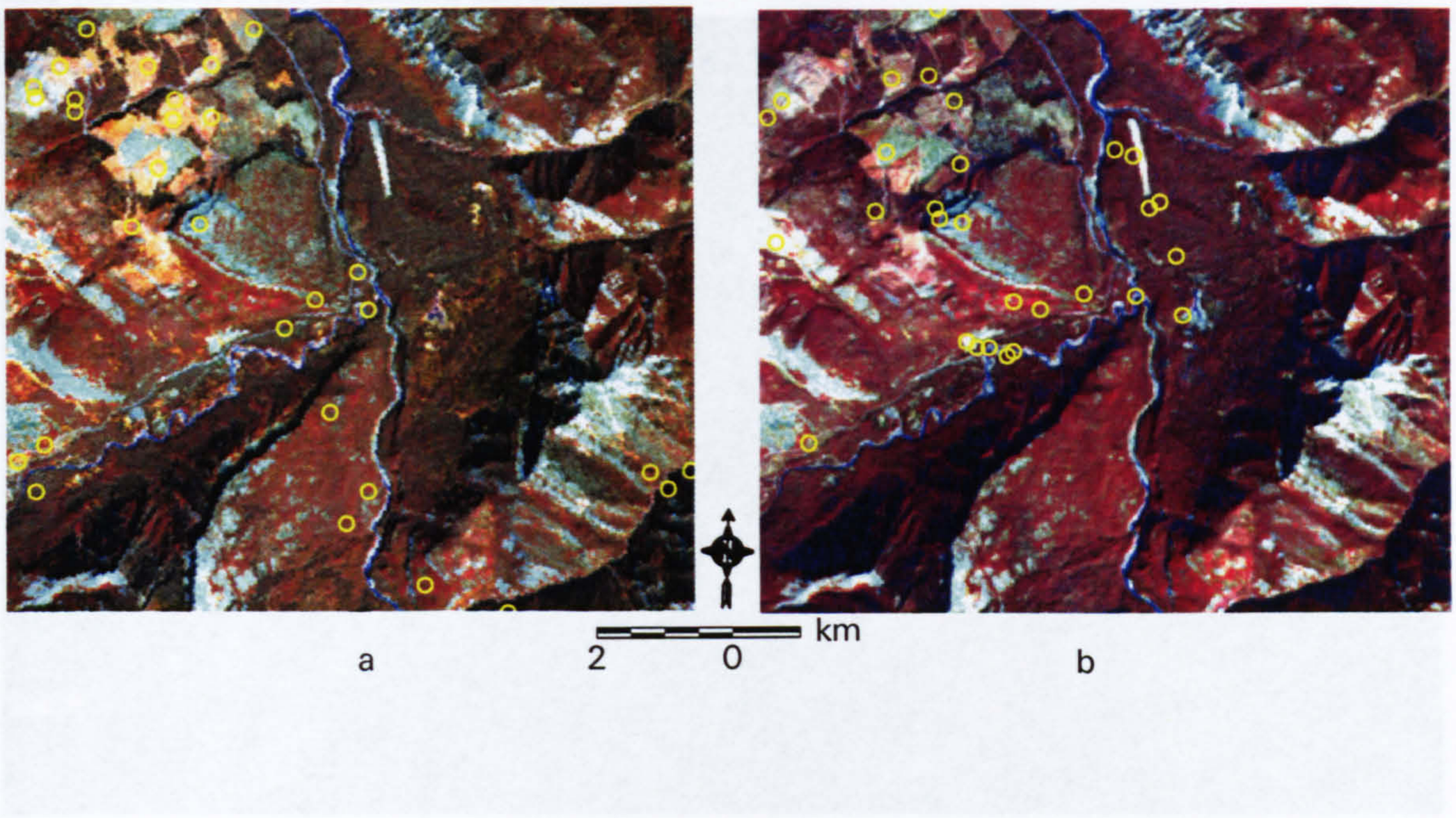


Figure 5.7: Subarea from TM false color composite (Bands 4, 5 and 3 in red, green and blue, respectively) from (a) 1988 with 30, 1988 ECODATA plot locations in open circles and b) 1993 with 30, 1994 ECODATA plot locations in open circles.



Fork Flathead confluence and the availability of a nearby cabin at Meadow Creek for lodging during survey work. Canopy cover at the ECODATA plots within the subarea (figure 5.9) included values from 0 to 60% with the highest values of cover measured in the Bob Marshall Wilderness (southern portion of images shown in figure 5.7) in 1988.

5.1.6 Aspatial regression

With six wavebands and multiple spectral index definitions (Teillet *et al.* 1997) from TM there is a large number of potential explanatory variables for an aspatial regression model. All of these variables have some causal relation to vegetation amount, since foliage absorbs and reflects differently than background does in all wavebands. A fundamental problem with using more than one spectral variable in a regression model is that spectral wavebands tend to be correlated with one another. Regression models are more likely to be robust when the explanatory variables are not highly correlated (Draper and Smith 1998). In the subarea, correlations between spectral reflectance in the six wavebands ranged from .51 to .98 (table 5.4). Using

Figure 5.8: 1993 TM image in figure 5.7 'draped' onto elevation model of the subarea. (a) Harrison Creek, (b) USFS landing strip, (c) Meadow Creek cabin, (d) Bunker Creek and (e) South Fork Flathead River.

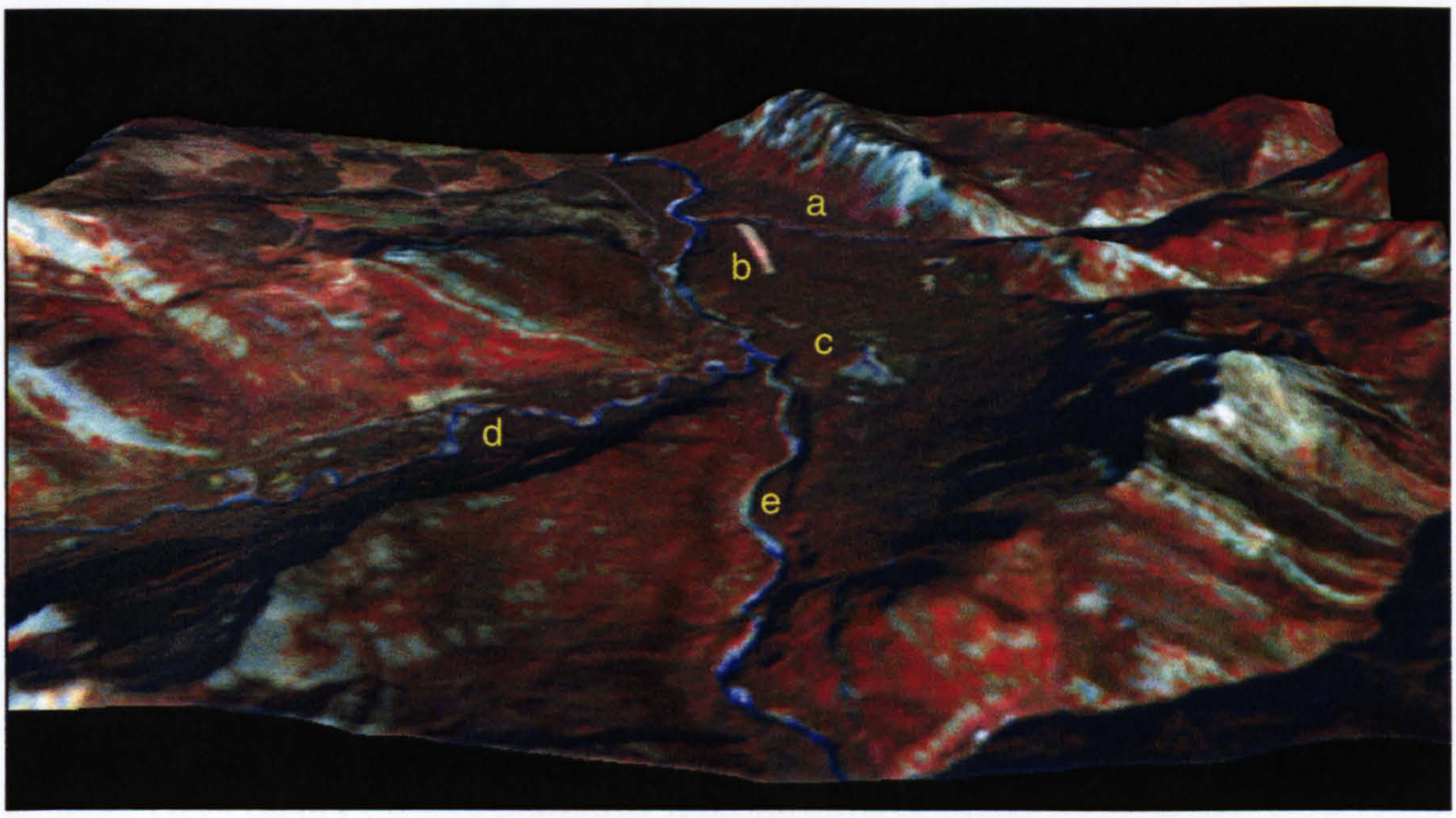
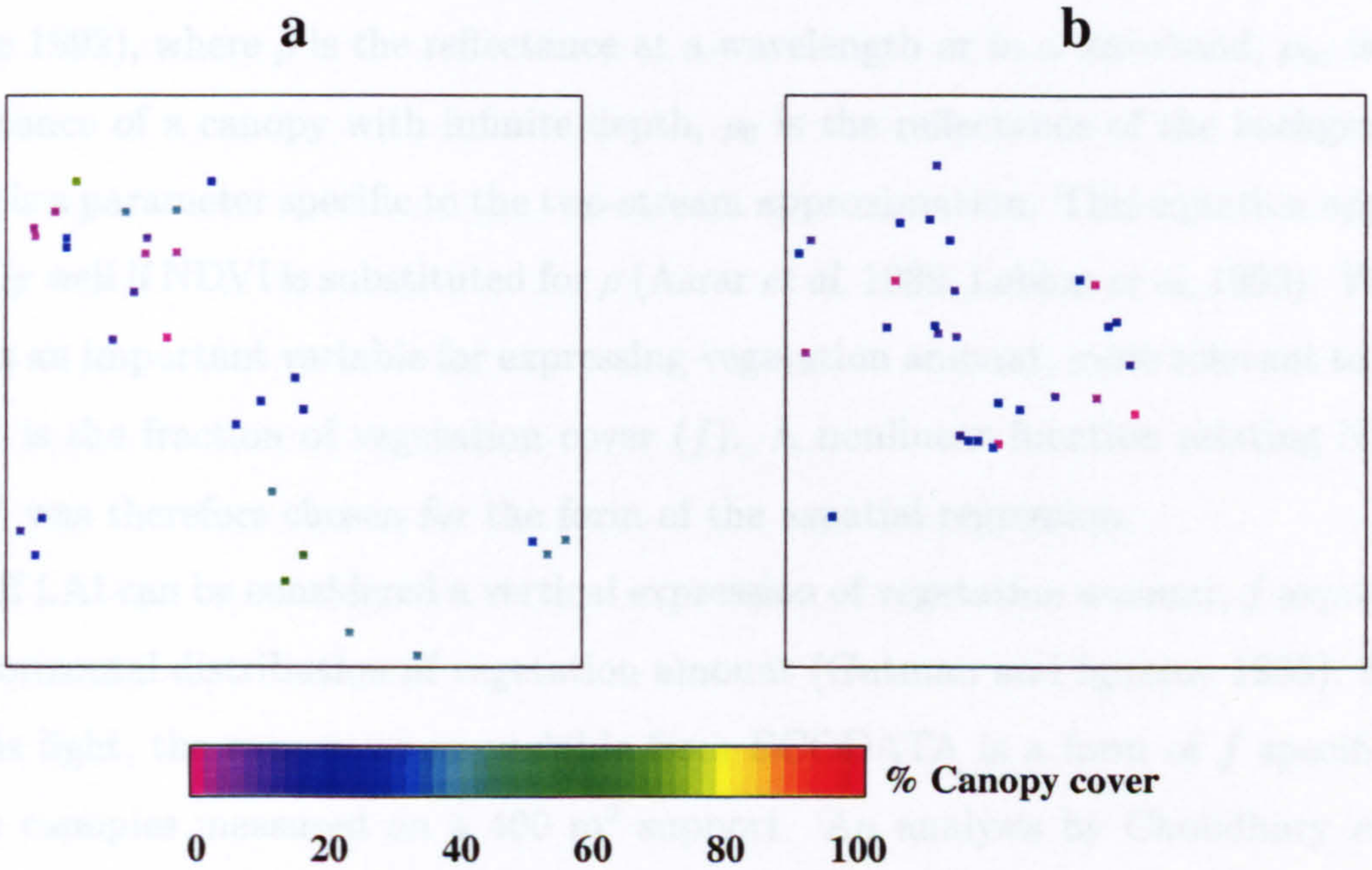


Figure 5.9: Values of canopy cover (%) at ECODATA plot locations in (a) 1988 and (b) 1994.



SOUTHAMPTON UNIVERSITY LIBRARY

Table 5.4: Linear correlation coefficients between reflectances in TM wavebands for pixels at 1988 ECODATA plot locations (n=335).

	ρ_{Band1}	ρ_{Band2}	ρ_{Band3}	ρ_{Band4}	ρ_{Band5}	ρ_{Band7}
ρ_{Band1}	1.00	0.97	0.97	0.45	0.86	0.92
ρ_{Band2}		1.00	0.98	0.58	0.91	0.94
ρ_{Band3}			1.00	0.51	0.92	0.96
ρ_{Band4}				1.00	0.70	0.53
ρ_{Band5}					1.00	0.96
ρ_{Band7}						1.00

more than one waveband or spectral index as an explanatory variable would have therefore introduced collinearity. Therefore, a single spectral variable was selected for the model and this was the NDVI.

Linear regression models are most commonly used in empirical studies of vegetation amount and optical remote sensing data (chapter 2). However, radiative transfer (RT) theory implies nonlinear relationships between spectral and vegetation amount variables (Asrar *et al.* 1989, Price 1992, Leblon *et al.* 1993). For example, the relationship between LAI and spectral reflectance based on a simple RT model is

$$\rho = \rho_\infty + (\rho_0 - \rho_\infty)e^{2cLAI} \tag{5.5}$$

(Price 1992), where ρ is the reflectance at a wavelength or in a waveband, ρ_∞ is the reflectance of a canopy with infinite depth, ρ_0 is the reflectance of the background and c is a parameter specific to the two-stream approximation. This equation applies equally well if NDVI is substituted for ρ (Asrar *et al.* 1989, Leblon *et al.* 1993). While LAI is an important variable for expressing vegetation amount, more relevant to this study is the fraction of vegetation cover (f). A nonlinear function relating NDVI and f was therefore chosen for the form of the aspatial regression.

If LAI can be considered a vertical expression of vegetation amount, f expresses the horizontal distribution of vegetation amount (Gutman and Ignatov 1998). Seen in this light, the canopy cover variable from ECODATA is a form of f specific to forest canopies measured on a 400 m^2 support. An analysis by Choudhury *et al.* (1994) suggests that NDVI is a power function of f:

$$NDVI' = [1 - f]^\epsilon \tag{5.6}$$

where $NDVI'$ is a 'scaled' NDVI, defined as

$$NDVI' = \frac{NDVI_\infty - NDVI}{NDVI_\infty - NDVI_0} \tag{5.7}$$

and ϵ is a ratio between the canopy extinction coefficient and a canopy geometry parameter (describing planophile, erectophile or randomly distributed leaves). Carlson and Ripley (1997) suggest that this is equivalent to

$$1 - NDVI' = f^{B_2} \tag{5.8}$$

where B_2 is about 0.5. The scaled NDVI is simply a linear transform of NDVI. If it is assumed that the $NDVI_0$ is near 0, then $\frac{NDVI_\infty}{NDVI_\infty - NDVI_0}$ is nearly 1 and the model simplifies to

$$NDVI = B_1 f^{B_2} \tag{5.9}$$

where B_1 is $NDVI_\infty - NDVI_0$. Interestingly, Chen and Cihlar (1996) also use a relation of this form for boreal forest stands and TM data, though they use 'effective LAI' rather than f as the explanatory variable.

Equation 5.9 was used to develop the regression model with canopy cover as the explanatory variable and NDVI as the response variable. Despite the fact that the ECODATA measurements and the TM data are subject to possibly large measurement error (Curran and Hay 1986, McGwire *et al.* 1993), this form was selected for two reasons: 1) the 'class' coding of canopy cover (table 5.1) facilitated estimation of the $E\{NDVI\}$ for each class of f and 2) a solution was numerically possible. When a direct approach was used without going through the equation 5.9 first, numerical errors occurred and equations were not consistent.

After the regression model was developed it needed to be inverted in order to predict canopy cover from NDVI. Data for 1988 and for 1993 were used to obtain separate models of the form

$$p = 100 * (\frac{a}{B_1})^{1/B_2} \tag{5.10}$$

where p is canopy cover, a is NDVI and B_1 and B_2 are coefficients. The software used to determine the coefficients of this regression model and calculate prediction

intervals was the package *nls2* (Huet *et al.* 1996) written for the Splus statistical software system (Mathsoft, Seattle, WA).

5.1.7 Cokriging

Separate LMCs were created for the 1988 and 1993 data. In both cases, LMCs were constructed by fitting the semivariograms and cross-semivariograms separately, balancing any lack of fit with the positive semi-definiteness criteria for the sills (equation 3.33). Models were fit using the S+SpatialStats function *model.variogram* (Mathsoft, Seattle, WA) and these criteria were checked. The LMCs were used to obtain the weights in the cokriging model, equation 3.23, where $\hat{p}(\mathbf{x}_o)$ is the prediction of canopy cover at location $\mathbf{x}_o$, $p(\mathbf{x}_j)$ are N_1 nearby ECODATA values at locations $\mathbf{x}_j$ weighted with factors λ_j and $a(\mathbf{x}_k)$ are N_2 nearby NDVI values at locations $\mathbf{x}_k$ weighted with factors ω_k. The cokriging weights were constrained using equation 3.28, effectively weighting the remotely sensed data more heavily than would the traditional biasedness constraints. The program used to accomplish the cokriging was *cokb3d* from GSLIB (Deutsch and Journel 1992). The search neighborhood for the primary data was set to 190 pixels or 5700 m, just beyond the range of the semivariogram. The search neighborhood for the secondary data was set to 50 pixels, or 1500 m. The minimum number of primary data was set to 1 so that pixels more than 5700 m away from a measurement location would not get a cokriged value. The maximum number of primary data used in the cokriging system was set to 8 and the maximum number of ancillary data was 4. A maximum of 4 meant that only the nearest neighbors of each pixel to be cokriged were used in the prediction.

5.1.8 Conditional simulation

As seen in chapter 3, numerous conditional simulation algorithms have been developed. For the study of synthetic data in chapter 4, probability field simulation was chosen for conditional simulation, with local ccdfs defined by cokriging. For this study, a special form of sequential Gaussian simulation was chosen to simulate canopy cover in the subarea. Instead of using full cokriging to construct the local

ccdf at each step, collocated cokriging (Almeida and Journel 1994, Goovaerts 1997) was used:

$$\hat{p}(\mathbf{x}_o) = \sum_{j=1}^{N_1} \lambda_j p(\mathbf{x}_j) + \omega_o a(\mathbf{x}_o) \tag{5.11}$$

Collocated cokriging uses only the ancillary data at the cell to be simulated ($a(\mathbf{x}_0)$), none of the surrounding ancillary data. The semivariogram model for the ancillary data was therefore not used in the collocated cokriging system. In addition, a Markov model (Almeida and Journel 1994) for the cross-semivariogram between primary and ancillary data was used. These options were available in the program *sgsim* in GSLIB (Deutsch and Journel 1998). Fifty realizations for 1988 and 1993 were generated.

5.2 Results and Discussion

The analysis of the over 300 ECODATA canopy cover measurements along with the collocated spectral data from TM and AVHRR in the study region yielded information about the spectral reflectance of conifer forests and how it relates to vegetation amount. The univariate, bivariate and spatial dependence characteristics of these data formed the basis for the predictions using aspatial regression, cokriging and conditional simulation. Global statistics were inferred from the study region and assumed to be applicable to the subarea. Canopy cover in the subarea, not the larger study region, was predicted using the three methods.

5.2.1 Reflectance of conifer forests

In the study region, reflectance in the red waveband (band 3) of TM at the ECODATA plot locations ranged primarily between 1 and 20% but with a number of pixels greater than 20%, especially in the 1988 data set (figure 5.10*a* and *b*). The mean values were 10% in 1988 and 8% in 1993; the standard deviations were 7% and 5% respectively. For the corresponding pixels from the AVHRR visible waveband (band 1), the spread of reflectance values was much narrower (figure 5.10*c* and d) with nearly all values falling below 15%. The AVHRR distributions lacked the long tail and appeared slightly more symmetric. Their mean values were 5% in 1988 and

4% in 1993, about half those of the TM distributions. The standard deviations were 6% and 1% respectively.

TM near infrared reflectance (band 4) at the ECODATA plot locations had a large range from 0 to over 60%, (figure 5.11*a* and *b*), with mean values of 34% in 1988 and 27% in 1993. The standard deviations were 13 and 9% respectively. The AVHRR band 2 distributions were much narrower (figure 5.11*c* and *d*), with standard deviations of 6% and 3% around means equivalent to those of the TM near infrared.

The NDVI from TM (figure 5.12*a* and *b*) at the ECODATA plot locations had a very similar mean value for the two dates (.57 in 1988 and .56 in 1993). The spread in values on both dates covered most of the positive range of NDVI, from .04 to .84, with equivalent standard deviations of .16. The mean NDVI values from AVHRR were .40 in 1988 and .32 in 1993. There was only a slight reduction in the spread of AVHRR NDVI in comparison to that of TM (figure 5.12) and the standard deviations of AVHRR NDVI were .14 and .08. Judging by the histograms, there is more variability and thereby 'information' in the AVHRR NDVI than in the AVHRR visible or near infrared reflectance alone.

TM and AVHRR can be regarded as measurements at two different supports (Atkinson 1995, Atkinson and Curran 1997), at 900 m^2 and 1.21 km^2. The change-of-support model (Matheron 1985) states that the mean value of a distribution of a random function should be conserved as support changes. The near infrared reflectance data from TM and AVHRR matched this model, since they had equivalent means. But the data from the red/visible wavebands and the NDVI did not coincide with this model. The difference in means between the TM band 3 and AVHRR band 1 may be explained by the difference in the sensors' spectral bandpasses and the unrelated radiometric preprocessing algorithms that were used (Frank *et al.* 1994). These discrepancies in sensor characteristics and preprocessing can help to explain the different means between TM and AVHRR NDVI as well. In addition, since NDVI is a nonlinear transform of reflectance, the areal mean NDVI is not the same thing as the NDVI of the areal mean (Aman *et al.* 1995, Hu and Islam 1997). Therefore, conservation of the mean with increasing support was not necessarily expected in this case.

Figure 5.10: Histogram of (a) TM band 3 reflectance at 1988 ECODATA plot locations, (b) TM band 3 reflectance at 1993 locations, (c) AVHRR band 1 reflectance at 1988 locations and (d) AVHRR band 1 reflectance at 1993 locations.

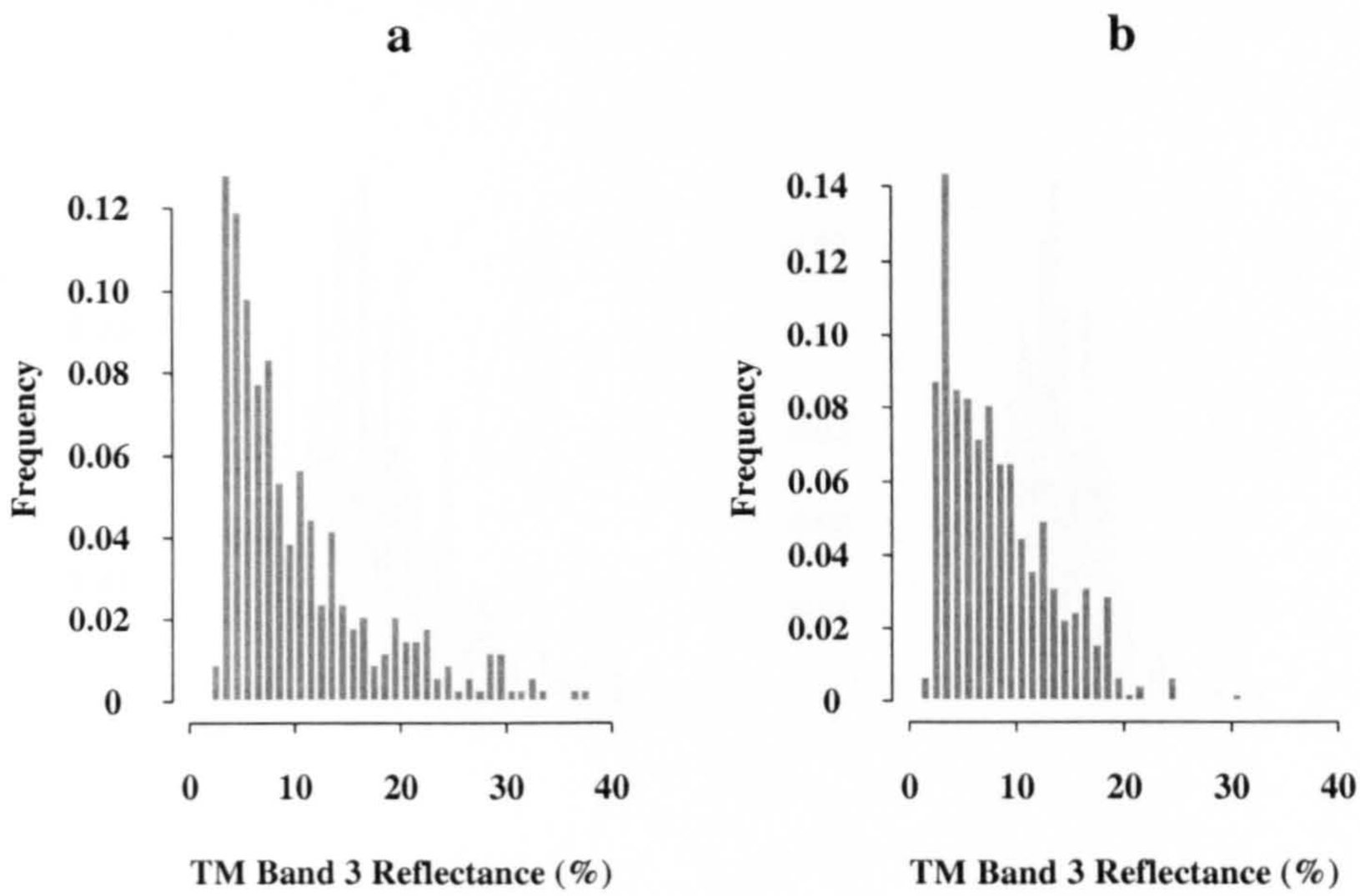


Figure 5.11: Histogram of (a) TM band 4 reflectance at 1988 ECODATA plot locations, (b) TM band 4 reflectance at 1993 locations, (c) AVHRR band 2 reflectance at 1988 locations and (d) AVHRR band 2 reflectance at 1993 locations.

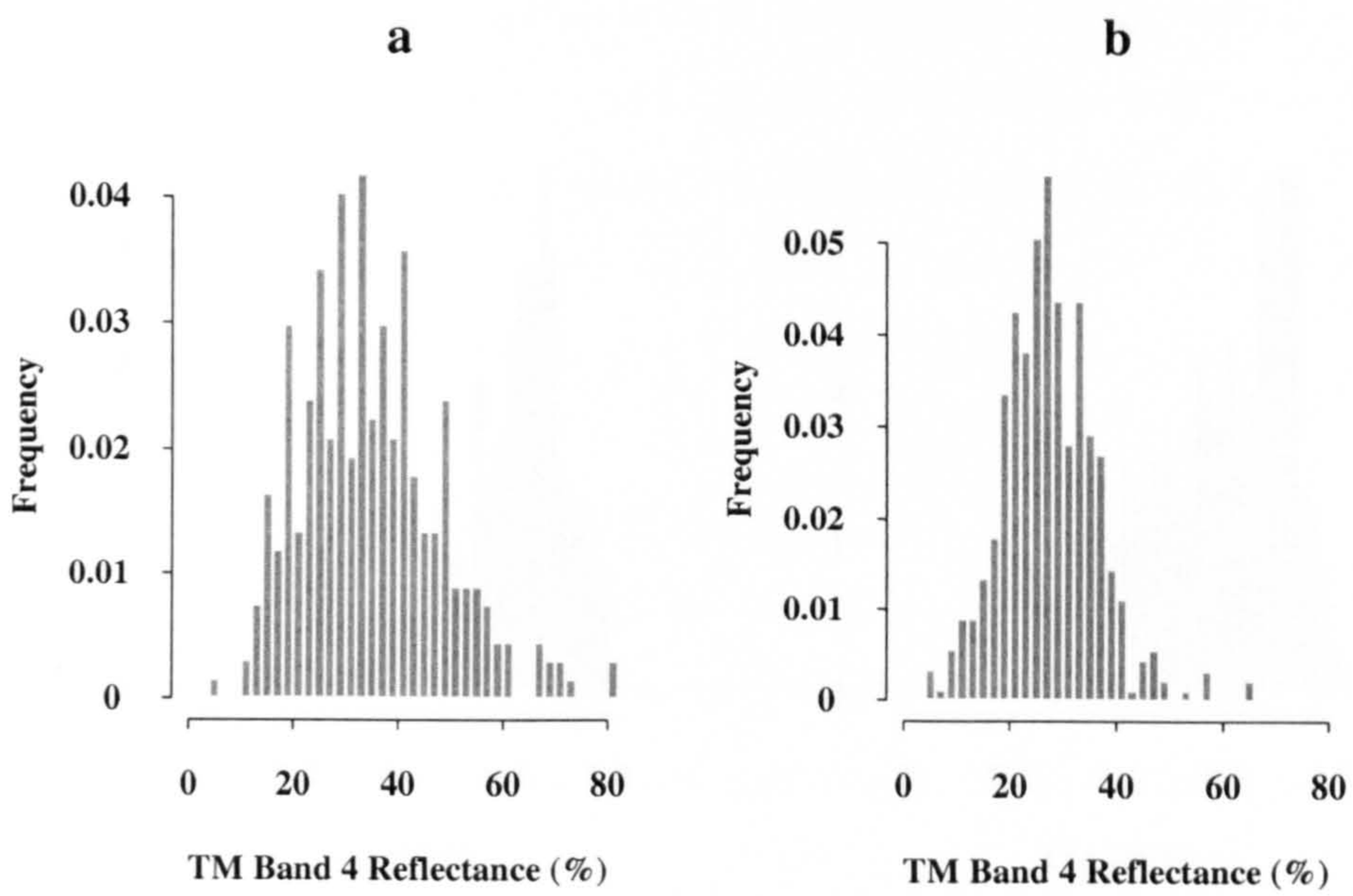
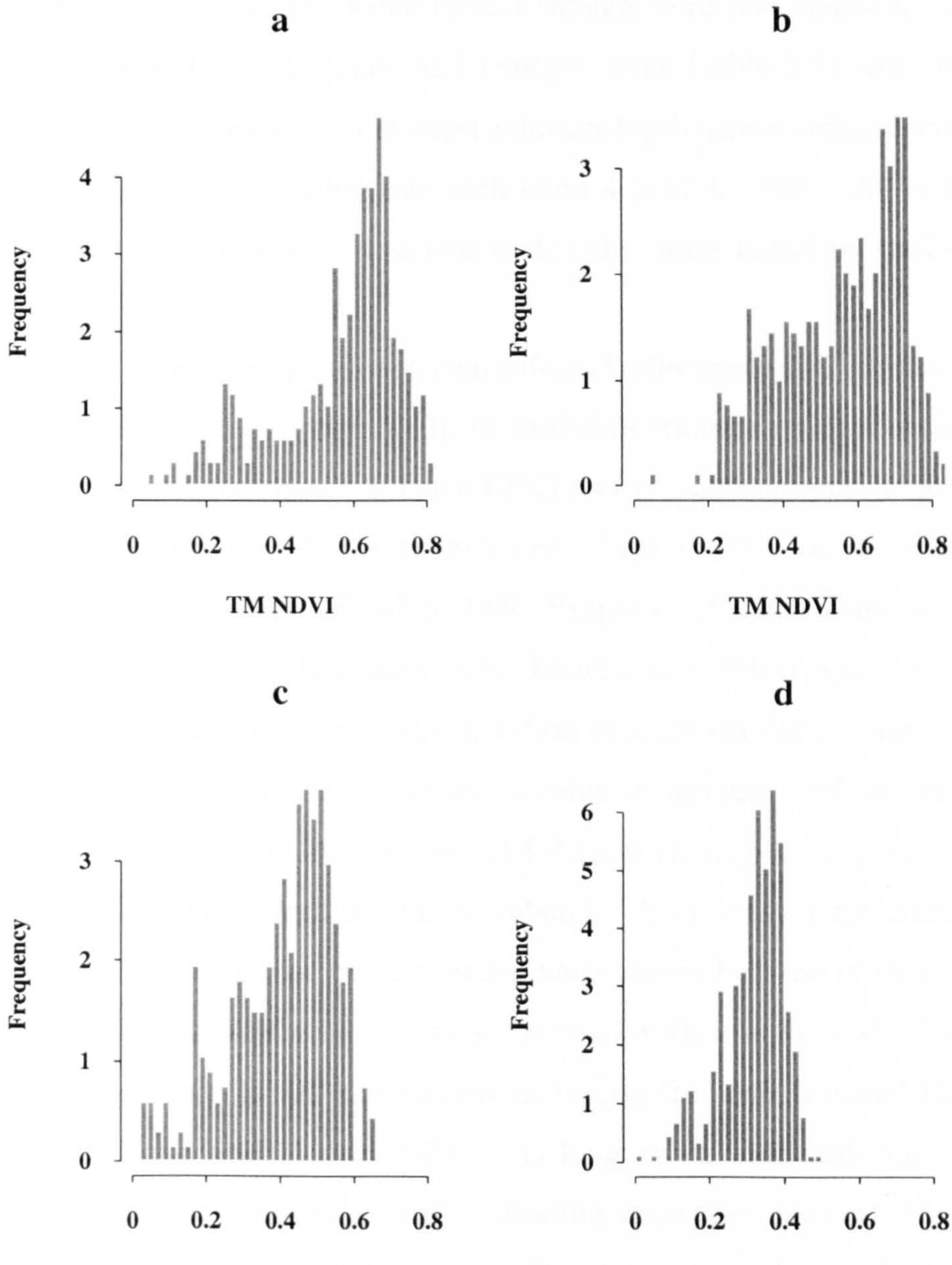


Figure 5.12: Histogram of (a) TM NDVI at 1988 ECODATA plot locations, (b) TM NDVI at 1993 locations, (c) AVHRR NDVI at 1988 locations and (d) AVHRR NDVI at 1993 locations.



Another aspect of the change-of-support model is decreasing variance as support increases (Matheron 1985, Zhang *et al.* 1990). Reflectance in visible, near infrared and NDVI all showed smaller variance at AVHRR support than at TM support, though the variance reduction in the 1988 AVHRR NDVI was slight.

The univariate distributions of spectral variables showed consistency between 1988 and 1993. The bivariate distributions of spectral data with the vegetation amount variable followed the same trends though were not identical. Linear correlations between TM reflectance and canopy cover (table 5.5) were all negative and statistically significant – an indirect relationship between reflectance and cover. The strongest linear correlation was with band 4 (-.67 in 1988; -.58 in 1993). The magnitude of these coefficients was well within the range found by studies reported in table 2.1.

A positive relationship between near infrared reflectance and vegetation amount exists at the leaf level (Curran 1980), in radiative transfer models considering homogeneous canopies (see figure 2, Price 1992) and geometrical-optical forest canopy models (Li and Strahler 1992, Gemmell and Varjo 1999). Yet several studies of conifers at the canopy scale (Franklin 1986, Ripple *et al.* 1991, Danson and Curran 1993, Nemani *et al.* 1993, Gemmell 1995, Jakubauskas 1996) report negative relationships between near infrared and vegetation amount similar to that found in this study. A negative relationship between vegetation amount and red reflectance is consistent with theory and experiments at field and canopy levels primarily because of absorption by chlorophyll in this waveband. A negative relationship between vegetation amount and near infrared reflectance occurs because of shadow effects – as biomass, height or leaf area increases the canopy throws more shadows, thereby decreasing reflectance in all wavelengths including the near infrared (Danson and Curran 1993, Boyd and Curran 1998). As long as infrared reflectance decreases more slowly than red reflectance with increasing vegetation amount, the NDVI will increase as vegetation amount increases. This is a reasonable explanation for the positive correlation between NDVI and canopy cover seen here.

Scatterplots between canopy cover and the spectral variables further revealed the trends summarized by the correlation coefficients (figure 5.13 and 5.14). For each canopy cover value there was high variance in reflectance and NDVI values. The

Table 5.5: Linear correlation coefficients between TM spectral variables and canopy cover at ECODATA plot locations.

Scene Date	ρ_{Band1}	ρ_{Band2}	ρ_{Band3}	ρ_{Band4}	ρ_{Band5}	ρ_{Band7}	NDVI
1988	-0.58	-0.63	-0.63	-0.48	-0.67	-0.63	0.53
1993	-0.32	-0.38	-0.43	-0.35	-0.58	-0.56	0.46

data indicated heteroscedasticity; that is, the variance of spectral values changed with canopy cover. This is consistent with high absorption by closed canopies and the diverse absorptances by the variety of soil backgrounds and understories in open canopies (Jasinski 1990). The infrared reflectance relationships (figure 5.13*b* and *c*) appeared slightly more homoscedastic than did the red (figure 5.13*a*).

Despite the low precision of the canopy cover data, the scatterplots in figures 5.13 and 5.14 suggested nonlinear relationships between canopy cover and the spectral variables. Standard boxplots (Hoaglin *et al.* 1983) summarizing the scatterplots make the nonlinearity more obvious (figure 5.15). The light bar inside each box indicates the median of each distribution. The medians showed the asymptotic trend usually associated with spectral indices and vegetation amount. Since the number of plots with canopy cover above 50% was quite small, fewer than 10% of the total number of measurements in each year, the characteristics of the relationship at the high end of canopy cover were less certain than at the low end.

The trends seen in the TM data were completely lost at the AVHRR support (figure 5.16) for visible (r=.05 in 1988 and r=-.25 in 1993), near infrared (r=.06 in 1988 and r=-.17 in 1993) and NDVI (r=-.02 in 1988 and r=.12 in 1993). Variance within each canopy cover class swamped variance among classes. Support of the ground data was not changed here, just the support of the reflectance variable. The sampling fraction of the ground data was not sufficient to average the plot values to predict canopy cover values at the 1.21 km^2 support. In general, support effects on bivariate distributions are less predictable than those on the univariate distributions (Cressie 1996).

The choice of the variable(s) to use in regression is commonly made based on which combination of variables yields the highest correlation. Therefore, in this case prediction would be accomplished with reflectance from TM band 5 as the

Figure 5.13: Relationship between canopy cover at the ECODATA plot locations and TM (a) reflectance in band 3 from 1988, (b) reflectance in band 4 from 1988, (c) reflectance in band 5 from 1988 and d) NDVI from 1988.

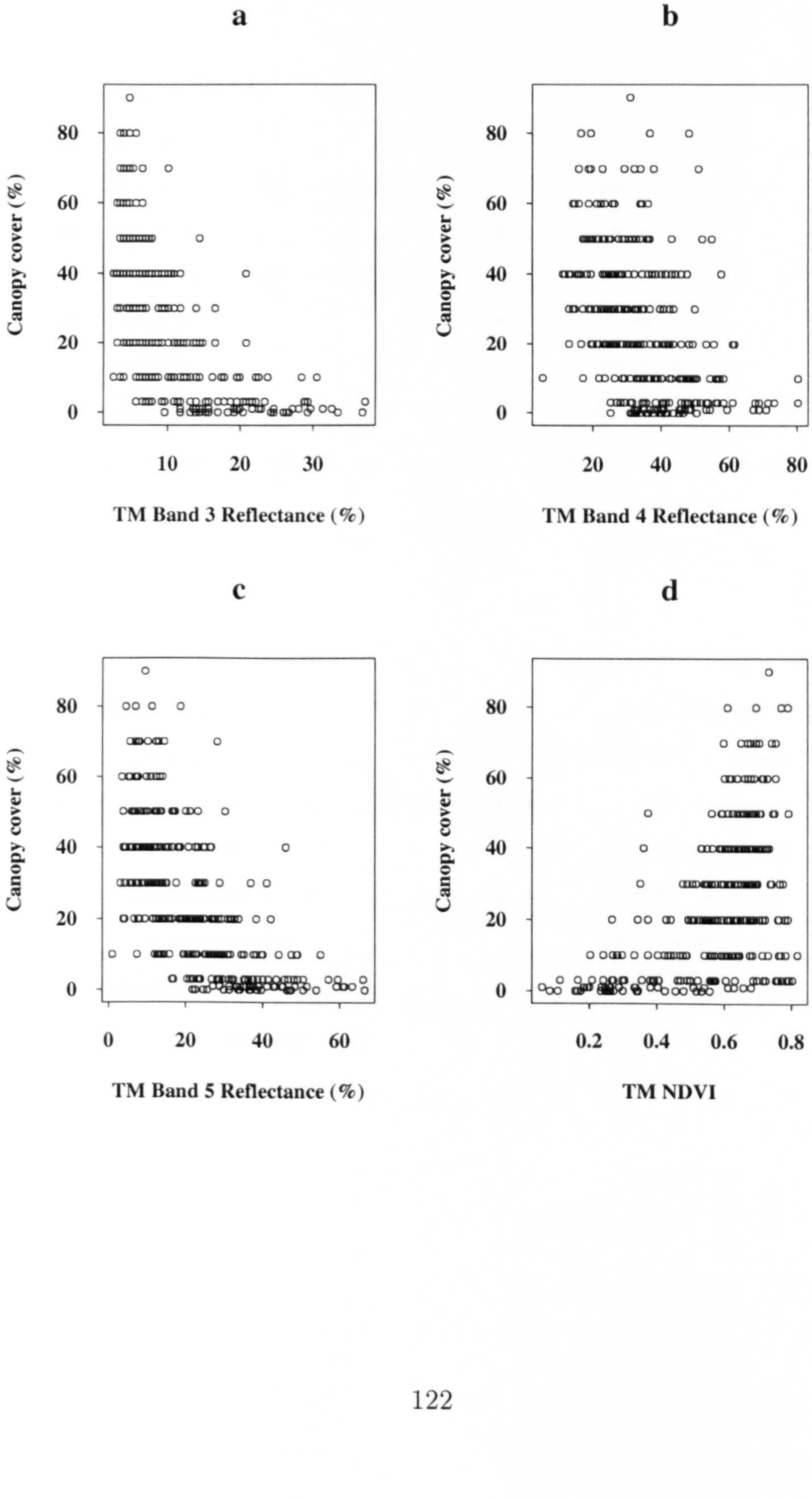
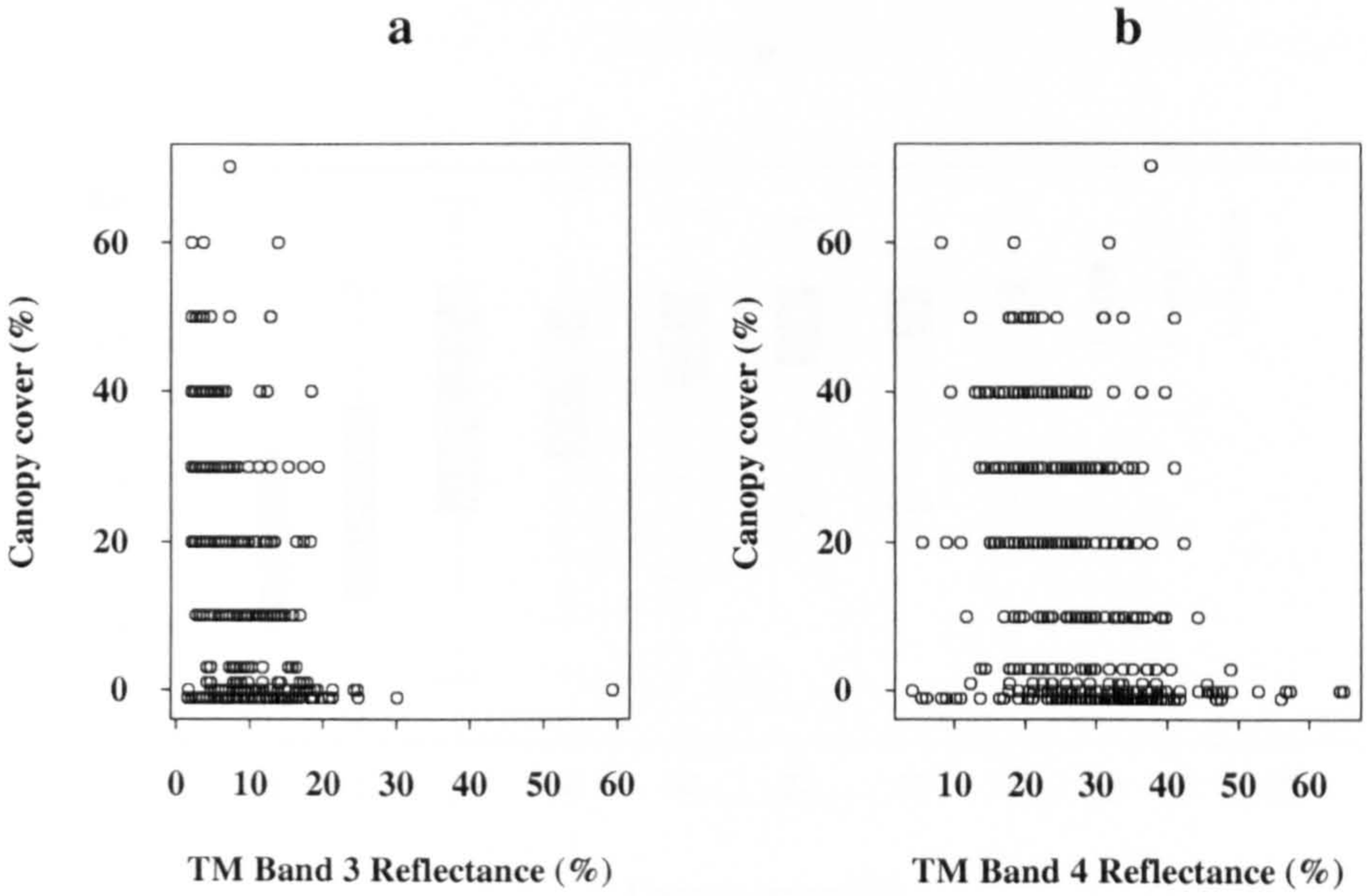


Figure 5.14: Relationship between canopy cover at the ECODATA plot locations and (a) reflectance in TM band 3 from 1993, (b) reflectance in TM band 4 from 1993, (c) band 5 from 1993 and (d) NDVI from 1993.



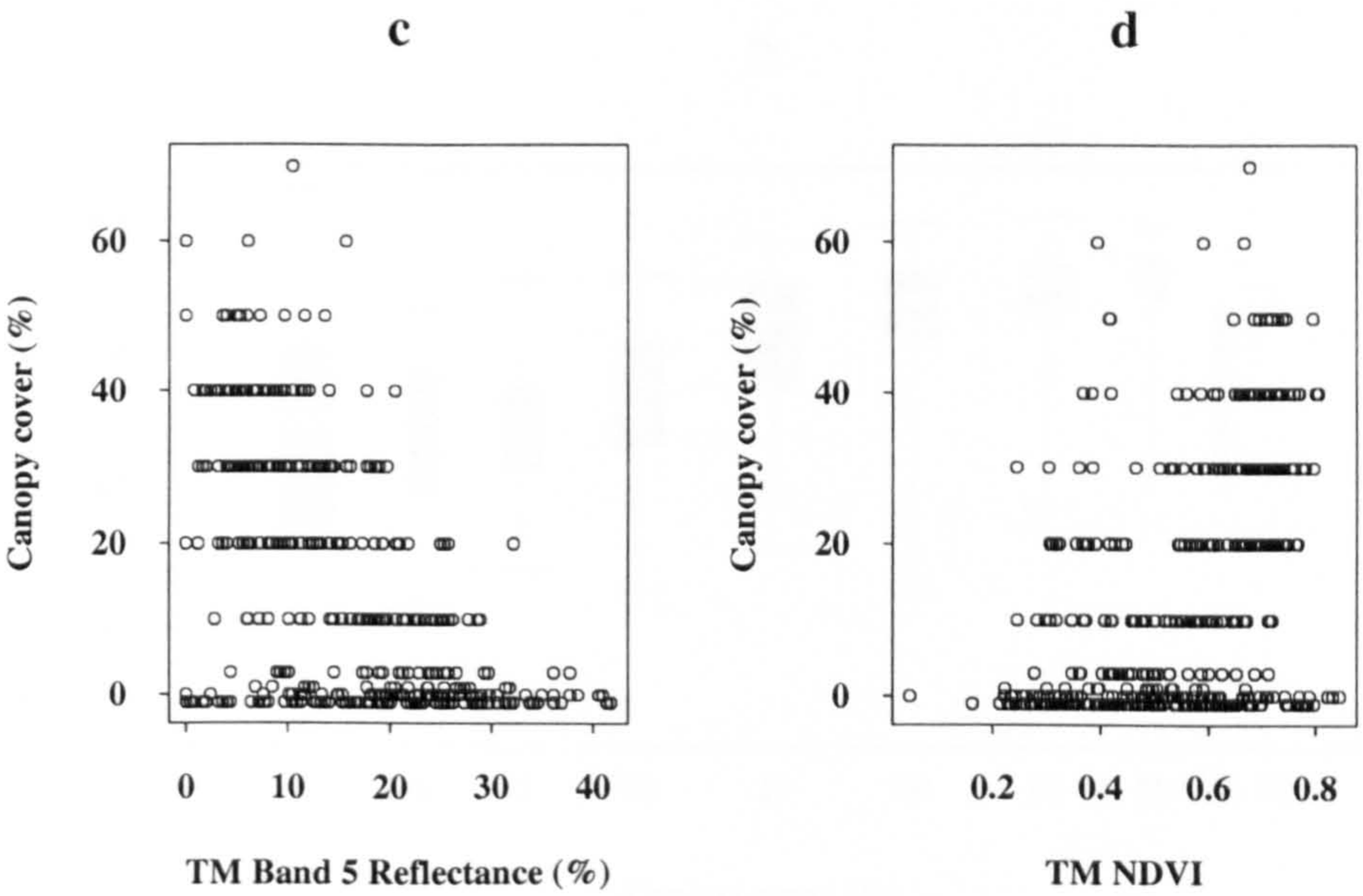


Figure 5.15: TM NDVI versus canopy cover from (a) 1988 and (b) 1993, where distribution of NDVI values for each value of canopy cover is summarized using a boxplot. The width of each box is proportional to the number of plots with the value of canopy cover on the abscissa.

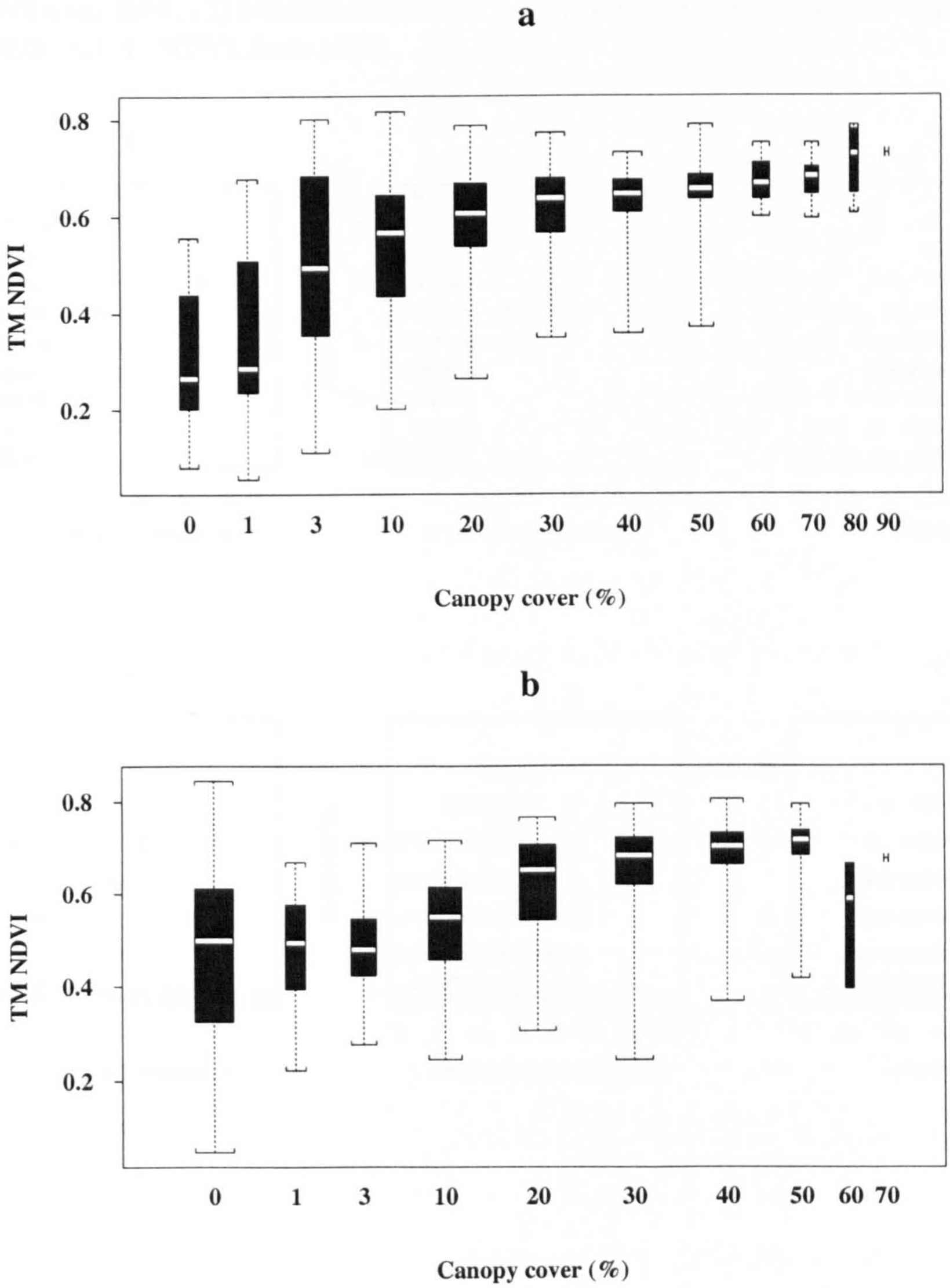


Figure 5.16: Relationship between canopy cover at the ECODATA plot locations and AVHRR (a) reflectance in band 1 from 1988, (b) reflectance in band 2 AVHRR from 1988, (c) NDVI from 1988, (d) reflectance in band 1 from 1993, (e) reflectance in band 2 AVHRR from 1993 and (f) NDVI from 1993.

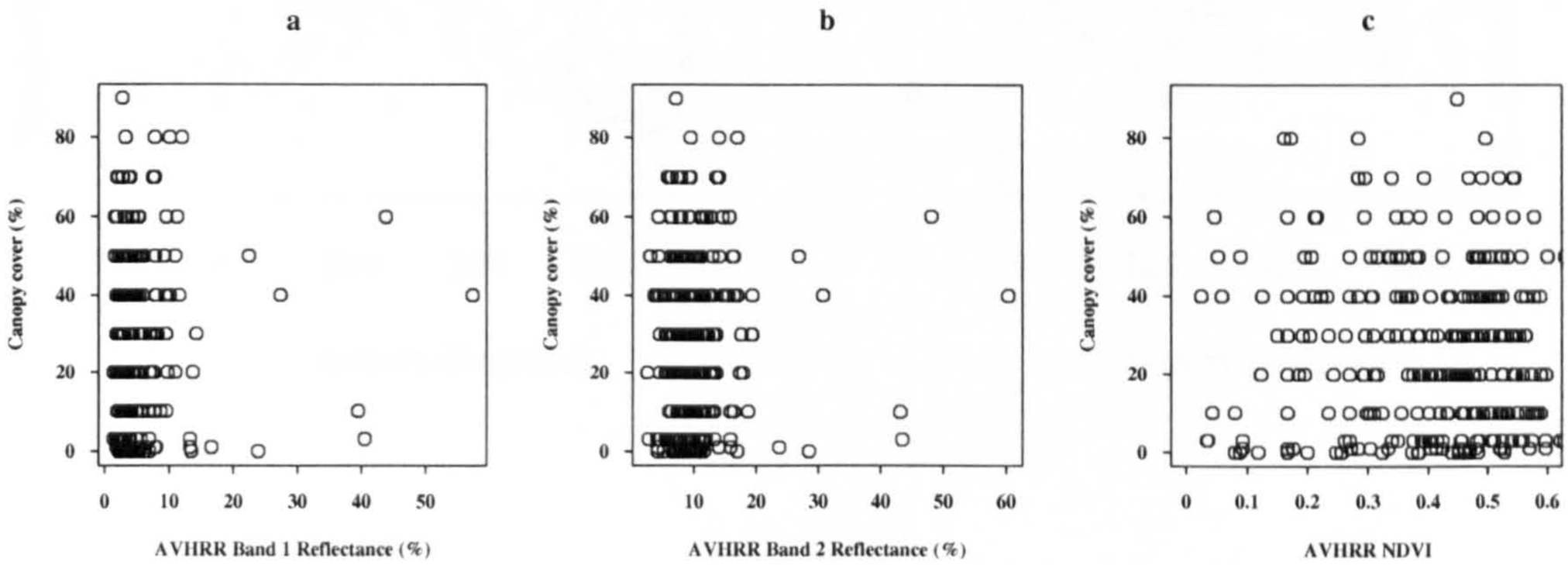
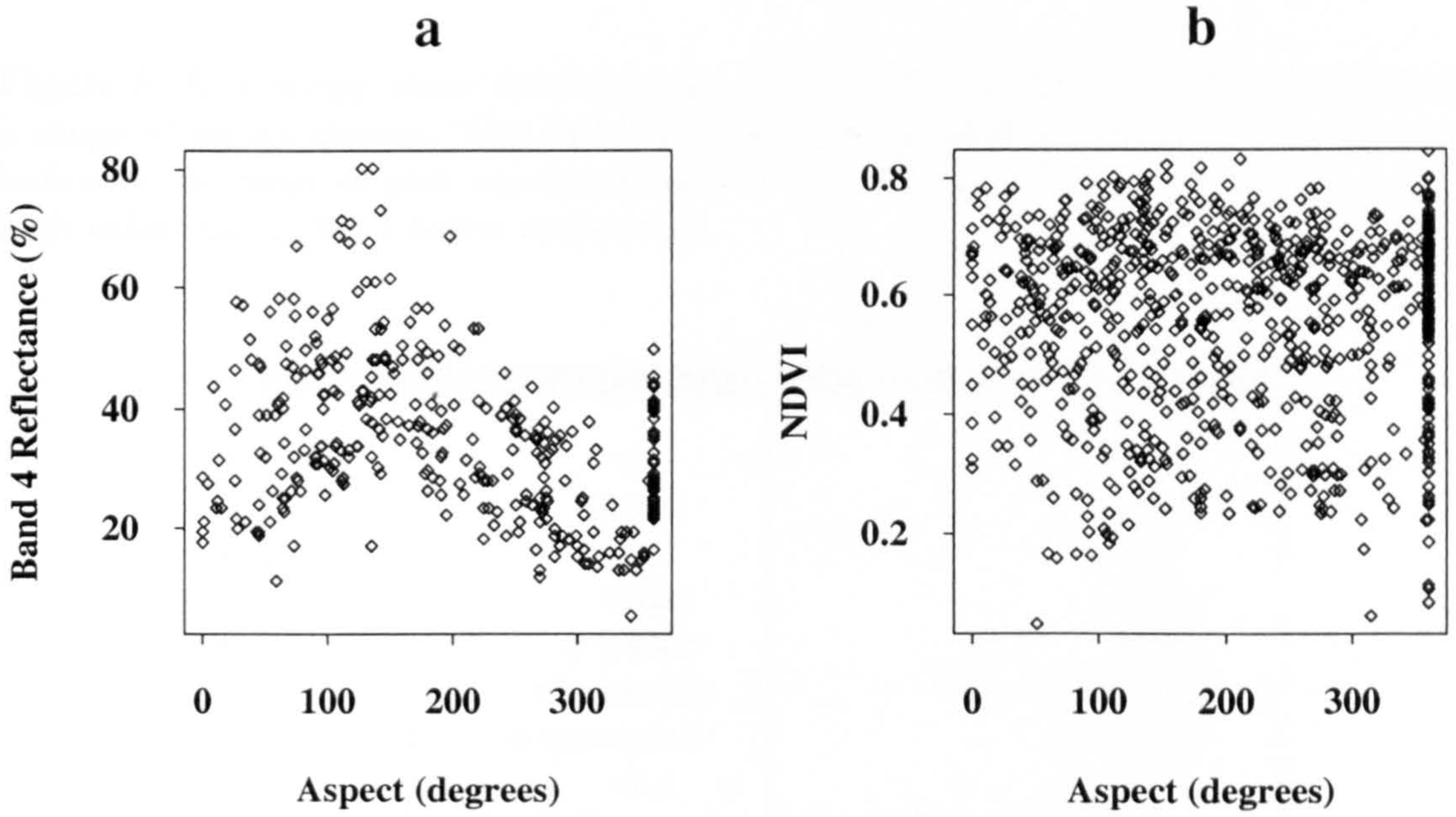


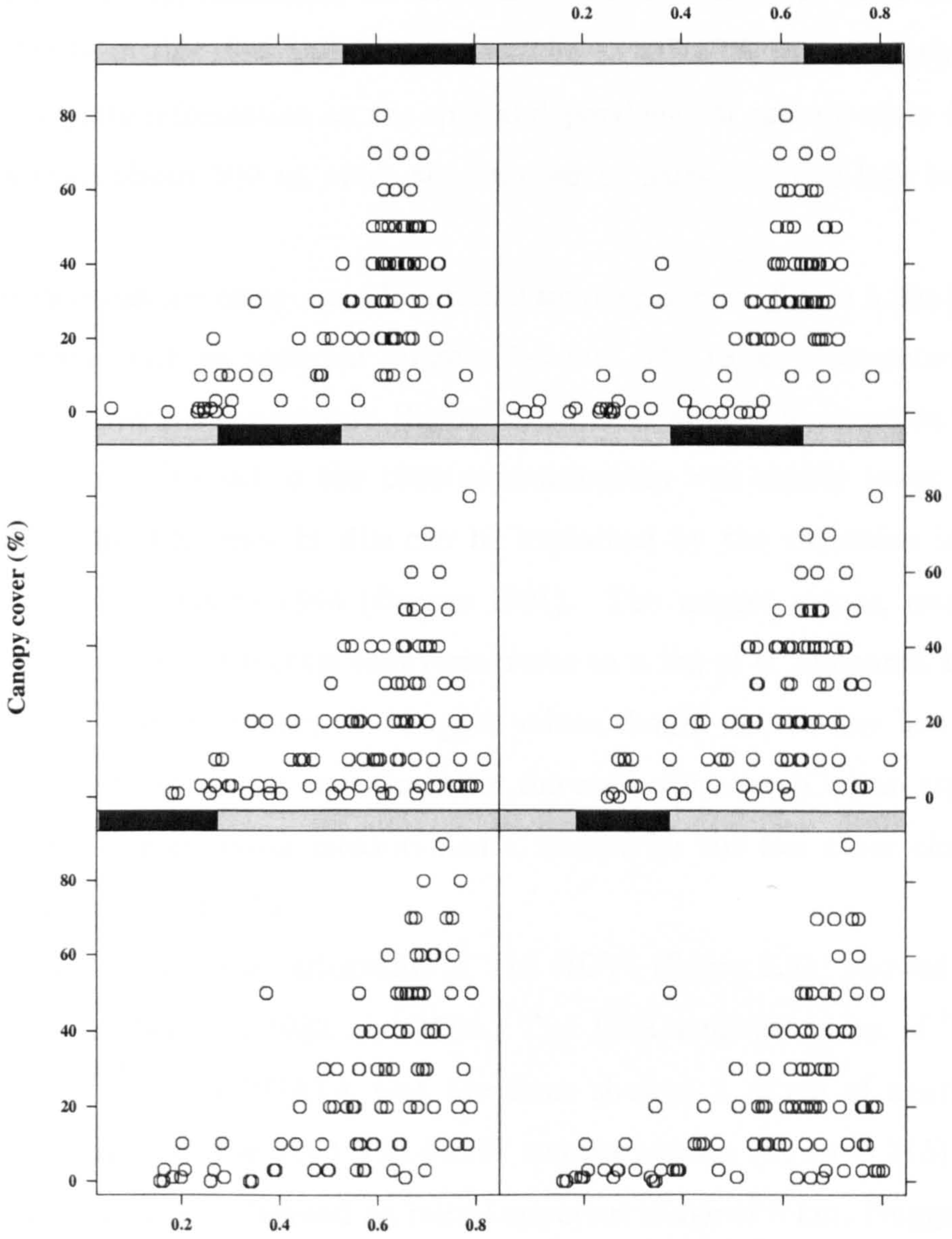
Figure 5.17: (a) Reflectance in TM band 4 versus aspect and b) TM NDVI versus aspect.



ancillary variable. But the lack of a terrain correction (Conese *et al.* 1993) in the image preprocessing steps meant the TM reflectance images contain illumination angle effects (figure 5.17*a*). In theory, spectral indices help to compensate for these illumination angle effects (Price 1987). The data indicated that illumination angle effects were weak or nonexistent in the NDVI transform (figure 5.17*b*). Scatterplots of canopy cover versus NDVI show that illumination did not seem to affect the relationship with canopy cover (figure 5.18). Therefore, NDVI was chosen as the ancillary spectral variable for all prediction methods.

One of the differences between the 1988 and 1993 TM scenes was the resampling algorithm used during the geometric correction process. Nearest neighbor resampling does not introduce new spectral vectors in the the image statistical distribution (Schowengerdt 1997), whereas cubic convolution does. Cubic convolution does add error to the image values and further acts as a smoother, affecting the image variance and semivariogram (Appendix A). The decrease in the variance of the (cubicly convoluted) 1993 TM data relative to that of the (nearest-neighbor

Figure 5.18: Canopy cover measurements versus NDVI at ECODATA plot locations in a range of aspect classes. The dark bar within the shaded panel above each scatterplot indicates the range of plot aspects, from low values (near 0°) in the lower left corner to high values (near 360°) in the upper right.



resampled) 1988 TM data may be partly caused by this effect.

5.2.2 Spatial statistics of primary and ancillary data

The 1988 and 1994 ECODATA measurements with their coincident image pixels provided enough data to examine and model spatial dependence statistics. A histogram of pair separation distance (figure 5.19) showed that with a lag width of 300 m (equivalent to approximately 10 TM pixels) there were more than 100 pairs at each lag for both the 1988 and 1994 data. The ECODATA measurements did not provide adequate information on the spatial dependence of canopy cover for *shorter* distances than about 300 m, since the number of pairs at some lags became too small.

The 1988 canopy cover omnidirectional semivariogram (figure 5.20*a*) showed a clear structure, with an apparent range of 5-7 km. The range interpreted from the 1994 canopy cover semivariogram (figure 5.20*b*) was about the same as that from 1988 or shorter. The sill in the 1994 semivariogram was clearly lower than that from 1988. The difference in sills can be explained by the reduction in variance (figure 5.4) from 1988 to 1994 (Barnes 1991). The nugget values, estimated by extrapolating the experimental semivariograms to a lag of 0, appeared to be both approximately 100. In theory, the nugget values should not be any less than 100, since the nugget variance is an estimate of the noise variance (Curran and Dungan 1989) and the canopy cover measurements, except for the low cover class, have a maximum precision of 10%.

Omnidirectional semivariograms of TM NDVI (figure 5.21) showed quite different structure between 1988 and 1994. The 1988 semivariogram of TM NDVI (figure 5.21*a*) from ECODATA plot locations showed a range of nearly 10 km, whereas the range in the 1993 TM NDVI semivariogram (figure 5.21*b*) showed a decrease in semivariance beyond an initial apparent range of 3 km. Nugget variance estimates consistent with both semivariograms ranged from about .002 (indicating an NDVI precision of .04, to about .01 (indicating an NDVI precision of .1 unit). Again, the sill from the later date (1993) was lower than that from 1988, though the 1993 semivariogram seemed to be linearly increasing beyond 10 km. The lower

Figure 5.19: Histogram of distances between plot locations from (a) 1988 and (b) 1994.

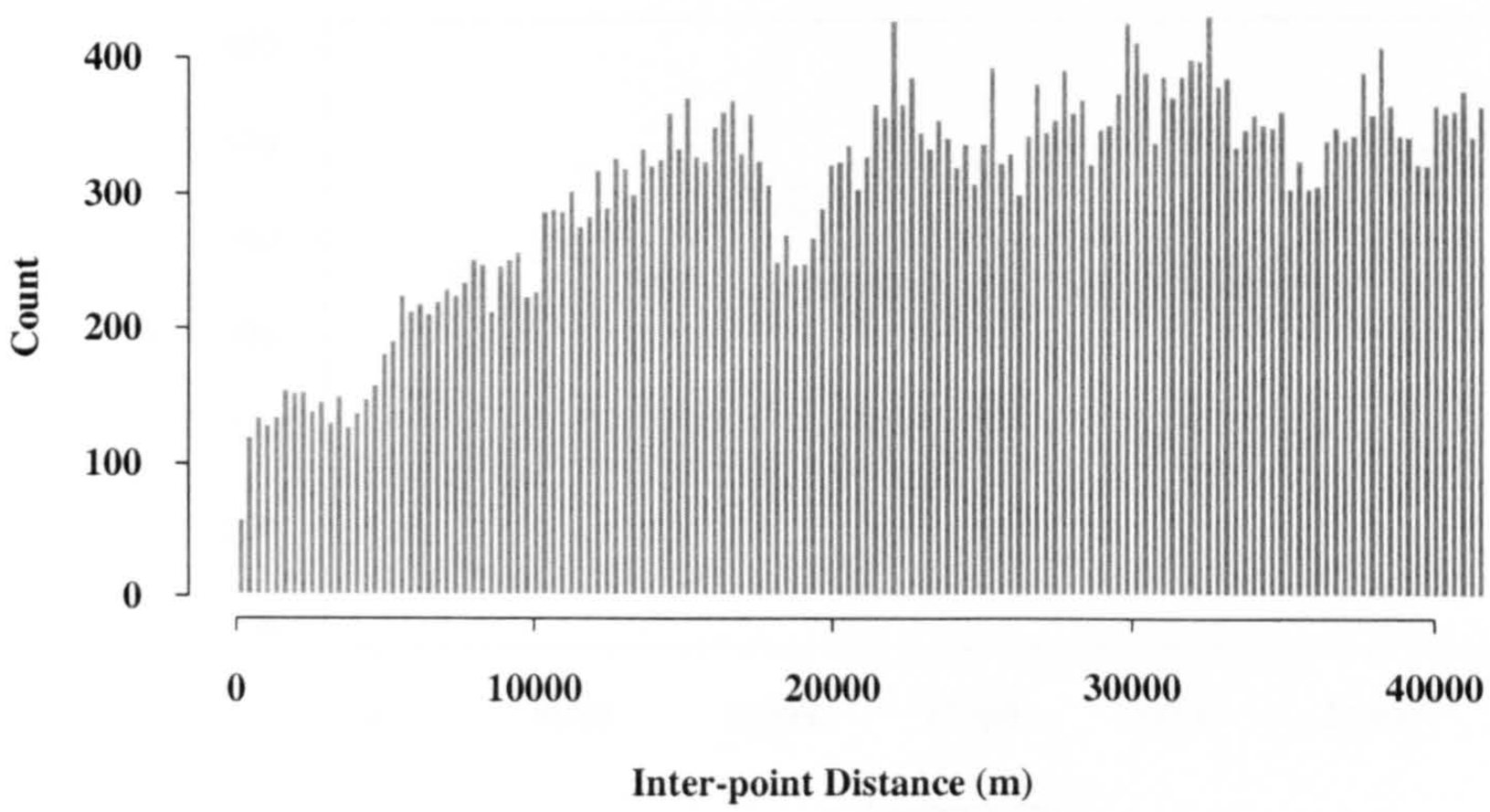
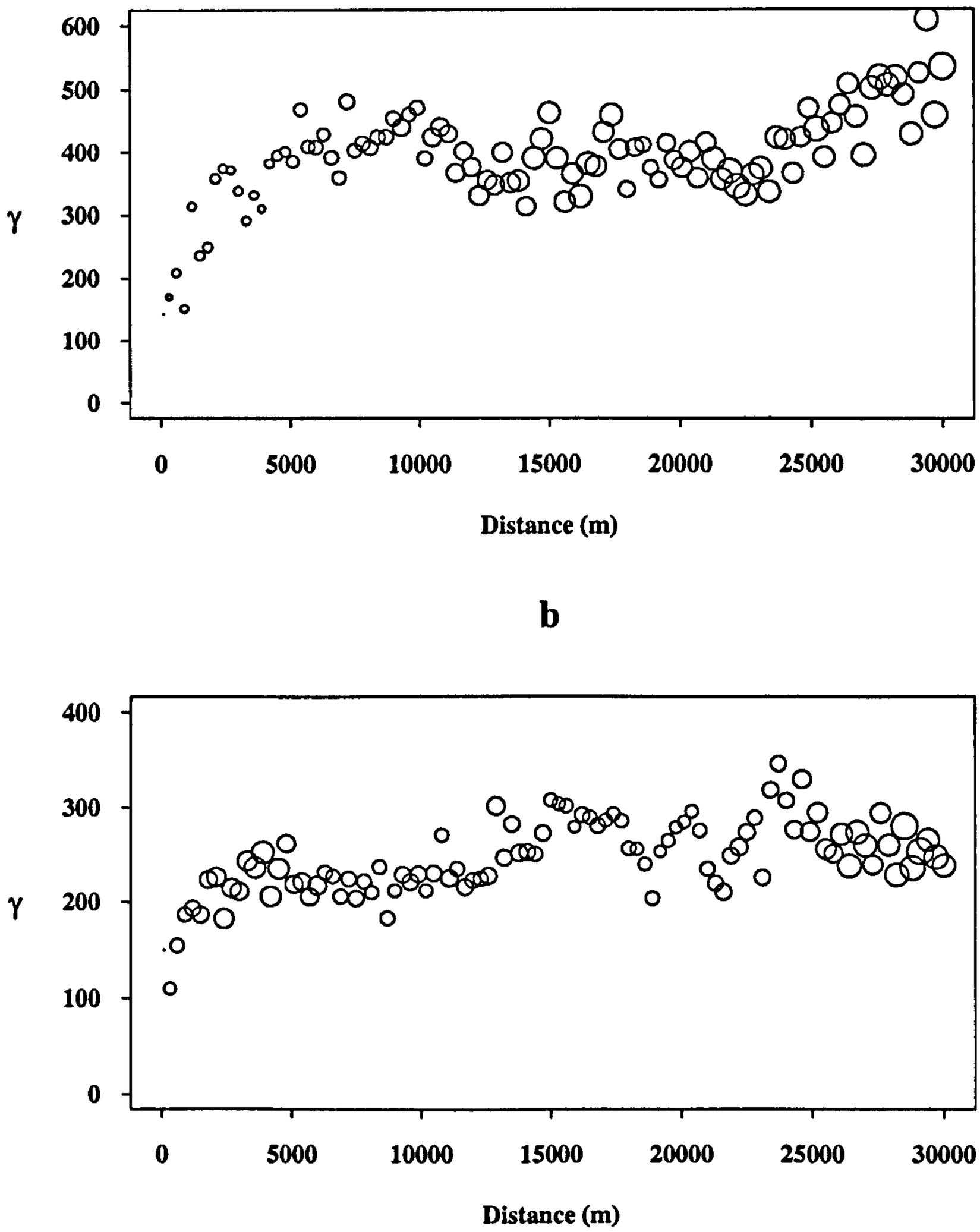


Figure 5.20: Omnidirectional semivariograms of canopy cover (%) from (a) 1988 and (b) 1994. Symbol size is proportional to the number of pairs at that distance.



semivariance values in the first several lags in the 1993 data could have been due in part to the cubic convolution resampling (Appendix A). Dungan (2000) showed that the semivariograms for the AVHRR NDVI had, as expected, much lower sills. They also had reduced ranges.

Fitting an LMC to the 1988 semivariograms and cross-semivariograms was straightforward, as simple single-structure spherical models coincided with the data. The LMC developed for the 1988 data was:

$$\gamma_{p1988}(\mathbf{h}) = 120 + 300 Sphr_{5500}(\mathbf{h}) \quad (5.12)$$

$$\gamma_{a1988}(\mathbf{h}) = 0.009 + .017 Sphr_{5500}(\mathbf{h}) \quad (5.13)$$

$$\gamma_{pa1988}(\mathbf{h}) = 0.4 + 2 Sphr_{5500}(\mathbf{h}) \quad (5.14)$$

The fit of this model to the experimental semivariograms (figure 5.22*a*, *c* and *e*) appeared reasonable.

The 1993 patterns presented a much less convincing fit to a strictly positive definite LMC (figure 5.22*b*, *d* and *f*). The model, similar to that for 1988 but with reduced sills and nuggets to represent the smaller variance in the primary and ancillary data, was:

$$\gamma_{p1993}(\mathbf{h}) = 100 + 150 Sphr_{5500}(\mathbf{h}) \quad (5.15)$$

$$\gamma_{a1993}(\mathbf{h}) = 0.003 + .017 Sphr_{5500}(\mathbf{h}) \quad (5.16)$$

$$\gamma_{pa1993}(\mathbf{h}) = 0.4 + Sphr_{5500}(\mathbf{h}) \quad (5.17)$$

The range from the 1993 NDVI data appeared substantially shorter than the 5500 m range of the canopy cover semivariogram. The lack of a sill in the NDVI semivariogram (figure 5.22*d*) means that no transitive model will fit well.

The reason for the range observed in these semivariograms, around 5000 m, is likely to be related to the terrain. Five kilometers is a typical distance from ridge to valley bottom, and from south-facing to north-facing slope in the study region (figure 5.5). Terrain, with its concomitant variation in insolation, soil and water availability, is one of the controlling factors of canopy cover in western Montana.

Figure 5.21: Omnidirectional semivariograms of TM NDVI from (a) 1988 and (b) 1993. Symbol size is proportional to the number of pairs at that distance.

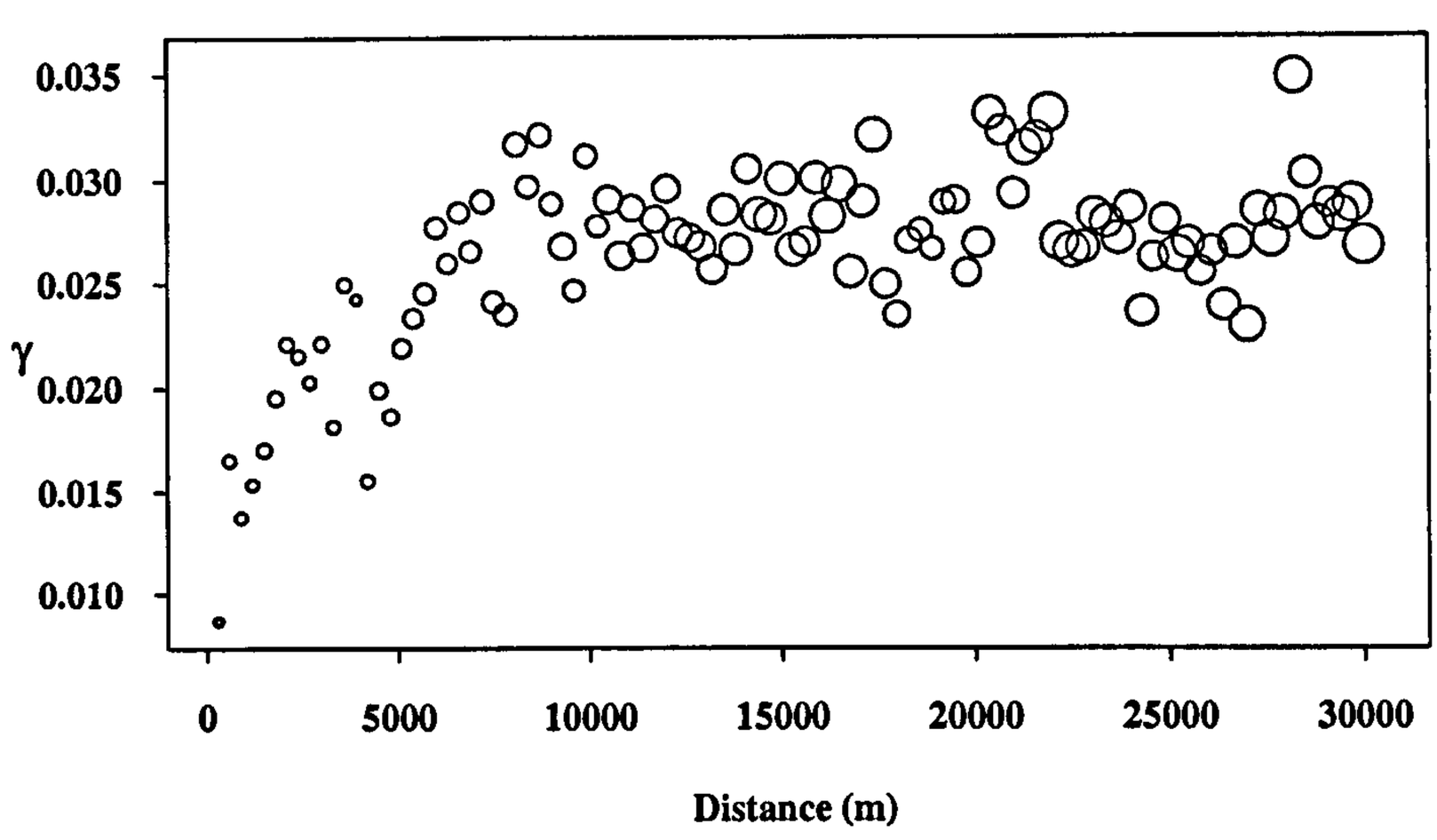


Figure 5.22: Omnidirectional semivariograms and models for (a) 1988 canopy cover (%), (b) 1994 canopy cover (%) and (c) 1988 NDVI and d) 1993 NDVI. Omnidirectional cross-semivariograms and models for (e) 1988 canopy cover and NDVI and (f) 1993 canopy cover and NDVI.

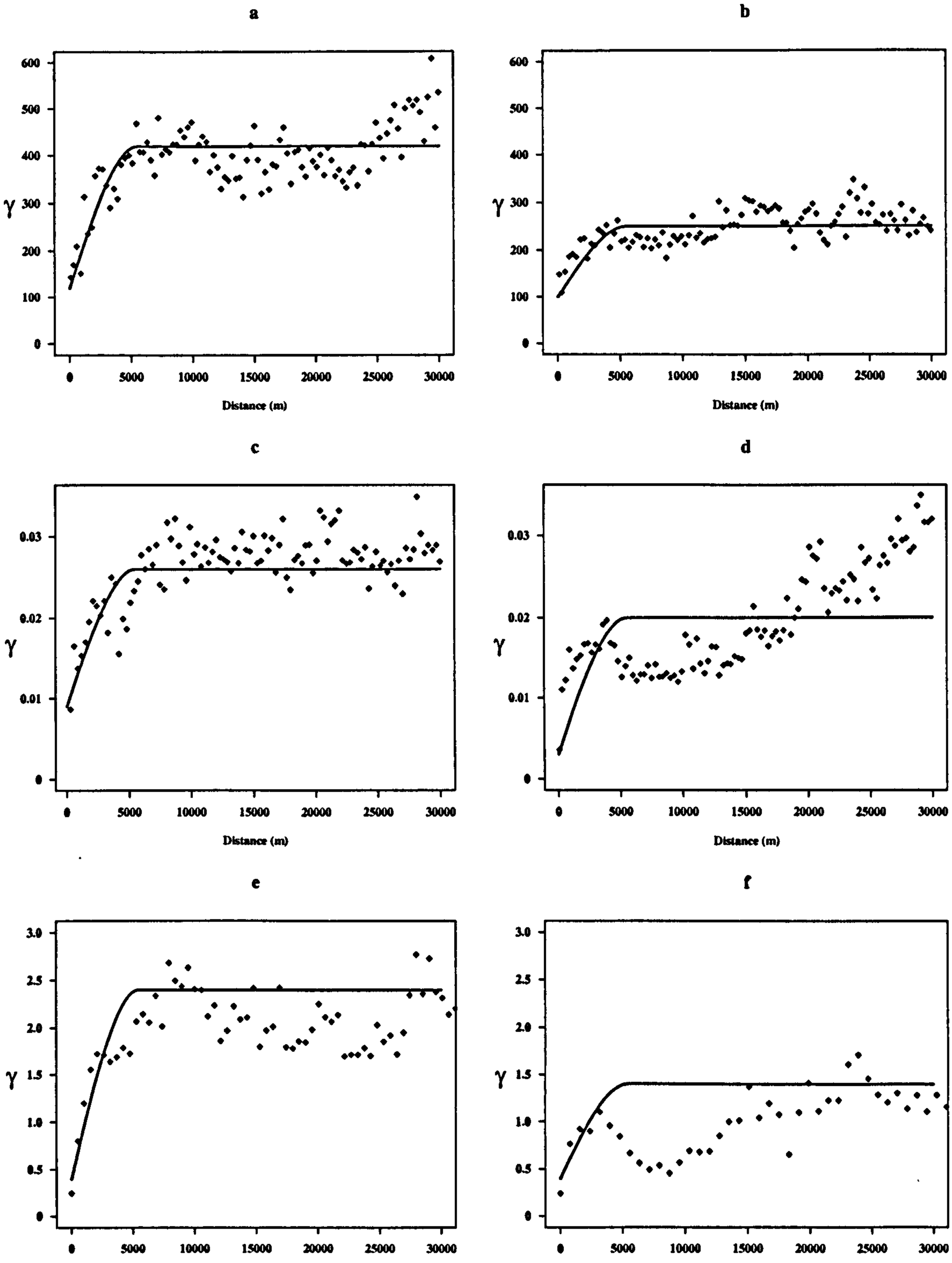


Table 5.6: Parameter estimates for aspatial regression model 5.9.

Date	B_1	B_2
1988	.72	.13
1993	.73	.12

5.2.3 Prediction of canopy cover

The results of applying the three prediction methods in the subarea were maps of canopy cover on a 348×310 cell grid, one for 1988 and one for 1993. It is unlikely that substantial changes in actual canopy cover occurred in the subarea between 1988 and 1994 because of the lack of major disturbance. Any changes that did occur would likely be within the range of precision of the ECODATA measurements. Given this assumption, it is possible to check the accuracy of predicted maps by comparing the 1994 ECODATA values with the values predicted from the 1988 model and, conversely, comparing the 1988 ECODATA values with the values predicted from the 1993 model. Further, the general appearance of the maps for the two dates should be similar, with regions of low values in 1988 remaining low in the 1993 maps and regions of high values in 1988 remaining high in 1993.

Because of the severe support differences between the ECODATA measurements and the AVHRR data (Atkinson 1997*b*), AVHRR NDVI was not used to make predictions.

Aspatial regression

The best-fit parameters for 1988 and 1993 (table 5.6) were very similar and statistically identical. The fits are shown in figure 5.23. If B_1 is interpreted as $NDVI_\infty$, it appears to be a lower value than expected. 13% of the 1988 TM image from the subarea and 34% of the 1993 image pixels had NDVI values larger than .73.

Equation 5.10 was used to yield predictive equations. When extrapolated to the subarea, the predicted maps (figure 5.24) showed the TM images' patterns, with all of their detail. This is expected, since the regression model is simply a 'rescaling' of the measurement units of the remotely sensed data. The mean value of the predicted maps was 52% and 68% in 1988 and 1993 respectively (table 5.7),

Figure 5.23: Nonlinear regression models from (a) 1988 and (b) 1993 data.

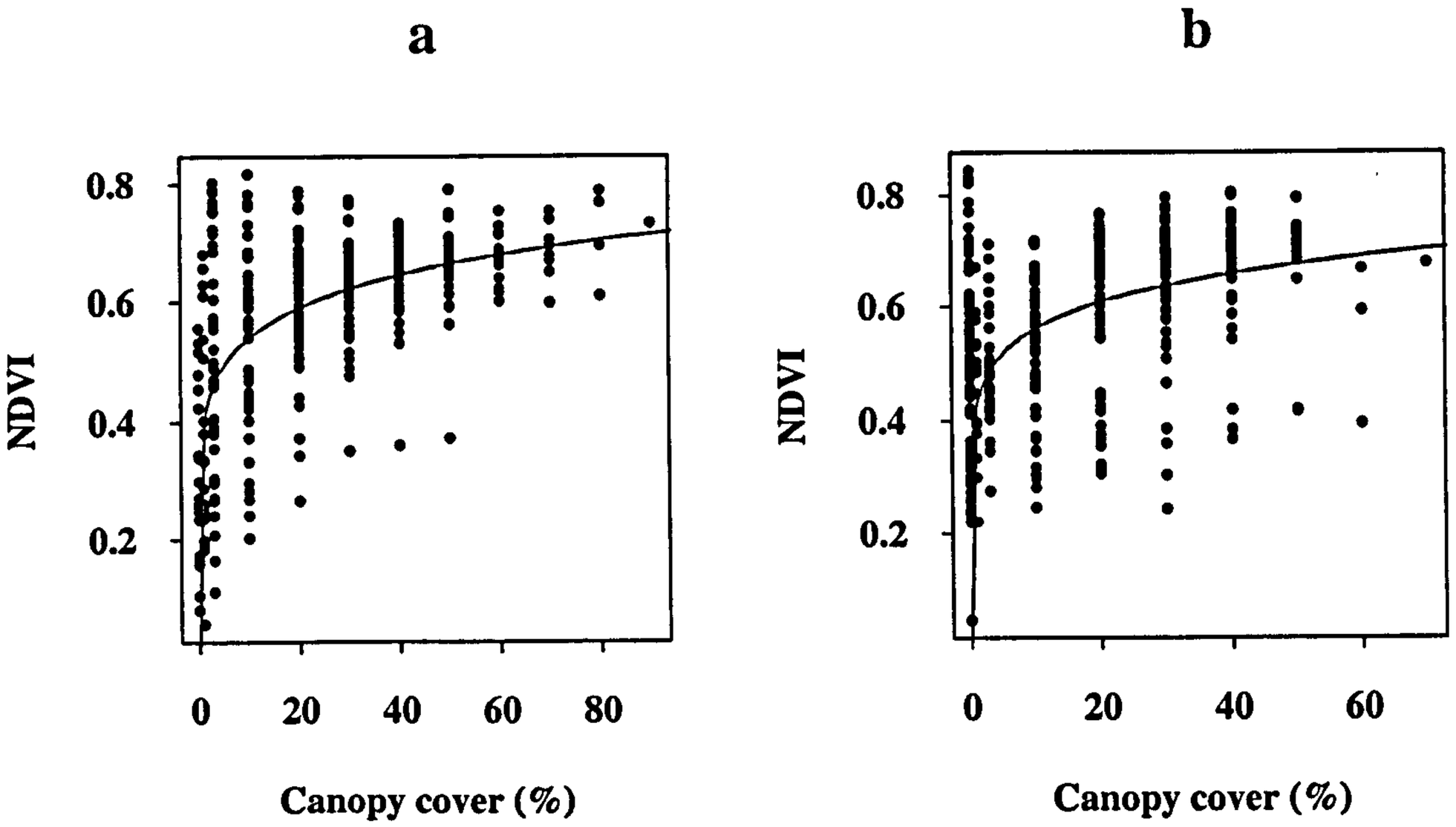
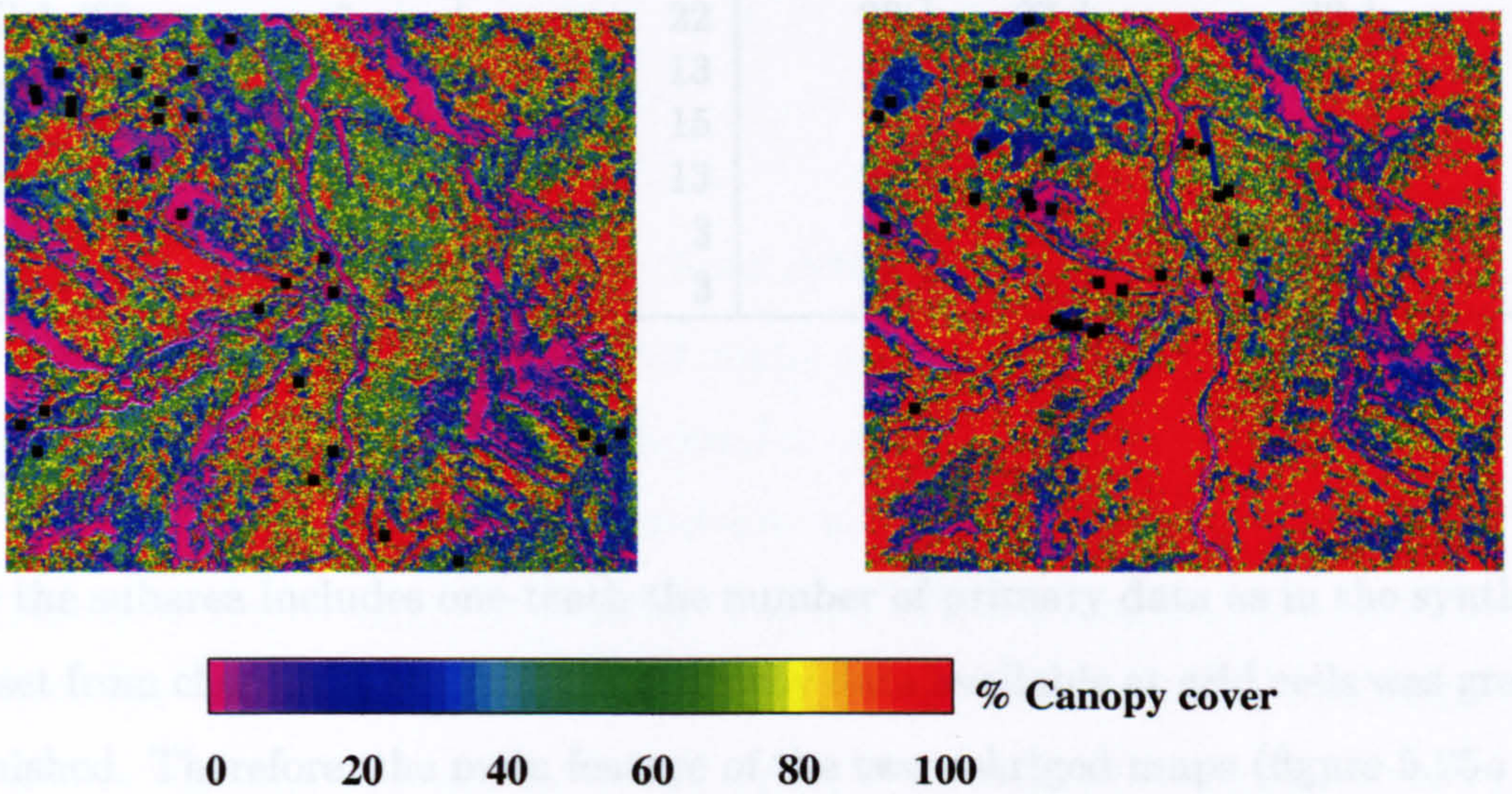


Figure 5.24: Predicted canopy cover (%) using regression from (a) 1988 data and (b) from 1993 data. ECODATA plot locations are shown as black squares.



much higher than the observed ECODATA mean. In fact, the entire distribution of predicted values is shifted upwards when compared to the measurements. The 1988 and 1993 maps looked somewhat similar, but values were generally higher in 1993 than in 1988. Point locations were checked by comparing 1988 predictions at 1993 measurements locations. This check yielded root-mean-square errors of 46 and 51 (in % units) respectively from the regression prediction.

The prediction intervals from the models were wider than the physical range of canopy cover and were therefore not mapped in figure 5.24. As was obvious from the scatterplots between the ECODATA measurements and NDVI (figures 5.13*d* and 5.14*d*), any aspatial regression model fit through these data will imply large uncertainty about predictions.

Table 5.7: Global statistics of canopy cover (%) maps predicted by aspatial regression (Reg), cokriging (Cok) and simulation (Sim) alongside the statistics from the 335 1988 ECODATA plot measurements and the 344 1994 plot measurements.

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
Reg '88	0	24	52	52	78	100
Reg '93	0	43	77	68	100	100
Cok '88	0	22	28	27	32	49
Cok '93	0	13	17	17	22	39
Sim '88	0	15	25	26	35	70
Sim '93	0	13	21	22	30	73
ECODATA '88	0	3	20	25	40	90
ECODATA '94	0	3	20	18	30	70

Cokriging

Since the subarea includes one-tenth the number of primary data as in the synthetic data set from chapter 4, the number of nearby data available at grid cells was greatly diminished. Therefore, the main feature of the two cokriged maps (figure 5.25*a* and *b*) was their spatial homogeneity. Because of the sparsity of the ground data and the long autocorrelation range (figure 5.20), values were similar over large areas. The influence of the ancillary data was evident, with the streams and bare-rock features predicted to have low values and some of the clearcut boundaries showing up.

The cokriged maps were also similar between 1988 and 1993, though the high values in the southern center of the 1988 image were not present in the 1993 image. The absence of ground measurements in the Bob Marshall Wilderness on the second date meant that these regions of higher canopy cover were missed. The RMSE of cokriging predictions from the 1988 model made at the 1993 plot locations was 12 (in % units) and that from the 1993 model made at the 1988 ECODATA locations was 18 (also in % units), a substantially smaller mean error than that from the aspatial regression maps. The global means of the cokriged maps were very similar to those of the ECODATA measurements (table 5.7) and the distribution was shifted towards lower values than maps predicted with aspatial regression. The spread was also smaller, with reduced maxima and third quartile, and increased first quartile. The color scale in figure 5.25 ranges from 0-60% (instead of 0–100% as in figure 5.24) to portray these differences in the distribution of predicted values.

The pattern of the cokriging variances (figure 5.25*c* and d) clearly showed the locations of the ECODATA plots. Since the ground measurements were the best available data on canopy cover, the cokriging variance maps are a useful representation of the relatively high reliability at plot locations. Away from plot locations, cokriging variance increased monotonically, with no obvious TM-image related patterns.

Conditional simulation

Multiple realizations of canopy cover from 1988 (figure 5.26) showed patterns that were much less smooth than cokriged maps yet without the detailed features of the TM image, as in the aspatial regression maps. The region in the upper left corner, where clearcuts existed, showed low values of canopy cover in most of the realizations. Regions of similar values changed shape somewhat from realization-to-realization, while remaining in the same general locations. The semivariograms from the realizations were similar to the model of canopy cover semivariance though with a somewhat shorter range.

The mean of simulated realizations (figure 5.27*a* and *b*) appeared similar to the cokriged maps, but not identical because of the different type of cokriging used for the sequential simulation. The RMSE of the mean map from the 1988 model made at the 1993 plot locations was 14% and that from the mean map from the 1993 model made at the 1988 plot locations was 16%. Accuracies of the 50 individual realizations measured in this way ranged from 15% to 22% and from 15% to 20%, respectively. Therefore, at least one realization happened to be just as accurate as the mean map using this limited evaluation method. Presumably more realizations could be found with this property and this would be a valid way to select a subset of realizations.

The variance of simulated realizations (figure 5.27*c* and d) in general showed low uncertainty near measurement locations, but also included patterns from the TM image (especially areas of low NDVI such as clearcuts and bare-rock outcrops).

Figure 5.25: Cokriged maps of canopy cover (%) from (a) 1988 data, (b) 1993 data and cokriging variance maps from (c) 1988 data and (d) 1993 data. ECODATA plot locations are shown as black squares on (a) and (b).

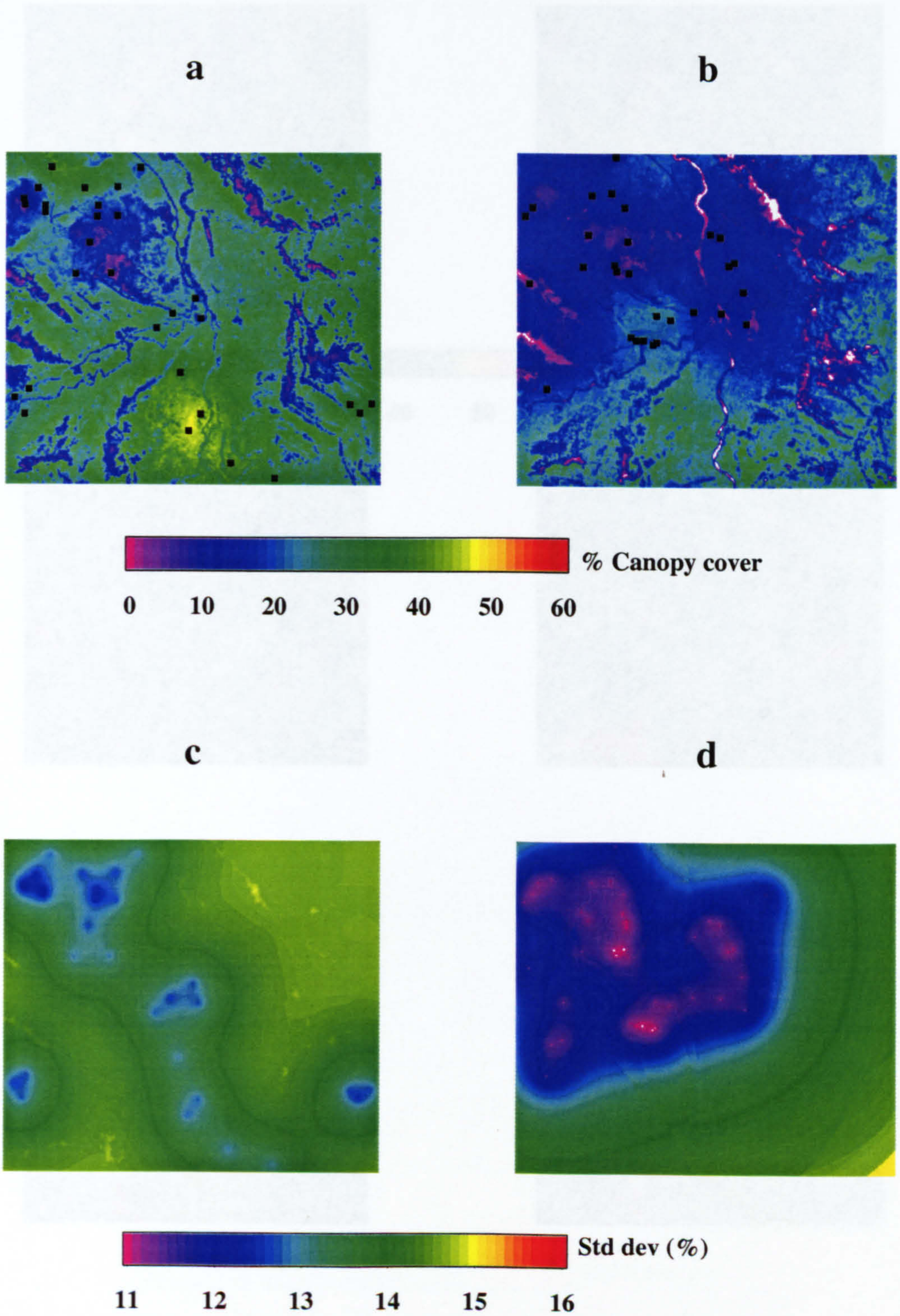


Figure 5.26: Six conditionally simulated realizations of canopy cover (%) from 1988 data.

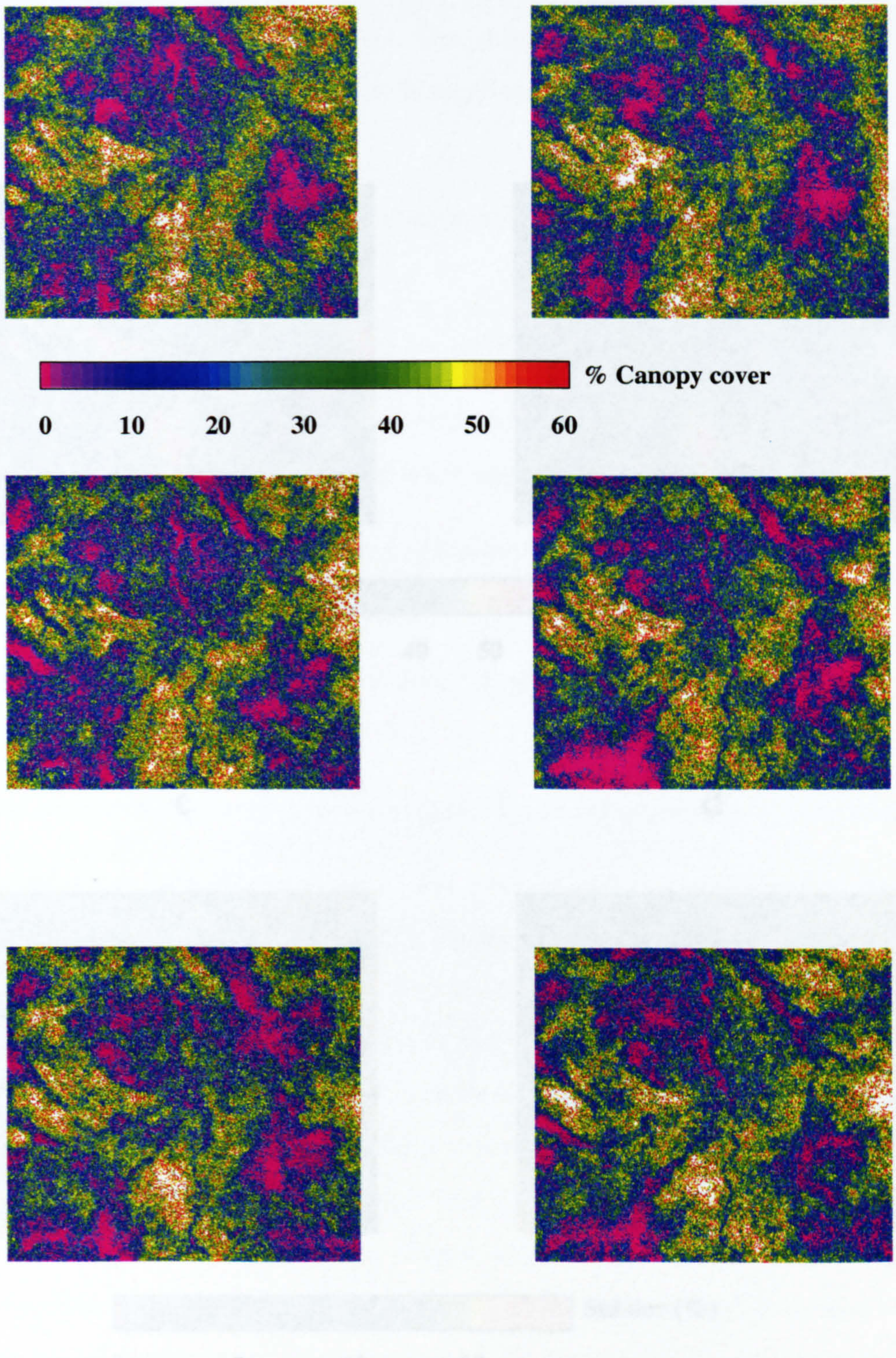
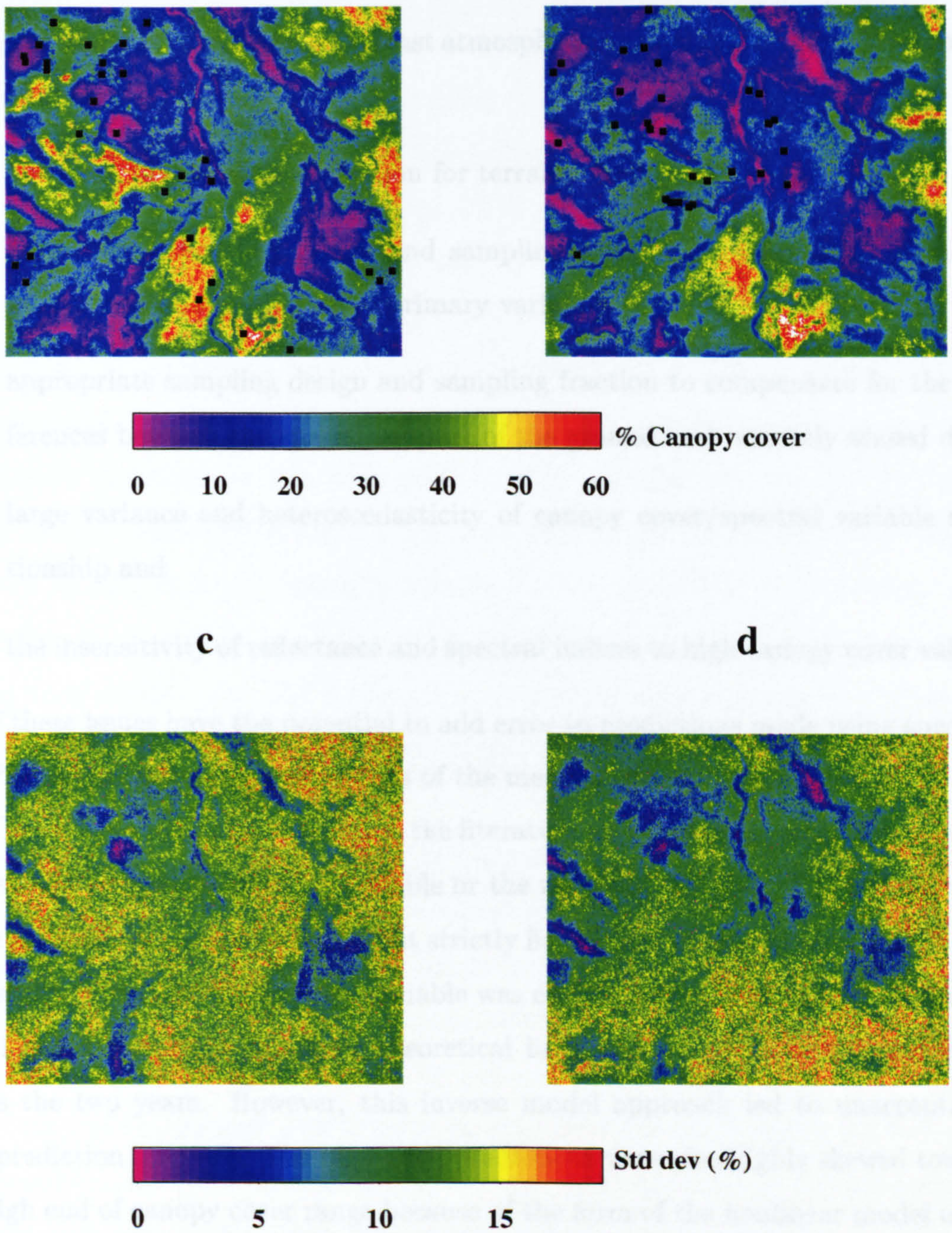


Figure 5.27: (a) Mean of 50 conditionally simulated realizations of canopy cover (%) from 1988 data, (b) mean map from 1993 data, (c) variance map from 50 conditionally simulated realizations of canopy cover (%) from 1988 data and (d) variance map from 1993 data. ECODATA plot locations are shown as black squares on (a) and (b).



5.3 Conclusions

The most common situation in remote sensing studies is the collection of ground measurements in a small region and extrapolation of results to a much larger region. In this study, the reverse approach was taken; data from larger extent were used to predict on values over a smaller extent. But this study had many issues in common with current state-of-the-art remote sensing of vegetation amount quantities. Those issues include:

1. atmospheric correction, or at least atmospheric normalization in multitemporal studies,

2. adequate illumination correction for terrain effects,

3. appropriate sampling design and sampling fraction for estimation of model parameters and prediction of primary variables,

4. appropriate sampling design and sampling fraction to compensate for the differences between the spatial support of the ground and remotely sensed data,

5. large variance and heteroscedasticity of canopy cover/spectral variable relationship and

6. the insensitivity of reflectance and spectral indices to high canopy cover values.

All of these issues have the potential to add error to predictions made using spectral variables as ancillary data, regardless of the method chosen.

There is no clear consensus from the literature about whether the ancillary variable should be the explanatory variable or the response variable. The assumptions of the aspatial regression model do not strictly fit the first or second scenario. In this case study, when the explanatory variable was canopy cover, the aspatial regression model had a form that has some theoretical basis and had consistent parameters across the two years. However, this inverse model approach led to unacceptably high prediction intervals. The regression predictions were also highly skewed toward the high end of canopy cover range because of the form of the nonlinear model used; a distribution not supported by the statistics from the study region. A similar bias was observed by White *et al.* (1997*a*) in another study in western Montana.

Cokriging was greatly affected by the sparsity and clumped distribution of ground measurements. Nonetheless in theory the cokriged maps were much more accurate than those from aspatial regression and they had a smaller error as judged at test locations. Because they also maintained univariate distributions more similar to those of the study area, they were considered preferable to the regression-predicted maps. Cokriging variance maps did not provide a useful description of uncertainty beyond emphasizing the locations of the ECODATA plots.

Conditionally simulated realizations using collocated cokriging seemed to display the desirable aspects of cokriging at the same time as they presented plausible spatial distributions of canopy cover given the data. The realizations appeared qualitatively reasonable, did slightly worse than cokriging at test locations, but reproduced the presumed global *and* spatial distributions. Therefore, conditionally simulated maps were considered preferable to the cokriging maps in this case.

With the satellite sensors and ground measurements available, large uncertainty exists about canopy cover when extrapolated across this landscape, whether using aspatial or geostatistical models. The lack of spatial sampling design in the ECODATA database made it impossible to judge whether the observed change in the global distribution from 1988 to 1993 was caused by an actual decrease in canopy cover or by bias in the sampling locations. Further, the affect of observer error is impossible to characterize retrospectively. Therefore, using this global distribution as the 'true' distribution, and making the assumption that it is equally representative of the subarea used for prediction, involves additional uncertainty beyond that imparted by measurement error.

Predictions can only be quantitatively checked at point locations which make up a very small proportion of the total predicted map. Expressing results of prediction algorithms as maps is an additional tool for evaluation that goes beyond summary expressions such as RMSE at point locations or differences in univariate distributions and semivariograms. Other methods of accuracy assessment, such as checking the consistency of the method over time in an area that has not changed (as was done in this study) may sometimes be necessary. This difficulty in accuracy assessment is inherent to remote sensing of continuous variables as well as to thematic classification of geographical data (Edwards *et al.* 1998).

Results with AVHRR data confirm the difficulties of relating coarse spatial resolution data to ground measurements. Increasing the support of a variable, including remotely sensed spectral variables, involves decreasing variance and increasing symmetry for univariate distributions and as yet unpredictable effects on multivariate distributions. These effects are clearly seen in an example of optical remote sensing of conifer canopy cover that is typical of the kinds of experiments that have been and continue to be conducted in Earth science. Prediction of vegetation amount variables from remotely sensed data and/or other regionalized variables are contingent on the specifications of support for all variables involved. Real remotely sensed reflectance data were compatible with a simple change-of-support model, but NDVI should not be expected to be compatible with this model. New methods to determine the accuracy of predictions made with AVHRR support data are receiving higher priority as these data yield new results and as new sensors with large support sizes are put into space (Justice *et al.* 1998, Verstraete *et al.* 1999, Milne and Cohen 1999).

6 Conclusions

In this chapter, the findings of this thesis are summarized according to the five research questions enumerated in chapter 1. Recommendations for future research based on these findings are then offered.

6.1 Summary

6.1.1 The implications of choosing geostatistical models

A review of the literature (chapter 2) confirmed that, throughout the 1990s, aspatial regression continued to be the most commonly used statistical method for developing prediction models of vegetation amount variables using remote sensing. Geostatistical models were very rarely used for prediction of continuous land surface variables and heretofore had not been directly compared to aspatial regression for mapping vegetation amount. Chapter 3 showed that aspatial and geostatistical regression methods can be unified under the rubric of optimal prediction from different types of random function models. Therefore, an implication of choosing geostatistical models is a decision to exploit the geographical aspects of variables using an autocorrelated random function model rather than a purely random function model.

For all prediction methods, sample size must be adequate to estimate statistics (correlation coefficient, spatial autocorrelation, and distribution parameters) of the primary and ancillary variables. Methods were tested on both synthetic and real data. A sample size of 300 was used for the study of synthetic data, chosen by noting when the mean, variance and semivariogram estimates stabilized. The real data, from a region in western Montana on two dates, was one of the largest data sets used to date for investigating relationships between vegetation amount and spectral data. There were over 300 observations on both dates.

In addition to the sample size, the sampling fraction was particularly relevant for the geostatistical methods, since these methods rely on nearby primary data values to make predictions. The sampling fraction in the synthetic data set, at .3%, appeared to be adequate for achieving accurate cokriging and conditional simulations. The sampling fraction for the Montana data in the subarea where canopy

cover was predicted was an order of magnitude lower. The relative scarcity of neighboring primary values caused high cokriging variances and contributed to regions of high variance across conditional simulations.

Geostatistical prediction requires modeling the spatial autocorrelation of both primary and ancillary variables in the form of semivariograms or other statistics of spatial dependence. Semivariogram models from the synthetic data fit linear models of coregionalization reasonably well. An LMC fit the 1988 Montana data, but was unsatisfactory for the 1993 data.

When canopy cover was predicted using the Montana data set, substantively different maps arose from aspatial regression, cokriging and conditional simulation. Differences included the global histogram, values at sampled locations, smoothness and spatial pattern. The maps resulting from aspatial regression had a mean much higher than the sample mean, implying either a biased result or that the sample from which population parameters were estimated was not representative of the subarea. The cokriging and conditionally simulated maps appeared more accurate as judged at a few point locations. The histogram and spatial autocorrelation of the cokriged maps did not match those of the sample data. Realizations from sequential Gaussian simulation more closely matched the histogram and semivariogram.

The sampling requirements (both number and fraction) for implementing a prediction method are also relevant to assessing the accuracy of mapped predictions. Besides local accuracy at individual locations, criteria for accurate maps include the accurate representation of global and spatial distributions, and potentially the multivariate distribution with other environmental variables. These multiple criteria can be addressed with conditional simulation.

6.1.2 The error and uncertainty components of prediction methods

A direct aspatial regression approach was taken with the synthetic data, where the ancillary data were used as the explanatory variable and the primary data were used as the response variable. For the Montana data, an inverse approach was used instead. The inverse approach discounted errors in the canopy cover data, but yielded consistent regression models for 1988 and 1993. The Montana study

showed that at TM support, no matter which prediction method was used, large uncertainties existed about maps of canopy cover. However, uncertainty descriptions produced by aspatial regression, cokriging and conditional simulation differed.

Results from both synthetic and real data showed that prediction intervals from aspatial regression were a function of the magnitude of the variable, not its location. Cokriging variances, in contrast, were a function of the location of the variable, not its value. Local variances from conditional simulation contained contributions from both value and location. Further, the suite of realizations output from a conditional simulation were useful depictions of global uncertainty.

6.1.3 Aspatial regression versus cokriging

A variety of correlations between ancillary and primary variables were represented in the synthetic data as a case study of this contest. The application of aspatial and geostatistical regression methods to these data suggested that cokriging was more accurate than aspatial regression unless the ancillary data accounted for about 80% or more of the variance in the primary data. This result cannot be used in an absolute sense, but implies that, given adequate sampling, cokriging will likely be more accurate than aspatial regression unless ancillary and primary data are very closely related. This has been an unstated assumption in much of the literature cited in chapter 2, but actual data have failed to bear out these high correlations.

The results of the synthetic data study implied that the NDVI data from Montana, which was significantly but not strongly correlated with canopy cover, would be more useful in a cokriging model to predict canopy cover than they would be in aspatial regression. This implication was supported by quasi-point validation at 30 locations where canopy cover most likely did not substantially change between the two dates.

6.1.4 Statistical relationships observed

In the Montana data set, forest canopy cover was the vegetation amount variable and Landsat Thematic Mapper reflectances and NDVI were the spectral variables. The expected inverse relationships between red reflectance and canopy cover was ob-

served. An inverse relationship between infrared reflectance and canopy cover was observed, confirming similar studies on smaller data sets but contradicting deterministic model-based predictions in the literature. A direct, asymptotic relationship with NDVI was observed, of a form supported by theoretical considerations. The parameters of this model were consistent in the two separate years for which plentiful data were available. Linear models did not account for more than 35% of the variance. Large uncertainties in this relationship were evident from the large scatter.

Autocorrelation was evident in canopy cover and spectral variables beyond a few lags, extending to around 5.5 km. This range or autocorrelation length was thought to represent the effect of terrain on all variables.

6.1.5 The effect of spatial support

The differences observed in the statistical distributions of spectral data from TM and AVHRR were interpreted to be primarily a function of support. The statistical distributions of near infrared reflectances matched the simple support effect model: conservation of the mean, a decrease in variance and an increase in symmetry. The distributions of NDVI did not match this simple model because NDVI is a nonlinear transform of reflectance.

Support was also a factor in the Montana study in the sense that the support of ground measurements differed from that of the remotely sensed measurements. The averaging of several pixels, a common method in studies reviewed in chapter 2, in effect increases the support of the spectral variable, thereby increasing the discrepancy in the experimental units being compared. Averaging may increase the likelihood that the pixel within which ground measurements were made is included in the analysis, but the properties of that pixel are diluted by those of the neighbors averaged in with it. No pixel averaging was done in the Montana study, but the ground measurement support was still less than half the size of the TM pixel.

The spatial heterogeneity of the environment (Myneni *et al.* 1995, Dubayah *et al.* 1997, Hu and Islam 1997, Chen 1999) leads to that fact that, at different supports different environmental factors influence reflectance. At supports in which vegetation is more likely to be homogeneous, red reflectance decreases and infrared

increases as vegetation amount increases. Therefore, NDVI should be more sensitive to vegetation amount than either red or infrared reflectance alone. These supports are likely to be measured in the laboratory or with field spectroradiometers using small spot sizes in the field. At larger supports, such as those from Landsat TM, infrared may decrease as vegetation amount increases because of the increased importance of shadowing at this support. NDVI is therefore less sensitive to changes in vegetation amount than either red or infrared alone. This explanation fits the observations described in chapter 5.

6.2 Recommendations

This thesis has shown that the *choice* of a random function model is critical to the prediction of vegetation amount using remote sensing. The choice of a method for any given prediction problem should be governed by theoretical and practical considerations.

6.2.1 Theoretical considerations

In aspatial regression, autocorrelation in the residuals violates the assumptions of the model. Yet spatial autocorrelation is likely to exist in the residuals of aspatial regression models because all related environmental variation, which tends to be spatially autocorrelated, is not factored in to the model. Therefore, the use of aspatial regression for accurate prediction depends upon verifying that ancillary variables do, in fact, explain almost all of the variance in the primary variable and that residuals are not spatially autocorrelated.

In geostatistical models, autocorrelation is an added source of information used to increase the accuracy of predictions. This is an advantage of geostatistical methods over aspatial regression. Correlated random functions seem more appropriate for geographically-expressed natural phenomena than do purely random functions. For this theoretical advantage to be realized, two aspects of sampling must be considered: adequate sample number and adequate sampling fraction.

The exactness property also distinguishes aspatial regression and geostatistical methods. Unless data are explicitly available (for example, repeated measurements

at the same location) to estimate measurement error (Atkinson *et al.* 1996), geostatistical prediction puts the error into the nugget variance. Aspatial regression, in contrast, puts the error into the residuals.

The stationarity properties of the geostatistical model have been thought by some to be an impediment to their use. However, when aspatial regression models are recognized as purely random functions, it can be seen that they require similar stationarity requirements (Fotheringham *et al.* 1996). This means that whichever method is chosen, conscious decisions must be made as to what region or domain the prediction model will be applied to.

While both aspatial regression and cokriging are locally optimal, they both ignore global considerations. Aspatial regression does not ensure that maps resulting from its application will have any particular statistical properties – if the map's semivariogram matches that of the primary variable it is a side-effect of the fact that the regression model simply rescales the measurement units of reflectance. Cokriging is a smoother, guaranteeing that the predicted map will have a different semivariogram than that measured from the primary data. Conditional simulation is the only prediction method that explicitly considers these global issues.

There is an apparent paradox in the idea that the optimality criteria used in all forms of kriging results in estimates that do not achieve the global statistics of the spatial field. Conversely, conditional simulation mimics the field properties but is locally less accurate – it has a variance twice that of kriging (Journel and Huijbregts 1978, Cressie 1991). In fact, there is some suggestion that local optimality and global accuracy cannot be achieved simultaneously. Some researchers have proposed compromise solutions intermediate between estimation and simulation (Olea and Pawlowsky 1996, Goovaerts 1998*a*). Csillag (1987) states, 'There is a contradiction between the requirements of constant attribute accuracy and constant spatial resolution.' Generally speaking, this dichotomy may be a geographer's analogue to the Heisenberg uncertainty principle. Whether or not this paradox represents something fundamental, the current state of the art in mapping quantitative variables makes it clear that objectives must be defined clearly, preferably *a priori* because the choice of a method affects the possibility of achieving those objectives and no one method allows the simultaneous achievement of all possible objectives.

When vegetation amount maps are used in combination with other spatial data layers in a GIS for the purposes of spatial analysis or use in a deterministic process model, it is the covariance of all these layers that will govern the output of that analysis (Berry 1993, Coughlan and Dungan 1996, Phillips and Marks 1996). Therefore, if the vegetation map is overly smoothed, not only will its global histogram be in error, its covariance with other data layers will be incorrect.

6.2.2 Practical considerations

Practical considerations for applying geostatistical methods include the requirement for plentiful data for estimating the spatial characteristics of ancillary and primary variables. Beyond the need for an adequate sample size, the need exists to pay greater attention to spatial sampling design. Examples of regional sampling efforts, such as those by the National Resources Conservation Service (NRCS) National Resources Inventory (Nusser and Goebel 1997) and the EPA's Environmental Monitoring and Assessment Program (Overton and Stehman 1993) are design-based. The Forest Service data from Montana used in this thesis were not design-based in this sense. This reduced the ability to use these data to estimate the global spatial distribution and its parameters. In cases in which mapping is the objective, it may be worthwhile to collect a larger number of perhaps lower-quality information, since the typically small ground data sample sizes used in remote sensing investigations are not useful for estimating global properties.

Investigations in this thesis used relatively small grids, around 100,000 cells. This is near the limit for geostatistical software available at low cost; use on much larger, more realistically sized grids will have need of specialized software. Algorithms that search large neighborhoods for the closest primary and ancillary data tend to be the most time-consuming to run; these could be speeded up by improvements in software and/or hardware.

The application of cokriging and some forms of conditional simulation requires modeling the semivariogram of the primary variable, the ancillary variable and their cross-covariance. Further, this model must fit stringent mathematical requirements. The data may not always coincide with such a model. To reduce the burden of

cross-semivariogram modeling, Ma and Journel (1999) have investigated simplifying assumptions that express the semivariogram of the ancillary variable (γ_a) and the cross-semivariogram (γ_{pa}) as functions of the semivariogram of the primary variable γ_p. These are worth exploring with remotely sensed data, though the simplification used in the Montana data set did not fit the observed semivariances particularly well.

The choice of conditional simulation algorithm remains an under-explored topic. For example, probability field simulation, applied to the synthetic data in chapter 4, required a slightly different set of implementation steps than sequential Gaussian simulation, applied in chapter 5. Further research is required to clarify this choice for practitioners.

6.2.3 Future research

Currently, scaling is an important research topic (Raffy 1994). Since the noun 'scale', has numerous, sometimes conflicting meanings (Turner and Gardner 1991, Quattrochi 1993, Curran and Atkinson 1999), the verb 'to scale' is equally ambiguous. This thesis identifies one specific aspect of scaling, the change-of-support problem, as highly relevant for the spatial prediction of vegetation amount. The parameters of any prediction model are expected to change with support. Because of the nonlinearity of NDVI with vegetation amount, straightforward regularization models (Rendu 1981) may not be adequate for scaling (Hu and Islam 1997). In fact, it is difficult to envision a case where support will not affect bivariate relations between ground-based and remote measurements, unless the scene were completely random (from the idealized, purely random function model). Simonett (1971) (page iv) stated almost three decades ago that 'Extrapolating from one scale or resolution to another ...turn[s] out to be a sleeper in the remote sensing field', a statement that seems equally applicable to the 1990s.

Given the difficulty in obtaining a large enough sampling fraction for truly regional mapping, a promising approach may be the combination of a deterministic approach with a geostatistical approach. That is, local ccdfs may be defined based on a deterministic, physically based radiative transfer model (i.e. Jasinski 1996)

that does not rely necessarily on local observations. Ccdfs could then be drawn from using a probability field technique. Maps from such an approach would not be conditional to data, but would have a spatial pattern reproducing a semivariogram deemed representative of the region.

Past research has not focused on questions of uncertainty, typically summarizing the quality of an aspatial regression relationship with a single standard error of estimate value. The well-known regression theory for prediction intervals has generally not been employed to describe uncertainty and has not been expressed spatially. A standard error estimate from aspatial regression is a limited figure of merit for the quality of predictions, and geostatistical models provide a rich set of tools that should be exploited in the future to more explicitly characterize uncertainty about remote sensing of vegetation amount.

In situations where an aspatial regression approach is taken, it is important to address the measurement error of both primary and ancillary variables, especially their relative magnitude, to make a wiser choice of explanatory and response variables. Reducing error in the spectral variables will be brought about minimizing sources of error and inconsistency caused by:

- misregistration in time and space,
- different supports between ground measurements and pixels,
- imprecise radiometric calibration,
- failure to remove the atmospheric contribution to the remotely sensed signal and
- failure to account for variation of reflectance with direction caused by terrain and non-Lambertian properties of vegetation

In comparison to the NOAA AVHRR and previous Landsat satellite sensors, these sources of error should be reduced in sensors recently launched or planned for the near future, including the Landsat 7 Enhanced Thematic Mapper (ETM+, Justice *et al.* 1998), the Ikonos radiometer (Space Imaging, Inc. Thornton, CO), the Moderate Resolution Imaging Spectroradiometer (MODIS, Justice *et al.* 1998) and the

Medium Resolution Imaging Spectrometer (MERIS, Verstraete *et al.* 1999); sensors that will surely be used to map vegetation amount.

This has been a broadly ranging thesis, touching on many of the pertinent issues of remote sensing from space, including atmospheric correction, geometric correction and most fundamentally, the issue of prediction. It therefore has relevance to the multitude of other quantitative remote sensing prediction problems, such as efforts to obtain maps of vegetation canopy biochemical content (Curran *et al.* 1997). The results highlight the importance of space in prediction methods, an aspect that in some ways is just beginning to be recognized.

A Effect of image resampling on statistics

During geometric correction of image data, the grid of spectral values must be interpolated to maintain a regular grid (Schowengerdt 1997). This interpolation step is referred to as 'resampling.' The three most common resampling algorithms are nearest neighbor, bilinear interpolation and cubic convolution. As with the geostatistical interpolation algorithms discussed in this thesis, resampling has the potential to change the statistics of the data. A simple illustration using the synthetic data from chapter 4 shows this to be the case.

The true image (Figure 4.1) was 'warped' using a first-order polynomial to simulate a typical geometric transformation. The polynomial was of the form:

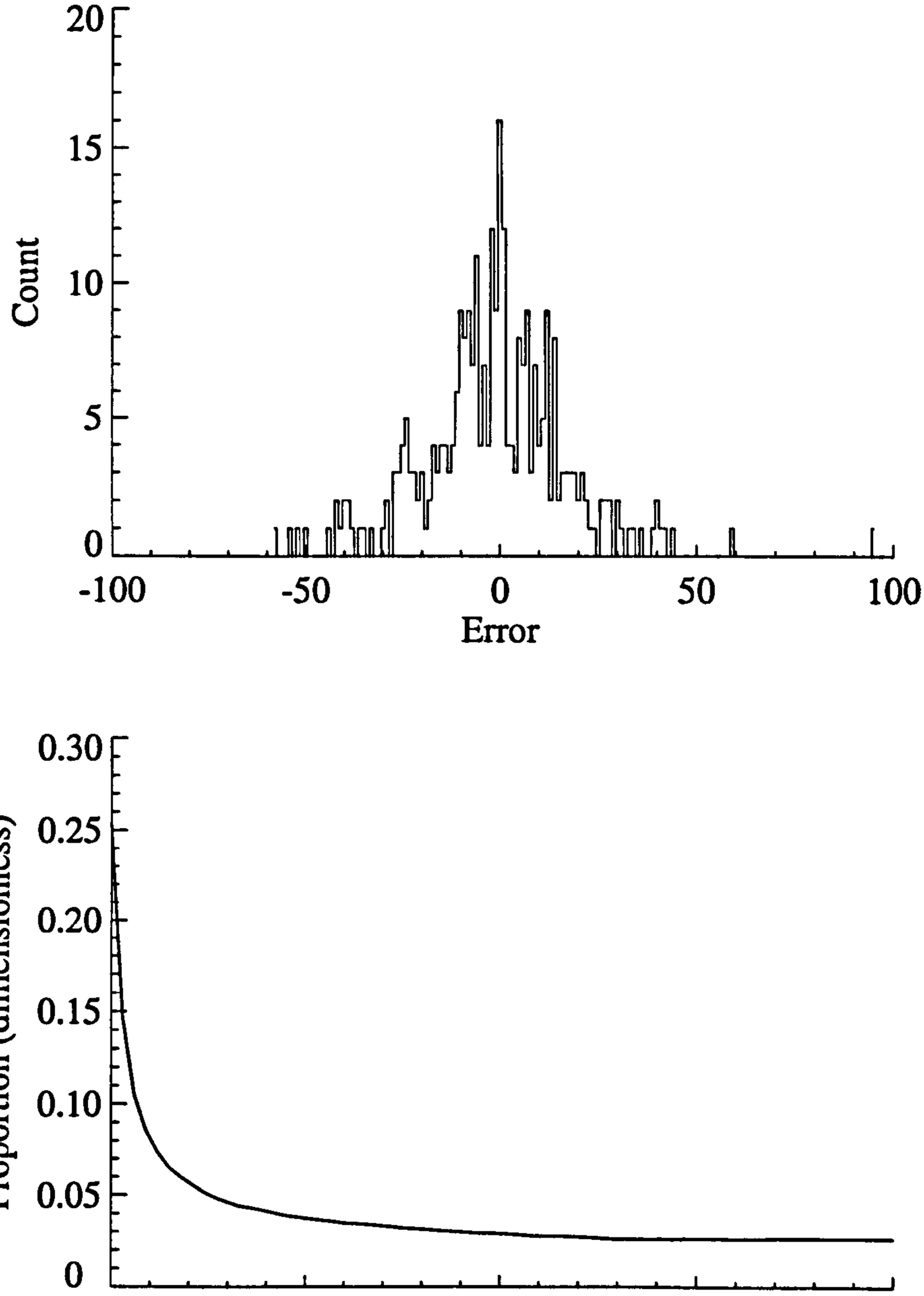
$$\begin{aligned} x' &= \sum_{i=0}^{1}\sum_{j=0}^{1} P_{i,j}x^i y^j \\ y' &= \sum_{i=0}^{1}\sum_{j=0}^{1} Q_{i,j}x^i y^j \end{aligned} \tag{A.1}$$

where x, y were the original coordinates, x', y' were the coordinates of the warped image and all the coefficients $P_{i,j}$ and $Q_{i,j}$ were 0 except $P_{1,0} = .02$, $P_{0,1} = 1.0$, $Q_{1,0} = 1.0$ and $Q_{0,1} = .02$. The resulting image was resampled using nearest neighbor and cubic convolution methods. The new values at the locations of the 300 samples in the resampled images were compared with the true values at their original locations. The nearest neighbor resampling added no error to these values. But the values from the cubicly convoluted image had an RMSE of 18, or 4% of the true mean value of 422, and error values ranged from -258 to 179 (Figure A.1a). The image variance was 3% lower in the cubicly convoluted image than the true image.

Resampling can also affect the semivariogram. Since cubic convolution weights nearby values, it acts as a smoother, leading to a reduction in the sill and nugget semivariances. In the illustration, the decrease in the semivariogram of the cubicly convoluted image as a proportion of the the semivariogram of the nearest-neighbor resampled image (Figure A.1b) showed that nugget and semivariances at small lags are particularly affected by cubic convolution resampling.

While the effects on statistics are small in the illustration, and likely to be similarly small for real data, they should be recognized as a factor in the image processing stream.

Figure A.1: (a) Histogram of differences (unitless) between 300 pixel values from an image resampled by cubic convolution and the corresponding pixels from an image resampled by nearest-neighbor method and (b) Difference between the semivariogram of the nearest-neighbor resampled image and the semivariogram of the cubicly convoluted image, expressed as a proportion of the nearest-neighbor resampled image's semivariogram.



References

Aase, J. K., F. H. Siddoway and J. P. Millard. 1984. Spring wheat-leaf phytomass and yield estimates from airborne scanner and hand-held radiometric measurements. *International Journal of Remote Sensing*, 5:771–781.

Abarca-Hernandez, F. and M. Chica-Olmo. 1999. Evaluation of geostatistical measures of radiometric spatial variability for lithologic discrimination in Landsat TM images. *Photogrammetric Engineering and Remote Sensing*, 65:705–711.

Abdulla, F. A. and D. P. Lettenmaier. 1997. Application of regional parameter estimation schemes to simulate the water balance of a large continental river. *Journal of Hydrology*, 197:258–285.

Addink, E. A. and A. Stein. 1999. A comparison of conventional and geostatistical methods to replace clouded pixels in NOAA-AVHRR images. *International Journal of Remote Sensing*, 20:961–978.

Ahern, F. J., T. Erdle, D. A. Maclean and I. D. Kneppeck. 1991. A quantitative relationship between forest growth rates and Thematic Mapper reflectance measurements. *International Journal of Remote Sensing*, 12:387–400.

Ahmed, S. and G. de Marsily. 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Water Resources Research*, 23:171–1737.

Alabert, F. G. 1987. The practice of fast conditional simulations through the LU decomposition of the covariance matrix. *Mathematical Geology*, 19:369–386.

Allee, R. J. and J. E. Johnson. 1999. Use of satellite imagery to estimate surface chlorophyll a and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA. *International Journal of Remote Sensing*, 20:1057–1072.

Almeida, A. S. and A. G. Journel. 1994. Joint simulation of multiple variables with a Markov-type coregionalization model. *Mathematical Geology*, 26:565–588.

Alt, D. and D. W. Hyndman. 1986. *Roadside Geology of Montana*. Missoula, MT: Mountain Press Publishing Company.

Aman, A., H. P. Randriamananatena, A. Podaire and R. Frouin. 1995. Upscale integration of Normalized Difference Vegetation Index: The problem of spatial heterogeniety. *IEEE Transactions on Geoscience and Remote Sensing*, 30:326–337.

Anderson, B. D. O. and J. B. Moore. 1979. *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall.

Anselin, L. and D. A. Griffith. 1988. Do spatial effects really matter in regression analysis? *Papers of the Regional Science Association*, 65:11–34.

Ardo, J. 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing*, 13:1779–1786.

Armstrong, M. and R. Jabin. 1981. Variogram models must be positive definite. *Mathematical Geology*, 13:455–459.

Asli, M. and D. Marcotte. 1995. Comparison of approaches to spatial estimation in a bivariate context. *Mathematical Geology*, 27:641–658.

Asrar, G., E. T. Kanemasu, R. D. Jackson and P. J. Pinter, Jr. 1985. Estimation of total above-ground phytomass production using remotely sensed data. *Remote Sensing of Environment*, 17:211–220.

Asrar, G., R. Myneni and E. T. Kanemasu. 1989. Estimation of plant canopy attributes form spectral reflectance measurements. In G. Asrar, editor, *Theory and Applications of Optical Remote Sensing*, pages 252–296. New York: John Wiley & Sons.

Atkinson, P. M. 1991. Optimal ground-based sampling for remote sensing investigations - estimating the regional mean. *International Journal of Remote Sensing*, 12:559–567.

Atkinson, P. M. 1993. The effect of spatial resolution on the experimental variogram of airborne MSS imagery. *International Journal of Remote Sensing*, 14:1005–1011.

Atkinson, P. M. 1995. A method for describing quantitatively the information redundancy and error in digital spatial data. In P. Fisher, editor, *Innovations in GIS 2*, pages 85–96. London: Taylor & Francis.

Atkinson, P. M. 1997*a*. On estimating measurement error in remotely-sensed images with the variogram. *International Journal of Remote Sensing*, 18:3075–3084.

Atkinson, P. M. 1997*b*. Scale and spatial dependence. In P. van Gardingen, G. M. Foody and P. J. Curran, editors, *Scaling-up: From Cell to Landscape*, pages 35–60. Cambridge: Cambridge University Press.

Atkinson, P. M. 1999. Geographical information science: Geostatistics and uncertainty. *Progress in Physical Geography*, 23:134–142.

Atkinson, P. M. and P. J. Curran. 1995. Defining an optimal size of support for remote sensing investigations. *IEEE Transactions on Geoscience and Remote Sensing*, 33:768–776.

Atkinson, P. M. and P. J. Curran. 1997. Choosing an appropriate spatial resolution for remote sensing investigations. *Photogrammetric Engineering and Remote Sensing*, 63:1345–1351.

Atkinson, P. M., P. J. Curran and R. Webster. 1990. Sampling remotely sensed imagery for storage, retrieval and reconstruction. *Professional Geographer*, 37:345–353.

Atkinson, P. M., R. Dunn and A. R. Harrison. 1996. Measurement error in reflectance data and its implications for regularizing the variogram. *International Journal of Remote Sensing*, 18:3735–3750.

Atkinson, P. M. and D. R. Emery. 1999. Exploring the relation between spatial structure and wavelength: implications for sampling reflectance in the field. *International Journal of Remote Sensing*, 20:2663–2678.

Atkinson, P. M., R. Webster and P. J. Curran. 1992. Cokriging with ground-based radiometry. *Remote Sensing of Environment*, 41:45–60.

Atkinson, P. M., R. Webster and P. J. Curran. 1994. Cokriging with airborne MSS imagery. *Remote Sensing of Environment*, 50:335–345.

Badwhar, G. D., R. B. MacDonald, F. G. Hall and J. G. Carnes. 1986. Spectral characterization of biophysical characteristics in a boreal forest: relationship between Thematic Mapper band reflectance and leaf area index for aspen. *IEEE Transactions on Geoscience and Remote Sensing*, 24:322–326.

Band, L. E. and I. D. Moore. 1995. Scale – landscape attributes and Geographical Information Systems. *Hydrological Processes*, 9:401–422.

Barnes, R. J. 1991. The variogram sill and the sample variance. *Mathematical Geology*, 23:673–678.

Baulies, X. and X. Pons. 1995. Approach to forestry inventory and mapping by means of multispectral airborne data. *International Journal of Remote Sensing*, 16:61–80.

Beaudoin, A., T. Letoan, S. Goze, E. Nezry, A. Lopes, E. Mougin, C. C. Hsu, H. C. Han, J. A. Kong and R. T. Shin. 1994. Retrieval of forest biomass from SAR data. *International Journal of Remote Sensing*, 15:2777–2796.

Belward, A. S. 1992. Spatial attributes of AVHRR imagery for environmental monitoring. *International Journal of Remote Sensing*, 13:193–208.

Berberoglu, S., C. D. Lloyd, P. M. Atkinson and P. J. Curran. 2000. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. *Computers and Geosciences*, 26:385–396.

Berry, J. K. 1993. Cartographic modeling: The analytical capabilities of GIS. In M. Goodchild, B. Parks and L. Steyaert, editors, *Environmental Modeling with GIS*, pages 58–74. New York: Oxford University Press.

Bhatti, A. U., D. J. Mulla and B. E. Frazier. 1991. Estimation of soil properties and wheat yields on complex eroded hills using geostatistics and Thematic Mapper images. *Remote Sensing of Environment*, 37:181–191.

Bicheron, P. and M. Leroy. 1999. A method of biophysical parameter retrieval at global scale by inversion of a vegetation reflectance model. *Remote Sensing of Environment*, 67:251–266.

Bierkens, M. F. P. and P. A. Burrough. 1993. The indicator approach to categorical soil data: Application to mapping and land use suitability analysis. *Journal of Soil Science*, 44:369–381.

Bilonick, R. A. 1988. Monthly hydrogen ion deposition maps for the northeastern US from July 1982 to September 1984. *Atmospheric Environment*, 22:1909–1924.

Bonan, G. B. 1993. Do biophysics and physiology matter in ecosystem models? *Climatic Change*, 24:281–285.

Bonham, C. D. 1989. *Measurements for terrestrial vegetation*. New York: Wiley.

Borgman, L. E. 1988. New advances in methodology for statistical tests useful in geostatistical studies. *Mathematical Geology*, 20:383–403.

Bourgault, G. 1994. Robustness of noise filtering by kriging analysis. *Mathematical Geology*, 26:733–752.

Box, E. O., B. N. Holben and V. Kalb. 1989. Accuracy of AVHRR vegetation index as a predictor of biomass, primary productivity and net CO$_2$ flux. *Vegetatio*, 80:71–89.

Boyd, D. and P. J. Curran. 1998. Using remote sensing to reduce uncertainties in the global carbon budget: The potential of radiation acquired in middle infrared wavelengths. *Remote Sensing Reviews*, 16:293–327.

Brincombe, A. 1997. A universal translator of linguistic hedges for the handling of uncertainty and fitness-for-use in GIS. In Z. Kemp, editor, *Innovations in GIS 4*, pages 114–126. London: Taylor & Francis.

Bruniquel-Pinel, V. and J. P. Gastellu. 1998. Sensitivity of texture of high resolution images of forest to biophysical and acquisition parameters. *Remote Sensing of Environment*, 65:61–85.

Brus, D. J. and J. J. deGruijter. 1994. Estimation of non-ergodic variograms and their sampling variance by design-based sampling strategies. *Mathematical Geology*, 26:437–454.

Butera, M. K. 1986. A correlation and regression analysis of percent canopy closure versus TMS spectral response for selected forest sites in the San Juan National Forest, Colorado. *IEEE Transactions on Geoscience and Remote Sensing*, 24:122–129.

Carlson, T. N. and D. A. Ripley. 1997. On the relationship between NDVI, fractional vegetation cover and leaf area index. *Remote Sensing of Environment*, 62:241–252.

Carr, J. R. 1990. Application of spatial filter theory to kriging. *Mathematical Geology*, 22:1063–1079.

Carr, J. R. 1996. Spectral and textural classification of single and multiple band digital images. *Computers & Geosciences*, 22:849–865.

Carr, J. R. and D. E. Myers. 1984. Application of the theory of regionalized variables to the spatial analysis of Landsat data. In *Proceedings of the Ninth Pecora Symposium*, pages 55–61, Silver Spring, MD. IEEE.

Carr, J. R. and D. E. Myers. 1985. COSIM: A FORTRAN IV program for co-conditional simulation. *Computers & Geosciences*, 11:675–705.

Carroll, S. S., G. N. Day, N. Cressie and T. R. Carroll. 1995. Spatial modeling of snow water equivalent using airborne and ground-based snow data. *Environmetrics*, 6:127–139.

Cassiani, G. and M. A. Medina. 1997. Incorporating auxiliary geophysical data into ground-water estimation. *Ground Water*, 35:79–91.

Chainey, S. and N. Stuart. 1998. Stochastic simulation: an alternative for digital geographic information systems. In S. Carver, editor, *Innovations in GIS 4*, pages 3–24. London: Taylor & Francis.

Chandrasekhar, S. 1960. *Radiative Transfer*. New York: Dover.

Chatterjee, S. and A. S. Hadi. 1988. *Sensitivity Analysis in Linear Regression*. New York: John Wiley & Sons.

Chavez, P. S. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24:459–479.

Chen, J. M. 1999. Spatial scaling of a remotely sensed surface parameter by contexture. *Remote Sensing of Environment*, 69:30–42.

Chen, J. M. and T. A. Black. 1992. Defining leaf-area index for non-flat leaves. *Plant, Cell and Environment*, 15:421–429.

Chen, J. M. and J. Cihlar. 1995. Quantifying the effect of canopy architecture on optical measurements of leaf area index using two gap size analysis methods. *IEEE Transactions on Geoscience and Remote Sensing*, 55:777–787.

Chen, J. M. and J. Cihlar. 1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, 55:153–162.

Chong, D. L. S., E. Mougin and J. P. Gastellu-Etchegorry. 1993. Relating the Global Vegetation Index to net primary productivity and actual evapotranspiration over Africa. *International Journal of Remote Sensing*, 14:1517–1546.

Choudhury, B. J., N. U. Ahmed, S. B. Idso, R. J. Reginato and C. S. T. Daughtry. 1994. Relations between evaporation coefficients and vegetation indexes studied by model simulations. *Remote Sensing of Environment*, 50:1–17.

Christakos, G. 1992. *Random Field Models in the Earth Sciences*. San Diego: Academic Press.

Christakos, G. and C. Panagopoulos. 1992. Space transformation methods in the representation of geophysical random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 30:55–70.

Clark, I. 1979. *Practical Geostatistics*. London: Applied Science.

Clevers, J. G. P. W. 1988. The derivation of a simplified reflectance model for the estimation of leaf area index. *Remote Sensing of Environment*, 25:53–69.

Clevers, J. G. P. W. 1989. The application of a weighted infrared-red vegetation index for estimating leaf area index by correcting for soil moisture. *Remote Sensing of Environment*, 29:25–37.

Cliff, A. D. and J. K. Ord. 1981. *Spatial Process: Models and Applications*. London: Pion Limited.

Cohen, W. B., T. A. Spies and G. A. Bradshaw. 1990. Semivariograms of digital imagery for analysis of conifer canopy structure. *Remote Sensing of Environment*, 34:167–178.

Collins, J. B. and C. E. Woodcock. 1999. Geostatistical estimation of resolution-dependent variance in remotely sensed images. *Photogrammetric Engineering and Remote Sensing*, 65:41–50.

Colwell, J. E. 1974. Vegetation canopy reflectance. *Remote Sensing of Environment*, 3:175–183.

Conese, C., M. A. Gilabert, F. Maselli and L. Bottai. 1993. Topographic normalization of TM scenes through the use of an atmospheric correction method and digital terrain models. *Photogrammetric Engineering and Remote Sensing*, 59:1745–1753.

Congalton, R. G. 1988*a*. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 54:593–600.

Congalton, R. G. 1988*b*. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 54:587–592.

Cook, E. A., L. R. Iverson and R. L. Graham. 1989. Estimating forest productivity with Thematic Mapper and biogeographical data. *Remote Sensing of Environment*, 28:131–141.

Coops, N., A. Delahaye and E. Pook. 1997. Estimation of eucalypt forest leaf area index on the south coast of New South Wales using Landsat MSS data. *Australian Journal of Botany*, 45:757–769.

Coughlan, J. C. and J. L. Dungan. 1996. Combining remote sensing and forest ecosystem modeling: An example using the Regional HydroEcological Simulation System (RHESSys). In H. L. Gholz, K. Nakane and H. Shinoda, editors, *The Use of Remote Sensing in the Modeling of Forest Productivity at Scales from the Stand to the Globe*, pages 139–158. Dordrecht: Kluwer.

Cox, D. R. and H. D. Miller. 1965. *The Theory of Stochastic Processes*. New York: John Wiley & Sons.

Craig, R. G. 1979. Autocorrelation in Landsat data. In *Proceedings of the 13th International Symposium on Remote Sensing of the Environment*, pages 1517–1524, Ann Arbor, MI. University of Michigan.

Cressie, N. 1990. The origins of kriging. *Mathematical Geology*, 22:239–252.

Cressie, N. 1991. *Spatial Data Analysis*. New York: John Wiley & Sons.

Cressie, N. 1996. Change of support and the modifiable areal unit problem. *Geographical Systems*, 3:159–180.

Crist, E. P. and R. J. Kauth. 1986. The tasseled cap de-mystified. *Photogrammetric Engineering and Remote Sensing*, 52:81–86.

Csillag, F. 1987. A cartographer's approach to quantitative mapping of spatial variability from ground sampling to remote sensing of soils. In *Proceedings, AUTO-CARTO 8*, pages 155–164, Falls Church, VA. ACSM/ASPRS.

Curran, P. J. 1980. Multispectral remote sensing of vegetation amount. *Progress in Physical Geography*, 4:315–341.

Curran, P. J. 1988. The semivariogram in remote sensing: An introduction. *Remote Sensing of Environment*, 24:493–507.

Curran, P. J. and P. M. Atkinson. 1998. Geostatistics and remote sensing. *Progress in Physical Geography*, 22:61–78.

Curran, P. J. and P. M. Atkinson. 1999. Issues of scale and optimal pixel size. In A. Stein, F. van der Meer and B. G. H. Gorte, editors, *Spatial Statistics for Remote Sensing*, volume 115-133. Dordrecht: Kluwer.

Curran, P. J. and J. L. Dungan. 1989. Estimation of signal-to-noise: a new procedure applied to AVIRIS data. *IEEE Transactions on Geoscience and Remote Sensing*, 27:620–628.

Curran, P. J., J. L. Dungan and H. L. Gholz. 1992. Seasonal LAI in slash pine estimated with Landsat TM. *Remote Sensing of Environment*, 39:3–13.

Curran, P. J. and A. M. Hay. 1986. The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing*, 52:229–241.

Curran, P. J., J. A. Kupiec and G. M. Smith. 1997. Remote sensing the biochemical composition of a slash pine canopy. *IEEE Transactions on Geoscience and Remote Sensing*, 35:415–420.

Curran, P. J., E. J. Milton, P. M. Atkinson and G. M. Foody. 1998. Remote sensing: from data to understanding. In P. A. Longley, S. M. Brooks, R. McDonnell and B. Macmillan, editors, *Geocomputation: A Primer*, pages 33–59. New York: John Wiley & Sons.

Curran, P. J. and H. D. Williamson. 1986. Sample size for ground and remotely sensed data. *Remote Sensing of Environment*, 20:31–41.

Curran, P. J. and H. D. Williamson. 1987. Airborne MSS data to estimate GLAI. *International Journal of Remote Sensing*, 8:57–74.

Danson, F. M. 1987. Preliminary evaluation of the relationships between SPOT-1 HRV data and forest stand parameters. *International Journal of Remote Sensing*, 8:1571–1575.

Danson, F. M. and P. J. Curran. 1993. Factors affecting the remotely sensed response of coniferous forest plantations. *Remote Sensing of Environment*, 43:55–65.

Daubenmire, R. 1968. *Plant Communities: A Textbook of Plant Synecology*. New York: Harper & Row.

Daughtry, C. S. T. 1990. Direct measurements of canopy structure. *Remote Sensing Reviews*, 5:45–60.

David, M. 1977. *Geostatistical Ore Reserve Estimation*. New York: Elsevier.

Davidson, E. A. and P. A. Lefebvre. 1993. Estimating regional carbon stocks and spatially covarying edaphic factors using soil maps at 3 scales. *Biogeochemistry*, 22:107–131.

Davis, B. M. 1987. Uses and abuses of cross-validation in geostatistics. *Mathematical Geology*, 19:241–248.

deGruijter, J. J. and C. J. F. terBraak. 1990. Model-free estimation from spatial samples: A reappraisal of classical sampling theory. *Mathematical Geology*, 22:407–415.

Demetriades-Shah, T. H., E. T. Kanemasu and I. D. Flitcroft. 1992. Comparison of ground- and satellite-based measurements of the fraction of photosynthetically active radiation intercepted by tallgrass prairie. *Journal of Geophysical Research – Atmospheres*, 97:18947–18950.

Deutsch, C. V. and P. W. Cockerham. 1994. Practical considerations in the application of simulated annealing to stochastic simulation. *Mathematical Geology*, 26:67–82.

Deutsch, C. V. and A. G. Journel. 1992. *GSLIB: Geostatistical Software Library and User's Guide*. New York: Oxford University Press.

Deutsch, C. V. and A. G. Journel. 1998. *GSLIB: Geostatistical Software Library and User's Guide*. New York: Oxford University Press, second edition.

Dewulf, R. R., R. E. Goossens, B. P. DeRoover and F. C. Borry. 1990. Extraction of forest stand parameters from panchromatic and multispectral SPOT-1 data. *International Journal of Remote Sensing*, 11:1571–1588.

Diallo, O., A. Diouf, N. P. Hanan, A. Ndiaye and Y. Prevost. 1991. AVHRR monitoring of savanna primary production in Senegal, West Africa: 1987-1988. *International Journal of Remote Sensing*, 12:1259–1279.

Dickinson, R. E., A. Henderson-Sellers, C. Rosenzweig and P. J. Sellers. 1991. Evapotranspiration models with canopy resistance for use in climate models – a review. *Agricultural and Forest Meteorology*, 54:373–388.

Dobson, M. C., F. T. Ulaby, T. Letoan, A. Beaudoin, E. S. Kasischke and N. Christensen. 1992. Dependence of radar backscatter on coniferous forest biomass. *IEEE Transactions on Geoscience and Remote Sensing*, 30:412–415.

Dowd, P. A. 1992. A review of recent developments in geostatistics. *Computers & Geosciences*, 17:1481–1500.

Draper, N. R. and H. Smith. 1998. *Applied Regression Analysis*. New York: John Wiley & Sons, third edition.

Dubayah, R., E. F. Wood and D. Lavallee. 1997. Multiscaling analysis in distributed modeling and remote sensing: An application using soil moisture. In D. Quattrochi and M. Goodchild, editors, *Scale in Remote Sensing and GIS*, pages 93–112. Boca Raton: CRC Lewis Publishers.

Dubrule, O. 1983. Two methods with different objectives: splines and kriging. *Mathematical Geology*, 15:245–257.

Dufrene, E. and N. Breda. 1995. Estimation of deciduous forest leaf area index using direct and indirect methods. *Oecologia*, 104:156–162.

Duncan, J., D. Stow, J. Franklin and A. Hope. 1993. Assessing the relationship between spectral vegetation indices and shrub cover in the Jornada Basin, New Mexico. *International Journal of Remote Sensing*, 14:3395–3416.

Dungan, J. L. 1998. Spatial prediction of vegetation quantities using ground and image data. *International Journal of Remote Sensing*, 19:267–285.

Dungan, J. L. 1999. Conditional simulation: An alternative to estimation for achieving mapping objectives. In A. Stein, F. van der Meer and B. G. H. Gorte, editors, *Spatial Statistics for Remote Sensing*, volume 1999, pages 135–152. Dordrecht: Kluwer.

Dungan, J. L. 2000. Scaling up and scaling down: Relevance of the support effect in remote sensing of vegetation. In N. Tate and P. Atkinson, editors, *Scales and GIS*, volume in press. Wiley and Sons.

Dungan, J. L., D. L. Peterson and P. J. Curran. 1994. Alternative approaches for mapping vegetation quantities using ground and image data. In W. K. Michener, J. W. Brunt and S. G. Stafford, editors, *Environmental Information Management and Analysis: Ecosystem to Global Scales*, pages 237–261. London: Taylor & Francis.

Easterling, W. E., A. Weiss, C. J. Hays and L. O. Mearns. 1998. Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains. *Agricultural and Forest Meteorology*, 90:51–63.

Edwards, T. C., Jr., G. G. Moisen and D. R. Cutler. 1998. Assessing map accuracy in a remotely sensed, ecoregion-scale cover map. *Remote Sensing of Environment*, 63:73–83.

Efron, B. 1982. The Jackknife, the Bootstrap, and other Resampling Plans. In *CBMS-NSF Regional Conference Series in Applied Mathematics*, 38, page 92, Philadelphia. Society for Industrial and Applied Mathematics.

Eid, T. and E. Naesset. 1998. Determination of stand volume in practical forest inventories based on field measurements and photo-interpretation: The Norwegian experience. *Scandinavian Journal of Forest Research*, 13:246–254.

Eidenshink, J. C. and J. L. Faundeen. 1994. The 1km AVHRR global land data set – 1st stages in implementation. *International Journal of Remote Sensing*, 15:3443–3462.

Emanuel, A. S., R. Behrens and T. Hewett. 1988. Reservoir performance prediction methods based on fractal geostatistics. *American Association of Petroleum Geologists Bulletin*, 72:181.

Emerson, C. W., N. S. N. Lam and D. A. Quattrochi. 1999. Multi-scale fractal analysis of image texture and pattern. *Photogrammetric Engineering and Remote Sensing*, 65:51–61.

ERDAS. 1997. *ERDAS Field Guild*. Atlanta, GA: ERDAS, Inc., fourth edition.

Everitt, J. H., M. A. Hussey, D. E. Escobar, P. R. Nixon and B. Pinkerton. 1986. Assessment of grassland phytomass with airborne video imagery. *Remote Sensing of Environment*, 20:299–306.

Fassnacht, K. S., S. T. Gower, M. D. MacKenzie, E. V. Nordheim and T. M. Lillesand. 1997. Estimating the leaf area index of North Central Wisconsin forests using the Landsat Thematic Mapper. *Remote Sensing of Environment*, 61:229–245.

Fedorov, V. V. 1989. Kriging and other estimators of spatial field characteristics (with special reference to environmental studies). *Atmospheric Environment*, 23:175–184.

Fields, R. W., editor. 1971. *Geology of Western Montana*, volume 2. Missoula, MT: Department of Geology.

Fischer, A. 1994. A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing of Environment*, 48:220–230.

Fisher, P. F. 1997. The pixel – a snare and a delusion. *International Journal of Remote Sensing*, 18:679–685.

Foody, G. M., M. B. McCulloch and W. B. Yates. 1995. Classification of remotely-sensed data by an artificial neural network – issues related to training data characteristics. *Photogrammetric Engineering and Remote Sensing*, 61:391–401.

Forshaw, M. R. B., A. Haskell, P. F. Miller, D. J. Stanley and J. R. G. Townshend. 1983. Spatial resolution of remotely sensed imagery: A review paper. *International Journal of Remote Sensing*, 4:371–383.

Fotheringham, A. S., M. Charlton and C. F. Brunsdon. 1996. The geography of parameter space: An investigation of spatial non-stationarity. *International Journal of Geographical Information Systems*, 10:605–627.

Fotheringham, A. S. and D. W. S. Wong. 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23:1025–1044.

Frank, T. D., S. A. Tweddale and D. E. Knapp. 1994. Variability of at-satellite surface reflectance from Landsat TM and NOAA AVHRR in Death Valley National Monument. *Photogrammetric Engineering and Remote Sensing*, 60:1259–1266.

Franklin, J. 1986. Thematic Mapper analysis of coniferous forest structure and composition. *International Journal of Remote Sensing*, 7:1287–1301.

Franklin, S. E., M. B. Lavigne, M. J. Dueling, M. Wulder and E. R. Hunt, Jr. 1997. Estimation of forest stand leaf area index using remote sensing and GIS data for modelling net primary production. *International Journal of Remote Sensing*, 18:3459–3471.

Franklin, S. E., M. B. Lavigne, B. A. Wilson and E. R. Hunt, Jr. 1994. Empirical relations between balsam fir (Abies balsamea) stand conditions and ERS-1 SAR data in western Newfoundland. *Canadian Journal of Remote Sensing*, 20:124–130.

Friedl, M. A., F. W. Davis, J. Michaelsen and M. A. Moritz. 1995. Scaling and uncertainty in the relationship between the NDVI and land surface biophysical variables – an analysis using a scene simulation model and data from FIFE. *Remote Sensing of Environment*, 54:233–246.

Friedl, M. A., J. Michaelsen, F. W. Davis, H. Walker and D. S. Schimel. 1994. Estimating grassland biomass and leaf area index using ground and satellite data. *International Journal of Remote Sensing*, 15:1401–1420.

Froidevaux, R. 1993. Probability field simulation. In A. Soares, editor, *Geostatistics Troia '92*, pages 73–83. Dordrecht: Kluwer.

Gao, B. K. 1993. An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers. *Remote Sensing of Environment*, 43:23–33.

Gates, D. M. 1980. *Biophysical Ecology*. New York: Springer-Verlag.

Gates, D. M., J. M. Keegan, J. C. Shleter and V. R. Weidner. 1965. Spectral properties of plants. *Applied Optics*, 4:11–20.

Geary, R. C. 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 5:115–145.

Gehlke, C. E. and K. Biehl. 1934. Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29:169–170.

Gemmell, F. and J. Varjo. 1999. Utility of reflectance model inversion versus two spectral indices for estimating biophysical characteristics in a boreal forest test site. *Remote Sensing of Environment*, 68:95–111.

Gemmell, F. M. 1995. Effects of forest cover, terrain, and scale on timber volume estimation with Thematic Mapper data in a Rocky Mountain site. *Remote Sensing of Environment*, 51:291–305.

Gholz, H. L. 1982. Environmental limits on aboveground net primary production, leaf area, and biomass in vegetation zones of the Pacific Northwest. *Ecology*, 63:469–481.

Gobron, H., B. Pinty and M. M. Verstraete. 1997. Theoretical limits to the estimation of the Leaf Area Index on the basis of visible and near-infrared remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 35:1438–1445.

Goel, N. S. and J. M. Norman. 1992. Biospheric models, measurements and remote sensing of vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 47:163–188.

Gohin, F. and G. Langlois. 1993. Using geostatistics to merge in situ measurements and remotely-sensed observations of sea surface temperature. *International Journal of Remote Sensing*, 14:9–19.

Goldberger, A. 1962. Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association*, 57:369–375.

Gong, P., R. L. Pu and J. R. Miller. 1995. Coniferous forest leaf area index estimation along the Oregon transect using compact airborne spectrographic imager data. *Photogrammetric Engineering and Remote Sensing*, 61:1107–1117.

Goodchild, M. F. 1980. Simulation of autocorrelation for aggregate data. *Environment and Planning A*, 12:1073–1081.

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

Goovaerts, P. 1998*a*. Accounting for estimation optimality criteria in simulated annealing. *Mathematical Geology*, 30:511–534.

Goovaerts, P. 1998*b*. Ordinary cokriging revisited. *Mathematical Geology*, 30:21–42.

Goovaerts, P. 1999. Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena*, 34:227–242.

Goovaerts, P. and A. G. Journel. 1995. Integrating soil map information in modeling the spatial variation of continuous soil properties. *European Journal of Soil Science*, 46:397–414.

Goward, S. N. and K. F. Huemmrich. 1992. Vegetation canopy PAR absorptance and the normalized difference vegetation index – an assessment using the SAIL model. *Remote Sensing of Environment*, 39:119–140.

Gower, S. T., C. C. Grier, D. F. Vogt and K. A. Vogt. 1987. Allometric relations of deciduous (Larix occidentalis) and evergreen (Pinus contorta and Pseudostuga menziesii) conifers of the Cascade Mountains in central Washington. *Canadian Journal of Forest Research*, 17:630–634.

Gower, S. T., C. T. Kucharik and J. M. Norman. 1999. Direct and indirect estimation of leaf area index, f(APAR) and net primary production of terrestrial ecosystems. *Remote Sensing of Environment*, pages 29–51.

Graetz, R. 1985. *Montana's Bob Marshall Country*. Helena, MT: Montana Magazine, Inc.

Gregotski, M. E. and O. Jensen. 1993. Fractal modeling techniques for spatial data. *IEEE Transactions on Geoscience and Remote Sensing*, 31:980–988.

Griffith, D. A. 1987. *Spatial Autocorrelation: A Primer*. Washington, DC: Association of American Geographers.

Gross, M. F., M. A. Hardisky, V. Klemas and P. L. Wolf. 1987. Quantification of biomass of marsh grass Spartina alterniflora Loisel using Landsat Thematic Mapper imagery. *Photogrammetric Engineering and Remote Sensing*, 53:1577–1583.

Gu, X. F., G. Guyot and M. Verbrugghe. 1992. Evaluation of measurement errors in ground surface reflectance for satellite calibration. *International Journal of Remote Sensing*, 13:2531–2546.

Gutman, G. and A. Ignatov. 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *International Journal of Remote Sensing*, 19:1533–1543.

Guyot, G. and X. F. Gu. 1994. Effect of radiometric corrections on NDVI-determined from SPOT-HRV and Landsat-TM data. *Remote Sensing of Environment*, 49:169–180.

Haining, R. P. 1990. *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: Cambridge University Press.

Haining, R. P., D. A. Griffith and R. Bennett. 1989. Maximum likelihood estimation with missing spatial data and with an application to remotely sensed data. *Communications in Statistics-Theory and Methods*, 18:1875–1894.

Hall, F. G., J. R. G. Townshend and E. T. Engman. 1995. Status of remote sensing algorithms for estimation of land surface state parameters. *Remote Sensing of Environment*, 51:135–156.

Haralick, R. M., K. Shanmugam and I. Dinstein. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3:610–621.

Hart, J. F. 1954. Central tendency in areal distributions. *Economic Geography*, 30:48–59.

Hay, A. M. 1979. Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, 45:529–533.

Heilman, J. L., E. T. Kanemasu, J. O. Bagley and V. P. Rasmussen. 1977. Evaluating soil moisture and yield of winter wheat in the Great Plains using Landsat data. *Remote Sensing of Environment*, 6:315–326.

Heller, R. C. and J. J. Ulliman. 1983. Forest resource assessments. In R. N. Colwell, editor, *Manual of Remote Sensing*, pages 2229–2324. Falls Church, VA: American Society of Photogrammetry, second edition.

Henderson-Sellers, A. 1993. Continental vegetation as a dynamic component of a global climate model - a preliminary assessment. *Climatic Change*, 23:337–377.

Herwitz, S., D. L. Peterson and J. R. Eastman. 1990. Thematic Mapper detection of changes in the leaf area of closed canopy pine plantations in central Massachusetts. *Remote Sensing of Environment*, 29:129–140.

Hoaglin, D. C., F. Mosteller and J. W. Tukey. 1983. *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley & Sons.

Hogg, R. V. and E. A. Tanis. 1983. *Probability and Statistical Inference*. New York: Macmillan.

Holm, R. G., M. S. Moran, R. D. Jackson, P. N. Slater, B. Yuan and S. F. Biggar. 1989. Surface reflectance factor retrieval from Thematic Mapper data. *Remote Sensing of Environment*, 27:47–57.

Holmgren, P. and T. Thuresson. 1998. Satellite remote sensing for forestry planning – A review. *Scandinavian Journal of Forest Research*, 13:90–110.

Hu, Z. L. and S. Islam. 1997. Scale invariant remote sensing algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 35:747–755.

Huet, S., A. Bouvier, M. A. Gruet and E. Joliet. 1996. *Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS Examples*. New York: Springer-Verlag.

Huete, A. R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25:295–309.

Hughes, J. P. and D. P. Lettenmaier. 1981. Data requirements for kriging – estimation and network design. *Water Resources Research*, 17:1641–1650.

Hyppanen, H. 1996. Spatial autocorrelation and optimal spatial resolution of optical remote sensing data in a boreal forest environment. *International Journal of Remote Sensing*, 17:3441–3452.

Inggs, M. R. and R. T. Lord. 1996. Interpolating satellite-derived wind-field data using ordinary kriging, with application to the nadir gap. *IEEE Transactions on Geoscience and Remote Sensing*, 34:250–256.

Irons, J. R. and G. W. Peterson. 1981. Texture transforms of remote sensing data. *Remote Sensing of Environment*, 11:359–370.

Isaaks, E. H. 1990. *The Application of Monte Carlo Methods to the Analysis of Spatially Correlated Data*. Ph.D. Dissertation. Stanford, California: Stanford University.

Isaaks, E. H. and R. M. Srivastava. 1988. Spatial continuity measures for probabilistic and deterministic geostatistics. *Mathematical Geology*, 20:313–341.

Isaaks, E. H. and R. M. Srivastava. 1989. *An Introduction to Applied Geostatistics*. New York: Oxford University Press.

Ishida, T. and H. Ando. 1999. Use of disjunctive cokriging to estimate soil organic matter from Landsat Thematic Mapper image. *International Journal of Remote Sensing*, 20:1549–1565.

Iverson, L. R., E. A. Cook and R. L. Graham. 1994. Regional forest cover estimation via remote sensing – the calibration center concept. *Landscape Ecology*, 9:159–174.

Jaggi, S., D. A. Quattrochi and N. S. N. Lam. 1993. Implementation and operation of 3 fractal measurement algorithms for analysis of remote-sensing data. *Computers & Geosciences*, 19:745–767.

Jakubauskas, M. E. 1996. Thematic Mapper characterization of lodgepole pine seral stages in Yellowstone National Park, USA. *Remote Sensing of Environment*, 56:118–132.

Jasinski, M. F. 1990. Sensitivity of the normalized difference vegetation index to subpixel canopy cover, soil albedo, and pixel scale. *Remote Sensing of Environment*, 32:169–187.

Jasinski, M. F. 1996. Estimation of subpixel vegetation density of natural regions using satellite multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 34:804–813.

Jensen, J. R. 1983. Biophysical remote sensing. *Annals of the Association of American Geographers*, 73:111–132.

Jensen, J. R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, NJ: Prentice Hall, second edition.

Jensen, J. R. and M. E. Hodgson. 1985. Remote sensing forest biomass: an evaluation using high resolution remote sensor data and loblolly pine plots. *Professional Geographer*, 37:46–56.

Jensen, M. A., W. J. Hann and R. Keane. 1992. *Ecosystem Inventory and Analysis Guide, Draft*. Missoula, MT: USDA, Forest Service Northern Region.

Johnson, J. A. and D. W. Wichern. 1982. *Applied Multivariate Statistical Analysis*. Englewood Cliffs, NJ Prentice-Hall.

Johnson, R. W. 1978. Mapping of chlorophyll a distributions in coastal zones. *Photogrammetric Engineering and Remote Sensing*, 44:617–624.

Johnston, R. J. 1978. *Multivariate Statistical Analysis in Geography: A Primer on the General Linear Model*. London: Longman.

Jones, H. E. 1937. Some geometrical considerations in the general theory of fitting lines and planes. *Metron*, 13:21–32.

Jordan, C. F. 1969. Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50:663–666.

Journel, A. G. 1974. Geostatistics for conditional simulation of ore bodies. *Economic Geology*, 69:527–545.

Journel, A. G. 1986. Geostatistics: Models and tools for the earth sciences. *Mathematical Geology*, 18:119–140.

Journel, A. G. 1990. *Fundamentals of Geostatistics in Five Lessons*. Washington, DC: American Geophysical Union.

Journel, A. G. 1993. Geostatistics: Roadblocks and challenges. In A. Soares, editor, *Geostatistics Troia '92*, pages 213–224. Dordrecht: Kluwer.

Journel, A. G. 1996. Modeling uncertainty and spatial dependence: Stochastic imaging. *International Journal of Geographical Information Systems*, 10:517–522.

Journel, A. G. and F. Alabert. 1989. Non-Gaussian data expansion in the earth sciences. *Terra Nova*, 1:123–134.

Journel, A. G. and C. J. Huijbregts. 1978. *Mining Geostatistics*. London: Academic Press.

Journel, A. G. and E. H. Isaaks. 1984. Conditional indicator simulation: Application to a Saskatchewan uranium deposit. *Mathematical Geology*, 16:685–718.

Journel, A. G. and M. E. Rossi. 1989. When do we need a trend model in kriging? *Mathematical Geology*, 21:715–738.

Jupp, D. L. B., A. H. Strahler and C. E. Woodcock. 1988. Autocorrelation and regularization in digital images. I. Basic theory. *IEEE Transactions on Geoscience and Remote Sensing*, 26:463–473.

Jupp, D. L. B., A. H. Strahler and C. E. Woodcock. 1989. Autocorrelation and regularization in digital images. II. Simple image models. *IEEE Transactions on Geoscience and Remote Sensing*, 27:247–258.

Justice, C. O., J. R. G. Townshend and V. L. Kalb. 1991. Representation of vegetation by continental data sets derived from NOAA-AVHRR data. *International Journal of Remote Sensing*, 12:999–1021.

Justice, C. O., E. Vermote, J. R. G. Townshend, R. Defries, D. P. Roy, D. K. Hall, V. V. Salomonson, J. L. Privette, G. Riggs, A. Strahler, W. Lucht, R. B. Myneni, Y. Knyazikhin, S. W. Running, R. R. Nemani, Z. M. Wan, A. R. Huete, W. vanLeeuwen, R. E. Wolfe, L. Giglio, J.-P. A. L. Muller, P. Lewis and M. J. Barnsley. 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, 36:1228–1249.

Kitanidis, P. K. 1997. *Introduction to Geostatistics: Applications to Hydrogeology*. New York: Cambridge University Press.

Kittredge, J. 1944. Estimation of the amount of foliage of trees and stands. *Journal of Forestry*, 42:905–908.

Knipling, E. B. 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of the Environment*, 1:155–159.

Kundzewicz, Z. W. 1995. Hydrological uncertainty in perspective. In Z. W. Kundzewicz, editor, *New uncertainty concepts in hydrology and water resources*. Cambridge: Cambridge University Press.

Kupfersberger, H. and G. Bloschl. 1995. Estimating aquifer transmissivities on the value of auxiliary data. *Journal of Hydrology*, 165:95–99.

Kuusk, A. 1998. Monitoring of vegetation parameters on large areas by the inversion of a canopy reflectance model. *International Journal of Remote Sensing*, 19:2893–2905.

Kyriakidis, P. C. and A. G. Journel. 1999. Geostatistical space-time models: A review. *Mathematical Geology*, 31:651–684.

Labovitz, M. L., D. L. Toll and R. E. Kennard. 1982. Preliminary evidence for the influence of physiography and scale upon the autocorrelation function of remotely sensed data. *International Journal of Remote Sensing*, 3:13–30.

Lacaze, B., S. Rambal and T. Winkel. 1994. Identifying spatial patterns of Mediterranean landscapes from geostatistical analysis of remotely-sensed data. *International Journal of Remote Sensing*, 12:559–567.

Lam, N. S.-N. 1990. Description and measurement of Landsat TM images using fractals. *Photogrammetric Engineering and Remote Sensing*, 56:187–195.

Lambert, N. J., J. Ardo, B. N. Rock and J. E. Vogelmann. 1995. Spectral characterization and regression-based classification of forest damage in Norway spruce stands in the Czech Republic using Landsat Thematic Mapper data. *International Journal of Remote Sensing*, 16:1261–1287.

Lance, K. 1993. Bringing technology down to earth. *Journal of Forestry*, 91:17–19.

Larsson, H. 1993. Linear regressions for canopy cover estimation in acacia woodlands using Landsat-TM, Landsat-MSS and SPOT HRV XS data. *International Journal of Remote Sensing*, 14:2129–2136.

Lathrop, R. G. and T. M. Lillesand. 1987. Calibration of Thematic Mapper thermal data for water surface temperature mapping: case study on the Great Lakes. *Remote Sensing of Environment*, 22:297–307.

Lathrop, R. G. and L. L. Pierce. 1991. Ground-based canopy transmittance and satellite remotely sensed measurements for estimation of coniferous canopy structure. *Remote Sensing of Environment*, 36:179–188.

Laurance, W. F., S. G. Laurance, L. V. Ferreira, J. M. R. de Merona, C. Gascon and T. E. Lovejoy. 1997. Biomass collapse in Amazonian forest fragments. *Science*, 278:1117–1118.

Leachtenauer, J. C. 1977. Optical power spectrum analysis: Scale and resolution effects. *Photogrammetric Engineering and Remote Sensing*, 43:1117–1125.

Leblon, B., H. Granberg, C. Ansseau and A. Royer. 1993. A semi-empirical model to estimate the biomass production of forest canopies from spectral variables. Part 1: Relationship between spectral variables and light interception efficiency. *Remote Sensing Reviews*, 7:109–125.

Leenaers, H., J. P. Okx and P. A. Burrough. 1990. Comparison of spatial prediction methods for mapping floodplain soil pollution. *Catena*, 17:535–550.

Legendre, P. 1993. Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74:1659–1673.

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

Lesch, S. M., D. J. Strauss and J. D. Rhoades. 1995. Spatial prediction of soil salinity using electromagnetic induction techniques. 1) Statistical prediction models – a comparison of multiple linear regression and cokriging. *Water Resources Research*, 31:373–386.

Lesieur, M. and K. B. Katsaros. 1981. Satellite-sensed turbulent ocean structure. *Nature*, 294:673.

Li, X. W. and A. H. Strahler. 1992. Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: effect of crown shade and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, 30:276–292.

Li, Z. Q., L. Moreau and J. Cihlar. 1997. Estimation of photosynthetically active radiation absorbed at the surface. *Journal of Geophysical Research – Atmospheres*, 102:29717–29727.

Liang, S. L., H. F. Adl, S. Kalluri, J. JaJa, Y. J. Kaufman and J. R. G. Townshend. 1997. An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. *Journal of Geophysical Research – Atmospheres*, 102:17173–17186.

Lillesand, T. M. and R. W. Kiefer. 1994. *Remote Sensing and Image Interpretation*. New York: John Wiley & Sons, third edition.

Lillesand, T. M., D. E. Meisner, A. L. Downs and R. L. Deuell. 1982. Use of GOES and TIROS/NOAA satellite data for snow cover mapping. *Photogrammetric Engineering and Remote Sensing*, 48:251–259.

Lloyd, C. D., S. Berberoglu, P. J. Curran and P. M. Atkinson. 2000. The use of geostatisical measures of spatial variability and the co-occurrence matrix for land cover mapping in the Mediterranean. *Remote Sensing of Environment*, page in press.

Lohani, B. and D. C. Mason. 1999. Construction of a Digital Elevation Model of the Holderness Coast using the waterline method and Airborne Thematic Mapper Data. *International Journal of Remote Sensing*, 20:593–607.

Lunetta, R. S., R. G. Congalton, L. K. Fenstermaker, J. R. Jensen, K. C. McGwire and L. R. Tinney. 1991. Remote sensing and Geographic Information System data integration – error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, 57:677–687.

Ma, X. and A. G. Journel. 1999. An expanded GSLIB cokriging program allowing for two Markov models. *Computers & Geosciences*, 25:627–639.

Mackay, D. S. and L. E. Band. 1997. Forest ecosystem processes at the watershed scale: Dynamic coupling of distributed hydrology and canopy growth. *Hydrological Processes*, 11:1197–1217.

Maitre, H. and M. Pinciroli. 1999. Fractal characterization of a hydrological basin using SAR satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 37:175–181.

Maling, D. H. 1989. *Measurement from Maps: Principles and Methods of Cartometry*. New York: Pergamon Press.

Marceau, D. J., P. J. Howarth and D. J. Gratton. 1994. Remote sensing and the measurement of geographical entities in a forested environment. 1) The scale and spatial aggregation problem. *Remote Sensing of Environment*, 49:93–104.

Matheron, G. 1971. The Theory of Regionalized Variables and its Applications. In *Cahiers du Centre de Morphologie Mathematique de Fountainbleau*, No 5, page 212. Fountainbleau, France: Ecole de Mines.

Matheron, G. 1985. Change of support for diffusion type random functions. *Mathematical Geology*, 17:137–165.

Matheron, G. 1989. *Estimating and Choosing*. Berlin: Springer-Verlag.

McBratney, A. B. and R. Webster. 1983. How many observations are needed to estimate the regional mean of a soil property? *Soil Science*, 135:177–183.

McBratney, A. B., R. Webster and T. M. Burgess. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. *Computers & Geosciences*, 7:331–334.

McCluney, W. R. 1976. Remote measurement of water color. *Remote Sensing of Environment*, 5:3–33.

McGwire, K., M. Friedl and J. E. Estes. 1993. Spatial structure, sampling design and scale in remotely-sensed imagery of a California savanna woodland. *International Journal of Remote Sensing*, 14:2137–2164.

Merrill, E. H., M. K. Bramblebrodahl, R. W. Marrs and M. S. Boyce. 1993. Estimation of green herbaceous phytomass from Landsat MSS data in Yellowstone National Park. *Journal of Range Management*, 46:151–157.

Milne, B. T. and W. B. Cohen. 1999. Multiscale assessment of binary and continuous land cover variables for MODIS validation, mapping, and modeling applications. *Remote Sensing of Environment*, 70:82–98.

Miranda, F. P., J. A. MacDonald and J. R. Carr. 1994. Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR B data on the Guiana Shield. *Remote Sensing Reviews*, 10:155–168.

Moellering, H. and W. Tobler. 1972. Geographical variances. *Geographic Analysis*, 4:35–50.

Moran, M. S. and R. D. Jackson. 1991. Assessing the spatial distribution of evapotranspiration using remotely sensed inputs. *Journal of Environmental Quality*, 20:725–737.

Moran, M. S., R. D. Jackson, T. R. Clarke, J. Qi, F. Cabot, K. J. Thome and B. L. Markham. 1995. Reflectance factor retrieval from Landsat TM and SPOT HRV data for bright and dark targets. *Remote Sensing of Environment*, 52:218–230.

Moran, M. S., R. D. Jackson, P. N. Slater and P. M. Teillet. 1992. Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. *Remote Sensing of Environment*, 41:169–184.

Moran, P. A. P. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10:243–251.

Morris, M. D. 1991. On counting the number of data pairs for semivariogram estimation. *Mathematical Geology*, 23:929–943.

Mowrer, H. T. 1997. Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modelling*, 98:73–86.

Mueller-Dombois, D. and H. Ellenberg. 1974. *Aims and Methods of Vegetation Ecology*. New York: John Wiley & Sons.

Musick, H. B. 1984. Assessment of Landsat multispectral scanner spectral indices for monitoring arid rangeland. *IEEE Transactions on Geoscience and Remote Sensing*, 22:512–519.

Myers, D. E. 1982. Matrix formulation of cokriging. *Mathematical Geology*, 14:249–257.

Myers, D. E. 1983. Estimation of linear combinations and cokriging. *Mathematical Geology*, 15:633–637.

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

Myers, J. C. 1997. *Geostatistical Error Management: Quantifying Uncertainty for Environmental Sampling and Mapping*. New York: Van Nostrand Reinhold.

Myers, V. I. and W. A. Allen. 1968. Electro-optical remote sensing method as non-destructive testing and measuring techniques in agriculture. *Applied Optics*, 7:1818–1838.

Myneni, R. B., S. Maggion, J. Iaquinta, J. L. Privette, N. Govron, B. Pinty, D. S. Kimes, M. M. Verstraete and D. L. Williams. 1995. Optical remote sensing of vegetation: Modeling, caveats and algorithms. *Remote Sensing of Environment*, 51:169–188.

Myneni, R. B. and D. L. Williams. 1994. On the relationship between FAPAR and NDVI. *Remote Sensing of Environment*, 49:200–211.

Nemani, R., L. L. Pierce, S. W. Running and L. E. Band. 1993. Forest ecosystem processes at the watershed scale – sensitivity to remotely-sensed leaf area index estimates. *International Journal of Remote Sensing*, 14:2519–2534.

Nemani, R. and S. W. Running. 1989. Testing a theoretical climate-soil-leaf area hydrological equilibrium of forests using satellite data and ecosystem simulation. *Agriculture and Forest Meteorology*, 44:245–260.

Nemani, R., S. W. Running, R. A. Pielke and T. N. Chase. 1996. Global vegetation cover changes from coarse resolution satellite data. *Journal of Geophysical Research – Atmospheres*, 101:7157–7162.

Nobel, P. S. 1983. *Biophysical Plant Physiology and Ecology*. San Francisco: W. H. Freeman.

Nusser, S. M. and J. J. Goebel. 1997. The National Resources Inventory: A long-term multi-resource monitoring programme. *Environmental and Ecological Statistics*, 4:181–204.

Olea, R. A. 1990. *Geostatistical Glossary and Multilingual Dictionary*. New York: Oxford University Press.

Olea, R. A. and V. Pawlowsky. 1996. Compensating for estimation smoothing in kriging. *Mathematical Geology*, 28:407–417.

Oliver, M. A. 1989. Geostatistics in physical geography. Part II: Applications. *Transactions of the Institute of British Geographers*, 14:270–286.

Oliver, M. A. and R. Webster. 1989. Geostatistical constrained multivariate classification. In M. Armstrong, editor, *Geostatistics*, pages 383–395. Dordrecht: Kluwer.

Oliver, M. A., R. Webster and J. Gerrard. 1989. Geostatistics in physical geography. Part I: Theory. *Transactions of the Institute of British Geographers*, 14:259–269.

Olsen, A. R. and H. T. Schreuder. 1997. Perspectives on large-scale natural resource surveys when cause-effect is a potential issue. *Environmental and Ecological Statistics*, 4:167–180.

Olsson, H. 1995. Reflectance calibration of Thematic Mapper data for forest change detection. *International Journal of Remote Sensing*, 16:81–96.

Openshaw, S. 1984. *The Modifiable Areal Unit Problem: Concepts and Techniques in Modern Geography*, volume 38. Norwich: GeoBooks.

Openshaw, S. 1992. Learning to live with errors in spatial databases. In M. F. Goodchild and S. Gopal, editors, *The Accuracy of Spatial Databases*, pages 263–277. London: Taylor & Francis.

Ormsby, J. P., B. J. Choudhury and M. Owe. 1987. Vegetation spatial variability and its effect on vegetation indices. *International Journal of Remote Sensing*, 8:1301–1306.

Overton, W. S. and S. V. Stehman. 1993. Properties of designs for sampling continuous spatial resources from a triangular grid. *Communications in Statistics-Theory and Methods*, 22:2641–2660.

Oza, M. P., V. K. Srivastava and P. K. Devaiah. 1992. Estimating the mean canopy diameter of teak plantations from Landsat MSS data. *International Journal of Remote Sensing*, 13:2363–2369.

Pala, V. and X. Pons. 1995. Incorporation of relief into geometric corrections based on polynomials. *Photogrammetric Engineering and Remote Sensing*, 61:935–944.

Pan, Y. D., J. M. Melillo, A. D. McGuire, D. W. Kicklighter, L. F. Pitelka, K. Hibbard, L. L. Pierce, S. W. Running, D. S. Ojima, W. J. Parton and D. S. Schimel. 1998. Modeled responses of terrestrial ecosystems to elevated atmospheric CO$_2$: a comparison of simulations by the biogeochemistry models of the Vegetation/Ecosystem Modeling and Analysis Project (VEMAP). *Oecologia*, 114:389–404.

Parton, W. J., J. W. Chason, D. D. Baldocchi and M. A. Huston. 1991. A comparison of direct and indirect methods for estimating forest canopy leaf area. *Agricultural and Forest Meteorology*, 57:107–128.

Peterson, D. L., M. A. Spanner, S. W. Running and K. B. Teuber. 1987. Relationship of Thematic Mapper simulator data to leaf area index of temperate coniferous forests. *Remote Sensing of Environment*, 22:323–341.

Peterson, D. L., W. E. Westman, N. J. Stephenson, V. G. Ambrosia, M. A. Spanner and J. A. Brass. 1986. Analysis of forest structure using Thematic Mapper Simulator data. *IEEE Transactions on Geoscience and Remote Sensing*, 24:113–121.

Phillips, D. L. and D. G. Marks. 1996. Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. *Ecological Modelling*, 91:213–229.

Phinn, S., J. Franklin, A. Hope, D. A. Stow and L. Huenneke. 1996. Biomass distribution mapping using airborne digital video imagery and spatial statistics in a semi-arid environment. *Journal of Environmental Management*, 47:139–164.

Phinn, S. R., D. A. Stow and D. VanMouwerik. 1999. Remotely sensed estimates of vegetation structural characteristics in restored wetlands, Southern California. *Photogrammetric Engineering and Remote Sensing*, 65:485–493.

Pinty, B., M. M. Verstraete and R. E. Dickinson. 1990. A physical model of the bidirectional reflectance of vegetation canopies: Inversion and validation. *Journal of Geophysical Research – Atmospheres*, 95:11767–11775.

Pollock, R. B. and E. T. Kanemasu. 1979. Estimating leaf-area index of wheat with LANDSAT data. *Remote Sensing of Environment*, 8:307–312.

Poole, M. A. and P. N. O'Farrell. 1971. The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 52:145–158.

Price, J. C. 1987. Calibration of satellite radiometers and the comparison of vegetation indices. *Remote Sensing of Environment*, 21:15–27.

Price, J. C. 1992. Estimating vegetation amount from visible and near infrared reflectances. *Remote Sensing of Environment*, 41:29–34.

Prince, S. D., Y. H. Kerr, J. P. Goutorbe, T. Lebel, A. Tinga, P. Bessemoulin, J. Brouwer, A. J. Dolman, E. J. Engman, J. H. C. Gash, M. Hoepffner, P. Kabat, B. Monteny, F. Said, P. Sellers and J. Wallace. 1995. Geographical, biological and remote-sensing aspects of the Hydrologic Atmospheric Pilot Experiment in the Sahel (HAPEX-Sahel). *Remote Sensing of Environment*, 51:215–234.

Quattrochi, D. A. 1993. The need for a lexicon of scale terms in integrating remote sensing data with geographic information systems. *Journal of Geography*, 92:206–212.

Raffy, M. 1994. Change of scale theory – a capital challenge for space observation of earth – Introduction. *International Journal of Remote Sensing*, 15:2353–2357.

Raffy, M. and C. Gregoire. 1998. Semi-empirical models and scaling: a least square method for remote sensing experiments. *International Journal of Remote Sensing*, 13:2527–2541.

Ramstein, G. and M. Raffy. 1989. Analysis of the structure of radiometric remotely-sensed images. *International Journal of Remote Sensing*, 10:1049–1073.

Rasmussen, M. S. 1998. Developing simple, operational, consistent NDVI-vegetation models by applying environmental and climatic information. Part II: Crop yield assessment. *International Journal of Remote Sensing*, 19:119–139.

Rendu, J.-M. 1981. *An Introduction to Geostatistical Methods of Mineral Evaluation.* Johannesburg: South African Institute of Mining and Metallurgy.

Richardson, A. J. and C. L. Wiegand. 1977. Distinguishing vegetation from soil background. *Photogrammetric Engineering and Remote Sensing*, 43:1541–1552.

Ripple, W. J. 1994. Determining coniferous forest cover and forest fragmentation with NOAA 9 AVHRR data. *Photogrammetric Engineering and Remote Sensing*, 60:533–540.

Ripple, W. J., S. Wang, D. L. Isaacson and D. P. Paine. 1991. A preliminary comparison of Landsat Thematic Mapper and SPOT-1 HRV multispectral data for estimating coniferous forest volume. *International Journal of Remote Sensing*, 12:387–400.

Robertson, G. P. 1987. Geostatistics in ecology: interpolating with known variance. *Ecology*, 68:744–748.

Rollin, E. M. and E. J. Milton. 1998. Processing of high spectral resolution reflectance data for the retrieval of canopy water content information. *Remote Sensing of Environment*, 65:86–92.

Rosema, A., W. Verhoef, H. Noorbergen and J. J. Borgesius. 1992. A new forest light interaction model in support of forest monitoring. *Remote Sensing of Environment*, 42:23–41.

Rosenthal, W. and J. Dozier. 1996. Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper. *Water Resources Research*, 32:115–130.

Rossi, R. E., J. L. Dungan and L. R. Beck. 1994. Kriging in the shadows: Geostatistical interpolation for remote sensing. *Remote Sensing of Environment*, 49:32–40.

Rossi, R. E., D. J. Mulla, A. G. Journel and E. H. Franz. 1992. Geostatistical tools for the modeling and interpretation of ecological spatial dependence. *Ecological Monographs*, 62:277–314.

Rouse, J. W., Jr., R. H. Haas, J. A. Schell and D. W. Deering. 1975. Monitoring vegetation systems in the Great Plains with ERTS. In *Proceedings of the Third Symposium on Significant Results Obtained with ERTS-1*, NASA SP-351, pages 309–317, Washington, DC. US Government Printing Office.

Running, S. W. 1990. Estimating terrestrial primary productivity by combining remote sensing and ecosystem simulation. In R. Hobbs and H. Mooney, editors, *Remote Sensing of Biosphere Functioning*, pages 65–86. New York: Springer-Verlag.

Running, S. W., R. R. Nemani, D. L. Peterson, L. E. Band and M. A. Spanner. 1989. Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. *Ecology*, 70:1090–1101 273–276.

Running, S. W., D. L. Peterson, M. A. Spanner and K. B. Teuber. 1986. Remote sensing of coniferous forest leaf area. *Ecology*, 67.

Russo, D. 1984. Design of an optimal sampling network for estimating the variogram. *Soil Science Society of America Journal*, 48:708–716.

Salvador, R. and A. Pons. 1998. On the reliability of Landsat TM for estimating forest variables by regression techniques: A methodological analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 36:1888–1897.

Schachter, B. 1980. Model-based texture measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:169–171.

Schowengerdt, R. A. 1997. *Remote Sensing: Models, and Methods for Image Processing*. San Diego: Academic Press, second edition.

Schultz, J. 1995. *The Ecozones of the World*. Berlin: Springer-Verlag.

Seaber, P. R., F. P. Kapinos and G. L. Knapp. 1987. *Hydrologic Unit Maps*. US Geological Survey Water-Supply Paper 2294. Denver, CO: US Government Printing Office.

Seeley, H. R. 1929. Computing tree heights from shadows on aerial photographs. *Forestry Chronicles*, 5:24–27.

Sellers, P. J. 1987. Canopy reflectance, photosynthesis and transpiration. II. The role of biophysics in the linearity of their interdependence. *Remote Sensing of Environment*, 21:143–183.

Sellers, P. J., D. E. Dickinson, D. A. Randal, A. K. Betts, F. G. Hall, J. A. Berry, G. J. Collatz, A. S. Denning, H. A. Mooney, C. A. Nobre, N. Sato, C. B. Field and A. Henderson-Sellers. 1997. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, 275:502–509.

Sellers, P. J., F. G. Hall, G. Asrar, D. E. Strebel and R. E. Murphy. 1992. An overview of the 1st International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE). *Journal of Geophysical Research – Atmospheres*, 97:18345–18371.

Sellers, P. J., Y. Mintz, Y. C. Sud and A. Dalcher. 1986. A Simple Biosphere model (SiB) for use with general circulation models. *Journal of Atmospheric Sciences*, 43:505–531.

Seo, D. J., W. F. Krajewski and D. S. Bowles. 1990. Stochastic interpolation of rainfall data from rain gages and radar using cokriging. 1. Design of experiments. *Water Resources Research*, 26:469–477.

Shaw, G. and D. Wheeler. 1994. *Statistical Techniques in Geographical Analysis*. London: David Fulton.

Shippert, M. M., D. A. Walker and B. E. Lewis. 1995. Biomass and leaf-area index maps derived from SPOT images for Toolik Lake and Imnavait Creek areas, Alaska. *Polar Record*, 31:147–157.

Shukla, G. K. 1972. On the problem of calibration. *Technometrics*, 14:547–553.

Simonett, D. 1971. Editorial. *Remote Sensing of Environment*, 2:iii–iv.

Simpson, D., A. Guenther, C. N. Hewitt and R. Steinbrecher. 1995. Biogenic emissions in Europe. 1. Estimates and uncertainties. *Journal of Geophysical Research – Atmospheres*, 100:22875–22890.

Smith, G. M. and P. J. Curran. 1996. The signal-to-noise ratio (SNR) required for the estimation of foliar biochemical concentrations. *International Journal of Remote Sensing*, 17:1031–1058.

Solow, A. R. 1990. Geostatistical cross-validation – a cautionary note. *Mathematical Geology*, 22:637–639.

Spanner, M. A., R. C. Wrigley, R. F. Pueschel, J. M. Livingston and D. S. Colburn. 1990. Determination of atmospheric optical properties during the First International Satellite Land Surface Climatology Project Field Experiment. *Journal of Spacecraft and Rockets*, pages 373–379.

Srivastava, R. M. 1987. Minimum variance or maximum profitability? *CIM Bulletin*, 80:63–68.

Srivastava, R. M. 1992. Reservoir characterization with probability field simulation. In *Annual Technical Conference of the Society of Petroleum Engineers*, SPE 24753, pages 927–938, Washington, DC. Society of Petroleum Engineers.

St-Onge, B. A. 1999. Topographic effects on the texture of high-resolution forest-stand images measured by the semivariogram. *Photogrammetric Engineering and Remote Sensing*, 65:923–935.

Stehman, S. V. 1992. Comparison of systematic and random sampling for estimating the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 58:1343–1350.

Stehman, S. V. 1999. Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, 20:2423–2441.

Stehman, S. V. and R. L. Czaplewski. 1998. Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sensing of Environment*, 64:331–344.

Stein, A., W. G. M. Bastiaanssen, S. Bruin, A. P. Cracknell, P. J. Curran, A. G. Fabbri, B. G. H. Gorte, J. W. van Groenigen, F. van der Meer and A. Saldona. 1998. Integrating spatial statistics and remote sensing. *International Journal of Remote Sensing*, 19:1793–1814.

Steven, M. D. 1987. Ground truth: An underview. *International Journal of Remote Sensing*, 8:1033–1038.

Stewart, G. H. 1988. The influence of canopy cover on understorey development in forests of the Western Cascade Range, Oregon, USA. *Vegetatio*, 76:79–88.

Stohlgren, T. J., T. N. Chase, R. A. Pielke, T. G. F. Kittel and J. S. Baron. 1998. Evidence that local land use practices influence regional climate, vegetation, and stream flow patterns in adjacent natural areas. *Global Change Biology*, 5:495–504.

Switzer, P. 1983. Some spatial statistics from the interpretation of satellite data. *Bulletin of the International Statistics Institute*, 50:962–972.

Switzer, P. 1993. The spatial variability of prediction errors. In A. Soares, editor, *Geostatistics Troia '92*, pages 261–272. Dordrecht: Kluwer.

Taylor, B. N. and C. E. Kuyatt. 1994. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. National Institute of Standards and Technology (NIST) Technical Note 1297. Washington, DC: US Government Printing Office.

Teillet, P. M. and B. N. Holben. 1993. Towards operational radiometric calibration of NOAA AVHRR imagery in the visible and infrared channels. *Canadian Journal of Remote Sensing*, 20:1–10.

Teillet, P. M., K. Staenz and D. J. Williams. 1997. Effects of spectral, spatial, and radiometric characteristics of remote sensing vegetation indices of forested regions. *Remote Sensing of Environment*, 61:139–149.

Teklehaimanot, Z. and P. G. Jarvis. 1991. Modelling of rainfall interception loss in agroforestry systems. *Agroforestry Systems*, 14:65–80.

Thomas, J. R. and A. H. Gerbermann. 1977. Yield-reflectance relations in cabbage. *Photogrammetric Engineering and Remote Sensing*, 43:1257–1266.

Thomas, J. R., C. L. Wiegand and V. I. Myers. 1967. Reflectance of cotton yields and its relation to yields. *Agronomy Journal*, 69:551–554.

Thome, K., B. Markham, J. Barker, P. Slater and S. Biggar. 1997. Radiometric calibration of Landsat. *Photogrammetric Engineering and Remote Sensing*, 63:853–858.

Tiefelsdorf, M. and B. Boots. 1995. The exact distribution of Moran's I. *Environment and Planning A*, 27:985–999.

Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46:234–240.

Todd, S. W., R. M. Hoffer and D. G. Milchunas. 1998. Biomass estimation on grazed and ungrazed rangelands using spectral indices. *International Journal of Remote Sensing*, 19:427–438.

Townshend, J. R. G. and C. O. Justice. 1990. The spatial variation of vegetation changes at very coarse scales. *International Journal of Remote Sensing*, 11:149–157.

Townshend, J. R. G. and C. O. Justice. 1995. Spatial variability of images and the monitoring of changes in the Normalized Difference Vegetation Index. *International Journal of Remote Sensing*, 16:2187–2195.

Treitz, P. M. and P. J. Howarth. 1999. Hyperspectral remote sensing for estimating biophysical parameters for forest ecosystems. *Progress in Physical Geography*, 23:349–390.

Tucker, C. J. 1977. Spectral estimation of grass canopy variables. *Remote Sensing of Environment*, 6:11–26.

Tucker, C. J., C. Vanpraet, E. Boerwinkel and A. Gaston. 1983. Satellite remote sensing of total dry matter production in the Senegalese Sahel. *Remote Sensing of Environment*, 13:461–474.

Turner, D. P., W. B. Cohen, R. E. Kennedy, K. S. Fassnacht and J. M. Briggs. 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, 70:52–68.

Turner, M. G. and R. H. Gardner. 1991. Quantitative methods in landscape ecology: An introduction. In M. G. Turner and R. H. Gardner, editors, *Quantitative Methods in Landscape Ecology*, Ecological Studies 82, pages 1–13. New York: Springer-Verlag.

Unwin, D. J. 1995. Geographical information systems and the problem of 'error and uncertainty'. *Progress in Human Geography*, 19:549–558.

Unwin, D. J. 1996. GIS, spatial analysis and spatial statistics. *Progress in Human Geography*, 20:540–551.

US Geological Survey. 1977. *EROS Digital Image Processing System – An Overview*. Sioux Falls, SD: US Geological Survey.

Ustin, S. L., C. A. Wessman, B. Curtiss, E. Kasischke, J. Way and V. C. Vanderbilt. 1991. Opportunities for using the EOS imaging spectrometers and synthetic aperture radar in ecological models. *Ecology*, 72:1934–1945.

van der Meer, F. 1994. Mapping the degree of serpentinization within ultramafic rock bodies using imaging spectrometer data. *International Journal of Remote Sensing*, 15:3851–3857.

van der Meer, F. 1996. Classification of remotely-sensed imagery using an indicator kriging approach – application to the problem of calcite-dolomite mineral mapping. *International Journal of Remote Sensing*, 17:1233–1249.

Vane, G., R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen and W. M. Porter. 1993. The airborne visible infrared imaging spectrometer (AVIRIS). *Remote Sensing of Environment*, 44:127–143.

VanMarcke, E. 1983. *Random Fields: Analysis and Synthesis*. Cambridge, MA: MIT Press.

Vermote, E. F., D. Tanre, J. L. Deuze, M. Herman and J. J. Morcette. 1997. Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 35:675–686.

Verstraete, M. M. and B. Pinty. 1996. Designing optimal spectral indexes for remote sensing applications. *IEEE Transactions on Geoscience and Remote Sensing*, 34:1254–1265.

Verstraete, M. M., B. Pinty and P. J. Curran. 1999. MERIS potential for land applications. *International Journal of Remote Sensing*, 20:1747–1756.

Verstraete, M. M., B. Pinty and R. B. Myneni. 1996. Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing. *Remote Sensing of Environment*, 58:201–214.

Vignesadler, M., A. Lepage and P. M. Adler. 1991. Fractal analysis of fracturing in African regions, from satellite imagery to ground scale. *Tectonophysics*, 196:69–86.

Vogelmann, J. E., T. L. Sohl, P. V. Campbell and D. M. Shaw. 1998. Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. *Environmental Monitoring and Assessment*, 51:415–428.

Vose, J. M., P. M. Dougherty, J. N. Long, F. W. Smith, H. L. Gholz and P. J. Curran. 1995. Factors influencing the amount and distribution of leaf area in pine stands. *Ecological Bulletins*, 43:102–114.

Vujakovic, P. 1987. Monitoring extensive 'buffer zones' in Africa: An application for satellite imagery. *Biological Conservation*, 39:195–208.

Wald, L. 1989. Some examples of the use of structure functions in the analysis of satellite images of ocean. *Photogrammetric Engineering and Remote Sensing*, 55:1487–1490.

Wardley, N. W. and P. J. Curran. 1984. The estimation of green-leaf-area index from remotely sensed airborne multispectral scanner data. *International Journal of Remote Sensing*, 5:671–679.

Watson, D. J. 1947. Comparative physiological studies on the growth of field crops. I. Variation in net assimilation rate and leaf area between species and varieties and within and between years. *Annals of Botany*, 11:41–76.

Webster, R. and T. M. Burgess. 1984. Sampling and bulking strategies for estimating soil properties in small regions. *Journal of Soil Science*, 35:127–140.

Webster, R., P. J. Curran and J. W. Munden. 1989. Spatial correlation in reflected radiation from the ground and its implications for sampling and mapping by ground-based radiometry. *Remote Sensing of Environment*, 29:67–78.

Webster, R. and M. A. Oliver. 1990. *Statistical Methods in Soil and Land Resource Survey*. New York: Oxford University Press.

Webster, R. and M. A. Oliver. 1992. Sample adequately to estimate variograms of soil properties. *Journal of Soil Sciences*, 43:177–192.

Wessman, C. A., J. D. Aber and D. L. Peterson. 1989. An evaluation of imaging spectrometry for estimating forest canopy chemistry. *International Journal of Remote Sensing*, 10:1293–1316.

Weszka, J. S., C. R. Dyer and A. Rosenfeld. 1976. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man and Cybernetics*, 6:269–285.

White, J. D., S. W. Running, R. Nemani, R. E. Keane and K. C. Ryan. 1997*a*. Measurement and remote sensing of LAI in Rocky Mountain montane ecosystems. *Canadian Journal of Forest Research*, 27:1714–1727.

White, K., J. Walden, N. Drake, E. Eckardt and J. Settle. 1997*b*. Mapping the iron oxide content of dune sands, Namib sand sea, Namibia, using Landsat Thematic Mapper data. *Remote Sensing of Environment*, 62:30–39.

Wickham, J. D. and K. H. Riitters. 1995. Sensitivity of landscape metrics to pixel size. *International Journal of Remote Sensing*, 16:3585–3594.

Wiegand, C. L., A. J. Richardson and E. T. Kanemasu. 1979. Leaf area index estimates for wheat from LANDSAT and their implications for evapotranspiration and crop modelling. *Agronomy Journal*, 71:336–342.

Woodcock, C. E., J. B. Collins, V. D. Jakabhazy, X. W. Li, S. A. Macomber and Y. C. Wu. 1997. Inversion of the Li-Strahler canopy reflectance model for mapping forest structure. *IEEE Transactions on Geoscience and Remote Sensing*, 35:405–414.

Woodcock, C. E., A. H. Strahler and D. L. B. Jupp. 1988*a*. The use of variograms in remote sensing: I. Scene models and simulated images. *Remote Sensing of Environment*, 25:324–348.

Woodcock, C. E., A. H. Strahler and D. L. B. Jupp. 1988*b*. The use of variograms in remote sensing: II. Real digital images. *Remote Sensing of Environment*, 25:349–379.

Wrigley, R. C., M. A. Spanner, R. E. Slye, R. F. Pueschel and H. R. Aggarwal. 1992. Atmospheric correction of remotely sensed image data by a simplified model. *Journal of Geophysical Research*, 97:18797–18814.

Wulder, M. 1998. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. *Progress in Physical Geography*, 22:449–476.

Xue, Y., P. J. Sellers, J. L. Kinter and J. Shukla. 1991. A simplified biosphere model for global climate studies. *Journal of Climate*, 4:345–364.

Yates, S. R. and A. W. Warrick. 1987. Estimating soil water content using cokriging. *Soil Science Society of America Journal*, 51:23–30.

Zhang, L., A. L. O'Neill and S. Lacey. 1996. Modelling approaches to the prediction of soil erosion in catchments. *Environmental Software*, 11:123–133.

Zhang, R., A. W. Warrick and D. E. Myers. 1990. Variance as a function of sample support size. *Mathematical Geology*, 22:107–121.

Zimmerman, D. A., G. deMarsily, C. A. Gotway, M. G. Marietta, C. L. Axness, R. L. Beauheim, R. L. Bras, J. Carrera, G. Dagan, P. B. Davies, D. P. Gallegos, A. Galli, J. Gomez-Hernandez, P. Grindrod, A. L. Gutjahr, P. K. Kitanidis, A. M. Lavenue, D. McLaughlin, S. P. Neuman, B. S. RamaRao, C. Ravenne and Y. Rubin. 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resources Research*, 34:1373–1413.