University of Nairobi

College of Architecture & Engineering

School of Engineering

SPATIAL PREDICTION OF SOIL PHYSICAL PROPERTIES USING INFRARED SPECTROSCOPY AND GIS

Achuti Mochama

B.Tech (Chemical & Process Engineering, Moi University)

Msc Student Reg No. F56/7231/2006

A thesis submitted in partial fulfillment for the degree of **Master of Science in Environmental and Biosystems Engineering** in the **Department of Environmental and Biosystems Engineering** in the **University of Nairobi**

June, 2011

Declaration

This thesis is my original work and has not been submitted for a degree in any other university.

.... 19/7/2011

Achuti Mochama Date

This thesis has been submitted for examination with our approval as university Supervisors

.... 19/7/2011

Prof. Elijah K. Biamah Date

.... 3rd/8/2011

Dr. Christian T. Omuto Date

Dedication

This study is dedicated to:

……… to the most important women in my life……

My dear mother,

Mrs. Teresia Mong'are Okechi for her effort in raising me up to this level.

Mum, may you live to see yourself a blessed mother among women.

My dear wife,

Edinah Naliaka Wepukhulu and my daughters

Dorothy, Jeslyn and Angeline for their patience and support

My Mentor

Mrs. Cecilia Irina, the Principal Academic Officer, Moi University

for her dear support to my professional and social development

………to the inspiring men towards my study……

My dear supervisors,

Dr. Christian Omuto, Lecturer, University of Nairobi

for his continuous professional support, mentoring and encouragement.

Prof. Elijah K. Biamah, Chairman, Department of Environmental & Biosystems Engineering, University of Nairobi

for his unceasing advice, professional help and encouragement

………to the support family towards my study……

Mr. Wilfred Mochama, Chief South Mugirango, his wife Drusila and family for their dedicated will and support towards my educational development.

Acknowledgements

I highly appreciate the technical guidance and inspiring support offered to me by my supervisors: Dr. Christian Omuto and Professor Elijah K. Biamah. They did guide me through entire study, notably thesis development, programming and literature search. They have been my mentors in this scholarly accomplishment.

I do extend gratitude to Dr. Keith Shepherd of ICRAF and his entire staff. They granted me an exposure to soil analysis instrumentation and access to a resourceful spectral library.

I recognize the foundations installed in me by my lecturers, especially Dr. Ayub Gitau, Mr. Ondicki and fellow students at the University of Nairobi during my coursework, their continual encouragement and technical guide.

Much appreciation to Mary Nelima of International Centre for Insect Physiology and Ecology (ICIPE) for reading my thesis and Mr. Bonface Muriro for his support in soil analysis.

I extent my appreciation to Mr. Dennis Makori of OnfonMedia for his financial support, Sophie Ondiaka of ICIPE for her encouragement and dedicated follow-ups towards accomplishment of the study.

Abstract

Soil physical properties include soil consistency limits and soil aggregate stability indices. These properties have influence on the soil agronomic characteristics as well as management of watershed hydrology, civil constructions and modeling environmental quality. The current methods for their determination are expensive and time consuming. There is a need for rapid and fairly accurate methods that can guarantee speed and allow comparison of point-measurements over time and space.

This study developed a new approach for predicting soil physical properties using infrared spectroscopy and GIS. The study involved measurement of the soil physical properties using near infrared spectroscopy and mechanical methods. The calibration of the spectral values with the mechanical values developed a calibration model coefficient of determination of 0.78, 0.81 and 0.88 for aggregate stability index, liquid limit and plastic limit respectively. Spectral reflectance was found to be a reliable surrogate predictor of the soil physical properties.

A prediction model was established for estimation of the soil physical properties using Kriging method and implemented in R software (R Development Core Team, 2008, www.cran.r-project.org). The findings give a visual and numerical prediction of the soil physical properties in the upper Athi river watershed, eastern Kenya. A map of the physical properties of soil was produced and is expected to provide useful insights for planning civil constructions, control of land degradation, and for environmental conservation in upper Athi river watershed.

This study should be tested with high resolution data sources and other related models which can improve the accuracy of the input data as well as prediction of the soil physical properties. Further testing and worldwide application with different models in different soil types and watersheds is highly recommended.

Table of Contents

List of tables

List of figures

List of abbreviations and symbols

ASTM	American Society for Testing and Materials
EBE	Environmental & Biosystems Engineering
cc	Compression index
C_1	Empirical coefficient varying from about 0.9 to 1.2
DBMS	Database Management Systems
DETRL	Department of Environment, Transport & Research Laboratory
D_{10}	Effective particle size or the 10% finer size expressed in mm
e	Void ratio
FTIR	Fourier-Transform Infrared
GIS	Geographical Information Systems
GPS	Global Positioning System
ICIPE	International Centre for Insect Physiology & Ecology
ICRAF	International Centre for Research in Agroforestry
IR	Infrared
K	Permeability
KARI	Kenya Agricultural Research Institute
Labid	Laboratory Identification Number
LL	Liquid limit
MPA	Multi Purpose Analyzer
MRE	Mean Reducing Error
NIR	Near Infrared
NIRS	Near Infrared Spectroscopy
OK	Ordinary Krigging
PI	Plasticity Index
PL	Plastic Limit
PLS	Partial Least Squares
PLSR	Partial Least Squares Regression
r^2	Coefficient of Determination
RMSE	Root Mean Square Error
SL	Shrinkage Limit
SPC	State Plane Coordinates System
UPS	Universal Polar Stereographic
US	United States
UTM	Universal Transverse Mercator

1. INTRODUCTION

1.1 Background

There is a huge challenge in agricultural and infrastructural development due to lack of sufficient data on soil physical properties (Shepherd and Walsh, 2007).

Soil physical properties include soil texture, soil structure and structural form. An understanding of interactions of these properties helps in agriculture and engineering applications of soil. They influence soil physical condition at various moisture contents and consequently the behavior of soil to mechanical stress.

Soil consistency is important to the engineer when making decisions on prudent use of the soil material. It provides a means for describing the degree and kind of cohesion and adhesion between the soil particles as related to the resistance of the soil to deform or rupture. Consistency largely depends on soil minerals and the water content (Sowers and George, 1961). Consistency gives information on soil physical properties such as rupture resistance, stickiness, plasticity and geophysical properties. The knowledge on soil physical properties is important in the management of watershed hydrology, civil constructions, and modeling environmental quality (Buol et al., 1980; Harpstead et al., 1997).

Soil physical properties have traditionally been determined by various methods such as Atterberg limit tests and aggregate stability (DETRL, 1974; Liu et al., 1984; Singh and Chowdhary, 1990). Although these methods are associated with high precision, they have been found to be labour intensive (Nanni and Dematte, 2006; Shepherd and Walsh, 2007). In order to capture the soil physical properties at landscape level, there is need for the development of a rapid technique with comparative accuracy. The recent developments in soil spectroscopy seem to offer a ray of hope towards this goal (McBratney et al., 2003).

Infrared spectroscopy is one of the methods used to measure soil properties. It has evolved with tremendous results on the development of pedotransfer functions. It provides a rapid soil assessment method that is simple, cost effective, non-destructive and capable of analyzing a wide range of materials. When combined with GIS, spectroscopy can facilitate estimation of many soil properties including those of physical characteristics (Janik et al., 2007).

Soil spectral reflectance is a cumulative property which is derived from the inherent spectral behavior of combinations of minerals, organic matter and soil water. Soil mineralogy is interrelated with soil texture, organic matter content and water absorption features. Qualitative and quantitative information on soil properties can be extracted from the analysis of the soil reflectance spectra. Qualitative information can be extracted from visual comparison of a reflectance spectrum of soil of unknown composition with spectra of soil of known properties. Quantitative information can be extracted through the implementation of numerical models which relate specific soil properties to soil spectral behavior (Calabro et al., 2010)

This study tested and confirmed the applicability of infrared spectroscopy and GIS in rapid assessment of physical properties of soil through the development of predictive models.

1.2 Problem Statement and Justification

The knowledge of soil physical properties is important for appropriate management of civil constructions and agronomical applications. In civil constructions, soil strength, stress deformation behaviour and fluid flow properties are of primary management concerns while in agronomic applications, the assessment of land degradation, environmental quality, prediction of the effects of green house gases and estimation of plant yield rely on knowledge of soil physical properties (Lal and Pierce, 1989; Harpstead et al., 1997).

Soil physical properties have been traditionally determined using Atterberg limits and aggregate stability. However, these methods are expensive and time consuming. In land management, decisions are made on either large-scale such as watershed level or micro-level within small land parcels. Thus, important information on soil properties is needed to support such decisions. This implies that useful knowledge of the physical properties of soil need to be determined in conformity to the scale of their demand.

The conventional methods for the determination of soil physical properties suffer huge challenges. They have limitations in terms of cost, time, and difficulty in estimation. There is need for a research to support the development of rapid, accurate, and efficient methods to augment the benefits of conventional methods for the determination of soil physical properties.

Recently, soil spectral reflectance method has been linked to the provision of information on various soil constituents with help of detectors and calibration techniques (Shepherd and Walsh, 2001). This method is rapid, relatively inexpensive and portable. It provides routine soil analysis, decision support, soil property classification, soil survey, precision agriculture, diagnosis of soil problems, contaminated site characterization and input data for models. Soil spectral reflectance seem to offer promise for the development of robust and accurate calibration models that can be useful in extrapolating the advantages of the conventional methods for the determination of soil physical properties.

1.3 Objectives

Overall Objective

The overall objective of the study was to develop a procedure for spatial prediction of physical properties of soil.

Specific Objectives

The specific objectives of this study were:

I. To develop calibration models between soil physical properties and soil infrared spectral reflectance

II. To develop a procedure for spatial prediction of physical properties of soils.

2. LITERATURE REVIEW

2.1 Physical Properties of Soil

Soil as a material needs to be studied and handled in the engineering manner design, construction and maintenance of infrastructure facilities. The design of foundations and settlement behaviour of the finished structure depend on the character of the underlying soil and on its action under the stresses imposed by the foundation. For underground and partially embedded structures, soil is important as a material upon which the structures are founded. It is also a major source of the loads to which they are subjected in service and which they must be designed to carry (Spancler, 1951).

All projects associated with soils must have data on engineering performance of soils. Consideration of the soils encountered plays a dominant role in each phase of the execution of any engineering project. In recent years locations of many airports, highways, earth dams and bridges have been established largely on the basis of the study of soils at alternate sites (Lambe, 1979; Singh and Chowdhary, 1990;). An important example of this modern procedure is the Washington National Airport at Gravelly point whose location, design and construction methods were established by the US corps of engineers only after an extensive soil survey (www.tc.faa.gov).

Describing spatial variability enables users of land resources to make better predictions of soil behaviour and performance (Nielsen and Wendroth, 2003). Unlike other engineering materials, soil physical properties may vary widely from place to place within a confine of a single engineering project. In most cases, the properties cannot be altered to the desired standard. Generally, soil must be used in the locality and in the original condition. An engineer's task is to see that the properties of the placed material correspond to those assumed in the design, or to change the design during construction to allow for any difference between the properties of the consummate fill and those employed in the design (Lambe, 1979).

A large part of soil engineering practice, thus, must be devoted to the location of the various types of soil encountered on a project, the determination of their physical properties, the correlation of those properties with the engineering requirements of the job, and the selection of the best available soil for use in the project. This involves precise and intensive investment in conventional soil survey, testing and location classification for discrete projects. However, with the advent of new soil analytical methods such as infrared spectroscopy and geographical information systems, location of various soils and time taken for determination of their properties can be tremendously improved.

Since landscape and soil vary greatly, land use and management must be steered towards fitting the soil's specific properties (Larlson and Robert, 1989). The widespread problem of soil degradation, declining environmental quality, greenhouse effect, need to preserve soil resources and the making of decisions on engineering uses of soils rely on the quantified knowledge of the properties of soils (Harpstead et al., 1997; Lal and Pierce, 1989).

Larlson and Robert (1989) have also noted that soil spatial variability within fields has been widely demonstrated by soil test results and crop yield differences. They have further noted that soil managed according to type of prevailing soil properties in a given geographical location provide higher yields than those managed uniformly over a large area without consideration of variability of soil physical properties.

The lack of specificity concerning soil variation and response to the specific needs of pedon units leads to unwise land use and land management. It may result into soil erosion, over or under application of chemicals, undue leaching of chemicals, undesirable tillage and low production on some plot units.

With modern soil survey, it is possible to assess management units (fields) to improve land use without losing efficiency in field operation. Computer-assisted landscape information is needed for best use of range land. There is a continued need for better soil maps that could aid rapid recalls, better soil-scape unit description and development of databases. When the information is applied to plot units as they occur in the field, the current concerns over cost of map production will be eased considerably (Larlson and Robert, 1989).

Soil classification may be based on soil's origin, mineralogy, grain size and its plasticity index. Adequate classification must group together soils which have characteristics that imply similar physical properties. Physical properties predict the effect of soil on an earth retaining system.

The following soil properties are significant. They form the basis of a complete soil description. They include: Shear strength, density, compressibility, permeability, colour, moisture content and composition.

Soil Colour

A change in colour encountered during excavation often means a different soil stratum with different properties. Colour is described with the aid of the Munsel Colour Charts. For approximate soil classification, colour may be used in identifying soil strata depths.

Soil Compressibility & Settlement

Soils subjected to increased load decrease in volume. The decrease in volume results into surface settlements. If the soil supports a structure and the settlement is large enough, damage may result. But if the soil compressibility and the loads to be imposed are known these settlements can be estimated. Good estimates aid in construction planning where settlements created by construction activities are potentially damaging (Schroeder, 1991). However, this requires in-situ measurements that may be inapplicable during design phase of the project.

The compression of the soil decreases with increasing stress. A major component in compression of cohesionless soil is the bending and distortion of the grains. Liu and Evett (1984) have developed an expression for the compressibility of clay soil (equation 2.1).

$$cc = 0.009(LL - 10). \qquad [2.1]$$

where cc is compression index and LL is liquid limit.

For soils with very low plasticity and porous rock, Sowers and George (1961) have found the compression index to be related to the undisturbed void ratio (equation 2.2).

$$cc = 0.75(e - a). \qquad [2.2]$$

where a is a constant whose value ranges between 0.2 for porous rock to 0.8 for highly micaceas soils and e is the void ratio. Soil compressibility is determined from direct laboratory tests or estimated from liquid limit as shown in table 2.1.

Table 2.1: Correlation of compressibility with liquid limits (Sowers and George, 1961).

Term	Compressibility Index	Liquid limit
Slightly/low compressibility	0-0.19	0-30
Moderate or intermediate	0.20-0.39	31-50
High compressibility	0.40 and over	51 and over

Soil Strength

The strength of a soil is a factor in the design of retaining walls, embankments and foundations. Shearing strength enables the soil to maintain equilibrium and sloping surface. It materially influences the bearing capacity of a foundation soil and lateral pressure which a soil backfill exert against a retaining wall and bulk head (Spancler, 1951; Sowers and George, 1961).

Shearing strength varies with varying soil conditions such as density, moisture content and degree of consolidation. If the shearing stress of a body of soil exceeds a certain design value the soil fails (Terzaghi and Peck, 1962).

Permeability

Permeability or hydraulic conductivity is the property of soil which permits appreciable movement of water through it when saturated and activated by hydrostatic pressure. This is facilitated by presence of continuous voids in the soil (Terzaghi and Peck, 1962; Singh and Chowdhary, 1990; Schroeder, 1991). Permeability depends on the property of both flowing water and the soil. These include: density and viscosity of water; void ratio, size, shape and arrangement of soil particles; degree of saturation; adsorption complex and clay water interaction.

Permeability can be determined in three ways: laboratory tests, field tests and empirical formulae. A widely used field test is the well pumping test. Some of the empirical equations used include (Singh and Chowdhary, 1990):

$$K = C_1(D_{10})_2 \qquad [2.3]$$

where K is permeability, D_{10} is the effective particle size or the 10% finer size expressed in mm and C_1 is an empirical coefficient varying from about 0.9 to 1.2.

Another property of soil related to permeability is the seepage. This is the percolation or slow movement of water through soil or rock. Field measurements are particularly appropriate in connection with investigations of seepage into excavation or foundation. The knowledge of permeability characteristics of soil is required in the computation of seepage through earth structures. It is also used in the estimation of pumpage-capacity requirements for unwatering cofferdams or excavations below the water table as well as determination the settlement rate.

Soil Consistency and Plasticity

The consistency of soils and therefore, their behaviour is greatly affected by the moisture condition. Moisture condition of soil reveals important characteristics of soil that has been correlated with physical behaviour of soil. Consistency describes the degree of coherence between particles of soil at any given moisture content. It is the resistance to deformation of a soil and is related to the force of attraction between particles or aggregates of particles.

Plasticity is the ability of a soil to deform continuously under applied stress and to retain the new shape on removal of stress.

Consistency and plasticity are dependent on moisture content of soil. The consistency limits have been used in classification of soils as well as indication of physical properties of soil (Liu and Evett, 1984).

There are four states of consistency: liquid, plastic, semi solid and solid. The boundaries between adjacent states are termed consistency limits or Atterberg limits. These are: liquid limit (LL), at which soil changes from liquid to plastic; plastic limit (PL), at which soil changes from plastic to semi solid and shrinkage limit (SL), at which solid state begins. The numerical difference between liquid limit and plastic limit is known as plasticity index (PI).

Plasticity tests give information on cohesive properties of soil and the amount of capillary water which it can hold. They further indicate the following characteristics as outlined in Table 2.2.

Table 2.2: Correlation of plasticity with physical properties of soil (DETRL, 1974)

Characteristic	Comparing soils of equal LL with PI increasing	Comparing soils of equal PI with LL increasing
Compressibility	About the same	Increases
Permeability	Decreases	Increases
Rate of change	Increases	-
Toughness	Increases	Decreases
Dry strength	Increases	Decreases

Table 2.2 clearly depicts the relative importance of Atterberg limits on soil physical properties. Thus, provided that these limits can be reliably and quickly determined, any interested person can comfortably approximate the soil properties prior to their utilization. The liquid limit can be used to compute compression index as well measure shear strength of soil at the same water content. It is analogous to shear test.

For any type of soil, there is a limit to the density that can be achieved by compaction. This maximum density can be attained only at a particular moisture content known as the optimum moisture content. Below the optimum moisture content the soil fabric resists re-arrangements of the particles into a state of close-packing, whereas above it the presence of water in the voids inhibits the achievements of closer inter-granular contact. Thus, moisture content determination is very important for any engineering installation on soil. Maximum dry densities will only be achieved by careful control of the moisture content and compactive effort during construction as shown in table 2.3 and table 2.4 (Hazelton et al, 2007)

Table 2.3: Rating for compressibility & shrinks-swell potential based on plasticity index (Hazelton et al, 2007).

Rating	PI (%)
Low	<25
Medium	25-35
High	35-40
Very high	>45

Table 2.4: Rating for compressibility and shrink-swell potential based on liquid limit (Hazelton et al. 2007).

Rating	Liquid limit (%)
Low	<45
Medium	45-55
High	55-75
Very high	>75

Soil Aggregate Stability

Soil aggregates are groups of soil particles that bind to each other more strongly than to adjacent particles. The space between the aggregates provides pore space for retention and exchange of air and water. Aggregate stability refers to the ability of soil aggregates to resist disruption when outside forces (usually associated with water) are applied.

Aggregation affects soil erosion, movement of water and plant root growth. Desirable aggregates are stable against raindrop impact and water movement. Aggregates that break down in water or fall apart when struck by raindrops release individual soil particles that can seal the soil surface and clog pores. Thus, determining the magnitude of soil stable aggregates will lead to the approximation of the ability of soil to withstand degradation as well as provide productive base for agronomical activities.

The stability of aggregates is affected by soil texture, clay mineralogy, extractable iron, cation exchange capacities, the amount and type of organic matter present, and the type and size of the microbial population. Clay expands as it absorbs water. Expansion and contraction of clay particles can shift and crack the soil mass and create or break apart aggregates (Le et al., 2007). With knowledge of spectral behaviour of ionic components of soils, the use of calibration model can aid the prediction of the soil properties. Marcques et al (2004) has proposed some aggregate stability indices that are important in assessing soil aggregate stability.

2.2 Soil Physical Properties Analysis

2.2.1 Determination of liquid limit

The most common method used to determine the liquid limit is the Casagrande apparatus method (Singh and Chowdhary, 1990). In this method, an air dried soil sample of at least 100 grams passing the sieve number 36 B.S is placed on a glass plate and mixed with distilled water using a spatula until the mass becomes stiff and thick. It is then left to mature. A small portion of the paste is put in a cup, leveled off with a spatula to give a mass depth of 1 cm. A clean straight groove is then cut by a grooving tool through the paste along the diameter of the cup. The V-shaped groove is about 2 mm wide at the bottom. The handle of the apparatus is turned at two revolutions per second while counting the number of blows until the two parts of soil come in continuous contact at the bottom of the groove along a distance of about 13mm. A small quantity of soil from the portions of the sample that have just flowed together as well as some of the soil removed by the grooving tool is then placed in a container for moisture content determination. The number of blows and respective moisture content is recorded. The remainder of the soil is re-mixed with a small amount of distilled water and the process repeated. From the test results, a graph is drawn on moisture content against number of blows which is then used to determine the liquid limit as the moisture content at the 25th blow.

2.2.2 Determination of soil plastic limit

The remainder of the main sample used for Liquid limit test is used for Plastic Limit test. This sample is mixed with distilled water in a mortar until it is plastic enough to be rolled into a ball. The ball of soil is then rolled between the hand and flat glass plate so as to form the soil mass into a thread. When the diameter of the thread becomes less than 1/8 inch, the soil is kneaded together and rolled out again. In this way the water in the sample is evaporated by the heat of the hand until the soil just ceases to be plastic and crumbles.

When crumbling of the thread occurs at a diameter of 3.1 mm, the portion of the crumbled soil are gathered together and placed in a container for moisture content determination. Duplicate determinations are made and the average value of these moisture contents is taken as the plastic limit of the soil (Singh and Chowdhary, 1990).

2.2.3 Determination of soil aggregate stability indices

Several methods have been proposed for the determination of soil aggregate stability. However, the suitability of these methods depends on the purpose of the study. The most widely used method is the wet-sieving method (Márquez et al., 2004). In this method, cyclically submerging and sieving soil in water emulates the natural stresses involved in the entry of water into soil aggregates. Unstable aggregates break down and pass through the sieve into the water filled beneath the water. The testing procedure results in an index for aggregate stability (Eijkelkamp, 2007).

2.3 Infrared spectroscopy

Infrared spectroscopy is a technique based on the vibrations of atoms of a molecule upon interaction with infrared (IR) electromagnetic radiation. Infrared spectrum is obtained by passing infrared radiation through a sample and determining the fraction of the incident radiation absorbed or transmitted. The energy at which a peak in absorption or transmittance spectrum appears corresponds to the frequency of vibration of the atoms or molecules of the sample analyzed (Stuart, 2004).

Infrared spectroscopy is a non destructive analytical technique where the interactions between incident light and a material's surface are studied. The emergent spectral reflectance can be calibrated with sample properties and be used to provide simple, rapid, and accurate results (Chang et al., 2001; Faithful, 2002).

Soil IR spectroscopy is usually interfered by weak overtones and fundamental vibrations and hence difficulty to interpret. These interference effects are minimized by use of mathematical transformations. Some of the two normally used transformations are: the standard normal variate (SNV) transformation, which standardizes the particle size effects and baseline drift, and the de-trending (D) transformation that removes curvilinear of the spectrum by use of a second order polynomial correction(Faithful, 2002).

There are a number of commercial IR spectrometers with specific operating software. One of these is a multipurpose analyzer (MPA). It uses OPUS and OPUS Lab software. Soil infrared reflectance is measured relative to a white reference using a MPA at wavelengths from 0.35 to 2.5 μm. this wavelength is used as the study specification range for optimum results. Air dried soils are used to control the effect of soil moisture on reflectance. Reflectance spectra are recorded at two positions by rotating the sample 90^0 to sample within dish variation. The average of two spectra is recorded at each position to minimize instrument noise. Before reading each sample, reference spectral is recorded using a calibrated spectral. Reflectance readings for each wavelength band are then expressed relative to the average of white reference readings.

Faithful (2002) illustrates that interpreting infrared spectra involves two approaches : relationship between observed bands and peaks with known absorbing functional groups or chemical compounds and chemometrics approach wherein what causes the peak is ignored and the absorbing wavelength is taken as an empirical basis to give the best correlation with traditional chemical analyses. Laboratory reference method values can also be calibrated to reflectance measurements using partial least squares regression

Recent applications of near infra-red spectroscopy (NIRS) have been published. It has found wide application in pharmaceutical, petroleum, soil, grain and foliage analysis. Shepherd and Walsh (2007) have provided a list of examples of applications of infrared spectroscopy in agriculture and environment. These applications offer a reliable evidence of the potential of the technology for analysis of materials. Moreover, other researchers have further utilized the potential of the spectroscopy technology in predicting material properties using statistical techniques.

Many researchers have used IR spectroscopy to study soil attributes, soil degradation and to carry out soil mapping. For example, Shepherd and Walsh (2007) have stated that in Madagascar study the IR spectroscopy predicted soil conditions. The soil condition index was calibrated to reflectance bands for corresponding pixels from the landsat thematic mapper satellite. The resulting calibration model was then applied to map continuous surface of the soil condition class in the watershed (500 km^2). This approach to soil survey is a better method than the traditional conventional soil survey approaches where policies and land resource management depended on a limited number of observations for overall soil mapping units. This approach further provide data on soil properties or problems to be used for quantifying risk factors and a datum for change detection, especially caused by human and natural processes such as cultivation, pollution and weathering.

Infrared spectral is influenced by soil particle size, surface properties, moisture content, texture and aggregation. The main absorption features in near infrared spectroscopy range are associated with clay lattice and water oxygen hydrogen bond, whereas organic matter influences the overall position and shape of the spectrum.

Assessment and monitoring of soil quality has been well predicted using spectroscopy (Shepherd et al., 2003; Matthew et al, 2005; Reevies et al., 2001; Yong et al.,2005 and Keith et al., 2001.). Shepherd et al. (2003) have used near infrared spectroscopy in the characterization of organic resource quality for soil and livestock management in tropical agro-ecosystems. In their study, they employed near infrared spectroscopy (NIRS) in determining the quality of a variety of organic resources such as trees, crops and manure through calibration between spectral library of finery ground organic material and their wet chemical analysis attributes. The study gave prediction efficiencies ranging from 74 to 92% for Phosphorus, Potassium, Calcium and Magnesium. Related studies have predicted soil carbon fractions, decomposition and mineralization of organic residues using infrared spectroscopy (Janik et al., 2007; Shepherd et al., 2005).

David et al. (2005) have predicted ordinal clay mineralogy levels for montmorillonite and kaolinite, with prediction efficiency of 88% and 96%, respectively. They have confirmed that the use of auxiliary predictors and spectroscopy have the potential to improve predictions. Their findings suggest that Visual NIR soil characterization has the potential to replace or augment standard soil characterization techniques where rapid and inexpensive analysis is required. The carbon mineralization rates from different soil physical fractions has also been predicted using diffuse reflectance spectroscopy (Mutio et al., 2006) and found to have 82% prediction efficiency. Field-Measured Infiltration Rates have been predicted using diffuse reflectance spectroscopy with 67% efficiency (Omuto et al., 2003).

2.4 Geographical Information Systems (GIS)

GIS is a computer assisted system for the capture, storage, retrieval, analysis and display of spatial data. GIS is a multi-faceted system of hardware, software, data, people and methods. Implementation of GIS is described as the act of combining technology with people and methods. This requires a flexible, context sensitive and systematic approach (Heywood et al., 2002; Longley et al., 2005).

Spatial data is geographical data characterized by information about positions, connections with other features and details of non-spatial characteristics. For example, spatial data about a weather station may include: latitude and longitude as a geographical reference, connection details such as which services roads, lifts and ski trails would allow the meteologist access to the weather station (Longley et al., 2005).

There are three types of GIS analysis procedures: those used for storage and retrieved e.g. display of soil of the area of interest; contained queries that allow the user to look at patterns in the data using queries, e.g. allowing only sandy soils from soil database to be selected for viewing or analysis and modeling procedures or functions for the prediction of what data might be at a different time and place. Predictions could be made about, for example, which soils would be highly vulnerable to erosion in high winds or during flooding or the type of soil presence in unmapped area (Longley et al., 2005).

According to Longley et al (2005), the simplified view of the real world adopted by GIS is termed a model. A GIS populated with data and ideas about how these data interact is a spatial model. They further illustrate that all primary & secondary data have three modes of dimensions: temporal, thematic & spatial. For example, data about a road accident which took place in three pines valley on 14/2/1995. The three modes are: temporal (the date during which the accident occurred); thematic (what happened i.e. description of the character of real world feature the data refer) and spatial (position on earth's surface where the event took place i.e. the three pines valley.

Table 2.5 gives some of the areas where GIS is applied (Heywood et al., 2002)

Table 2.5: Examples of GIS application areas

Activity	Application
Socio-economic/Government	Health
	Local government
	Transport Planning
	Service Planning
	Urban management
Defense agencies	Target site identification
	Tactical support planning
	Mobile command modeling
	Intelligence data integration
Commerce and business	Market share analysis
	Insurance
	Fleet management
	Direct marketing
	Target marketing
	Retail site location
Utilities	Telecommunications
	Emergence repairs
	Service provision
	Network management
Environmental management	Landfill site selection
	Mineral mapping potential
	Pollution monitoring
	Natural hazard assessment
	Resource management
	Environmental impact assessment

In agriculture, GIS has been utilized to carry out land suitability assessment for Musa ABB group plantation (Boonyanuphap et al., 2004) wherein land use types, environmental conditions, soil characteristics, and the possibility for adjusting environmental conditions to make bananas more suitable for future growth were used to determine possible areas for new banana plantations under land management practices in Thailand.

Predictive soil mapping has been reviewed by Scull et al (2003) as a development of a numerical or statistical model of the relationship among environmental variables, soil properties and GIS. This has improved soil map production. Bien et al (2004) have employed a GIS-based approach for the long term prediction of human health risks at contaminated sites to enhance efficiency of contaminated land management. Brown (2007) has also shown how the use of Visual NIR soil spectral library can be used for local soil characterization and landscape modeling in a watershed.

2.4 Kriging

To determine spatial distribution of geographical phenomenon one needs to estimate its values at unsampled points. Estimation of the values gives rise to predictive mapping that utilizes geostatistical methods, one of which is Kriging (Geoff Bohling, 2005). Kriging Is an optimum interpolation based on regression against observed z values of surrounding data points weighted according to spatial covariance values.

Kriging involves a set of methods: simple Kriging that utilizes known mean; ordinary Kriging that works with unknown mean; co-Kriging that works well where two or more variables are spatially interdependent with the one whose values are to be estimated; universal Kriging for non-stationary variations; probability and indicator Kriging that are non parametric and based on indicator functions and Bayesian Kriging.

2.5 Soil Mapping

The soil map provide a framework within which to carry out a detailed sampling and testing required for final engineering design. In large-area survey, small scale pedological maps may provide invaluable basic data which can sometimes be augmented and interpreted for engineering application. It is necessary to undertake a regional soil engineering survey to provide information on the principle advantages, inherent constraints and distribution of soils in order to aid the location of engineering development (Brink et al., 1984).

Maps serve as storage medium for information. They help us understand the spatial patterns, relationships, and complexity of the environment in which we live (Robinson et al 1995).

According to Robinson et al (1995), a major shortcoming in current cartography is the lack of data in readily usage forms. Much available information lacks portability because it is coded in incompatible hardware or software specific format. Thus, cartographic development lags significantly behind technological potential. More people are needed to be trained to use and maintain digital database & visualization.

Robinson et al. (1995) have described two main types of coordinate systems commonly used in mapping: Geographical coordinate system which uses latitudes & longitudes for curved surfaces and the rectangular coordinate or plane coordinates which locate position on a flat map. The following are the commonly used coordinate systems ((Robinson et al., 1995 and Longley et al., 2005): Universal transverse mercator (UTM) grid system that is used for topographic maps, satellite imagery, natural resources data bases and other applications that require precise positioning; Universal Polar Stereographic (UPS) grid System that covers polar area (south of 80° S latitude & north of 84° N latitude); State plane coordinates system (SPC) that is based on UTM and tied to location in the national geodetic surgery system and Public Land Survey (US Rectangular Land Survey) System that is defined on ground.

Unlike paper maps, GIS has a zoom facility that allows mapping to be viewed at a range of scales. GIS provides a seamless medium for viewing space and 3-dimensional view of the world. Geo-visualization in GIS is beneficial in making the growing complexity of land use planning, resource use and community development intelligible to communities. It unlocks the potential of many digital data sources. It helps communities to shift land use decision from regulatory processes to performance-based strategies and community decision making process more proactive.

In addition to geo-visualization, spatial analysis in GIS has demonstrated the ability to turn raw data into useful information in pursuit of scientific discovery. It has been used to further the aim of science by revealing patterns that were not previously recognized and that hide at undiscovered generalization and laws.

2.5 Conclusion

Describing spatial variability of soil physical properties is important as it enables users of land resources to make better predictions of soil behavior and performance. It is evident that soil physical properties have been valuable in soil utilization.

Soil physical properties analysis methods have been in existence for many years and are well documented These methods need refinement and/or replaced with more robust methods. Infrared spectral reflectance has been tested and confirmed to be able to analyze pharmaceutical products, petroleum, soil, grain and foliage.

The fore-mentioned GIS potentials shall be utilized to present the final output of the study as user-friendly as possible through presentation of characteristic interpolated soil physical properties map. This is anticipated to necessitate rapid and reliable development and environmental management decisions (Longley et al., 2005).

The relationship of the various properties of soil with spectral reflectance implies that spectral reflectance is a possible predictor of soil physical properties. Moreover, the ability of GIS to produce spatial maps of distribution of features on the earth's surface provides a reliable evidence that soil physical properties can be easily mapped using the GIS technology.

It is evident that new soil physical assessment methodology and their spatial prediction which is proposed in this study is a timely undertaking that would add value to precision agriculture, infrastructural development and environmental management. It is important to note that the approach proposed has a solid evidence for its success.

3. MATERIALS AND METHODS

3.1 Study Area

The soil samples used in this study were obtained from the upper Athi river watershed, Machakos district, Kenya (Figure 3.1). The watershed covers about 5000 km^2, 85% of which is semi-arid to arid.

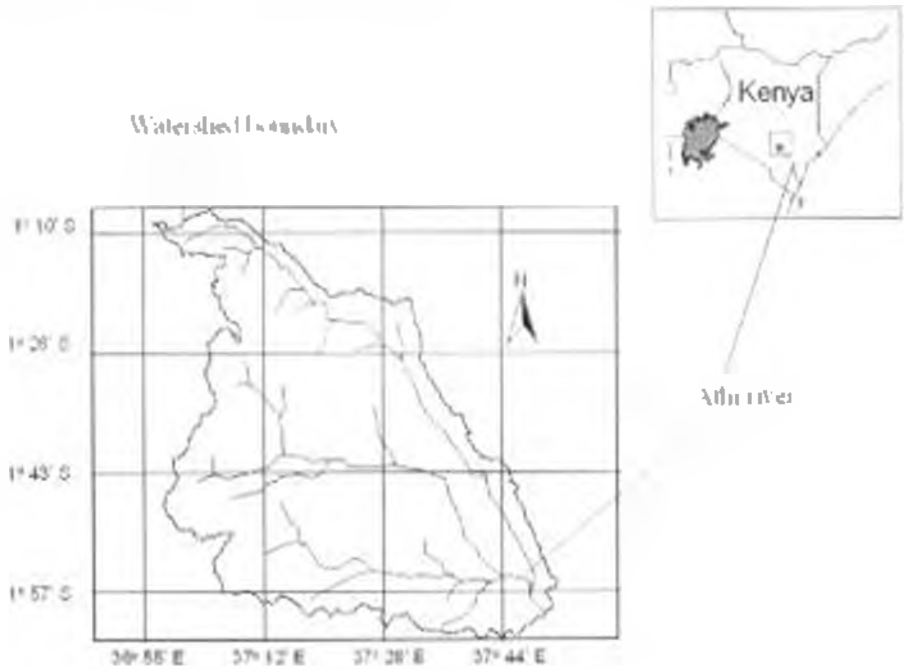

Figure 3.1: Study area (Omuto and Shrestha, 2007)

3.2 Spatial Sampling Design

The procedures used for soil sampling and spectral reflectance determination have been described in Omuto and Shrestha (2007). The sampling frame consisted of sampling points placed within square plots. The plots grouped into clusters of 4 plots each.

The plots in each cluster were arranged in a Y-frame to represent all parts of a cluster. One plot at the centre and the other 3 at each ends of the Y-arms, 480m from the centre plot. In total 45 clusters were randomly placed, representing the whole study area.

The plots were georeferenced at their centre with a *Garmin*® GPS (GARMIN International, 2002). Each plot had 3 sampling points: P1, P2, P3 such that P1 was 5m from the surface, P3 was 25m from the surface and P2 was between P1 and P3. Samples from the P2 point were picked using $100cm^3$ stainless steel core-rings.

The samples were carefully packed in polythene bags and transported to the laboratory. They were air-dried and sieved to pass through 2mm sieve.

A total of 696 soil samples were collected and scanned for spectral reflectance at the World Agroforestry Centre (ICRAF) - Nairobi.

About 40 g of each soil sample were taken for laboratory determination of physical properties at the University of Nairobi. The corresponding spectral reflectance signatures were also collected for pedo-transfer model development with physical properties.

Sample processing

This is the process of documenting samples received into the laboratory. The samples were first allocated laboratory identification numbers (labid). The labid was used to trace the sample through the analysis steps.

All the samples were first sorted out and arranged in a sequential order on the wooden trays. Information on the packaging material was recorded down neatly and clearly. Packages that had no samples were noted. The wooden trays were labeled with batch number, first and last labid of the samples in the tray.

Sub sampling was done to select soils from the foresaid batches for both spectral and wet analysis.

3.3 Analysis of Physical Properties of Soil

The physical properties of the soil samples determined were consistency limits and aggregate stability indices.

3.3.1 Determination of liquid limit

The American Society for Testing and Materials (ASTM) D 423-66 procedure was used (Liu and Evett, 1984); using the Casagrande apparatus as illustrated in figure 3.2. This involved taking a sample of about 40g and placing it in the evaporating dish and thoroughly mixing with 2 to 3 ml of distilled water. The sample was then alternately and repeatedly shred, kneaded, and chopped with a spatula. Further additions of water were done at increments of 1 to 3 ml to produce a thick paste with consistency that required a certain number of drops of the cup (blows) to cause closure of the standard groove. The moisture content of the paste at the closure of the groove was used to calculate the liquid limit of the sample by plotting moisture content versus the number of blows (figure 3.3). The moisture content at the 25^{th} drop of the cup constituted the liquid limit of the soil sample.

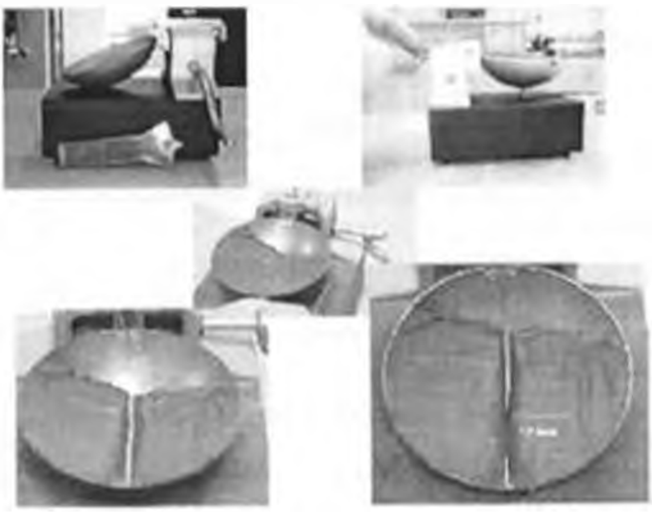

Figure 3.2: Liquid limit determination apparatus

In carrying out the experiment, the following precautions were put in place: all clay lumps were broken down to provide uniform flow; the soil samples were dried under room temperature that did not exceed 30 ^{0}C to reduce effect of drying on results. The soil and water were thoroughly mixed to ensure uniformity of liquid absorption; the liquid limit device was frequently checked and adjusted to ensure good working condition and position; the moisture content was immediately determined after the soil had flown together and safety measures were instituted to avoid injury from the experiment.

$$Moisture\ content = \frac{(Weight\ of\ moist\ soil) - (weight\ of\ oven\ dry\ soil)\ X\ 100}{(weight\ of\ oven\ dry\ soil)} \qquad [3.1]$$

The moisture content for three portions of the same sample was determine and plotted against the number of drops of the cup. Figure 3.3 shows a typical curve that illustrates how the liquid limit was determined. The figure shows that the moisture content at the 25th blow is 58%. The 58% constituted the moisture content of the sample.

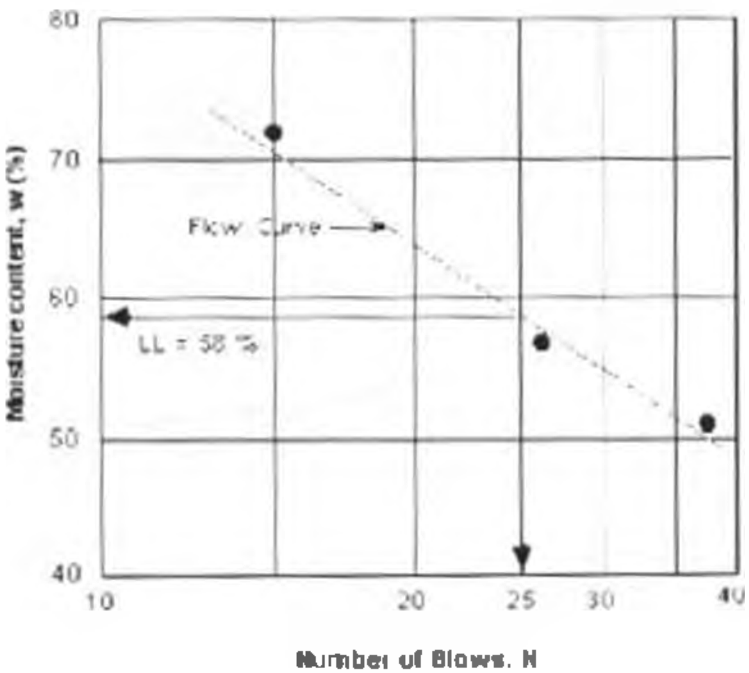

Figure 3.3: Liquid limit determination chart

3.3.2 Determination of plastic limit

The ASTM D 424-59 procedure was used in this process (Liu and Evett, 1984). In this procedure, about 40g of soil were mixed with just enough water to necessitate rolling of the soil paste into a ball. This wet soil was shaped into a ball and rolled into a thread on a flat glass plate with tips of the fingers of one hand as illustrated in figure 3.4. The soil was further remolded again into a ball upon reaching a diameter of approximately 3.2mm. Rolling and remolding were repeated until the thread just started to crumble at a diameter of approximately 3.2mm. The moisture content at this point was determined by oven drying the soil within 105-110^{0}C for 24 hours. The plastic limit of the soil sample was determined using the following equation (Liu and Evett, 1984):

$$Plastic\ limit = \frac{(Weight\ of\ moist\ soil + can) - (weight\ of\ oven\ dry\ soil + can)\ X\ 100}{(weight\ of\ oven\ dry\ soil + can\ ) - (weight\ of\ can)} \quad [3.2]$$

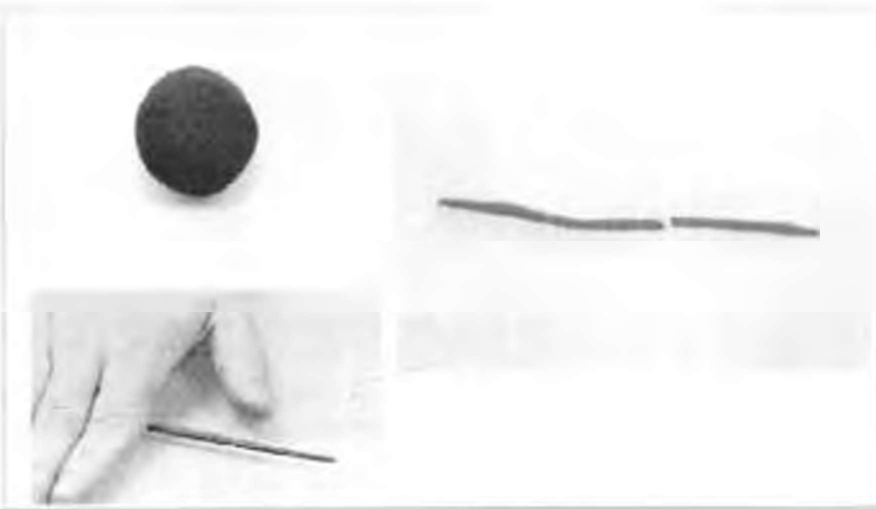

Figure 3.4: Plastic limit determination illustration

To ensure reproducibility of the results, the tests were done exactly the same way each time.

3.3.3 Determination of Plasticity Index (PI)

Plasticity index was calculated as follows:

$$PI = LL - PL \quad [3.3]$$

Where:

LL is the Liquid limit

PL is the Plastic Limit and

PI is the plasticity Index

3.3.4 Aggregate stability Index determination

Aggregate stability was determined using the wet sieving apparatus (described in www.eijelkamp.com, 16th July 2008). The procedure involved gradual mixing of about 4g of soil with sodium hexametaphosphate in a shaking rack of the wet sieve apparatus (Figure 3.5).

Figure 3.5: Complete set of wet sieve apparatus

The fraction of the aggregates passing through the sieve and the fraction retained were determined and used to obtain aggregate stability index as follows.

$$AI = \frac{fR}{fR + fP} \qquad [3.4]$$

Where

AI= Aggregate stability index

fR= Fraction of aggregates retained

fP= Fraction of aggregates passed

3.4 Soil Analysis Using Near Infra-Red Spectral Reflectance

The soil samples were finely ground to 2mm size and put into specially cleaned Petri dishes for scanning. The dishes were identified with dish numbers. The Petri dishes were arranged in a sequential and consistent manner to avoid confusion that might have arose during sample scanning.

Scanning the samples was then carried out based on the Standard Operating Procedure for measurement of soils and plants, available at the soil analysis lab, ICRAF.

3.5 Calibration of Soil Physical Properties and Infra-Red Spectral

3.5.1 Reflectance

The first step in the development of pedotransfer functions were spectral transformation to remove mainly spectral noise caused by soil moisture, sample preparation and optical equipment factors (Faithful, 2002). This involved Savitzkyt-Golay spectral transformations using second order polynomial derivative, the Savitzky-Golay function (Faithful, 2002; Savitzky and Golay, 1964):

$$\frac{dR}{\lambda} = \frac{1}{2}\left(\frac{f_{i+1} - f_i}{(i+1) - i} + \frac{f_{i+2} - f_{i+1}}{(i+2) - (i+1)}\right) \qquad i = 1,216 \qquad [3.5]$$

where λ is the spectral wavelength and f is the spectral reflectance. Equation (3.3) was implemented using the Unscrambler® v9.0 software (CAMO Technologies, 2007). The smoothened spectral signatures were inspected to identify spectral regions with low signal-to-noise ratio. The regions that showed high spectral noise were deleted from the spectral signature before calibration since they have been shown to bear no meaningful calibration strengths (Shepherd and Walsh, 2007).

The spectral signatures of the soil samples were then matched to their liquid limits, plastic limits and aggregate stability indices using query tables developed in Microsoft® Access. The link was established.

Calibration of soil physical properties with soil spectral reflectance was done using partial least squares (PLS) regression method (Haaland and Thomas, 1988). This method generalizes and combines features from principal component analysis and multiple regression. It was used to predict a set of soil physical properties, the dependent variables from a large set of soil spectral reflectance, the independent variables observed on 696 samples. The independent and dependent variables were denoted respectively as X and Y variable sets.

The PLS regression analysis was carried out to predict Y from X. With Y as a vector and X as a full rank matrix, the goal was accomplished using a multiple regression analysis. The PLSR analysis was conducted for each soil physical property investigated with the help of Unscrambler® v9.0 software (CAMO Technologies, 2007). The model was calibrated on the selected calibration set. A cross validation was performed to avoid PLSR over fitting. The predicted outputs were compared with measured soil physical properties by means of the determination coefficient (r^2).

In order to ensure robust and accurate model development, the samples were divided into two sets: one set had 50% of the samples and was used to develop calibration models while the remaining 50% of the data was used to test the models. Coefficient of determination (r^2) and root mean square error (RMSE) were used to assess the predictive accuracy of the models developed.

3.6 Procedure for Spatial Prediction of Physical Properties of Soil

3.6.1 Spatial structure for the soil physical properties

A normal probability plot for the soil physical properties was used to assess whether the data collected was normally distributed as required by spatial models (Omuto and Biamah, 2008).

3.6.2 Spatial prediction of physical properties of soil

Prediction of spatial distribution of the soil physical properties was determined using geostatistical techniques. Spatial variation with interdependence was described with a variogram.

In geostatistics, the concept of variance from statistics is extended to semivariance. A semivariogram is a mathematical quantity for describing the spatial structure of a property. It is derived from the concept that the occurrence of a soil property in the field is not completely disordered but has a relationship with its neighbours (Omuto and Biamah, 2008)

The spatial structure of each property was characterized by experimental semivariogram using the equation 3.6 (Nielsen and Wendroth (2003). Semivariogram, which is required by this method, were determined as follows:

$$\lambda(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[z(x_i) - z(x_i + h) \right]^2 \qquad [3.6]$$

where $N(h)$ is the number of pair of georeferenced soil samples $[Z(x_i), Z(x_i + h)]$ separated by a distance h. $\lambda(h)$ is the experimental semivariogram value at distance interval h; z(xi) and z(xi+h) are sample values at two points separated by the distance interval h. The two points form a pair of coordinates obtained with global positioning system (GPS).

The equation 3.6 was used to calculate semivariance values. These values were further analysed for mapping in R software (R Development Core Team, 2008, www.cran.r-project.org).

Ordinary Kriging (OK) method solely utilizes primary data of soil physical properties measured at sampled locations u_a to estimate soil physical properties at unsampled locations u.

The constancy of the mean is assumed only within a local neighborhood *W(u)* centered at the location u being estimated. Here, the mean is deemed to be a constant but unknown value, that is m(u')=constant but unknown, W(u).

The OK estimator is written as a linear combination of the n(u) data $Z(u_a)$ with a single biasedness constraint (Toktam et al., 2010) in the form:

$$z^*(u) = \sum_{a=1}^{n(u)} \lambda_a(u) z(u_a) \qquad [3.7]$$

where

N(u) is the number of values $Z(u_a)$ involved in the estimation of the unrecorded point u. λ_a are the weights. $Z(u_a)$ is the laboratory based analytical estimate of soil physical properties at a point.

The assumption was that the estimator is unbiased based on condition that the expression of the difference between the estimator and the estimated values should be zero and that the variance between the two should be minimum. Thus,

$$E\{z^*(u)-z(u)\} = 0 \qquad [3.8]$$

and

$$\sigma_k^2(u) = E\{[z^*(u) - z(u)]^2\} = \min imum \qquad [3.9]$$

Substituting equation 3.6 into equation 3.7 and differentiating yields

$$\frac{\partial}{\partial z(u_a)}\left[\sum_{a=1}^{n(u)} \lambda_a z(u_a) - z(u_a)\right] = 0 \qquad [3.10]$$

Or

$$\sum_{a=1}^{n(u)} \lambda_a = 1 \qquad [3.11]$$

Expression 3.9 is known as the Kriging system.

In order to solve the Kriging system, N needs to be known. N was determined by Mean Reducing Error (MRE) defined by

$$MRE = \frac{1}{n}\left[\sum_{a=1}^{n} \lambda_a z(u_a) - z^*(u_a) \right] \big/ \sigma_r(u_a) \qquad [3.12]$$

Where n is the total number of sample points which are close together. In this study MRE values were plotted as a function of neighbourhood values N. the value of N for which MRE was approximately zero was chosen. Solutions of equation 3.10 and 3.9 were established using the R software (R Development Core Team, 2008, www.cran.r-project.org). The map of the study area was then produced using the R software .

R is an elegant and comprehensive statistical and graphical programming language. It provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification and clustering) and graphical techniques, and is highly extensible. In R software, prediction of soil physical properties involved design of a prediction program (Appendix 1). Once the program was ready, data processing was done to fit the designed program coding system.

Data processing involved combining of soil physical properties with geographical coordinates of the soil sampling points using the Microsoft Excel software. The resultant data sheet was then exported to notepad software. The data in notepad for was now fitted into the prediction program in R.

The spatial prediction was done following the procedure described in section 3.6.2, using the Kriging Method that was implemented using the R software. The program codes were designed that were run in the R software environment. The program codes are shown in appendix 1, 2 and 3. These programs utilized the procedure in section 3.6 to produce the prediction maps (figures 4.6, 4.7 and 4.8).

4. RESULTS AND DISCUSSION

4.1 Results of Laboratory Soil Analysis

As described in the methodology section 696 samples of soil were analyzed to determine their physical properties. The following sub-sections show extracts of the results obtained for liquid limit, plastic limit and aggregate stability indices.

4.1.1 Liquid limit

The procedure in section 3.3.1 was followed. Table 4.1 shows a sample of the results obtained.

Table 4.1: Sample moisture content results

Sample no.	No. of Drops Drops	Can Number	Can weight (grams)	Weight of Moist Soil + Can (grams)	Weight of Dry Oven Soil + Can (grams)	Weight of water (grams)	% moisture Content
1a	24	13	13.24	14.08	13.85	0.23	37.70
1b	20	14	8.39	9.19	8.93	0.26	48.15
1c	14	15	7.38	8.28	7.98	0.3	50.00

$$\text{Moisture content} = \frac{(\text{Weight of moist soil+can}) - (\text{weight of oven dry soil+can})\ X\ 100}{(\text{weight of oven dry soil+can}) - (\text{weight of can})}$$

For sample 1a,

$$\text{Moisture content} = \frac{(14.08) - (13.85)\ X\ 100}{(13.85) - (13.24)} = 37.70\ \%$$

The full set of the liquid limit results was used to plot moisture content versus number of blows for each sample. Figure 4.1 is a typical representation of the plots.

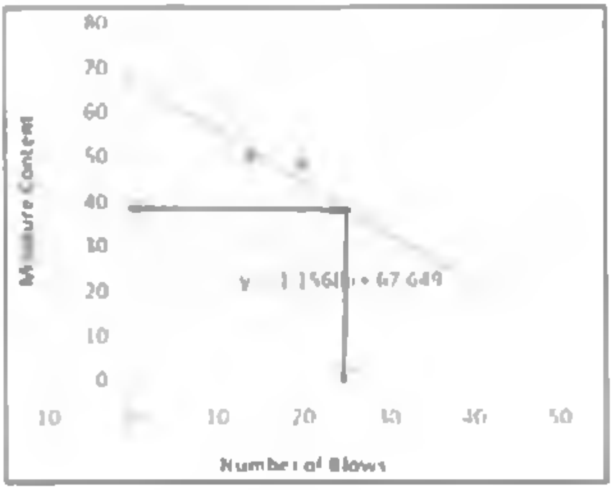

Figure 4.1: Sample liquid limit determination plot

Figure 4.1 shows that the moisture content of the sample represented by the plot was 38% at the 25[th] blow. This moisture content represented the liquid limit of the sampled soil (Liu and Evett, 1984).

4.1.2 Plastic limit

The plastic limit was determined as described in section 3.3.2. Table 4.2 shows sample results.

Table 4.2 Sample plastic limits

Samples	Can Number	Can weight (grams)	Moist Soil + Can weight (grams)	Dry oven Soil + Can weight (grams)	% Moisture Content
1	22	6.5	11.21	10.65	13.49
2	23	6.86	9.83	9.47	13.79
3	24	13.14	17.17	16.68	13.84
4	25	13.17	14.82	14.66	10.74
5	26	9.61	12.1	11.79	14.22
6	27	9.62	11.84	11.57	13.85
7	28	6.19	11.64	10.83	17.46

From table 4.2, the plastic limit for the soil sample number 1 was obtained as follows:

$$\text{Plastic limit} = \frac{(Weight\ of\ moist\ soil + can) - (weight\ of\ own\ dry\ soil + can)\ X\ 100}{(weight\ of\ own\ dry\ soil + can) - (weight\ of\ can)}$$

For sample number 1 (table 4.2),

$$\text{Plastic limit} = \frac{(11.21 - 10.65)X\ 100}{(10.65 - 6.5)} = 13.49$$

4.1.3 Plasticity index

Using equation 3.1, the plasticity index for soil samples were calculated as follows:

Plasticity Index = LL – PL

For the typical sample number 1 (table 4.1 and table 4.2), the plastic limit was calculated as

Plasticity Index = LL – PL

= 38% -13.49%

= 24.51%

The consistency limits determined herein are very useful for identifying soils. They give a measure of the plasticity of the soil.

4.1.4 Aggregate stability index

The soil aggregate stability indices was obtained as described in section 3.3.4, the results of which are represented by 10 samples as shown in table 4.3. All weights were in grams.

Table 4.3: Sample aggregate stability indices

Samples	Batch	Labid	Initial dry soil	Dry soil + Can	Can No.	Can weight (grams)	Weight of retained soil + Can (grams)	Retained soil weight (grams)	Aggregate stability index
1	1784	6466	4.14	86.85	1	89.45	90.57	1.12	0.27
2	1784	6467	5.26	86.61	2	81.38	83.04	1.66	0.32
3	1784	6477	4.95	76.22	3	82.24	83.46	1.22	0.25
4	1784	6464	3.76	77.87	4	84.2	85.15	0.95	0.25
5	1784	6479	4.10	96.26	5	94.59	96.97	2.38	0.58
6	1784	6468	3.90	88.61	6	81.43	83.3	1.87	0.48
7	1784	6459	5.36	73.98	7	86.48	89.48	3.00	0.56
8	1784	6435	3.87	82.13	8	95.40	96.76	1.36	0.35
9	1784	6284	4.28	97.94	17	81.65	83.63	1.98	0.46
10	1784	6440	4.4	81.58	12	95.08	96.62	1.54	0.35

4.2 Near Infrared Soil Analysis

The experimental work was done independently at ICRAF soil science laboratories to obtain the near infra-red (NIR) Spectral data discussed in section 3.4. The typical sample results are shown in table 4.4 below.

Table 4.4: Extract of soil spectral reflectance

Samples	Labid	W350	W360	W370	W380	W390	W420
1	6080	0.0526	0.0519	0.0497	0.0530	0.0529	0.0599
2	6081	0.0510	0.0468	0.0502	0.0444	0.0492	0.0526
3	6082	0.0548	0.0620	0.0619	0.0621	0.0643	0.0713
4	6083	0.0738	0.0715	0.0697	0.0741	0.0769	0.0866
5	6084	0.0815	0.0628	0.0641	0.0692	0.0702	0.0784
6	6085	0.0763	0.0611	0.0669	0.0660	0.0678	0.0759
7	6086	0.0578	0.0471	0.0455	0.0451	0.0437	0.0457

The spectral readings were from wavelength 350 to 2500nm. The table only gives an extract of the first few wavelengths. The reflectance data for all samples produced unique spectral plots. Figure 4.2 shows typical curves for five samples.

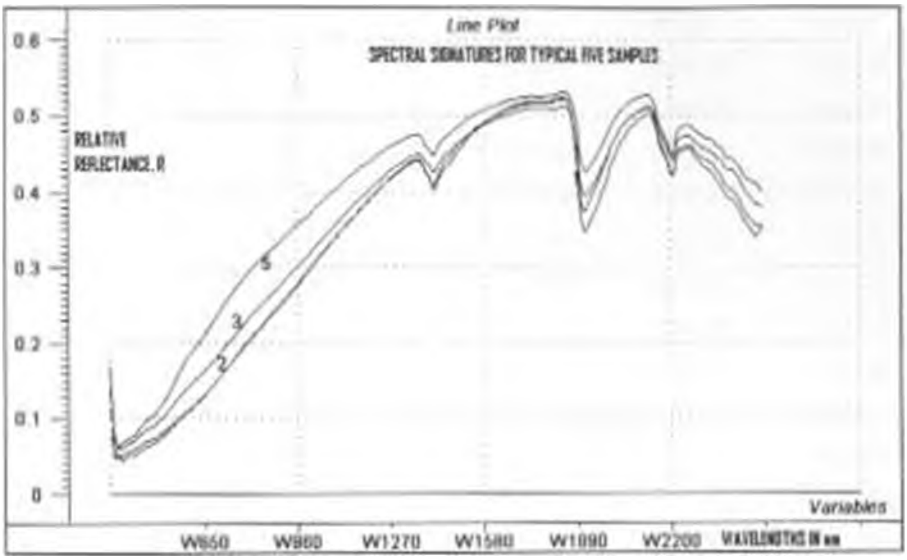

Figure 4.2: Typical spectral signature for five soil samples.

4.3 Calibration of physical properties with soil spectral reflectance

The near infrared reflectance (NIR) spectral data was transformed using Savitzkyt-Golay 2nd derivative to remove spectral noise. The data was then matched with their respective liquid limit, plastic limit and aggregate stability indices. Both NIR and soil physical properties results were tabulated as represented by table 4.5.

Table 4.5: Soil Atterberg limits, aggregate stability indices and spectral reflectance

Samples	AI	LL	PL	W350	W360
1	0.53232	0.23196	0.14593	0.05255	0.05188
2	0.59572	0.25158	0.15563	0.05096	0.04681
3	0.47825	0.25743	0.16080	0.05484	0.06200
4	0.34358	0.22246	0.15366	0.07381	0.07150
5	0.55376	0.21397	0.16956	0.08152	0.06282
6	0.43469	0.23693	0.15249	0.07628	0.06110
7	0.63507	0.24091	0.13978	0.05782	0.04710
8	0.59109	0.26054	0.16322	0.04380	0.04116
9	0.24058	0.20515	0.14025	0.07078	0.07111
10	0.34699	0.21546	0.14326	0.08143	0.06439
11	0.29973	0.21632	0.14402	0.07582	0.06636
12	0.39873	0.23363	0.15132	0.05767	0.05914
13	0.55156	0.25615	0.15840	0.04606	0.04185

Key

LL Liquid limit

PL Plastic limit

AI Aggregate index

W350 350nm wavelength

The tabulated NIR and Soil physical properties data was split into two sets with equal number of samples. One set was used for calibration while the other set was used for validation.

Table 4.6 shows an extract of the calibration data while table 4.7 shows an extract of the validation data.

Table 4.6: Calibration data extract for soil Atterberg limits, aggregate stability indices and spectral reflectance.

Samples	AI	LL	PL	W350	W360
97	0.5161	0.2603	0.1486	0.0452	0.0385
594	0.5616	0.2491	0.1600	0.0370	0.0420
371	0.5598	0.2767	0.1592	0.0480	0.0415
484	0.6544	0.2551	0.1656	0.0617	0.0569
495	0.4891	0.2853	0.2112	0.0625	0.0431
686	0.6292	0.2561	0.1572	0.0502	0.0414
129	0.5589	0.2476	0.1448	0.0581	0.0505
205	0.7082	0.2952	0.1794	0.0494	0.0511
384	0.6238	0.2454	0.1312	0.0479	0.0378
488	0.7664	0.2682	0.1574	0.0674	0.0535
624	0.4205	0.2244	0.1689	0.0611	0.0536
62	0.3191	0.2698	0.1620	0.0584	0.0473
361	0.3182	0.2101	0.1467	0.1022	0.0965
675	0.3504	0.2427	0.1580	0.0572	0.0423
329	0.5397	0.2291	0.1439	0.0646	0.0498
620	0.5419	0.2882	0.2195	0.0432	0.0379
547	0.5939	0.2723	0.1513	0.0434	0.0369
173	0.2850	0.2689	0.1394	0.0301	0.0373
546	0.7052	0.2771	0.1389	0.0466	0.0332

Table 4.7: Validation data extract for soil Atterberg limits, aggregate stability indices and spectral reflectance.

Samples	AI	LL	PL	W350	W360
287	0.2790	0.2779	0.0795	0.0594	0.0539
476	0.6039	0.2473	0.1517	0.0800	0.0637
521	0.6421	0.2554	0.1913	0.0516	0.0436
188	0.4989	0.2879	0.1811	0.0437	0.0356
514	0.6348	0.2481	0.1463	0.0327	0.0251
454	0.6344	0.3040	0.1787	0.0620	0.0495
71	0.6068	0.2595	0.1927	0.0567	0.0556
133	0.3741	0.2735	0.2156	0.0291	0.0408
112	0.7493	0.2925	0.1726	0.0574	0.0478
309	0.4728	0.2907	0.2016	0.0638	0.0494
684	0.4926	0.2572	0.1620	0.0610	0.0538
165	0.5982	0.2541	0.1589	0.0347	0.0299
597	0.7186	0.2391	0.1366	0.0585	0.0516
169	0.6094	0.2898	0.1617	0.0308	0.0341

Both validation and calibration was done as described in section 3.5, the results of which are illustrated in the summary of calibrations between infrared spectral reflectance and soil physical properties in table 4.8.

Table 4.8: Spectral calibration and validation statistics

Soil Eng. Property	Transform	Index	Calibration model		Validation model	
			r^2	RMSE	r^2	RMSE
AI	BC	0.85	0.78	0.075	0.69	0.077
LL	ln		0.81	0.019	0.77	0.021
PL	BC	0.5	0.88	0.042	0.85	0.044
Cases			697		232	

The relationship of the physical properties with spectral reflectance implies that spectral reflectance is a possible predictor of soil physical properties. The predictive capacity is high for plastic limit (r^2>0.85), followed by liquid limit (r^2>0.77), and aggregate stability index (r^2>0.69).

These results are quite impressive, depicting the relatively ease of using the spectral signatures to predict the soil physical properties. Since soil spectral reflectance can be rapidly and easily obtained with high repeatability, many soil samples can be analyzed rapidly for their Atterberg limits and aggregate stability by use of infra-red spectroscopy.

Nagaraj (2005) has developed a method for determining liquid limit using absorption water content and equilibrium water content under Ko-stress. Nagaraj (2005) has also developed a prediction of plasticity index using the flow index by the cone penetration cup method. Related prediction studies have been developed that correlate one property of soil to another. Thus, correlation with infrared spectral data provides another rapid method for determination of soil physical properties (Lal and Pierce, 1989; Harpstead et al., 1997).

The soil physical properties form the most important inferential soil tests with very wide universal acceptance and have provided a basis for explaining most engineering properties of soil including compressibility, permeability, toughness, dry strength and rate of volume change of soil as well as provision of reliable information concerning the cohesive properties of soil and amount of capillary water which it can hold (DETRL, 1974).

In soil classification, the Casagrande (1947) used consistency limits to develop the Casagrande soil classification system. In this system, the soil is classified into two major classes: those with coarse grains and those with fine grains. Those with coarse grains are classified by sieving, visual inspection and rubbing between fingers while those with fine grains are classified by consistency limits (Casagrande, 1947).

4.4 Spatial Prediction of soil physical properties

Normal probability plots for the soil physical properties were carried out using the Unscrambler software (Figures 4.3, 4.4 and 4.5) to test whether the soil data was normally distributed. Spatial modeling requires that the data used should be normally distributed (Omuto and Biamah, 2008).

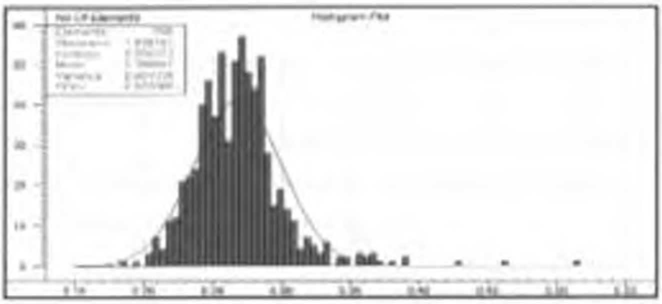

Figure 4.3: Distribution of soil liquid limit

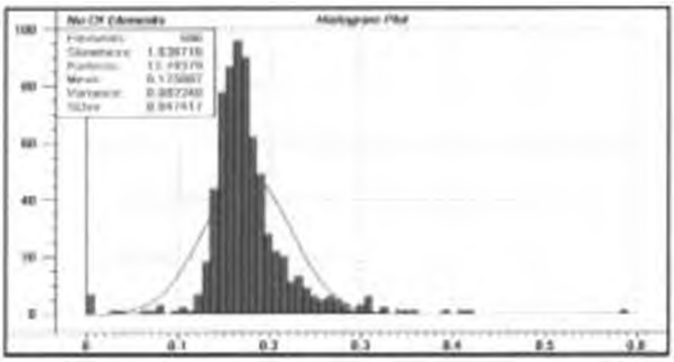

Figure 4.4: Distribution of soil plastic limit

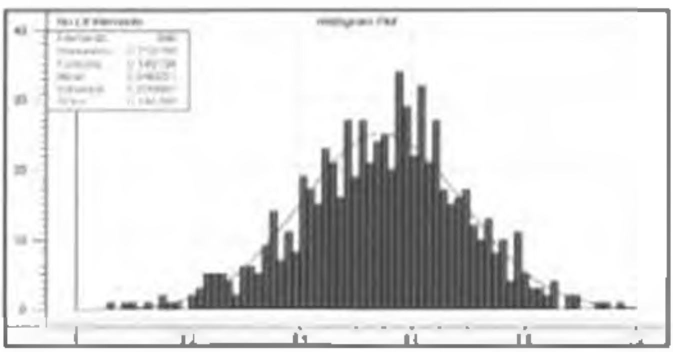

Figure 4.5: Distribution of soil aggregate stability index

The plots depict normal distribution curves. The results of the probability plots confirm that the soil samples were randomly sampled to represent the whole soils in the study area, the observations were independent from one another and that the analysis was accurate and hence represents a reliable data.

Using the R software (R Development Core Team, 2008, www.cran.r-project.org), prediction maps for soil physical properties were produced (figures 4.7, 4.8 and 4.9). The figures compare well with those produced by Geoff Bohling (2005) during the study of porosity (figure 4.6).

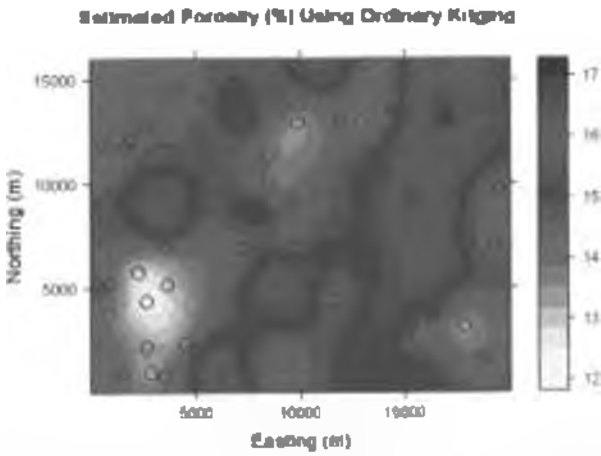

Figure 4.6: Spatial prediction map of the soil porosity (bohling, 2005)

Using Kriging method, Bohling estimated values of porosity on unsampled points. For example, using his prediction map (figure 4.6) the estimating porosity at point (1500, 10000) based on porosity values at nearest six data points was found to be approximately 16%.

The same Krigging procedure was used to produce prediction maps shown in figures 4.7, 4.8 and 4.9.

4.4.1 Spatial prediction of soil liquid limit

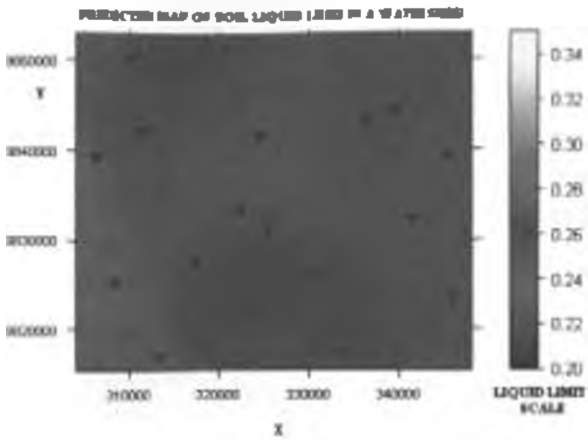

Figure 4.7: Spatial prediction map of the soil liquid limit

Figure 4.7 shows that at point (330000, 9820000), the LL is approx 0.27 while at point (330000, 9835000), the LL is 0.29. With this data, a developer can make an informed decision on the suitability of the soil for his/her project.

4.4.2. Spatial prediction of soil plastic limit

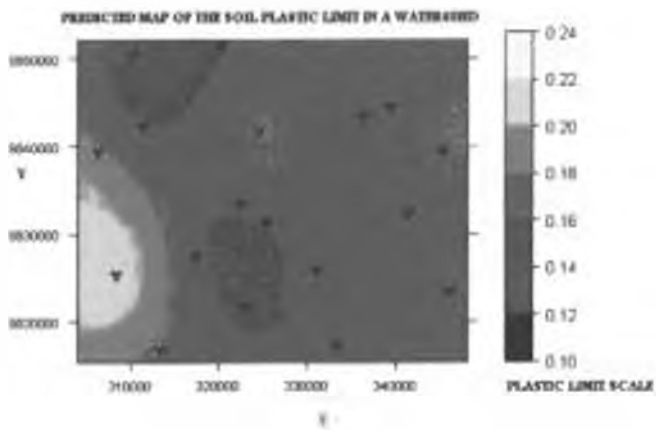

Figure 4.8: Spatial prediction map of the soil plastic limit

This map shows at point (325000, 9820000), the PL is approx 0.15 while at point (310000, 9825000), the PL is 0.21.

4.4.3 Spatial prediction of soil aggregate stability indices

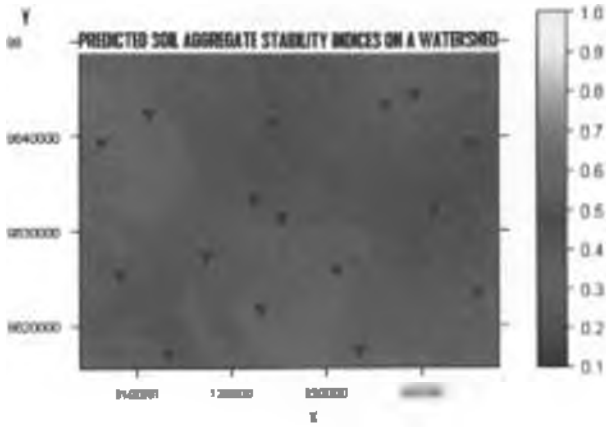

Figure 4.9: Spatial prediction map of the soil aggregate stability index

This map shows at point (310000, 9840000), the AI is approx 0.65 while at point (330000, 9840000), the AI is 0.5.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The combined application of infrared spectroscopy and geographical information systems was found to be successful in predicting the soil physical properties through the development of calibration and spatial prediction models.

The calibration model developed is relatively accurate and stable in predicting soil physical properties. The study found a calibration coefficient of determination of 0.78, 0.81 and 0.88 and validation coefficient of determination of 0.69, 0.77 and 0.85 for aggregate index, liquid limit and plastic limit respectively. All the three properties were well fitted into the model giving reliable accurate prediction with low RMSE of between 0.019 and 0.077.

Moreover, the spatial prediction maps developed show that interested parties, including agronomists and structural engineers could reliably estimate soil physical properties of a given location with ease. This gives promise for rapid utilization of the soil without undergoing rigorous point to point soil sampling and testing. By integrating improved modeling techniques, freely accessible data and limited ground sampling, the approach presented in this study show hope in generating the necessary information for predicting soil physical properties.

5.2 Recommendations

There is more room for further studies given that there are inadequate soil physical properties databases developed. The study should be tested with high resolution data sources and other related models to improve the accuracy of the input data as well as prediction of the soil physical properties in a watershed. Further testing and worldwide application with different models in different soil types and watersheds is highly recommended.

6. REFERENCES

Bien J. D., J. Ter Meer, W. H. Rulkens and H. H. M. Rijnaarts. (2004). A GIS based approach for the long term prediction of human health risk at contaminated sites. Environmental Modeling and Assessment 9: 221-226.

Boonyanuphap J., Det W. and Katsutoshi S. (2004). GIS-based land suitability assessment for Musa (ABB group) plantation. J. Appl. Hort., 6(1): 3-10

Brink, A. B. A.., T. C. Partridge and A. A. B. Williams. (1984). Soil survey for engineering: monographs on soil and resources survey. Oxford University Press. New York, USA.

Brown J. D. (2007). Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd - order Uganda watershed. Geoderma 140.444-453.

Buol, S. W., F. D. Hole and R. J. McCracken. (1980). Soil genesis and classification. Iowa State University Press, Iowa, USA.

Calabro G., Leone A.P. and Amenta P. (2010). Prediction of soil properties from reflectance spectroscopy using multivariate statistical analysis and artificial intelligence methods: a comparative study. 11th European symposium on Statistical methods for the food industry 23-26 February 2010, Benevento.

Camo Technologies. (2007). The Unscrambler® v9.0 software, Aspen Corporate Park 1, 1480 Route 9 North, Suite 209, Woodbridge, NJ 07095, USA.

Chang, C. W., D. A. Laird, M. J. Mausbach and C. R.Hurburgh. (2001). Near-Infrared reflectance spectroscopy: Principal components regression analyses of soil properties. Soil Science Society of America Journal 65: 480–490.

Cohen, J. M., P. P. Joseph and F. D. William.(2005). Visible-near infrared reflectance spectroscopy for rapid nondestructive assessment of wetland soil quality. J Environ Qual 34:1422-1434 (2005).

David, J. B., K. D. Shepherd, M. G. Walsh, M. D. Mays and T. G. Reinsch. (2005). Global soil characterization with VNIR diffuse reflectance spectroscopy. Elsevier B.V. Online edition.

DETRL. (1974). Soil mechanics for road engineers. Ministry of Defense, England.

GARMIN International. (2002). GARMIN® 12 XL Hanheld GPS. Olathe, Kansas.

Haaland, D. M., V. T. Thomas. (1988). Partial least squares methods for spectral analyses. Analytical Chemistry 60, 1193 – 1202.

Heywood I., Sarah C. and Steve C. (2002). An introduction to geographical information systems. 2nd ed. Pearson education Limited, Edinburg Gate, England.

Faithful, N. T. (2002). Methods in agricultural chemical analysis: a practical handbook. CABI publishing, New York, USA.

Harpstead, M. I., T. J. Sauer and F. B. William. (1997). Soil science simplified 3rd Ed. Iowa State University Press, Iowa, USA.

Hazelton, P. A. and Murphy B.W. (2007). Interpreting soil test results: what do all the numbers means? (2nd Ed). Csiro publishing, oxford street, Australia.

Hengl, S., C. Kreutz, J. Timmer and T. Maiwald. (2007). Database identifiability analysis of non-linear models. Bioinformatics, 23 (19) 2612-2618

Janik, L. J., J. O. Skjemstad, K. D. Shepherd and L. R. Spouncer. (2007). The prediction of soil carbon fractions using mid-infrared partial least square analysis. Australian Journal of Soil Research, 45, 73-81.

Lal, R. and F. J.Pierce (Eds). (1989). Soil management for Sustainability. Soil and Water Conservation Society, Iowa, USA.

Keith D. S. and G. W. Markus. (2001). Light reflectance provides rapid assessment of soil quality". Natural resource problems, priorities and policies programme working paper 2001-1, International centre for research in agroforestry, Nairobi. Kenya.

Lambe, T. William. (1979). Soil mechanics: SI version Series in soil engineering. John Wiley and Sons, New York, USA.

Larlson, E. W. and P. C. Robert (1989). Soil management for sustainability: farming by soil in soil management for sustainability. Soil and Water Conservation Society, Iowa, USA.

Le, Y., B. D. Blavet, G. D. Noni, J. Y. Laurent, J. Asseline, and C. Chenu. (2007). Erodibility of mediterranean vineyard soils: relevant aggregate stability methods and significant soil variables. European Journal of Soil Science 58 (1), 188 195.

Liu, C and J. H. Evett. (1984). Soil properties: testing, measurements and evaluation. Prentice Hall, Englewood Cliffs, New Jersey.

Longley, A. Paul, M. F. Goodchild, D. J. Magure and D. W Rhind. (2005). Geographical information systems and science 2nd Edition. John Wiley Ltd, West Sussex, England.

Márquez, C. O., V. J. Garcia, C. A. Cambardella, R. C. Schultz and T. M. Isenhart. (2004). Aggregate-size stability distribution and soil stability. Soil Sci. Soc. Am. J. 68:725-735.

McBratney, A. B., S. M. L. Mendoca and B. Minasny. (2003). Digital soil mapping. Geoderma, 117 (1-2): 3-52.

Mc Graw-Hill Encyclopedia. (1977). Mc Graw-Hill, R. R. Donnelley and Sons company, Pennsylvania.

Mutio, P. K., D. S. Keith, A. Alain and C. George. (2005). Prediction of carbon mineralization rates from different soil physical fractions using diffuse reflectance spectroscopy. Soil Biology & Biochemistry 38 (2006) 1658–1664.

Nagaraj, H. B. (2005). Prediction of engineering properties of fine grained soils from their index properties. http//hdl.handle.net/2005/209.

Nanni, R. M. and J. A. M. Dematte. (2006). Spectral reflectance in comparison with traditional soil analysis. Soil Science society of America Journal, 70, 393-407.

Nielsen, D. R. and O. Wendroth. (2003). Spatial and temporal Statistics: sampling field soils and their vegetation. Reiskirchen, Germany.

Omuto, C. T. and D. P. Shrestha. (2007). Remote sensing techniques for rapid detection of soil Physical degradation. International Journal of Remote Sensing, Vol 28. 21: 4785–4805.

Omuto, C. T. and E. K. Biamah. (2008). Mapping infiltration and water retention characteristics using mixed – effect modeling and co-Kriging with infrared spectral reflectance. www.elsevier.com/locate/hydrol.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., in: http// www.cran.r-project.org).

Robinson, H. Arthur, Joel L. Morrison ,Philip C Muehrcke ,A.jon Kimerling and stephwen c guptill. (1995). Elements of cartography. John wiley & Sons inc., USA.

Savitzky, M. J. E. Golay. (1964). Smoothing and differentiation of data by simplified least-squares procedures. Anal. Chem. 36 (8) 1627-1639.

Schroeder, W. L. (1991). Soils in construction. John Wiley and Sons Ltd, New York, USA.

Scull, P., J. Franklin, O. A. Chadwick and D. McArthur. (2003). Predictive soil mapping: a review. Progress in Physical Geography 27, 2 (2003) pp. 171-197

Shepherd, D. K and M. G. Walsh. (2001). Light reflectance provides rapid assessment of soil quality. Natural Resources Problems, Priorities and Policies Programme Working Paper, International Center For Research In Agroforestry, Nairobi, Kenya.

Shepherd, D. K., C. A. Palm, C. N. Gachengo and B. Vanlauwe. (2003(1)). Rapid characterization of organic resource quality for soil and livestock management in tropical agroecosystems using near infrared spectroscopy. Agronomy Journal 95:1314-1322

Shepherd, D. K., M. G.Walsh, A. Awiti. (2003(2)). Use of soil spectral indicators for assessing and monitoring soil quality. Paper presented at ASA-CSSA-SSSA annual meetings, 2-6 November 2003, Denver, Colorado, USA. Annual Meeting Abstracts [CD-ROM]. ASSA, CSSA, and SSSA, Madison, WI. USA.

Shepherd, D. K., B. Vanlauwe , C.N. Gachengo and C. A. Palm. (2005). Decomposition and mineralization of organic residues predicted using near Infrared spectroscopy. Plant and Soil (2005) 277:315–333

Shepherd, D. K and M. G. Walsh. (2007). Infrared Spectroscopy: enabling evidence based diagnostic surveillance approach to agricultural and environmental management in developing countries. Journal of Near Infrared Spectroscopy 15, 1-19.

Singh. A. and G. R. Chowdhary. (1990). Soil engineering in theory and practice 2nd edition: geotechnical testing and instrumentation. CBS Publishers and Distributors Pvt Ltd. New Delhi. India.

Sowers. B. G. and S. F. George. (1961). Introductory soil mechanics and foundations 2nd edition. Macmillan Company. New York. USA.

Spancler. G. Merlin. (1951). Soil engineering. International textbook Company. Scranton. USA

Stuart. B. (2004). Infrared spectroscopy: fundamentals and applications. John Wiley and Sons Ltd. New York. USA.

Terzaghi. K. and R. B. Peck. (1962). Soil mechanics in engineering practice. John Wiley and Sons. New York. USA.

Thine. C. O., M. G. Walsh. K. D. Shepherd. R. Coe. (2003). Prediction of field-measured infiltration rates using diffuse reflectance spectroscopy. Poster presented at ASA-CSSA-SSSA annual meetings. 2-6 November 2003. Denver. Colorado. USA. Annual Meeting Abstracts [CD-ROM]. ASSA. CSSA. and SSSA. Madison. WI. USA.

Toktam. T., Shamsollah A., Shaban S. and Kanwar L.S. (2010). Soil surface salinity prediction using ASTER Data: comparing statistical and geostatistical models. Australian Journal of Basic and Applied Sciences. 4(3): 457-467. 2010

Reeves. J. B and G. W. McCarty. (2001). Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and a fibre-optic probe. J. Infrared Spectrosc.9. (25-34).

Yong. H., S. Hai-yan. G. P. Annia. and A. Gómez. (2005). Measurement and analysis of soil nitrogen and organic matter content using near-infrared spectroscopy techniques. J Zhejiang Univ Sci B. 6(11): 1081–1086.

7. APPENDICES

Appendix 1: Program in R for spatial prediction of soil liquid limit

```
#1. upload the packages
library(foreign)
library(nlme)
library(sp)
library(maptools)
library(spatstat)
library(mgcv)
library(rgdal)
library(mgcv)
library(gstat)


#2. import the sampled data. Give it a name e.g. "soilq.txt" as shown below
soilq=read.table("soilq.txt", header=T)
str(soilq)
points=soilq


#3. upload the packages for mapping
library(sp)
library(maptools)
library(spatstat)
library(mgcv)
library(rgdal)
library(Hmisc)
library(gstat)
library(colorspace)
library(class)
library(MASS)
library(grid)
library(nnet)
library(mda)
library(vcd)
library(nlme)

#4. Convert the samples into map layer
coordinates(points)=~X+Y
```

```
#7 Fit linear model
liquidlimit=lm(ll~(ydist+dem+landuse),points)
cor(fitted(fc.lm),points$ll)^2
summary(fc.lm)
hist(residuals(liquidlimit))


#8. Fit the ordinary variograms for the residuals
plot(variogram(residuals(fc.lm)~1, points), plot.nu=T,pch="+")
fc.v=variogram(residuals(fc.lm)~1, points)
fc.ovgm=fit.variogram(fc.v, vgm(nugget=0.0, model="Gau",range=10200, sill=0.0071))
plot(fc.v, fc.ovgm, plot.nu=F, xlab="Distance (metres)", ylab="Variance (vol/vol)^2",
col="black")
str(fc.ovgm)

#9. Fit regression Kriging models for spatial mapping of the attributes
```

```
proj4string(points)=CRS("+proj=utm +ellps=WGS84")
str(points)

#5. import predictors for the data (make sure they are in UTM and ASCII formats
predictors=readGDAL("xdist.asc")#This is X coordinate
predictors$ydist=readGDAL("ydist.asc")$band1#This is Y coordinate
predictors$dem=readGDAL("dem.asc")$band1# This is DEM
predictors$landuse=readGDAL("landuse.asc")$band1 # This is land use
str(predictors)
predictors$ydist=predictors$band1
predictors$band1=NULL
object.size(predictors)
proj4string(predictors)=CRS("+proj=utm +ellps=WGS84")
str(predictors)


#6. overlay the predictors
predictors.ov=overlay(predictors, points)
points$ydist=predictors.ov$ydist
points$dem=predictors.ov$dem
points$xdist=predictors.ov$xdist
points$landuse=predictors.ov$landuse
str(points)
```

```
fc.rev=variogram(residuals(fc.lm)~1, points)
plot(fc.rev)
fc.rvgm=fit.variogram(fc.rev, vgm(nugget=0.0,model="Gau",range=10200, sill=0.0071))
plot(fc.rev, fc.rvgm, plot.nu=F)
fc.rk=krige(residuals(fc.lm)~1, points, predictors, fc.rvgm)
str(fc.rk)
predictors$lanfc=predict(fc.lm, predictors)
predictors$Pred=predictors$lanfc+fc.rk$var1.pred
str(predictors)



#10. Produce the maps of teh attributes for export
fc.rkpred.plt=spplot(fc.rk["var1.pred"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), sp.layout=list("sp.points",pch="+",col="black", fill=1, points))
fc.rkvar.plt=spplot(fc.rk["var1.var"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), at=seq(0.2,0.35,0.002),sp.layout=list("sp.points",pch="+",col="black", fill=1,
points))
predictors$lnfc1=predictors$Pred+fc.rk$var1.pred
str(predictors)
fc.rkpred.plt=spplot(predictors["Pred"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), at=seq(0.2,0.35,0.002),sp.layout=list("sp.points",pch="+",col="black", fill=T,
points))
print(fc.rkpred.plt, split=c(1,1,2,1), more=FALSE)
#print(fc.rkvar.plt, split=c(2,1,2,1), more=FALSE)

#11.     Validate the prediction
pointas=read.table("peri.txt", header=T)
coordinates(pointas)=~X+Y
proj4string(pointas)=CRS("+proj=utm +ellps=WGS84")
predictors.ov=overlay(predictors, pointas)
pointas$Pred=predictors.ov$Pred
str(pointas)
cor(pointas$Pred,(pointas$ll))^2
plot((pointas$Pred)^2~(pointas$ll), xlab="Measured Soil Liquid limit", ylab="Predicted Soil
Liquid limit")
write.table(pointas, file="valid.txt")

#export to ASCII
write.asciigrid(predictors["Pred"], "predll.asc")
write.asciigrid(fc.rk["var1.var"], "predllvar.asc")
```

Appendix 2: Program in R for spatial prediction of soil plastic limit

```
#1. upload the packages
library(foreign)
library(nlme)
library(sp)
library(maptools)
library(spatstat)
library(mgcv)
library(rgdal)
library(mgcv)
library(gstat)
```

```
#2. import the sampled data. Give it a name e.g. "soilp.txt" as shown below
soilp=read.table("soilp.txt", header=T)
str(soilp)
points=soilp
```

```
#3. upload the packages for mapping
library(sp)
library(maptools)
library(spatstat)
library(mgcv)
library(rgdal)
library(Hmisc)
library(gstat)
library(colorspace)
library(class)
library(MASS)
library(grid)
library(nnet)
library(mda)
library(vcd)
library(nlme)
```

```
#4. Convert the samples into map layer
coordinates(points)=~X+Y
proj4string(points)=CRS("+proj=utm +ellps=WGS84")
str(points)
```

```
#5. import predictors for the data (make sure they are in UTM and ASCII formats
predictors=readGDAL("xdist.asc")# This is X coordinate
predictors$ydist=readGDAL("ydist.asc")$band1# This is Y coordinate
predictors$dem=readGDAL("dem.asc")$band1# This is DEM
predictors$landuse=readGDAL("landuse.asc")$band1 # This is land use
str(predictors)
predictors$ydist=predictors$band1
predictors$band1=NULL
object.size(predictors)
proj4string(predictors)=CRS("+proj=utm +ellps=WGS84")
str(predictors)


#6. overlay the predictors
predictors.ov=overlay(predictors, points)
points$ydist=predictors.ov$ydist
points$dem=predictors.ov$dem
points$xdist=predictors.ov$xdist
points$landuse=predictors.ov$landuse
str(points)


#7 Fit linear model
PlasticLimit=lm(pl~(ydist+dem+landuse),points)
cor(fitted(fc.lm),points$pl)^2
summary(fc.lm)
hist(residuals(PlasticLimit))


#8. Fit the ordinary variograms for the residuals
plot(variogram(residuals(fc.lm)~1, points), plot.nu=T,pch="+")
fc.v=variogram(residuals(fc.lm)~1, points)
fc.ovgm=fit.variogram(fc.v, vgm(nugget=0.0, model="Gau",range=10200, sill=0.0071
plot(fc.v, fc.ovgm, plot.nu=F, xlab="Distance (metres)", ylab="Variance (vol/vol)^2",
col="black")
str(fc.ovgm)

#9. Fit regression Kriging models for spatial mapping of the attributes
fc.rev=variogram(residuals(fc.lm)~1, points)
plot(fc.rev)
fc.rvgm=fit.variogram(fc.rev, vgm(nugget=0.0,model="Gau",range=10200, sill=0.007
```

```
plot(fc.rev, fc.rvgm, plot.nu=F)
fc.rk=krige(residuals(fc.lm)~1, points, predictors, fc.rvgm)
str(fc.rk)
predictors$lanfc=predict(fc.lm, predictors)
predictors$Pred=predictors$lanfc+fc.rk$var1.pred
str(predictors)



#10. Produce the maps of teh attributes for export
fc.rkpred.plt=spplot(fc.rk["var1.pred"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), sp.layout=list("sp.points",pch="+",col="black", fill=T, points))
fc.rkvar.plt=spplot(fc.rk["var1.var"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), at=seq(0.1,0.7,0.02),sp.layout=list("sp.points",pch="+",col="black", fill=T, points))
predictors$lnfc1=predictors$Pred+fc.rk$var1.pred
str(predictors)
fc.rkpred.plt=spplot(predictors["Pred"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), at=seq(0.1,0.25,0.02),sp.layout=list("sp.points",pch="+",col="black", fill=T,
points))
print(fc.rkpred.plt, split=c(1,1,2,1), more=FALSE)
#print(fc.rkvar.plt, split=c(2,1,2,1), more=FALSE)

#11.     Validate the prediction
pointas=read.table("peri.txt", header=T)
coordinates(pointas)= ~X+Y
proj4string(pointas)=CRS("+proj=utm +ellps=WGS84")
predictors.ov=overlay(predictors, pointas)
pointas$Pred=predictors.ov$Pred
str(pointas)
cor(pointas$Pred,(pointas$pl))^2
plot((pointas$Pred)^2~(pointas$pl), xlab="Measured Soil Plastic Limit", ylab="Predicted soil
Plastic Limit")
write.table(pointas, file="valid.txt")

#export to ASCII
write.asciigrid(predictors["Pred"], "predpl.asc")
write.asciigrid(fc.rk["var1.var"], "predlnvar.asc")
```

Appendix 3: Program in R for spatial prediction of soil aggregate stability

```
#1. upload the packages
library(foreign)
library(nlme)
library(sp)
library(maptools)
library(spatstat)
library(mgcv)
library(rgdal)
library(mgcv)
library(gstat)

#2. import the sampled data. Give it a name e.g. "solai.txt" as shown below
porosity=read.table("solai.txt", header=T)
str(solai)
points=solai

#3. upload the packages for mapping
library(sp)
library(maptools)
library(spatstat)
library(mgcv)
library(rgdal)
library(Hmisc)
library(gstat)
library(colorspace)
library(class)
library(MASS)
library(grid)
library(nnet)
library(mda)
library(vcd)
library(nlme)

#4. Convert the samples into map layer
coordinates(points)=~X+Y
proj4string(points)=CRS("+proj=utm +ellps=WGS84")
str(points)

#5. Import predictors for the data (make sure they are in UTM and ASCII formats
predictors=readGDAL("xdist.asc")#This is X coordinate
```

```
predictors$ydist=readGDAL("ydist.asc")$band1# This is Y coordinate
predictors$dem=readGDAL("dem.asc")$band1# This is DEM
predictors$landuse=readGDAL("landuse.asc")$band1 # This is land use
str(predictors)
predictors$ydist=predictors$band1
predictors$band1=NULL
object.size(predictors)
proj4string(predictors)=CRS("+proj=utm +ellps=WGS84")
str(predictors)

#6. overlay the predictors
predictors.ov=overlay(predictors, points)
points$ydist=predictors.ov$ydist
points$dem=predictors.ov$dem
points$xdist=predictors.ov$xdist
points$landuse=predictors.ov$landuse
str(points)

#7. Fit linear model
aggregateindex=lm(ai~(ydist+dem+landuse),points)
cor(fitted(fc.lm),points$ai)^2
summary(fc.lm)
hist(residuals(aggregateindex))

#8. Fit the ordinary variograms for the residuals
Page 2
plot(variogram(residuals(fc.lm)~1, points), plot.nu=T,pch="+")
fc.v=variogram(residuals(fc.lm)~1, points)
fc.ovgm=fit.variogram(fc.v, vgm(nugget=0.0, model="Gau",range=10200, sill=0.0071))
plot(fc.v, fc.ovgm, plot.nu=F, xlab="Distance (metres)", ylab="Variance (vol/vol)^2",
col="black")
str(fc.ovgm)

#9. Fit regression Kriging models for spatial mapping of the attributes
fc.rev=variogram(residuals(fc.lm)~1, points)
plot(fc.rev)
fc.rvgm=fit.variogram(fc.rev, vgm(nugget=0.0,model="Gau",range=10200, sill=0.0071))
plot(fc.rev, fc.rvgm, plot.nu=F)
fc.rk=krige(residuals(fc.lm)~1, points, predictors, fc.rvgm)
str(fc.rk)
predictors$lanlc=predict(fc.lm, predictors)
```

```
#export to ASCII
write.asciigrid(predictors["Pred"], "predai.asc")
write.asciigrid(fc.rk["var1.var"], "predbvar.asc")
```

```
predictors$Pred=predictors$lnfc+fc.rk$var1.pred
str(predictors)

#10. Produce the maps of teh attributes for export
fc.rkpred.plt=spplot(fc.rk["var1.pred"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), sp.fc.rkvar.plt=spplot(fc.rk["var1.var"],
col.regions=bpy.colors(),scales=list(draw=TRUE, cex=0.7),
at=predictors$lnfc1=predictors$Pred+fc.rk$var1.pred
str(predictors)
fc.rkpred.plt=spplot(predictors["Pred"], col.regions=bpy.colors(),scales=list(draw=TRUE,
cex=0.7), at=print(fc.rkpred.plt, split=c(1,1,2,1), more=FALSE)

#print(fc.rkvar.plt, split=c(2,1,2,1), more=FALSE)

#11Validate the prediction
pointas=read.table("peri.txt", header=T)
coordinates(pointas)=~X+Y
proj4string(pointas)=CRS("+proj=utm +ellps=WGS84")
predictors.ov=overlay(predictors, pointas)
pointas$Pred=predictors.ov$Pred
str(pointas)
cor(pointas$Pred,(pointas$ai))^2
plot((pointas$Pred)^2~(pointas$ai), xlab="Measured soil aggregate stability indices",
ylab="Predicted write.table(pointas, file="valid.txt")
```