THE USE OF MULTIVARIATE CLASSIFICATION IN THE DEFINITION OF SOIL SERIES IN LAKE NAIVASHA AREA

By

SIMFUKWE PAUL

February, 2001
Enschede, The Netherlans

THE USE OF MULTIVARIATE CLASSIFICATION IN THE DEFINITION OF SOIL SERIES IN LAKE NAIVASHA AREA

By

Paul Simfukwe

SUPERVISOR

Assoc. Prof. Dr. D. G. Rossiter

CO-SUPERVISOR

Dr. W. Siderius

Thesis submitted to the International Institute for Aerospace Survey and Earth Sciences (ITC) in partial fulfilment of the requirements for the degree of Master of Science in Geoinformation Science and Earth Observation, with specialisation in Land Degradation Conservation and Rehabilitation, Enschede, The Netherlands.

Degree Assessment Board
Assoc. Prof. Dr Rossiter (Head of Division)
Prof. Dr. Ir. A.K. Bregt (External examiner)
Dr. O.C. Spaargaren (Internal Examiner)
Dr. D.G. Rossiter (Main supervisor)
Dr. W. Siderius (2^{nd} Supervisor)
Dr. A. Farshad (Student advisor)

INTERNATIONAL INSTITUTE FOR AEROSPACE SURVEY AND EARTH SCIENCES ENSCHEDE, THE NETHERLANDS

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Aerospace Survey and Earth Sciences. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Dedication

A dedication to my lovely fiancée Kamo Msimuko
Thank you for loving me, I love you too.
And to my family

ABSTRACT

The knowledge of soil properties and their response to management practices is an essential requirement by any land user. Much of the Lake Naivasha area is under intensive horticultural and flower cultivation. Such land is susceptible to degradation when badly managed. Soil information is therefore on demand by a wide range of land users including farmers, planners, conservationists, developers, engineers, lenders and investors in the area. In other intensively used areas of the world, the soil series has proven to be an appropriate information carrier for particular management options. Soil series are relatively homogeneous units enabling specific management decisions, and typically occupy well-defined landscape positions and form compact spatial entities.

An unequal nested sampling design was adopted for soil characterisation and determination of the spatial variability of the soils in the area. Six centres with 600m spacing were established along a catena in the upper lacustrine and low-mid volcanic plains in the SULMAC farm. Closer sub-centres were established, with the geometric series with a basis of 600m, namely 190m, 60m, 19m, and 6m. This yielded 72 sampling points whose important soil properties were determined. The depth to the weathered zone, the thickness of the A-horizon, the depth of the B-horizon and the gravel percentage in the A and B-horizon were used to perform the multivariate analysis.

The data was analysed by descriptive statistics, one-way ANOVA, clustering of observations, and fully nested ANOVA. All observations were grouped into clusters in property space by similarity, using the minimum-distance-to-class-centroid method. The result was a dendrogram of different clusters. The spatial distribution of these clusters was investigated by mapping and by comparaison with geo-pedological maps units. Spatial variability was investigated by the fully nested ANOVA.

Soil series were defined by two approaches: based on map units defined by geo-pedological mapping, and by multivariate cluster analysis. In the first approach, observations were grouped by map unit and their central concepts and range were established. This resulted in three proposed series. In the second approach to defining series, multivariate analysis defined clusters of individuals which are generally similar to one another, and which form statistically different classes. The clusters are relatively homogeneous and can be used to define individual series with unique central concepts. However, these proposed series are not always mappable. In some cases, observations from different clusters were intimately mingled. Other clusters formed compact units, but these were smaller than the minimum legible delineation at detailed map scale (1:10 000).

The fully-nested ANOVA showed that the range of spatial dependence of soil properties varied from very short, i.e. <6m to 19m (thickness of the A horizon) through medium, i.e. 60m to 190m (weathering depth) to moderately long, i.e. 60m to 600m (gravel content). The longer-range properties should be more useful in defining mappable series.

All the series defined by both approaches fell within the boundaries of one class at the lowest levels of two hierarchical taxonomic systems. The soils classified as *Ashy-pumiceous Mixed, Isothermic Vitriandic Haplustept,* and *Ashy, Glassy, Isothermic Aridic Ustpsammen*t (USDA family).

Synthesising the two approaches, three series were proposed, all in the same family: The Naivasha, the Sulmac, and the Longonot series. Thus soil series that are homogenous enough for intensive management can be defined in Lake Naivasha area.

ACKNOWLEDGEMENT

I would like to thank all the organisations and persons who in one way or another contributed to the successful writing of this thesis. Space and time would not allow me list them all, save for the following.

Above all I thank God for giving me this opportunity and strength to do this program. Glory to the most high.

In the first place I would like to thank and express my appreciation to the Associate Prof. Dr. D G Rossiter for the guidance rendered towards this research from the start to the end. His comments and suggestions were constructive and timely and always encouraging. Many thanks also to Dr. W Siderius and Dr A Farshad for their comments and advice through out my writing of this thesis.

Many thanks to Dr C M M Mannaerts, and Drs R Becht for their field work arrangements and guidance in Kenya. My special thanks to Mr R Hennemann for his invaluable input in the logistical arrangements, help and advice before, during and after fieldwork and also for the laboratory arrangements and follow-ups at ISRIC.

Thanks to Sulmac Co. management for the permission to work on their farm and the co-operation and assistance I received during my field data collection. Thanks for the transport and laboratory facility. I am greatly indebted particularly to Ruth Vaughan (Technical Manager) and Monica, who went an extra mile making sure I lacked nothing during my data collection at their farm. Thanks to Martin Otieno Agak (field Assistant) and Joseph M. Waithanji (driver) who patiently worked with me even overtime to make sure I got what I needed.

I would like to thank the Kenya Wildlife Service, Hells Gate National Park for their permission and co-operation to work in their area. Thanks to Sara Higgins, the secretary of the Lake Naivasha Riparian Owners Association, who prepared an introductory letter to enable me to work in the area smoothly. Thanks to Mr. Van Oostrum and ISRIC staff at ISRIC laboratory for working tirelessly on my samples.

Many thanks to all my cluster mates Andre Nagelhout (Netherlands), Artkilt G, (Ethiopia) Disanayakamudianselage Bandula Ranatunga (Sri Lanka), Fabiola Padilla (Bolivia), Hoang (Vietnam) and Tilaye B. B (Ethiopia) for the co-operation, both in field and office.

Finally by all means not the least, I thank Kamo my lovely Fiancée for her love letters. Thank you for your love, your emotional, spiritual and moral support I have received from you during my absence from you. I am sorry for being away, may this write-up bring some joy. Thanks to my family and friends for their love and support.

May The Lord Bless You All.

"…I will strengthen you and help you; I will uphold you with my righteous right hand." (Isaiah 41: 10).
"When Anxiety was great within me, your consolation brought joy to my soul" (Psalm 94: 10).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: INTRODUCTION

1.1 Introduction

1.1.1 Knowledge of soils

An assessment of the soil properties and their response to management is required in agriculture and forestry, for informed decision making in rural and urban planning for feasibility and design studies in land development projects and for many engineering works. Soil information is in great demand by a wide range of users including farmers, planners, conservationists, developers, engineers, lenders and investors in the study area. "The practical purpose of soil surveys is to enable more numerous, more accurate and more useful predictions to be made for specific purposes than could have been otherwise" (Dent and Young, 1980).

A soil map delineates areas occupied by different kinds of soils, each of, which has unique sets of interrelated properties and characteristics of the material from which it was formed, its environment and history. The geographical distribution of many individual soil properties or soil qualities can be extracted from soil maps for special purposes such as showing predicted soil behaviour for particular use. The objective of the soil survey is the same for all kinds of land i.e. to produce a map of high predictability, though the number of mapping units with their composition and detail (and scale) varies with complexity of soil pattern and the specific needs of users.

1.1.2 High Cost of surveys

Systematic soil survey is a lengthy and costly operation. It involves university-educated specialists, field and laboratory equipment and determinations, several source documents such as remotely sensed data -, which altogether contribute to making a soil information an expensive product. Opportunity cost is often high and decisions must be made without it. However the value of the multipurpose soil interpretation from the same basic survey map is not yet really evaluated, which in fact could decrease the cost per unit area (Zink, 1992).

1.1.3 Constraints to reckon with in modern challenges

The rapid development of information technology such as GIS, and the increasing use of modelling procedures introduce some constraints on soil data. For such operations to be applied, more and better (quality) data is needed. The enhancement of the pedotransfer functions can only be possible and efficient with the availability of an appropriate and accurate primary data. Often the danger is that sophisticated techniques and technologies will be applied on poor data sets, which will diminish the quality of data interpretation and land use decision making (Zink, 1992).

With regard to the digital soil mapping, users in the developing countries complain of the commercially motivated hardware and software developers. Equipment and programmes become out dated in a short time and their restrictive budgets cannot allow them to adjust in pace.

1.1.4 Demand-supply problems in soil survey

The users often demand (from the producer) for soil maps of high accuracy with very little effort. In reality soil boundaries and mapping units are complex, and not always visible. And map units are not that homogenous, though a relatively accurate map can be made but the task is time demanding and expensive. The user therefore must be willing to pay for a good soil map.

Sometimes the users demand for a nicely and uniquely graded land from the 'best' to the 'poorest' quality. It must be understood that what may be bad for one use may be the best for another. Soils and other land requirements are specific for each use.

In some cases the producer is expected to tell the user what to do with the land, which is not his task, though he can suggest options. Soil surveys and evaluation studies only provide information to certain aspects of land management decisions. The role of the producer (Butler, 1977) is merely to predict, that is to provide a means or a tool with which the problem solver can inform himself about the relevant soil conditions at the site in the area.

On the other hand, in some cases the producer is not sufficiently trained to meet the challenges of producing the desired soil information. This is especially true in the developing countries where there are less commercial and intensive land users and less motivation from their government's restrictive budgets in training staff.

Generally drawbacks also affect soil mapping, inasmuch as the spatial-vertical and spatial-lateral distribution of soils are not always properly displayed. Soils in nature are three-dimensional continuum, but cartographically, discrete units are aggregated to produce a mere human abstraction (Ibanez, J. J. etal).

Not withstanding the above problems this research tries to develop the soil series concept in the lake Naivasha area in Kenya. It is an attempt to provide the needed soil information at a level where relatively precise management decisions could be made. The soil series concept will be highly useful among the highly developed farmers who practice ranching, dairy farming, irrigation, horticulture and flower cultivation in the area.

A multi-purpose survey approach was carried out which is expected to provide the basis of interpretation for many different uses, some of which may not be known at present. And the use of the multivariate classification, the combination of both numerical and the conventional qualitative classification approaches in the development and proposition of the series will be handy. This is also an attempt in the application of new concepts, methods, techniques and technologies to soil survey as conducted on an operational basis in real institutional context.

1.2 Research Problem

There is a high demand for soil information by land users in the Lake Naivasha area, some of which is latent, therefore there is a great need for detailed soil information to be made available in the area. This soil information is for both to know where particular management options should be exercised to make the best use of the area and to avoid expensive failures especially that this area is prime for horticulture and flower cultivation. At the same time the soils in this area are susceptible to degradation (e.g. soil erosion) when badly managed. The soil series could be the desired level of detail to satisfy the demand. However it is an open question, for Lake Naivasha, if series can be defined, and if so, at what scale they can be mapped as consociations or associations.

1.3 Research Questions

- Is a 1:10,000 scale soil series map justifiable with regard to its homogeneity
- What are the realistic class limits for properties that define soil series?
- Do series defined by the multivariate analysis always fall within categorical limits of higher-level classification?

Q What is the spatial variability of the soils in the area?

1.4 Hypothesis

1. Soil series that are homogenous enough for intensive management can be defined in Lake Naivasha area
2. Soil series are more homogeneous than (narrow property class limits) higher level of 2nd level WRB and the US Soil Taxonomy family classification, and the majority of the pedons classified to a series also classify to one higher level.
3. Soil series can be mapped as consociations at 1: 10 000, i.e. in polygons of minimum size 0.4ha. That is, the named series and similar soils occupy at least 75% of the mapped area; no more than 15% of the area can contain soils that differ significantly from the named soil and which are more limiting to major land uses.
4. Series mapped by the multivariate analysis (numerical classification) do not always fall within the range of the hierarchical classification (Taxonomic Systems).
5. The spatial variability of soil properties varies within short distances, but mostly within the range of a series.

1.5 Objectives

1.5.1 General Objective

The general objective is to test the series concept, as to how much data is needed to define them and to produce a soil map with high predictability of soil attributes at series level using modern classification methods, specifically numerical classification and cluster analysis.

1.5.2 Specific Objectives

1. To produce a detailed soil map at the scale of 1:10,000, using the geopedologic approach, which will in turn be a soil series map.
2. To collect a multiple of soil attributes (multivariate properties) and by using numerical classification and cluster analysis, assign class limits for important attributes to be used to differentiate soils series.
3. To establish to what degree the soil series defined by multivariate analysis fit within class limits of higher Taxa in both the WRB and the USDA Taxonomic system, i.e. is there a 'taxonomic hiatus' in the concepts?
4. To construct the field key to the series proposed
5. To propose the soil series in the study area
6. To establish the nature of the spatial variability of the soils in the area.

Chapter 2: Literature Review

2.1 The Soil Series Concepts

2.1.1 Evolution Of The Soil Series Concept

The soil series was introduced in 1903 (Yaalon and Berkowicz, 1997), primarily to relate soils of one survey to those of another. A soil series was to consist of the all soil types that were formed on the same regoliths derived in the same way during the same period of time. The soil series were meant to group soils of similar genesis. Soil series could cover soil types with the full range of gravel to clay inclusive.

By 1899 the soil type concept was in use and was considered as a label of the mapping units in the US. According to the Bureau of Soils in 1902, the soil types were localised soils that differ in one or both crop productions and yields. They were place names with a texture term in it. The soil types were understood as localised kinds of soils developed from the same genetic origin whose distinctions were made by differences in the crop adaptations or yields, on conspicuous difference in colour and rock bed depths.

Soil series gradually replaced soil type as the pre-eminent class in soil surveys. Soil types then became subdivision of the series based on texture, later restricted to texture of the surface layer. In 1950's the soil type was dropped as a category in the system. Instead, separations base on textures of the surface were to be recognised as texture phases, paralleling those of slope, erosion, and stoniness. In the 1920's and 1930's the series concept consisted of a set of soils closely similar in morphology and composition. Presumably the soils were also similar in genesis though the understanding of it had changed greatly in three decades. The emphasis was on the character of the soil profile rather than on the nature and origin. This concept prevailed up to 1960's.

Then came the third concept outlined in Soil Taxonomy in 1992, which was first outlined in the 7th approximation (Soil survey staff, 1960) and provided that the soil series were to record pragmatic distinctions, i.e., to be keyed to the soil usefulness. No emphasis was made on the genesis. Series were subdivision of families, classes of the next higher rank in a multiple-category system.

From the time the series was concept introduced in 1903, the total number recognised in the US has increased irregularly but continuously. The increased numbers have been due to additions in the criteria and to the narrowing ranges of properties permitted within series, and also in the increase in the mapping areas. Between 1903 and 1992, the numbers have changed as follows; 190 in 1904, 534 in 1912, 2000 in 1938, 5500 in 1951, and 14,200 in 1992. (Yaalon and Berkowicz, 1997).

2.1.2 Two Modern Concepts

There are two very important concepts of soil series worth consideration here;

2.1.3 By Dent and Young

The first is that of Dent and Young (1981), who consider soil series as groupings of soils with the same sequence of horizons developed on similar parent material under similar environmental or external conditions, exhibiting likeness in their characteristics and behaviour in the landscape. Soils within a series are developed on the same parent material in the sane environment and have profiles that are almost alike, with horizons that are similar in their vertical sequence, thickness and morphological properties.

Here the concept defines the series in terms of the site characteristics and medial profile with a restricted range of variability with respect to the horizons present, their thickness, colour, structure, presence of carbonates or soluble salts, mineralogy and sometimes other properties. Each is intended to have unique combination of site features and morphological characteristics, which can be assessed, for the most part, in the field.

Once the series have been defined and mapped according to its differentiating characteristics, many accessory properties of practical importance can be predicted.

2.1.4 By USDA

The second is that of the USDA Natural Resources Conservation Service (NRCS), that is, the American soils survey. In the USDA system, the series is part of the Soil Taxonomy and is the lowest category whose differentiae are mostly the same as that for higher categories, but the limits and ranges permitted in one or more properties is more restrictive than is permitted in the family or in some higher category.

In the Soil Survey Manual (Soil Survey Staff, 1993, p. 20) the series concept is amplified as follows; "The soil series category is the most homogenous category in the taxonomy used in the United States. As a class, a series is a group of soils or polypedon that have horizons similar in arrangement and in differentiating characteristics. The soils of a series have a relatively narrow range in sets of the properties. The surface layer and such features as slope, stoniness, degree of erosion, and topographic position may vary unless these factors are associated with significant differences in the arrangement of horizons. Soil series are differentiated on all the higher categories plus those additional and significant characteristics in the series control section. Some of the characteristics commonly used to differentiate series are the kind, thickness, and arrangement of horizons and their structure, colour, texture, reaction, consistence, content of carbonates and other salts, content of humus, content of rock fragments and mineralogical composition. A significant difference in any of one these can be the basis for recognising a different series. Very rarely, however, do two soils series differ in just one of these characteristics. Most characteristics are related and generally several change together."

2.1.5 What's the Difference?

The basic difference in the two concepts is that the former is a bottom–up while the latter is a top-down definition. In this study the Dent and Young's definition of the series was followed as far as possible in addition to the conventional classification. In the bottom-up approach, soils are grouped into 'natural' classes whose limits are set by soil forming processes.

Under both definitions, the soil series names are usually place names, commonly local people's or place names and are usually used for detailed soil mapping. This serves to communicate detailed information about the soils at specific places which in-turn can aid transfer of research knowledge and technology from one place to another especially that at this level definite, specific statements about the use and response to management of the soil can be made. Soil series are also useful units both for general-purpose interpretation and as a basis for research on soil-plant relationship.

2.1.6 Central Concept: Representative Profile

"The series should have a central concept of the 'typical' pedon; this is used by the surveyor and land user alike to reduce the complexity. Often modellers use this central concept to represent the entire mapped area. And if we believe that specific combinations of soil forming factors lead to specific soil properties, and that certain combinations of soil-forming factors are 'typical' in a certain landscape, we believe, in the 'typical' result…" (Rossiter DG, 2000. P. 68).

The 'modal profile' can be synthetic, where actual observations are combined based on the surveyor's experience and information on the series in a database or it can be an actual representative profile selected to represent the series. Presumably the experienced correlators should select this representative profile as a good example of the series.

2.1.7 Official Soil Series Descriptions

Each soil series must be defined as fully and accurately as existing knowledge permits, whether it is a proposed series in an individual survey or an established one. In the Soil survey manual, by Soil Survey Staff (1993, p. 45), there is a standard format for recording specific kind of information to be used to ensure inclusion of essential information and to enable comparisons of series definitions.

The official soil series descriptions record definitions of the series and other relevant information about each series. The detailed definition and series interpretation record are essential. The general descriptive information is also needed to aid the reader to identify the soil in the landscape and relating to it other soils.

According to the Soil Survey Manual, an official series description should include at least the following:

1. Full Taxonomic name of the family taxon for which it is a member. This indicates the classes that provide limits of properties that are diagnostic for the series at all categorical levels, except for those between series of the same family.
2. A description of a typical pedon and its horizons, describing each in as detail as necessary to recognise its taxonomic class. Horizons that are diagnostic for the pedon must be described.
3. A statement of the ranges of properties of the series. This section also contains statements about the relationship of the series control section and diagnostic horizons to vertical subdivisions of the typical pedon
4. A statement distinguishing the series from "competing series" with which it might be confused. Competing series are mainly those that share common limits with the series described or are members of the same family.
5. A statement that identifies at least one specific place that represents a norm for the series - a "type location." A type location should be described accurately enough so that it can be located in the field.

2.2 Numerical Classification

More and more measurements are being made now at many sites to describe the soil and other land resources. So many data can be difficult to interpret and comprehend, and classification often helps us to provide a simpler picture. "It is also convenient to be able to talk about one or more groups of individuals using names rather than lists of characters and their values. Further a classification might enable an investigator to economise on expensive measurements." (Webster and Oliver, 1990. p. 168).

Webster and Oliver, (1990) describes a simple numerical way of classifying soil from measured values. The process is called dissection. In this process, the measured range of properties of interest is divided at critical or convenient points. If two or three properties are judged to be important then all their scales can be divided to produce a classification still with manageable few groups. When many groups are relevant, however, simultaneous dissection of every scale is not feasible: far too many groups result. Another disadvantage is that it takes no account of the natural discontinuities in the population.

In the ‘natural’ classification, a suitable dividing point needs to be discovered. In dealing with multivariate classification, its easier to think of those individuals in a more closely occupied parts of the character space as constituting clusters which are hoped to be isolated in the classification. The second aspect of the concept of ‘natural’ classification is that individuals in any one group are generally similar to one another. Members of the ‘natural’ group should be similar in many respects, useful especially to population possessing many characters of interest. Such a grouping should also be useful for many purposes. For it to be so, the choice of the properties in the numerical classification, should tell both the intrinsic and the extrinsic variables of the soil (de Gruiter J J, 1977).

2.3 Multivariate Classification

The concept of multivariate grouping is one in which the individuals share many attributes, but for which no single attribute is either sufficient or necessary to confer class membership. Such a grouping is not confined to formal taxonomy; they are a norm in everyday life and language. Webster and Oliver (1990), states that though this grouping may be generally useful, they cannot expect to be the most useful for any particular purpose; indeed they may not be useful for any desired purpose because the classes may not conform to any practical difference. Another problem is that it can be very difficult to create keys for identifying their members and for allocation of the new members.

Most attempts to classify individuals mathematically or numerically have sought to create classes within which the different members are generally alike and substantially different from the members of other classes. Such methods for creating classifications are known by the general name of cluster analysis. Their aim according to Webster and Oliver (1990) is to identify clusters in the population. However they say that the name is misleading because most methods will create classes whether or not clusters exist in any true sense.

2.3.1 Hierarchical Classification

Webster and Oliver describes the hierarchical classification as one in which individuals belong to small groups, the small groups belong to larger groups and so on. It may also be thought as a division of character space in which the whole space is divided into smaller compartments, which in turn are divided into yet smaller compartments. Usually the groupings are done at a few distinct levels of generalisation, known as categories. Generally in any one category the classes are disjoint, since any appreciable overlap of classes is incompatible with a hierarchy. The method of creating hierarchies are either agglomerative, putting individuals together into larger and larger groups, or divisive, creating smaller and smaller groups from a single population.

2.3.2 Agglomerative grouping

As described by Webster and Oliver (1990), the starting point for most agglomerative methods is a set of similarities (or dissimilarities) between individuals. A measure is chosen to represent relations between individuals and their values are calculated for all pairs to form a matrix. Usually the measure is expressed as similarity and scaled so that identity is represented by the value 1, and maximum dissimilarity by 0. Such a matrix contains information that is needed. The values are scrutinised to find the smallest, i.e. the distance between the closest pair of individuals. The pairs are fused to form groups. The simplest method is the single-linkage grouping.

There are many other linking and averaging strategies which exists, with their strengths and weaknesses. Among them are complete linkage and on averaging clustering are unweighted arithmetic average, weighted arithmetic average, unweighted centroid, weighted centroid and ward's method. The single linkage is mostly chosen because of its computational simplicity, and that it contracts space to one or few dimension. It was selected in order to provide a comparison to the centroid method. The unweighted centroid method has been chosen in this research for averaging because it fuses clusters of the closest centroids, and has been preferred for simple random or systematic sampling situations by many ecologist, (P. Lengendre and L. Legendre, 1998). The MINITAB software, which is being used for the analysis of the data, has these options; centroid, weighted centroid, grouped-avarage, complete linkage, wards method, median and Mc Quitty method.

2.3.3 Single linking grouping and the centroid methods

(a) Single linking

This method is similar to the nearest neighbour strategy in ecology (Lance and Williams, 1967b). After fusion of the closest pair of individuals, grouping proceeds as follows (Webster and Oliver, 1990). The matrix of the distances is re-scanned; individuals with the shortest distance join to form one group. The second shortest distance is found. If this is between the member of the first group and a third individual then the latter joins the group. If it is between two other individuals then they are fused to form a second group. The process is repeated a third and further times, and fusion is decided as above. If however, the two individuals in question are in different groups then the two are fused. Then the process continues until all individuals are contained in a single group.

Webster and Oliver (1990) illustrate the procedure, in nine sites from a survey of west Oxfordshire for which principal components were calculated. The table shows the values, and the scatter diagram shows the relative position of the sites in the plane of the first principal components.

The Table and scatter diagram

Site	X	Y
6	-2.630	-2.809
14	-2.089	0.726
34	-2.396	-0.661
45	-1.861	-0.729
47	0.367	0.266
57	5.200	-3.023
59	2.481	-1.728
63	2.594	-0.807
70	4.354	-0.732

Table 2.1 values of the scatter

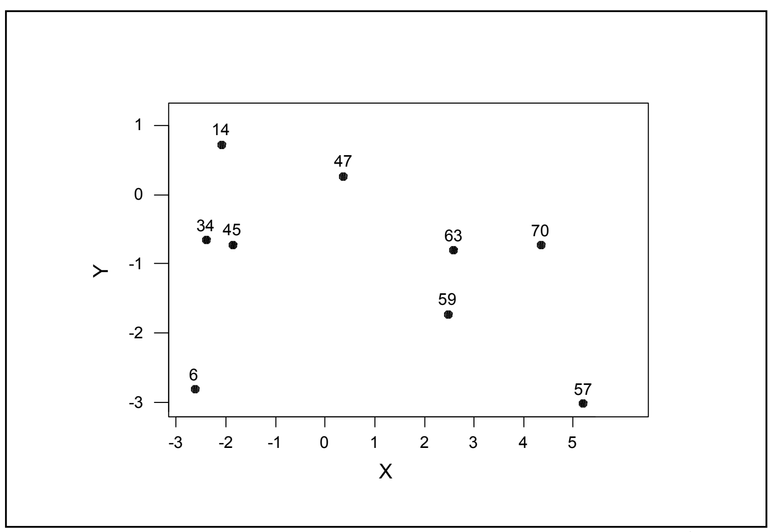


Figure 2.1 The table of values (Table 2.1) of the two principle components and the scatter plot of the nine points on the plane.

The first fusion is between sites 34 and 45, separated by 0.54 units of the distance, the next is of sites 59 and 63 at distance of 0.93 units and the third joins sites 14 and 34, which already belongs to the group with sites 45. The spanning tree diagram below figure 2.2 shows the fusion links. The results can be shown as a tree diagram or dendrogram fig. 2.3

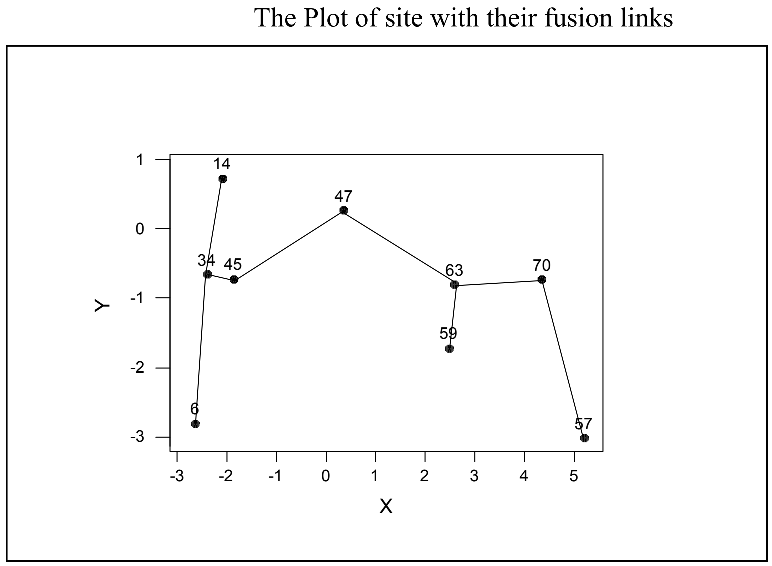


Figure 2.2. Minimum spanning tree fitted to scatter of nine points.

A classification with some desire number of classes can be obtained readily from the dendrogram by drawing a horizontal line to cut just that number of vertical stems.

A feature of the single-linkage single strategy is that each fusion is decided solely on the distance between a pair of individuals. Thus the distance between a group and another individual is a distance between the individual and the nearest member of the group; similarly, the distance between groups is the distance between their nearest members. The dispositions of the groups as a whole are not taken into account, and this is undoubtedly a weakness of the method. Single linkage is a good choice when clusters are clearly separated. When observations lie close together, single linkage tends to identify long chain-like clusters that can have a relatively large distance separating observations at either end of the chain.

(b) The centroid Method

The centroid method is one of the attractive fusion strategy from the geometric point of view. The groups are considered synthetic individuals whose positions in the Euclidean character space is its centroid. The centroid is the middle of a cluster. Distances between one centroid and other individuals are then calculated and substituted for the distances from the individual points included in the pair. Then all the distances are re-examined and the procedure is repeated as described above for the single linkage.

The centroid can be used as a measure of cluster location. For a given cluster, the average distance from the centroid is the average of the distances between observations and the cluster centroid.

The centroid method seems to separate the main groups more clearly than the single-linkage grouping.

(c) Dendrograms

Dendrogams clearly illustrates the clusters formed at each particular level. They do not allow the identification of what exact similarity links among objects are. With some clustering methods, this information is not directly available and must be found a posteriori when needed. In any case, for a synoptic clustering which only aims at organising major clusters of objects, connecting links are required (Legendre P and Legendre L, 1998).

Dendrograms

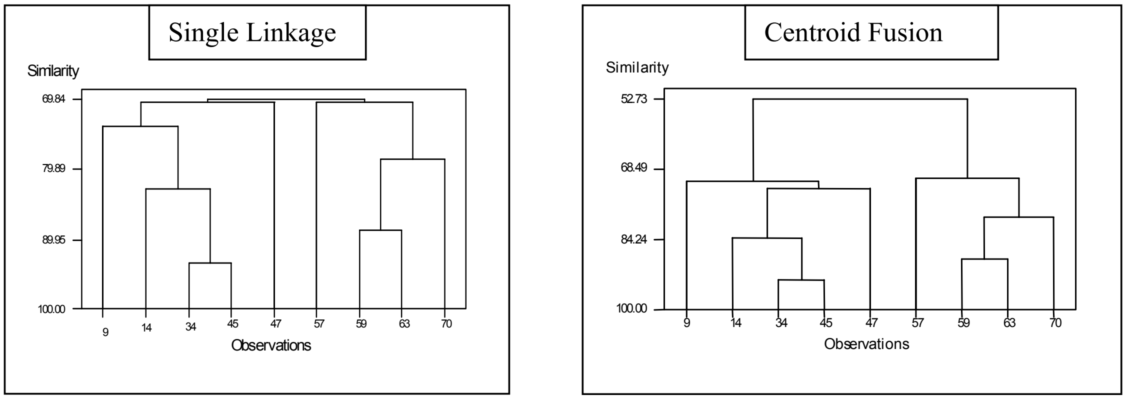


Figure 2.3. Dendrograms resulting from single-linkage and the centroid fusion of the nine sites (respectively left to right).

2.3.4 Minimum Spanning Tree (MST)

This is the technique that helps to clarify the relationship among the individuals. When there is a set of *n* individuals distributed in character space, they can be joined together by a network of *n-1* links in such a way that individuals are connected to every other individual through the network and there are no closed loops. The result is a tree spanning all the individuals. That tree in which the total length of the network is least is the minimum spanning tree (MST). “Topographically, the MST is the exact equivalent of the dendrogram formed by single linking grouping” (Webster and Oliver, 1990). Apart from being more efficiently derived than the dendrogram, the MST shows positions occupied by the individuals. Thus it reveals not only the pair or pairs of individuals, which are most alike, but also which pairs of individual in different branches of the tree are most similar. See fig. 2.2.

2.4 Nested sampling and analysis

Webster and Oliver (1990), describe the model for nested variation. It is based on the idea that a population can be divided into classes at a number of distinct stages to form a hierarchy. An initial division of population at stage 1 into classes can be divided at stage 2 into sub-classes to form a two-stage nested. This is a hierarchical classification. These classes can be subdivided further at stage 3 to give finer classes and so on. Each stage is a category and any member observation belongs to only one class in the category. The theory here being that a single observation embodies variation constituted from every stage in the hierarchy, including an unresolved variance in the smallest subdivision.
For the design with *m* stages, the model of variation is;

$$Z_{ij..m} = u + A_i + B_{ij} + C_{ijk..m} + E_{ijk...m}$$

Where $Z_{ijk...m}$ is the value of the *m*th unit in, *k*th class in stage 3, in the *j*th class at stage 2, and the *i*th class in stage 1. The general mean is u; A*i* is the difference between u and the mean of class i in the first category; B*ij* is the difference between the mean of the class of the *j*th sub class in class i and the mean of the class *i;* and so on. The $E_{ijk...m}$ quantity represents the deviation of the observed value from its class at the last stage of the subdivision. The quantities Ai, Bj, Ck,…, Eijk…m are assumed to be independent random variables associated with stages 1, 2, 3, …,m respectively, having means of zero and variances v_1^2, v_2^2, v_3^2, …,v_m^2.The latter are components of variance. The individual component for a given stage and measures the variations attributable to that stage, and together they sum to the total variance

$$v_1^2 + v_2^2 + v_3^2 + \ldots, v_m^2$$

According to Wilding, (1984), random sampling is unbiased and statistically sound. It is commonly preferred, among statisticians but may cluster data spatially unless a large number of observations are taken. This may confound the systematic and random errors.

2.4.1 Nested design.

The components of variance can be estimated by a hierarchical analysis of variance enabled by using a multi-stage or nested sampling design with replication at each level. The nested sampling scheme subdivides a population into classes at two or more distinct stages. At each stage there are small classes within larger classes apart from the lowest. The hierarchy may comprise stages in in a systematic classification scheme such as order, main group, sub group and family of the soil.

Webster and Oliver (1990) wrote that Youden and Mehlich (1937) were the first to apply this kind of design. These pioneers saw that for values of variables distributed in space, the stages could be represented by different sampling intervals, provided they are suitably nested. In this instance the components of variance represent the variations associated with the different separating distances. This forms the foundation for designing more efficient sampling in future.

The principle was illustrated on two soil series; *Culvers* and *Sassafras* series. Nine primary stations 1.6km apart were selected on each series (stage 1). At each station two sub-stations were chose at 305m apart, stage 2, then at each sub-station, two sampling areas were selected 30.5m apart, stage 3 and finally at stage 4, two sampling points at 3.05m apart were sampled. The progression of spacing was geometric so the components of variance might reasonably be independent there by allowing confidence limits to be determined.

The merit of the nested sampling and analysis is in the link between the component of variance and the spatial autocorrelation of the regionalised variable. If the components of variance were accumulated, starting with the smallest spacing, they were equivalent to the semi-variances obtained by the equation below over the same range of distances. In practice, they are only a rough estimate of the true semi-variances because each is based on the degree of freedom.

$$\boldsymbol{Z}(x) = u_v + e(x),$$

Where $z(x)$ is the value of property $\boldsymbol{Z}$ at $\boldsymbol{x}$ within the region, $\boldsymbol{u}_v$ is the mean of the within the region, $\mathbf{e}(x)$ is spatial random component, with the mean zero and a variance defined by,

$$2y = e(x) - e(x + h), \quad y = \text{variance}$$

Semi-variance is: $$2y(h) = \{z(x) - z(x + h)\}^2$$

$\mathbf{x}$ and $(\mathbf{x} + \mathbf{h})$, denotes the co-ordinates of the two positions in one, two or three dimensions, h is the vector distance (lag) between them.

The analysis proceeds by first computing the sums of squares of deviations of the means of classes in stage 1 from the general mean and multiplying each by the number of observations that make up the class mean. For each class in stage 2, the difference between its mean and the mean of the class to which it belongs in stage 1 is squared and multiplied by the number of observations in that class. The sum of these values is the appropriate sum of squares. This is repeated for each stage and the sum of squares of the individual stages sum up to the total sum of the squares. The sum is obtained dividing the sum of squares of each stage by the appropriate degrees of freedoms (Webster and Oliver, 1990).

2.4.2 Unequal sampling

When surveyors cannot always ensure that there are same numbers of individuals in every class; they may wish to vary the numbers. For example some classes may cover more ground than others, and it is reasonable that the more extensive classes are better represented in the sample. Another point is that one may want to achieve a good spatial resolution over a wide range of distances demands many stages. For each additional stage, the size of the samples doubles, in a balanced design. The nested sampling can become prohibitively expensive if many stages were required. For four stages, 72 sampling points will be required for 9 stations, i.e. 9 * 2 * 2 * 2. If a fifth and sixth stages were added it would bring the sampling points to 144 and 288 respectively. At lower stages full replication is not necessary, it can be done only on a proportion to economise.

Thus either by chance or design the sampling may be unequal. The analysis of variance is the same whether equal or unequal. While the sample size would significantly reduce. Considering the above example, at stage 5 in the unequal survey only half of the fourth stage points would be taken into account giving only 108 sample sites instead of 144 and further at stage 6 only 153 sites would be observed instead of 288 sites. This design achieves economy in the sampling effort of 25 per cent and 47 per cent respectively. See the spatial configuration of the unequal sampling below.

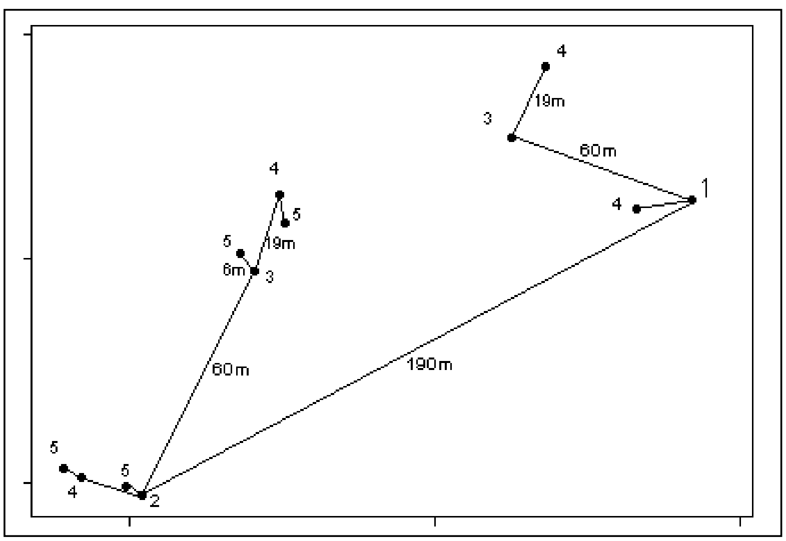
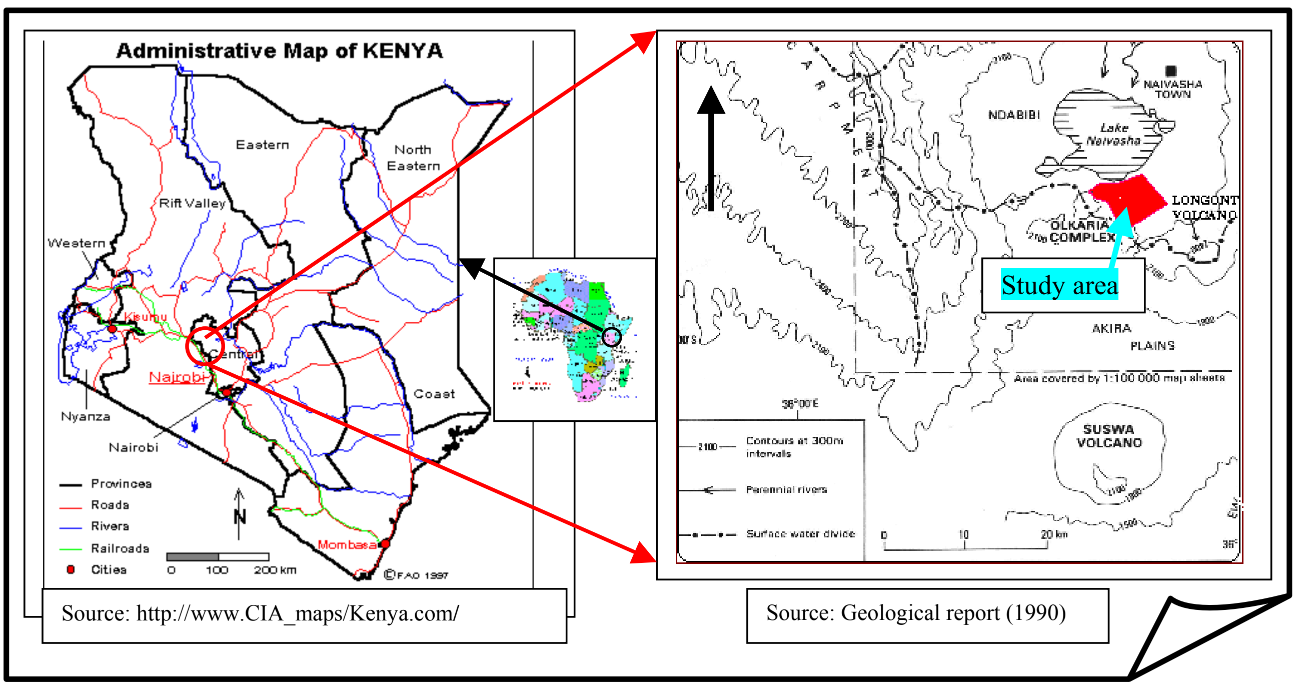


Figure 2.4. The spatial configuration of the unequal nested sampling

Chapter 3: DESCRIPTION OF THE STUDY AREA

3.1 Geographic location

The study area is located in the East African Rift valley, about 80km north west of Nairobi. It falls in the Nakuru district in the Rift Valley Province in Kenya. It is within the Naivasha catchment area which is approximately between latitude 00^0 to 1^0 00' S and 36^0 00' to 36^0 45' E. According to the LNROA report (1995), Lake Naivasha is located at 00^0 46' S and longitude 36^0 22' E. The sample area of the study area is called Sulmac Farm on the Longonot Volcanic plain, situated in the South South-east of Lake Naivasha. In relation to Naivasha town centre, it is in the South-west. The area is bouldered by Lake Naivasha in the north, the Longonot volcanic mountain in the south, the Olkaria volcanic complex hills in the south west and western part and in the east is the Obsidian ride. The major landowners of the area are; Sulmac Company, Sher Agencies, the Longonot horticulture, and the Kenya Wildlife Services.



Source: http://www.CIA_maps/Kenya.com/

Source: Geological report (1990)

Figure 3.1. The Location map of the study area (Marked in red colour)

3.2 Population and communication

Naivasha town is fast growing due to recent rapid ramification of industrial activities especially in the flower and horticultural production, the tourist industry and other human activities around the lake shores. The lake being the only fresh water lake in the area has attracted a lot of people. Due to the drought that has hit the area in the recent past the lake and the Malewa River are the only source fresh water.

According to the report of the Lake Naivasha Management plan workshop, (LNRAO, 1996) the horticultural industry employs directly some 20 000 unskilled, semi-skilled and skilled workers. About 200 000 people depend on the horticultural industry around the Lake for a regular income directly or indirectly. Population statistics for naivasha show that in 1979 the Naivasha district population was 233 302 in 43 197 households with the population increase of 3.5. At this increase in 1993 it was estimated at 450 000 people (LNRAO, 1993).

Lake Naivasha being an important national and international lake, it attracts a lot of tourists especially for bird and game viewing.

There is a well-maintained tarmac road and a railway line connecting to Nairobi and Kisumu. There also a major truck that passes through the town to from Mombassa to Kampala. Quite a number of privately owned airstrips also exist. Within the Naivasha basin, there is a good network of gravel roads and track and a few paved roads.

3.3 Climate

The climate in the valley varies according to altitude differences. Generally the altitude is high there by experiencing cool climate instead of the typical tropical hot wet conditions since it lies close to equator. The balance among precipitation, evapotranspiration and temperature plays a significant role on plant growth.

The area has a semi arid type of climate, the average rainfall around lake Naivasha is 600mm and is on the rain shadow caused by the Nyandarua mountain range which receives as much as 1525mm. From the climatic data the evapo-transpiration is about 1360mm around Naivasha area which is more than twice the amount it receives, (Kamoni, 1988)

The mean monthly maximum temperature ranges between 23^0C to 28.2^0C, the highest occurring in January and February, while the minimum is between 6.80C to 8.00C the coolest months being in July and August. The average range is between 8^0C and 11.5^0C. The mean annual temperature range is between 15.9^0C and 18.6^0C. The coolest months are July and August.

On the general note the temperature and the evapotranspiration is very high, where the rainfall is less than evapo-transpiration except for April and May. Seasonal distribution of rain shows that there is a long period of rain from March to May and a short period during October to November. December to February is the driest part of the year, with sunny days and cool, clear nights. Rain-fed crops can only thrive in March, April and May.

Month	Tmax	Tmin	Tmean	P	Eo	Et
Jan	27.6	8	17.9	22	118	117.8
Feb	28.2	8.1	18.3	35	178	114.8
Mar	27.2	9.7	18.6	59	190	120.9
Apr	25.1	11.5	18.3	107	149	102
May	23.7	11.2	17.5	86	132	93
Jun	23	9.8	16.6	41	120	90
Jul	22.5	9.2	15.9	32	125	89.9
Aug	22.8	9.3	16.2	44	142	99.2
Sep	24.5	8.7	16.7	44	158	108
Oct	24.5	9	17.3	47	183	114
Nov	25.5	9.2	17	58	134	99
Dec	24.6	8.6	17.2	36	158	102.3
Mean	*24.9*	*9.4*	*17.3*	*50.9*	*148.9*	*104.2*

Table 3.1 Climatic data for National animal husbandry station – Naivasha
Country: Kenya, Meteostation: Naivasha, Altitude: 1900m,
Coordinates: 0^0 46'S 36^0 26' E

The source Kamoni (1988) and FAO cropwat, the temperature records are similar to those of Baraton sub-station, no. 90.36/002, 0^0 43'S 36^0 25'E. The climatic data for Baraton station has been taken for the years of 1937-1955 and rainfall data for the years 1913-1972 (Siderius and Muchena, 1977).

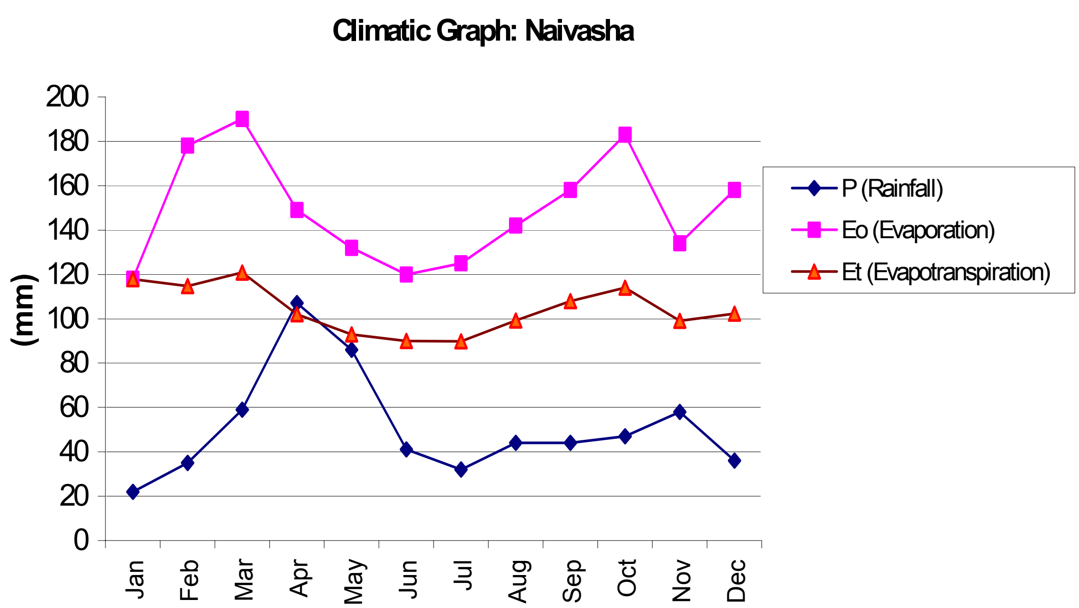


Figure 3.2. Graph of the rainfall distribution in Naivasha generated from the above table.

3.4 Soil climate

The soil in the study area has an Ustic/aridc moisture regime and an *isothermic* temperature regime (Siderius, 1977). In the 'Africa soil moisture regimes', Naivasha is classified as *Ustic* (Van Wambeke 1982). *Ustic* soil moisture regime is one in which the mean annual soil tempareture is 22^0C or higher and the mean summer and winter soil temperature differ by less than 6^0C at the depth of 50cm below the soil surface. The soil moisture control section in the areas of the of the *ustic* moisture regimes is dry in some or all parts for 90 or more cumulative days in the normal years. It is moist, however, in some parts either for more than 180 cumulative days per year or for 90 or more. The *Isothermic* temperature regime has the mean annual temperature between 15^0C and 22^0C. The difference between the mean summer and winter soil temperature is less than 6^0C (USDA, 1998).

3.5 Geology

Geologically the age of the volcanic activity that led to the formation of the study is is between 0.4 and 0.45Ma (Clarke etal 1990). The study area is in the middle of the Great Rift Valley.

The Longonot and the Okaria Volcanic complex centres mainly influence the geology of the study area. On the southeast is the Longonot Volcano. The Longonot caldera formed at the summit of the large volcanic shield with slopes frequently less than 5%. Steeper slopes 10-15% occur in the immediate vicinity of the caldera rim. Ephemeral streams have dissected the middle and the upper slopes of the shield beyond the caldera rim. Gullies as deep as 20m are common.

The products of the Longonot volcano occupy about 350 km^2, and to the north of the volcano it spreads close to the lake. Arcuate lava flow fronts up to 40m high form distinct topographic features in the north, east and southern slopes. According to Clarke etal (1990), the Longonot Volcano group incorporates seven formations. The following are some of the important events in history; (i) Building of the early shield, (ii) Caldera formation, (iii) Building of a pyroclsatic and lava cone, (iv) Formation of the summit crater and (v), the flank and crater floor lava eruption. The third formation is important to the study area. It is represented by the Akira Pumice formation. Followed by stages of cone building and dominated by lava- the Longonot Trachyte Formation. The Longonot Mixed Lava Formation was erupted on the northern lower flanks at this time also.

The Okaria Volcanic complex is on the south west of Lake Naivasha and its centres occur as either steep sided domes, formed of lava and/ or pyroclastic rock, or as thick lava flows of restricted lateral extent. Individual domes range from being small topographic features less than 500m in basal diameter and no more than 50m high to prominent feature such as Olkaria Hill, which is 340m high, and has a basal diameter of 2 km (Clarke etal, 1990).

The influence of the Olkaria Volcanic complex is on the western part of the study area dominated by hills.

Lacustrian plain has “Gambrian lake” sediments with considerable amounts of pyroclastic sediments.

Geological sketch location map

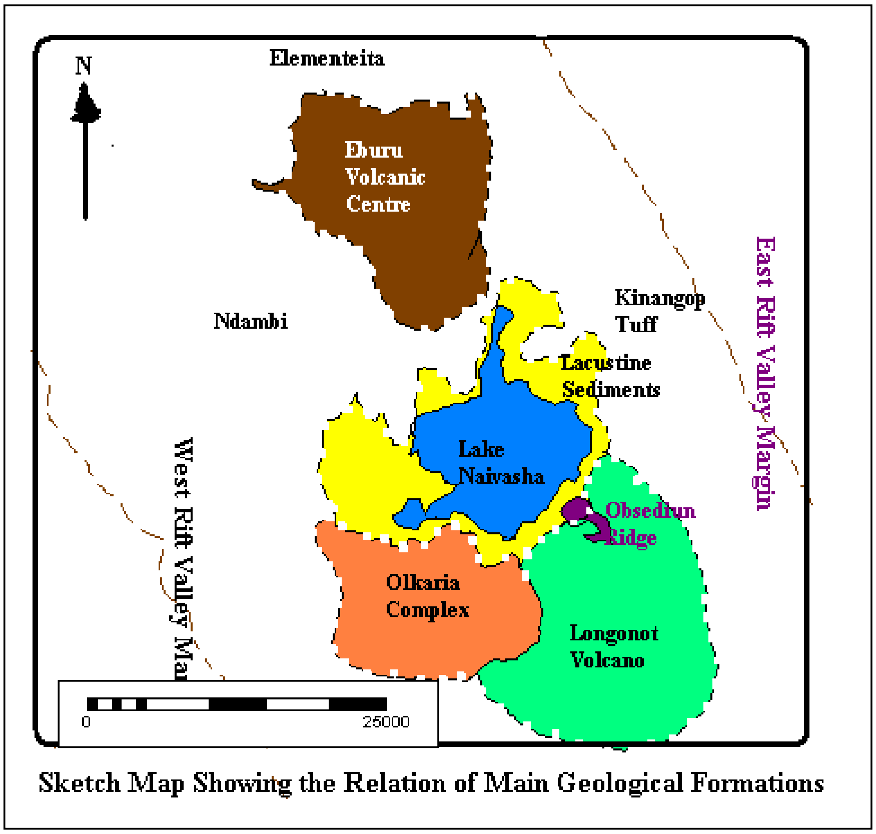


Figure 3.3. Sketch map showing the general position of the major geological features

3.6 Geomorphology

The study area occurs in the rift valley floor and three major landscapes have been identified accoding o the geopedological , i.e. the lacustrian plain, Volcanic plain and the highland. Thomson (1958) distinguishes three types of landscape in Naivasha as the Kinangop plateau, the Mau escarpment and the rift valley floor.

3.6.1 Lacustrian plain

This area surrounds the lake and extends from about 1850-1920 m above sea level. Differentiation was made on the relief level. Three levels have been identified named as low Lacustrine plain, mid Lacustrine plain and the high Lacustrine plain. These were formed due to fluctuation of the lake level. The Lacustrine plain is composed of "Gamblian lake" sediments and some pyroclastic material (Siderius, 1977). Landforms on this plain are: Riperian zone (Pl.-411), tread and riser complexes, (Pl-111, Pl-211, Pl-311) for the three levels. The general slope is flat to gently sloping (less than 4%), with no distinct channels.

3.6.2 Volcanic plain

The volcanic plain boulders the Lacustrine plain. In the study area, three relief levels were identified, but several levels exist. The lowest level is named as low volcanic plain, bouldering the Lacustrine plain. Then followed by the mid volcanic plain, and the high volcanic plain. There are many land forms on this plain among them are; tread riser complexes (Pv-111, Pv-211, Pv-311,) bottom side complex, Pv-411).

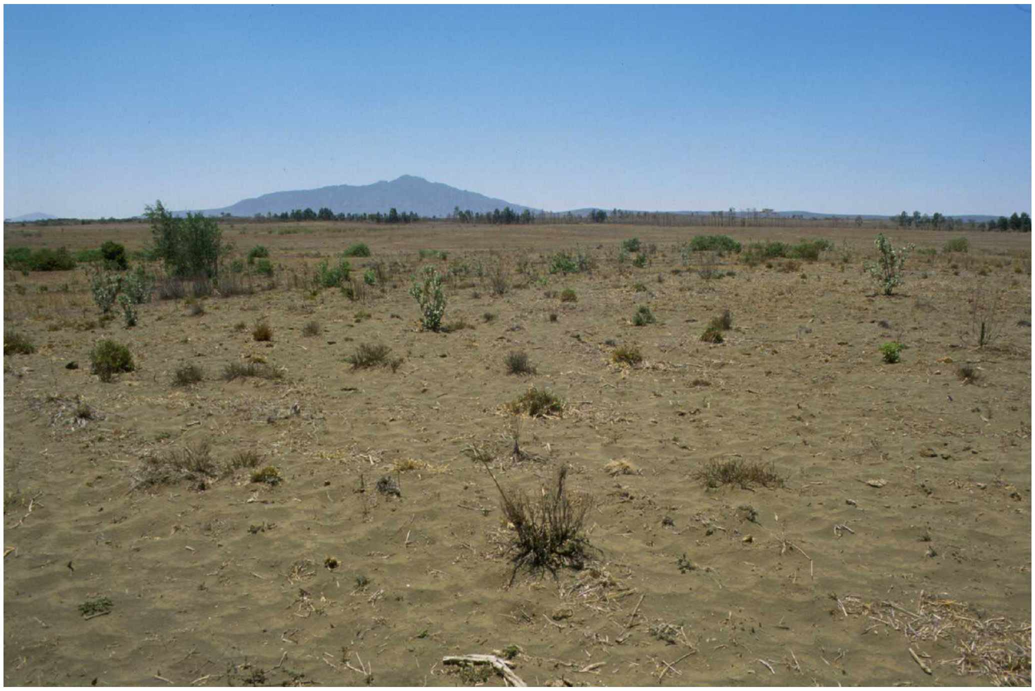


Figure 3.4 Exposed soil surface on the mid volcanic Plain (Pv-211).
The effects of over-grazing and extreme drought are obvious.

Lacustrine Plain

Figure 3.5 Vegetation on the low lacustrine plain (Pl-311). [Reference profile 2]

3.6.3 Hilland

In hilland meets the volcanic plain at its foot. These are elevated areas due to volcanic activities both on the sides and the Rift Valley floor. The hilland has been sub-divided into the foot slope, slope facet complex, and the summit at landform level in the mapped area (Hi-114, Hi-111, Hi-112).

3.7 Soils

Sombroek et al (1980) indicates that the distribution of the soils in the area is complex and influenced by intensive variation in relief, climate, and volcanic activities and the underlying rocks. The output map of the area comprehensively available is at the scale of 1:1,000,000 (Sombroek et al.,1980). According to this map, the soils in lake Naivasha basin classified according to the FAO legend (1974) are; Calcic GLEYSOLS, Ando-Haplic PHEOZEMS, Gleyic CAMBISOLS, Ando-calcic REGOSOLS, LITHOSOLS, and calcic XEROSOLS, from the Lacustrine plain through the volcanic plain, to the volcanic hills respectively.

Gathahi (1988), classified the soils developed on the Lacustrine plain deposits as FLUVISOLS which are somewhat excessively well drained to well drained. Extremely deep to very deep, dark greyish brown to dark greyish brown, loamy sands to sand loam, calcareous, slightly sodic to strongly sodic. Those developed on the pyroclastic deposits as CAMBISOLS, well-drained, very deep, dark greyish brown to very pale brown, friable sandy loam to laom, calcarous, slight to strongly sodic.

Siderius (1980), classified soil on the lacustrian deposites as *Eutric* CAMBISOLS and that generally the soil in the study area have high supply of phosphorus, calcium and magnesium. The potassium levels are high while those of nitrogen and carbon are low.

3.8 Landuse and Vegetation

3.8.1 Agriculture

According to Harper et al, (1990), the area was occupied by pastoralists (the Maasai tribe) from the 18^{th} centuary, grazing and watering their livestock around the lake Naivasha. When the white settlers arrived, land use changes stated to occur. They introduced beef and dairy farming, irrigated agriculture and later horticultural, flower and vegetable crops were introduced. Nowadays, flower and horticulture are the main activities around the lake Naivasha. All these crops are sustained by intensive irrigation mostly in green houses.

3.8.2 Vegetation

In the study area, the main vegetation types are: large communities of papyrus and grassland vegetation on the riparian zone map (unit Pl-411). On the lacustrian plain (map unit Pl-111-311) dominantly are the accacia trees. Mostly they are conserved forests, while in some areas they are scattered trees mixed with grass. On the volcanic plain (Map unit Pv-111-311) is a wooded grassland vegetation, with scattered shrubs and short grass. Mainly grazed by livestock and wildlife. Most part of the natural vegetation has been cleared for agriculture except for the National Park. Harvesting of the indigenous trees is prohibited, (Nakuru Plan, 1994/96).

3.8.3 Wildlife

There is a number of privately owned game reserves and national parks in the area, the most famous being the Hells gate National Park. Along the shores of the lake Naivasha are a number of conservation areas both private and government owned. A variety of animals include bufallos, waterbuks, giraffe, hippos, impala, zebra, warthog and many others.

Chapter 4: Materials and Methods

4.1 Research Methods

The research was conducted in three phases namely: Pre-field, field, and post-field work.
During the pre-field period literature review about the soil series concept, multivariate and numerical classification concepts were done. The collection of the available data included the following.
Aerial photos of scale 1:12 500 of 1984 and 1:10 000 of 1991 taken by the Kenya map Surveys for the Olkaria Geothermal project. Two satellite images TM (1996 and 2000), topographical by the DOS and Kenya Government 1975, and geological map by Clarke etal (1990) of the study area both at scale of 1:50 000. Exploratory soil map and agro climatic zone of Kenya by Sombroek etal (1980) both at 1:1,000,000 and the Sulmac farm map at the scale of 1:5 000 by map surveys limited January 2000, and other reference materials were collected and studied.
The aerial photos and images were interpreted using the geopedologic approach, resulting in the preliminary map and a legend, at the scale of 1:10,000 and 1:12,500.
Fieldwork comprised nested sampling, to determine the scale of spatial variability and to collect data in the delineated mapping units for the detailed soil map of the study area. The choice of the transect was determined to capture the changes in the soil along the catena, and also to see if any differences occur in the soils developed on the Longonot volcano hills and the Olkaria hills. See the map in figure 4.1 showing the study area with the two Sulmac sub-farms.

4.1.1 Sampling Scheme

The observation line for the main independent sampling centres (stations) at the bearing of approximately 112 degrees from the Northwest corner of the farm. The transect was made diagonally on the Longonot farm area. Note that there is a clear catena from the Longonot volcano through the farm to the lake. So the diagonal crosses the geomorphic variability and also any possible trends parallel to this direction ie. SW -> NE. The starting point was set near the NW corner X, Y coordinate 206400, 9907350. The first station was set 300m from the starting point, then 600m between stations and 300m to the end of the farm. Six centres were established, and the total diagonal distance was 3,600m. The calculations of the positions are;

$$E2 = E1 + (D1 * \sin(a)) \text{ where } a = \text{radial angle}$$
$$N2 = N1 + (D1 * \cos(a)) \quad ,$$

where E2 is the X-coordinate of the second point.

Result:

Easting	Northing	code
206678.155	9907237.62	10000
207234.466	9907012.85	20000
207790.776	9906788.09	30000
208347.086	9906563.33	40000
208903.396	9906338.56	50000
209459.707	9906113.80	60000

Table 4.1 Table of coordinate of stations

Then closer sampling centres were established, with the distances from a geometric series starting with a basis of 600m, spacing as in Webster & Oliver (1980): 600, 190, 60, 19, 6
These were computed as follows:
First, Series 1 is a descending factor of 10 from the station spacing: (600, 60, 6)
Then, Series 2 is the geometric means of series 1: (190, 19)
Series 2 is calculated as follows:

sqrt(600 * 60) = 189.7367 == 190 etc.

Note that the logarithms are linear:(2.7782, 2.2782, 1.7782, 1.2782, 0.7782).
At each stage after the first, points were located at random direction from the previous points. For example, stage 2 was calculated as:

E2 = E1 + (D1 * sin(a))
N2 = N1 + (D1 * cos(a))

Where (E2, N2) = 2nd point (190m from centre)
D1 = 190m
a = random angle from N (0 - 2*pi)
and so forth for other distances. However, at 6m (stage 5), only points near stage 2 centres were established, i.e. this is an unbalanced nested sampling scheme. The scheme yielded 72 sample points.

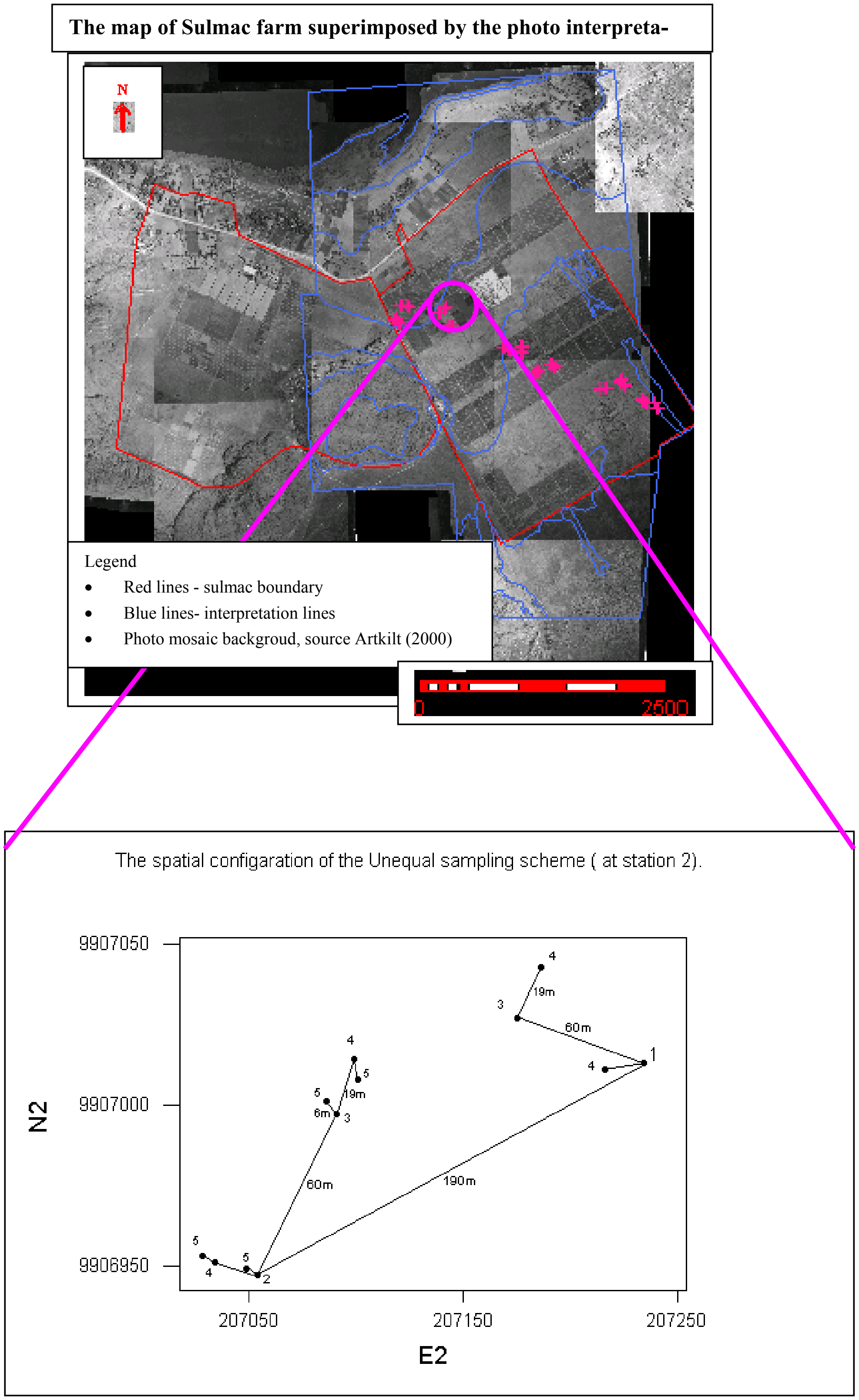


Figure 4.1. The spatial configuration of samples from the nested design. Below is a detailed picture at station 2.

4.2 Field sampling:

The 72 points were loaded into the GPS and were navigated to.
The points were checked for obvious errors in distance (especially at 19m and 6m spacing), since the GPS may have a single-reading error of 20m or more. In most cases, a tape was used at these distances to confirm the GPS reading. In cases where the points fell on inappropriate sites such as roads, in ditches, or where the soil was obviously disturbed (excavated or piled up), other random angles were applied subjectively in the field. All efforts were done to ensure objectivity, such as; if one angle changes the angles that depended on that location were not necessarily changed. The sample point locations that depended on the changed positions were adjusted accordingly.
At the sample points, mini-pits to at least 50cm depth were opened, and further auguring reached the bottom of the weathered zone (a clear colour change from weathered to an unweathered ash) or to at least 100cm depth. In this way the differences and similarities could be established between soil characteristics at each sample point. The description of the soil was done according to FAO-UNESCO (1990) guidelines. Tentative classification of the soil was done in the field to be confirmed by the laboratory results.
The following characteristic were described:
Horizons, thickness, and boundaries (sharpness and topology), depth to the weathered zone (grey) +/- 5cm. Depth and thickness of any compacted/cemented layers. For each horizon: colour (moist or dry according to the field conditions), structure, material: ash, reworked / weathered ash, pumice gravel.
Samples were collected by bulking in layers 0-20cm and 20-50cm: about 200g of soil were collected for laboratory analysis in the Netherlands. Litter and undecomposed organic matter were not sampled. In bulking the entire layer was mixed and sub-sampled.
The following properties were sent for laboratory analysis at ISRIC, using their standard methods):

1. For all the 72 samples, the following properties were determined:

- Course fragments (not passing through the 2mm sieve): amount (weight percentage), size and shape (this test was done at ITC laboratory).
- Particle size Distribution including all sands sizes.
- pH_{-H2O}
- Electrical conductivity (ECe)

2. The following extra property measurements were requested from the ISRIC laboratory on the stage 1 and 2 samples (a total of 24 samples) and the samples 61110-C and 11000-C: Mineralogy on the 50 – 420 microns sand sizes.

- CEC (NH4OAc)
- Exchangeable cations (Ca, Mg, K and Na)
- pH_{-KCl}
- Phosphate Retention
- Al_{-OX} + Fe_{-OX} (Aluminium Oxalate and Iron Oxalate)
- Organic Carbon (OC)

Unfortunately, they were not completed in time
All the sampling points in the unequal nested scheme were in the form of mini pits. These pits were, described and sampled for soil chemical and physical analysis. Three other mini pits were opened on the high Lacustrian plain, high volcanic plain and on the swale. Two-reference profile pits on the representative Lacustrian plain and the volcanic plain were opened where more properties were described and sampled for Laboratory analysis.

4.3 Post-fieldwork

4.3.1 Map preparation

After fieldwork the re-examination of the aerial photos was done to change the original interpretation where necessary fitting the acquired field knowledge. The geopedologic approach was maintained. The soils were described and classified by the standard procedures outlined in the FAO-UNESCO guidelines (1990), and the WRB taxonomy.

The base map was obtained by digitising the interpretation done on the 1:12,500 and the 1:10,000 scaled aerial photos. Under the stereoscope the photos were re-examined separating different land-forms using the geopedologic approach.

The photos were scanned, and geo-referenced using the GPS tie points, which were collected during fieldwork. At least 9 tie points were used for geo-referencing each photo interpretation (for each photo). In some cases were few points were collected, the control points were also obtained from the 1:50,000 topographic map. Each photo's interpretation was digitised separately. ILWIS was used in the digitising and map processes. Affine transformation was used in the Map Reference operation during digitising of the interpretation lines. The sigma values were all less than 0.4mm.

The individual interpretations were sub-mapped and were glue in ILWIS to obtain a full coverage interpretation of the study area. The sub-map was edited whilst being overlaid on the appropriate Geo Referenced photos. By gluing re-sampled photos created the mosaic was created, which was also used for the same purpose. The study area is fairly flat; therefore there were no serious distortions due to relief displacement.

4.3.2 Data Analysis

The analyses of the data include cluster analysis, histograms, and one way ANOVA and the nested ANOVA analysis.

One way ANOVA

ANOVA was used to test the classification justification by comparing the variances within the classes to that which exist between classes, whether there is significant difference or not.

The important assumption in the One way ANOVA is that the nature of the parent population is normally distributed. The Null hypothesis (Ho) the sample means is not different from the parent population mean. The Alternative hypothesis (H1) is the sample means are different. At a selected confidence interval (CI) level e.g.

$$Ho: u_1 = u_2 = u_3$$
$$P = 0.05 \text{ i.e. at } 95\% \text{ CI}$$

Fully Nested ANOVA

The Fully Nested ANOVA is used to perform fully nested (hierarchical) analysis of variance and to estimate variance components for each response variable. All factors are implicitly assumed to be random. The results from the fully nested samples were analysed to determine the nature of the spatial variability of soil properties in question.

Descriptive Statistics

Descriptive statistics were also performed on all numerical data using MINITAB program from which the mean, median, standard deviation and the coefficient of variance including the five-number summary were recorded to describe the distributions.

Cluster Observations

By clustering observations in MINITAB the observations were classified into groups which are similar with regard to the properties in question. This method is used when the groups are initially not known. This procedure uses an agglomerative hierarchical method that begins with all observations being separate, each forming its own cluster. In the first step, the two observations closest together are joined. In the next step, either a third observation joins the first two, or two other observations join together into a different cluster. This process will continue until all clusters are joined into one, but one must decide how many groups are logical for the data at hand. The dendrogram or tree diagram is used to show the amalgamation steps.

Final grouping

The decision about final grouping is also called "cutting the dendrogram". The complete dendrogram is a graphical depiction of the amalgamation of observations into one eventual cluster. Cutting the dendrogram is akin to drawing a line across the dendrogram to specify the final grouping.

To know where to cut the dendrogram, the cluster analysis may be performed without specifying a final partition. The similarity and distance levels are examined in the session window output and in the dendrogram. The similarity level at any step is the percent of the minimum distance at that step relative to the maximum inter-observation distance in the data. The pattern of how similarity or distance values change from step to step can help one to choose the final grouping. The step where the values change abruptly may identify a good point for cutting the dendrogram, provided it makes sense with regard to the data values. Looking at dendrograms for different final groupings can also help to decide which one makes the most sense for the data (MINITAB Inc, 1998).

In this research, numerical variables of weathered depth, depth of the surface A horizon and that of the subsurface Bw1 horizons, gravel percentage for the surface and subsurface horizons were used to perform the cluster analysis. In MINITAB all the variables were standardised by subtracting the means and dividing by the standard deviation before the distance matrix is calculated, to minimise the effect of scale differences since the variables are mostly in different units. If they are standardised, cluster centroids and distance measures are in standardised variable space). By clustering observations in MINITAB (Stat > Multivariate > Cluster Observations), the dendrogram and the tables showing the similarity level and the clusters are generated. The dendrogram was examined to determine the resulting clusters in the final partition. See the dendrogram in figure 5.9

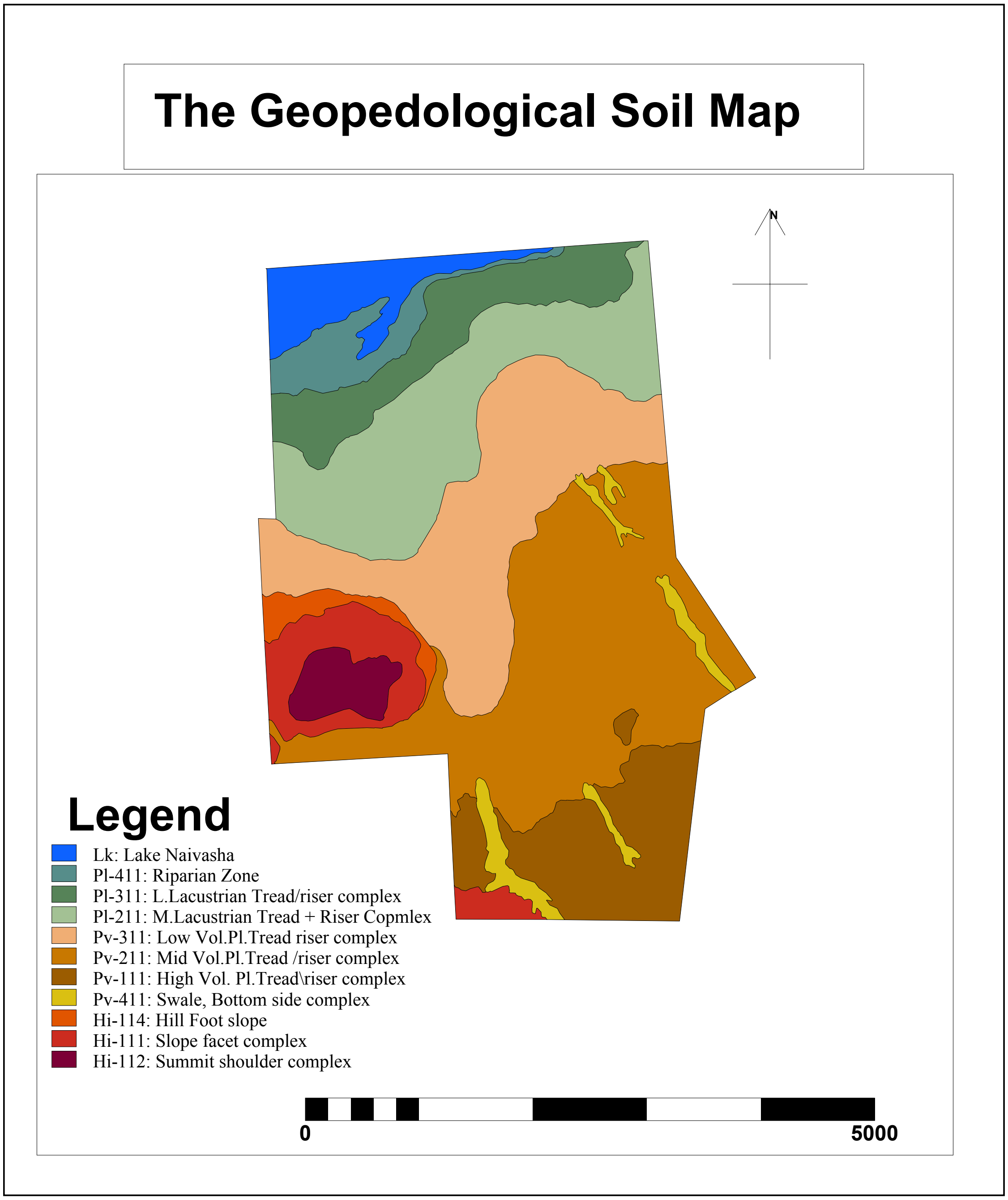
The geopedological soil series Map

The final geopedologic soil series map at the scale of 1:10,000 was prepared.

The central concept was selected for each series. Descriptions of the proposed series according to the USDA format was made with the help of the laboratory results.

Chapter 5: RESULTS AND DISCUSSIONS

5.1 Detailed Geopedological soil map.



The original interpretation was a detailed map at the scale of 1: 10,000.

The geopedological Legend of the soil map in figure 5.1

Landscape	Relief	Lithology	Land form		Soils
Hilland (Hi)	High Hills	Lava flow and domes pyroclastics	Hi-111	Slope facet complex	Leptosols ? (not sampled)
		Akira pumice	Hi-112	Summit/sholder complex	Leptic Andosols ? (not sampled)
		Lacustrine sediments and reworked pyroclastics	Hi-114	Foot slope	Skeleti-Vitric Andosol
Volcanic plain (Pv)	High Volcanic plain	Longonot Ash and Akira pumice	Pv-111	Tread/riser complex	Arenic-Vitric Andosol (Eutric)
	Mid Volcanic plain	Longonot Ash and Akira pumice	Pv-211	Tread/riser complex	Areni-Vitric Andosol (Eutric)
	Low Volcanic plain	Longonot Ash and Akira pumice and Lacustrine sediments	Pv-311	Tread/riser complex	Areni- Vitric Andosol (Eutric)
	Swale	Longonot Ash and Akira pumice	Pv-411	Bottom side complex	Areni- Vitric Andosol (Dystric)
Lacustrine Plain (Pl)	Mid terrace	Lacustrine sediments	Pl-211	Tread/riser complex	Arenic Vitric Andosol
	Low terrace	Lacustrine sediments	Pl-311	Tread/riser complex	Sodi-Fluvic Cambisol, (Skeletic, Eutric)
	Riparian Zone	Lacustrine sediments	Pl-411	Undifferentiated	not sampled
Lake (Lk)	Lake Naivasha	Water	Lk	Lake Naivasha	

Table 5.1 The legend for the soil map

In general the soils in the area are moderately deep to very deep, though there are a few indurated pumice gravel layers, which inhibit root growth. In a few places were also cemented hard layers of about 5-8cm thick at the depth of 60-65cm on the volcanic plain and at 46-55cm at the Lacustrine plain. The cementing agent was not quite established, but appeared to be fine silt, which had been washed to the wetting depth of irrigation. Both the surface and the subsurface layers are distinctly hydrophobic when dry.

<u>On the volcanic plain</u>, soils are developed on the volcanic ash deposits from Longonot. The A-horizon is generally poorly developed, and Bw-horizon also shows very little pedogenesis. The underlying C-horizon is very poorly sorted sand mostly of unweathered volcanic ash mixed with pumice gravel. In some case the C-horizon is a layer of very dark gray sands (2.5Y3/1) and gray to olive gray (5Y5/1-2). The soils are well drained deep to very deep sand loams to laomy sands, Very dark greyish brown to olive Brown (2.5Y3/2 - 2.5Y4/4) *Areni- Vitric Andosol (Eutric)* by the WRB (FAO, ISRIC and ISSS 1998) classification. In a few cases the texture was gravely sand loam to gravely loamy sands.

<u>The soils on the Lacustrine plain</u> are mainly developed from the lacustrine sediments, which include re-worked volcanic ash. On the lower terrace the soils are heavier than those on the upper terrace. They vary from silty clay to sandy loam. They are generally deep with a few pumice gravel layers, very dark grayish brown to grayish brown Sodi-Fluvic Cambisol, (Skeletic, Eutric) by the WRB - FAO, ISRIC and ISSS (1998) classification.

<u>USDA Soil Taxonomy (1998)</u>

On the Lacustrine plain the soils belong to the order of *Inceptsols*. The central concept of *Inceptisols* is that of soils that are of cool to very warm, humid and subhumid regions and that have a *cambic* horizon and an *ochric* epipedon. *Inceptisols* are soils with minimal development.
These soils had *ochric* epipedon over a *cambic* horizon. They have more than 30% (by volume) course fragments. They were classified as *Ashy-pumiceous Mixed, Isothermic Vitriandic Haplustept*.

On the volcanic plain, the soils belong to the order of Entisols. The central concept of Entisols is that soils have little or no evidence of the development of pedogenesis. These soils have no diagnostic horizons other than an *ochric* epipedon. The soils were classified as *Ashy, Glassy, Isothermic Aridic Ustpsamment Haplic Ustarents* on the family level. They were almost classified as *Inceptisol*, because they nearly fulfilled the requirements of the *cambic* horizon if it was not for the colour, which did not show any difference in the hue. The mineralogy family class could be confirmed to see whether it is glassy or mixed.

WRB - FAO, ISRIC and ISSS (1998)

On the Lacustrine plain, the soils belong to the *Cambiso*l group. On the WRB 2nd Level the soils were classified as *Sodi-Fluvic Cambisol*, *(Skeletic, Eutric*). A cambic Horizon according to the WRB definition is a subsurface horizon showing evidence of alteration relative to the underlying horizons. It lacks the set of properties diagnostic for a *ferralic, argic, natric* or *spodic* horizon and the dark colours, organic matter content and structure of a *histic, folic*, *mollic* or *umbric* horizon.

On the volcanic plain the soils were classified as *Areni-Vitric Andosol (Eutric)* and on the foothills/footslopes they were classified as *skeleti-vitric Andosol (Dystric*). These soils lack a *cambic* horizon, as the subsoil is very weakly differentiated from the parent material.
On the hills one would expect *Leptosols*, and Leptic *Andosols*. Figure 5.1 and table 5.1 shows the map of the study area and the legend respectively showing the soils in each mapping unit. Refer to the reference pit descriptions in appendix A for a detailed description and classification.

Reference profile pit 1

Figure 5.2 Reference profile pit 1 on the mid volcanic plain (Pv-211)

Figure 5.3 Vegetable field and carnations green houses on the low volcanic plain (Pv-311) Notice fault separating lava flow in background.

5.2 Statistical analysis

5.2.1 Descriptive Statistics and One-way ANOVA by station and map unit

Each variable was characterized with stratification by geopedological map unit. In addition, one variable (weathered depth) was characterized with stratification by station (sample centre) for comparison. In this way, geographical differences due to soils should be revealed, and can form the basis for defining series.
One-way ANOVA test the following hypothesis:

The null hypothesis is that there are no differences in the mean for each variable among the groups i.e. among map units or among the stations

Depth of the weathered zone

(1) By station
Descriptive statistics for depth of the weathered zone

station	N	N*	Mean	Median	TrMean
1	12	0	48.25	50.00	49.40
2	12	0	61.83	62.00	62.40
3	8	4	63.88	65.00	63.88
4	9	3	56.44	48.00	56.44
5	8	4	53.25	54.00	53.25
6	11	1	57.45	50.00	55.89

station	StDev	SE Mean	Minimum	Maximum	Q1	Q3
1	7.53	2.17	30.00	55.00	45.25	55.00
2	5.69	1.64	50.00	68.00	58.50	67.00
3	7.47	2.64	50.00	75.00	60.00	69.00
4	18.47	6.16	40.00	85.00	40.00	77.50
5	10.17	3.59	35.00	65.00	46.25	62.25
6	14.93	4.50	44.00	85.00	48.00	75.00

Table 5.2. The summary of the descriptive statistics of the weathered depth by station.

Boxplots of the weathered depth by station

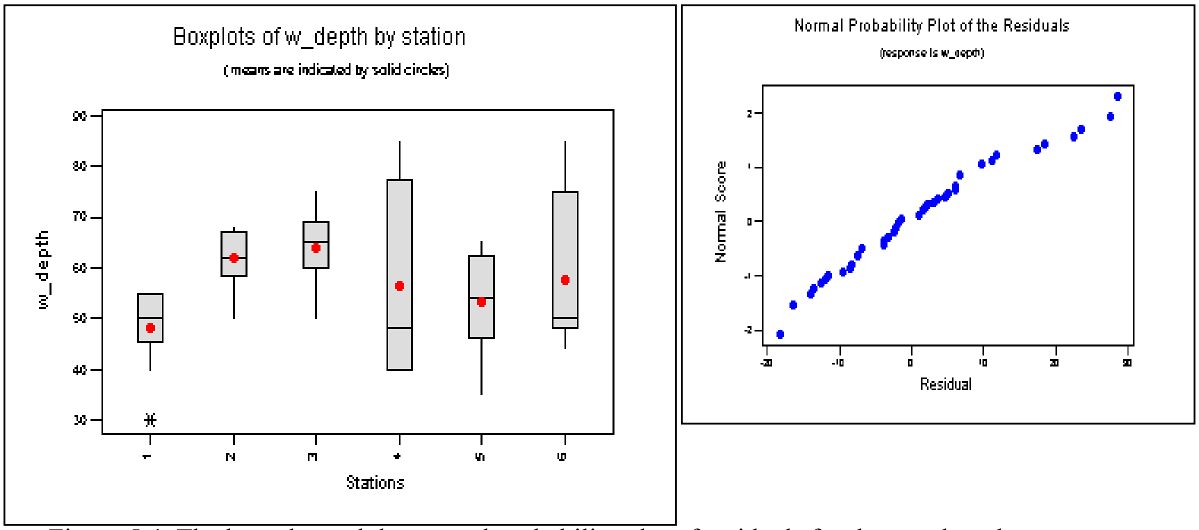


Figure 5.4. The box plot and the normal probability plot of residuals for the weathered zone

Analysis of Variance for weathered depth					
Source	DF	SS	MS	F	P
Station	5	1687	337	2.58	0.036
Error	54	7053	131		
Total	59	8740			

Table 5.3. One-Way ANOVA for weathered depth

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
				---------+---------+---------+-------
1	12	48.25	7.53	(-----*------)
2	12	61.83	5.69	(------*-----)
3	8	63.87	7.47	(-------*-------)
4	9	56.44	18.47	(------*-------)
5	8	53.25	10.17	(-------*-------)
6	11	57.45	14.93	(-----*------)
				---------+---------+---------+-------
Pooled StDev = 11.80				50 60 70

Table 5.4. Individual 95% CIs for Mean

The F-test p-value of 0.036 value analysis indicates enough evidence at $\alpha = 0.05$ to reject the null hypothesis. Thus, there are significant differences among the means of weathered depth at the six stations.

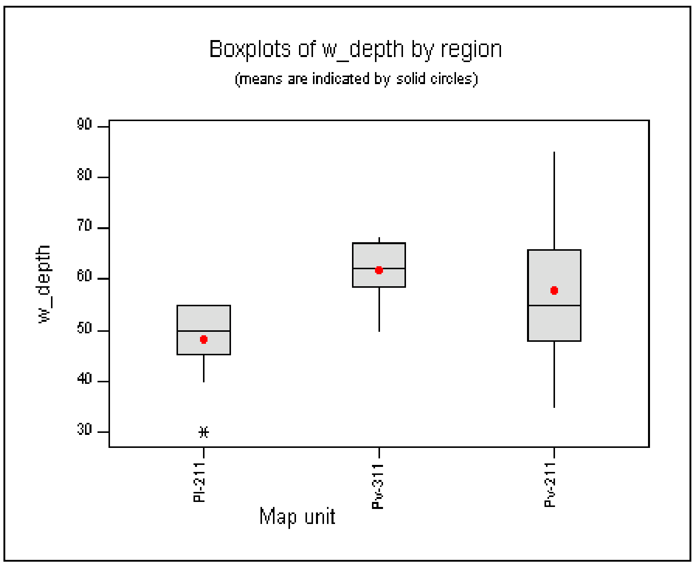
The box-plot figure 5.2 shows that the station 1 has the lowest depth with its mean at 48cm station 2 and 3 are not visually different from each other but distinctly different from station 1. Station 4-6 cannot be separated because they over lap. They also overlap all others. Station 4 has the widest range. The general trend is that the weathered depth increases from the station 1 through 3 then it decreases slightly through 6, i.e. ascending to towards the volcano.

(2) By map unit

SMUcode	N	N*	Mean	Median	TrMean	
Pl-211	12	0	48.25	50.00	49.40	
Pv-311	12	0	61.83	62.00	62.40	
Pv-211	36	12	57.69	55.00	57.25	
SMUcode	StDev	SE Mean	Minimum	Maximum	Q1	Q3
Pl-211	7.53	2.17	30.00	55.00	45.25	55.00
Pv-311	5.69	1.64	50.00	68.00	58.50	67.00
Pv-211	13.68	2.28	35.00	85.00	48.00	65.75

Table 5.5. Descriptive statistics: weathered depth by map unit.

Box plots of the weathered depth by map unit.



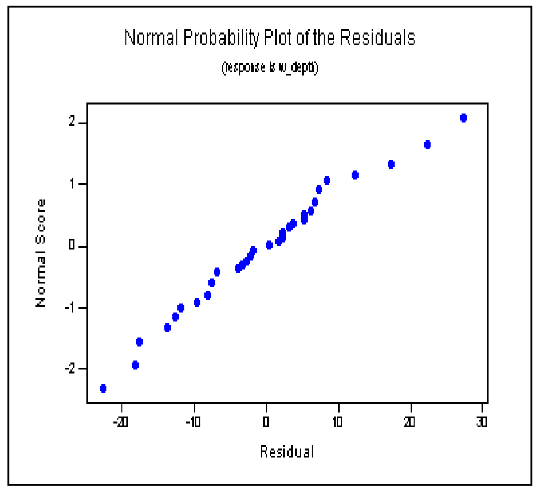


Figure 5.5. The Boxplots of the weathered depth by the Mapping units and the normal probability plot of the residuals.

Analysis of Variance for w_depth					
Source	DF	SS	MS	F	P
SMU ID	2	1208	604	4.57	0.014
Error	57	7532	132		
Total	59	8740			

Table 5.6. One way ANOVA of the weathered depth in the map units

```
                                             Individual 95% CIs For Mean
                                             Based on Pooled StDev
Level        N      Mean     StDev  --------+---------+---------+--------
Pl-211      12     48.25      7.53  (-------*--------)
Pv-311      12     61.83      5.69                   (-------*--------)
Pv-211      36     57.69     13.68                 (----*----)
                                    --------+---------+---------+--------
Pooled StDev =     11.49                 48.0      56.0      64.0
```

Table 5.7. Individual 95% CIs

There are significant differences in the means of the weathered depths among the mapping units. F-test P = 0.014 at $\alpha = 0.05$

Here again there is a clear visual differences in the weathered depth between Lacustrine plain (Pl-211) and low volcanic plain (Pv-311), but the mid volcanic plain (Pv-211) overlaps both the low volcanic plain and the Lacustrine plain. The mean weathered depth is shallowest in the Lacustrine plain and deepest in the low volcanic plain.
Summary table

	Mean depth (cm)	Range (cm)	Calculated figures from mean and StDev (cm)
Pl-211	48	41 - 56	40.72 – 55.78
Pv-311	62	56 – 68	56.14 – 67.52
Pv-211	58	44 – 71	44.01 – 71.37

Table 5. 8. The summarry table of the depth of the weathered.

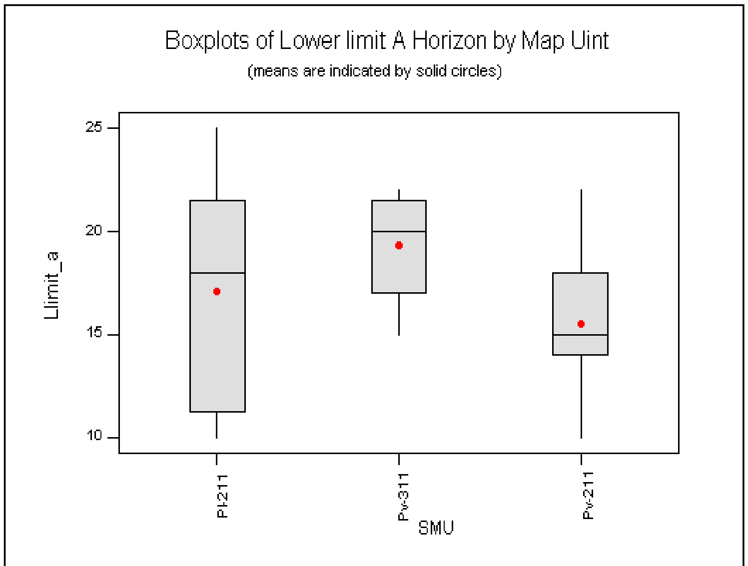
Thickness of the A horizon

The Descriptive statistics on the depth of the lower limit of the A horizon by mapping units:

Variable	SMUcode	N	Mean	Median	TrMean	StDev
Llimit_a	Pl-211	12	17.08	18.00	17.00	5.12
	Pv-311	12	19.333	20.000	19.500	2.570
	Pv-211	48	15.521	15.000	15.477	2.881
Variable	SMUcode	SE Mean	Minimum	Maximum	Q1	Q3
Llimit_a	Pl-211	1.48	10.00	25.00	11.25	21.50
	Pv-311	0.742	15.000	22.000	17.000	21.500
	Pv-211	0.416	10.000	22.000	14.000	18.000

Table 5.9. Descriptive statistics on the depth of the lower limit of the A horizon by mapping units:

The box plot of the depth of the A horizon.



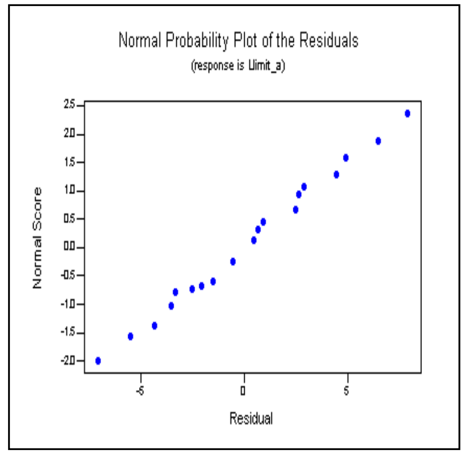


Figure 5.6. The box plot of the depth of the A horizon and the normal plot of residuals

One-way Analysis of Variance

Analysis of Variance for Llimit_a					
Source	DF	SS	MS	F	P
SMUcode	2	145.9	73.0	6.70	0.002
Error	69	751.6	10.9		
Total	71	897.5			

Table 5.10. Analysis of Variance for thickness of A horizon

```
                                      Individual 95% CIs For Mean
                                         Based on Pooled StDev
Level         N       Mean      StDev  --------+---------+---------+--------
Pl-211       12     17.083      5.125      (--------*---------)
Pv-311       12     19.333      2.570                 (---------*--------)
Pv-211       48     15.521      2.881   (----*---)
                                       --------+---------+---------+--------
Pooled StDev =       3.300                    16.0      18.0      20.0
```

Table 5.11. Individual 95% CIs

There is statistical evidence that there are significant differences in the means of the depth of the A horizon among the map units, F test, $\alpha = 0.05$ and $P = 0.002$.

There is a clear separation of the depth range between Pv-211 and Pv-311, but Pl-211 overlaps both. The following summary table shows the limits. The box plot also shows the differences.

	Mean depth (cm)	Range (cm)	Calculated figures from mean and StDev (cm)
Pl-211	17	12 - 22	11.93 – 22.24
Pv-311	19	17 - 22	16.76 – 21.9
Pv-211	16	13 - 18	12.64 - 18.40

Table 5.12. Summary table depth of A horizon

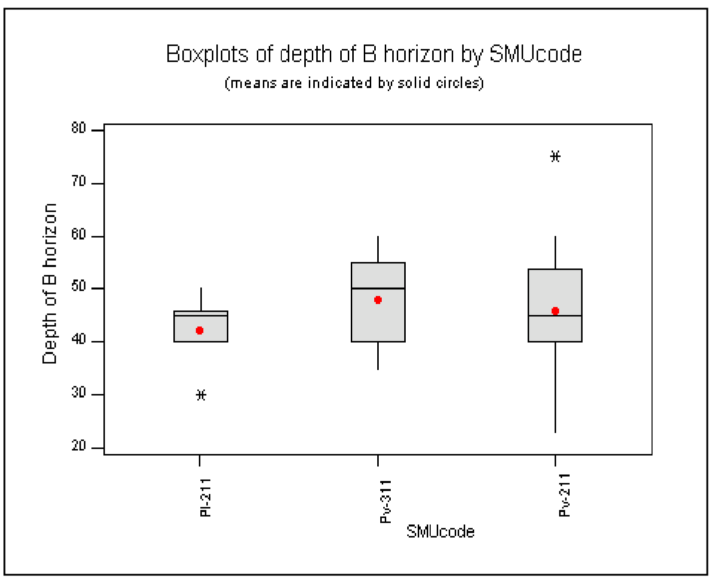
The Depth of the B horizon

The descriptive statistics on the depth of the B horizon by the mapping units

Variable	SMUcode	N	Mean	Median	TrMean	StDev
Llimit_b	Pl-211	12	42.33	45.00	42.80	6.36
	Pv-311	12	47.92	50.00	48.00	8.69
	Pv-211	48	46.04	45.00	46.07	9.78
Variable	SMUcode	SE Mean	Minimum	Maximum	Q1	Q3
Llimit_b	Pl-211	1.84	30.00	50.00	40.00	46.00
	Pv-311	2.51	35.00	60.00	40.00	55.00
	Pv-211	1.41	23.00	75.00	40.00	53.75

Table 5.13. Descriptive statistics on the depth of the B horizon

The box plot of the B horizon depth.



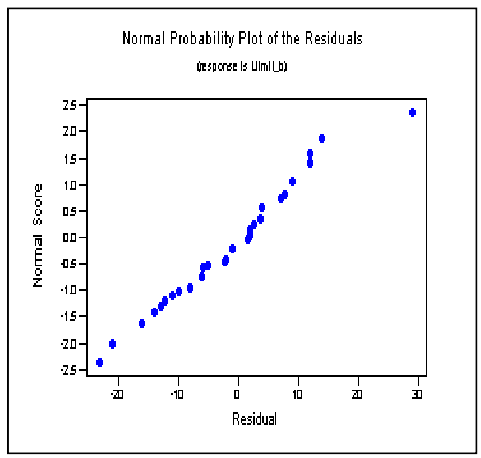


Figure 5.7. The box plot of the depth of the B horizon with the normal plot of residuals

Table 5.14. Analysis of Variance

Analysis of Variance for depth of B horizon					
Source	DF	SS	MS	F	P
SMUcode	2	200.5	100.2	1.20	0.308
Error	69	5773.5	83.7		
Total	71	5974.0			

```
                                           Individual 95% CIs For Mean
                                           Based on Pooled StDev

Level       N      Mean     StDev  ------+---------+---------+---------+
1          12    42.333     6.358  (----------*---------)
2          12    47.917     8.691            (----------*---------)
3          48    46.042     9.783              (----*----)
                                   ------+---------+---------+---------+
Pooled StDev =    9.147               40.0      45.0      50.0      55.0
```

Table 5.15. Individual 95% CIs

There is no statistical evidence to reject the null hypothesis. There are no significant differences in the means of the depth of the B horizon, F test, a = 0.05, P = 0.308.

Their ranges overlap among themselves as can be seen in the box plot and in the summary table below.

	Mean depth (cm)	Range (cm)	Calculated figures from mean and StDev (cm)
Pl-211	42	36 - 48	35.98 – 48.691
Pv-311	48	36 - 57	36.23 – 56.61
Pv-211	46	36 - 56	36.06 – 56.03

Table 5.16. Summary table depth of B horizon

Amounts of Gravel in A horizon

Descriptive statistics on percentage gravel in the A horizon by the mapping units.

SMUcode	**N**	**Mean**	**Median**	**TrMean**	**StDev**
1	12	3.530	3.258	3.345	1.231
2	12	7.37	6.92	6.63	4.63
3	48	0.4478	0.0625	0.4155	0.5484
SMUcode	**SE Mean**	**Minimum**	**Maximum**	**Q1**	**Q3**
1	0.355	2.020	6.894	2.773	4.050
2	1.34	2.78	19.34	3.41	9.60
3	0.0792	0.0000	1.6135	0.0000	0.8834

Table 5.17 Descriptive statistics on percentage gravel in the A horizon

Box plot of the percentage gravel in A horizon

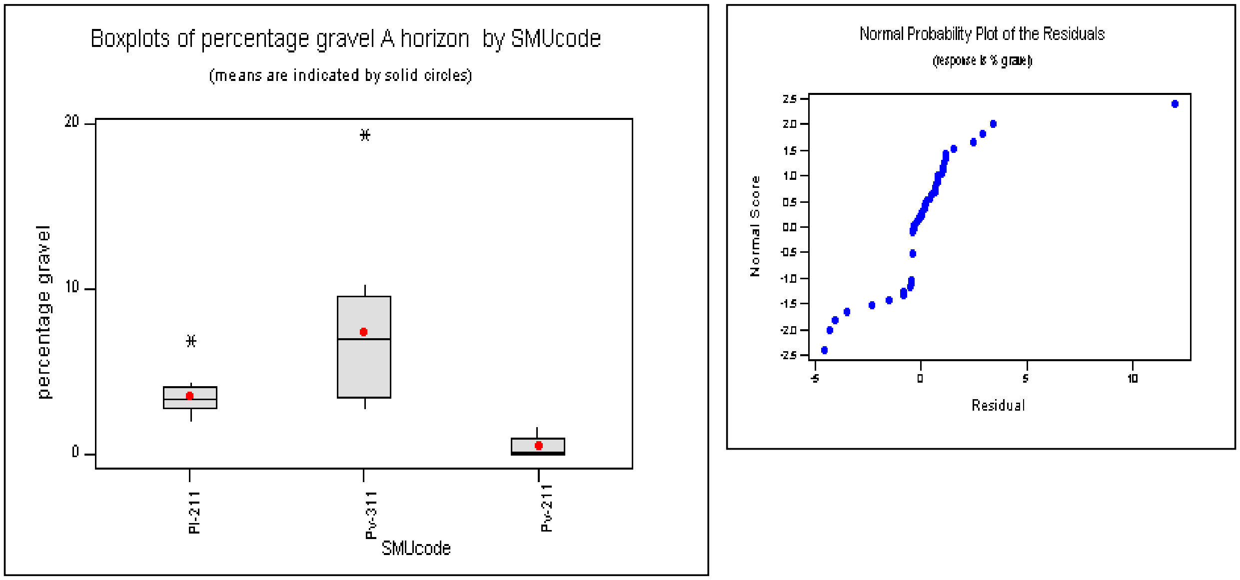


Figure 5.8. The boxplot of Gravel percentage in A horizon and the normal probability plot of the residuals

Analysis of Variance for gravel percentage in A horizon

Source	DF	SS	MS	F	P
SMUcode	2	488.28	244.14	63.17	0.000
Error	69	266.69	3.87		
Total	71	754.97			

Table 5.18. Analysis of Variance for gravel percentage in A horizon

```
                                              Individual 95% CIs For Mean
                                              Based on Pooled StDev
Level        N      Mean     StDev  -+---------+---------+---------+-----
1           12     3.530     1.231            (---*----)
2           12     7.366     4.631                         (---*----)
3           48     0.448     0.548   (-*-)
                                    -+---------+---------+---------+-----
Pooled StDev =     1.966            0.0       2.5       5.0       7.5
```

Table 5.19. Individual 95% CIs

The mean gravel percentages by weight in the different mapping units are significantly different (P = 0.000). The boxplot also confirm the difference visually.

Note that Pv-211 has a very short range and having the least amounts, Pv-311 has the largest range has the greatest amounts. The ranges of Pl-211 and Pv-311 slightly overlap. Thus gravel percentage is consistently very low in the mid volcanic plain, moderate and also fairly consistent in the low volcanic plain, but then again increases slightly in the lacustrine plain. The summary table below shows the ranges.

The summary table

	Mean gravel (%)	Range (%)	Calculated figures from mean and StDev (cm)
Pl-211	3.5	2.3 – 4.8	2.30 – 4.76
Pv-311	7.4	2.8 - 12	2.75 – 12.00
Pv-211	0.4	0 – 1	0 – 1.00

Table 5.20. The summary table for the percentage gravel in A horizon.

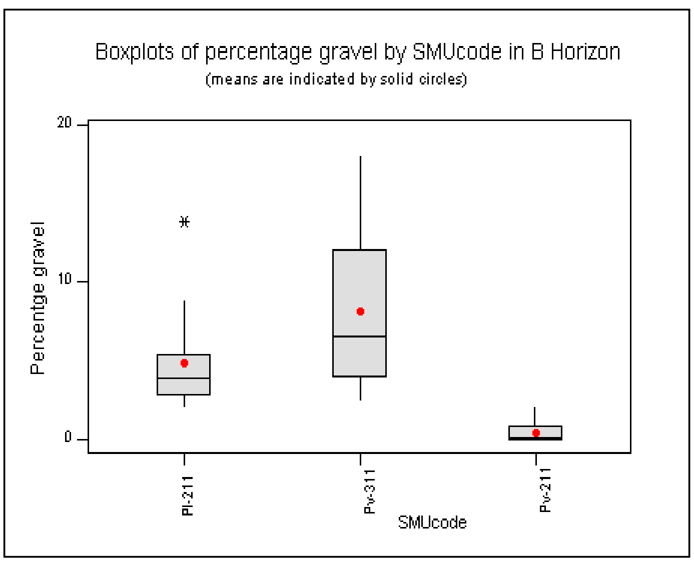
Amounts of Gravel n the B Horizon

Descriptive statistics for the Percentage gravel in the B horizon by the mapping units

SMUcode	N	Mean	Median	TrMean	StDev
Pl-211	12	4.892	3.954	4.278	3.331
Pv-311	12	8.14	6.59	7.72	4.99
Pv-211	48	0.4221	0.0655	0.3789	0.5455
SMUcode	SE Mean	Minimum	Maximum	Q1	Q3
Pl-211	0.961	2.120	13.814	2.820	5.365
Pv-311	1.44	2.53	17.97	4.03	12.06
Pv-211	0.0787	0.0000	1.9611	0.0000	0.8009

Table 5.21. Descriptive statistics for the Percentage gravel in the B horizon

The box plot of the percent gravel in the B horizon



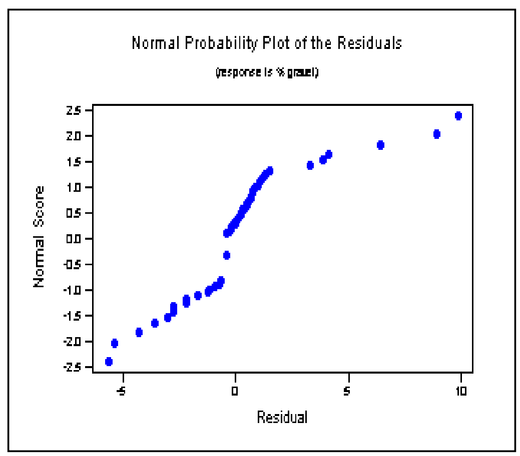


Figure 5.9. The boxplot of the gravel percentage and the normal probability plot of residuals for the B horizon

Table 5.22. ANOVA for Gravel B horizon

Analysis of Variance for percentage gravel in B horizon					
Source	DF	SS	MS	F	P
SMUcode	2	657.33	328.67	55.37	0.000
Error	69	409.54	5.94		
Total	71	1066.87			

```
                                           Individual 95% CIs For Mean
                                           Based on Pooled StDev

Level        N      Mean     StDev  -+---------+---------+---------+-----
Pl-211      12     4.892     3.331              (---*----)
Pv-311      12     8.139     4.987                        (----*----)
Pv-211      48     0.422     0.546  (-*--)
                                    -+---------+---------+---------+-----
Pooled StDev =     2.436           0.0       3.0       6.0       9.0
```

Table 5.23. Individual 95% Cis

There are significant differences among the means of the percentages of gravel in the B horizon. F test, a = 0.05, and P = 0.000. The results are very similar to that of percentage gravel in the A horizon above and so is the interpretation. This suggests that the gravels are from the original parent material, and not as a result of differential erosion (lag concentration).
The box plot above (Figure 5.8) also shows this difference.

Summary table of the gravel percentage in the B Horizon.

	Mean gravel (%)	Range (%)	Calculated figures from mean and StDev (cm)
Pl-211	4.9	1.6 – 8.2	1.56 – 8.22
Pv-311	8.1	3.2 – 9.9	3.15 – 9.88
Pv-211	0.4	0 – 1.0	0 - 0.97

Table 5.24. The summary table showing means and ranges.

5.2.2 Tentative series defined by GP mapping units

Summarising the above results, and including field observations of texture, colour and parent material, we find three tentative series. The representative value is taken as the mean, and the normal range is taken as +/- one standard deviation, as shown above. The table below shows the class limit of the tentative series defined by the GP units of the conventional Soil Taxonomic systems classification. Note that these descriptions do not cover every aspect of the series, because several defining variables were not determined, for example mineralogy.

The summary of all the variable ranges defined by the map units

Means

	Means for the Response variables				
	W_Depth (m)	*LimitA (cm)*	*LimitB (cm)*	*GravelA (% wt)*	*GravelB (% wt)*

Pl-211	48	17	42	3.5	4.9
Pv-311	62	19	48	7.4	8.1
Pv-211	58	16	46	0.4	0.4

Ranges

	Class limits for the Response variables				
	W_Depth (m)	*LimitA (cm)*	*LimitB (cm)*	*GravelA (% wt)*	*GravelB (% wt)*
Pl-211	41 - 56	12 - 22	36 - 48	2.3 – 4.8	1.6 – 8.2
Pv-311	56 – 68	17 - 22	36 - 57	2.8 - 12	3.2 – 9.9
Pv-211	44 – 71	13 - 18	36 - 56	0 – 1	0 – 1.0

Tables 5.26. The Summary tables for the means and ranges of the variables defining the map units

Pl-211: (**Naivasha Series**) thin solum, moderately thick sandy loam or loamy sand A horizon, shallow lower limit of sandy loam or loamy sand Bw-horizon, slightly-gravely, over deep unweathered lacustrine sediments derived from volcanic ash.

Pv-311: (**Sulmac Series**) moderately thick solum, gravely sandy loam to gravelly loamy sand A throughout, over volcanic ash, moderately thick A horizon, shallow low limit of Bw horizon.

Pv-211: (**Longonot Series**) moderately thick solum, Sandy loam to loamy sand, glassy, over volcanic ash, moderately deep, moderately thick A and very shallow B-horizons, non-gravely.

Note: the term thin, mod thick, thick and very thick applied to horizons refers to <10cm, 10-20cm, 20-30cm, 30-40cm >40cm respectively, applied to the solum the limits are <30, 30-50, 50-100, and 100-150 and > 150 as very shallow, shallow, moderately deep, deep and very deep, respectively.

Notice that the series defined by the geopedological units overlap in most of the properties. Only the Longonot series is clearly separated from the others by almost complete lack of gravel. The central concepts of Naivasha and Sulmac series are clearly differentiated.

5.2.3 The field key to the series proposed (Conventional)

The key to series can be used find the series name of a soil for which the definition is known and also to locate the names of all soil series so far recognised on parent material types with the same lithology and/or same mineralogy (Clayden and Hollis, 1984). In our case we shall use the properties under consideration.

The field Key

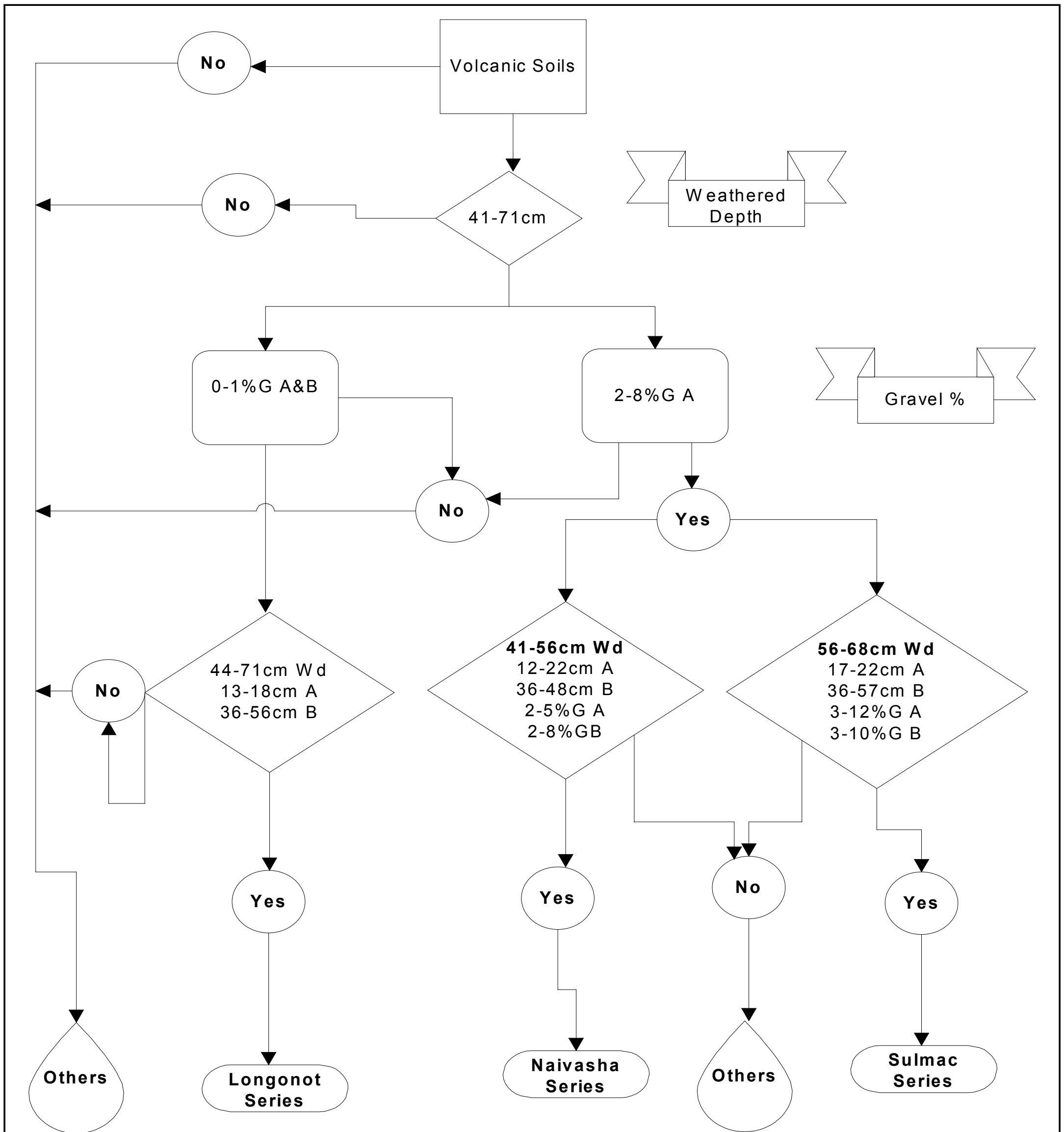


Figure 5.10 The field Key for the proposed series

5.2.4 Numerical classification

5.2.5 Cluster analysis, assign class limits for important attributes to be used to differentiate soils series.

Here clustering of observations to classify observations into groups was used. The following figure shows the dendogram determined by cluster analysis on all numerical variables.

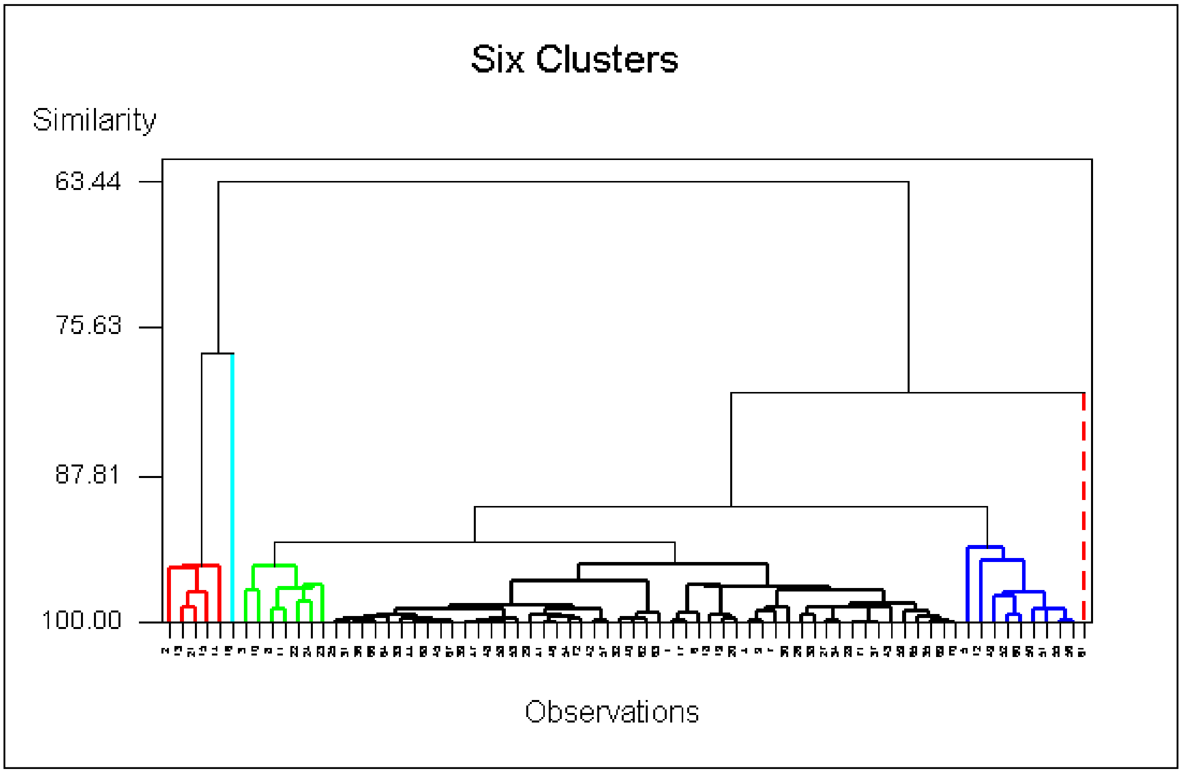


Figure 5.11 The dedrogram showing six clusters
(Cluster 1=black, cluster 2=red, Cluster 3=green, cluster 4=blue, dotted red=5, cyan=6)

By examining the resulting clusters in the dendrogram and the distance levels in the Session window output, six clusters were decided to be the final reasonable partitioning. The table below shows the final clusters. Cluster 1 contains most of the observations, while Clusters 2, 3, and 4 have a small number. Clusters 5 and 6 each contain only one unusual observation, which can be considered to fall outside of any series to be defined. We thus tentatively identify four series, to be associated with cluster 1 – 4.

Clusters	No. of Observa-tions	Within Clusters Sum of Squares	Avag Distance From Centroid	Max distance From Centroid
Cluster1	49	68.054	1.095	1.976
Cluster2	5	5.015	0.952	1.266
Cluster3	7	8.022	1.046	1.381
Cluster4	9	9.419	0.965	1.627
Cluster5	1	0.000	0.000	0.000
Cluster6	1	0.000	0.000	0.000

Table 5.27. The final clusters by considering six clusters

The cluster membership for each observation was stored in the observation table. Thus the clusters may be mapped and compared to map units.

Descriptive statistics and One way ANOVA

Descriptive statistics and One-way analysis of variance was performed on the dependent variables listed above in one column, subscripts being taken from the ‘6_cluster’column, which is the cluster number. This was done to see the differences between means of the groups with respect to the variables. The boxplots were also generated to give a visual impression of the differences. Clusters 5 and 6, containing one observation each, were graphed but not considered in the grouping.

The Null hypothesis:

There are no differences in the soil properties among the group means i.e. in the four tentative series proposed

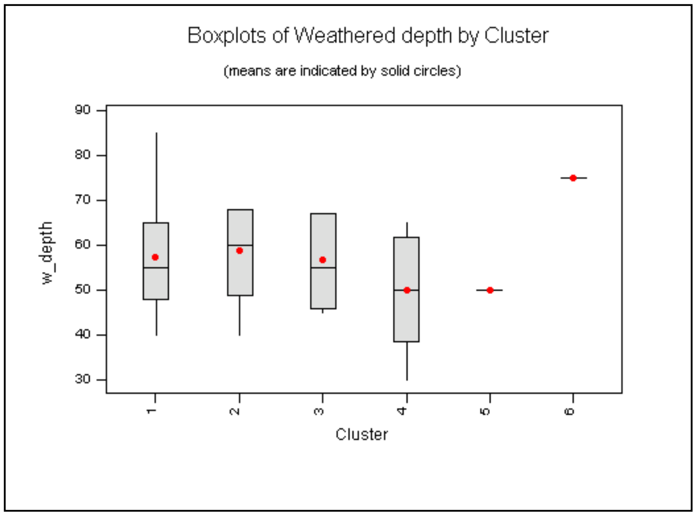
The depth of the weathered zone

Descriptive Statistics of the weathered depth by clusters

Cluster	N	N*	Mean	Median	TrMean
1	38	11	57.39	55.00	56.79
2	5	0	58.80	60.00	58.80
3	7	0	56.71	55.00	56.71
4	8	1	50.13	50.00	50.13
5	1	0	50.000	50.000	50.000
6	1	0	75.000	75.000	75.000
cluster	StDev	SE Mean	Minimum	Maximum	Q1
1	12.53	2.03	40.00	85.00	48.00
2	11.45	5.12	40.00	68.00	49.00
3	9.64	3.64	45.00	67.00	46.00
4	12.44	4.40	30.00	65.00	38.75
5	*	*	50.000	50.000	*
6	*	*	75.000	75.000	*

Table 5.28 Descriptive Statistics of the weathered depth by clusters

The boxplot and the Normal probability plot.



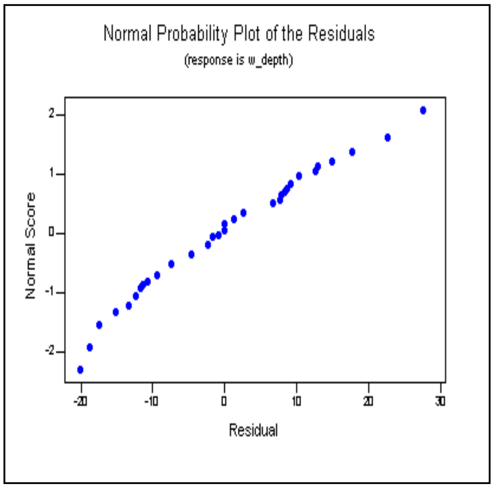


Figure 5.12. The Box plot and the normal probability plot of the residuals for the weathered depth.

Analysis of Variance for weathered depth

Source	DF	SS	MS	F	P
Cluster_6	5	766	153	1.04	0.405
Error	54	7974	148		
Total	59	8740			

Table 5.29. Analysis of Variance for weathered depth

```
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
1          38     57.39     12.53                 (-*-)
2           5     58.80     11.45              (----*-----)
3           7     56.71      9.64              (---*----)
4           8     50.12     12.44           (---*---)
5           1     50.00      0.00   (-----------*-----------)
6           1     75.00      0.00               (------------*-----------)
                                   --------+---------+---------+--------
Pooled StDev =    12.15                   40        60        80
```

Table 5.30. Individual 95% CIs

There is no statistical evidence to reject the null hypothesis F test, $\alpha = 0.05$, P= 0.405. It is also visually clear from the boxplot in figure 5.10 that the means of weathered depth among the clusters are not significantly different from each other.
The ranges overlap and can not easily separate them. See the summary below.

The summary table for the means and the ranges of the weathered zone

	Mean depth (cm)	Range (cm)	Calculated figures from mean and StDev (cm)
Series 1	57	45 – 70	44.84 – 69.9
Series 2	59	47 – 70	47.35 – 70.3
Series 3	57	47 – 66	47.1 – 66.4
Series 4	50	38 – 63	37.7 – 62.6

Table 5.31. Means and ranges for the weathered depth.

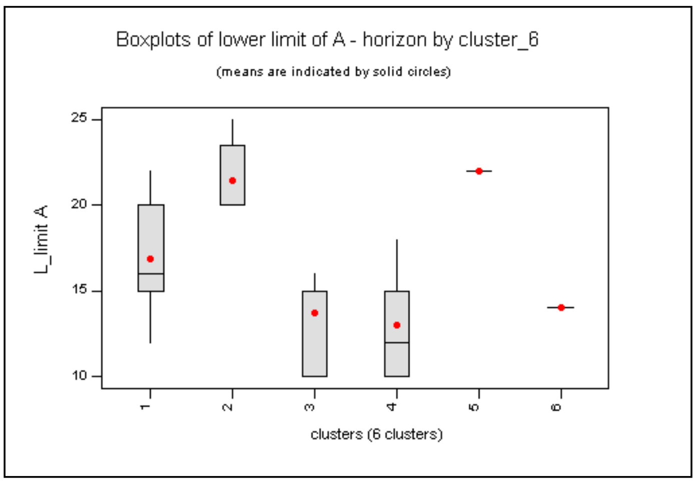
The thickness of the A horizon

Descriptive statistic on thickness of A horizon by clusters.

cluster_	N	Mean	Median	TrMean	StDev
1	49	16.857	16.000	16.844	3.048
2	5	21.400	20.000	21.400	2.191
3	7	13.714	15.000	13.714	2.563
4	9	13.000	12.000	13.000	2.872
5	1	22.000	22.000	22.000	*
6	1	14.000	14.000	14.000	*
cluster_	SE Mean	Minimum	Maximum	Q1	Q3
1	0.435	12.000	22.000	15.000	20.000
2	0.980	20.000	25.000	20.000	23.500
3	0.969	10.000	16.000	10.000	15.000
4	0.957	10.000	18.000	10.000	15.000
5	*	22.000	22.000	*	*
6	*	14.000	14.000	*	*

Table 5.32 Descriptive statistic on thickness of A horizon by clusters

Box plot for A horizon thickness by clusters



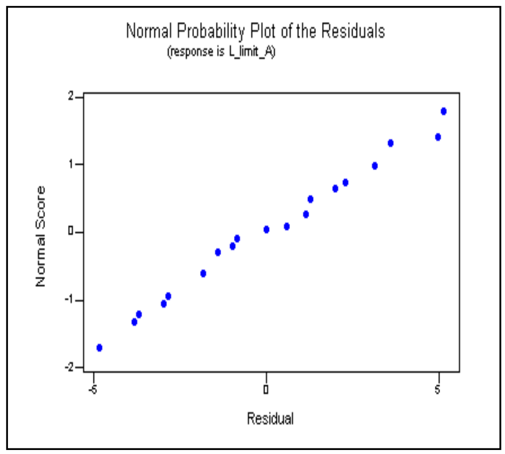


Figure 5.13. The box plot of the and the normal probability plot of the residual and the lower limit of A hriszons

Analysis of Variance for lower limit of A Horizon

Source	DF	SS	MS	F	P
Cluster_6	5	326.87	65.37	7.56	0.000
Error	66	570.63	8.65		
Total	71	897.50			

Table 5.33. Analysis of Variance for lower limit of A Horizon

```
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -------+---------+---------+---------
1          49    16.857     3.048                (*)
2           5    21.400     2.191                    (----*---)
3           7    13.714     2.563        (---*---)
4           9    13.000     2.872       (---*--)
5           1    22.000     0.000                (---------*--------)
6           1    14.000     0.000   (--------*---------)
                                   -------+---------+---------+---------
Pooled StDev =    2.940                12.0      18.0      24.0
```

Table 5.34 Individual 95% CIs

There are significant differences in the group means classified in the six clusters. There are essentially three groups, i.e. clusters 1 (moderately thick, in a narrow range centred around 17cm), 2 (thicker, central value 21.4cm), and (3 plus 4) (thinner, in the range of 13 to 14cm)

<u>The summary table for the A horizon depth:</u>

	Mean depth (cm)	Range (cm)	Calculated figures from mean and StDev (cm)
Series 1	17	14 - 20	13. 8 – 19.9
Series 2	21	19 - 24	19.2 - 23.6
Series 3	14	11 - 16	11.2 – 16.3
Series 4	13	10 - 16	10.1 – 15.9

Table 5.35. Summary table for the A horizon depth

<u>The depth of B horizon's</u>

Descriptive statistics table

cluster_	N	Mean	Median	TrMean	StDev
1	49	48.694	48.000	48.578	6.049
2	5	38.600	40.000	38.600	2.191
3	7	46.29	46.00	46.29	4.99
4	9	29.44	30.00	29.44	3.54
5	1	50.00	50.00	50.00	*
6	1	75.00	75.00	75.00	*
cluster_	SE Mean	Minimum	Maximum	Q1	Q3
1	0.864	40.000	60.000	45.000	55.000
2	0.980	35.000	40.000	36.500	40.000
3	1.89	40.00	55.00	42.00	50.00
4	1.18	23.00	35.00	27.50	31.00
5	*	50.00	50.00	*	*
6	*	75.00	75.00	*	*

Table 5.36. Descriptive statistics table

The boxplots and one way ANOVA

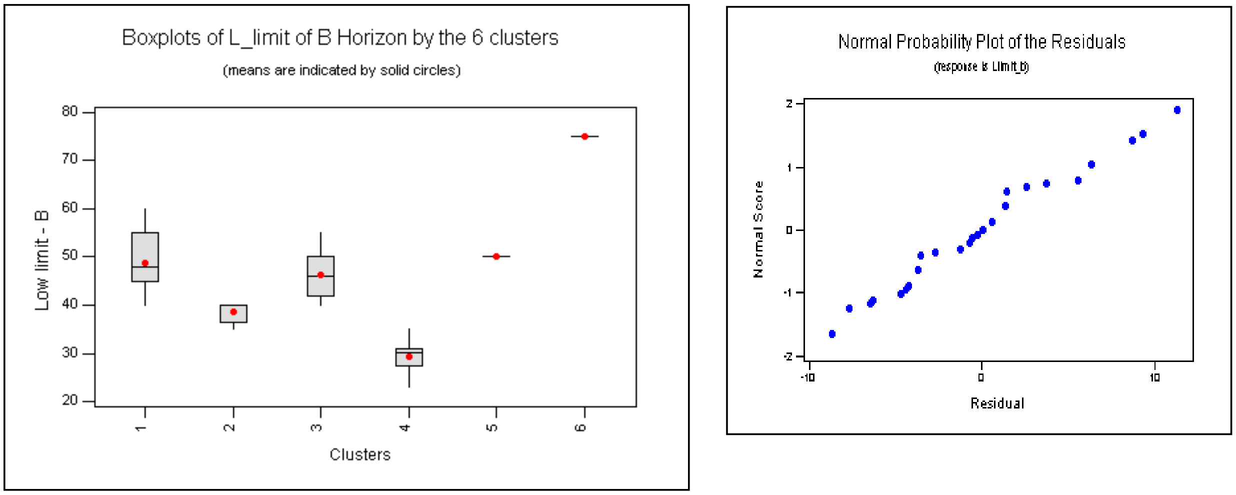


Figure 5.14. The Boxplot and the normal probability plot of the residuals for the lower limit of the B-Horizon.

Source	DF	SS	MS	F	P
Cluster_6	5	3948.7	789.7	25.74	0.000
Error	66	2025.3	30.7		
Total	71	5974.0			

Table5.37 Analysis of Variance for lower limit of B-horizon

```
                                         Individual 95% CIs For Mean
                                         Based on Pooled StDev
Level        N       Mean      StDev     --------+---------+---------+--------
1           49     48.694      6.049                 *)
2            5     38.600      2.191          (-*--)
3            7     46.286      4.990              (-*-)
4            9     29.444      3.539      (-*-)
5            1     50.000      0.000           (-----*-----)
6            1     75.000      0.000                        (-----*----)
                                         --------+---------+---------+--------
Pooled StDev =      5.539                        40        60        80
```

Table 5.38 Individual 95% CIs

Again here there are three significantly different groups of clusters 1, 2 and 4, can easily be distinguished from each other. Cluster 3 slightly overlaps 1 and 2. For the summary of the means and there range see the summary table (Table 5.38.) These results differ from the thickness of the A horizon in that clusters 3 and 4 are significantly different (they were in one group for the A horizon), in fact here cluster 3 and 1 are grouped. The difference between clusters 1 and 4 is similar to the result for the A horizon, but cluster 2 is now thinner than cluster 1, the opposite of the case in the A. Thus cluster 3 differs from cluster 4 in the thickness in the B only.

The summary table for the B horizon depth:

	Mean depth (cm)	Range (cm)	Calculated figures from mean and StDev (cm)
Series 1	49	43 - 55	42.6 – 54.7
Series 2	39	36 - 41	36.4 – 40.8
Series 3	46	41 - 51	41.3 – 51.3
Series 4	29	26 - 33	25.9 – 33.0

Table 5.39. Summary table for the tentative series by the B horizon depth

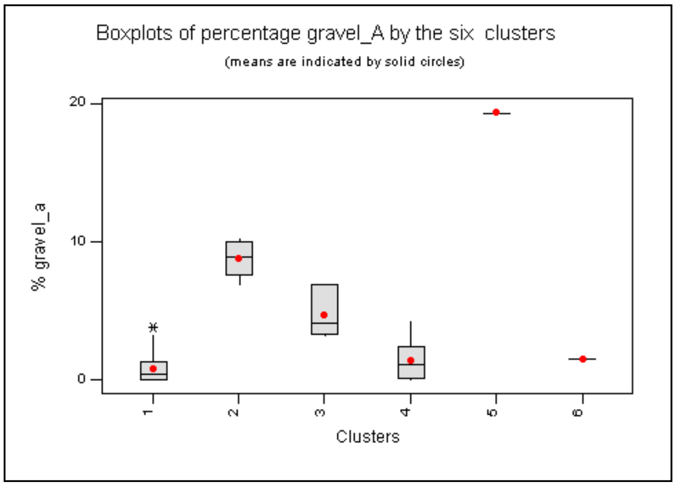
Amounts of Gravel in A horizon

Descriptive statistics on the Gravel percentage in the A horizon by clusters

cluster_	N	Mean	Median	TrMean	StDev
1	49	0.842	0.448	0.759	1.117
2	5	8.833	8.873	8.833	1.323
3	7	4.764	4.151	4.764	1.582
4	9	1.403	1.086	1.403	1.464
5	1	19.342	19.342	19.342	*
6	1	1.4903	1.4903	1.4903	*
cluster_	SE Mean	Minimum	Maximum	Q1	Q3
1	0.160	0.000	3.863	0.000	1.338
2	0.592	6.894	10.235	7.608	10.038
3	0.598	3.244	6.948	3.362	6.885
4	0.488	0.000	4.249	0.158	2.442
5	*	19.342	19.34	*	*
6	*	1.490	1.490	*	*

Table 5.40 Descriptive statistics on the Gravel percentage in the A horizon by clusters

The boxplots of the gravel percentage



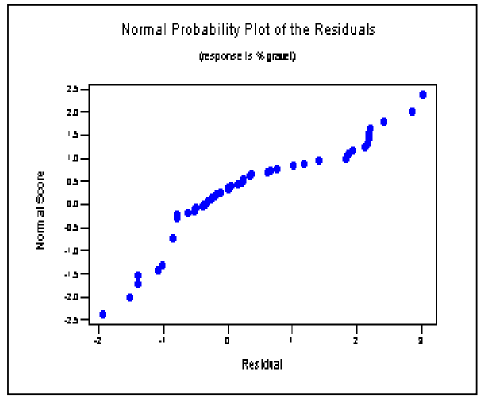


Figure 5.15. The box plot and the normal probability plot of the residuals for percentage gravel in the A_horizon.

Source	DF	SS	MS	F	P
Cluster_6	5	655.87	131.17	87.36	0.000
Error	66	99.10	1.50		
Total	71	754.97			

Table 5.41 Analysis of Variance for percentage gravel

```
                                  Individual 95% CIs For Mean
                                     Based on Pooled StDev
Level         N       Mean      StDev  --+---------+---------+---------+----
1            49      0.842      1.117     *)
2             5      8.833      1.323               (-*)
3             7      4.764      1.582         (-*)
4             9      1.403      1.464     (*)
5             1     19.342      0.000                            (---*--)
6             1      1.490      0.000   (--*---)
                                       --+---------+---------+---------+----
Pooled StDev =      1.225              0.0       7.0      14.0      21.0
```

Table 5.42 Individual 95% CIs

Both the boxplot and the analysis of variance show that the means of the groups are significantly (P = 0.000) different..

The four clusters can be grouped into three groups, of clusters (1 plus 4), 2 and 3. Cluster 2 has the highest gravel content, followed by cluster 3 and then clusters 1 and 4. The following table shows the mean and the range of each cluster.

The summary table for the gravel percentage in A horizon

	Mean Gravel Perc (%)	Range (%)	Calculated figures from mean and StDev (%)
Series 1	1	0 - 2	-0.3 - 2.0
Series 2	9	8 - 10	7.5 - 10. 2
Series 3	5	3 - 6	3.2 - 6.3
Series 4	1	0 - 3	-0.06 - 2.9

Table 5.43. Summary table for the mean and range of the percentage grave in the A horizon.

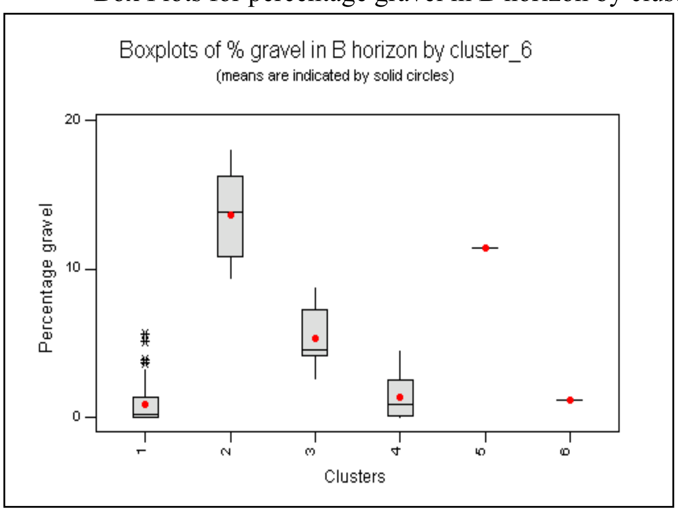
Amounts of Gravel in the B hrizon

Descriptive statistics for gravel percentage in the B horizon

cluster_	N	Mean	Median	TrMean	StDev
1	49	0.936	0.240	0.779	1.400
2	5	13.61	13.81	13.61	3.13
3	7	5.375	4.535	5.375	2.073
4	9	1.388	0.881	1.388	1.618
5	1	11.437	11.437	11.437	*
6	1	1.1903	1.1903	1.1903	*
cluster_	SE Mean	Minimum	Maximum	Q1	Q3
1	0.200	0.000	5.647	0.000	1.359
2	1.40	9.43	17.97	10.85	16.26
3	0.783	2.687	8.758	4.201	7.237
4	0.539	0.000	4.521	0.140	2.574
5	*	11.437	11.437	*	*
6	*	1.1903	1.1903	*	*

Table 5.44 Descriptive statistics for gravel percentage in the B horizon

Box Plots for percentage gravel in B horizon by cluster



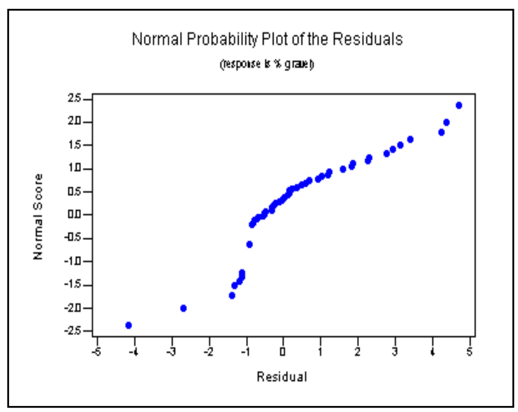


Figure 5.16. The Box plot and the normal probability plot of the residuals for the percentage gravel in the B horizon

Table 5.45 Analysis of Variance for % gravel B horizon

Source	DF	SS	MS	F	P
Cluster_6	5	886.87	177.37	65.04	0.000
Error	66	180.00	2.73		
Total	71	1066.87			

```
                                     Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -----+---------+---------+---------+-
1           49     0.936     1.400        (*)
2            5    13.605     3.131                              (--*--)
3            7     5.375     2.073               (--*-)
4            9     1.388     1.618        (-*-)
5            1    11.437     0.000                      (------*-----)
6            1     1.190     0.000   (-----*------)
                                    -----+---------+---------+---------+-
Pooled StDev =     1.651                0.0       5.0      10.0      15.0
```

Table 5.46 Individual 95% CIs

As can be observed from the boxplot and the analysis of variance test, The mean gravel percentage of the B horizon in the clusters are significantly (P = 0.000) different from each other.

Again here as was observed in the A horizon, only clusters 1 and 4 overlap each other while all others are distinctly defined. The results agree completely with those for the A horizon, thus the relative content of gravels is consistent between the A and B horizons in the clustering. Table below shows the Summary of the means and the range.

The summary table for the percentage gravel in the B Horizon

	Mean Gravel Perc (%)	Range (%)	Calculated figures from mean and StDev (%)
Series 1	1	0 - 2	-0.5 - 2.3
Series 2	14	10 - 17	10.5 - 16.7
Series 3	5	3 - 6	3.3 - 7.4
Series 4	1	0 - 3	-0.2 - 3.0

Table 5.47. Summary table for the mean and ranges in percentage gravel in the B Horizon.

5.2.6 Tentative series defined by clustering

Summarising the above results, we find four tentative series. The representative value is taken as the mean, and the normal range is taken as +/- one standard deviation, as shown above. The table below shows the class limit of the tentative series defined by multivariate classification. The process of cluster analysis more or less ensures that the central concepts of the four series are significantly different, if the dendrogram was cut at a realistic point, as here.

	Central values for the Response variables				
	W_Depth (m)	*LimitA (cm)*	*LimitB (cm)*	*GravelA (% wt)*	*GravelB (% wt)*
Series 1	57	17	49	1	1
Series 2	59	21	39	9	14
Series 3	57	14	46	5	5
Series 4	50	13	29	1	1

	Class limits for the Response variables				
	W_Depth (m)	*LimitA (cm)*	*LimitB (cm)*	*GravelA (% wt)*	*GravelB (% wt)*
Series 1	45 - 70	14 - 20	43 - 55	0 – 2	0 - 2
Series 2	47 - 70	19 - 24	36 – 41	8 – 10	10 - 17
Series 3	47 - 66	11 - 16	41 - 51	3 – 6	3 - 6
Series 4	38 - 63	10 - 16	26 - 33	0 – 3	0 - 3

Tables 5.48. The Summary tables of the means and ranges of the soil properties in the proposed series

We can describe these tentative series informally as follows:

Series 1: moderately thick solum, moderately thick A horizon, shallow to mod. deep lower limit Bw-horizon, non-gravely
Series 2: moderately thick solum, thick A horizon, shallow Bw horizon, gravely
Series 3: Similar to Series 1, but moderately gravelly and thinner A horizon, still in the 'moderately thick' class
Series 4: thin solum, moderately thick A Horizon and very shallow B-horizons, non-gravely

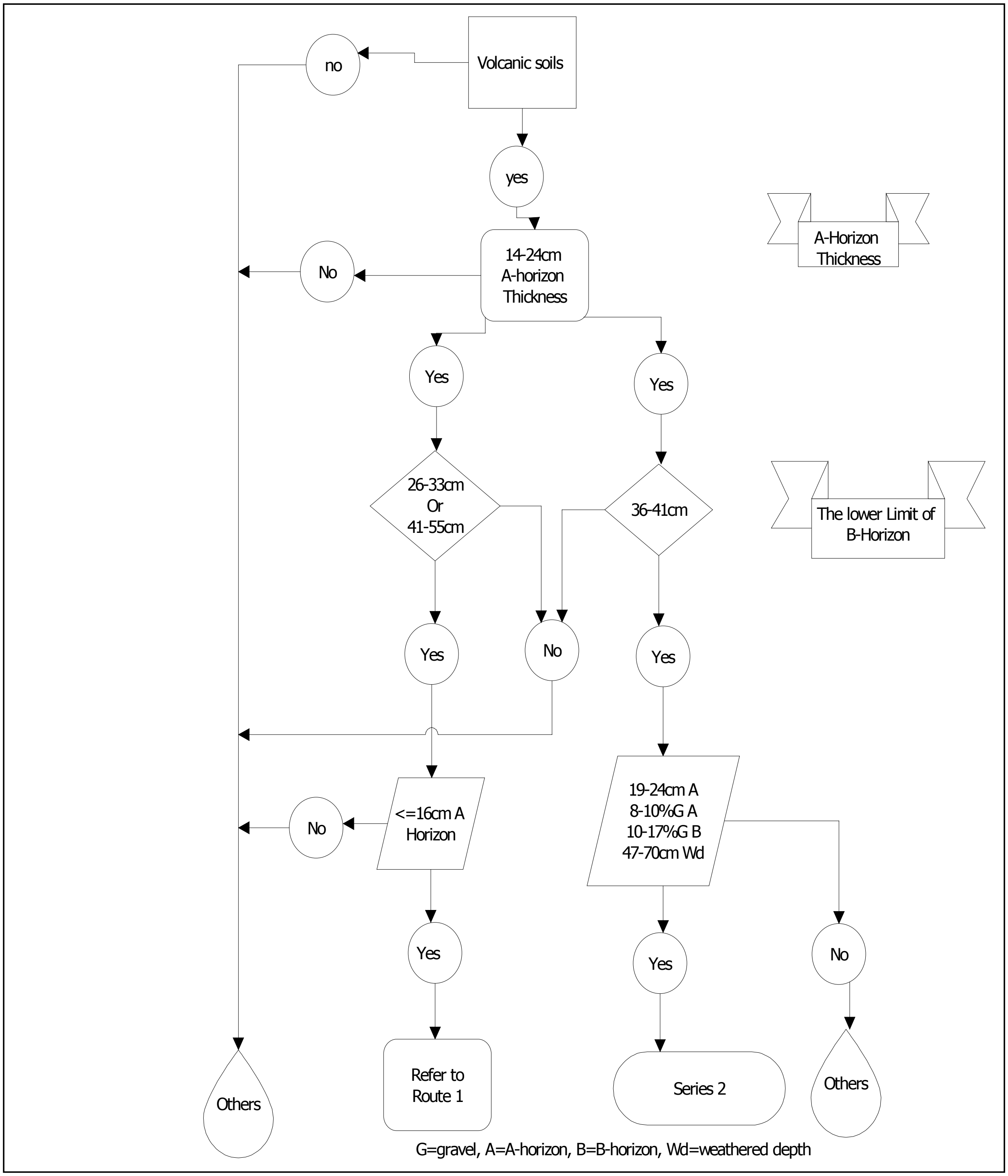
Note that:
The series 4 occurs in unit Pv-211 (Longonot series) and is a distinct cluster but could not be separated from series 1 under the GP map units. Series 4 also occurs in Pl-211 (Naivasha series)

Note: the term thin, mod thick, thick and very thick applied to horizons refers to <10cm, 10-20cm, 20-30cm, 30-40cm >40cm respectively, applied to the solum the limits are <30, 30-50, 50-100, and 100-150 and > 150 as very shallow, shallow, moderately deep, deep and very deep, respectively.

5.2.7 The field key to the series proposed (Cluster)

Here the key is developed based on the above numerical classification.

Route 1



A= thickness of A-horizon, B= depth of B-horizon, Wd= weathered depth, G A or G B= gravel % in A or B Horizon

Field Key continua

Route 2

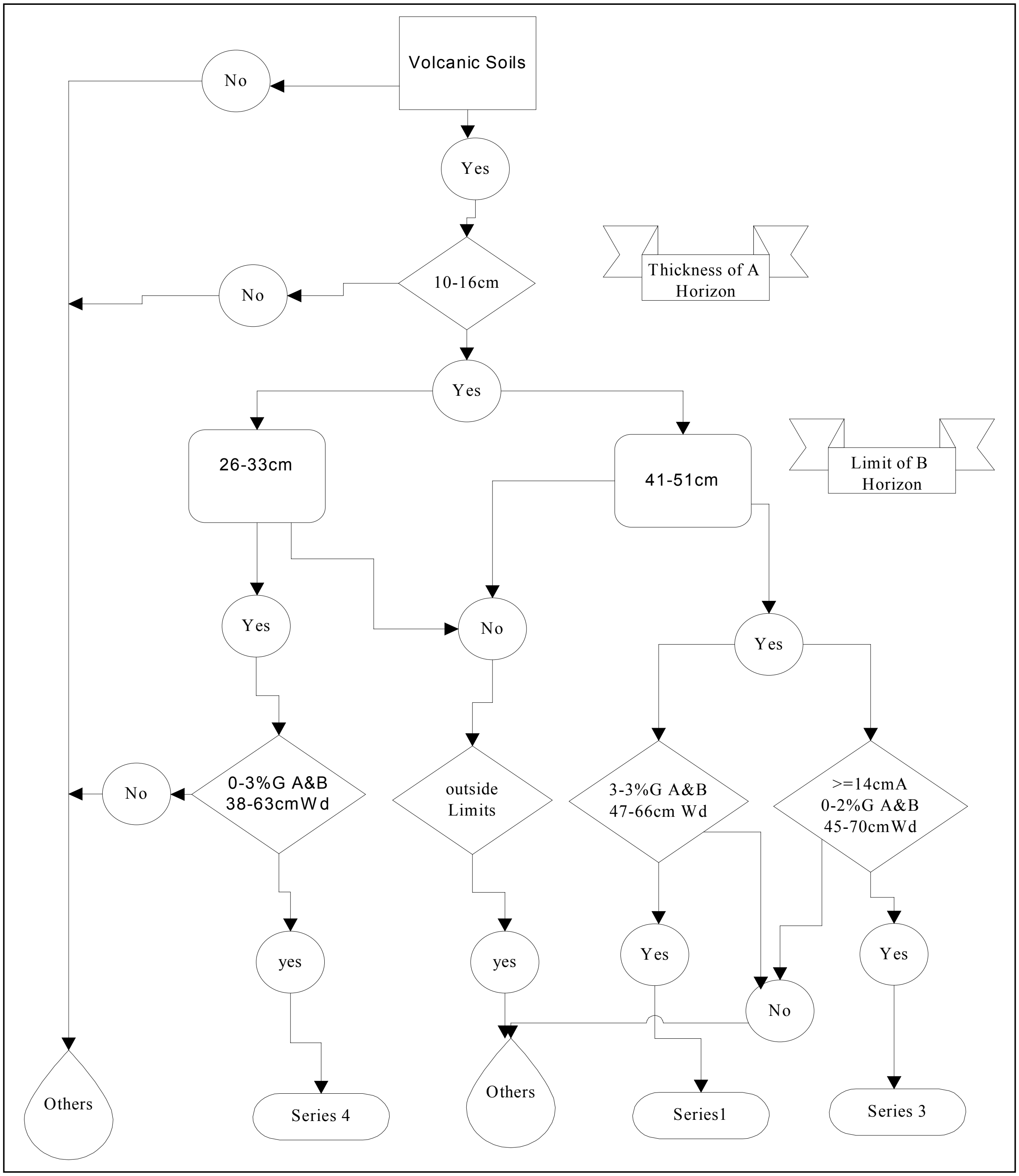
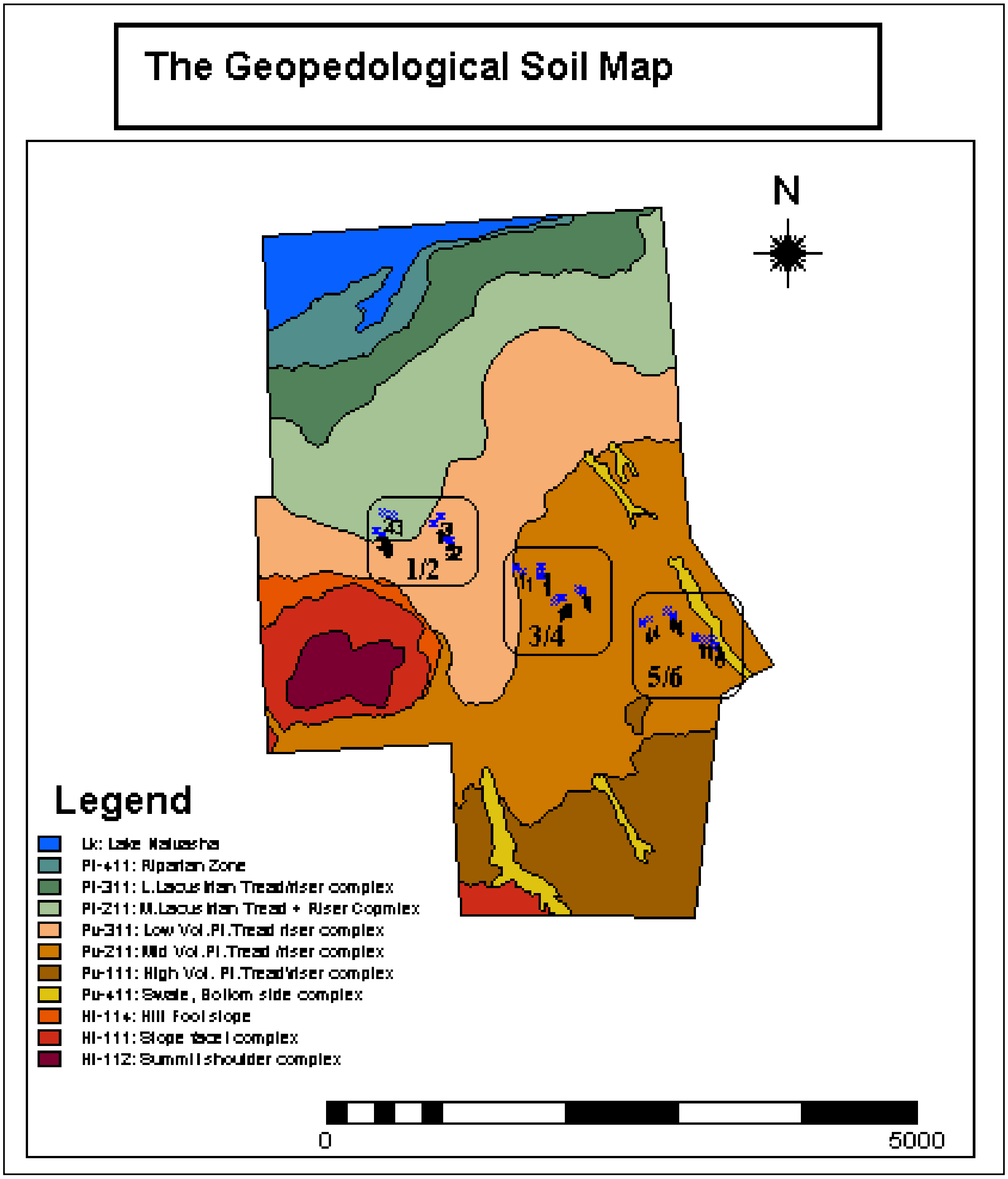


Figure 5.18 The Field key for the Series proposed (Route 2)

5.3 Spatial distribution of tentative series defined by cluster analysis

Do the tentative series (clusters) have a geographic expression? That is, could they be mapped? To answer that question, the clusters that are defined by the multivariate analysis were displayed on the geopedological (GP) soil map to the spatial location of the clusters among the GP unit. In ILWIS the attribute point map was created using the membership column. The map below shows the overlay of the GP soil map and the Cluster attribute map of the clusters. The clusters are labelled 1 to 6. Centres 1/2, 3/4 and 5/6 are highlighted and maps are displayed on the next page, figures 5.18-20 to give details of the clusters.



The detailed maps for the highlighted centres (Centres 1/2, 3/4, and 5/5).

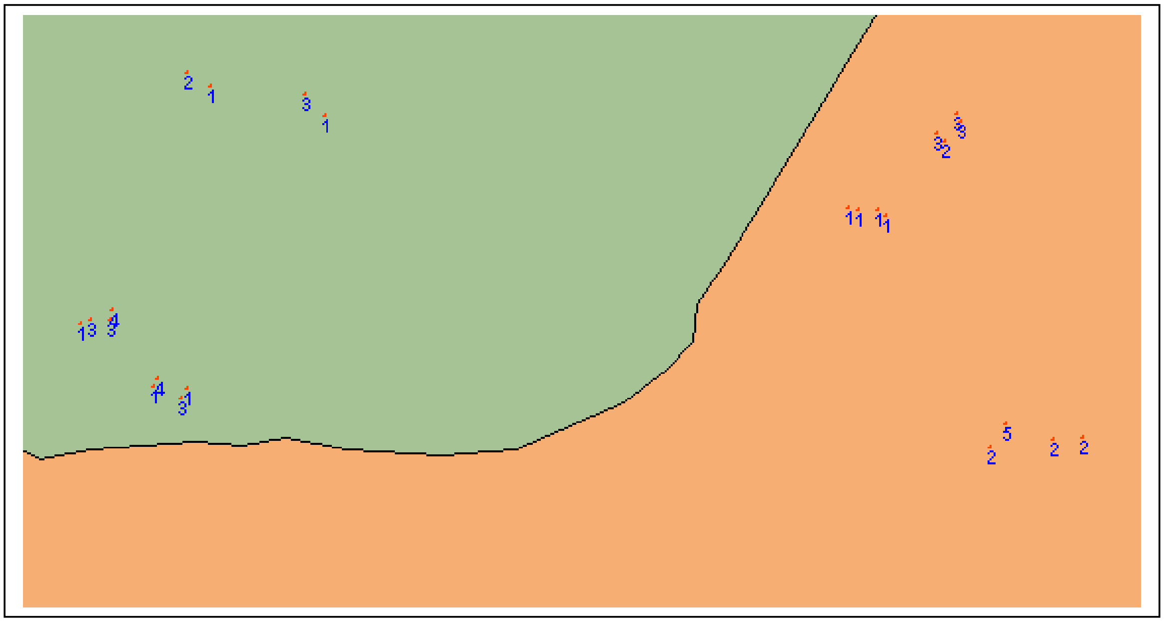


Figure 5.20 Centres 1/2

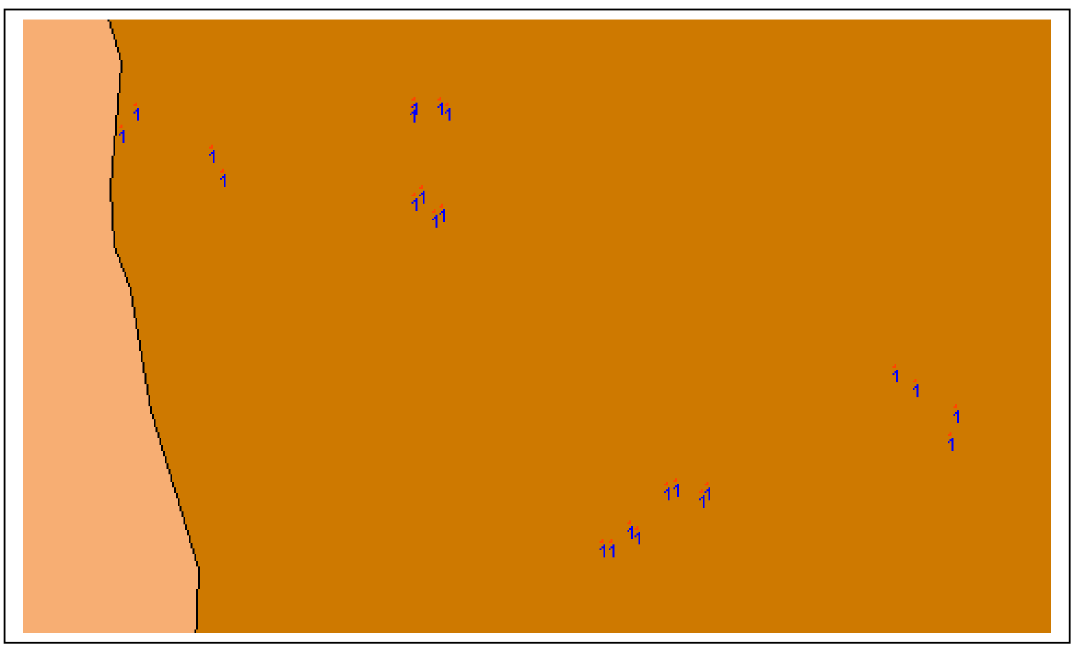


Figure 5.21 Centres 3/4

The detailed maps for the highlighted centres.

Figure 5.22 Centre 5/6

Figure 5.20-22. The detailed maps at the six centres, the numbers indicate the cluster to which it belongs.

Tabulated Statistics

		Clusters				
		1	**2**	**3**	**4**	**All**
Stations	**1**	5	1	4	2	**12**
	2	4	4	3	0	**11**
	3	12	0	0	0	**12**
	4	12	0	0	0	**12**
	5	5	0	0	7	**12**
	6	11	0	0	0	**11**
	All	**49**	**5**	**7**	**9**	**70**

Table 5.49 Cross Tabulation of stations by clusters displaying counts

The Cross Table with map units

		Clusters				
		1	***2***	***3***	***4***	**All**
Map units	***Pl-211***	5	1	4	2	**12**
	Pv-311	4	4	3	0	**11**
	Pv-211	37	0	0	7	**44**
	Pv-411	3	0	0	0	**3**
	All	**49**	**5**	**7**	**9**	**70**

Table 5.50. The cross table of Map units against Clusters

The results show that the soil series defined by multivariate analysis are distributed as follows:
Cluster 1 falls mainly in map unit Pv-211, but is also significant (more than 1/3 of the total samples) in Pv-311, Pv-411 and Pl-211. In Pv-211 it is dominant, but in spots it is closely mingled with Cluster 4 (within 6 to 19m) and could not be mapped at any realistic scale, even with grid sampling. It is the only cluster in Pv-411, but the small sample size leaves open the question of mappability. In map unit Pv-211 it is concentrated in two sub-stations, only 60m from another sub-station with soils from other clusters. Based on this evidence, a grid spacing of 60m or so would be needed to reliably separate it in Pv-211. In map unit Pl-211 it is intimately mingled with the other clusters at a range of 6 to 19m so it is not mappable at this scale.
Cluster 2 falls mainly within Pv-311 and a bit in Pl-211, it intimately associated with cluster 3 and 5 in Pv-211 at the ranges of 6m to 19m an so cannot be mapped at our scale. May be it can be mapped at the grid sampling of 19m since few observations were done in the map unit.
Cluster 3 falls in Pv-311 and Pl-211 but cannot be separated from the others spatially becaue it is intimately intermingled with other clusters and it occurs across two map units.
Cluster 4 also fall within map unit Pv-211 at the distance of 190m from cluster 1 and can possibily be mapped at the scale of 1:10 000. Here the grid sampling of 60m may clearly separate it from cluster 1.
Cluster 5 and 6 are single observations and they occur in the middle of other clusters. They are not considered as meaningful clusters for our purpose. It is interesting to see this evidence of inclusions within map units, which differ substantially from the bulk of the observations.

With this evidence, it may be advisable to add Cluster 3 to Cluster 1. This spatial evidence is in accordance with the closeness in multivariate space (see above).

Based on this evidence, it appears that the clusters defined by multivariate classification, while well separated in property space, do not form geographic clusters at any realistic scale, so cannot be mapped as such. Cluster 1 is indeed found in all four-map units, and could thus be mapped as the only series in the study area, with the others as inclusions. Considering cluster 3 as ‘similar’ to cluster 1, the map unit purity would meet the standards for a consociation.

5.4 Proposed Series (Synthesis)

Two classification methods were used to try to find series. An obvious question is to what extent do these methods give similar results The cross table of map units with the clusters above shows the count of the observations in each map unit, to help us to correlate the two.

The Cross Table with map units

		Clusters				
		1	*2*	*3*	*4*	**All**
Map units	***Pl-211*** ***(Naivasha)***	5	1	4	2	**12**
	Pv-311 ***(Sulmac)***	4	4	3	0	**11**
	Pv-211 ***(Longonot)***	37	0	0	7	**44**
	Pv-411	3	0	0	0	**3**
	All	**49**	**5**	**7**	**9**	**70**

Table 5.51. The cross table of Map units against Clusters

Naivasha Series

Observe that the Naivasha series is mostly a combination of cluster 1 and 3. On the map they occur intimately together and cannot be mapped separately. The clusters 1 and 3 are very similar except that cluster 3 is gravelier. This series has a mixed composition gravely patches. The ranges for the Naivasha series are therefore altered to accommodate the cluster 3 for the gravel range. This series is confirmed by the fact observed above that the clusters 1 and 3 intimately intermingled in the multivariate space in Pl-211.

Naivasha Series class limits

Volcanic soils
Over Gamblian Lake sediments
Sandy Loam or loamy sand (field),
Slightly-gravely (2 – 4) % A horizon and (3 – 8)% in B horizon, pumiceous gravel
Shallow weathering depth to mod. deep (41 – 70) cm
Moderately thick A horizon, (12 – 22) cm
Shallow to marginally mod. deep lower limit of Bw-horizon (36 - 55)cm

*{**Naivasha Series:** Sandy loam or loamy sand over Lacustrine sediments Shallow to mod. deep weathering depth, moderately thick A horizon, shallow to marginally mod deep lower limit of Bw-horizon, slightly-gravely}.*

Sulmac Series

The table shows that the Sulmac series Pv-311 have mainly cluster 2. Few of clusters 1 and 3 observations are also present in this series. Looking on the geographical distribution, cluster 1 observation is concentrated within 19m range which is not mappable under this scale. Neither can cluster 1 nor 3 observations be mapped, since they are few and are intimately occurring together with cluster 2. This also confirms that these two clusters are closely related in the multivariate space especially 2 and 3 i.e. they are both gravely. So again the Sulmac series can be justified.

Therefore the Sulmac series is a distinct series associated with low volcanic plain. The ranges for the gravel percentage and the depth of the weathered zone need to be altered to accommodate Series 2 from the numerical classification. So the Sulmac series is:

Sulmac Series class limits

Volcanic soils

Over volcaniclastic material

Gravely sandy loam/loamy sand

Gravely (4 – 12) % A horizon and (3 – 17)% in B horizon, pumiceous gravel

Slightly shallow to moderately deep, (47 – 70) cm

Moderately thick A horizon, (17 – 22) cm

Shallow to mod. deep lower limit of B(36 - 57) cm

*{*Sulmac Series***:*** *Gravely sandy loam/loamy sands to loam sands, over volcanic ash Slightly shallow to moderately deep, moderately thick A horizon, shallow low to mod. deep limit of Bw horizon, gravely}.*

Longonot Series

This series is associated with the mid volcanic plain (Pv-211) and mainly of cluster 1. Most observations of cluster 4 also occur in this series. But note from the limits that cluster 4 limit is within cluster 1. Cluster 4 has narrower ranges except of the gravel content, which is also almost the same. Note from the map (figure 5.18) also that cluster 4 can most likely be mapped separately with a grid sampling of 60m though it may just turn out to be an inclusion in cluster 1 even at that sampling grid. Since at on one side that they are mingled with cluster 1. The multivariate classification supports the Longonot series, since cluster 1 is the only one cluster, which can clearly be mapped. Though the boundaries may not necessarily conform to the GP lines. But this goes to show that this series as mapped has inclusions and is itself may occur as an inclusion in map units dominated by other series. Note that the Longonot series is very much similar to the Naivasha series by the properties under consideration but they develop on different parent material. To accommodate the cluster 4 from the numeric classification the gravel range has been widened a bit.

Longonot Series class limits

Volcanic soils

Over volcaniclastic material

Volcanic glass (very dark grey – greyish brown sand, 2.5Y3-5/1 and 5/2)

Sandy loam /loamy sand

Non gravely to slightly gravely < 3% gravel in A and B horizon.

Moderately deep (38 – 71) cm

Moderately thick A-horizon (13 – 18) cm

Shallow to mod. deep lower limit of B-horizon (26 - 56) cm

*{**Longonot Series:** Sandy loam to loamy sand, glassy, over volcanic ash moderately deep, moderately thick A and very shallow B-horizons, non-gravely to slightly gravely.}*

Note: the term thin, mod thick, thick and very thick refers to <10cm, 10-20cm, 20-30cm, 30-40cm >40cm respectively.

The summary of all the variable ranges defined by the proposed series

Means

	Means for the Response variables				
	W_Depth (m)	*LimitA (cm)*	*LimitB (cm)*	*GravelA (% wt)*	*GravelB (% wt)*
Naivasha	48	17	42	3.5	4.9
Sulmac	62	19	48	7.4	8.1
Longonot	58	16	46	0.4	0.4

Ranges

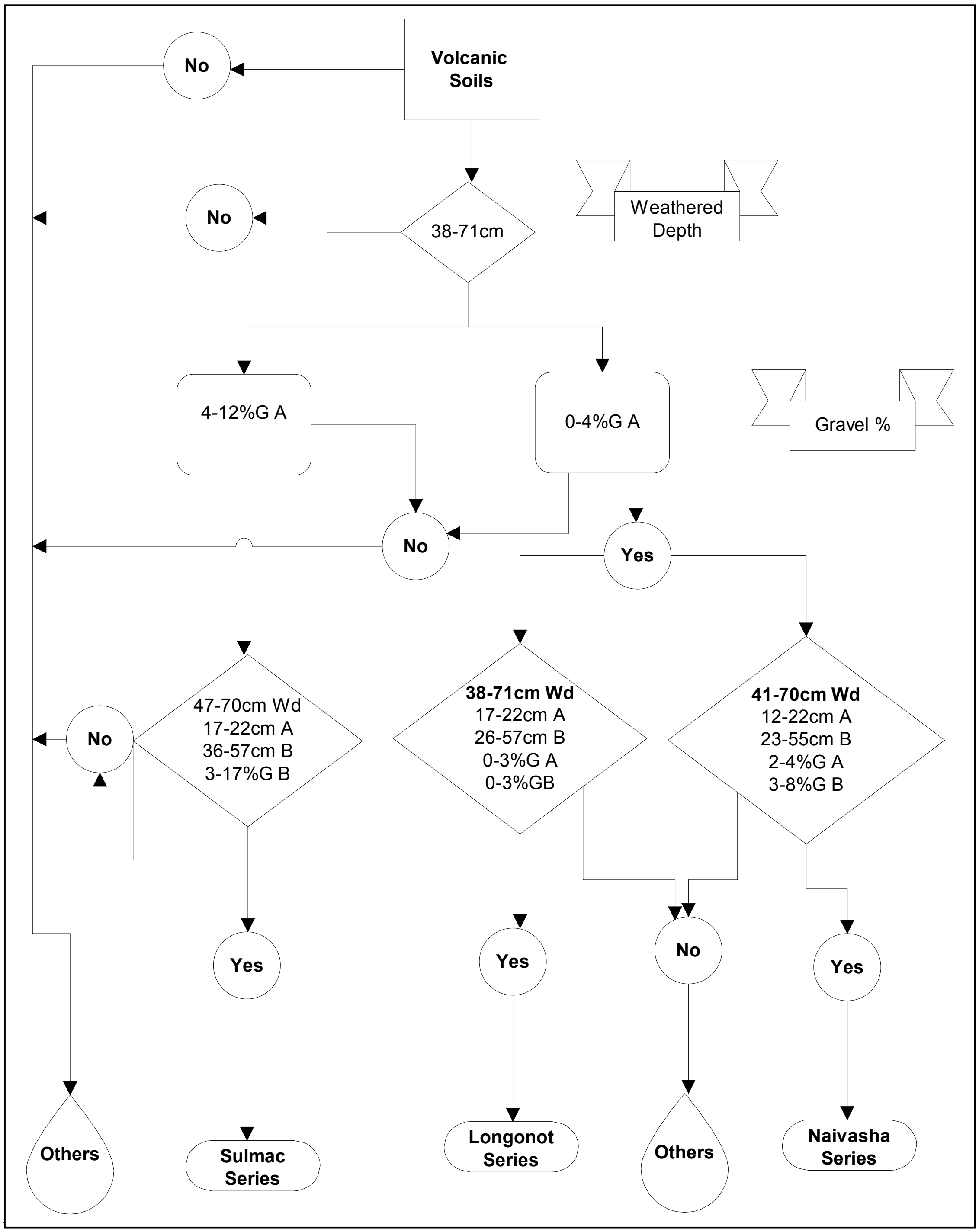
	Class limits for the Response variables				
Series	***W_Depth (m)***	***LimitA (cm)***	***LimitB (cm)***	***GravelA (% wt)***	***GravelB (% wt)***
Naivasha	41 - 70	12 - 22	36 - 55	2. – 4	2 – 8
Sulmac	47 – 70	17 - 22	36 - 57	4 - 12	3 – 17
Longonot	26 – 71	13 - 18	36 - 56	0 – 3	0 – 3

Tables 5.52. The Summary tables for the means and ranges of the variables defining the proposed series

5.4.1 Field Key of the proposed series

Notice that some ranges have been adapted to accommodate or separate unrealistic differences within and between the series.

Field Key for the proposed (synthesized) series.



5.4.2 Relation between soil series defined by multivariate analysis and Taxonomic systems

These series defined by multivariate analysis do not cross the class limits of the taxonomic lower level system i.e. the 2nd level WRB or the family level ST, at least on account of the variables under consideration. So under the *Arenic-Vitric Andosol (Eutric)* second-level WRB or *Ashy, Glassy, Isothermic Aridic Ustpsamment* ST family, three Series have been identifield and are within the lowest levels of the two hierarchical taxonomic systems.

There are very few numeric properties that are available to check this relationship. Samples were collected to determine particle size distribution and all sand separates, mineralogy, cation exchange activity, organic matter, pH and the electrical conductivity, but these laboratory results could not come in time for analysis.

It is clear that the lowest level of the taxonomic systems cover a wide range of feature space and hence more than one clusters defined by the multivariate analysis find themselves in one second-level group (WRB) or family (ST) defined by the taxonomic systems. For example the Naivasha series had the non-gravely and the slightly gravely clusters together. The Longonot series also had clusters 1 and 4, but cluster 4 is almost a subset of cluster 1. But its also worth noting that the multivariate analysis goes into details trying to separate even unmappable clusters as is the case in the Sulmac series where cluster 1 was fairly concentrated but too small a cluster to be mapped.

The multivariate classification assigns clusters based on similarity levels in the feature space and may not bear any practical implication, while the limits set in the hierarchical taxonomic systems are not natural clusters of soils but empirical limits for major differentiation in the land management practices. We therefore expect a taxonomic 'cross' or 'hiatus' in these two classification systems. However, in the present case, such a hiatus was not identified.

5.5 The structure of the spatial variability of soil properties.

The Fully Nested ANOVA was performed (hierarchical) analysis of variance and to estimate variance components for each response variable in at various spacings. Response variables on which this was done are weathered depth, lower limit of A horizon, lower limit of B horizon, gravel percentage of the A horizon and the gravel percentage of the B horizon. The Factors were stages 1 through 5, i.e successively closer spacings. The following are the tables showing the components of variance, cumulative variance and percentage variance and other necessary aspects.

Analysis of Variance for weathered depth									
Source	Distance	DF	SS	MS	Est Com Var	perc CompVar	StDev	Cum Var	% Cum Var
stage1	600	5	1686.692	337.3384	-63.028	0	0	214.556	100
stage2	190	6	4662.733	777.1221	162.437	75.71	12.745	214.556	100
Stage3	60	12	732.8417	61.0701	2.255	1.05	1.502	52.119	24.29
Stage4	19	17	940.6667	55.3333	12.127	5.65	3.482	49.864	23.24
Stage5	6	19	717	37.7368	37.737	17.59	6.143	37.737	17.59
Total		**59**	**8739.933**		**214.556**	**14.648**			

Table 1

Analysis of Variance for Lower limit of A horizon									
Source	Dist (m)	DF	SS	MS	Est Com Var	perc CompVar	StDev	Cum Var	% Cum Var
Stage1	600	5	203	40.6	2.197	14.17	1.482	15.502	100
Stage2	190	6	101.0167	16.8361	-0.336	0	0	13.305	85.82
Stage3	60	12	222.9833	18.5819	4.594	29.63	2.143	13.305	85.82
Stage4	19	17	128.8333	7.5784	-0.78	0	0	8.711	56.19
Stage5	6	19	165.5	8.7105	8.711	56.19	2.951	8.711	56.19
Total		**59**	**821.3333**		**15.501**	**3.937**		**15.501**	**100**

Table 2

Analysis of Variance for Lower limit of B Horizon									
Source	Dist (m)	DF	SS	MS	Est Com Var	perc CompVar	StDev	Cum Var	% Cum Var
Stage1	600	5	2019.6	403.92	19.444	21.3	4.41	91.269	100
Stage2	190	6	1046.75	174.4583	22.699	24.87	4.764	71.825	78.7
Stage3	60	12	907.1667	75.5972	20.179	22.11	4.492	49.126	53.83
Stage4	19	17	462.4167	27.201	-1.204	0	0	28.947	31.72
Stage5	6	19	550	28.9474	28.947	31.72	5.38	28.947	31.72
Total		59	4985.933		91.269	9.554		91.269	100

Table 3

Analysis of Variance for gravel percentage in A Horizon									
Source	Dist (m)	DF	SS	MS	Est Com Var	perc CompVar	StDev	Cum Var	% Cum Var
Stage1	600	5	450.5228	90.1046	6.58	48.94	2.565	13.445	100
Stage2	190	6	119.1663	19.861	3.16	23.51	1.778	6.865	51.06
Stage3	60	12	76.1104	6.3425	1.145	8.51	1.07	3.705	27.55
Stage4	19	17	60.1373	3.5375	2.167	16.12	1.472	2.56	19.04
Stage5	6	19	7.4646	0.3929	0.393	2.92	0.627	0.393	2.92
Total		59	713.4014		13.446	3.667		13.446	100

Table 4

Analysis of Variance for gravel percentage in B Horizon									
Source	Dist (m)	DF	SS	MS	Est Com Var	perc CompVar	StDev	Cum Var	% Cum Var
Stage1	600	5	597.7971	119.5594	7.669	39.09	2.769	19.62	100
Stage2	190	6	209.6961	34.9493	6.976	35.55	2.641	11.951	60.91
Stage3	60	12	55.1563	4.5964	-0.833	0	0	4.975	25.36
Stage4	19	17	110.5694	6.5041	3.388	17.27	1.841	4.975	25.36
Stage5	6	19	30.1624	1.5875	1.587	8.09	1.26	1.587	8.09
Total		59	1003.381		19.62	4.429		19.62	100

Table 5

Table 5.53-57 Tables 1-5 shows the Analysis of variance of each response variables with their components of variance at each stage.

The following table shows the summary of the components of variance for response variables mentioned.

Cumulated Components of variance for the response variables with the Lag											
		Cumulated Components of variance for the Response variable					Cumulated Percentage variance explained for the variables				
Stages	Dist (m)	W_depth	LimitA	LimitB	GravelA	GraveB	W_depth	LimitA	LimitB	GravelA	GraveB
stage1	600	214.56	15.50	91.27	13.45	19.62	100	100	100	100	100
stage2	190	214.56	13.31	71.83	6.87	11.95	100	85.8	78.7	51.1	60.9
stage3	60	52.12	13.31	49.13	3.71	4.98	24.3	85.8	53.8	27.6	25.4
stage4	19	49.86	8.71	28.95	2.56	4.98	23.2	56.2	31.7	19.0	25.4
stage5	6	37.74	8.71	28.95	0.39	1.59	17.6	56.2	31.7	2.9	8.1
Total		214.56	15.50	91.27	13.45	19.62	100	100	100	100	100

Table 5.58. Components of variance and percentage variance of weathered depth, Lower limit of A and B horizons and percentage gravel of A and B horizons contributed at each stage in the survey of the of Sulmac farm – Naivasha.

The table above shows the accumulated components of variance for each variable at the separating distance indicated. The plot of the above components variance against the separating distance on the logarithm scale is performed to give the first approximations of the variograms. See the graph below for the plot.

Variograms for all variables

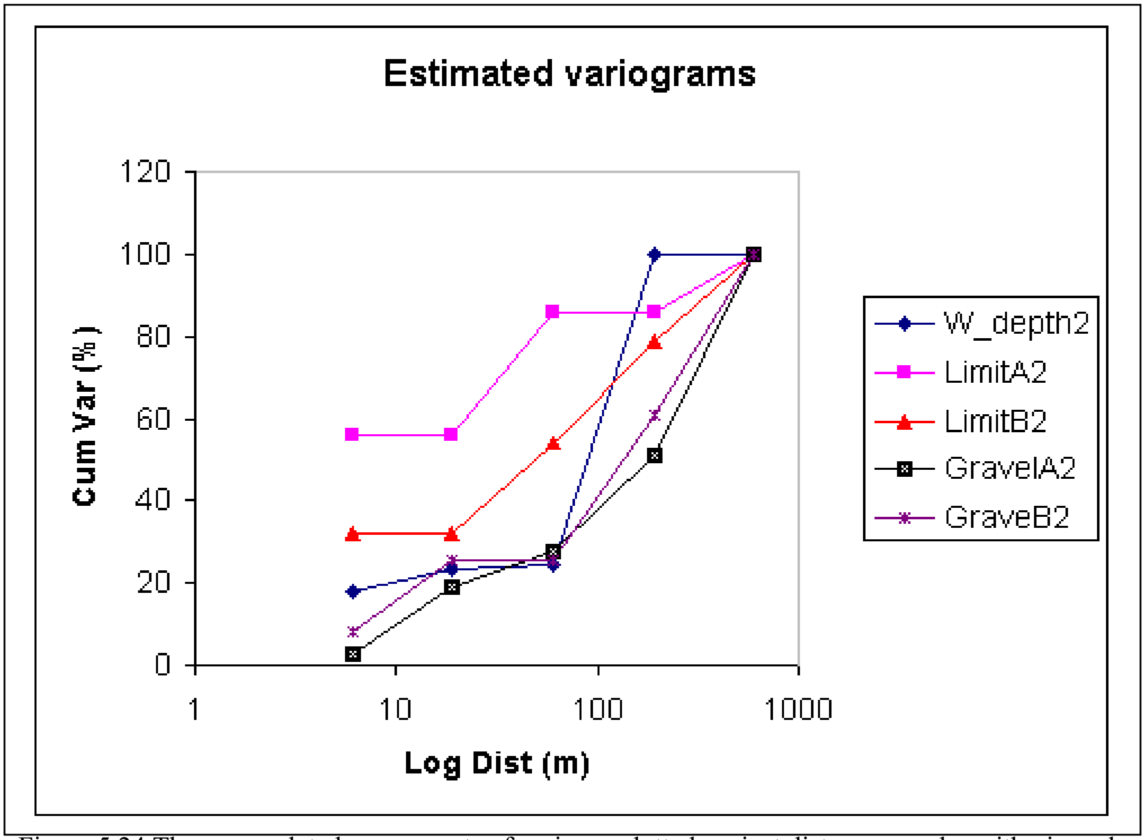


Figure 5.24 The accumulated components of variance plotted against distance on a logarithmic scale for the weathered depth, Lower limit depth of A and B horizons and gravel percentage of A and B horizon of the soils at Sulmac farm- Naivasha.

The results of the analysis for the response variable in question show that:

Weathered Depth

The analyses of the weathered depth in table and the plot shows that largest component of variance comes from the spacing of between 60m and 190m. At the spacing of 190m contributes 76 per cent of the total variance. The variogram shows that the variance is fairly constantly low from 6m through 60m spacing, and then it increases and reaches its maximum at 190m, (its sill, at h = 190m). The full extent of variation at this range seems to have been included Note that as the separating distance approaches zero; there is an unresolved variance, which is not zero. This is the called the **nugget** value.

Lower limit A horizon

For the lower limit of A horizons depth, 56% of the total variance occurs within the 6m separating distance stage 5. This is a very large nugget effect, indicating that local management (plowing, ridge-and-furrow pattern) has a very large effect on topsoil thickness. At 19m separating distance stage 4, there is no positive contribution. Then 30% variation is contributed at stage 3. At stage 2 there is again no positive to the variance. Stage 1 contributes 14%. So most of the non-nugget variance occurs within the distances of 19m and 60m.

The large unresolved variance at stage 5 and 4 may indicate that shorter distances would be needed to resolve that variance or it may just as well be that the nature of the property has that natural variability which is not spatially dependant.

Lower limit B Horizon

The lower limit of the B-horizon generally follows the pattern shown by the thickness of the A horizon, but with not such an extreme nugget. In this case, 29% variance was at stage 5 and it remained steady through stage 4, then increases steadily until it reaches its maximum at stage 1. Most variations occur within the distance of 19m to 600m. Note that 79% of the total variation is covered within 190m range.

Percentage gravel in A and B horizons

For the gravel percentage in A and B horizons, the variograms are similar, most of the variations are contributed between stage 3 and stage 1 i.e. the 60m to 600m, there is 73% contribution in the A horizon and 75% in B within the same distance. There is almost no nugget effect. From the graph they seems continue without limit, that is, wider spacing might reveal some residual spatial dependency.

Variable	Range of spatial variability
Weathered depth	Long
Thickness of A Horizon	Moderately long, very high nugget
Depth of B Horizon	Moderately long to long, high nugget
Percentage Gravel A horizon	Long to very long, no nugget
Percentage Gravel B horizon	Long to very long, very low nugget

Table 5.59

Key to distance (range) description

Range(m)	Description
< 6	Very short
6 - 19	Short
19 - 60	Moderately long
60 - 190	Long
190 - 600	Very long
> 600	Extremely long

Table 5. 60

Table 5.59-60 Variable variability range and key tables

Chapter 6: CONCLUSIONS AND RECOMMENDATIONS

6.1 Adequacy of proposed soil series for intensive management

With respect to the hypothesis, "Soil series that are homogenous enough for intensive management can be defined in Lake Naivasha area"

This study shows that homogeneous units can be defined for intensive land management if the right sampling strategy is employed for mapping the soil properties. The multivariate approach adequately and successfully separates homogenous units for this purpose as has been demonstrated on the variables that have been considered. For intensive management purposes the gravelier Sulmac series can be managed differently from the non-gravely Naivasha and the Longonot series.

6.2 Relation between soil series and taxonomic systems.

With respect to the hypothesis, "Soil series are more homogeneous than the higher levels of 2nd level WRB and the US Soil Taxonomy family classification, and the majority of the pedons classified to a series also classify to one higher level"

The results show that all three series classify to one family as *Ashy, Glassy, Isothermic Aridic Ustipsamment* on the US Soil Taxonomy classification, assuming there is less than 30% unweathered volcanic glass in the sand and silt fraction. The WRB 2nd level classification classifies all the three series under *Areni- Vitric Andosol (Eutric).* Therefore it is true that series are more homogeneous than higher level taxonomic classes. The series by the multivariate classification can produce even more homogeneous clusters of series within one higher level classification, as the ranges at this level can be even narrower.

6.3 Relation between soil series and consociations mapped at 1: 10 000.

With respect to the hypothesis, "Soil series can be mapped as consociations at 1: 10 000, i.e. in polygons of minimum size 0.4ha. That is, the named series and similar soils occupy at least 75% of the mapped area; no more than 15% of the area can contain soils that differ significantly from the named soil and which are more limiting to major land uses."

This study shows that the series defined by both conventional methods and the multivariate classifications are very homogenous and the soils within the series do not contain significantly different soils i.e. which is more limiting to major land uses. From the multivariate classification, some series had inclusions, which intimately mingled with the major clusters, but the differences between them could not significantly inhibit major land uses. This is due to the fact that that the ranges permitted in these series are very narrow, sometimes are without any practical significance and also that the inclusions may not form any cluster at all as was the case with cluster 3 in the Naivasha series.

6.4 Relation between ranges of series mapped by the multivariate analysis and that of hierarchical classification.

With respect to the hypothesis, "series mapped by the multivariate analysis (numerical classification) do not always fall within the range of the hierarchical classification (Taxonomic classification)."
From the results we do not have many numeric properties with which we could use to characterise the soils and to perform the multivariate analysis to compare the ranges to see if indeed the ranges from numerical classification do not respect the range of the hierarchical system.
(It was hoped that the particle size distribution classes, mineralogy classes, cation exchange activity classes and organic carbon classes pH and Electrical conductivity classes would have revealed this fact if they were available for analysis. But unfortunately the laboratory results were late).
However one would guess that the case is most likely to be true as we have seen that the multivariate classification uses similarity levels to form clusters, while the Taxonomic systems pick a range which has been empirically tested for practical differentiation in management.

6.5 The range of spatial variability of soil properties in the study area

With respect to the hypothesis, "The spatial variability of soil properties varies within short distances, but mostly within the range of a series."
As can be observed from the estimated variograms figure 5.24 (of the weathered depth, lower limit of A and B horizons, and gravel percentage of A and B Horizon that), the spatial variability of these properties in geographic space varies. The spatial variability ranges from 19m to 190m as can be seen from the table below. Indeed these results show that the variability is within the range of a mappable series of 1:10 000 scale (i.e. minimum delineation area 0.4 ha i.e. $\sqrt{(4000)} = 63.2$m minimum range) and even on a larger scale especially on the A horizon thickness. Therefore in this area sampling intervals can be designed wisely considering these findings with respect to the above properties.

Variable	Range of variability (m)
Weathered depth	Long (60-190)
Thickness of A Horizon	Moderately long (19 –60)
Depth of B Horizon	Moderately long to long (19 – 190)
Percentage Gravel A horizon	Long to very long (60 – 600)
Percentage Gravel B horizon	Long to very long (60 – 600)

Table 6.1 Variable variability ranges

6.6 Recommended mapping strategy for intensive soil management

As has already been established the spatial variability of the properties considered reveals that the spatially dependent variation is mostly between 19m and 190m. For the mapping strategy for intensive soil management, the sampling grid of 60m will successfully and economically map out the thickness of the A-horizon and the depth of the B-horizon. For weathered depth and the gravel content a grid sampling of 190m is efficient and recommendable.
The series are homogeneous units, which are ideal for the intensive land managers.

Chapter 7: REFFERENCES AND APPENDICES

7.1 References

1. Butler, B. E., (1980). Soil classification for soil survey. Monographs on soil survey, Oxford University Press. Oxford.
2. Burrough, P. A., (1986). Sampling designs for quantifying map unit composition, in Spatial variabilty of soils and land forms, M. J. Mausbach and L. P. Wilding, Editors. 1991, Soil Science Society of America. Madison.
3. Clarke, M. C. G., Woodhall, D. G., Allen, D. and Darling, G. (1990). Geological, volcanological and Hydrological controls on the occurrence of geothermal activity in the area surrounding Lake Naivasha, Kenya, British geological survey, Nottingham.
4. Clayden, B and J. M. Hollis, (1984). Soil survey: Criteria for differentiating soil series. Technical Monograph No. 17., Rothambsted Experimental Station, Harpenden.
5. de Gruijter, J. J., (1977). Numerical classification of soils and its application in soil survey. Agricultural Research reports. Vol. 855., PUDOC. Wageningen.
6. de Gruiter, J. J. D. J. J. Walvoort, et al. (1997). "Continuous soil maps – a fuzzy set approuach to bridge the gap between aggregation levels of process and distribution models." Geodema 77. Elsevier. Amsterdam.
7. Dent, D and A. Young, (1981). Soil survey and land evaluation; George Allen & Unwin., London, England
8. FAO (1984). Guidelines: levels evaluation for rainfed Agriculture FAO soil bulletin no. 52.
9. FAO and ISRIC (1990). Guidelines for soil description, 3rd edition. FAO, Rome.
10. Frank, F. (1984). Statistics for land managers, Paeony Press, Jedburgh (Scotland).
11. Hamududu B. H (1998). Erosion assessment for large basins using remote sensing and GIS, Naivasha basin, Kenya. MSc report, ITC, Enschede.
12. Harper D. M., Mavuti, K. M. and Muchiri, S M. (1990). Ecology and Management of Lake Naivasha, Kenya in relation to the climatic change, Alien species, Introduction and Agricultural Development. Environmental Conservation, Vol. 17, No. 4. Kenya Government.
13. Ibanez, J. J, Zinck, A. J, and Jimenez-Ballesta, R. (1992). Soil survey: old and new challenges. In: Soils survey: perspectives and strategies for the 21st century (workshop). Report No. 21, ITC, Enschede.
14. ILACO B.V (1981). Agricultural compendium, Elsevier, Amsterdam.
15. ILWIS Dept (1997). ILWIS 2.1 user's guide, ITC, Enschede.

16. Kamoni P. T (1988). Detailed soil survey of a part of quarantine farm-National Husbandry Research station, Naivasha, Nakuru District, Kenya. KSS report. Kenya.
17. Kauffman, J. H, Minzhu, W, and Bin, Z. (1995). Reference soil of tropical China (Hainan Island), ISRIC, Wageningen.
18. Kwacha, C. P. H. J (1998). Vulnerability of soils to change in Agricultural use around Lake Naivasha, MSc thesis Report, ITC, Enschede.
19. Lance, G. N and Williams, W. T (1966). A Generalised sorting for computers classifications Nature, London.
20. Legendre P and Legendre L. (1998). Numerical Ecology, 2nd Edition, Elsevier Science B. V. Amsterdam.
21. LNROA, (1995). Lake Naivasha Riparian Owners Association report. Naivasha.
22. LNRAO and KWSTI, (1996) Lake Naivasha Management implementation committee stakeholders workshop report, 8 October 1996. Kenya Wildlife Service Training Institute. KWSTI, Naivasha.
23. LNRAO, (1993). Three phase Environmental Impact study of recent developments around Lake Naivasha, Phase 1 Report, John Goldon Associates, Nairobi.
24. Mausbach, M. J and Wilding L. P (1991). Spatial variability of soils and landforms. SSSA spacial publication number 28, SSSA Inc, Madison.
25. MINITAB Co (1998). MINITAB 12.22 statistical program, MINITAB Inc.
26. Moore, S. D. (1997). The Active Practice of Statistics. W. H. Freeman and Company, New York.
27. Nauta, D. (1971). Discussion paper on the Alkaline characteristcs of the soils at the lower farm and problems caused thereby regarding irrigation practices – Seminar. 7 January, 1971, National Animal Husbandry Reseach Station, Naivasha.
28. Rossiter, D. G., (2000). Methodology for soil resourse inventory, lecture notes and references. 2nd Revised version, SOL. 27., ITC, Enschede.
29. Siderius W. (1980). Soils condition at Kulya Farm (Naivasha). KSS Report, Kenya.
30. Siderius W. and Muchena F. N. (1977). Soils and Environmental conditions of Agriculture Research Stations in Kenya. KSS Report M5. Kenya.
31. Soil Survey Division staff, (1993). Soil survey manual, USDA Handbook No. 18. US Department of Agriculture. Washington DC.
32. Soil Survey Division Staff, (1975). Soil taxonomy: a basic system of soil survey classification for making and interpreting soil surveys. Agricultural Handbook 436. US Department of Agriculture Soil Conservation Service, Washington DC.
33. Soil Survey Division Staff, (1998). Keys to Soil taxonomy, US Department of Agriculture Natural resource Conservation Service, Washington DC.
34. Sombroek. W.G., H.M.H.Broun and B.J.A.Van der Pouw, (1980). Exploratory soil survey map and agro-climatic zone map of Kenya, Scale 1:1,000,000. Exploratory soil survey reports No. E1, 1982, Kenya Soil Survey, Nairobi.

35. Sumner M. E. (2000). Handbook of soil science, CRC Press, Washington DC.
36. Thompson A. O. and Dodson R. G (1963). Geology of Naivasha area, Government of Kenya. Nairobi.
37. Van Wambeke A (1982). Africa soil Moisture regimes, SSMS technical monograph, No 3, SMSS & AID, New York.
38. Webster, R. and M.A Oliver, (1990). Statistical methods in soil and land resource survey. Oxford University Press. Oxford.
39. Wilding, L. P (1984). Spatial variability: its documentation, accommodation and implication to soil surveys in Soil spatial variability workshop of the ISSS and the SSSA in Las Vegas, USA, Pudoc, Wageningen, The Netherlands.
40. Wilding, L. P. and Drees, L. R. (1983). Spatial Variability and pedology. In pedologenesis and soil Taxonomy I. Development and soil science 11A Elsevier publishing co. Wageningen. The Netherlands.
41. Youden, W. J. and Mehlich, A. (1937). Selection of efficient methods for soil sampling. Contributions of the Boyce Thompson Institute for Plant Research 9: 59-70. New York.
42. Zinck, J. A (1988/89) Physiography and soils. Lecture notes, SOL 41, ITC, Enschede.
43. Zinck, J. A (1992). Soil surveys: Perspectives and strategies for the 21st century (workshop), No. 21, ITC, Enschede.

7.2 APPPENDICES

7.2.1 Appendix A: Profile descriptions of Reference Profiles and Minipits

1. Reference profiles

Site and Profile Description

Profile ID:	**PT 01**
Classification:	FAO: Areni-Vitric Andosol (Eutric) USDA: Ashy, Glassy, Isothermic Aridic Ustpsamment Diagnostic criteria FAO: Ochric /Andic horizon, vitric, and arenic properties Diagnostic criteria USDA: Ochric epipedon Soil moisture regime: Ustic Soil temperature regime: Isothermic
Author:	Simfukwe P.
Date:	13/09/00
Location:	Field 733 middle of the farm of Sulmac Longonot Farm UTM 207718, 9906308 on the low volcanic plain
General Landform:	The physiographic position is on the tread riser complex of the low volcanic plain, gently undulating topography and is gently sloping gradient- slope 2%, Soil erosion not observed. No micro relief.
Parent Material:	Volcanic ash and pyroclastics.
Landuse/Vegetation:	Formally four years ago was used for Carnations, under the ploughing and drip irrigation. At the time of survey it was fallow. About 50% grass cover,
Drainage:	Excessively drained, rarely saturated, the soil was very dry during the time of survey. No evidence of flooding, nor ground water.
Presence of rocks/stones	No stones/rock fragments observed on the surface.
Effective Depth:	Very deep more than 150 cm deep
Remarks:	

Horizon	**Depth(cm)**	Description
Ah	0-14	Very dark greyish brown (2.5Y3/2) moist, and olive brown (2.5 Y4/3) dry loamy sand (field); very weak fine, sub angular blocky structure; soft consistence when dry; none sticky none plastic; few fine and common very fine roots; penetration 1kg/cm^2; clear and smooth boundary to
Bw1	14-35	Dark olive brown (2.5Y5/3) moist and olivine brown 2.5Y4/3 dry loamy sand (field); very weak course, sub angular blocky structure; soft consistence when dry; none sticky none plastic; few fine and few very fine roots; penetration >4.5 kg/cm^2; clear

		and smooth boundary to
Bw2	35-62	Dark olive brown (2.5Y5/3) moist and olivine brown 2.5Y4/3 dry loamy sand (field); single grain structure, soft consistence when dry; none sticky none plastic; few fine and few very fine roots; Very few channels (θ, 1-3cm) penetration 1.5 kg/cm^2; abrupt and smooth boundary to
Bm	62-68	Light brownish grey (2.5Y6/2) moist and olivine brown 2.5Y4/3 dry loamy sand (field); massive structure, very hard consistence when dry; none sticky none plastic; massive weakly cementation/induration by silica; few fine roots; penetration >4.5 kg/cm^2; abrupt and smooth boundary to
C	68-115	Very dark grey (5Y3/1) moist and grey (5Y5/1) sand (field); single grain structure, loose consistence when dry; none sticky none plastic; few fine roots; Very few channels (θ, 1-3cm) penetration 0.25 kg/cm^2; abrupt and wavy boundary to
2Cm	115-155+	Very dark grey (5Y3/1) moist and grey (5Y6/1) course sand (field); massive structure, hard consistence when dry; none sticky none plastic; slightly weathered very few fine angular pumice and obsediun rock fragments; massive weakly cementation/induration by silica; few fine roots; penetration >4.5 kg/cm^2; no boundary recorded downwards.
		Remarks: Laomy sand gets courser as you go down to Bw, in 2Cm there are a few pumice gravels. (θ =7cm) obsediun boulders were found in Bw. In 2Cm were chips of glassy materials.

Reference Profile no.2

Site and Profile Description

Profile ID:	PT 02
Classification:	FAO: Andic Fluvic Cambisol, (Skeletic, Eutric) USDA: Ashy-pumiceous Mixed, Isothermic Vitriandic Haplustept. Diagnostic criteria FAO: Ochric epipedon, cambic, sub surface fluvic, and skeletic properties. Diagnostic criteria USDA: Ochric epipedon Soil moisture regime: udic /Ustic. Soil temperature regime: Isothermic
Author:	Simfukwe P,
Date:	14/09/00
Location:	Near the Sulmac pump station on the Lacustrine plain UTM 204648, 9908546.
General Landform:	The physiographic position is on the tread riser complex of the low Lacustrine plain, almost flat topography, nearly level gradient- slope 0.8%, Soil erosion not observed. No micro relief.
Parent Material:	Lake sediments mainly clayey with considerable amounts of pyroclastic material, silts and gravel.
Landuse/Vegetation:	Nature protected area without interference, grasses are grazed by hippos/ wild life.
Drainage:	Well drained, saturated for short periods in most years, with slow run-off, no evidence of flooding, ground water table was > 2m deep. It was moist during the time of survey.
Presence of rocks/stones	No stones/rock fragments observed on the surface.
Effective Depth:	Very deep more than 150 cm deep
Remarks:	

Horizon	**Depth(cm)**	Description
O	0-4	Very dark greyish brown (10YR3/3, dry) loam (field); single grain structure; loose consistence when dry; none sticky none plastic; very few fine slightly weathered sub-rounded pumice stones; many fine channel pores; many very fine to fine roots; abrupt and smooth boundary to
Ah	4-9	Very dark greyish brown (2.5Y3/2, moist) and olive brown 2.5Y4/3, dry) loam (field); single grain structure; loose consistence when dry; none sticky none plastic; few fine to medium slightly weathered sub-rounded pumice stones; many fine channel pores; many fine to medium roots; abrupt and smooth boundary to

A2	9-25	Olive brown (2.5Y4/3, moist) gravely loam (field); weak, fine to medium sub-angular blocky structure; very friable moist consistence; slightly sticky none plastic; common, fine to medium slightly weathered rounded pumice gravel. Many very fine interstitial/ channel pores; few very fine to fine roots. Clear and wavy boundary to
Bw1	25-47	Greyish brown (2.5Y5/2, moist) gravely sandy loam (field); weak, fine to medium sub-angular blocky structure; very friable moist consistence; slightly sticky none plastic; common, fine to medium slightly weathered rounded pumice gravel. Many very fine interstitial/ channel pores; very few very fine roots. Clear and wavy boundary to
Bw2	47-67	Greyish brown (2.5Y5/2, moist) very gravely sandy loam (field); weak, fine to medium sub-angular blocky structure; very friable moist consistence; slightly sticky none plastic; many, fine to medium slightly weathered sub rounded pumice gravel. Many very fine interstitial/ channel pores; very few very fine roots. Clear and smooth boundary to
Bw3	67-86	Light olive brown (2.5Y5/3, moist) gravely sandy loam (field); very weak, medium sub-angular blocky structure; very friable moist consistence; slightly sticky none plastic; dominant, fine to medium slightly weathered rounded pumice gravel. Many very fine interstitial/ channel pores; very few very fine roots. Clear and smooth boundary to
2C	86-93	Greyish brown (2.5Y5/2, moist) sandy loam (field); weak, medium platy structure; very friable moist consistence; slightly sticky none plastic;. many very fine channel pores; very few very fine roots; abrupt and smooth boundary to
3Bw	93-103	Greyish brown (2.5Y5/2, moist) gravely sandy loam (field); weak, fine to medium sub-angular blocky structure; very friable moist consistence; none sticky none plastic; abundant, medium size slightly weathered sub rounded pumice gravel; very few very fine roots; abrupt and smooth boundary to
4C	103-118	Light olive brown (2.5Y5/3, moist) sandy loam (field); strong, very fine platy structure; very friable moist consistence; none sticky none plastic; abundant, medium size slightly weathered sub rounded pumice gravel; many very fine channel pores; olive yellow (2.5Y6/6) few to many, fine, distinct mottles; very few,

		medium to course roots; abrupt and wavy boundary to
5Bw	118-132	Light olive brown (2.5Y5/3, moist) gravely sandy loam (field); weak, medium sub-angular structure; very friable moist consistence; none sticky none plastic; dominant, fine to medium size, slightly weathered sub rounded pumice gravel; many very fine channel pores; very few, fine roots; abrupt and wavy boundary to
6Bw	132-152+	Olive brown (2.5Y4/3, moist) sandy loam (field); weak, fine to medium sub-angular blocky structure; very friable moist consistence; none sticky none plastic; common, fine to medium size, slightly weathered sub rounded pumice gravel; very few, coarse roots.
		Remarks: A2 had extrusive organic stains on the faces of the peds. Animal barrows filled with Ah material and roots of diameter between 5 to 40 cm.

Site and Profile Description

Profile ID:	10000
Classification:	FAO: Areni-Vitric Andosol (Eutric) USDA: Ashy, Glassy, Isothermic Aridic Ustpsamment (possibly ashy particle-size class and glassy mineralogy)
Author:	Simfukwe P.
Date:	19/09/00
Location:	Field 313 lower farm of Sulmac Longonot Farm UTM 206678, 9907232
Landform:	The physiographic position is on the middle of the tread with, nearly level lacustrian plain; slope 0.5%
Parent Material:	Gamlian lake sediments with possible admixture of volcanic ash
Landuse/Vegetation:	Ploughed for vegetable cultivation and currently under vegetable cultivation.
Drainage:	Well drained, the soil was moist during the time of survey.
Presence of rocks/stones	None
Effective Depth:	Very deep more than 150 cm deep

Horizon	**Depth(cm)**	Description
Ap	0-22	Dark olive brown (2.5Y3/3) moist colour and olive grey (5Y4/2) dry colour sand loam; weak fine to medium subangular blocky; loose when dry and very friable when moist; none plastic and none sticky; very few medium size pumice rock fragments, rounded and slightly weathered; few fine roots; clear and smooth boundary to
Bw	22-50	Very dark greyish brown (2.5Y3/2) moist colour sand loam. Weak, fine to medium subangular blocky in structure; loose when dry and very friable when moist; none plastic and none sticky; very few medium size pumice rock fragments, rounded and slightly weathered; few fine roots.
		Remarks: at 55cm there was a layer of black sands (volcanic glass), these contributed to the naming of the soil as *sandy, glassy,* The soils in this mapping unit had in some places many fine to medium rock fragments. And classified as *sandy-skeletal.* This pit description is representative of observation points 10000-11111

Site and Profile Description

Profile ID:	20000
Classification:	FAO: Areni-Vitric Andosol (Eutric) USDA: Ashy, Glassy, Isothermic Aridic Ustpsamment
Author:	Simfukwe P.
Date:	19/09/00
Location:	Field 423 lower farm of Sulmac Longonot Farm UTM 207234, 9907013
Landform:	The physiographic position is on the middle of the tread riser complex with, nearly level volcanic plain; slope1%
Parent Material:	Volcanic ash and pyroclastics with possible admix of lake sediments of the Gamlian Lake
Landuse/Vegetation:	Ploughed for vegetable cultivation and currently under vegetable cultivation.
Drainage:	Well drained, the soil was moist during the time of survey
Presence of rocks/stones	15-40% abundance, course gravel and sub rounded in shape both on the surface and in the horizons
Effective Depth:	Very deep more than 150 cm deep

Horizon	**Depth(cm)**	Description
Ap	0-20	Olive brown (2.5Y4/4) moist colour gravely sandy loam; weak fine to course sub angular blocky, friable when moist; slightly plastic and slightly sticky; common, fine to course pumice gravel, sub rounded and slightly weathered; common fine roots; clear and smooth boundary.
Bw1	20-40	Olive brown (2.5Y4/3) moist colour gravely sandy loam; weak fine to medium sub angular blocky, friable when moist; slightly plastic and slightly sticky; common, fine to medium pumice gravel, sub rounded and slightly weathered; few fine roots; clear and smooth boundary.
Bw2	40-50	Dark olive brown (2.5Y3/3) moist colour gravely sandy loam; weak fine to medium sub angular blocky, friable when moist; slightly plastic and slightly sticky; many, fine to medium pumice gravel, sub rounded and slightly weathered; very few fine roots; clear and smooth boundary.
Bw3	50-60	Dark greyish brown (2.5Y4/2) moist colour loamy sand; weak fine to medium sub angular blocky, friable when moist; none plastic and none sticky; many, fine to medium pumice gravel, sub rounded and slightly weathered; very few fine roots; clear and smooth boundary.
Bw4	60+	Olive brown (2.5Y4/4) moist colour loamy sand; weak fine to medium sub angular blocky, friable when moist; none plastic and none sticky; many, fine to medium pumice gravel, sub rounded and slightly weathered; very few fine roots; clear and smooth boundary.
		Remarks: *This* pit description is representative of observation points 20000-21111

Site and Profile Description

Profile ID:	30000
Classification:	**FAO:** Areni-Vitric Andosol (Eutric) **USDA:** Ashy, Glassy, Isothermic Aridic Ustpsamment
Author:	Simfukwe P.
Date:	21/09/00
Location:	Field 633 middle of the farm of Sulmac Longonot Farm UTM 207791, 9906788
Landform:	The physiographic position is on the lower position of the tread riser complex of the mid volcanic plain, gently andulating, slope 3%
Parent Material:	Volcanic ash and pyroclastics.
Landuse/Vegetation:	Formally used for the carnations, but at the time of survey it as fallow.
Drainage:	Well drained, the soil was dry during the time of survey
Presence of rocks/stones	No rock fragments observed on the surface.
Effective Depth:	Very deep more than 150 cm deep

Horizon	**Depth(cm)**	Description
O	0-2	**Small layer of organic matter, mainly grass and roots**
Ah	2-15	Dark greyish brown (2.5Y4/2) dry colour; loamy sand; very weak, very fine to fine, sub angular blocky, soft when dry consistence; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; common very fine to fine roots; clear and smooth boundary.
Bw1	15-45	Olive (5Y4/2) dry colour loamy sand; very weak very fine to fine, sub angular blocky, slightly hard when dry consistence; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; fine roots; clear and smooth boundary.
Bw2	45+	Olive (5Y4/2 moist colour; loamy sand; very weak, very fine to fine, and sub angular blocky in structure; very friable moist consistence; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; no roots; clear and smooth boundary.
		Remarks: At the time it was use to graze cows and zebras. This pit description is representative of observation points 30000-31111

Site and Profile Description

Profile ID:	40000
Classification:	FAO: Areni-Vitric Andosol (Eutric) USDA: Ashy, Glassy, Isothermic Aridic Ustpsamment
Author:	Simfukwe P.
Date:	21/09/00
Location:	Field 743 middle of the farm of Sulmac Longonot Farm UTM 208347, 9906563 on the mid volcanic plain
Landform:	The physiographic position is on the middle position of the tread riser complex of the mid volcanic plain, undulating, slope 6%
Parent Material:	Volcanic ash and pyroclastics.
Landuse/Vegetation:	Formally used for the carnations, but at the time of survey it as fallow.
Drainage:	Well drained, the soil was dry during the time of survey
Presence of rocks/stones	No rock fragments observed on the surface.
Effective Depth:	Very deep more than 150 cm deep

Horizon	**Depth(cm)**	Description
O	0-3	**Small layer of organic matter, mainly grass and roots**
Ah	3-18	Olive brown (2.5Y4/3) dry colour loamy sand; weak very fine to medium, sub angular blocky, soft consistence when dry; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; many very fine to fine roots; clear and smooth boundary.
Bw1	18-55	Light olive brown (2.5Y5/2) dry colour loamy sand; weak very fine to course, sub angular blocky structure; soft consistence when dry; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; many very fine to fine roots; clear and smooth boundary.
Bw2	55+	Light olive brown (2.5Y5/2) dry colour loamy sand; weak very fine to medium, sub angular blocky structure; soft consistence when dry; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; few very fine to fine roots; clear and smooth boundary.
		Remarks: Also used to graze cows and zebras. The unwhethered zone was observed at 85 cm. This pit description is representative of observation points 40000-41111.

Site and Profile Description

Profile ID:	50000
Classification:	FAO: Areni-Vitric Andosol (Eutric) USDA: Ashy, Glassy, Isothermic Aridic Ustpsament
Author:	Simfukwe P.
Date:	21/09/00
Location:	Field 851 middle of the farm of Sulmac Longonot Farm UTM 208903, 9906339 on the mid volcanic plain
Landform:	The physiographic position is on the tread riser complex of the mid volcanic plain, gently undulating, slope 4%
Parent Material:	Volcanic ash and pyroclastics.
Landuse/Vegetation:	Formally used for the carnations, but at the time of survey it was fallow. There were several bushes indicating that it had been fallow for along time
Drainage:	Well drained, the soil was dry during the time of survey
Presence of rocks/stones	No rock fragments observed on the surface.
Effective Depth:	Very deep more than 150 cm deep

Horizon	**Depth(cm)**	Description
Ah	0-15	Olive brown (2.5Y4/3) dry colour loamy sand; very weak very fine to fine, sub angular blocky structure; very soft consistence when dry; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; many very fine to fine roots; clear and smooth boundary.
Bw1	15-23	Light olive brown (2.5Y5/3) dry colour loamy sand; weak fine medium, sub angular blocky structure; slightly hard consistence when dry; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; few very fine to fine roots; clear and smooth boundary.
Bw2	23+	Olive (5Y5/3) dry colour loamy sand; weak very fine to medium, sub angular blocky structure; soft consistence when dry; none plastic and none sticky; very few, fine to medium pumice gravel, sub rounded and slightly weathered; few very fine to fine roots; clear and smooth boundary.
		Remarks: on the ridges, the soils were manly sandy and had the black sands. The land is used to graze cows and zebras. The grass cover was less than 10% but had bushes. This pit description is representative of observation points 50000-51111

Site and Profile Description

Profile ID:	60000
Classification:	FAO: Areni-Vitric Andosol (Eutric) USDA: Ashy, Glassy, Isothermic Aridic Ustpsament
Author:	Simfukwe P.
Date:	21/09/00
Location:	Field 1063 upper part of the Sulmac Longonot Farm UTM 209460, 9906114 on the mid volcanic plain
Landform:	The physiographic position is on the tread riser complex of the mid volcanic plain, gently undulating, slope 4%
Parent Material:	Volcanic ash and pyroclastics.
Landuse/Vegetation:	Formally used for the carnations, but at the time of survey it was fallow. There were few bushes indicating that it had been fallow for along time
Drainage:	Well drained, the soil was dry during the time of survey
Presence of rocks/stones	No rock fragments observed on the surface.
Effective Depth:	Very deep more than 150 cm deep

Horizon	**Depth(cm)**	Description
O	0-2	Small layer of organic matter, mainly of grass and roots.
Ah	2-14	Olive brown (2.5Y4/3) dry colour; loamy sand; weak very fine to medium, sub angular blocky structure; soft consistence when dry; none plastic and none sticky; very few (less than 2%), fine to medium pumice gravel, sub rounded and slightly weathered; common very fine to medium roots; clear and smooth boundary.
Bw1	14-75+	Light olive brown (2.5Y5/3) dry colour; loamy sand; weak very fine to medium, sub angular blocky structure; soft consistence when dry; none plastic and none sticky; very few (less than 2%), fine to medium pumice gravel, sub rounded and slightly weathered; common very fine to fine roots; clear and smooth boundary
		Remarks: on the ridges, the soils were manly *glassy, Isothermic, Typic Haplustands*. The land is used to graze cows and zebras. The grass cover was substantial about 40% cover but had bushes. This pit description is representative of observation points 60000-61111.

7.2.2 Appendix B: List of the GPS Points for the Nested sampling scheme

stage	Sample ID	E	N	Angle	distance
1	**10000**	206678	9907238	0	600
4	**10010**	206662	9907247	300	19
3	**10100**	206738	9907233	94	60
4	**10110**	206750	9907219	141	19
2	**11000**	206644	9907051	190	190
5	**11001**	206647	9907044	140	6
4	**11010**	206662	9907044	112	19
5	**11011**	206659	9907038	205	6
3	**11100**	206698	9907078	58	60
5	**11101**	206703	9907079	70	6
4	**11110**	206714	9907084	84	19
5	**11111**	206713	9907090	11	6
1	**20000**	207234	9907013	0	600
4	**20010**	207216	9907011	265	19
3	**20100**	207175	9907027	186	60
4	**20110**	207186	9907043	38	19
2	**21000**	207054	9906947	250	190
5	**21001**	207049	9906949	290	6
4	**21010**	207034	9906951	280	19
5	**21011**	207028	9906953	286	6
3	**21100**	207091	9906997	38	60
5	**21101**	207086	9907001	317	6
4	**21110**	207099	9907014	25	19
5	**21111**	207101	9907008	155	6
1	**30000**	207791	9906788	0	600
4	**30010**	207781	9906772	212	19
3	**30100**	207841	9906750	126	60
4	**30110**	207847	9906733	156	19
2	**31000**	207980	9906786	91	190
5	**31001**	207984	9906785	123	6
4	**31010**	207995	9906793	73	19
5	**31011**	207998	9906800	17	6
3	**31100**	207982	9906726	178	60
5	**31101**	207984	9906731	34	6
4	**31110**	207995	9906714	128	19
5	**31111**	207789	9906712	235	6

stage	Sample ID	E	N	Angle	distance
1	**40000**	208347	9906563	0	600
4	**40010**	208363	9906571	60	19
3	**40100**	208309	9906609	320	60
4	**40110**	208324	9906599	125	19
2	**41000**	208160	9906531	260	190
5	**41001**	208154	9906529	252	6
4	**41010**	208178	9906524	112	19
5	**41011**	208181	9906529	35	6
3	**41100**	208116	9906490	227	60
5	**41101**	208110	9906490	266	6
4	**41110**	208133	9906499	61	19
5	**41111**	208128	9906502	296	6
1	**50000**	208903	9906339	0	600
4	**50010**	208921	9906346	68	19
3	**50100**	208848	9906314	246	60
4	**50110**	208848	9906333	358	19
2	**51000**	209073	9906424	63	190
5	**51001**	209075	9906430	17	6
4	**51010**	209059	9906411	228	19
5	**51011**	209061	9906406	157	6
3	**51100**	209133	9906424	93	60
5	**51101**	209131	9906420	208	6
4	**51110**	209132	9906405	188	19
5	**51111**	209130	9906406	102	6
1	**60000**	209460	9906114	0	600
4	**60010**	209473	9906127	46	19
3	**60100**	209449	9906173	350	60
4	**60110**	209431	9906178	286	19
2	**61000**	209295	9906208	300	190
5	**61001**	209300	9906212	47	6
4	**61010**	209310	9906196	131	19
5	**61011**	209304	9906193	241	6
3	**61100**	209346	9906177	122	60
5	**61101**	209347	9906182	10	6
4	**61110**	209332	9906161	220	19
5	**61111**	209333	9906167	351	6

7.2.3 Appendix C: List of other observation points.

Number	Location ID	Eastings	Northings	Date	Obs_Type
1	1	203619	9905685	8-Sep-00	GeoRef Point
2	2	194919	9913641	9-Sep-00	GeoRef Point
3	3	213594	9914322	26-Sep-00	GeoRef Point
4	1011	207948	9904763	29-Sep-00	GeoRef Point
5	1043	209059	9905451	29-Sep-00	GeoRef Point
6	1053	209335	9905631	29-Sep-00	GeoRef Point
7	1073	209907	9905964	29-Sep-00	GeoRef Point
8	312	206468	9907336	28-Sep-00	GeoRef Point
9	313	206654	9907030	28-Sep-00	GeoRef Point
10	323	207106	9907309	28-Sep-00	GeoRef Point
11	333	207394	9907486	28-Sep-00	GeoRef Point
12	341	207192	9907794	29-Sep-00	GeoRef Point
13	353	207653	9908086	29-Sep-00	GeoRef Point
14	412	206852	9906714	28-Sep-00	GeoRef Point
15	433	207589	9907171	28-Sep-00	GeoRef Point
16	443B	207868	9907344	29-Sep-00	GeoRef Point
17	453B	208147	9907515	29-Sep-00	GeoRef Point
18	461	207942	9907831	29-Sep-00	GeoRef Point
19	471	208218	9908005	29-Sep-00	GeoRef Point
20	473	208499	9908180	29-Sep-00	GeoRef Point
21	512	207055	9906394	28-Sep-00	GeoRef Point
22	523	207517	9906684	28-Sep-00	GeoRef Point
23	553	208344	9907198	29-Sep-00	GeoRef Point
24	561	208337	9907202	29-Sep-00	GeoRef Point
25	573	208892	9907547	29-Sep-00	GeoRef Point
26	623	207712	9906359	28-Sep-00	GeoRef Point
27	653	208548	9906879	29-Sep-00	GeoRef Point
28	663	208824	9907054	29-Sep-00	GeoRef Point
29	673	209103	9907232	29-Sep-00	GeoRef Point
30	711	207362	9905713	28-Sep-00	GeoRef Point
31	723	207914	9906043	28-Sep-00	GeoRef Point
32	753	208745	9906556	29-Sep-00	GeoRef Point
33	773	209302	9906919	29-Sep-00	GeoRef Point
34	811	207550	9905392	29-Sep-00	GeoRef Point
35	853	208940	9906251	29-Sep-00	GeoRef Point
36	873	209498	9906599	29-Sep-00	GeoRef Point
37	953	209141	9905934	29-Sep-00	GeoRef Point
38	973	209697	9906289	29-Sep-00	GeoRef Point
39	A1	204980	9907965	8-Sep-00	GeoRef Point
40	CR1	194654	9914688	9-Sep-00	GeoRef Point
41	CR2	194921	9913640	9-Sep-00	GeoRef Point
42	HEL01	207078	9905978	15-Sep-00	GeoRef Point
43	HEL02	206335	9905253	15-Sep-00	GeoRef Point
44	HEL03	208022	9904006	15-Sep-00	GeoRef Point
45	HEL04	208308	9903997	15-Sep-00	GeoRef Point
46	HEL05	207278	9904582	15-Sep-00	GeoRef Point
47	HELBH1	208457	9903942	15-Sep-00	GeoRef Point
48	MA1	212020	9927630	9-Sep-00	GeoRef Point
49	MA2	211893	9927712	9-Sep-00	GeoRef Point
50	MO02	212107	9911439	11-Sep-00	GeoRef Point

51	MO03	207948	9904756	11-Sep-00	GeoRef Point
52	MO05	207783	9905463	11-Sep-00	GeoRef Point
53	MO06	207538	9905805	11-Sep-00	GeoRef Point
54	MO07	207512	9906680	11-Sep-00	GeoRef Point
55	MO08	206654	9907022	11-Sep-00	GeoRef Point
56	MO1	213320	9912695	11-Sep-00	GeoRef Point
57	MOB	385257	9889462	23-Sep-00	GeoRef Point
58	OS1	214180	9925766	9-Sep-00	GeoRef Point
59	**PIT01**	**207718**	**9906310**	**13-Sep-00**	**Ref pit**
60	**PIT02**	**204648**	**9908546**	**14-Sep-00**	**Ref pit**
61	R1	203620	9905683	8-Sep-00	GeoRef
62	**SWA**	**208684**	**9907714**	**29-Sep-00**	**Mini pit**
63	TA01	206250	9906495	28-Sep-00	GeoRef
64	TU01	204524	9927507	12-Sep-00	GeoRef
65	TU02	204314	9927122	12-Sep-00	GeoRef

The list of GPS points used for georeferencing the Aerial photos and for a few pits.

The Location of the GPS point dispayed on the gereferenced satellite image.

Figure 7.1The display of GPS point from table in appendix E (above)

7.2.4 Appendix D: The table for the Nested Analysis

obs_ID	M_unit	Stg No	stage 1	stage 2	stage 3	stage 4	stage 5	Depth _w	Limit_ a	Limit_ b	grave-lA	gra-velB	Cluster 6
10000	Pl-211	1	1	0	0	0	0	50	22	50	2.7	5.6	1
10010	Pl-211	4	1	0	0	1	0	40	25	40	6.9	13.8	2
10100	Pl-211	3	1	0	1	0	0	45	10	45	4.2	8.8	3
10110	Pl-211	4	1	0	1	1	0	55	20	45	2.7	2.1	1
11000	Pl-211	2	1	1	0	0	0	30	18	30	3.3	3.6	4
11001	Pl-211	5	1	1	0	0	1	55	22	46	3.0	3.7	1
11010	Pl-211	4	1	1	0	1	0	55	18	44	2.0	3.2	1
11011	Pl-211	5	1	1	0	1	1	55	15	40	3.4	2.7	3
11100	Pl-211	3	1	1	1	0	0	46	20	46	3.0	2.1	1
11101	Pl-211	5	1	1	1	0	1	46	10	46	3.2	4.3	3
11110	Pl-211	4	1	1	1	1	0	52	15	46	3.7	4.2	3
11111	Pl-211	5	1	1	1	1	1	50	10	30	4.2	4.5	4
20000	Pv-311	1	2	0	0	0	0	68	20	40	8.9	9.4	2
20010	Pv-311	4	2	0	0	1	0	68	20	35	8.3	18.0	2
20100	Pv-311	3	2	0	1	0	0	58	22	40	10.2	14.5	2
20110	Pv-311	4	2	0	1	1	0	50	22	50	19.3	11.4	5
21000	Pv-311	2	2	1	0	0	0	64	22	50	2.8	5.2	1
21001	Pv-311	5	2	1	0	0	1	55	20	55	3.0	2.8	1
21010	Pv-311	4	2	1	0	1	0	60	20	60	3.9	3.9	1
21011	Pv-311	5	2	1	0	1	1	60	20	60	3.3	2.5	1
21100	Pv-311	3	2	1	1	0	0	60	20	38	9.8	12.3	2
21101	Pv-311	5	2	1	1	0	1	65	16	50	6.9	5.9	3
21110	Pv-311	4	2	1	1	1	0	67	15	42	6.9	7.2	3
21111	Pv-311	5	2	1	1	1	1	67	15	55	5.0	4.5	3
30010	Pv-211	4	3	0	0	1	0	70	20	50	1.6	2.0	1
30100	Pv-211	3	3	0	1	0	0	65	22	55	0.1	0.1	1
30110	Pv-211	4	3	0	1	1	0	66	18	55	0.0	0.0	1
31000	Pv-211	2	3	1	0	0	0	75	13	45	1.2	1.4	1
31001	Pv-211	5	3	1	0	0	1	60	18	40	1.5	1.1	1
31010	Pv-211	4	3	1	0	1	0	60	15	45	0.7	1.3	1
31100	Pv-211	3	3	1	1	0	0	50	15	50	0.7	0.6	1
31101	Pv-211	5	3	1	1	0	1	65	22	55	0.4	0.4	1
40000	Pv-211	1	4	0	0	0	0	85	18	55	0.1	0.2	1
40010	Pv-211	4	4	0	0	1	0	80	20	55	1.1	1.6	1
40110	Pv-211	4	4	0	1	1	0	75	15	55	0.0	0.0	1
41000	Pv-211	2	4	1	0	0	0	45	12	45	0.0	0.0	1
41001	Pv-211	5	4	1	0	0	1	40	12	40	0.0	0.0	1
41100	Pv-211	3	4	1	1	0	0	48	16	48	0.0	0.0	1
41101	Pv-211	5	4	1	1	0	1	55	12	45	0.0	0.0	1
41110	Pv-211	4	4	1	1	1	0	40	15	40	0.0	0.0	1
41111	Pv-211	5	4	1	1	1	1	40	15	40	0.0	0.0	1
50000	Pv-211	1	5	0	0	0	0	50	15	23	0.9	0.6	4
50100	Pv-211	3	5	0	1	0	0	50	12	32	0.3	0.3	4
50110	Pv-211	4	5	0	1	1	0	65	15	30	0.0	0.0	4
51010	Pv-211	4	5	1	0	1	0	63	10	30	1.6	1.5	4
51011	Pv-211	5	5	1	0	1	1	58	10	30	1.1	0.9	4
51100	Pv-211	3	5	1	1	0	0	60	12	40	0.6	0.4	1
51110	Pv-211	4	5	1	1	1	0	45	20	45	0.0	0.0	1
51111	Pv-211	5	5	1	1	1	1	35	15	35	0.0	0.0	4
60000	Pv-211	1	6	0	0	0	0	75	14	75	1.5	1.2	6
60100	Pv-211	3	6	0	1	0	0	80	16	60	1.2	1.1	1
60110	Pv-211	4	6	0	1	1	0	85	16	45	1.5	0.9	1
61000	Pv-211	2	6	1	0	0	0	45	18	45	0.0	0.0	1
61001	Pv-211	5	6	1	0	0	1	44	15	44	0.0	0.0	1

61010 Pv-211	4	6	1	0	1	0	50	16	50	0.0	0.0	1
61011 Pv-211	5	6	1	0	1	1	50	15	50	0.0	0.0	1
61100 Pv-211	3	6	1	1	0	0	50	18	50	0.0	0.0	1
61101 Pv-211	5	6	1	1	0	1	50	18	50	0.0	0.0	1
61110 Pv-211	4	6	1	1	1	0	55	18	55	0.0	0.0	1
61111 Pv-211	5	6	1	1	1	1	48	12	48	0.0	0.0	1