Ecology and Development Series No. 17, 2004

Wilson Agyei Agyare

Soil characterization and modeling of spatial distribution of saturated hydraulic conductivity at two sites in the Volta Basin of Ghana

Cuvillier Verlag Göttingen

To my mother (Grace Asiedu) for educating me, and my wife (Juliana Agyare) for standing by me.

ABSTRACT

Soil data serve as an important input parameter for hydro-ecological and climatological modeling of water and chemical movement, heat transfer or land use change. Soil properties, most especially the hydraulic properties are highly variable spatially and measuring them is time-consuming and expensive. For that matter efficient methods, for estimating soil hydraulic properties are important. The purpose of this study is to characterize the spatial variation of soil physical properties, identify suitable models and important parameters for estimating saturated hydraulic conductivity (K_s).

The study was carried out at two locations in the Volta Basin of Ghana, near Tamale (9°28'N and 0°55'W) and Ejura (7°19'N and 1°16'W) sites. Data was collected from an area of 6-km^2 and 0.64-km^2 at Tamale and Ejura pilot sites, respectively. Data collected include soil diagnostic horizon, texture, color, mottles, structure, roots, gravel concretion fraction, particle size distribution, pH, organic carbon, cation exchange capacity (CEC), bulk density and K_s at the topsoil (0-15 cm) and subsoil (30-45 cm). Semivariogram analysis and kriging interpolation were used to develop digital elevation model (DEM) and eight terrain attributes at 30 ×30 m^2 grid size.

Stepwise multiple regression (SMR) and generalized linear model (GLM) were used to evaluate different independent variables for estimating K_s. Statistical evaluation procedures used include: coefficient of determination (R^2), normalize mean square error (NMSE), ANOVA, non-parametric median test, geometric mean error ratio (GMER) and geometric standard deviation of error ratio (GSDER). Different pedo-transfer functions (PTFs) were evaluated and compared. Also, artificial neural network (ANN) was used to model K_s using varying data sets and the results compared.

Saturated hydraulic conductivity is highly variable with coefficient of variation more than 100 %. The soils at Ejura have comparatively high sand content (> 69 %) and high clay content, which does not change much from 23 % (topsoil) to 21 % (subsoil), compared to 7 % (topsoil) to 23 % (subsoil) for the Tamale soil. A higher spatial dependency (range) was observed for most parameters in the subsoil compared to the topsoil at both sites.

The two sites have about the same mean elevation (169 m), with the Tamale site having a higher range and covering a larger area compared to the Ejura site. The Tamale pilot site is virtually flat with a mean slope gradient varying from 0.0° - 3.1° compared to 0.0-10.7° at Ejura site. The terrain parameters had poor relationship with K_s, which resulted in poor performance in using terrain parameters for estimating K_s.

Different soil types were mapped by digitizing areas of uniform soil morphological properties into eight soil types (namely: Haplic Luvisol, Lithic Leptosol, Ferric Acrisol, Plinthic Acrisol, Dystric Plinthosol, Eutric Plinthosol, Eutric Gleysol, and Dystric Gleysol)) at the Tamale site and five soil types (namely: Ferralic Cambisol, Ferric Acrisol, Haplic Acrisol, Gleyic Acrisol, and Gleyic Fluvisol) at the Ejura site. Non-parametric median test indicated differences in sand, silt, and clay content and pH for the different soil types at both sites at the two soil depths mainly as a result of differences in soil translocation and leaching at different landscape positions. Relationship between soil type and land use type was observed as specific crops dominate on certain soil types (such as, rice cultivation on Eutric and Dystric Gleysol).

In using SMR and GLM it was observed that the most important data for K_s modeling are site, soil depth, particle size distribution (sand, silt and clay content) and

bulk density. Terrain attributes, soil type and land use type parameters may be used to improve on model performance but can not be relied upon as the basis for modeling K_s.

Comparing eleven existing PTFs for estimating K_s the models of Campbell, Brutsaert, Ajuja and Rawls outperformed the remaining ones, thus indicating their wider domain of applicability. The models of Campbell (R^2=0.38) and Brutsaert (R^2=0.35) were outstanding in terms of correlating and deviations from the measured K_s based on the GMER.

With adequate sensitive data ANN can be used to estimate K_s using soil physical properties with improved estimation when terrain attributes are included. In ANN it was shown that the use of terrain parameters alone can not yield appreciable estimation of K_s. The topsoil K_s was found to be significantly influenced by source of training data but the subsoil is not affected by training data source.

In general, it was found that soil physical properties vary spatially and that through soil mapping it is possible to put soils in groups of uniform texture (sand, silt and clay content) and pH as these properties vary across the catena for the topsoil and subsoil. Amongst the evaluated PTFs the models by Campbell and Brutseart were more suitable for estimating K_s at the two sites. The ANN method can be used to model K_s with improved results compared to PTFs. The soil parameters; sand, silt, clay content, and bulk density were found to be the most important for modeling K_s.

Bodencharakterisierung und Modellierung der räumlichen Verteilung der gesättigten hydraulischen Leitfähigkeit an zwei Standorten im Voltabecken in Ghana

KURZFASSUNG

Bodendaten sind wichtige Parameter für die hydro-ökologische und klimatologische Modellierung von Wasser und Chemikalien, Wärmetransfer oder Landnutzungsveränderungen. Bodeneigenschaften, insbesondere die hydraulischen Eigenschaften, sind stark variable und ihre Messung ist zeitaufwändig und teuer. Daher sind effiziente Methoden zur Bestimmung bodenhydraulischer Eigenschaften wichtig. Das Ziel dieser Studie ist die Charakterisierung der räumlichen Variabilität der bodenphysikalischen Eigenschaften, die Identifizierung geeigneter Pedotransferfunktionen (PTF) und die Entwicklung eines Modells für die Bestimmung der gesättigten hydraulischen Leitfähigkeit (K_s).

Die Studie wurde an zwei Standorten im Voltabecken in Ghana in Tamale (9°28'N und 0°55'W) und Ejura (7°19'N und 1°16'W) durchgeführt. Die Untersuchungsgebiete umfassen eine Fläche von 6-km^2 in Tamale bzw. 0.64-km^2 in Ejura. Folgende Bodenparameter wurden untersucht: bodendiagnostischer Horizont, Textur, Farbe, Marmorierung, Struktur, vorhandene Wurzeln, Kies-/Schotterkonkretion, Korngrößenverteilung, pH-Wert, organischer Kohlenstoff, Kationenaustauschkapazität (KAK), Lagerungsdichte und hydraulische Leitfähigkeit (K_s). Eine Semivariogrammanalyse bzw. Kriginginterpolation diente zur Erstellung eines digitalen Höhenmodells (DEM) mit acht Geländeattributen und einer Rastergröße von 30 ×30 m$^{2.}$

Verschiedene Pedotransferfunktionen (PTF) wurden bewertet und anschließend miteinander verglichen. Außerdem wurde ein künstlich-neuronales Netzwerk (ANN) zur Modellierung von K_s mit verschiedenen Datengruppen eingesetzt und die Ergebnisse verglichen. Die schrittweise multivariante Regression (SMR), das generalisierte lineare Modell (GLM) und künstliche neuronale Netzwerk (ANN) wurden zur Modellierung der gesättigten hydraulischen Leitfähigkeit verwendet.. Die eingesetzten statistischen Bewertungsverfahren umfassen: Bestimmungsmaß (R^2), normalisierter mittlerer Quadratfehler (NMSE), ANOVA, nicht-parametrischer Mediantest, GMER und GSDER.

Die gesättigte hydraulische Leitfähigkeit ist mit einem Variationskoeffizient von über 100 % stark variable. Die Böden in Ejura besitzen einen relativ hohen Sandgehalt (> 69 %) und hohen Tongehalt, der sich innerhalb des Bodenprofils von 23% im Oberboden bis 21% im Unterboden- kaum verändert, verglichen mit 7% im Oberboden und 23% im Unterboden für Böden in Tamale. Dagegen wurde an beiden Standorten eine höhere räumliche Abhängigkeit zwischen Unter- und Oberböden bei anderen Bodenparamern beobachtet.

Beide Standorte liegen ungefähr auf gleicher Höhe (169 m), wobei in Tamale die Höhenunterschiede über eine größere Fläche ausgeprägter sind als in Ejura. In Tamale ist das Gelände fast flach mit einem mittleren Hanggrandient zwischen 0.0° und 3.1°, verglichen mit 0.0-10.7° in Ejura. Die Geländeparameter weisen nur eine schwache Beziehung zu K_s auf; dies bedeutet, dass die Geländeparameter nur bedingt zur Bestimmung von K_s.genutzt werden können.

Anhand der Digitalisierung von Gebieten mit gleichen morphologischen Eigenschaften werden acht verschiedene Bodentypen in Tamale (Haplic Luvisol, Lithic Leptosol, Ferric Acrisol, Plinthic Acrisol, Dystric Plinthosol, Eutric Plinthosol, Eutric Gleysol und Dystric Gleysol) und fünf Bodentypen in Ejura (Ferralic Cambisol, Ferric Acrisol, Haplic Acrisol, Gleyic Acrisol, und Gleyic Fluvisol) klassifiziert. Der nicht-parametrische Mediantest deutet auf Unterschiede in Sand- und Schluffgehalt, pH-Wert und Lagerungsdichte für die Bodentypen beider Standorte und beider Bodentiefen hin und kann hauptsächlich als Ergebnis einer unterschiedlichen Bodenverlagerung und Versickerung an verschiedenen Stellen im Gelände gesehen werden. Eine Beziehung zwischen Bodentyp und Landnutzungstyp konnte durch die Dominanz bestimmter Anbaupflanzen auf bestimmten Bodentypen nachgewiesen werden (z.B. Reis auf Eutric und Dystric Gleysol).

Ein Vergleich zwischen elf PTFs zur Bestimmung von K_s zeigt, dass die Modelle von Campbell, Brutsaert, Ajuja und Rawls besser geeignet sind als die anderen Modelle, was auch durch ihre verbreitete Anwendung zum Ausdruck kommt. Die Modelle von Campbell (R^2=0.38) und Brutsaert (R^2=0.35) sind hervorragend hinsichtlich der Korrelierung und weisen die geringsten Abweichungen von der gemessenen K_s auf der Grundlage von GMER auf.

Mit ausreichend sensitiven Daten kann ANN zur Vorhersage von K_s mit bodenphysikalischen Eigenschaften genutzt werden, wobei eine Einbeziehung der Geländeeigenschaften die Ergebnisse verbessert. Durch ANN wird gezeigt, dass der Einsatz von Geländeparametern allein nicht zu einer akzeptablen Vorhersage der K_s führen kann. Während die K_s im Oberboden significant durch die Trainingsdaten beeinflusst wird, trifft dies nicht auf die K_s im Unterbodens zu.

Durch den Einsatz von SMR und GLM wird deutlich, dass Standort, Bodentiefe, Korngrößenverteilung (Sand-, Schluff-, Lehmgehalt) die wichtigsten Daten für die Modellierung von K_s und Lagerungsdichte darstellen. Zwar können Parameter wie Geländeeigenschaft, Bodentyp und Landnutzungstyp zur Verbesserung der Modelleistung eingesetzt werden, sollten aber nicht als Basis für die K_s-Modellierung genutzt werden.

Im Allgemeinen kann festgestellt werden, dass die bodenphysikalischen Eigenschaften räumlich stark variieren und, dass es durch Bodenkartierung möglich ist, Böden hinsichtlich einheitlicher Struktur (Sand-, Schluff-, Lehmgehalt) und pH-Wert zu gruppieren, da diese Eigenschaften entlang der Catena im Ober- bzw. Unterboden variieren. Unter den bewerteten PTFs waren für beide Standorte die Modelle von Campbell und Brutseart am besten für die Vorhersage von K_s geeignet. Die ANN-Methode kann zur Modellierung von K_s eingesetzt werden; dabei sind die Ergebnisse besser als die Ergebnisse mit PTF. Die für die Modellierung von K_s wichtigsten Bodenparameter sind Sand-, Schluff- und Lehmgehalt sowie Lagerungsdichte.

TABLE OF CONTENTS

ABBREVIATIONS

ANN	Artificial Neural Network
ANOVA	Analysis of variance
AS	Aspect
BD	Bulk density
BMBF	German Federal Ministry of Education and Research
BP	Backpropagation
BS	Base Saturation
CEC	Cation exchange capacity
Cl	Clay
CSIR	Council for Scientific and Industrial Research
CSS	Course structural size
CV	Coefficient of variation
Cv	Curvature
DEM	Digital Elevation Model
DGPS	Differential Global Positioning System
DSS	Decision Support System
ELEV	Elevation
GC	Gravel/concretion
GCM	Global Circulation Models
GLM	Generalized Linear Model
GLOWA	Globaler Wandel des Wasserkreislaufes (Global Change in Hydrologic Cycle)
GPS	Global Positioning System
LS	Length-Slope factor
MLP	Multi-Layer Perceptron
MR	Multiple Regression
MSE	Mean Square Error
MSSG	Moderately strong structure
NMSE	Normalized Mean Square Error
OC	Organic Carbon

PFC	Profile curvature
PNC	Plan curvature
PSD	Particle Size Distribution
PTF	Pedo-Transfer Function
r	Coefficient of correlation
R^2	Coefficient of determination
RMSE	Root mean square error
SARI	Savannah Agricultural Research Institute
Sd	Sand
Si	Silt
SL	Slope
SMR	Stepwise Multiple Regression
SPI	Stream power index
SRI	Soil Research Institute
SS	Subsoil
SSG	Strong structure
TIN	Triangulated Irregular Network
TS	Topsoil
UA	Upslope area
UCC	University of Cape Coast
WGS	World Geodetic System
WI	Wetness index
ZEF	Zentrum für Entwicklungsforschung

1 INTRODUCTION

Soil is susceptible to misuse and mismanagement that often results in its degradation or loss of soil quality, i.e., reduction in the soil's ability to perform its ecosystem function and food productivity (Lal et al., 2003). There are a number of physical, chemical and biological causes of soil degradation (Lal et al., 2003), but the most important ones considered in the Volta Basin of Ghana – our area of interest – are deforestation, bush burning, removal and/or burning of crop residue, and mining activities, that are spurred by socio-economic and political issues such as population density, land tenure and policies. The small-scale system of farming practices in Ghana leads to soil nutrient mining with little or no use of fertilizer as illustrated for Ghana (Rhodes, 1995) and for sub-Saharan Africa (Vlek 1993 and Vlek et al., 1997).

The soil layer is the source and sink of heat and moisture to and from the atmosphere, thus underscoring the importance of land surface processes and therefore soil in climate modeling. Soil properties are the most important determinants in land surface processes as they influence the soil's potential ability to receive and/or store heat and moisture to and from the atmosphere.

The other most important natural resource is water, which is mainly used for consumption, agriculture and hydropower generation in the Volta Basin. Ghana among other West African countries, is estimated to experience water scarcity by 2025 due to climate change resulting in reduced rainfall, increased evaporation and advancing rates of desertification. Combined with the existing high rate of deforestation and degradation of vegetative cover this may have a serious effect on soil and water resources (UNEP, 2003). Water is an essential commodity used to generate cheap hydropower to fuel Ghana's industrial growth initiated in the early 1960's. In the early 1980s, drought affected the water level in the Akosombo hydropower dam. In addition, an increased demand for domestic electric power in Ghana and irrigation water in Burkina Faso has given rise to concern about the future of hydropower in Ghana (van de Giesen et al., 2001). Water (in)security as it relates to availability and usage, particularly in the dry season when some water sources (e.g., streams and wells) dry up, is a widespread problem in Ghana. Most people depend on such water for consumption and on rainwater for agriculture.

Soil degradation affects water quality through the transport of suspended and dissolved loads in surface water and agricultural chemicals into ground water (Lal et al., 2003). Soil degradation also affects climate change by its effect on greenhouse gases, rapid mineralization of organic carbon, increase in emission of N_2O and decreases in biomass productivity, thereby affecting the quality and quantity of biomass returned to the soil. These highlight the importance of sustainable use of land and water resources in the Volta Basin of Ghana and Burkina Faso. This study looks at soil characteristics and their variation as they are of great importance to soil behaviour, land degradation and climate with their effect on human life.

This study was carried out in the Volta Basin of Ghana. The Volta River drains about three-quarters of Ghana with a network of sources,– Black Volta, White Volta and the Oti Rivers – mainly from Burkina Faso, that flow into the Atlantic Ocean. The Volta Basin of Ghana has a population of about 7 million (GSS, 2002a). Crop production is mainly done under rain-fed conditions, with water being the most important limiting factor both in amount and distribution (Ofori-Sarpong, 1985), followed by inherent soil fertility. The fraction of precipitation that is available for *in situ* evapotranspiration is of great importance as it is the primary determinant for crop yield. In the Volta Basin as a whole, the precipitation that does not evapotranspire feeds the rivers, either directly through surface runoff or by recharging the ground water. The partitioning of water into the fraction that evapotranspires, runs off or deep percolates depends to a greater extent on the soil physical properties, which go a long way in influencing the land use pattern.

This study was carried out to characterize the spatial distribution of soil physical properties and to examine the variation in soil types and land use types in terms of soil properties at the topo-scale level. Furthermore, to identify suitable procedures and key parameters relevant for estimating saturated hydraulic conductivity. The following sections in this chapter review the importance and sources of soil data – particularly soil hydraulic data – and the state-of-the-art procedures for estimating saturated hydraulic conductivity.

1.1 Importance of soil data

Soil data are important for sound natural resource management (McKenzie et al., 2000). The prediction or modelling of runoff and land use change or validation of soil vegetation atmosphere transfer (SVAT) models, depends heavily on accurate data on soil physical properties and the understanding of these data. Soil properties such as texture, organic carbon, structure, aggregate size and stability influence soil erodibility, soil water storage, infiltration, particle detachability, water and sediment transport and chemical interaction.

As indicated by Ellison (1947); Dangler et al. (1975) and Elliot et al. (1988) these parameters are important in erosion models (such as Revised Universal Soil Loss equation (RUSLE) or Water Erosion Prediction Project (WEPP) model). Soil hydraulic properties such as saturated hydraulic conductivity, field capacity, drainable porosity, and water retention parameters are important input data for runoff estimation (Grayson et al., 1992).

In order to adequately describe the interaction of land surface and atmospheric boundary layer, one must effectively describe heat and moisture movement at the surface and within the soil. With the increasing spatial resolution of meso-scale models there is renewed interest in land surface models (LSM) with increasing need for information on spatial soil characteristic (Chen and Dudhia, 2001). The use and importance of soil parameters such as texture, porosity, matric potential, saturated hydraulic conductivity, slope of the retention curve, field capacity and wilting point in LSM were outlined by Chen and Dudhia (2001) and Wilson et al. (1987), and especially for bare soil by Ek and Cuenca (1994).

Concerns about the quality of soil and water resources have motivated the development of empirical and simulation models for evaluating the movement of water and chemicals into and through the soil media (Kohler et al., 2003 and Gaur et al., 2003). Furthermore, concerns about climate change and its impact on human life have resulted in a number of climatic models (global circulation models (GCM)), which all require spatial soil data to initialize. The use of such models is limited, because they need detailed data on soil physical properties. The key question is how to obtain soil physical parameters; in particular, the hydraulic properties that are not only difficult to measure but also highly spatially variable.

1.2 Source and spatial variability of soil parameters

In the past, soil surveying has played a major role in the development of pedology (Simonson, 1991) and soil maps have contributed immensely to natural resource management (Moore et al., 1993a). However, standard soil surveys, by design, do not provide detailed (high resolution) soil information for environmental modeling (Moore et al., 1993a). For instance, the existing soil map of the Ghanaian part of the Volta Basin is at a very low scale of 1:250,000 and provides very little information on soil physical properties. Even regional soil survey reports (such as Soils of Afram basin (Adu and Mensah-Ansah, 1995); and Soils of Bole-Bamboi area (Adu, 1995)) do not include soil hydraulic functions with the details required for comprehensive hydro-ecological modeling at watershed levels. Survey maps usually show a high variation in soil properties, most especially for soil hydraulic properties, as these are only measured at few selected points in soil survey studies. According to Burrough (1986), the major limitations associated with conventional soil survey maps are due to their limited coverage, uncertainties or errors as a result of locating class boundaries, non-uniformity of soil attributes, and insufficiency in information as it relates to details on soil properties at a given location. The soil scientist is normally aware of these constraints from his knowledge on soil-landscape relation.

One of the most important soil hydraulic properties is saturated hydraulic conductivity, which gives an indication of a soil's ability to transmit water (Klute and Dirksen, 1986). It is a function of particle size distribution, pore size distribution, continuity and configuration, bulk density and chemical properties such as organic carbon content and soil reaction (Hillel, 1998). Saturated hydraulic conductivity together with other soil hydraulic properties, are very important soil parameters used for determining infiltration, irrigation practice, drainage design, runoff, erosion, groundwater recharge, and leaching of soil nutrients (Rawls et al., 1992 and Vereecken et al., 1990). They can, to some degree of accuracy, be inferred from the state of other more easily measurable entities and knowledge of their relationship (Bouma, 1989).

The importance of soil properties stems from the important role they play in ecological modeling. The first step towards modeling is the collection of input data, which will be used to set the initial conditions for the model. These soil properties, most especially the hydraulic properties, are highly spatially variable (Wilding, 1984;

Wilding and Drees, 1983 and Warrick and Nielson, 1980) and measuring them is time-consuming and expensive (Schaap et al., 1999). In the past, much attention has been given to parametization of hydraulic properties; the spatial distribution of these properties has however rarely been considered due the difficulty in measuring them in the field. However, knowledge of the spatial variability (heterogeneity) of hydraulic properties is important in the quantification of flow and transport processes in soil at field or regional scales. In their investigation of soil hydraulic properties (Zhu and Mohanty, 2002), found the saturated hydraulic conductivity (K_s) to be the most variable compare to the "van Genuchten parameters". Saturated hydraulic conductivity significantly influences water flux in terms of infiltration and evaporation as their patterns follow that of K_s. Therefore, this study focuses on K_s.

1.3 Estimating soil parameters using environmental correlation

Soil is the result of interaction among soil forming factors (climate, relief, organisms, parent material, and time) (Jenny, 1941 and Jenny, 1980). Its spatial variability is therefore considered to be the causative realization of the complex combinations of soil-forming processes as influenced by the soil-forming factors. Until recently, most soil scientist emphasized the vertical relationships of soil horizons and soil-forming processes rather than horizontal relationships that characterize traditional soil survey (Buol et al., 1989). Characterizing spatial variability of soil parameters must link patterns to processes. Quantitative interpolation techniques (e.g., kriging) often ignore pedogenesis, while methods based on landscape position lack a consistent quantitative framework. Soil properties such as organic matter, A- and B-horizon thickness and degree of development, soil mottling, pH, depth of carbonates and soil water storage have all in the past been correlated to landscape position (Kreznor et al., 1989) using qualitative mapping units that delineate head slopes, linear slopes and foot slopes.

The use of environmental correlation to determine soil properties offers a suitable alternative to measuring soil parameters, more so with the advent of high accuracy terrain mapping systems such as the Global Positioning System (GPS). Past work has confirmed the existence of relationships between topographic attributes, such as elevation, slope, aspect, specific catchment area, and plan and profile curvature on the one hand and hydrological and erosion processes on the other (Speight 1974 and

Moore et al., 1991). Odeh et al. (1991) and Moore et al. (1993b) found that slope, plan and profile curvature, upslope distance and area accounted for much of the soil variation. Environmental correlation takes into account the spatial variation of the soil, which is essential for ecological and environmental modeling of the landscape.

Correlations between terrain attributes (such as slope, wetness index, sediment transport capacity index) and soil attributes (e.g., A-horizon, organic matter, silt and sand content) support the hypothesis that the soil catena develops in response to the way water flows through and over the landscape. The surface soil properties are mostly modified by land management while lower horizons may show greater response to topographic attributes (Moore et al., 1993a). Many previous investigations have already proved that there is a strong correlation between soil variability and upslope area calculated from digital elevation models (DEMs), because the landform configuration frequently governs the movement of materials and water on the landscape (Moore et al., 1993a; Gessler et al., 1995; Park and Vlek, 2002).

Basically, there are four main soil landscape or environmental correlation approaches that have been used in the past to characterize and estimate the spatial distribution of soils using readily available terrain or environmental attributes. These are the statistical correlation, geostatistical, semi-deterministic and the rule-based approaches, which may be used complementarily.

The statistical correlation approach is based on functional correlation of statistical analysis (regression and multivariate ordination) between soil attributes and one or several selected terrain and environmental attributes that are fairly easy to measure and are also physical meaning (Gessler et al., 1995).

The geostatistical approach uses the theory of regionalized variables (Matheron, 1971), which considers spatial variability of a soil property as a realization of a random function represented by a stochastic model (McBratney et al., 2000). Major limitations of the univariate geostatistical technique of kriging are due to the assumptions of stationarity, which is not often met by the field-sampled data sets, and the large data size requirement needed to define the spatial autocorrelation. Geostatistical procedures differ from classical procedures (statistical methods) in that the locations of the measurements are taken into account through the spatial coordinates.

The semi-deterministic approach utilizes landscape position to categorize soil properties based on the proposition that determination of soil distribution is done most efficiently by separation of pedogeomorphological units where similar hydrological, geomorphological and pedological processes occur (Conacher and Dalrymple, 1977; Kreznor et al., 1989 and Park et al, 2001).

The rule-based approach is analogous to the conventional soil survey in that it can use a wide range of evidence data based on prior knowledge and data availability. The conventional method, which has unspecified uncertainty introduced when knowledge is applied, conveys very little knowledge about the variation of the individual soil properties or the quantitative nature of the variation. However, the rule-based system builds on the surveyor's ability to construct quantitative statements about the individual soil properties through the development of a network of rules (Cook et al., 1996 and Zhu et al., 1997).

As an example of the soil landscape approach, Gessler et al. (1995) developed statistical models (multiple and logistic regressions) between terrain attributes (plan curvature, wetness index and upslope area) and soil attributes with R^2 of 63 % and 68 %, respectively, for A horizon and solum depth, respectively. Young and Hammer (2000) used cluster analysis to identify pedological and geologically distinct groups of soil thus revealing patterns of soil homogeneity and relationships among soil properties and landforms. Park et al. (2001), in using the quantitative approach delineated a 1.3 × 0.68 km area into soil landscape units that had a very good agreement with thickness of A horizon (R^2 = 52 %) and thickness of loess (R^2 = 63 %). For further reading see McBratney et al. (2000).

1.4 Estimation of saturated hydraulic conductivity

As a result of the high variability associated with soil hydraulic properties (Wilding, 1984 and Warrick and Nielson 1980), most work carried out in the past has been limited to the use of empirical and physical relationships and recently the use of artificial neural network. Many studies explored the possibility of estimating soil hydraulic functions from data available from soil surveys. A common approach is the use of pedotransfer functions (PTFs), which estimate the hydraulic properties through correlation with

comparatively easy to measure or widely available soil parameters (Bouma and van Lenen, 1987; Bouma, 1989 and Rawls et al., 1992).

The vast majority of PTFs, however, are empirically based on linear regression equations (Rawls, 1992), while others are physically based (Campbell, 1985 and Brutsaert, 1967). Although PTFs use at least some information about the particle-size distribution, considerable differences exist amongst PTFs in terms of the required input data. Varieties of PTFs with different mathematical concepts, estimation properties and input data requirements have been developed in the past. Williams et al. (1992) and Schaap et al. (1998) used hierarchical approaches to estimate saturated hydraulic conductivity, which are useful since they permit more flexibility toward the required input data when estimating hydraulic properties. Improvement in accuracy can be obtained with additional data. Developing PTFs utilizing soil texture, organic matter, soil structure, and bulk density as the common surrogates to estimate hydraulic properties are appropriate, but have some limitations such as large data requirement and site specificity and thus require local calibration.

Neural network models are a special class of PTFs, which use feed-forward back propagation or radial basis functions to approximate continuous (non-linear) functions. They have been used to estimate soil hydraulic properties (Schaap et al. 1999; Schaap and Bouten, 1996 and Pachepsky et al., 1996). An advantage of neural networks compared with traditional PTFs, is that neural networks require no *a priori* model concept. The optimal relation that links input data (basic soil properties) to output data (hydraulic parameters) is obtained and implemented in an iterative calibration procedure. However, the neural network has some disadvantages, such as the large volume of data required for training, the inability to extrapolate and the difficulty to implement compared to the traditional regression models (Schaap et al., 1998).

The use of terrain attributes for modeling K_s may serve as a suitable alternative, as terrain data are fairly easy to collect compared to intensive soil sampling. Terrain plays a fundamental role in modulating the earth surface and atmospheric processes. Thus, an understanding of the nature of terrain can directly lead to the understanding of the nature of these processes, in both subjective and analytical terms. DEM generation methods and a steadily increasing range of techniques for DEM interpretation and visualisation support these interactions. DEM data has numerous applications, most of

which are dependent on surface roughness and shape, with the exception of surface temperature and rainfall representation that are directly dependent on elevation (Hutchinson and Gallant, 2000).

Since Ruhe and his colleagues (e.g., Ruhe and Walker, 1968) first attempted to establish a functional correlation between certain soil properties and selected topographical parameters on loess-covered hillslopes in Iowa, many similar studies have followed. This approach has become the backbone for modern soil-landscape analysis (McBratney et al., 2000; Park and Vlek, 2002). In a soil-landscape analysis framework, the upslope area and its derivates (e.g., specific catchment area, wetness index, and stream power index) are the most widely used terrain parameters (Park and Vlek, 2002). Previous investigations have proved that there is a strong correlation between soil variability and upslope area calculated from DEMs, because the landform configuration frequently governs the movement of materials and water on the landscape (Burt and Butcher, 1986; Moore et al., 1993a; Gessler et al., 1995; Western and Blöschl, 1999; Park and Vlek, 2002).

The fundamental role of flowing water in controlling or explaining many environmental processes has resulted in the development of many flow routing and contributing area algorithms with varying limitations. Reflecting its importance, many different algorithms for calculating the upslope area are reported in current literature (O´Callaghan and Mark, 1984; Bauer et al., 1985; Fairfield & Leymarie 1991; Freeman 1991; Quinn et al., 1991, 1995; Costa-Cabral and Burges, 1994; Tarboton, 1997; Wilson et al., 2000). These algorithms may be classified as single (such as the D8 (deterministic eight-node) and Rho8 (random eight-node)) or multiple flow algorithms (such as the MFD (multiple flow direction) and DEMON method). The multiple flow algorithms allow flow into multiple cells (flow divergence) whiles the single flow algorithms are limited in that perspective.

Studies have been carried out in the past on the effect of data source, data structure and cell size on the terrain attribute in various applications such as in agricultural non-point source pollution model (Panuska et al., 1991), surface runoff models (Vieux, 1993), and in watershed model predictions (Zhang and Montgomery, 1994). Topographic attributes have been used to improve our understanding of hydrological, geomorphological, and ecological systems.

Most available empirical assessments of flow algorithms are limited to hydrological responses in a specific hydrological modeling framework (Wolock and McCabe, 1995; Wilson et al., 2000. Desmet and Govers (1996) present a notable exception by comparing different algorithms to predict the spatial occurrence of gully positions in Belgium. Park et al. (2003) compared different flow routing algorithms for estimating the spatial distribution of selected soil properties (soil moisture, soil pH, clay content).

In the numerous soil-landscape studies carried out in the past, the relationship between terrain attributes and K_s has not been addressed. The important question is: can K_s be estimated using any or a combination of these soil landscape approaches? Alternatively, will the inclusion of terrain attributes in estimating K_s help improve model performance?

1.5 Objective

This study is important in that it seeks to provide a soil data base for hydro-ecological and climatological modeling in the Volta Basin in Ghana through field work carried at two important locations in the basin. It is also to provide information on the spatial variation of soil properties and important parameters for estimating saturated hydraulic conductivity that can be used for assessing and managing land and water resources in the Volta Basin.

The general objective of this study is to characterize the spatial variation of soil physical properties, and to identify a suitable model and important parameters for estimating saturated hydraulic conductivity based on soil, land management and terrain data from two pilot sites in the Volta Basin of Ghana.

The specific objectives are as outlined below:

- To characterize the spatial distribution of soil physical parameters at the two pilot sites in the Volta Basin
- To quantify the variation in soil properties for different soil types and their relation to land use types at the two pilot sites
- To identify soil, land management and terrain parameters that are important for estimating saturated hydraulic conductivity using Stepwise Multiple Regression (SMR) or a Generalize Linear Model (GLM).

- To evaluate, compare and quantify the validity of existing pedotransfer functions for estimating saturated hydraulic conductivity against measured data
- To evaluate the potential of Artificial Neural Network (ANN) models for estimating saturated hydraulic conductivity

2 STUDY AREA

The study was carried out at two locations in the Volta Basin of Ghana at 9°28'N and 0°55'W (Tamale site), and 7°19'N and 1°16'W (Ejura site) (Figure 2.1).

Ghana is a West African state that is bordered by the Ivory Coast in the west, Togo in the East, and Burkina Faso to the north and the Atlantic Ocean in the south. The Volta Basin in Ghana has an area of 165,712 km² with a population of about 38 % of the total 18.4 million inhabitants of the country with a growth rate of 2.5 % (2000 census). The population in the basin has been steadily increasing, putting pressure on the natural resources in the basin. It is estimated that between 1984 and 2000, the population increased by 57 %. Figure 2.1 shows the lower boundary of the Volta Basin in Ghana and the first administrative (regional) boundary in Ghana.

Table 2.1. Environmental conditions at Tamale and Ejura study sites

Parameter	Tamale site	Ejura site
Latitude (° ' "N)	9°28'N	7°19'N
Longitude (° ' "W)	0°55'W	1°16'W
Vegetation type	Guinea savannah	Transitional (derived) savannah
Climate (Köppen Class)	Dry hot low latitude (Aw)	Tropical Monsoon (Am)
Mean annual rainfall (mm)	1095	1264
Mean temperature (°C)	28.1	26.6
Altitude (m)	152 m – 193 m	161 m – 178 m
Human activities	Cultivation	Cultivation
Parent material	Voltaian sandstone	Voltaian sandstone

The sites were selected so as to represent the diverse geomorphological, relief, climatic and socio-cultural conditions in the basin across the country (Table 2.1). The two sites lie in the main geological formation, in the two largest ecological zones in the Volta Basin (Figure 2.1), and represents two distinct relief areas, i.e., with flat (Tamale) and moderately steep (Ejura) slopes. These differences are highlighted in the subsequent sections.

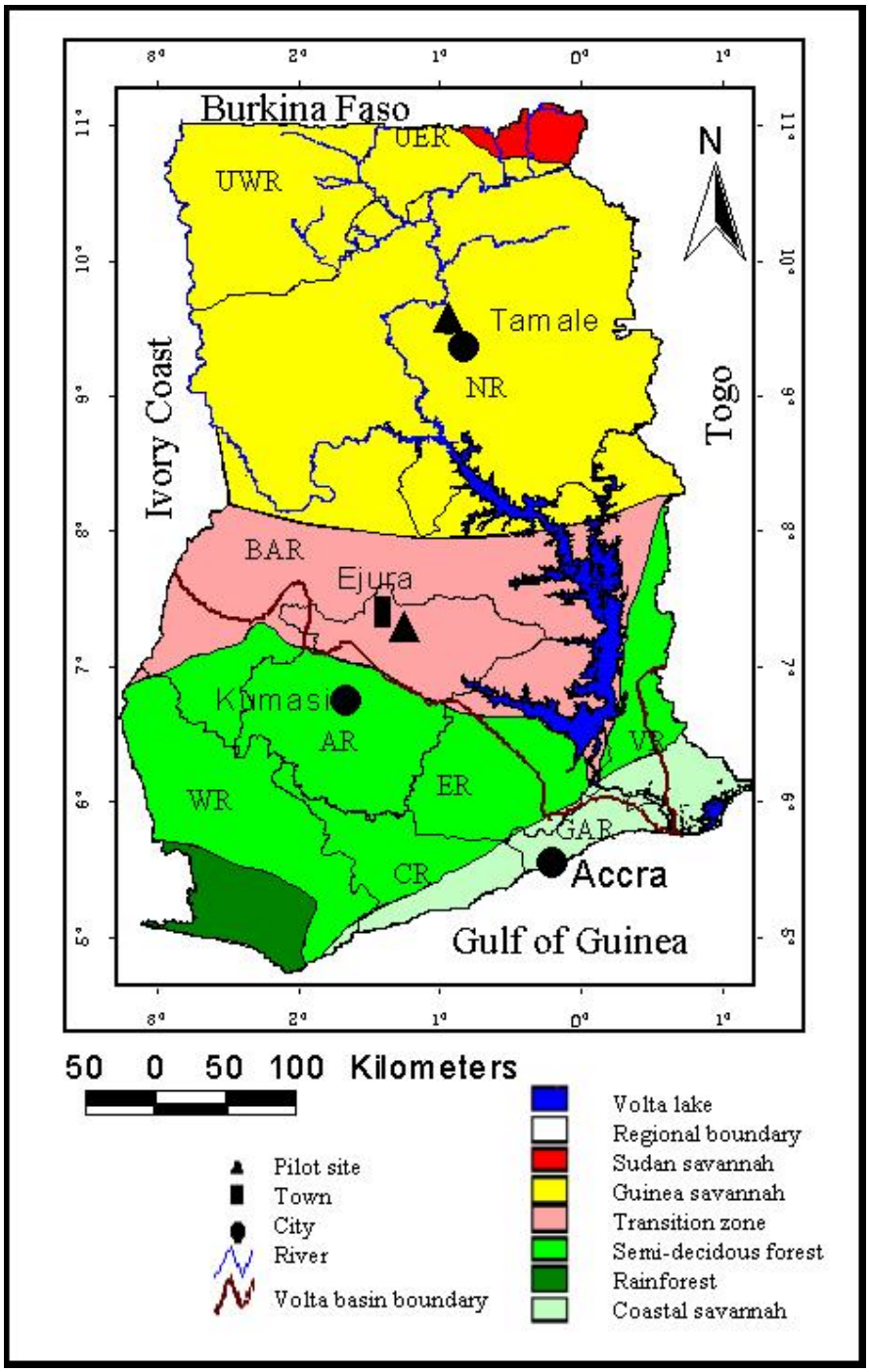


Figure 2.1. Map showing the ecological zones of Ghana, regional boundaries, lower limit of the Volta Basin and pilot sites. The regions as indicated are Ashanti Region (AR), Brong Ahafo Region (BAR), Central Region (CR), Eastern Region (ER), Greater Accra Region (GAR), Northern Region (NR), Western Region (WR), Upper-East Region (UER), Upper-West Region (UWR) and Volta Region (VR). (Source: modified from Taylor (1952))

2.1 Geomorphology, relief and drainage

Ghana is generally a low relief country, except in the east with the Akwapim-Togo mountain ranges. The landscape is characterized by several zones, from a narrow coastal belt that widens from the west to the east in southern Ghana, with a maximum elevation of about 150 m. This coastal belt stretches about 80 km northwards and then rises through a sequence of rolling hill ranges till the Kwahu plateau in the mid-west to mid-east. The Kwahu plateau with an average elevation of 450 m forms the most important physical divide in the country. It separates the basin of the River Volta and its tributaries (such as Afram, Pru, Black and White Volta) northwards from the basins of Pra, Birim, Ofin, Tano and other rivers that flow directly towards the Atlantic sea west of the River Volta. Most of the well-developed alluvial flood plains and terraces of the river Volta are now under the man-made lake established in 1963. Along the upper reaches, narrow stretches of alluvial are encountered. These flood plains are normally 8-13 m high and the terraces may be 20-25 m above the riverbed (Adu and Mensah-Ansah, 1995).

Most of the area in Tamale is gently undulating with broad valleys and isolated low-lying hills and inselbergs (Adu, 1995). The actual site where data collection took place covers three villages (namely: Nwodua, Nyougu, and Kochim). This site is located about 2 km off the main Tamale-Kumbungu road and is about 20 km away from Tamale. The Ejura area is one of the high elevation areas within the Volta Basin. An elevation of about 244 m can be found on the Ejura scarp. The comparatively high elevations are found on Voltaian sandstones that are flat bedded and have gently to steeply sloping topography (Smith, 1962). Ejura is the main town close to the second pilot site at a village called Samari Nkwanta, about 20 km from Ejura through Baabaso.

2.2 Geology

The geology found in the Volta Basin of Ghana (Figure 2.2) consists mainly of the (1) Granite (Middle Pre-Cambrian) formation of granite and granodiorite covering most of the Upper East, Upper West regions and the western part of the Northern region; (2) Birrimian rock formation consisting mainly of phyllite, schist, tuff, and greywacke covering the western portions of the Brong Ahafo and Northern region and along the granitic zone in the Upper regions with phases of metamorphosed sedimentary (d),

sedimentary (e), and volcanic (f); and (3) Voltaian sediments made up of mainly sandstone, quartzite, shale arkose and mudstone, covering the largest portion of the basin from the upper parts of Ashanti and Eastern regions, and most part of the mid-west to the eastern parts of Brong Ahafo and Northern regions (Bates, 1962a). It is made-up of different classes of Voltaian rocks; upper Voltaian (a), Obosum beds (b), and Oti beds of basal sandstone (b) (see Figure 2.1 for regional boundaries).

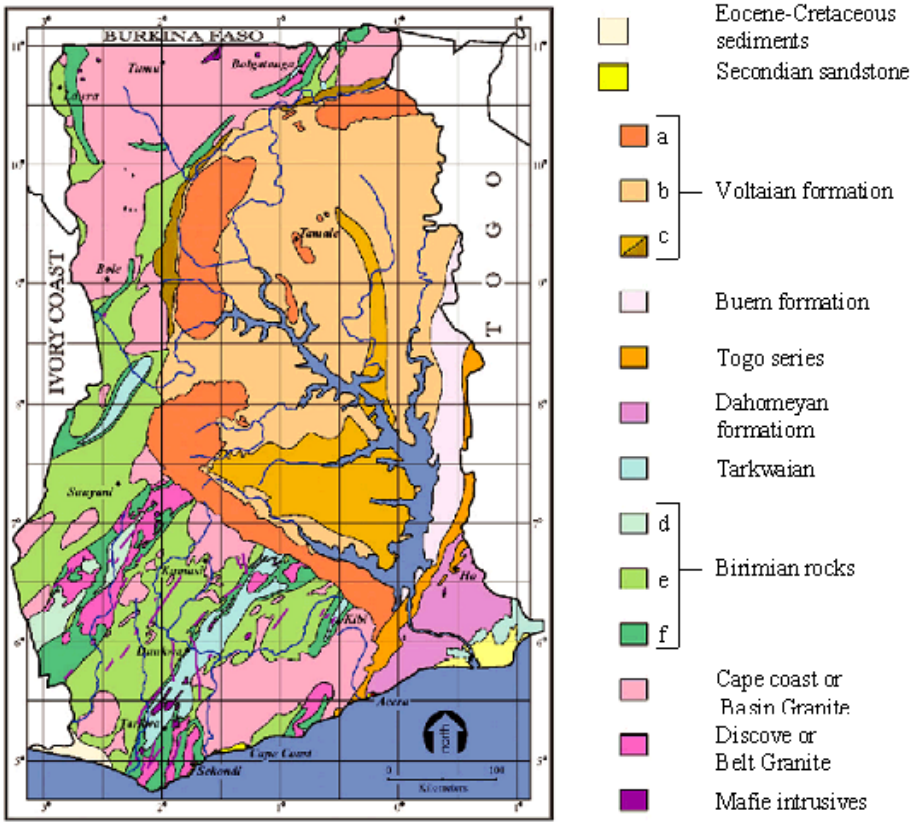


Figure 2.2. Geology map of Ghana (source: modified from Grubaugh, 2003)

The two pilot sites in this study both lie in the Voltaian sandstone basin. The Voltaian system is sedimentary and the most extensive formation in Ghana (covering about 45% of the country). The Voltaian rocks are essentially flat bedded and have not been deformed to any appreciable extent by orogenic movements, except near the eastern and southern margins of the basin. The Voltaian rocks contain few minerals of economic value. The age of the Voltaian sediments is doubtful, as they lack many

fossils, but it is estimated that the rocks may be late Devonian or early Carboniferous Age (approximately 250 million to 290 million years old) (Smith, 1962).

2.3 Soils

The soils found in the Volta Basin of Ghana are predominantly Lixisols, Leptosols, Plinthosols, Acrisols and Luvisols (Figure 2.3). The soils formed from the Voltaian

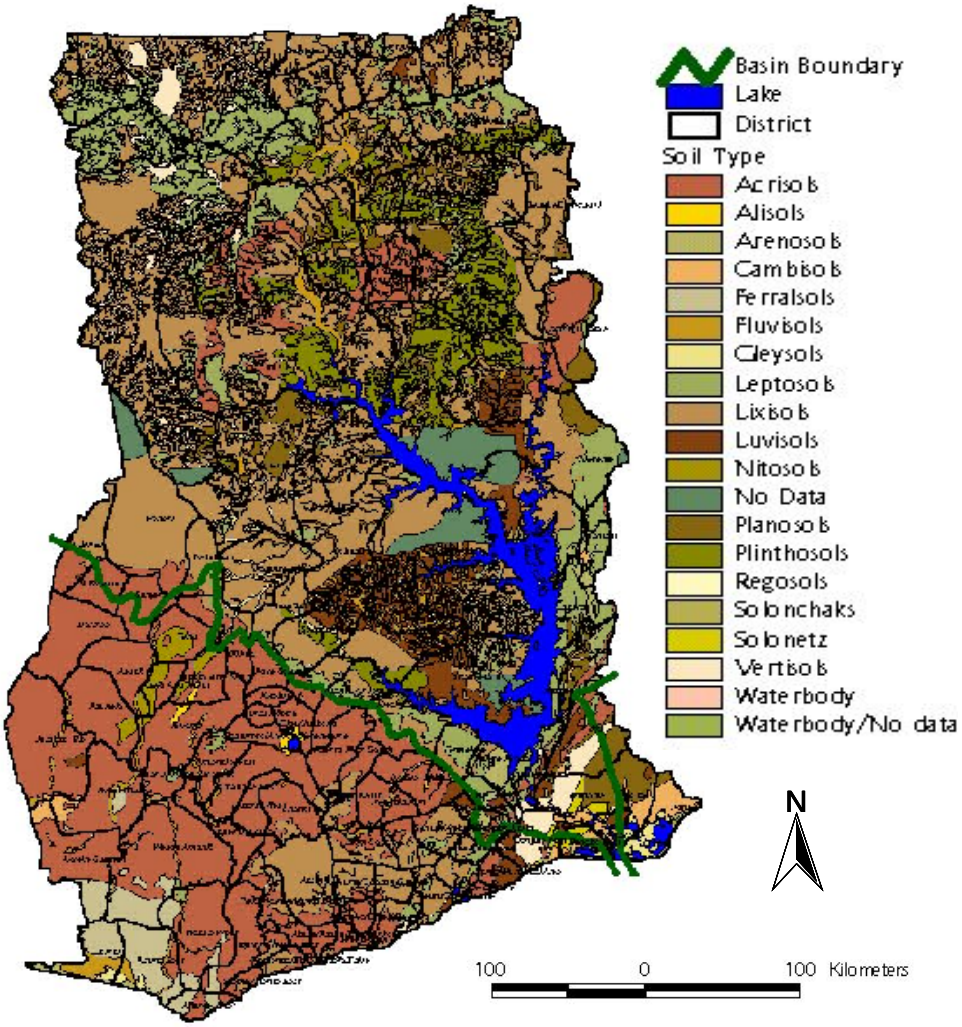


Figure 2.3. FAO soil map of Ghana (Source: Soil Research Institute (SRI), Kumasi, Ghana)

sediments vary widely in terms of soil texture and productivity and generally have low inherent soil fertility compared to the granite derived soils. Furthermore, the

groundwater table in the Voltaian sediments is shallow as compared to that of the granitic soils (Hauffe, 1989). Different hydrological conditions along the slopes have led to the development of different soils from the upland to lowland, resulting in various soil associations (such as Kpelesawgu-changnalili-Lima series and Ejura-Sene soil associations at the Tamale and Ejura sites, respectively) (See Appendix 1 for FAO classification). For further details on the soil associations see Smith (1962); Adu and Mensah-Ansah (1995) and Adu (1995).

Over vast areas, concretionary and shallow iron-pan soils occur, which render the soils unsuitable for cultivation (Asiamah, 2002), while the majority of the soils is highly susceptible to erosion. At both pilot sites, it is not uncommon to find iron and manganese concretions in the mid-upper slope, although it is a more characteristic feature of the soils at the Tamale site. The soils at the Tamale site are usually of low aggregate stability and are therefore limited by unfavorable pore size distribution, low infiltration and high susceptibility to surface sealing by rain water and soil erosion (Hauffe, 1989).

2.4 Hydrology and water supply

The water supply within the Volta Basin is related mainly to the underlying rock, vegetation and land management. In most cases, the nature of the ground surface cover and soil texture are more important than the amount of rainfall, as these influence the amount of water that can percolate into the underlying rock (Bates, 1962b). The Voltaian rocks are compacted and cemented, and with the possible exception of some of the arkoses (Obosum sandstone), referred to as sandy rocks, which form a long belt between Kete Krachi and Sang, with the rocks themselves not being permeable (Bates, 1962b). Greater central parts of the Voltaian formation – covering areas of Atebubu, Salaga, Tamale and Nasia – is underlain by shale, except for few patches of sandstone (Tamale area). In some parts, the soil is underlain by mudstone, or are shallow soils underlain by 30 cm or more of lateritic hardpan (Bates, 1962b). Percolation below this hardpan is practically zero, thereby making the area prone to shallow floods in the wet season.

2.5 Climate

The climate in Ghana is controlled primarily by the tropical continental air mass or North East Trade Winds and the tropical maritime air mass or South West Monsoon wind (Dickson and Benneh, 1988). The distinct seasons experienced in the country are the result of the movement and position of the convergence zone - referred to as the Inter-Tropical Convergence Zone (ITCZ) - between the two air masses. Figure 3.4 illustrates the long-term monthly rainfall (Rf) and temperature (Tp) distribution at Tamale (1939-2000) and Ejura (1973-1993) pilot sites.

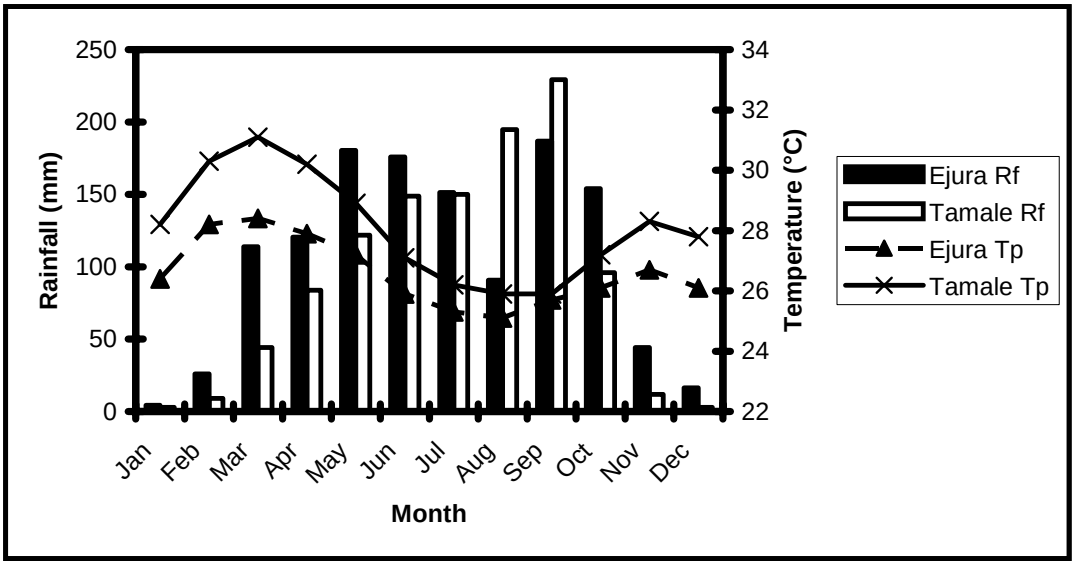


Figure 2.3. Monthly rainfall (Rf) and temperature (Tp) distribution at Tamale (1939-2000) and Ejura (1973-1993) pilot sites (Source: drawn using data from Ghana Meteorological Service)

The Tamale site in the Northern Region has a tropical continental or interior savanna climate, mostly influenced by the tropical continental air mass. It is classified as dry hot low latitude climate (Aw) according to the Köppens climatic classification. The rainy season lasts about seven months (April – October) and the dry season the remaining five months (Figure 2.3). This study area receives about 1000 – 1200 mm of rainfall a year (Kasei, 1993). The rains are very erratic in terms of onset, frequency, duration and amount (Adu, 1969, Kasei, 1993). The onset of rains is highly unpredictable, as it may commence anytime between March and June. The region is

also characterized by a mid-season drought or short dry spells, usually in the month of July, that adversely affect crop production. Annual temperature is about 28°C with potential evaporation of 1500 – 1800 mm (Kasei, 1993). Rainfall in the months of July to September is usually in excess of the monthly potential evaporation.

The Ejura site lies within the transitional ecological zone (Figure 2.1), which is between the Guinea savanna zone to the north characterized by a single rainy season and the forest vegetation to the south characterized by two rainy seasons. However, the influence of the tropical maritime air mass (South West Monsoon) is dominant. It is classified as tropical monsoon climate (Am) according to the Köppens climatic classification. The rainy season is characterized by monthly rainfall totals rising slowly from March, with a slight decrease in July and/or August, until a maximum is reached in September or October (Figure 2.3), thus resulting in a main and minor rainy seasons. Mean annual rainfall is 1264 mm with a very high annual and monthly variability, and no clear-cut beginning and end of rains (Smith, 1962, Adu and Mensah-Ansah, 1995). Rains are of short duration, lasting an hour or two. The average temperature is 26.6 °C, with the highest mean monthly temperature in March just before the onset of the rains.

Generally, relative humidity across the country decreases northwards. Relative humidity in the south is usually more than 90% during the night and early morning. In the north, between April or May to October (rainy season), night and early morning relative humidity may average 95 % falling to about 70% in the afternoon. During the dry season, average relative humidity is below 80% and falls to as low as 25% in the extreme north as a result of high day temperature (Walker, 1962). Seasonal and diurnal variation in relative humidity increases northwards with the seasonal variation being more pronounced. Wind speeds are generally low, averaging below 8 km per hour inland and on the coast between 8 and 16 km per hour measured at about 9 m above ground in a reasonably open area.

2.6 Vegetation and land use

Ghana can be put into two broad agro-ecological zones, i.e., savannah and forest. The savannah zone comprises of the Guinea savannah covering most of Northern Ghana, the Sudan savannah covering the north eastern corner of the country and the coastal savannah the south eastern part of the country. The forest zone consists of the rain forest

in the southwestern corner of the country and the deciduous forest covering the major part of southwest area and a narrow corridor along the eastern border with Togo. Between these two broad categories of agro-ecological zones is the forest-savannah Transitional zone. Illustrated in Figure 2.1 are the ecological zones in Ghana across the regions. The Volta Basin in Ghana covers mainly the transitional, Guinea and Sudan savannah zones. The Tamale and Ejura sites are located in the Guinea savannah and Transitional zones, respectively.

In Northern Ghana all land belongs to the chief and is allocated to lineages in a patrilineal system. With the permission of the chief, all other people who do not belong to the local ethnic group can obtain access (usufruct) rights to land. However, the government has an overriding power over the use of land. This complicated communal land tenure discriminates against women in terms of access to land, encourages farm fragmentation at the expense of large scale mechanize farming, constrains the use of land as a collateral security for loans, and discourages the good management of land. The Guinea savannah zone is very important to the agricultural economy of Ghana, as it produces the bulk of the country's food crops and livestock (MoFA, 1991). Increasing human population in the zone has led to increasing pressure on the most important farm resource (land). This has made the original farming system of shifting cultivation with long fallow periods less viable. The main crops cultivated on uplands are yam (*Dioscorea spp.*), sorghum (*Sorghum bicolor*), maize (*Zea mays*), groundnut (*Arachis hypogaea*), cowpea (*Vigna unguiculata*), and cotton (*Gossypium spp.*) with rice (*Oryza spp.*) in valley bottoms. Livestock production is very important to the economic life of the people in Northern Ghana. Mostly, the family owns cattle with the family head having the direct responsibility for the animals. The major problems facing livestock keeping in the Northern region are lack of drinking water, grazing areas, feed for the animals during the dry season, diseases and overstocking. Sheep, goat and poultry, such as guinea fowl and fowls, which are owned individually, can be found in almost every household. Mainly as a result of the annual bush burning, wildlife is presently rare except in forest reserves and in some instances in woodland along rivers.

The principal occupation in the Ejura area is agriculture. Individuals mainly hold a land title. The majority of the farms are small-scale subsistence farms with a few commercial farms. The principal food crops cultivated in the area include: maize,

cassava (*Manihot esculenta*), yam, groundnut and cowpea; mango (*Mangifera indica*), cashew (*Anacardium occidentale*) and *Citrus spp*. are the main tree crops. Traditional shifting cultivation with land rotations is widely practiced. Most households own some livestock, mainly sheep and goat, fowls and guinea fowls. Hunting plays an important role in the daily lives of the people in the area. It provides abundant meat for most households with some hunters consequently becoming professionals.

2.7 Social conditions

Though Ghana is made up of many ethnic groups, there are four main tribes, namely: Akan, Moshi-Dagomba, Ewe, and Ga. The dominant religions in the country are Christianity, Islam, and indigenous beliefs. The official language is English, with Akan being the most widely spoken African language. On average, life expectancy is 57.4 years and infant mortality is 57.4 deaths to 1000 live births.

With respect to tribes, the Northern region, where the Tamale site is located is very heterogeneous. The main ethnic groups are the Dagomba, Gonja and Mamprusi. The Gonjas occupy the largest land area but the Dagombas are the highest in number. The 2000 Census gives the population of the region as 1.82 million with a growth rate of 2.8 % (Table 2.1). A large portion of the population is found along the main south-north road through the regional capital Tamale with a population of 202,317 (GSS, 2000a). The primary occupation of the Dagombas and Mamprusis is farming while the Gonjas are mainly hunters. With about 29.5 % of the total land area in Ghana, the Northern region is the least densely populated out of the ten regions (25.9 people per square kilometer). About 67.2 % of the population in the region works in agriculture and related industries. The average household size is about 7.4, which is the largest in the country. The region can boast of a university and an agricultural research institute. The actual site where data collection took place near Tamale covers three villages (namely: Nwodua, Nyougu, and Kochim). The people in this area are mostly Dagombas and mostly Moslems. The area has a primary, junior secondary and vocational school. However, for senior secondary school the children have to travel to Tamale, Kumbungu or Tolon. The people enjoy pipe-borne drinking water as the area is located along the main water supply line to the regional capital. For major health needs, the people have to visit Tamale hospital, but for minor ailments they may visit clinics in neighboring

villages. Nwodua has electric power, but the two adjoining villages are yet to be connected. Houses in this area are mainly built of mud with thatch or iron roofs. The people in these villages are predominantly Moslem. The road linking these villages and also to the main road are untarred, but motorable all year round. With the Savannah Agricultural Research Institute (SARI) about 12 km from the site the people stand to benefit from agricultural technologies. The main economic activities are crop production and animal rearing. Other activities may take place in Tamale due to its proximity.

Table 2.2. Basic population data for Ashanti region, Northern region and Ghana

	Ashanti region	*Northern region*	*Ghana*
Population (million)	3.6	1.8	18.4
Growth rate (%)	3.4	2.8	2.7
Population density per km^{-2}	148.1	25.9	79.3
*Pop. in agric. and related industries (%)	67.2	41.6	47.9
Household size	5.3	7.4	5.1

Source: GSS, 2000a and *GSS, 2000b.

The Ashanti region, where the Ejura site is located, has a population of 3.61 million with a growth rate of 3.4 (2000 census) and an average household size of 5.3 as compared to the national average 5.1 (Table 2.2). The region has a population density of 148.1 people per square kilometer with 41.6 % of the active working people in agriculture and related industries. Ejura is the main town close to the sampling site at Samari Nkwanta. Ejura has a population of 29,478 (GSS, 2000a) consisting mainly of Ashantis who are the landlords and Dagombas and other northern tribes who are mainly migrant workers. In the main Ejura township, houses are constructed from blocks, but at the sampling site most houses are built of pole-frame mud houses. The road network linking Ejura to Kumasi (second largest city) and to Yeji (inland port on the Volta lake) are good, but the road linking Ejura to Techiman through Nkorasa to the north is in a poor state. The road linking Ejura to the sampling site is untarred and very difficult to use during and just after heavy rains. Samari Nkwanta has a primary school and a borehole. For further education at the level of Junior Secondary School and Senior Secondary School the children have to travel to Ejura. Ejura as a third level administrative capital (i.e. district capital) has pipe-borne water, electricity, churches

and mosques, postal and telecommunication services, banking services and a hospital. The Hausa or Dagomba community is predominantly Moslem, while the Ashantis are mainly Christians, animists or plain heathens. Busangas from the Upper East region and other tribes from northern Ghana, who are predominantly Moslem settler farmers inhabit Samari Nkwanta.

3 MATERIALS AND METHODS

This chapter outlines the materials and methods used in carrying out this study at Tamale and Ejura pilot sites. Here, the methodology used in collecting field data (such as cutting of transects, soil sampling, Differential Global Positioning System (DGPS) point data collection and soil classification or mapping) and laboratory analysis of soil samples for particle size distribution (i.e. sand, silt and clay content), organic carbon, cation exchange capacity (CEC), pH, bulk density, and saturated hydraulic conductivity are elaborated. Also the procedures used to generate point data, create DEM and subsequently develop terrain attributes are explained. Procedures such as ANOVA and correlation coefficient (r) used to evaluate the potential of estimation are outlined.

3.1 Transects creation

To enable soil data collection at a regular intervals for geostatistical analysis, transects were cut at the two sites through the vegetation cover to create grid points at which sampling was carried out. A courser grid was used at the Tamale site compared to the Ejura site (Figure 3.1) to a larger area, taking into account the high cost of soil analysis and the fact that the Tamale site has a gentler sloping terrain compared to the Ejura site.

A 3-km baseline was first made through the centre of the 6-km^2 area at the Tamale site. This was then divided into strips by running parallel traverses at right angles to the baseline at 100-m intervals. Stakes or pegs were put at 200-m intervals along the traverses up to 1000 m on each side of the baseline, resulting in rectangular grid of 100 × 200 m^2.

At the Ejura site, an 800-m baseline was made through the center of 0.64-km^2 area. Then the area was divided into strips by running parallel traverses at right angles to the baseline at 40-m intervals. Stakes were put at 40-m intervals along the traverses up to 400 m on each side of the baseline. The traverses were staked in such a way that even-numbered traverses were located at the centre of odd-numbered ones, resulting in a triangular grid (with a base of 40 m and height of 40 m).

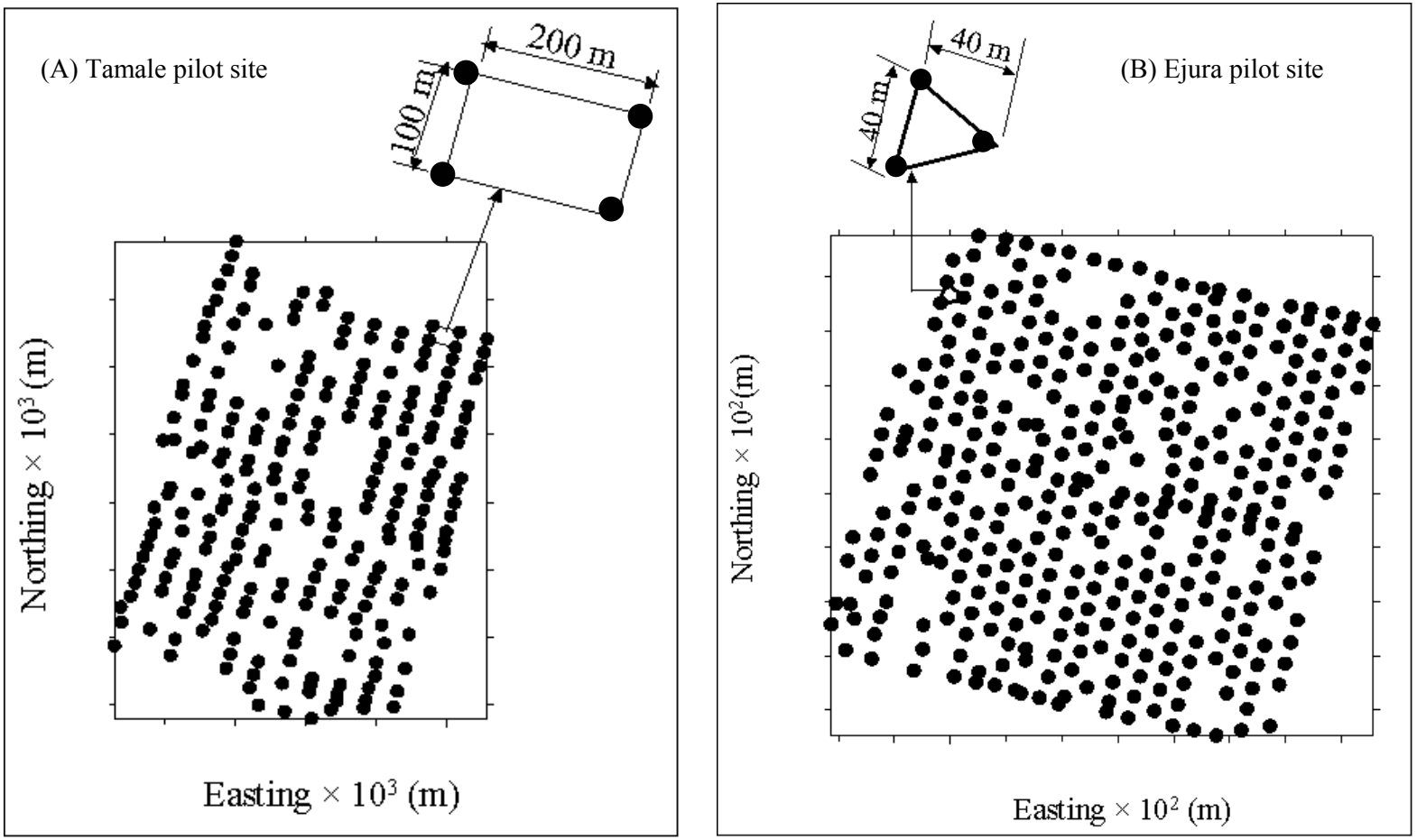


Figure 3.1. Sampling grid pattern at (A) Tamale and (B) Ejura pilot site

3.2 Soil sampling and analysis

At the staked points, disturbed and undisturbed soil samples were collected from 0 – 15 cm (topsoil) and 30 – 45 cm (subsoil) depths. These two soil zones represent the most disturbed and least disturbed for the rooting zone of most crops. At each sampling point, soil samples were first taken from the topsoil, and then the 15 – 30 cm was removed with a hoe – taking care not to unnecessarily disturb the soil – to take the subsoil sample.

The undisturbed samples were taken using a 10.0 cm long by 8.3 cm diameter cylindrical metal core with the help of a ring holder. The sampling core was inserted into the ring holder, which was then inverted onto the soil. The handle of the holder was then tapped gently with a mallet until the top of the soil core was about 0.5 cm below the soil surface. The soil around the holder was dug, the soil sample core brought out and excess soil cut off with a soil knife.

All disturbed soil samples were air dried and sieved (2 mm and 0.5 mm), analyzed for particle size distribution, pH, organic carbon and CEC. The undisturbed soil samples were analyzed for bulk density and saturated hydraulic conductivity. The same analysis was carried out on samples collected from different horizons in the profile pits.

Laboratory analysis was carried out to determine soil pH, CEC, organic carbon, particle size distribution, bulk density and saturated hydraulic conductivity (see below).

The soil properties studied in this work include: particle size distribution (i.e. sand, silt and clay), soil structure, CEC, organic carbon, pH, bulk density and saturated hydraulic conductivity (K_s). These properties influence the behavior of soil in relation to water movement, retention and loss, and nutrient dynamics.

3.2.1 Soil pH

The pH of the soil as an indicator of soil acidification is mainly dependent on the type of parent material and the level of soil leaching in the environment. It was determined by suspending the soil in 0.01M $CaCl_2$ solution in a 1:2.5 soil to solution ratio. The suspension was stirred intermittently for 30 minutes. The pH was taken using a pH-meter and a combined glass electrode (Thomas, 1996). Soil pH classification values are provided on Table 3.1.

Table 3.1. Soil pH classification values

Description	*pH*
Extremely acid	< 4.5
Very strongly acid	4.5 – 5.0
Strongly acid	5.1 – 5.5
Medium acid	5.6 – 6.0
Slightly acid	6.1 – 6.5
Neutral	6.6 – 7.3
Mildly alkaline	7.4 – 7.8
Moderately alkaline	7.9 – 8.4
Strongly alkaline	8.5 – 9.0

Source: Smith, 1962.

3.2.2 Soil cation exchange capacity (CEC)

The CEC is the soil's ability to retain or store and fix cationic nutrients. It is influenced by the amount and activity of clay mineralogy, organic matter content and soil reaction or pH.

The silver-thiourea (AgTu) extraction procedure was used (Chhabra et al., 1975; Pleysier and Juo, 1980 and Searle, 1984). The AgTu complex behaves as a large cation with a single positive charge (see equation 3.1); its high polarizability partially explains its very high ability to exchange cations from the negative surface charge. The free energy of exchange of Na and other cations in soil by AgTu is about 5 kCal/equivalent at 25°C. For example:

$$(Ca, Mg, K, Na, Al) - Soil + AgTu^{(+)} \rightarrow AgTu - Soil + Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Al^{3+} \quad [3.1]$$

The unbuffered solution has a pH value of about 6. All exchangeable cations (basic and acidic) are displaced by a small amount of AgTu. Therefore, a dilute solution of AgTu can be used to extract all exchangeable cations and saturate the soil with $AgTu^{(+)}$, thus allowing for a one step extraction. In this procedure, due to the absence of a silver lamp to use on the Atomic Absorption Spectrometer (AAS), the individual cations of Ca^{2+}, Mg^{2+}, K^{+}, H^{+} and Al^{3+} were measure by AAS and flame photometer, and by titration using and automatic titration after extracting the soil with 0.01 M AgTu. The extraction ratio was 10 g of soil to 25 mls of 0.01M AgTu solution, shaking for 30

minutes and filtrating with Whatman 42 filter paper. The individual cations were summed up after calculating the $Cmol^{(+)}kg^{-1}$ soil to get the CEC:

$$CEC = \frac{Ca\ mg/kg\ soil}{200.4} + \frac{Mg\ mg/kg\ soil}{121.6} + \frac{K\ mg/kg\ soil}{391.0} + Al^{3+}\ Cmol + H^{+}\ Cmol \quad [3.2]$$

Note: meq/100 g soil ≡ $Cmol^{(+)}kg^{-1}$ soil
Thus (Al^{3+} + H^{+}) meq/100 g soil ≡ (Al^{3+} + H^{+}) $Cmol^{(+)}kg^{-1}$ soil

3.2.3 Soil organic carbon

Soil organic carbon influences water holding capacity, cation exchange capacity and erodibility. Soil carbon stock is one of the indices of sustainable agriculture. It affects storage and release of nutrient ions, buffer system for soil pH, nutrient carrier, mobilization of phosphorous, and soil structure.

The Walkley Black procedure was used. The organic carbon was oxidized by a known concentration of Potassium dichromate (1.0 N or 0.166 M) solution added in excess. The excess unreacted dichromate was titrated with 0.5 M Ammonium iron (II) Sulphate in a redox reaction using diphenylamine indicator.

$$2Cr_2O_7^- + 3C + 16H^+ \rightarrow 4Cr^{3+} + 3CO_2 + 8H_2O \quad [3.3]$$

$$Cr_2O_7^- + 14H^+ + 6Fe^{2+} \rightarrow 6Fe^{3+} + 2Cr^{3+} + 7H_2O \quad [3.4]$$

$$\%\,organic\,carbon = \frac{M(V_1 - V_2) \times 0.39}{S} \quad [3.5]$$

where M = molarity of Potassium dichromate; V_1 = blank titration; V_2 = sample titration; and S = weight of soil sample.

The constant 0.39 takes into account the incomplete combustion of organic matter in this procedure. To obtain % organic matter, the organic carbon content is multiplied by the van Bemmelen factor of 1.724.

3.2.4 Particle size distribution: The hydrometer method

The soil particle size together with its mineralogical composition largely determines the nature and behavior of soil, i.e., internal geometry and porosity, its interactions with fluids and solutes, as well as compressibility, strength, and thermal regime (Hillel, 1998). The three main fractions are clay (<2 μm), silt (2-50 μm) and sand (50-2000 μm). The clay fraction as the colloidal fraction amongst the particle sizes has the most influence on soil behavior as a result of its greater surface area per unit mass leading to a strong physicochemical activity. It is responsible for soil carbon stabilization, anion and cation retention, and the physical cohesion that influences the erodibility of the soil.

Gravel, stones or boulders are solid particles of sizes greater that 2 mm. Though not considered to be part of the soil in situations where the gravels and stones occupy enough of the soil's volume, it can influence the soil physical processes significantly. The Bouyoucos (1962) method was used to determine the particle size distribution on 2-mm sieved air-dry soil.

Soil texture

The soils were classified into different textural classes using the USDA textural triangle. This was done using the computer program for soil textural classification developed by Gerikis and Baer (1999) that uses percent sand and clay.

3.2.5 Bulk density

Bulk density gives an indication of soil compaction. It affects resistance to root growth and development, soil porosity and infiltration capacity. It is affected by the structure of the soil, i.e., looseness or degree of compaction, as well as by swelling and shrinking characteristics (Hillel, 1998).

Soil bulk density (gcm^{-3}) was determined by the core method (Blake and Hartage, 1986) using a 10.0 cm long by 8.3 cm diameter cylindrical metal core. Samples were dried at 105 °C for 24 hours in a forced air oven, weighed and density calculated as sample dry weight (g) divided by sample volume (cm^3).

3.2.6 Saturated hydraulic conductivity

The saturated hydraulic conductivity (K_s) measurements were made on 10.0 cm long by 8.3 cm diameter core samples in the laboratory using the falling head method developed by Klute and Dirksen (1986). The soil in the core is held in place with a fine nylon cloth, tied with a rubber band and soaked in water for 24 hours or until saturated. The soaked soil is fitted with another cylinder of the same diameter but of 20 cm height at the top of the core to allow imposition of a hydraulic head. A large metallic box with perforated false bottom is filled with fine gravel (<2 cm). A fast filtration filter paper is place on the soil core. With the core placed on the gravel box, water is gently added to the core to give a hydraulic head in the extended cylinder. The water then flows through the soil and is collected in the box and drained off by plastic pipe tubing. The fall of the hydraulic head H_t at the soil surface was measured as a function of time t using a water manometer with a meter scale. Note: Enough time must be allowed for water to flow through the soil to ensure uniform flow. Saturated hydraulic conductivity was calculated by the standard falling head equation:

$$K_s = \left(A_c L \Big/ A_s t \right) \ln\left(H_t \Big/ H_0 \right) \qquad [3.9]$$

where A_c is the surface area of the cylinder, A_s surface area of the soil, H_0 and H_t are the hydraulic head at time t = 0 and t = t, respectively, and L is the length of the soil sample. Since $A_c = A_s$, rearranging equation 3.9, a regression of $\ln(H_0/H_t)$ against t with slope b:

$$K_s = bL \qquad [3.10]$$

Table 3.2 presents the classification for saturated hydraulic conductivity.

Table 3.2. Saturated hydraulic conductivity classification values

Description	Saturated hydraulic conductivity (cmh^{-1})
Very slow	< 0.8
Slow	0.8 – 2.0
Moderate	2.0 – 6.0
Moderately rapid	6.0 – 8.0
Rapid	8.0 – 12.5
Very rapid	> 12.5

Source: Landon, 1991.

3.3 Environmental variable (terrain attributes) generation

In this section, the various procedures used in generating DEM and terrain attributes are explained. The behavior of the DEM and terrain attributes in relation to soil parameters is outlined in Chapter 5. The terrain attributes are used as part of the input parameters to model saturated hydraulic conductivity.

Most important biophysical processes operating on the earth's surface are influenced by past events, contemporary controls, interactions and thresholds (Montgomery and Dietrich, 1995). The influence of surface morphology on catchment hydrology and the impact of slope, aspect and horizon shading on insolation probably represent the most important controls operating at the topo-scale (Wilson and Gallant, 2000). Many studies have shown how the shape of the land surface can affect the lateral movement and accumulation of water, sediments, and other constituents (Moore et al., 1988 and Gessler et al., 1995).

There are three main sources of topographic data: surface-specific point elevation, contour and streamline, and remotely sensed elevation data. These data can be organized into one of three data structures, i.e., (1) regular or square grids (digital elevation model (DEM)), (2) triangulated irregular network (TIN), and (3) contours. The choice of one method over the other is usually a question of preference and / or source of data (Moore et al., 1991 and Wilson and Gallant, 2000).

Topographic (terrain) attributes or indices, such as slope, aspect, plan curvature and profile curvature can be derived from digital elevation data sets. This is done through the derivation of each element as a function of its surroundings or neighbors (Moore et al., 1991). These attributes can be classified into two main groups, i.e. (1) primary attributes that are physically based indices derived empirically to spatially characterise the variability of specific processes occurring in the landscape and (2) secondary attributes that are computed from two or more primary attributes that allow the description of a pattern as the function of a process (Moore et al., 1991, 1993a, and Wilson and Gallant, 2000). These attributes are used to describe the morphometry, catchment position, and surface attributes of hillslopes and stream channels forming a drainage basin.

Most primary terrain attributes can be computed using a second-order finite difference scheme or by fitting a bivariate interpolation function (such as: $z = f(\chi, y)$)

to the DEM and then calculating the derivatives of the function (Moore et al., 1993c; Mitasova et al., 1996). Many authors, including Dikau (1989), Giles (1998), and Burrough et al. (2000, 2001) have used computed topographic attributes to generate formal landform classification.

The attributes that quantify the roles played by topography in redistributing water in the landscape and modifying the amount of solar radiation received at the surface have important hydrological, geomorphological, and ecological consequences in many landscapes. These attributes may affect soil characteristics because pedogenesis of the soil catena is affected by the way water moves through the environment in many landscapes, the distribution and abundance of soil water, the susceptibility of the landscapes to erosion by water, and the distribution and abundance of flora and fauna (Wilson and Gallant, 2000).

Slope influences the flow rate of water and sediment by controlling the rate of energy expenditure or stream power available to drive the flow. It influences the amount of sediment that moves downslope and also the amount of water available for eluviation and illuviation, and leaching in the soil profile. Consequently, it influences the process of soil formation both laterally and vertically.

Aspect defines the slope direction and orientation. It therefore determines the flow direction and also the slope exposure to climate, for instance the amount of solar radiation and subsequently the rate of weathering.

Plan curvature – i.e., curvature of the land surface transverse to the slope direction – influences the convergence and divergence of flow. It therefore gives an indication of areas liable to high flow accumulation and poor drainage.

Profile curvature – i.e., the curvature of the land surface in the direction of steepest descent – is an indication of the rate of change of slope. It affects flow acceleration and deceleration and therefore soil aggradation and degradation.

Upslope contributing area gives an indication of the amount of flow accumulating in a particular cell or grid. It therefore has a direct relationship with sediment deposition as well as leaching and water transport through the soil, which influences soil formation.

Wetness index describes the spatial distribution and extent of zones of saturation. A high wetness index indicates regions likely to be of poor soil drainage and

high water availability, whereas regions with a low wetness index are susceptible to temporary soil drought. It has a direct influence on soil formation and subsequently on soil hydraulic conductivity.

The stream power measures the erosive power of the flowing water. It gives information on the ability of the flowing water to detach and move particles downhill. The higher this value the more prone is the area to erosion.

Length-slope factor or sediment transport capacity index is a form of stream power index. It is very useful in identifying locations of net erosion and net deposition.

3.3.1 DEM generation

A Differential Global Positioning System (DGPS), Ashtech equipment (Ashtech, 1998a), was used to generate point measurements. This was done by first creating a control site at which one of the satellite antennas or receivers was mounted on a tripod. The second antenna was then mounted on a bipole. The two antennas were synchronized using a handheld device and data recorded at 15-sec intervals (see Ashtech manual for details (Ashtech, 1998). Each data set was checked and the procedure repeated until the entire site was covered. During the field campaign, at specific locations where soil data was collected, the site identification codes were noted.

The data was processed using the Locus processor 1.2 (Ashtech, 1998b). The processing was carried out using Clark 1880 Modified parameters as provided by the Survey Department of Ghana, Accra (Table 3.3). In the absence of World Geodetic System (WGS) 1884 control points, a control point was created and data taken at this point in the static mode for three or more hours to insure accuracy. The data quality was checked, and site coordinates with failed quality assurance (QA) and high standard error (i.e., > 10) were not used for DEM development. The data was reprocessed and then adjusted to improve the site coordinate quality and reliability.

Semivariogram analyses were conducted prior to a kriging interpolation of the elevation data. Using the appropriate model function, the point measurements were interpolated at a grid size of 30 m. A 30-m grid resolution was chosen in accordance with work done by Park et al. (2003). The semivariogram analysis was performed using S-Plus software (Mathsoft Inc., 1999) and interpolation performed in Surfer 7

(Goldensoftware Inc., 1999) program. The grid data was then imported into DiGeM for terrain analysis (Conrad, 2001).

Table 3.3. Parameters for coordinate system

Parameter	Value
Projection type	Transverse Mercator
Projection parameters	
Longitude of central meridian	001°00'00.00" W
Scale factor of central meridian	0.999750
Latitude of grid origin	04°40'00.00" N
False Easting	274320.000m
False Northing	0.000m
Geodetic datum transformation parameters	
X shift	-159.506 m
Y shift	262.531 m
Z shift	221.791 m
X rotation	1.828 arc sec
Y rotation	2.399 arc sec
Z rotation	7.46 arc sec
Scale difference	1.0000039 ppm
Ellipsoid:	
Ellipsoid type	C1880MOD
Semi-major axis	6378306.604
Inverse flattening	296.000
Projection accuracy desired:	
Horizontal	1 m + 1000 ppm
Vertical	0.1 m + 1000 ppm

3.3.2 Terrain attribute generation

The imported grid data in DiGeM 2.0 (Conrad, 2001) - a terrain analysis program – was then used to generate eight different terrain parameters. In DiGeM, a pre-processing was first carried out through which all sinks were filled. The morphometric properties (aspect and slope gradient) with the option to create curvatures (plan curvature, profile curvature and curvature) using the Zevenbergen and Thorne (1987) method were generated. Also, an upslope contribution area was generated from flow accumulation using the multiple flow direction algorithm (Freeman, 1991). Finally, topographic indices (wetness index, stream power and LS factor) were generated using DiGeM. Details on the relevance of these generated terrain attributes (aspect, slope gradient, plan curvature, profile curvature, curvature wetness index, stream power and LS factor) are provided in section 4.2. DiGeM is a free stand-alone program for calculating various

terrain parameters (see www.geogr.uni-goettinen.de/pg/saga/digem/). The data was saved as an ASCII file and imported into ArcView 3.2 for visualization. The resulting grid point data was used for further statistical analysis.

3.4 Statistical methods

This section describes the general statistical procedure used in investigating soil variability, compares soil and environmental attributes, and spatial distribution of soil. Here the descriptive statistics used, method of data preparation for further analysis, and correlation analysis carried out between the different parameters is explained. Where a specific statistical procedure different from what is outlined here is used it is clearly stated in the test.

3.4.1 Descriptive statistics

The summary statistics of all data used in the analysis was done using SPSS 10.0 (SPSS Inc., 1999) and presented in section 4.1. These data provide the mean, minimum and maximum, standard deviation, skewness, kurtosis, and coefficient of variation for all the data.

3.4.2 Data preparation and transformation

The highest and lowest 5 % of the saturated hydraulic conductivity data weer not used ensured quality data. This level was set to take care of field and experimental errors that may result in extremely high or low values of K_s, as the accuracy of this parameter is very critical for model development. All continuous data was assessed for normality, following the steps outlined below:

1. The distribution of each data was examined the histogram density graph of SPLUS for skewness and kurtosis
2. The skewness and kurtosis, which are expected to be zero for both measures in a normally distributed data set using SPSS, were analyzed
3. A q-q plot for each parameter was observed for normality.

Data that were not normally distributed were transformed into the normal distribution (or close to normal distribution) using the ladder of powers (Table 3.4)

(Hamilton, 1990) before further analysis. Normalization was necessary because most multivariate and parametric statistical tests are based on the assumption of multivariate normality, i.e., all the variables used and their combinations are normally distributed. Since multivariate normality testing is extremely difficult, univariate normality as used here reduces heteroscedasticity (unequal variability in one variable compared to another) that may weaken the analysis. Though it does not ensure multivariate normality it increases the likelihood. This normalization was done to obtain the same data distribution for better comparison.

Table 3.4. Nonlinear transformation for changing distributional shapes (ladder of powers

Power	Transformation	Name	Effects	Inverse function
3	X^3	Cube	Reduces extreme negative skewness	$(X^*)^{1/3}$
2	X^2	Square	Reduces negative skewness	$(X^*)^{1/2}$
1	$X^1=X$	Raw	No effect	X^*
-0.5	$X^{1/2}$	Square root	Reducess mild positive skewness	$(X^*)^2$
-1	$Log_{10}(X)$	Log	Reduces positive skewness	10^{X^*}
-1.5	$(X^{-1/2})$= -1/SQRT(X)	Negative reciprocal root	Reduces extreme positive skewness	$(-X^*)^{-2}$
-2	$-(X^{-1})$=-1/X	Negative reciprocal	Reduces very extreme positive skewness	$(-X^*)^{-1}$

X*: transformed variable, Source: Hamilton (1990)

3.4.3 Correlation analysis

Correlation analysis was used in this study to determine the relationship between different pairs of variables and the trend of this relationship if it exists. In order to identify associations among measured and estimated soil and environmental parameters, Pearson's r correlation matrix was constructed after a min-max transformation with the range of zero-one (0-1) for normalized data. This transformation is a method of standardization that gives data with a minimum value of zero and a maximum value of one. Using the following function, the 0-1 min-max transformation was made:

$$y' = \left(y - \min 1 \Big/ \max 1 - \min 1 \right)(\max 2 - \min 2) + \min 2 \qquad [3.11]$$

where y is the original value, y′ is the new value, min 1 is the original minimum value, and max 1 is the original maximum value. Also, min 2 = 0 and max 2 = 1 are the new minimum and maximum values, respectively, thus

$$y' = \left(y - \min 1 \Big/ \max 1 - \min 1 \right) \qquad [3.12]$$

Min-max transformation was used to put all data on a common scale. As a linear transformation, standard deviation, skewness, and kurtosis are not affected by the transformation. The correlation matrix gives information on the magnitude of association (i.e. intensity), direction of association, and the statistical significance (p) between a given pair of variables. For ease of communication, the intensity of association follows the rules of thumb described in Hamilton (1990) (Table 3.5). Care must be taken in interpreting correlation results, as the correlation yields misleading results when the relationship between variables being considered is non-linear.

Table 3.5. Rules of thumb for interpreting correlation coefficient and coefficient of determination (modified from Hamilton, 1990)

Correlation coefficients (r)	Coefficient of determination (R^2)	Interpretation
(-)1.0	1.00	Perfect positive (negative) correlation
(-)1.0 > r >(-)0.8	1.00 > R^2 > 0.64	Strong positive (negative) correlation
(-)0.8 > r >(-)0.5	0.64 > R^2 > 0.25	Moderate positive (negative) correlation
(-)0.5 > r >(-)0.2	0.25 > R^2 > 0.04	Weak positive (negative) correlation
(-)0.2 > r >0	0.04 > R^2 > 0.00	No correlation

3.4.4 Semivariogram analysis and interpolation

Semivariogram analysis was carried out using S-PLUS (Mathsoft Inc., 1999) on all continuous data to generate spatial data at 30 m grid. This was carried out so that the spatial continuity of the different data sets could be defined. It was done on the premise that for any pair of random variables, those that are closer together are more likely to have similar values than those that are farther apart from each other (Isaaks and Srivasta, 1989) and that neither the expectation of the process nor its variogram are

location dependent (i.e. stationary). The spatial variability of the different variables is then described in terms of the perceptible distance of spatial dependence (range), process variance (sill), and the spatially independent or random error (nugget). It was done following the steps outlined below:

1. Exploratory data analysis for outliers and normality
2. Check for global trend due to any physical processes such as hydrologic flow or transport processes
3. Check for anisotropy, i.e., the influence of direction on the semivirogram in terms of range and sill
4. Empirical semivariogram construction using 20 lags and a maximum distance of 1500 m and 400 m for the Tamale and Ejura sites, respectively
5. Model semivariogram fitting to the empirical variogram. This was done using an omni-directional variogram, i.e., assuming isotropic conditions as the range and sill do not change so much with direction. Note however that, some anisotropy was recorded (Table 4.15 and 4.16).

Interpolation was done to estimate parameters over the landscape so that data could be obtained at all points over the surface. In order to create a continuous map, a weighted interpolation method (ordinary kriging) was carried out in Surfer 7 program using the modeled semivariogram with the least objective. The resulting estimated data was imported to DIGEM, saved as an ASCII file and then imported into ArcView.

3.4.5 Statistical comparison

The coefficient of variation (CV) was used to assess the variability of the different data sets, based on the following equation (Beckett and Webster, 1971):

$$CV = (s\tan dard\ deviation / mean) \bullet 100 \qquad [3.13]$$

The main limitation of CV to assess variability is the strong influence of normal distribution. Care must be taken, since most of the parameters do not possess a normal distribution.

Statistical evaluation procedures were used to test the performance of the different models and to compare them to the measured data. Methods used included coefficient of determination (R^2), which measures the degree of change between the measured and estimated values, and analysis of variance (ANOVA) to identify significant differences between the measured and estimated value and mean of different estimation models.

4 SOIL PROPERTIES, TERRAIN ATTRIBUTES AND THEIR SPATIAL DISTRIBUTION

In this chapter, the general characteristics of soil properties, terrain attributes and distribution of soil properties of the study sites are outlined. The soil properties considered here include particle size distribution (sand, silt and clay), cation exchange capacity (CEC), organic carbon, pH, bulk density and saturated hydraulic conductivity (K_s); the terrain parameters are elevation, slope gradient, aspect, Length-Slope (LS) factor, profile curvature, plan curvature, curvature, wetness index, and stream power index.

The focus of this chapter is the comparison of the soil properties, terrain attributes and distribution of soil properties in terms of their similarity and differences at the two sites. The variation of the different soil and terrain parameters is evaluated using coefficient of variation (CV). Furthermore, the relationship among the soil parameters and terrain attributes are highlighted with emphasis on the relationship between K_s and the other parameters using Pearson's correlation (r). The emphasis is placed on K_s, because it is the key soil property of concern (response variable) in this study.

The chapter is divided into five sections with the first three sections covering the general characteristics of soil properties, terrain parameters or spatial distribution of soil properties. The forth section presents relationship between soil and terrain parameters and the final section presents conclusions.

4.1 General characteristics of soil properties

This section presents the general characteristics of soil properties with their variation (CV) at each of the sites. Furthermore, the relationship among the soil properties is discussed. Finally, a comparison is made between the soil properties at the two sites. For purposes of comparison Table 4.1 gives recommended bulk density values and K_s-values for specific textural classes.

Table 4.1. Recommended bulk density and saturated hydraulic conductivity for different textural classes

Texture	Bulk density	K_s[a]	K_s[b]	K_s[c] (topsoil)	K_s[d] (subsoil)
Clay	1.20[2]	0.06	0.20		
Clay loam	1.55[1]	0.23	0.26		
Loam	1.60[2]	1.32	1.00		
Loamy sand	1.85[1]	6.11	14.58	17.1	
Sandy clay loam		0.43	0.89		1.14
Sandy loam	1.80[1,2]	2.59	4.41		1.76
Silt loam	1.65[1]	0.68	0.25		

[1]Bulk density by (Bowen 1981), [2]bulk density by Kar et al. (1992); Saturated hydraulic conductivity by [a]Rawls et al. (1982), [b]Clapp and Hornberger (1978), and [cd]Bonsu (1992)

4.1.1 Soil properties at the Tamale site

Presented in Table 4.2 are the descriptive statistics of soil properties at the Tamale site. The Tamale site has a mean texture (USDA method) of sandy loam in the topsoil and loam in the subsoil. The soil texture is dominated by sandy loam, loam, silt loam and loamy sand in the topsoil, and loam, sandy loam, clay loam and clay in the subsoil in decreasing order of distribution (see Tables 4.3 and 4.4). This shows that the texture in the subsoil is generally heavier than that in the topsoil. The range for sand, silt and clay content for both topsoil and subsoil is wide with a sharp increase in clay content from the topsoil to the subsoil, which may be due to soil translocation from the topsoil to the subsoil.

The mean organic carbon content of 0.46 % and 0.38 % for the topsoil and subsoil, respectively, is below the value of 0.5 % given by Landon (1991) as average for tropical soils. The topsoil average carbon content is less than the range (0.6-1.2 %) given by Young (1976) as desirable for tropical crop production. The low organic carbon content may be due to the fact the sampled points are in intensively cultivated fields or young fallow areas (2-5 years), where primary carbon production is low. Contributing to this low carbon content may also be the continual removal of plant material for human and animal consumption with relatively little returned to the land, respiration losses due to high temperature (Brady and Weil, 1996), and erosion losses due to high intensity rains. The lower level of organic carbon in the subsoil compared to the topsoil is in agreement with work done by Nelson et al., 1994. This is mainly the result of organic matter accumulation due to continuous plant growth.

Table 4.2. Descriptive statistics of soil properties in the topsoil and subsoil at Tamale site

Parameter	Soil level	Minimum	Maximum	Mean	Standard Deviation	CV[a] (%)	Skewness	Kurtosis
Sand (%)	Topsoil	23.0	83.6	55.8	12.9	23.1	0.03	-0.73
	Subsoil	17.9	79.6	43.8	13.6	31.1	0.18	-0.55
Clay (%)	Topsoil	0.36	37.6	7.05	5.50	78.0	2.16	6.78
	Subsoil	3.12	53.1	22.9	12.2	53.3	0.69	-0.26
Silt (%)	Topsoil	7.76	60.4	37.1	10.8	29.1	-0.17	-0.66
	Subsoil	13.6	60.3	33.3	8.08	24.2	0.29	0.45
Carbon	Topsoil	0.04	1.37	0.46	0.26	55.5	0.91	0.59
(%)	Subsoil	0.05	1.09	0.38	0.19	50.4	1.01	1.10
CEC	Topsoil	0.54	6.24	2.17	1.21	55.8	1.18	0.77
($cmol(+)kg^{-1}$)	Subsoil	0.40	6.39	2.55	1.40	55.1	0.81	-0.25
pH	Topsoil	3.73	7.34	4.87	0.51	10.5	0.88	2.22
	Subsoil	3.71	7.22	4.60	0.57	12.3	1.69	4.32
K_s	Topsoil	0.02	17.3	2.20	2.81	128	2.65	8.49
(cmh^{-1})	Subsoil	0.01	13.9	1.41	2.16	153	2.93	10.6
BD[b]	Topsoil	1.15	1.86	1.49	0.13	8.8	0.46	0.62
(gcm^{-3})	Subsoil	1.10	1.93	1.50	0.11	7.3	0.03	2.25

Topsoil sample number is 238; subsoil sample number is 183; [a]CV: coefficient of variation; BD[b]: Bulk density

The CEC in both topsoil and subsoil have wide ranges (Table 4.2). The mean CEC of 2.17 $cmol(+)kg^{-1}$ and 2.52 $cmol(+)kg^{-1}$ for the topsoil and subsoil, respectively, are low compared to the lower limit of average CEC of 5 $cmol(+)kg^{-1}$ (Landon, 1991), and the mean for ranges of 3-15 $cmol(+)kg^{-1}$ (Müller-Sämann and Kotschi, 1994) and 1-10 $cmol(+)kg^{-1}$ (Foth and Ellis, 1997), which are common for Kaolinitic clay. The main clay mineralogy in the area is Kaolinitic.

The soil pH varies from extremely acidic to neutral reaction in the topsoil (3.73 – 7.34) and subsoil (3.71 – 7.22) (Table 4.2). The mean is very strongly acid in both topsoil (4.87) and subsoil (4.60), and is below the average value of 6.5 cited by Foth and Ellis (1997) as being ideal for good availability of plant nutrients in mineral soils. The low soil pH observed may be due to the high leaching resulting from high intensity rains.

The Tamale soil had a mean K_s of moderate flow (2.20 cmh^{-1}) in the topsoil and a slow flow (1.41 cmh^{-1}) in the subsoil (see Table 3.2 for K_s classification). The mean of the measured saturated hydraulic conductivity values for the different textural classes in Tamale (Table 4.3) are within ranges given by Rawls et al. (1982), Clapp and

Hornberger (1978), and K_s obtained by Bonsu (1992) for work done in the savannah zone of northern Ghana as presented in Table 4.1.

Table 4.3. Saturated hydraulic conductivity in the topsoil and subsoil based on USDA textural classification for Tamale site

	Mean		Standard deviation		CV (%)	
Texture	Topsoil	Subsoil	Topsoil	Subsoil	Topsoil	Subsoil
Clay		0.36 (21)*		0.46		128
Clay loam	0.72 (3)	0.76 (36)	1.19	1.17	165	154
Loam	1.61 (46)	1.49 (70)	2.51	2.19	156	147
Loamy sand	4.39 (22)	4.56 (1)	4.47		102	
Sandy clay loam	0.99 (2)	1.26 (7)	1.06	1.31	107	104
Sandy loam	2.38 (133)	2.18 (43)	2.67	2.90	112	133
Silt loam	1.02 (32)	2.26 (5)	1.16	2.28	114	101

*Number of samples in parenthesis

The bulk density range of 1.15-1.89 gcm^{-3} in the topsoil and 1.10-1.93 gcm^{-3} in the subsoil are both below the critical value of 2.1 gcm^{-3}, beyond which plant growth is severely limited. With the exception of the clay texture in the subsoil, the bulk densities for the different textural classes as presented in Table 4.4 are within the limits given by Bowen (1981) and Kar et al. (1976) (see Table 4.1), beyond which root growth is impeded and crop production thus reduced.

Table 4.4. Bulk density in the topsoil and subsoil based on USDA textural classification for soils at Tamale site

	Mean		Standard deviation		CV (%)	
Texture	Topsoil	Subsoil	Topsoil	Subsoil	Topsoil	Subsoil
Clay		1.51 (21)*		0.11		7.28
Clay loam	1.57 (3)	1.52 (36)	0.14	0.10	8.92	6.58
Loam	1.45 (46)	1.51 (70)	0.11	0.12	7.59	7.95
Loamy sand	1.52 (22)	1.45 (1)	0.15		9.87	
Sandy clay loam	1.68 (2)	1.57 (7)	0.20	0.15	11.9	9.55
Sandy loam	1.50 (133)	1.48 (43)	0.14	0.11	9.33	7.43
Silt loam	1.44 (32)	1.45 (5)	0.10	0.05	6.94	3.45

* Number of samples in parenthesis

The clay content has the largest coefficient of variation (CV) amongst the different particle size distributions. The high CV of clay in the topsoil may be the result of different land use practices and erosion resulting in sediment deposition in low-lying

areas. The high CV of organic carbon is mainly due to the large coverage area of the site with very different land use types (fallow, grazing and cropping at different stages). The pH has a low CV. The K_s had the highest CV of 128 % and 153 % in the topsoil and subsoil, respectively (Table 4.2). However this is within the range of 86 – 190 % observed by Warrick and Nielson (1980) when they compared the results of measured saturated hydraulic conductivity from different authors. Expressing the K_s in terms of textural classes still shows a high CV, ranging from 102 % to 165 % in the topsoil and 104 % to 154 % in the subsoil (Table 4.3). Due to this high variability of K_s, its accuracy is best determined in terms of order of magnitude (Tieje and Hennings, 1996 and McKenzie et al, 2000). In general, bulk density exhibited the lowest CV and was within ranges cited by Warrick and Nielsen (1980).

Table 4.5 presents the Pearson's correlation among the normalized Tamale soil data. There are significant cross correlations among most of the soil parameters. Focusing on K_s (i.e. the dependent variable for further analysis), significant correlation at the 0.01-level was observed between K_s and sand (0.48), clay (-0.42), CEC (0.31) silt (-0.22), and bulk density (-0.13) in decreasing order. The correlation was positive for sand and CEC and negative for the other soil properties. Significant correlation are also observed amongst other parameters such as sand and clay (-0.70), sand and silt (-0.66), and clay and CEC (0.38) at the 0.01-level.

Table 4.5. Correlation coefficient with significance level for point soil data from Tamale site

Parameter	Ks [L]	Sand	Clay	Silt	CEC	Carbon	Bulk density
Sand[‡]	0.48**						
Clay[†]	-0.42**	-0.70**					
Silt [L]	-0.22**	-0.66**	-0.5				
CEC[†]	0.31**	-0.34**	0.38**	0.08			
Carbon[†]	-0.08	-0.19**	0.11*	0.15**	0.18**		
BD[‡]	-0.13**	0.08	0.11*	-0.19**	0.16**	0.05	
pH [L]	0.00	0.08	-0.05	-0.06	0.29**	0.14	0.11*

[L] Log, [‡] square and [†]square root transformed parameter
** Correlation is significant at the 0.01 level (2-tailed)
*Correlation is significant at the 0.05 level (2-tailed)

4.1.2 Soil properties at the Ejura site

The soils at Ejura have a high sand content with mean values of about 72 % in the topsoil and 69 % in the subsoil, with a minimum value of about 40 % in both soil levels

(Table 4.6). Like the sand content, the clay content, though high, does not change so much from the topsoil (23 %) to the subsoil (21 %), indicating a similar soil textural condition within the 0 – 15 cm and 30 – 45 cm soil depth. The soil has a mean texture of sandy clay loam at both soil levels. Soils at this site are dominated by sandy loam, loamy sand, and sandy texture in the topsoil and subsoil in decreasing order of distribution (Tables 4.7 and 4.8).

Table 4.6. Descriptive statistics of soil properties in the topsoil and subsoil at Ejura site

Parameter	Soil level	Minimum	Maximum	Mean	Standard Deviation	CV[a] (%)	Skewness	Kurtosis
Sand (%)	Topsoil	40.4	90.8	72.5	7.16	9.9	-0.39	0.59
	Subsoil	40.4	89.0	69.0	8.57	12.4	0.14	-0.13
Clay (%)	Topsoil	7.56	46.5	23.0	5.68	24.7	0.18	0.58
	Subsoil	7.00	47.1	21.1	5.89	28.0	0.22	0.40
Silt (%)	Topsoil	0.08	18.8	4.49	2.88	64.1	1.67	4.27
	Subsoil	0.56	22.0	9.94	4.83	48.6	0.19	-0.31
Carbon	Topsoil	0.11	1.87	0.73	0.36	48.7	0.95	0.48
(%)	Subsoil	0.11	1.46	0.64	0.27	42.0	0.71	0.27
CEC	Topsoil	0.46	4.41	2.72	0.82	30.3	-0.42	-0.25
($cmol(+)kg^{-1}$)	Subsoil	0.48	4.34	2.66	0.77	28.9	-0.12	-0.36
pH	Topsoil	3.68	7.50	5.31	0.64	12.0	0.80	1.36
	Subsoil	3.70	7.52	4.74	0.82	17.3	1.14	0.97
K_s	Topsoil	0.06	61.4	13.4	14.1	105	1.41	1.18
(cmh^{-1})	Subsoil	0.06	42.6	4.4	6.88	157	3.16	11.26
BD[b]	Topsoil	1.02	1.74	1.41	0.13	9.1	-0.16	-0.09
(gcm^{-3})	Subsoil	1.09	1.99	1.58	0.11	7.0	-0.80	2.20

Topsoil sample number is 238; subsoil sample number is 183; [a]CV: coefficient of variation; BD[b]: Bulk density

The carbon content gave mean values of 0.73 % and 0.64 % in the topsoil and subsoil, which is average for tropical soils (Landon, 1991).

The mean CEC values for topsoil (2.72 $cmol(+)kg^{-1}$) and subsoil (2.66 $cmol(+)kg^{-1}$) are low (Landon, 1991). Even the maximum CEC for the topsoil (4.41 $cmol(+)kg^{-1}$) and subsoil (4.34 $cmol(+)kg^{-1}$) were lower than the lower limit of 5 $cmol(+)kg^{-1}$ given by Landon (1991) for average CEC in tropical soils.

The soil pH varied from extremely acid (3.68 (topsoil) and 3.70 (subsoil)) to mildly alkaline (7.50 (topsoil) and 7.52 (subsoil)). The average pH in the topsoil is

strongly acid (5.31) and very strongly acid (4.74) in the subsoil. This is not ideal for plant growth (Foth and Ellis, 1997).

The Ejura soils, on average, have a high flow characteristic with a very rapid (13.4 cmh^{-1}) and moderate (4.38 cmh^{-1}) flow in the topsoil and subsoil, respectively. The K_s-values as presented in Table 4.7 are comparable to those given by Rawls et al. (1982), Clapp and Hornberger (1978), and Bonsu (1992) (see Table 4.1). For example the very rapid (16.2 cm^{-1}) for topsoil and moderately rapid flow (5.64 cmh^{-1}) for subsoil loamy sand are comparable to that of moderately rapid (6.11 cmh^{-1}) reported by Rawls et al., (1982), very rapid (14.58 cmh^{-1}) by Clapp and Hornberger (1978), and moderately rapid (7.13 cmh^{-1}) for topsoil by Bonsu (1992).

Table 4.7. Saturated hydraulic conductivity in the topsoil and subsoil based on USDA textural classification for Ejura site

	Mean		Standard deviation		CV (%)	
Texture	Topsoil	Subsoil	Topsoil	Subsoil	Topsoil	Subsoil
Loam	3.0 (1)	0.97 (4)*		1.6		165
Loamy sand	16.2 (152)	5.64 (58)	14.8	6.3	91.7	112
Sand	25.4 (9)	8.45 (13)	20.6	9.4	81.3	112
Sandy clay loam		0.89 (10)		1.2		130
Sandy loam	10.9 (212)	4.06 (246)	12.6	6.9	115	170

* Number of samples in parenthesis

The bulk density of the topsoil (1.41 gcm^{-3}) is lower than that of the subsoil (1.58 gcm^{-3}). The maximum bulk density of 1.99 gcm^{-3} for the subsoil is close to the limit that is not conducive for cultivation (Table 4.6). Table 4.8 presents the bulk density classified according to the textural classes (USDA). It gives a mean bulk density in the range of 1.37 gcm^{-3} (loam) to 1.46 gcm^{-3} (sand) in the topsoil and 1.55 gcm^{-3} (sand) to 1.69 gcm^{-3} (sandy clay loam) in the subsoil, thus indicating a more compact subsoil condition.

The silt content shows the highest CV among the different particle size distributions followed by the clay content, with sand content having the lowest (Table 4.6). Organic carbon has a CV of 49 % and 42 % in the topsoil and subsoil, respectively. CEC has 30 % (topsoil) and 29 % (subsoil) CV. The pH has a CV of 12 % and 17 % in the topsoil and subsoil, respectively. The K_s has a high CV in the topsoil (105 %) and subsoil (157 %). The comparatively lower CV for K_s in the topsoil

compared to the subsoil may be the result of land use activities resulting in mixing of the soil. Grouping the saturated hydraulic conductivity data based on textural classes (USDA) still gives a high CV (81 – 115 % (topsoil) and 112 – 170 % (subsoil)) as presented in Table 4.7. The CV of less than 10 % at both soil depths for bulk density is similar to that given by Warrick and Nielson (1980).

Table 4.8. Bulk density in the topsoil and subsoil based on USDA textural classification for soils at Ejura site

	Mean		Standard deviation		CV (%)	
Texture	Topsoil	Subsoil	Topsoil	Subsoil	Topsoil	Subsoil
Loam	1.37 (1)	1.64 (4)		0.07		4.27
Loamy sand	1.41 (152)	1.57 (58)	0.14	0.10	9.93	6.37
Sand	1.46 (9)	1.55 (13)	0.10	0.10	6.85	6.45
Sandy clay loam		1.69 (10)		0.07		4.14
Sandy loam	1.41 (212)	1.58 (246)	0.12	0.11	8.51	6.96

* Number of samples in parenthesis

Table 4.9. Correlation coefficient with significance level for point soil data from Ejura site

Parameter	Ks[L]	Sand	Clay	Silt	CEC	Carbon	Bulk density
Sand[‡]	0.33**						
Clay[†]	-0.12**	-0.81**					
Silt[L]	-0.40**	-0.69**	0.20**				
CEC[†]	-0.00	-0.08*	0.11**	0.05			
Carbon[†]	0.10**	-0.03	0.08*	-0.08*	0.12**		
BD[‡]	-0.53**	-0.14**	-0.16**	0.38**	-0.07	-0.05	
pH[L]	-0.27**	0.33**	-0.14**	-0.38**	0.01	0.24**	-0.21**

[L] Log, [‡] square and [†]square root transformed parameter
** Correlation is significant at the 0.01 level (2-tailed)
*Correlation is significant at the 0.05 level (2-tailed)

Pearson's correlation among the normalized Ejura soil data is presented in Table 4.9. A significant correlation at the 0.01-level can be observed between saturated hydraulic conductivity and bulk density (-0.53), silt (-0.40), sand (0.33), pH (-0.27), clay (-0.12), and carbon (0.10) in decreasing order. Significant correlation was also observed amongst other parameters such as sand and silt (-0.69), and pH and sand (0.33) at the 0.01-level.

4.1.3 Comparison of soil data from Tamale and Ejura

The results presented from the two sites indicate that the soils at the two sites vary over space irrespective of the topography of the area. The soils at the Ejura site have a comparatively high sand content in both soil zones (0-15 cm and 30-45 cm). Consequently, they exhibit a more uniform texture in both soil zones compared to the Tamale soils, which show a drastic increase in clay content from topsoil (7 %) to subsoil (23 %).

The mean organic carbon content at the Ejura site is average compared to the low content at Tamale. The higher carbon content in the Ejura soils is mainly due to the high carbon content observed along the river (see Figure 4.6 (C)). The CEC is low at both sites and soil levels with a mean of less or equal 2.7 cmol(+)kg^{-1}. Mean soil pH had a strong to very strong acid reaction, which is not ideal for crop production as it may affect nutrient availability.

Though the maximum bulk density at both sites and soil depths exceeded the ideal value of 1.6 gcm^{3} for crop production, the average value of 1.6 gcm^{-3} or less is good. The flow characteristics at the two sites were very different, with moderate (topsoil) to slow flow (subsoil) at Tamale compared to very rapid (topsoil) to moderate flow (subsoil) at the Ejura site. Notwithstanding this mean flow characteristic, the flow obtained for different soil texture and soil types varies in accordance with these properties (see Sections 4.1.1 and 4.1.2). The high variation in saturated hydraulic conductivity based on soil properties for the different sites and soil depths is illustrated in Figures 4.5 and 4.6.

Bulk density gave the lowest CV and K_s the highest CV at both sites. The CV for clay at Tamale is highest among the particle size distribution values, while at Ejura, silt gives the highest CV.

4.1.4 Combined soil data from Tamale and Ejura

For certain analyses described in Chapters 6 and 8, the data from Tamale and Ejura were combined and the descriptive statistics for the combined soil parameters of sand, clay, silt, organic carbon, CEC, pH, K_s, and bulk density are presented in Table 4.10. The parameters were transformed (normalized) based on the steps outlined in Section 3.4.2.

Table 4.10. Descriptive statistics of combined soil properties for Tamale and Ejura sites

Soil Parameter	Min.	Max.	Mean	Std. Dev.	CV (%)	Skew-ness	Kurtosis	Transformed Skew-ness	Transformed Kurtosis
Sand‡	17.9	90.8	63.3	14.7	23.2	-0.83	0.18	-0.28	-0.59
Clay†	0.36	53.1	19.0	9.52	50.0	0.24	0.31	-0.59	0.06
Silt L	0.08	60.4	17.4	15.5	87.6	0.79	-0.66	-0.45	-0.35
Carbon†	0.04	1.87	0.59	0.32	53.9	1.04	1.14	0.29	-0.01
CEC†	0.40	6.39	2.56	1.04	40.5	0.44	0.19	-0.11	-0.31
pH L	3.68	7.52	4.93	0.72	14.6	0.83	0.91	0.42	-0.02
Ks L	0.01	61.4	6.43	10.4	161	2.65	7.25	-0.31	-0.31
BD‡	1.02	1.99	1.49	0.14	9.17	-0.22	0.19	0.07	0.18

L Log, ‡ square and †square root transformed parameter
Total number of sample is 1126

Table 4.11 presents the correlation among the normalized combined (Tamale and Ejura) soil data. Significant cross correlations (Pearson) among most of the soil parameters can be observed. However, the discussion here is centered on the correlation between saturated hydraulic conductivity and the other soil parameters, since saturated hydraulic conductivity is the main focus of this work.

Table 4.11. Correlation coefficient with significance level for the combined point soil data from Tamale and Ejura sites

Parameter	*Ks L*	*Sand*	*Clay*	*Silt*	*CEC*	*Carbon*	*Bulk density*
Sand‡	0.55**						
Clay†	0.01	-0.14**					
Silt L	0.53**	-0.80**	-0.36**				
CEC†	-0.03	-0.02	0.35**	-0.14**			
Carbon†	0.21**	0.22**	0.28**	-0.36**	0.22**		
BD‡	-0.36**	-0.04	-0.01	0.17**	0.03	-0.02	
pH L	0.25**	0.29**	0.02	-0.34**	0.15**	0.26**	-0.12**

L Log, ‡ square and †square root transformed parameter. ** Correlation is significant at the 0.01 level (2-tailed); *correlation is significant at the 0.05 level (2-tailed)

Significant correlation at 0.01-level can be observed between saturated hydraulic conductivity and the other soil parameters in decreasing order: sand (0.55), silt (-0.53), bulk density (-0.36), pH (0.25) and organic carbon content (0.21) (Table 4.11). The correlation was positive for sand, pH, and organic carbon and negative for the other parameters. Significant correlation can also be observed amongst other parameters such as sand and silt (-0.80), sand and organic carbon content (0.22), and sand and pH (0.29)

at the 0.01-level. This cross correlation can lead to severe multicollinearity problems in multiregressional analysis, if the variables are used together in a model to explain a phenomenon (Hair et al., 1998).

4.2 Characteristics of terrain attributes at Tamale and Ejura sites

The semivirogram parameters for the DEMs at Tamale and Ejura (Figures 4.1 and 4.2) are presented in Table 4.12. The DEM for the Tamale site shows non-stationarity for the distance considered, as it does not sill at a range. The figures illustrate the virtually flat undulating surface at Tamale and the gentle to moderate slopes at Ejura.

Table 4.12. Semivariogram model parameters for DEM at Tamale and Ejura sites

Site	Anisotropy	Model	Range	Sill	Nugget	Slope	Objective
Tamale	[1]NA	Linear	-	-	0.0	0.052	807
Ejura	NA	Linear	-	-	49.8	118	86.3

[1]NA: No anisotropy

Tables 4.13 and 4.14 present the descriptive statistics for terrain attributes from a 30 m × 30 m grid for the Tamale and Ejura sites, generated as outlined in section 3.3. With the exception of elevation, all other attributes have medium to extremely large (> 10 %) CVs. Elevation was normally distributed and therefore there was no need for transformation. All remaining parameters were normally transformed as indicated in Tables 4.13 and 4.14, with the exception of the three curvature parameters (profile curvature, plan curvature, and curvature) for which appropriate normalization functions could not be found despite the number of trials, as they contain both negative and positive interval data. Each terrain attribute is interpolated using the appropriate model semivariogram method to visualize the distribution of its spatial pattern. To facilitate interpretation and communication terrain maps are presented using the untransformed data and re-aligned to enhance visualization.

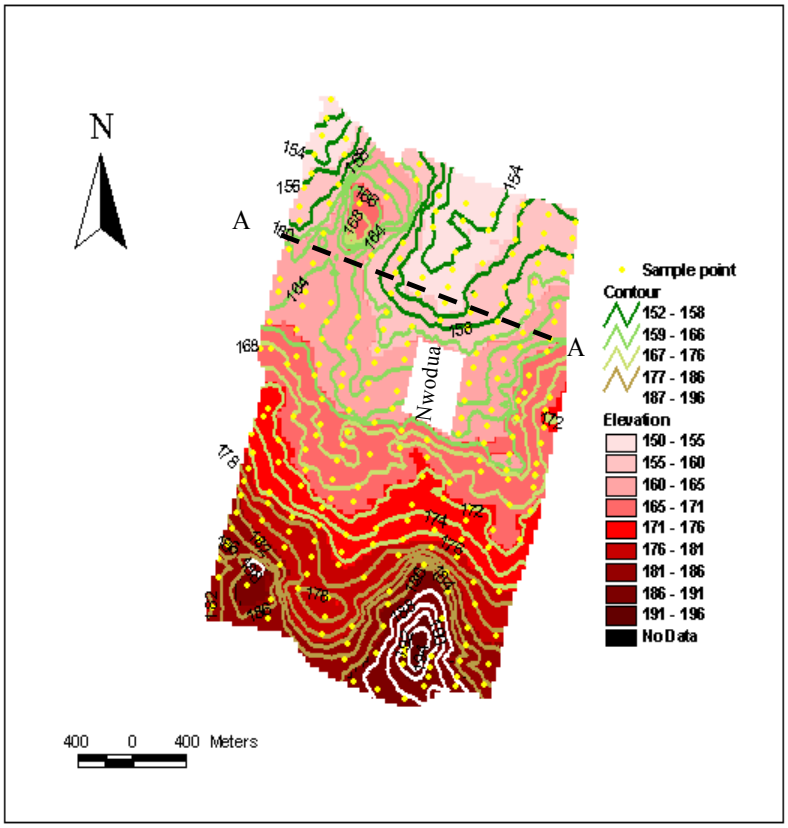


Figure 4.1. Digital elevation model (DEM) of Tamale site in the Volta Basin of Ghana

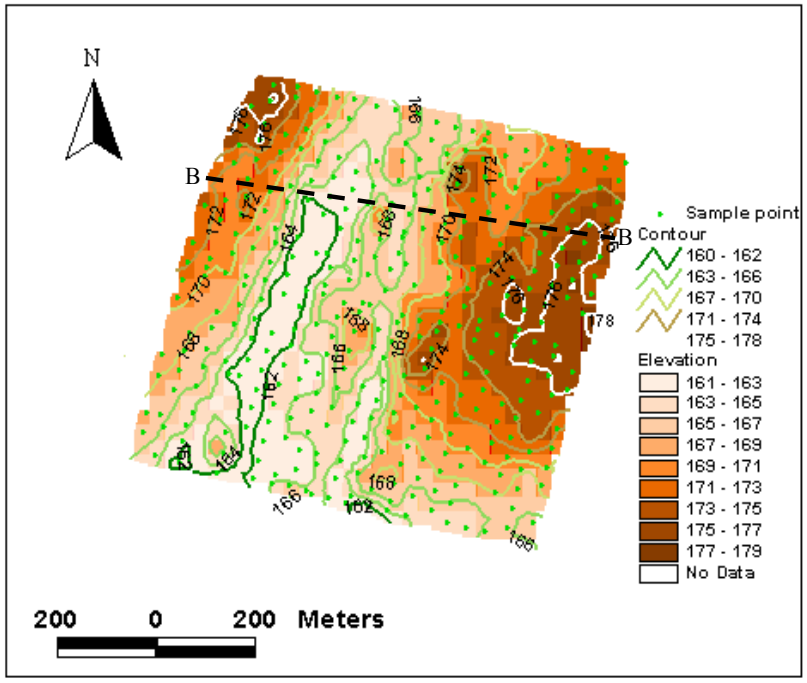


Figure 4.2. Digital elevation model (DEM) of Ejura site in the Volta Basin of Ghana

4.2.1 Spatial distribution of terrain attributes: Tamale pilot site

Figure 4.3 illustrates the terrain parameters for the Tamale site, which have been re-aligned at 180° to enhance visualization. The re-alignment means that the maximum aspect of 240-360 ° is obtained in the direction facing west (Figure 4.3(A)). Aspect has a significant correlation with only two terrain parameters (elevation (-0.13) and slope (-0.17)) (see Table 4.17a).

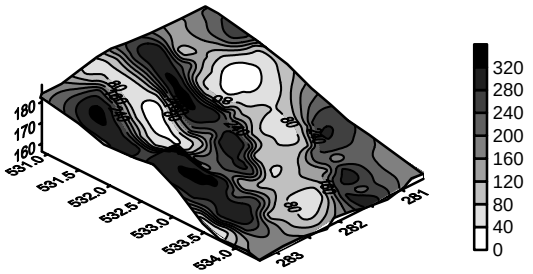
The Tamale site is virtually flat with a mean slope gradient of 1.08°, varying from 0.00° to 3.10° (Table 4.13). The slope gradient correlates significantly with elevation, wetness index, upslope area, stream power index, length-slope (LS) factor, and aspect. Figure 4.3(B) shows the distribution of the slope gradient. The slope is highest on the shoulder and backslope, and intermediate on the footslope and toeslope.

Taking the re-alignment into consideration, the LS factor is lowest in the lower elevation area in the north and increases southwards to the higher elevation area (Figure 4.3(C). It has mean value of 0.63.

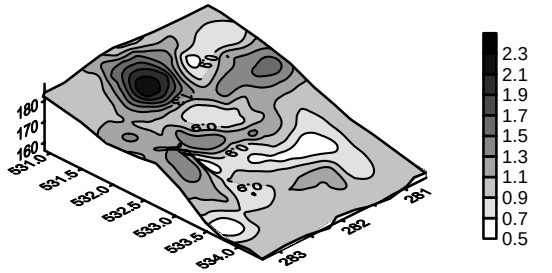
Profile curvature, plan curvature and curvature are dimensionless parameters that indicate convexity (positive value) and concavity (negative value) and its magnitude on the slope. The plan curvature is convex over the entire terrain, varying from 0.018-0.046 (Figure 4.3(E)). The profile curvature has both convex and concave slopes (Figure 4.3(D)). The convex slopes are mainly on the interfluve, shoulder and backslope. All three curvature parameters show no significant relationship with elevation, indicating that the land is virtually flat or no particular shape (convex or concave) is dominant at high or low elevations. This lack of correlation may also be due to limited sampling points, thus resulting in a ‘smoothen’ surface curvature pattern. Unlike the profile curvature, plan curvature and curvature had convex slope shapes at the foot and toeslopes (Figure 4.3(E) and (F)).

The stream power index exhibits some ‘noisy’ patterns around certain areas, which may be due to limited data points (Figure 4.3(G)). It correlates significantly with curvature parameters, wetness index, upslope area, slope and LS factor (Table 4.17a). A higher index is observed closer to the stream channel and decreases away from it.

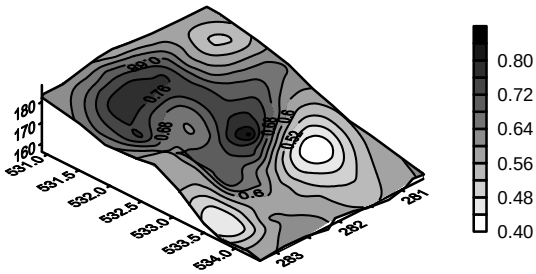
(A) Slope aspect (degree)



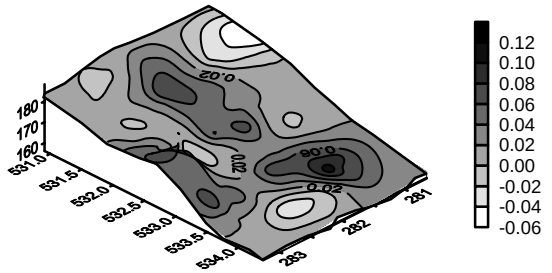
(B) Slope gradient (degree)



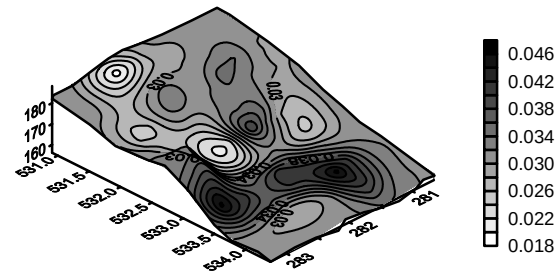
(C) Length-slope factor (dimensionless)



(D) Profile curvature (10^{-3} dimensionless)



(E) Plan curvature (10^{-3} dimensionless)



(F) Curvature (10^{-3} dimensionless)

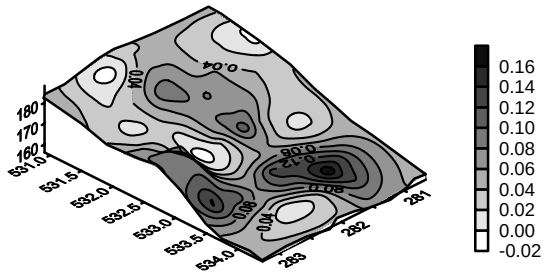
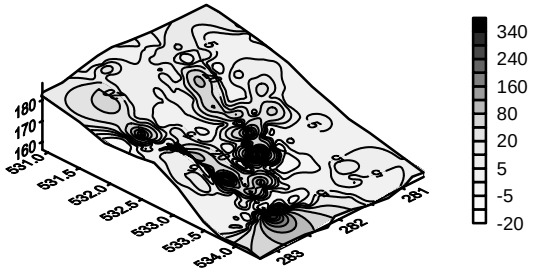
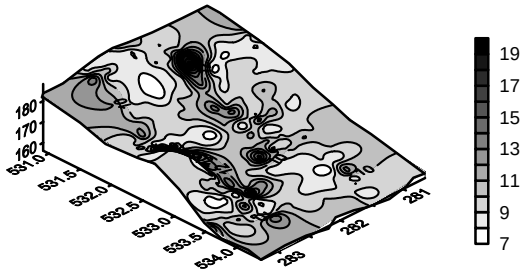


Figure 4.3a. Spatial distribution of terrain parameters at Tamale site: (A) Aspect; (B) Slope gradient; (C) Length-slope factor; (D) Profile curvature; (E) Plan curvature and (F) Curvature

(G) Stream power index (dimensionless)



(H) Wetness index (dimensionless)



(I) Upslope area (m^2)

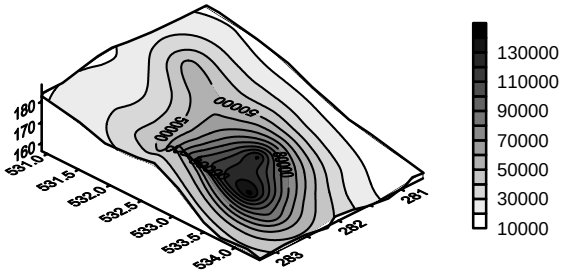


Figure 4.3b. Spatial distribution of terrain parameters at Tamale site: (G) Stream power index; (H) Wetness index and (I) Upslope contributing area

The wetness index is high along the stream channel and toeslope and lowest on high elevation slopes (Figure 4.3(H)). With a mean value of about 10, this dimensionless parameter varies from about 7-20. The low values are observed on the interfluve, shoulder and backslope. Like the stream power index, it has a significant relationship with the curvature parameters, upslope area, stream power index, slope, and LS factor. Also, it has a negative significant correlation with elevation, thus indicating low values at high elevations (Table 4.17a).

The upslope areas have smoothened patterns that may be the result of limited data points (Figure 4.3(I)). These areas have a high value around Nwodua (see Figure 4.1), which decreases radially outward. With the exception of aspect, it has a significant correlation (at 0.05-level or better) with all other terrain parameters.

Table 4.13. Descriptive statistics of terrain parameters (using 30 m × 30 m grid) at Tamale site

Terrain Parameter	Min.	Max.	Mean	Std. Dev.	CV(%)	Skew-ness	Kur-tosis	Transformed Skew-ness	Transformed Kur-tosis
Profile curvature	-0.0005	0.0010	0.00003	0.00021	815	1.84	1.81		
Plan curvature	-0.0008	0.0012	0.00003	0.00020	624	1.81	11.49		
Curvature	-0.0012	0.0022	0.00006	0.00038	660	2.16	9.96		
Elevation	152	193	169	10.4	6.15	0.41	-0.81		
Wetness index [L]	6.98	20.26	9.96	1.96	19.6	1.76	4.61	1.03	1.55
Upslope area [L]	900.00	2E+06	53557	193524	316	7.09	59.5	1.01	1.15
Stream P. index[L]	0.02	388.06	20.57	56.67	275	4.84	25.3	0.14	0.95
Slope[†]	0.00	3.11	1.08	0.01	52.9	0.85	0.10	-0.31	1.02
L-S factor[†]	0.00	5.62	0.63	0.60	95.6	3.58	22.2	0.97	3.01
Aspect[†]	1.98	360	170	2.31	77.7	0.18	-1.71	-0.10	-1.56

[L] Log, [‡] square and [†]square root transformed parameter
Total number of sample is 238

4.2.2 Spatial distribution of terrain attributes: Ejura pilot site

The interpolated maps for terrain parameters at the Ejura site were re-aligned 10° anti-clockwise to enhance visualization (Figure 4.4).

The aspect varied from 0-360° with the high values occurring on slopes facing west (Figure 4.4(A)). It had no significant correlation with any of the other terrain parameters (see Table 4.18a).

The slope gradient is highest on the backslope followed by the shoulder and toeslope (Figure 4.4(B)). It has values in the range of 0.00-10.71° with a mean of 2.66° (Table 4.14). The slope gradient has significant correlation (0.01-level) with LS factor, wetness index, stream power index, and upslope area – in order of decreasing coefficient of correlation – and significant negative correlation with elevation (-0.11). Thus the slope decreases with increasing elevation.

The LS factor has high values on the backslope, similar to the slope gradient parameter (Figure 4.4(C)). As indicated earlier, the LS factor has a strong positive correlation with slope gradient (r = 0.79) (see Table 4.18a and Figures 4.4(B) and 4.4(C)). The LS factor also has significant negative correlation (0.01-level) with profile

curvature, plan curvature, curvature, and elevation, and a positive correlation (0.01-level) with upslope area and stream power index.

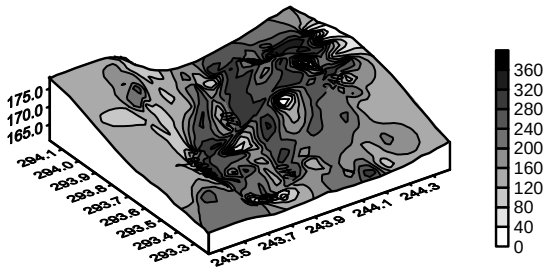
Table 4.14. Descriptive statistics of terrain parameters (using 30 m × 30 m grid) at Ejura site

Terrain Parameter	Min.	Max.	Mean	Std. Dev.	CV (%)	Skew-ness	Kur-tosis	Transformed Skew-ness	Transformed Kur-tosis
Profile curvature	-0.0049	0.0039	-0.00006	0.0011	1977	-0.20	3.02		
Plan curvature	-0.0023	0.0039	0.00008	0.0008	988	0.81	3.03		
Curvature	-0.0069	0.0078	0.00002	0.0017	7208	0.58	3.14		
Elevation	161	178	169	4.64	3	0.15	-1.15		
Wetness index [L]	5.16	19.79	8.19	1.76	21	2.41	10.99	1.12	2.70
Upslope area [L]	900.00	374107	13878	40786	294	6.24	43.76	1.09	0.78
Stream P. index[L]	0.06	377.47	15.00	36.59	244	6.12	45.87	0.29	0.26
Slope[†]	0.00	10.71	2.66	0.03	60	1.41	3.52	0.20	0.86
L-S factor[†]	0.00	7.31	1.40	1.25	89	1.88	3.91	0.73	0.61
Aspect[†]	0.00	360	193	1.58	47	-0.13	-1.05	-0.81	0.41

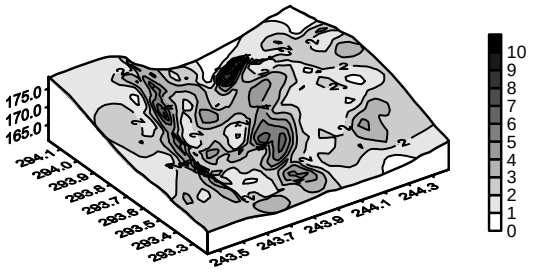
[L] Log, [‡] square and [†]square root transformed parameter
Total number of sample is 374

The profile curvature is predominantly convex on the interfluve, shoulder and back slope – in order of decreasing magnitude – and changes to a concave slope on the footslope, toeslope, and channel – in order of increasing magnitude (Figure 4.4(D)). With the exception of slope and aspect, the profile curvature has a significant correlation with all other terrain parameters (see Table 4.18). The correlation with elevation is positive, which is an indication of increasing profile curvature with elevation. The plan curvature is predominantly convex with the exception of a small area along the channel (Figure 4.4(E)). Like the profile curvature, it has significant positive correlation (0.27) with elevation, indicating that it also increases with elevation. Plan curvature had a significant correlation with profile curvature, curvature, wetness index, upslope area, stream power index, and LS factor. With a strong positive correlation of 0.92 with profile curvature, curvature behaves similarly to the profile curvature (see Table 4.18a). None of the curvature parameters correlate with slope gradient.

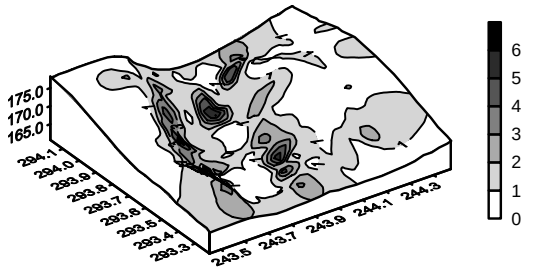
(A) Slope aspect (degree)



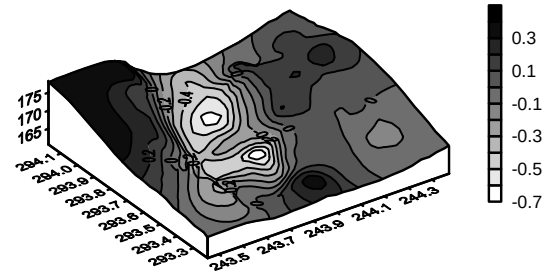
(B) Slope gradient (degree)



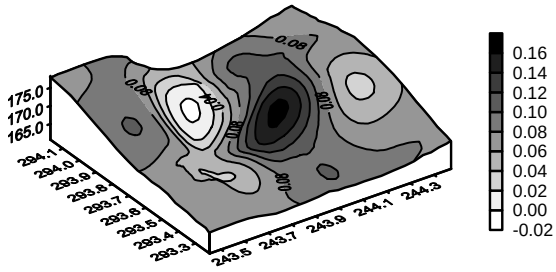
(C) Length-slope factor (dimensionless)



(D) Profile curvature (10^{-3} dimensionless)



(E) Plan curvature (10^{-3} dimensionless)



(F) Curvature (10^{-3} dimensionless)

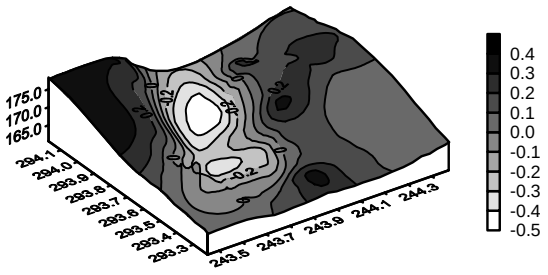
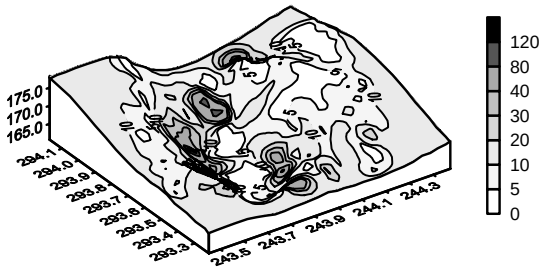
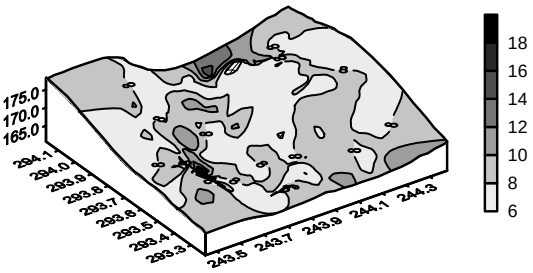


Figure 4.4a. Spatial distribution of terrain parameters at Ejura site: (A) Aspect; (B) Slope gradient; (C) Length-slope factor; (D) Profile curvature; (E) Plan curvature and (F) Curvature

(G) Stream power index (dimensionless)



(H) Wetness index (dimensionless)



(I) Upslope area (m^2)

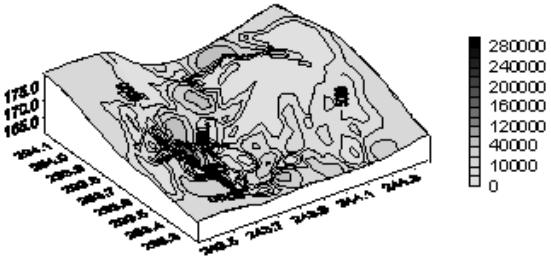


Figure 4.4b. Spatial distribution of terrain parameters at Ejura site: (G) Stream power index; (H) Wetness index and (I) Upslope contributing area

As a result of high variation (0.06-377) (see Table 4.14) with a mean of 15, the stream power index is represented using an unequal interval (Figure 4.4(G)). The index increases from the interfluves towards the channel (Figure 4.4(G)). This is explained by a significant positive correlation (0.40) of stream power index with slope gradient. The stream power decreases with increasing elevation as indicated by the negative significant correlation with elevation (-0.44) (Table 4.18a).

Locations with a high wetness index are found along the toeslope, with all other places having a more or less evenly distributed wetness index (Figure 4.4(H)). The wetness index has a strong positive correlation (0.89) with upslope area. It has significant negative correlation (-0.57) with slope, as it decreases with increasing slope.

The upslope contribution area decreases with elevation (-0.46) (Figure 4.4(I)).

4.2.3 Comparison of terrain attributes distribution at the Tamale and Ejura sites

The terrain attributes elevation, aspect, slope, profile curvature, plan curvature, curvature, wetness index, upslope area, length-slope (LS) factor were evaluated for their spatial distribution and their relationship with saturated hydraulic conductivity. The two sites have about the same mean elevation (169 m), with the Tamale site having a higher range and covering a larger area compared to the Ejura site. The Tamale site has gentle slopes compared to the moderately steep shoulder to footslope at the Ejura site (Figures 4.1, 4.2, 4.3(B) and 4.4(B)). The mean aspect at the Ejura site (193°) is higher compared to that at the Tamale site (170°). Mean LS factor at Ejura (1.4) is more than twice that of Tamale (0.6) indicating higher sediment transport capacity at the Ejura site. Curvature parameters at Tamale are predominantly convex with few portions having a concave profile curvature. At the Ejura site, the curvature parameters are mainly concave slopes along the toe and footslopes and convex slopes at the backslope and shoulder slope (Figures 4.3(D-F) and 4.4(D-F)). The stream power index exhibited wide variations at the two sites, ranging from low values at the interfluve to high values at the channel. The range of variation of stream power index is higher at the Tamale site compared to that at Ejura. Thus, the erosive power of flowing water is higher at Tamale mainly due to the larger specific catchment area at that site. The range of variation and mean of wetness index at the two sites were about the same. Upslope contributing area at both sites decreases from the interfluve to the channel.

The terrain parameters had in general a poor relationship with saturated hydraulic conductivity (K_s) (see Tables 4.17 and 4.18). The relationships were also not consistent. For instance, the LS factor had a mixed relationship with a significant negative correlation with topsoil K_s at Tamale (-0.14) and positive correlation with topsoil K_s at Ejura (0.19). The other parameters such as elevation, upslope area, stream power index, slope and aspect show some relationship at one site at a particular level but not at the other. No relationship was observed for the curvature parameters with saturated hydraulic conductivity. These poor and mixed relationships may have a negative impact on the use of terrain parameters for estimating K_s. However, significant relationships are observed between some of the terrain parameters and between some terrain and other soil properties.

4.3 Spatial distribution of soil properties

The spatial soil parameters were generated using DGPS points and geostatistical interpolation tools as outlined in section 3.3.

The following sub-sections present the spatial interpolation and distribution of soil parameters at the two pilot sites. First, the interpolation parameters are presented, followed by the subsections on the spatial distribution of the soil properties. With the exception of K_s, which is mapped using the logarithmic values because of its large range, for all other parameters the un-transformed data is used for ease of visualization and explanation. For all statistical analyses, with the exception of descriptive statistics that may be referred to, the appropriate transformed data is used for all soil parameters. Each parameter is presented twice, illustrating the topsoil (0-15 cm) and subsoil (30-45 cm), symbolized by (1) and (2), respectively – e.g. Figure 4.5(A1) for topsoil K_s and Figure 4.5(A2) for subsoil K_s.

4.3.1 Spatial interpolation of soil properties

After investigating for anisotropy and determining the empirical semivariogram using the normalized data, the suitable model semi-variogram with the least objective was selected for interpolation of the data. For example, the spherical model semivariogram was selected for interpolation of K_s of the Tamale subsoil. Tables 4.15 and 4.16 present the selected semivariogram models for soil data at Tamale and Ejura sites, respectively.

The pH (topsoil) at Tamale site and the silt content (topsoil) and sand content (subsoil) at Ejura site show spatial non-stationarity for the distance being considered, as it does not sill at a range. Sand has the highest nugget for the Tamale soil and the subsoil of Ejura, which may be due to measurement error and fine-scale variability (i.e. variation at spatial scales too fine to detect or too fine based on the sampling grid size).

Spatial correlation (range) of soil properties varied from 161 m to 1209 m (topsoil), and 185 m to 2100 m (subsoil) for the Tamale and 82 m – 325 m (topsoil), and 115 m – 555 m (subsoil) for the Ejura site. Beyond these ranges there is no spatial dependency. The higher range or spatial dependency in the subsoil compared to the topsoil is an indication of minimal soil disturbance in the subsoil. The higher range in the subsoil indicates a longer distance over which properties such as saturated hydraulic conductivity are closely related compared to the topsoil. Saturated hydraulic

conductivity, which is the main parameter for this study, showed a spatial dependency of 475 m (topsoil), and 520 m (subsoil) at the Tamale pilot site; and 82 m (topsoil), and 225 m (subsoil) at the Ejura pilot site (Tables 4.15 and 4.16).

Table 4.15. Semivariogram model parameters for topsoil and subsoil properties at Tamale

Parameter	Aniso-tropy	Model	Range	Sill	Nugget	Slope	Objec-tive
			Topsoil				
Sand‡	[1]NA	Exponential	161	1.9E+07	1E+05		8.3E+10
Clay†	NA	Gaussian	155	0.746	0.000		0.046
Silt[L]	NA	Spherical	592	0.011	0.009		0.000
Carbon†	NA	Spherical	309	0.032	0.000		0.000
CEC†	NA	Exponential	1209	0.122	0.049		0.006
pH[L]	NA	Power	0.05		0.000	0.001	0.000
Ks[L]	NA	Spherical	475	0.132	0.198		0.003
BD‡	NA	Spherical	595	0.042	0.098		0.001
			Subsoil				
Sand‡	NA	Exponential	185	1.3E+06	751082		1.7E+12
Clay†	NA	Spherical	972	0.781	0.791		0.238
Silt[L]	0°(2), 135°(2)	Spherical	743	0.003	0.009		0.000
Carbon†	90° , 135	Spherical	309	0.006	0.016		0.000
CEC†	NA	Exponential	61.2	0.136	0.055		0.002
pH[L]	NA	Spherical	2100	0.001	0.002		0.000
Ks[L*]	NA	Spherical	519	0.158	0.300		0.013
BD‡	NA	Spherical	637	0.036	0.065		0.000

[L] Log transformed, ‡ square transformed and †square root transformed parameter
* 15 lags, [1]NA: No anisotropy

Table 4.16. Semivariogram model parameters for topsoil and subsoil properties at Ejura

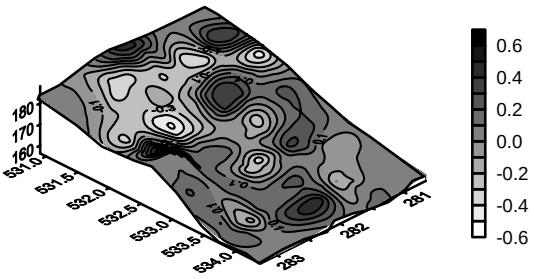
Parameter	Aniso-tropy	Model	Range	Sill	Nugget	Slope	Objec-tive
			Topsoil				
Sand‡	[1]NA	Gaussian	283	27.460	5.120		11.920
Clay†	45°, 90° (1.25, 2)	Exponential	325	0.310	0.130		0.001
Silt[L]	NA	Power	0.52		0.064	0.001	0.000
Carbon†	NA	Spherical	303	0.027	0.027		0.000
CEC†	NA	Spherical	117	0.029	0.043		0.000
pH[L]	NA	Spherical	158	0.001	0.002		0.000
Ks[L]	NA	Gaussian	82.4	0.040	0.301		0.002
BD‡	NA	Spherical	171	0.037	0.089		0.000
			Subsoil				
Sand‡	NA	Linear			636492	2130	1.9E+10
Clay†	0°(2), 90°, 135°(2)	Spherical	333	0.190	0.196		0.002
Silt[L]	NA	Spherical	335	0.044	0.046		0.000
Carbon†	0°, 135°(2)	Spherical	376	0.010	0.016		0.000
CEC†	45°, 90°, 135°	Spherical	133	0.027	0.034		0.000
pH[L]	90°, 135°(2)	Spherical	555	0.001	0.003		0.000
Ks[L*]	NA	Spherical	225	0.057	0.360		0.003
BD‡	NA	Spherical	116	0.042	0.065		0.002

[L] Log transformed, ‡ square transformed and †square root transformed parameter
* 15 lags, [1]NA: No anisotropy

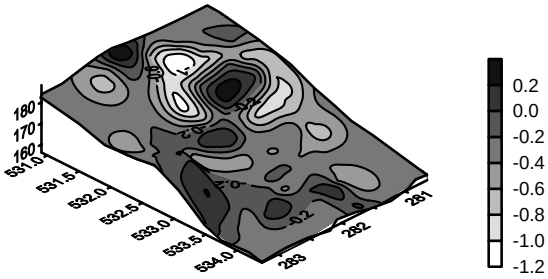
4.3.2 Spatial distribution of soil properties: Tamale pilot site

The flow in the topsoil is dominated by slow and moderate flow characteristics (saturated hydraulic conductivity) with some places (between rising elevations) having a very slow flow (Figure 4.5(A1); see Table 3.2 for flow classification). At the subsoil level, the flow is mainly very slow with some high elevation points having a slow flow (Figure 4.5(A2)). A relationship exists between the conductivity in the topsoil and subsoil, as the same locations show similar flow patterns. Conductivity in the subsoil is better defined than in the topsoil (Figure 4.5(A)). This is expressed in the higher range (520 m) and sill (0.16) of the subsoil compared to that of the topsoil (475 m, 0.13) (Table 4.15).

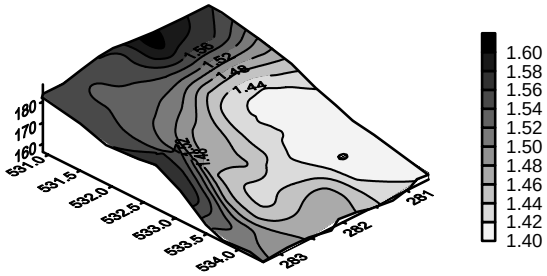
(A1) $LogK_s$ (log $cmhr^{-1}$) (Topsoil)



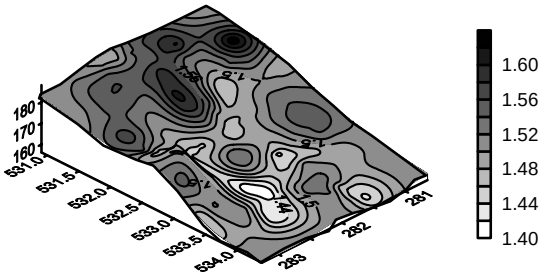
(A2) $LogK_s$ (log $cmhr^{-1}$) (Subsoil)



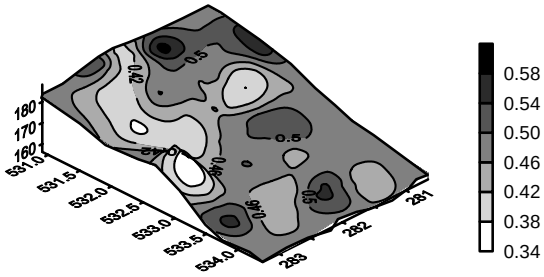
(B1) Bulk density (gcm^{-3}) (Topsoil)



(B2) Bulk density (gcm^{-3}) (Subsoil)



(C1) Organic carbon (%) (Topsoil)



(C2) Organic carbon (%) (Subsoil)

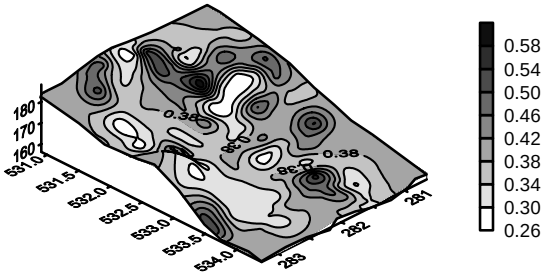
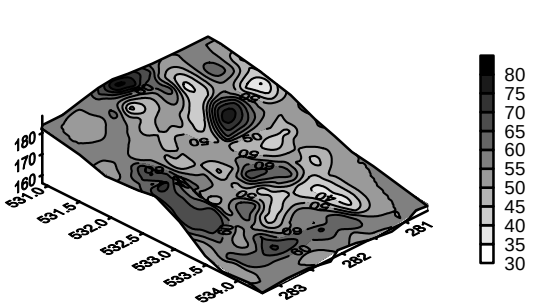
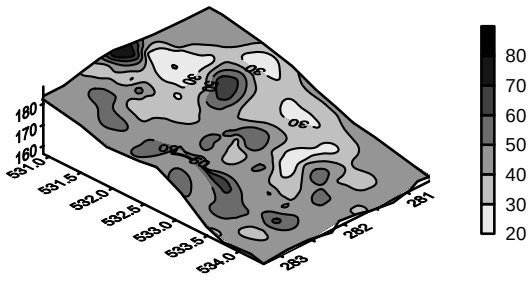


Figure 4.5a. Spatial distribution of soil properties at Tamale pilot site: (A1) log of topsoil saturated hydraulic conductivity (K_s); (A2) log of subsoil saturated hydraulic conductivity (K_s); (B1) topsoil bulk density; (B2) subsoil bulk density; (C1) topsoil organic carbon content and (C2) subsoil organic carbon content

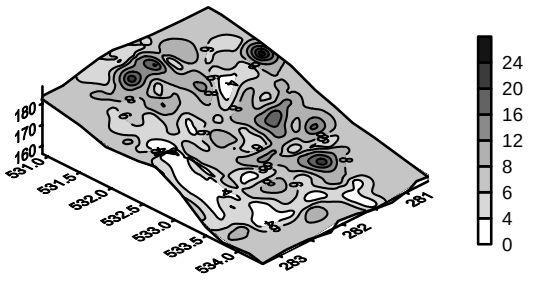
(D1) Sand content (%) (Topsoil)



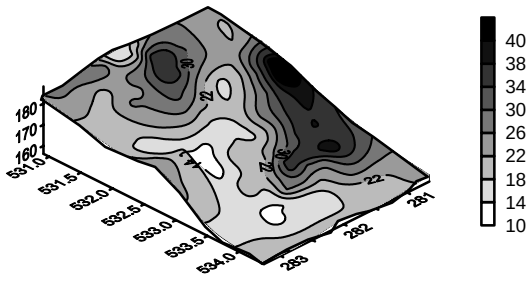
(D2) Sand content (%) (Subsoil)



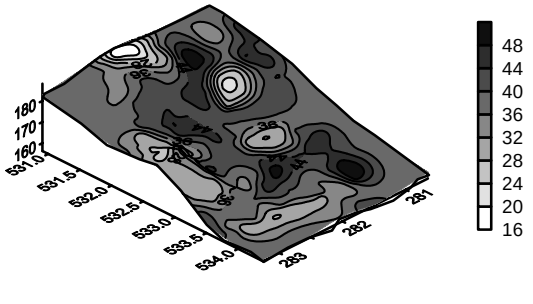
(E1) Clay content (%) (Topsoil)



(E2) Clay content (%) (Subsoil)



(F1) Silt content (%) (Topsoil)



(F2) Silt content (%) (Subsoil)

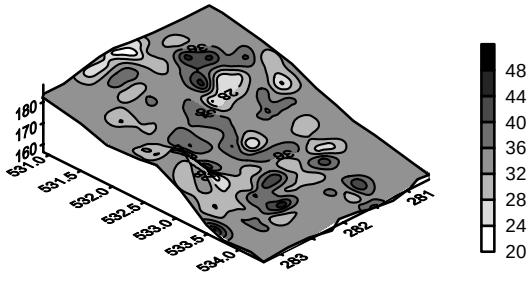


Figure 4.5b. Spatial distribution of soil properties at Tamale pilot site: (D1) topsoil sand content; (D2) subsoil sand content; (E1) topsoil clay content; (E2) subsoil clay content; (F1) topsoil silt content and (F2) subsoil silt content

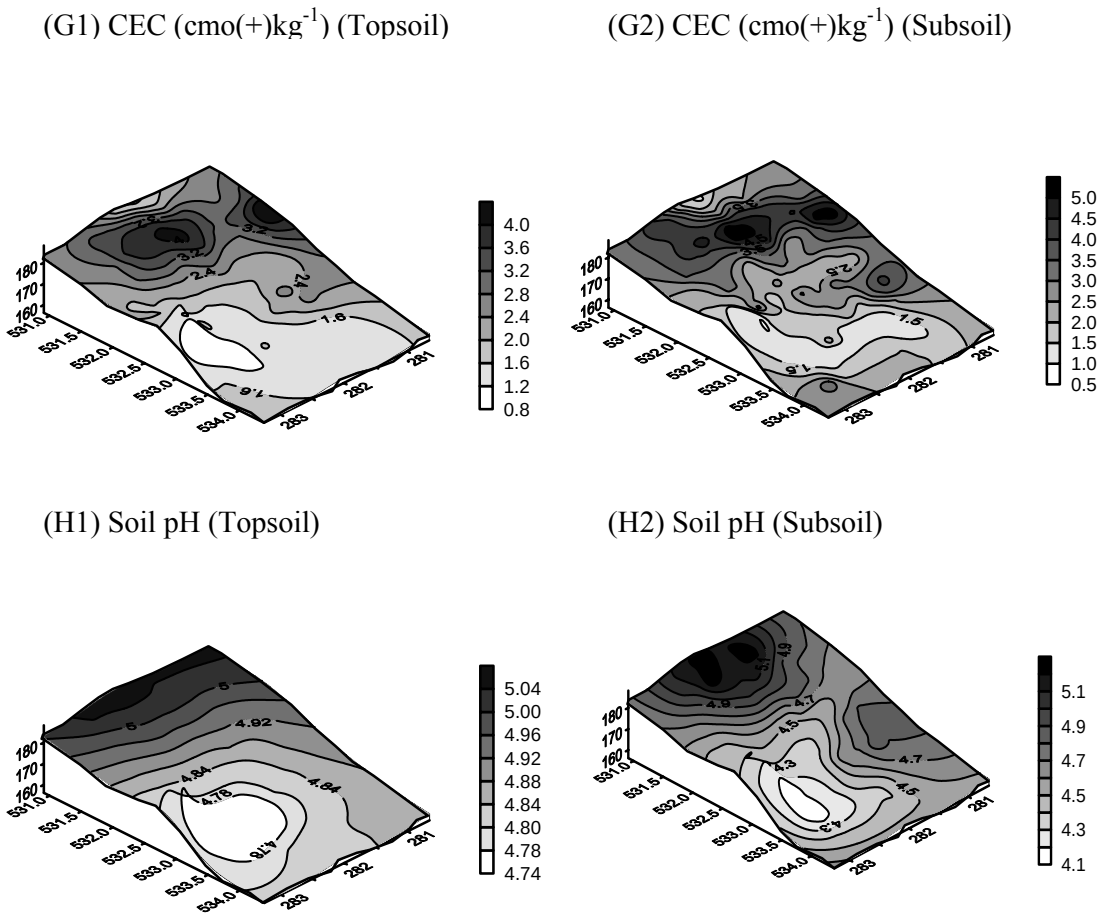


Figure 4.5c. Spatial distribution of soil properties at Tamale pilot site: (G1) topsoil CEC; (G2) subsoil CEC (H1); topsoil pH and (H2) subsoil pH

Bulk density increases with elevation in both the topsoil and subsoil (Figure 4.5(B)). There is a strong positive correlation at the 0.01-level between bulk density and elevation with r-values of 0.38 (topsoil) and 0.64 (subsoil) (Table 4.17). Bulk density also increases with slope in the topsoil (r =0.15) and subsoil (r =0.25).

The distribution of organic carbon at the Tamale site exhibits similar distribution characteristics in the topsoil and subsoil, with areas of high values in the topsoil showing high values in corresponding areas in the subsoil (Figure 4.5(C)). The organic carbon pattern is poorly related to terrain parameters. The pattern observed may be due to the influence of different land use types such as different fallow regimes, crops and crop management.

The spatial distribution of the soil textures evaluated in terms of particle size distribution (sand, clay and silt content) is illustrated in Figure 4.5(D), 4.5(E), and 4.5(F). The clay content (Figure 4.5(E)) is high at high elevation and decreases from the slope shoulder to the channel, as indicated by the positive correlation with elevation (Table 4.17). The mean (23 %) and range of variation (51 %) of clay content in the subsoil is higher than that in the topsoil (see Table 4.2). The spatial distribution of sand and clay in the topsoil is patchy as indicated by the short semivariogram range of 161 m for sand and 155 m for clay (see Table 4.13). The silt content increases from the footslope to the channel in the topsoil and subsoil (Figure 4.5(F)). In general the silt content has a significant negative correlation at 0.01-level with elevation, with values of -0.22 (topsoil) and -0.38 (subsoil) (Table 4.17).

Comparing CEC distribution patterns in the topsoil and subsoil, similar patterns are observed, with regions of high CEC in the topsoil corresponding to regions of high CEC in the subsoil (Figure 4.5(G)). The CECs in the topsoil and subsoil have a similar spread with a low mean CEC (less than the critical level of 5 cmol(+)kg^{-1} set by Landon, (1991)). As presented in Table 4.17, CEC has a significant correlation (at 0.01-level) with elevation, slope gradient, LS factor, and aspect for both soil depths. This implies that CEC increases with elevation and slope.

Figures 4.5(H1) and 4.5(H2) illustrate that for both soil depths, the soil pH increases with increasing elevation. This is reflected in the strong positive correlation between soil pH and elevation for the topsoil (0.80) and subsoil (0.89) (Table 4.17). Soil pH also has a positive correlation with slope gradient. This may be the result of high leaching levels experienced at the higher elevation.

There is a significant negative correlation only between topsoil K_s and elevation (0.01-level) and LS factor (0.05-level) (see Table 4.17a), and not in the subsoil (Table 4.17b).

4.3.3 Spatial distribution of soil parameters: Ejura pilot site

The distribution pattern of saturated hydraulic conductivity in the topsoil and subsoil increases from the backslope to channel (Figures 4.6(A1) and 4.6(A2)). The distribution in both soil depths are not uniform, reflecting the highly variable nature of K_s. Flow in the topsoil and subsoil varies from very rapid to very slow, with very high flow areas in

the topsoil corresponding to similar flow areas in the subsoil (see Figures 4.6(A1) and 4.6(A2)). Figure 4.6(B1) and 4.6(B2) illustrate the similarity in bulk density distribution at the topsoil and subsoil level, respectively, at the Ejura site. The subsoil has higher values, which is due to limited disturbance of soil at that soil depth.

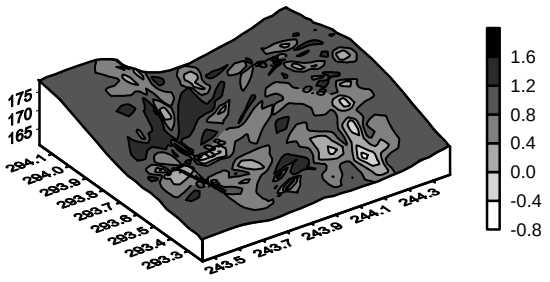
Bulk density is high on the backslope and footslope. The subsoil bulk density has a significant positive correlation (0.32) with elevation, implying that it increases with elevation. For both soil depths there is a significant negative correlation (0.01-level) between bulk density and wetness index and upslope area (see Table 4.18).

The organic carbon distribution at the Ejura site is illustrated in Figure 4.6(C) with a high organic carbon content observed along the stream channel. This is mainly due to forest vegetation along the river. Otherwise there is not much variation of organic carbon across the landscape. Apart from the area along the stream, organic carbon content in the topsoil is higher compared to the subsoil.

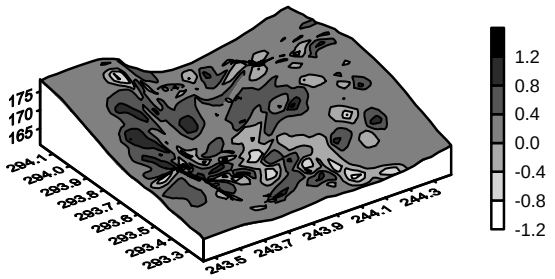
Figures 4.6(D), 4.6(E), and 4.6(F) show distribution patterns of the sand, clay, and silt content spatial, respectively, at the Ejura site. The interpolated maps indicate similar distribution patterns for topsoil and subsoil. There is a comparatively high clay and silt content along the toeslope and channel. Topsoil sand has a positive correlation with elevation (0.14) and silt has significant negative correlation with elevation (-0.20), LS factor (-0.14), and slope gradient (-0.11) (see Table 4.18a). In the subsoil, silt has a significant negative correlation (0.01-level) with upslope area, stream power index, slope gradient, and LS factor and a positive correlation (0.16) with slope aspect (Table 4.18b). At the subsoil level, sand content is positively correlated with slope gradient while the clay content has a negative correlation with elevation (see Table 4.18b).

The distribution of CEC in the topsoil is without any well-defined pattern (Figure 4.6(G1)). However, in the subsoil there is a well-laid pattern with high CEC along the stream channel. All the other areas have values less than 2.8 cmol(+)kg^{-1} as shown in Figure 4.6(G2). CEC has a significant correlation (0.05-level or better) with all the terrain parameters except slope gradient and aspect for both soil depths and with plan curvature for the subsoil CEC (see Table 4.18).

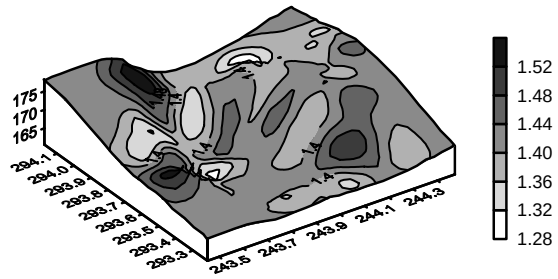
(A1) LogK_s (log cmhr^{-1}) (Topsoil)



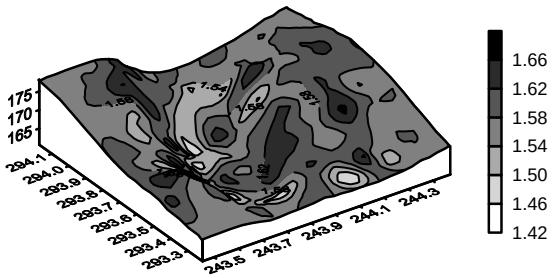
(A2) LogK_s (log cmhr^{-1}) (Subsoil)



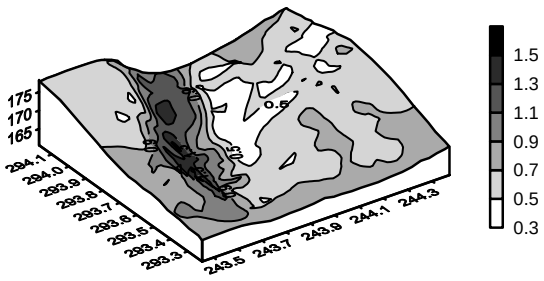
(B1) Bulk density (gcm^{-3}) (Topsoil)



(B2) Bulk density (gcm^{-3}) (Subsoil)



(C1) Organic carbon (%) (Topsoil)



(C2) Organic carbon (%) (Subsoil)

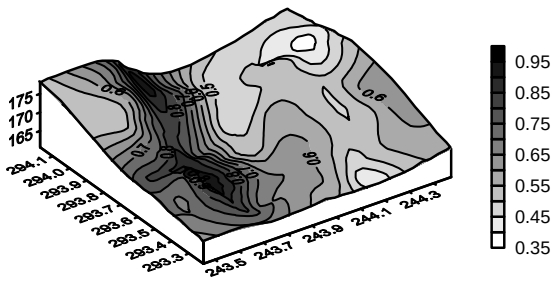
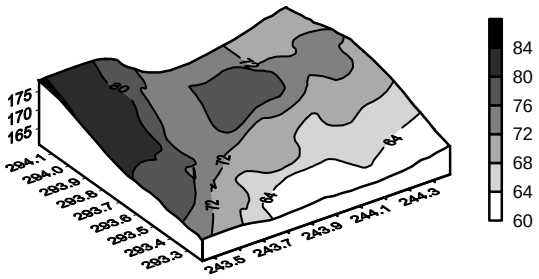
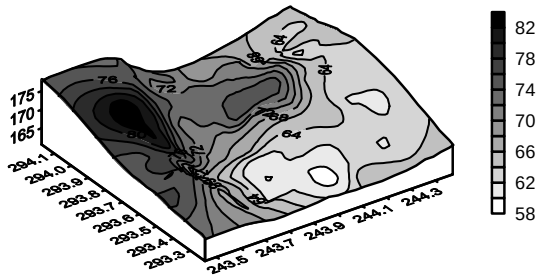


Figure 4.6a. Spatial distribution of soil properties at Ejura pilot site: (A1) log of topsoil saturated hydraulic conductivity (K_s); (A2) log of subsoil saturated hydraulic conductivity (K_s); (B1) topsoil bulk density; (B2) subsoil bulk density; (C1) topsoil organic carbon content and (C2) subsoil organic carbon content

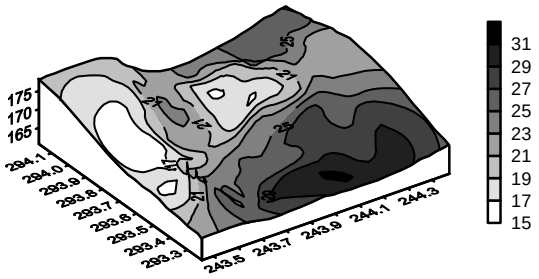
(D1) Sand content (%) (Topsoil)



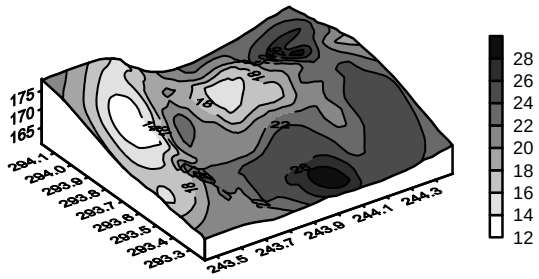
(D2) Sand content (%) (Subsoil)



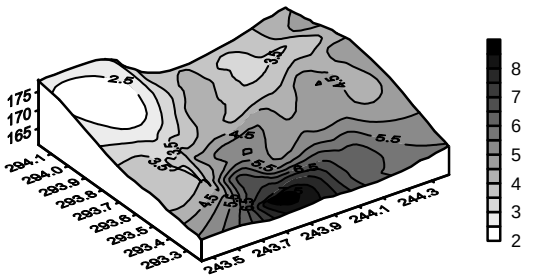
(E1) Clay content (%) (Topsoil)



(E2) Clay content (%) (Subsoil)



(F1) Silt content (%) (Topsoil)



(F2) Silt content (%) (Subsoil)

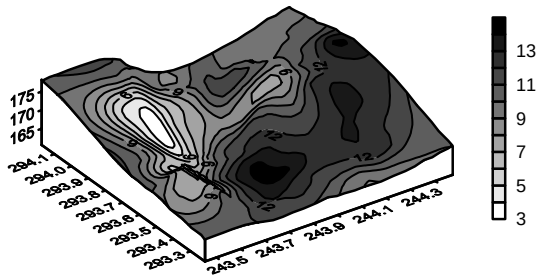
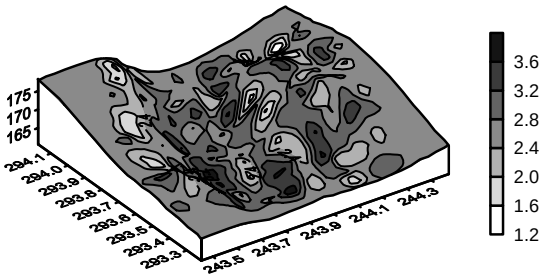
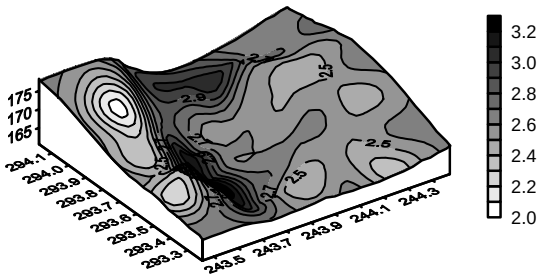


Figure 4.6b. Spatial distribution of soil properties at Ejura pilot site: (D1) topsoil sand content; (D2) subsoil sand content; (E1) topsoil clay content; (E2) subsoil clay content; (F1) topsoil silt content and (F2) subsoil silt content

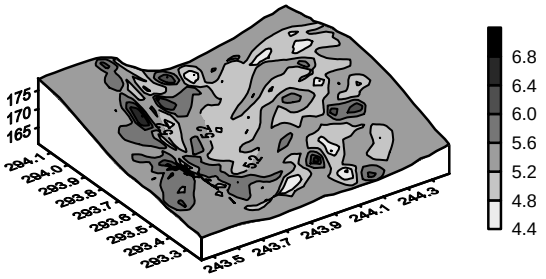
(G1) CEC (cmol(+)kg^{-1}) (Topsoil)



(G2) CEC (cmol(+) kg^{-1}) (Subsoil)



(H1) Soil pH (Topsoil)



(H2) soil pH (Subsoil)

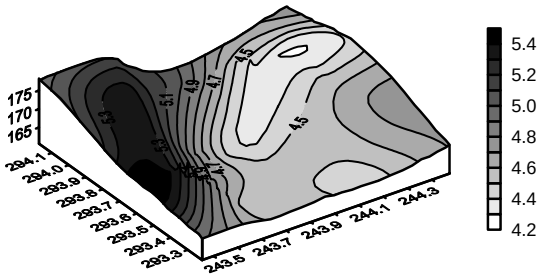


Figure 4.6c. Spatial distribution of soil properties at Ejura pilot site: (G1) topsoil CEC; (G2) subsoil CEC; (H1) topsoil pH and (H2) subsoil pH

A high correspondence exists in the topsoil and subsoil pH, as illustrated by the pH distribution map in Figure 4.6(H). However, the pH distribution patterns in the topsoil have higher range of variation than in the subsoil (Figures 4.6(H1) and 4.6(H2)). The pH has a negative correlation (0.01-level) with elevation and a significant positive correlation with the upslope contributing area and stream power index at both soil depths (see Table 4.18). For the topsoil, there is a significant correlation between soil pH and slope aspect (Table 4.18a).

4.3.4 Comparison of spatial distribution of soil properties at Tamale and Ejura sites

The soil properties exhibited varying degrees of spatial variability as illustrated in Figures 4.5 and 4.6 for the two pilot sites. This variability is an indication of the

heterogeneous nature of the soil and therefore the need for detailed studies at the topo-scale for effective calibration of hydro-ecological models. The soil properties sand content, silt content, bulk density, pH and organic carbon content have significant and high correlations with the saturated hydraulic conductivity, underlining their importance for use as surrogates in pedotransfer functions.

4.4 Spatial correlation of K_s with soil and terrain parameters

Tables 4.17 and 4.18 present the Pearson's correlation (r) and significance level for terrain parameters, and interpolated soil data for the Tamale and Ejura sites. Here, the focus is on relationship of saturated hydraulic conductivity (K_s) and other parameters, as the former is the focus of this work.

At the Tamale site, a significant correlation at the 0.01-level was observed between K_s and topsoil soil parameters (sand content, silt content, CEC, clay content, and pH – in decreasing order). For terrain parameters, only elevation had correlation with K_s (Table 4.17a). At the 0.05 significance level, K_s correlated negatively with the LS factor. The absence of correlation between K_s and topsoil bulk density at the Tamale site may be due to farming activities that disturb the topsoil. It is worth noting the significant correlation amongst other parameters, (such as CEC and elevation (0.76), pH and elevation (0.80), sand content and silt content (-0.86), and sand content and clay content (-0.54)), and therefore the possible existence of multicollinearity in further analysis. For the Tamale subsoil, significant correlations (0.01-level) exist between K_s and all the soil parameters but none of the terrain parameters (Table 4.17b). The higher r-values between K_s and soil parameters in the subsoil are the result of minimal soil disturbance at that soil depth.

Saturated hydraulic conductivity in the topsoil was significantly correlated at the 0.01-level with bulk density, sand content, silt content, clay content, and LS-factor – in decreasing order of magnitude – at the Ejura site (Table 4.18a). At this soil depth, stream power index, upslope contributing area, slope gradient, and soil pH each had a significant correlation with K_s in the range of 0.11-0.14. The subsoil level K_s and all soil properties (sand, silt, and clay content, pH, organic carbon and bulk density – in decreasing order of magnitude) were correlated. Of the terrain parameters, significant correlation (-0.17) exists only between the K_s and aspect. With the exception of bulk

Table 4.17a. Pearson coefficient of correlation (r) for topsoil spatial data at Tamale (using 30 m × 30 m grid)

	K_s^L	PFC	PNC	Cv	ELEV	WI^L	UA^L	SPI^L	$SL^\dagger$	$LS^\dagger$	$AS^\dagger$	$SAND^\ddagger$	$CLAY^\dagger$	$SILT^L$	$CEC^\dagger$	$OC^\dagger$	$BD^\ddagger$
Profile curvature (PFC)	0.05																
Plan curvature (PNC)	0.07	0.73**															
Curvature (Cv)	0.06	0.93**	0.93**														
Elevation (ELEV)	-0.18**	0.00	-0.07	-0.04													
Wetness index (WI) ^{L}	-0.04	-0.39**	-0.42**	-0.44**	-0.20**												
Upslope area (UA) ^{L}	-0.09	-0.39**	-0.47**	-0.46**	-0.14*	0.91**											
Stream power index (SPI) ^{L}	-0.12	-0.33**	-0.45**	-0.42**	-0.04	0.53**	0.83**										
Slope (SL) †	-0.12	0.10	-0.02	0.04	0.30**	-0.55**	-0.20**	0.31**									
Length-slope factor (LS) †	-0.14*	-0.14*	-0.30**	-0.23**	0.12	0.15*	0.50**	0.82**	0.70**								
Aspect (AS)†	0.01	-0.12	-0.09	-0.12	-0.13*	0.12	0.06	0.00	-0.17**	-0.14							
Sand‡	0.49**	0.03	0.06	0.05	0.03	-0.17*	-0.14*	-0.05	0.10	-0.01	0.12						
Clay†	-0.26**	-0.12	-0.18**	-0.16*	0.23**	0.10	0.12	0.10	0.02	0.09	-0.12	-0.54**					
Silt^L	-0.49**	0.00	0.03	0.02	-0.22**	0.15*	0.10	0.01	-0.16*	-0.06	-0.08	-0.86**	0.22**				
CEC†	-0.39**	-0.08	-0.14*	-0.12	0.76**	-0.04	0.01	0.06	0.25**	0.17**	-0.26**	-0.27**	0.37**	0.15*			
Organic carbon (OC)	-0.03	-0.06	-0.02	-0.04	-0.02	-0.03	-0.03	-0.01	-0.05	-0.07	-0.01	-0.26**	0.28**	0.19**	0.08		
Bulk density (BD) ‡	0.01	0.00	-0.04	-0.02	0.38**	-0.15*	-0.14*	-0.10	0.15*	0.00	0.20**	0.29**	-0.06	-0.37**	0.15*	0.04	
pH^L	-0.18**	-0.07	-0.10	-0.09	0.80**	-0.20**	-0.16*	-0.07	0.26**	0.06	-0.12	0.02	0.19**	-0.17**	0.69**	0.03	0.38**

^{L} Log, ‡ Square and †Square root transformed parameter

** Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

Table 4.17b. Pearson coefficient of correlation (r) for subsoil spatial data at Tamale (30 m × 30 m grid)

	K_s[L]	PFC	PNC	Cv	ELEV	WI[L]	UA[L]	SPI[L]	SL[†]	LS[†]	AS[†]	SAND[‡]	CLAY[†]	SILT[L]	EC[†]	OC[†]	BD[‡]
Profile curvature (PFC)	0.06																
Plan curvature (PNC)	0.11	0.62**															
Curvature (Cv)	0.10	0.91**	0.89**														
Elevation (ELEV)	-0.18	0.04	-0.02	0.01													
Wetness index (WI)[L]	-0.09	-0.36**	-0.42**	-0.43**	-0.29**												
Upslope area (UA)[L]	-0.10	-0.34**	-0.44**	-0.43**	-0.20*	0.93**											
Stream power index (SPI)[L]	-0.12	-0.30**	-0.45**	-0.41**	-0.06	0.71**	0.91**										
Slope (SL)[†]	-0.03	0.11	0.00	0.06	0.37**	-0.54**	-0.23**	0.18*									
Length-slope factor (LS)[†]	-0.11	-0.13	-0.30**	-0.24**	0.15	0.21*	0.51**	0.79**	0.67**								
Aspect (AS)[†]	-0.06	0.00	0.03	0.01	-0.20*	0.06	-0.02	-0.11	-0.22**	-0.24**							
Sand[‡]	0.61**	0.08	0.14	0.12	0.01	0.00	0.01	0.00	-0.02	0.01	0.05						
Clay[†]	-0.64**	0.02	-0.06	-0.02	0.30**	-0.16	-0.16*	-0.10	0.12	-0.02	-0.04	-0.75**					
Silt[L]	-0.31**	-0.04	0.00	-0.02	-0.38**	0.16	0.14	0.10	-0.10	0.03	-0.14	-0.55**	0.11				
CEC[†]	-0.41**	-0.02	-0.09	-0.06	0.72**	-0.10	-0.02	0.08	0.28**	0.21**	-0.32**	-0.35**	0.49**	0.07			
Organic carbon (OC)	-0.42**	-0.02	0.00	-0.01	0.20*	-0.07	-0.08	-0.08	0.03	-0.04	-0.15	-0.45**	0.46**	0.23**	0.47**		
Bulk density (BD)[‡]	-0.43**	0.01	-0.05	-0.02	0.64**	-0.20*	-0.15	-0.04	0.25**	0.09	0.14	-0.23**	0.42**	-0.08	0.55**	0.26**	
pH[L]	-0.28**	0.01	-0.08	-0.04	0.89**	-0.22**	-0.15	-0.03	0.29**	0.13	-0.16*	-0.15	0.43**	-0.20	0.75**	0.27**	0.59**

[L] Log, [‡] Square and [†]Square root transformed parameter

** Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

Table 4.18a. Pearson coefficient of correlation (r) for topsoil spatial data at Ejura (30 m × 30 m grid)

	K_s [L]	PFC	PNC	Cv	ELEV	WI [L]	UA [L]	SPI [L]	SL [†]	LS [†]	AS [†]	SAND [‡]	CLAY [†]	SILT [L]	CEC [†]	OC [†]	BD [‡]
Profile curvature (PFC)	-0.01																
Plan curvature (PNC)	-0.02	0.57**															
Curvature (Cv)	-0.01	0.92**	0.84**														
Elevation (ELEV)	-0.01	0.39**	0.27**	0.39**													
Wetness index (WI) [L]	0.06	-0.40**	-0.51**	-0.50**	-0.33**												
Upslope area (UA) [L]	0.12*	-0.49**	-0.54**	-0.57**	-0.46**	0.89**											
Stream power index (SPI) [L]	0.14*	-0.47**	-0.46**	-0.52**	-0.44**	0.43**	0.79**										
Slope (SL) [†]	0.11*	-0.02	0.10	0.03	-0.11*	-0.57**	-0.16**	0.40**									
Length-slope factor (LS) [†]	0.19**	-0.32**	-0.25**	-0.33**	-0.36**	0.00	0.43**	0.83**	0.79**								
Aspect (AS) [†]	-0.04	0.05	-0.06	0.00	-0.03	0.02	0.05	0.03	0.05	0.04							
Sand [‡]	0.44**	0.00	0.01	0.00	0.14**	-0.03	-0.02	0.01	0.09	0.08	-0.10						
Clay [†]	-0.36**	0.00	-0.02	-0.01	-0.07	0.08	0.06	0.02	-0.10	-0.05	0.09	-0.94**					
Silt [L]	-0.42**	0.07	-0.02	-0.06	-0.20**	-0.02	-0.07	-0.09	-0.11*	-0.14**	0.01	-0.69**	0.57**				
CEC [†]	-0.05	-0.19**	-0.17**	-0.20**	-0.43**	0.16**	0.21**	0.18**	0.03	0.17**	0.07	-0.27**	0.23**	0.27**			
Organic carbon (OC)	0.07	-0.09	-0.10*	-0.10*	-0.54**	0.22**	0.33**	0.34**	0.14**	0.32**	-0.11*	0.15**	-0.10	-0.25**	0.22**		
Bulk density (BD) [‡]	-0.51**	0.05	0.04	0.05	0.09	-0.15**	-0.14**	-0.06	0.09	0.00	-0.16**	0.11*	-0.17**	-0.02	-0.15**	0.08	
pH [L]	0.11*	-0.01	-0.05	-0.03	-0.31**	0.11*	0.12*	0.09	-0.02	0.06	-0.13*	0.32**	-0.35**	-0.24**	-0.02	0.46**	.15**

[L] Log, [‡] Square and [†]Square root transformed parameter

** Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

Table 4.17b. Pearson coefficient of correlation (r) for subsoil spatial data at Ejura (30 m × 30 m grid)

	K_s[L]	PFC	PNC	Cv	ELEV	WI[L]	UA[L]	SPI[L]	SL[†]	LS[†]	AS[†]	SAND[‡]	CLAY[†]	SILT[L]	CEC[†]	OC[†]	BD[‡]
Profile curvature (PFC)	0.11																
Plan curvature (PNC)	0.01	0.59**															
Curvature (Cv)	0.08	0.93**	0.85**														
Elevation (ELEV)	-0.01	0.40**	0.32**	0.41**													
Wetness index (WI)[L]	-0.06	-0.39**	-0.51**	-0.49**	-0.32**												
Upslope area (UA)[L]	-0.07	-0.48**	-0.55**	-0.56**	-0.43**	0.89**											
Stream power index (SPI)[L]	-0.06	-0.46**	-0.47**	-0.52**	-0.39**	0.39**	0.75**										
Slope (SL)[†]	0.04	-0.02	0.10	0.04	-0.07	-0.59**	-0.18**	0.41**									
Length-slope factor (LS)[†]	0.01	-0.31**	-0.25**	-0.32	-0.31**	-0.04	0.40**	0.82**	0.80**								
Aspect (AS)[†]	-0.17**	0.04	-0.10	-0.02	-0.03	0.03	0.05	0.03	0.05	0.04							
Sand[‡]	0.64**	0.05	0.03	0.05	0.01	-0.05	0.01	0.09	0.17**	0.16**	-0.17**						
Clay[†]	-0.58**	-0.06	-0.06	-0.07	-0.13*	0.15**	0.12*	0.03	-0.16**	-0.07	0.16**	-0.92**					
Silt[L]	-0.57**	-0.05	0.02	-0.02	0.11	-0.07	-0.17**	-0.22**	-0.18**	-0.26**	0.16**	-0.83**	0.59**				
CEC[†]	-0.14*	-0.17**	-0.07	-0.15*	-0.40**	0.15*	0.22**	0.18**	0.10	0.20**	0.10	-0.06	0.10	0.00			
Organic carbon (OC)	0.23**	0.00	-0.03	-0.02	-0.50**	0.06	0.14*	0.16**	0.12*	0.20**	-0.16**	0.31**	-0.15**	-0.37**	0.30**		
Bulk density (BD)[‡]	-0.21**	0.08	0.07	0.08	0.32**	-0.16**	-0.17**	-0.09	0.09	0.00	-0.10	-0.02	-0.03	0.08	-0.17**	0.05	
pH[L]	0.51**	0.02	-0.05	-0.01	-0.18**	0.09	0.12*	0.13*	0.06	0.12*	-0.25**	0.76**	-0.63**	-0.68**	-0.05	0.61**	0.09

[L] Log, [‡] Square and [†]Square root transformed parameter

** Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

density, the magnitude of correlation between K_s and the soil parameters is higher for the subsoil than for the topsoil. Significant correlations amongst the other parameters also exist as explained earlier in this section.

4.5 Conclusion

The texture of the Ejura soil is more uniform with depth than that of Tamale. The Tamale soils have a sharp increase in clay content from the topsoil to the subsoil. This difference may be due to differences in parent material at the two sites, with the Tamale soils having a higher inherent clay content, which may have translocated down the profile over time.

The soil properties at the two sites were spatially highly variable, especially the saturated hydraulic conductivity at both soil depths. The high variability of K_s underlines the difficulty in accurately assessing this soil property and therefore the need for suitable estimation methods. The least variable parameter was bulk density.

The terrain attributes at the Ejura site were more pronounced compared to those at the Tamale site. The site is characterized by steeper slopes compared to gentler topography near Tamale. This implies that the Ejura site is more prone to high runoff and erosion if the land cover is removed.

It was observed that the spatial dependencies for soil properties in the subsoil had higher ranges than in the topsoil. This may be the result of less soil disturbance in the subsoil zone. Also, the range of spatial dependency for soil properties at Tamale was generally higher than that at the Ejura site. This is the result of a gentler slope at the Tamale site and therefore the greater likelihood that parameters have been less influenced by abrupt an terrain effect.

The K_s was observed to be related with spatially distributed soil properties, but the relationship with terrain attributes was poor. Consequently, it will be more suitable to use soil properties to estimate K_s than terrain attributes.

5 SPATIAL DISTRIBUTION OF SOIL TYPES AND LAND USE TYPES AND THEIR RELATION TO SOIL PROPERTIES

Most soil physical parameters are highly variable (Wilding, 1985 and Warrick and Nelson, 1980). In the past, soil surveys through the use of mental models based on point observations of soil profiles have been extended to larger areas to provide maps of pre-classified soil types for which it is very often difficult to separate evidence from interpretation (Hudson, 1992). Soil survey is used to classify soils into areas and location of supposedly uniform soil properties. The principal constraint is our ability to represent spatially variable attributes (Cook et al., 1996) and primary soil properties and their variation within the soil types. Such representation is an approximation of a more complex pattern of soil variation (Lagacherie et al., 1996).

This chapter is divided into four sections with three main objectives: (1) identify the soil type distribution at the two pilot sites in the Volta Basin of Ghana and the possible processes that may lead to such distribution of the soil types, (2) compare the soil types at each site in terms of the soil properties, and (3) compare the different land use types observed during sampling in terms of soil properties. In each of the first three sections, the results obtained for the different sites are compared. Finally, in the last section, conclusions are drawn based on the findings in the previous sections.

5.1 Spatial distribution of soil types

In this section, the main objective is to identify the spatial distribution of the different soil types and the processes that may lead to the soil types at the Tamale and Ejura sites and to compare the results from the two sites. To achieve this, the section is divided into sub-sections highlighting the method used in mapping the soils at the pilot sites, the spatial distribution of the soil types and a comparison of results from the two sites.

5.1.1 Methods of soil identification and mapping

A preliminary survey was first carried out to give a broad picture of the possible soil series (types) that may be encountered at the site, followed by the detailed survey. The soils were classified into soil series or types using mini-pits, the common soil surveying practice in Ghana (Adu and Mensah-Ansah, 1995).

Mini-pits of about 30 cm diameter were dug up to 70 cm depth at each stake point along the transects (Figure 3.1) using an earth chisel. Data was collected f on soil diagnostic horizons, texture, color (Munsell color chart), mottles, structure, roots, concretional fraction, topographical position, and land use.

A soil series comprises soils that have been developed under similar conditions from similar parent material and therefore exhibit similar profile morphology or characteristics. This means that soils in the same series should have profiles that have been developed under similar drainage and climatic conditions; they should have the same parent rock, same number of horizons or layers and corresponding layers should have similar color, texture, consistency, structure and content of secondary minerals. For convenience of reference, upland soil series are named after towns (e.g., Techiman series) and lowland soils are named after rivers or streams (e.g., Sene series) within the locality where they were first identified (Adu and Mensah-Ansah, 1995).

Eight (8) representative sample soil profile pits were dug, three (3) at the Tamale site and five (5) at the Ejura site. The representative profile locations were selected after visiting all sample points and an initial sketch of the different soil types at the site was made. The profiles were then sited within the identified soil types. Table 5.1 provides the environmental conditions at the soil profile points. For consistency, all soils were described using the guidelines for soil description by FAO (1990).

With the coordinates for all sampling (mini-pit) points known from Differential Global Positioning System (DGPS) readings (see section 3.3.1), the records were critically examined and put in a GIS database using ArcView. Similar records indicating the same soil types were digitized by interpolation. Note here that in this study all sampling points were visited based on grid patterns (see Figure 3.1) and the area mapped. This differs from the usual practice in soil survey, where soils are identified along selected transects and then, based on one's expert knowledge and information on topography and vegetation, the soil map is extended to unsampled sites.

In all, 13 soil types (series) were identified, eight (namely: Haplic Luvisol, Lithic Leptosol, Ferric Acrisol, Plinthic Acrisol, Dystric Plinthosol, Eutric Plinthosol, Eutric Gleysol, and Dystric Gleysol)) at the Tamale site and five (namely: Ferralic Cambisol, Ferric Acrisol, Haplic Acrisol, Gleyic Acrisol, and Gleyic Fluvisol) at the Ejura site (FAO, 1988) (See Appendix 1 for series (local) names). The descriptions of

the different soil types are presented in sub-sections 5.1.2 and 5.1.3 with detailed profile description in Appendices 2 and 3.

Table 5.1. Soil types at the profile pits at the Tamale and Ejura sites

[a]Profile code	FAO (1988) classification	Series (local) name	Nearest town	Latitude (° ‘ “N)	Longitude (° ‘ “W)	Altitude (m)	Relief position
TP1	Ferric Acrisol	Kumayili	Tamale	9 28 10.5	0 55 48.3	190.3	Upper slope
TP2	Dystric Plinthosol	Kpelesawgu	Tamale	9 28 17.8	0 55 53.0	186.7	Mid-slope
TP3	Eutric Gleysol	Lima	Tamale	9 28 25.9	0 55 4.8	177.0	Lower slope
EP1	Ferralic Cambisol	Kintampo	Ejura	7 19 25.6	1 16 17.3	180.3	Upper slope
EP2	Ferric Acrisol	Techiman	Ejura	7 19 18.6	1 16 21.7	175.0	Upper slope
EP3	Haplic Acrisol	Amantin	Ejura	7 19 26.4	1 16 29.2	163.3	Mid-slope
EP4	Gleyic Acrisol	Denteso	Ejura	7 19 18.3	1 16 38.0	160.7	Lower slope
EP5	Gleyic Fluvisol	Sene	Ejura	7 19 20.2	1 16 34.8	163.3	Lower slope

[a]Profile code as in Figures 5.1 and 5.3

The soils at each site may be grouped into three main divisions as follows: upland soils, mid-slope, and soils of flat valley bottoms and lower slopes. The upland soils comprise free draining soils such as: Luvisol, Leptosol and Acrisol at the Tamale site; and Cambisol, and Ferric Acrisol at the Ejura site. The mid-slope soils consist of Dystric Plinthosol, Eutric Plinthosol at Tamale and Haplic and Gleyic Acrisol at the Ejura site. The soils on flat valley bottoms and lower slopes are seasonally waterlogged such as Gleysol at the Tamale site and Fluvisol at the Ejura site.

5.1.2 Spatial distribution of soil types at the Tamale pilot site

Figure 5.1 illustrates the different soil types identified at the Tamale site. The figure also provides the location of the sample profiles dug at the site and illustrates the variation in soil types in terms of drainage, texture, depth and amount of iron concretion. It is evident from the figure that the grid size of 100 m × 200 m selected for this site may be too large to capture variations for soil types covering smaller areas, thus the patchy pattern observed for smaller soil types.

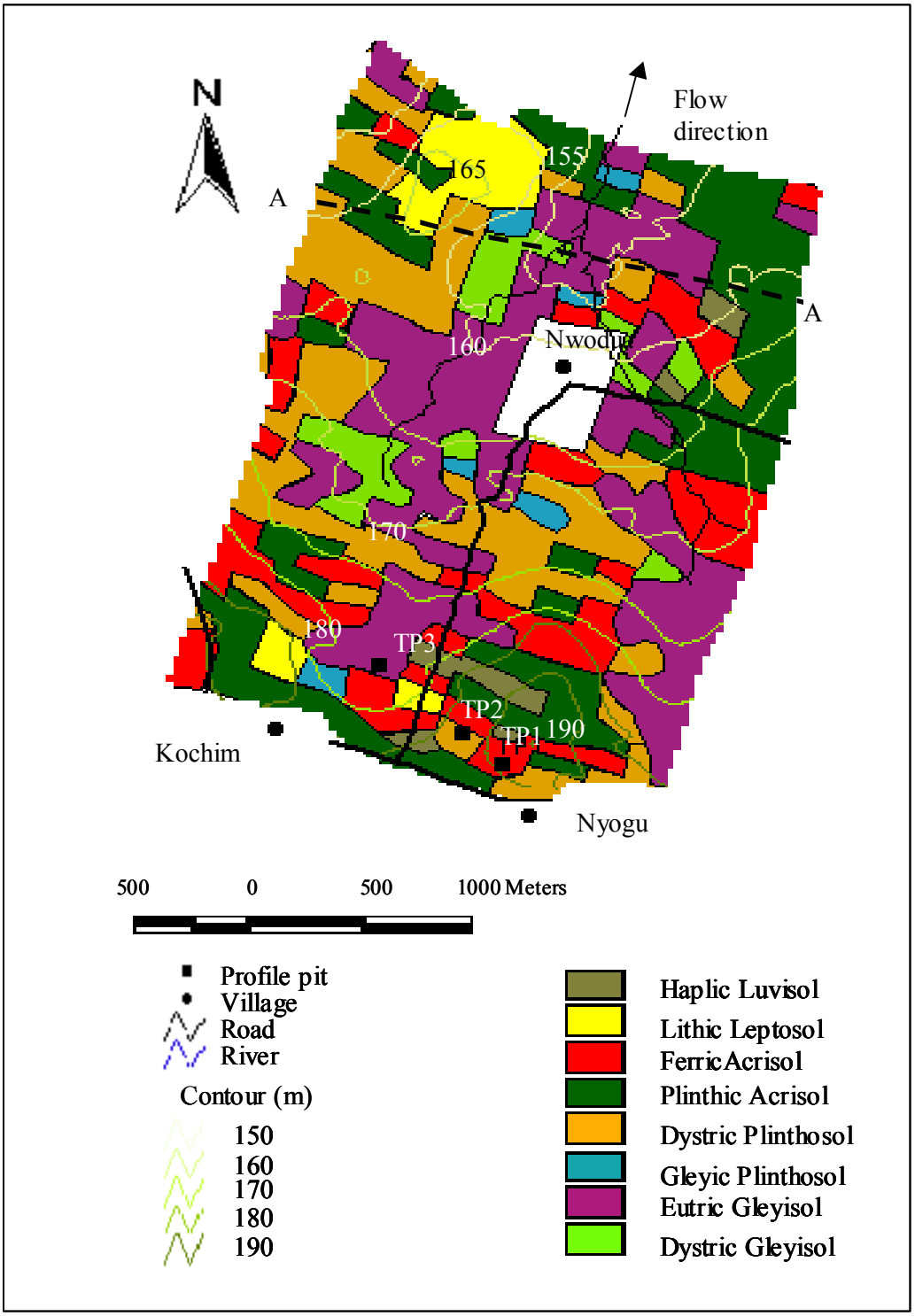


Figure 5.1. Soil map showing soil types, contours and location of profile pits at Tamale pilot site

Soils found around Tamale were developed over Voltaian clay shale under savannah vegetation. Plinthic Acrisol and Dystric Plinthosol in the uplands and Eutric Gleysol in the flat valley bottoms dominate in the site. A detailed sample profile description for the Ferric Acrisol, Dystric Plinthosol and Eutric Gleysol is provided in Appendix 2.

The Tamale soils are characterized by many, i.e., >15 % by volume gravel concretion (Table 5.2). Soils such as Lithic Leptosol and Plinthic Acrisol have more than 60 % gravel concretions at both soil depths (0-15 cm and 30-45 cm).

Table 5.2. Soil types with many or higher (i.e., >15% by volume) gravel and/or concretion at the Tamale and Ejura sites

	% area with gravel or concretion	
Soil type	Topsoil (0-15 cm)	Subsoil (30-45 cm)
	Tamale site	
Haplic Luvisol	20.0	40
Lithic Leptosol	66.7	75
Ferric Acrisol	15.2	19.4
Plinthic Acrisol	64.8	61.0
Dystric Plinthosol	7.7	25.6
Eutric Plinthosol	0.0	0.0
Eutric Gleysol	4.6	0
Dystric Gleysol	0	7.7
	Ejura site	
Ferralic Cambisol	0.0	21.0
Ferric Acrisol	0.0	3.6
Haplic Acrisol	0.0	0.0
Gleyic Acrisol	0.0	1.4
Gleyic Fluvisol	0.0	0.0

Plinthite and its hardened form petroplinthite (gravel concretion) in the past were thought to be formed on flat terrains with hydromorphic conditions and that the present occurrence of petroplinthite in upland soils is the results of earth inversion. A recent study by Asiamah (2002) has shown that plinthite and petroplinthite formation is an on-going process in all agro-ecological zones and landscape positions in Ghana. This process is attributed to the continuous enrichment of the main mineral component (iron oxide) when vegetation cover is removed. Once the Ferric ion content reaches the threshold level of 80 $mgkg^{-1}$ and moisture content is less than 10 %, irreversible hardening of the material occurs (Asiamah, 2002).

Figure 5.2 illustrates the soil types encountered across the section A-A in Figures 4.1 and 5.1. It shows the variation in soil morphological properties as one moves from high to low elevation. Following are the characteristics of the eight different soil types identified at the Tamale site, with information on their probable location in the landscape and the possible usage to which they can be put.

Haplic Luvisol (HL) usually consist of sandy soil with few scattered small-sized, spherical, black and shiny ironstone concretions (petroplinthite) in the topsoil of about 15 cm or more thickness overlying a sandy clay subsoil horizon of tightly packed irregular ironstone gravel and ferruginized sandstone. It is well drained and occupies the summit and upper slopes with high K_s in the subsoil. It is usually used for the cultivation of cereals (such as sorghum, maize) and yams.

Lithic Leptosol (LL) is an upland soil found mostly on the summit. These soils are shallow to very shallow, well to moderately well drained, highly concretionary soils, or erosion remnants. Ironpan outcrops are rampant over the surface, which is often completely devoid of dense vegetation. These soils are not suitable for cultivation as they are too shallow and droughty.

Ferric Acrisol (FA) is a moderately well drained soil found on low summits, upper and middle slopes (Figure 5.2). The topsoil is often disturbed by cultivation. The texture usually varies from loamy sand topsoil overlying sandy loam to sandy clay loam in the subsoil. It is usually a moderately deep soil (about 90 cm or more), below which tightly packed ironstone concretions or an ironpan may be encountered (see representative profile in Appendix 2a). It is comparatively favored for cultivation because it is easy to work.

Plinthic Acrisol (PA) is moderately well drained, shallow, with abundant concretion and occurs on upper and middle slopes. It consists of slightly humus topsoil with concretions up to a depth of about 30 cm. It usually has lumps of soft ironstone that have a high tendency to develop ironpan in the subsoil and as a consequence form a shallow rooting media. The subsoil is usually heavy textured compared to the topsoil leading to the holding up of seepage water in the wet season between the two horizons (see Table 5.3). This soil type is considered relatively unproductive because of its high concretional gravel content and the presence of clay shale at shallow depths. It is usually used for the cultivation of maize and groundnuts through manual tillage.

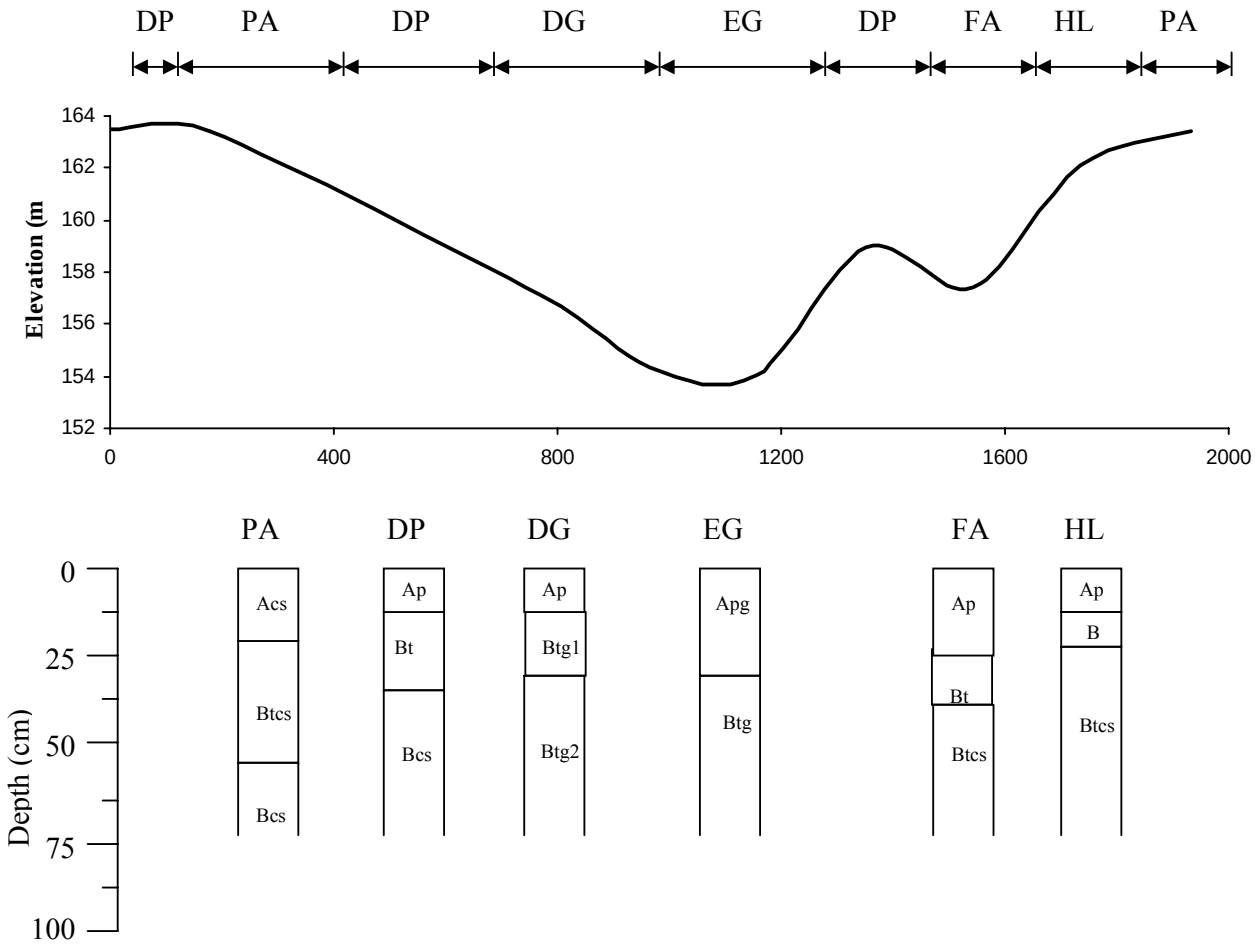


Figure 5.2. Soil types across the Tamale pilot site along section A-A as shown in Figures 4.1 and 5.1

Dystric Plinthosol (DP) occurs on very gentle slopes usually below Plinthic Acrisol and above Eutric and Dystric Gleysol (Figure 5.2). It is mainly formed from colluvium and is found on the lower slopes of undulations transition between the well-drained soils of the upper slopes and the poorly drained soils of the flat valley bottoms. They consist of about 30 cm more or less sandy loam to loam overlying clayey subsoil with many to abundant concretions (see representative profile in Appendix 2b). They are imperfectly drained because of impedance within the clay substratum and the middle-lower slope position on the topography. This soil type is unattractive for cultivation, because it dries out rapidly early in the dry season and easily erodes.

Eutric Plinthosol (EP) is a groundwater laterite soil, transitional between the imperfectly drained soils of the Dystric Plinthosol and the floodplains soils of the Eutric and Dystric Gleysol. It is shallow and poorly drained with massive ironstone, and consists of sandy loam topsoil over massive, compact and manganiferous ironpan, which often extends downwards 30-70 cm or more. Eroded areas usually have outcrops of massive ironpans at the surface. They are saturated with water during the rainy seasons and exceedingly droughty during the dry seasons. They are unsuitable for agriculture except in some limited areas where the topsoil is fairly deep. However, they only support poor pasture for grazing.

Eutric Gleysol (EG) occurs in sloughs of clay shale formation often in association with Dystric Gleysol. It usually consists of a sandy loam or loamy sand over sandy loam or loam that may contain frequent polished ironstone concretions. This soil usually overlies layers with an abruptly high clay content that are very hard and very compact, developing a clay pan when dry and becoming plastic and massive when wet (see representative profile description in Appendix 2c). Eutric Gleysols are poorly drained and subjected to seasonal water logging or flooding, but generally become very dry during dry seasons. With good water control measures – to control excessive flooding and drought – it is highly recommended for rice production.

Dystric Gleysol (DG) is an alluvial soil found at the valley bottoms. It usually consists of slightly yellow mottled, porous, silty clay of up to about 15 cm overlying orange or red mottled silty clay. Dystric Gleysol usually has a lighter soil texture at depth (about 120 cm or deeper), which is characteristic of the series that suggest its development on old riverbed alluvium overlain by subsequent in-filling (Adu, 1995).

Like the Eutric Gleysol, it is also liable to water logging and flooding for varying periods, but generally becomes thoroughly dry during the dry season. This soil has potential for mechanized flooded rice cultivation and dry season vegetable growing.

Description of soil properties of soil types at Tamale site

Table 5.3 presents data on mean and coefficient of variation (CV) for soil properties (sand, clay, organic carbon, CEC, saturated hydraulic conductivity (K_s), and bulk density for topsoil and subsoil of different soil types at the Tamale site. In general, the soils have a sand content of between 42-63 % in the topsoil and 40-55 % in the subsoil. The topsoils have a low clay content (10 % or less). All soil types have subsoils with higher clay content than the topsoil in the range of 17 - 29 %; this may be a result of translocation of clay particles down the soil profile.

In the topsoil, all soil types have an average carbon content and in the subsoil, with the exception of Haplic Luvisol and Lithic Leptosol, a low organic carbon (< 0.5 %). The above average organic carbon obtained for the subsoil of Lithic Leptosol may be related to its limited land use due to its shallowness. The CEC values are low for all soil types, i.e., less than the 5 cmol(+)kg^{-1} limit given by Landon (1991). All soil types are acidic with pH less than or equal to 5.1 and 4.8 in the topsoil and subsoil, respectively, as a result of the soil being highly weathered and leached.

The bulk density for the different soil types at the Tamale site is in the range of 1.2-1.6 gcm^{-3}. The K_s is highly variable with CVs as high as 225 and the lowest mean value of 0.5 cmh^{-1} (very slow) for the subsoil of Dystric Gleysol and the highest mean value of 7.5 cmh^{-1} (moderately rapid) for Ferric Acrisol in the subsoil (Table 5.3). In general, the Dystric Gleysol had the lowest K_s of 1.2 cmh^{-1} in the topsoil and 0.5 cmh^{-1} in the subsoil (see Table 3.2 for classification values of saturated hydraulic conductivity or flow rate).

The CVs indicate high variation of soil properties within the various soil types, except for soil pH and bulk density. This confirms the difficulty in putting soil into uniform groups (soil type) with sharp boundaries (Burrough, 1986).

Table 5.3. Means and coefficients of variation of soil properties (topsoil and subsoil) for different soil types at the Tamale site

Soil type	n	Sand (%) Mean (CV)[a]	Silt (%) Mean (CV)	Clay (%) Mean (CV)	Carbon (%) Mean (CV)	CEC[b] Mean (CV)[a]	pH Mean (CV)	BD[c] (g cm^{-3}) Mean (CV)	K_s (cm h^{-1}) Mean (CV)
Topsoil									
Haplic Luvisol	5	63.0 (24.0)	31.9 (37.7)	5.1 (74.6)	0.6 (97.0)	2.1 (80.7)	4.8 (10.9)	1.5 (12.0)	1.9 (72.9)
Lithic Leptosol	9	62.3 (22.4)	31.4 (29.8)	6.3 (143.1)	0.6 (57.1)	1.9 (56.8)	5.1 (7.1)	1.5 (6.4)	4.8 (80.8)
Ferric Acrisol	46	56.8 (26.9)	35.8 (32.8)	7.4 (87.4)	0.5 (64.9)	2.6 (53.7)	5.1 (9.5)	1.5 (7.1)	2.2 (140.4)
Plinthic Acrisol	54	62.3 (20.2)	30.6 (32.1)	7.1 (86.3)	0.5 (63.1)	2.3 (61.2)	5.0 (8.4)	1.6 (9.6)	3.4 (116.5)
Dystric Plinthosol	52	54.5 (21.3)	38.6 (27.5)	6.9 (65.5)	0.5 (50.0)	2.1 (53.9)	4.8 (9.3)	1.5 (7.4)	2.0 (138.6)
Eutric Plinthosol	4	62.2 (17.6)	33.3 (36.9)	4.5 (61.5)	0.2 (54.5)	1.8 (53.3)	4.7 (5.8)	1.4 (11.0)	1.4 (118.3)
Eutric Gleysol	65	53.5 (19.3)	40.3 (21.5)	6.2 (61.4)	0.5 (40.6)	2.0 (50.6)	4.7 (12.5)	1.4 (7.7)	1.6 (121.9)
Dystric Gleysol	15	42.3 (15.3)	48.1 (16.3)	9.6 (61.6)	0.6 (52.0)	1.7 (41.9)	4.6 (11.4)	1.4 (5.5)	1.2 (81.9)
Total	250								
Subsoil									
Haplic Luvisol	5	44.6 (39.9)	26.7 (14.7)	28.7 (53.5)	0.5 (87.8)	2.6 (78.6)	4.7 (11.7)	1.6 (13.8)	4.7 (129.7)
Lithic Leptosol	4	44.0 (27.9)	38.6 (18.2)	17.4 (80.5)	0.6 (60.4)	1.6 (34.7)	4.4 (12.0)	1.5 (3.8)	1.8 (174.9)
Ferric Acrisol	36	46.0 (39.3)	30.8 (29.4)	23.2 (62.9)	0.4 (53.3)	2.7 (57.6)	4.8 (13.1)	1.5 (8.1)	2.3 (148.9)
Plinthic Acrisol	41	46.4 (28.1)	29.3 (20.7)	24.3 (43.7)	0.4 (45.2)	2.4 (53.0)	4.6 (15.1)	1.6 (10.7)	7.5 (102.4)
Dystric Plinthosol	43	39.8 (31.8)	32.6 (21.0)	27.6 (46.1)	0.4 (46.7)	2.7 (52.7)	4.6 (8.0)	1.5 (7.5)	2.8 (197.6)
Eutric Plinthosol	2	54.9 (37.3)	23.8 (9.9)	21.3 (107.4)	0.3 (6.3)	1.2 (6.7)	4.3 (4.4)	1.2 (8.3)	0.6 (20.2)
Eutric Gleysol	61	44.9 (25.8)	35.6 (26.7)	19.5 (50.9)	0.4 (55.1)	2.4 (52.5)	4.5 (11.2)	1.5 (5.9)	1.1 (225.0)
Dystric Gleysol	13	40.7 (30.3)	36.6 (14.9)	22.7 (49.4)	0.3 (33.8)	1.8 (63.6)	4.4 (11.5)	1.5 (6.0)	0.5 (145.8)
Total	205								

[a]CV: Coefficient of variation; [b]CEC (cmol(+)kg^{-1}); [c]BD: bulk density

5.1.3 Spatial distribution of soil types at the Ejura site

The soils found at the Ejura site are formed over fine-grained Voltaian feldspathic (arkose) sandstones (Adu and Mensah-Ansah, 1995). They occur under forest-savannah transitional vegetation, forming a Cambisol - Acrisol - Fluvisol soil association. The area has predominantly Acrisols on the uplands and midslopes and Fluvisols in the lowland and valley bottom. As illustrated in Figure 5.3, the dominant soil types at this site are the Ferric Acrisol, Haplic Acrisol and Gleyic Acrisol. A detailed description of the sample profile for the soil types at the Ejura site is presented in Appendix 3.

Five soil types were identified with varying morphological properties from the upland to the lowland. Figure 5.4 illustrates the soils encountered for section B-B across the mapped area at Ejura as shown in Figures 4.2 and 5.3.

A brief description of these soil types is provided in this sub-section. Figure 5.5 shows pictures of three profile pits at Ejura: Ferralic Cambisol – EP1; Ferric Acrisol – EP2; and Gleyic Acrisol – EP4. The pictures illustrate the shallowness of the Ferralic Cambisol, the presence of iron concretions in the Ferric Acrisol and the presence of an ironpan at depths in the Gleyic Acrisol.

Ferralic Cambisol (FC) is an upland soil found mostly on the summit (Figure 5.4). These soils are shallow to very shallow, well to moderately well drained, with some concretionary or erosion remnants (Table 5.4 and Appendix 3a). The soil is formed over ironpan layers. Because it is shallow in nature (30 cm or less deep to ironpan sheet), it is best left under natural vegetation (Figure 5.5).

Ferric Acrisol (FA) is a moderately shallow to moderately deep soil, which often has concretion in the lower subsoil (Figure 5.5 and Appendix 3b). It occurs on summits and upper slopes of gentle and moderate sloping topography and has high internal drainage, thus drying up readily during the dry season (Table 5.4 and Figure 5.7). It is moderately susceptible to erosion and marginal for mechanical cultivation because topsoil erosion can lead to the exposure of the concretionary subsoil. It is fairly good for cultivation of shallow rooted crops such as maize, legumes and vegetables.

Haplic Acrisol (HA) is a moderately well to imperfectly drained soil in the association (Table 5.4 and Figure 5.7). It occurs on the upper portion of the mid-slope. It is a deep soil, free of concretions and gravel (see representative profile in Appendix

3c). It is medium texture with low natural fertility and erosion prone. It is ideal for cultivation of yams, cassava, maize, tobacco, and vegetables.

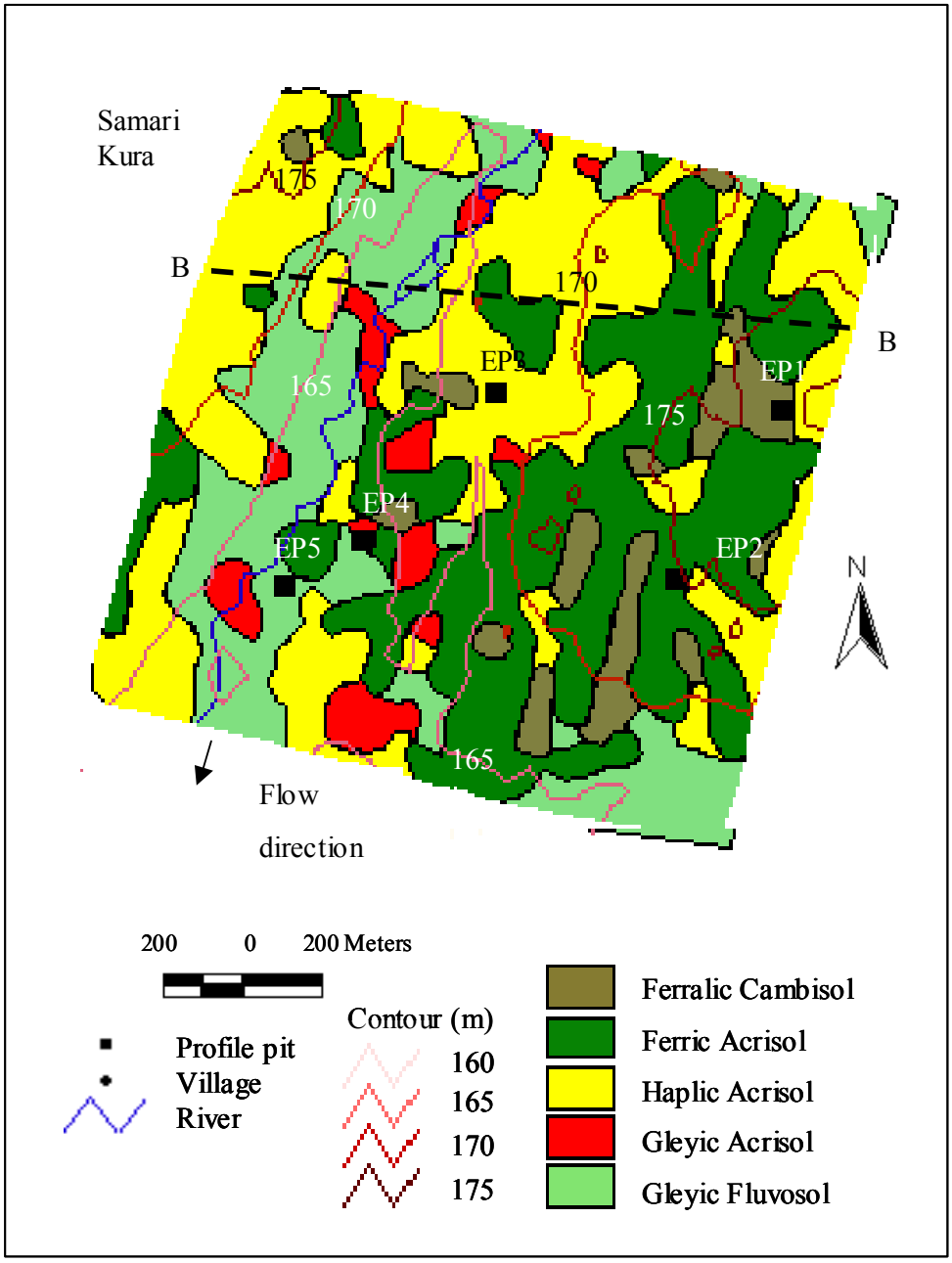


Figure 5.3. Soil map showing soil types, contours and location of profile pits at Ejura pilot site

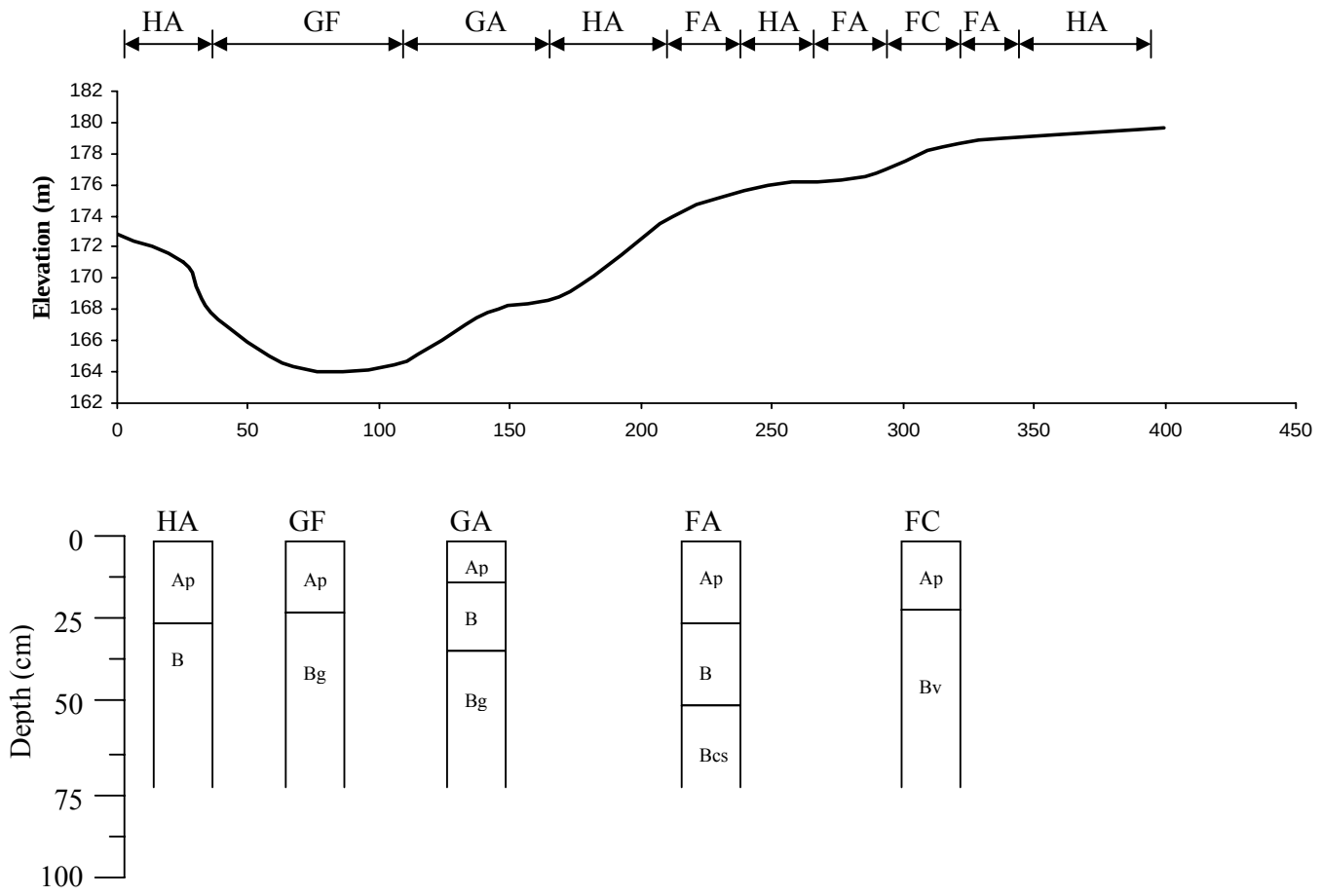


Figure 5.4. Soil types across the Ejura pilot site along section B-B as shown in Figures 4.2 and 5.3

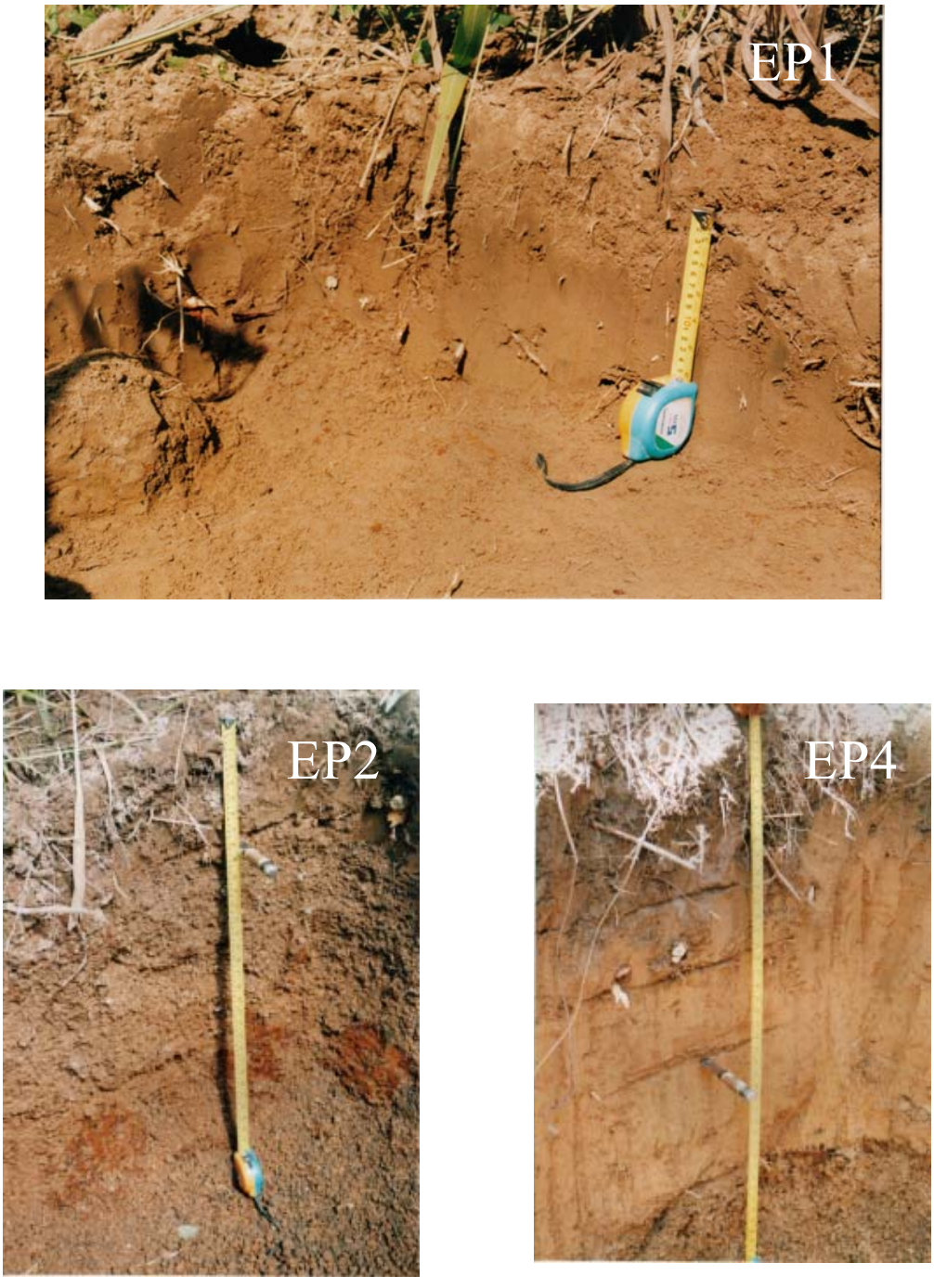


Figure 5.5. Selected profile pits at Ejura pilot site: Ferralic Cambisol – EP1; Ferric Acrisol – EP2; and Gleyic Acrisol – EP4.

Gleyic Acrisol (GA) is a moderate to poorly drained soil that occurs at the mid-lower slope on gently sloping topography and may extend close to the streambed. It is mainly derived as a wash from Haplic Acrisol and other soil types on upper and mid-slopes. They are usually deep and sandy and may overlie a hard ironstone pan (see Figure 5.5). The soil is usually saturated during the wet season but dries out rapidly during the dry season (see representative profile description in Appendix 3d).

Gleyic Fluvisol (GF) is the poorly drained valley bottom member of the soil association found at the Ejura site (Table 5.4 and Figure 5.7). It occurs mainly in valley bottoms that are seasonally waterlogged or flooded (Figure 5.4). It is basically an alluvial soil, mainly found on level bottomlands bordering the stream. As presented in Appendix 3e it is a moderately deep to deep soil.

Description of soil properties of soil types at Ejura site

Table 5.4 presents data on mean and coefficients of variation (CV) for soil properties (sand, silt, clay, organic carbon, CEC, bulk density and saturated hydraulic conductivity (K_s) for the different soil types at the top and subsoil at the Ejura site. The soils types have a high sand content with mean values of 65 % and 62 % or more in the topsoil and subsoil, respectively. Clay content is low, between 4-8 % for the topsoil and 7-14 % for the subsoil.

The topsoil organic carbon content of lowland soils of Gleyic Acrisol (1.1 %) and Gleyic Fluvisol (0.8 %) is higher than that of the upland soils (0.6-0.7 %). The CEC is low for all the soil types with values less than 3 cmol(+)kg^{-1}, which is below the critical value of 5 cmol(+)kg^{-1} indicated by Landon (1991). The soils are acidic with topsoil pH in the range of 5.1-5.5 and that of the subsoil 4.5-4.8.

The mean bulk densities of 1.4-1.5 gcm^{-3} for the topsoil and 1.6 gcm^{-3} for the subsoil are within the acceptable range for cultivation. All the soils have a very rapid flow (> 12 cmh^{-1}) in the topsoil except the Ferralic Cambisol and Gleyic Fluvisol a with rapid flow (8-12 cm^{-1}). In the subsoil, Gleyic Fluvisols have a slow flow (< 2 cmh^{-1}), the remaining soil types have a moderate flow (2-6 cmh^{-1}) or better (see Table 3.2 for classification values of K_s or flow rate).

Table 5.4. Means and coefficients of variation of soil properties (topsoil and subsoil) for different soil types at the Ejura pilot site

Soil type	n	Sand (%) Mean (CV)[a]	Silt (%) Mean (CV)	Clay (%) Mean (CV)	Carbon (%) Mean (CV)	CEC[b] Mean (CV)[a]	pH Mean (CV)	BD[c] (g cm^{-3}) Mean (CV)	K_s (cm h^{-1}) Mean (CV)
				Topsoil					
Ferralic Cambisol	27	70.8 (8.2)	24.4 (17.5)	4.8 (59.3)	0.7 (32.4)	2.7 (35.9)	5.3 (12.1)	1.4 (6.9)	9.6 (110.9)
Ferric Acrisol	122	71.3 (8.7)	24.2 (21.3)	4.5 (51.0)	0.6 (39.7)	2.8 (29.9)	5.2 (10.6)	1.4 (8.0)	12.4 (102.9)
Haplic Acrisol	134	74.5 (8.3)	21.3 (22.6)	4.2 (61.2)	0.7 (45.8)	2.6 (32.8)	5.3 (11.3)	1.4 (9.4)	12.5 (102.7)
Gleyic Acrisol	74	73.1 (11.8)	22.9 (30.8)	4.0 (81.9)	1.1 (39.6)	2.8 (25.5)	5.5 (13.3)	1.4 (10.8)	19.0 (96.9)
Gleyic Fluvisol	17	65.3 (14.2)	26.4 (27.1)	8.3 (51.0)	0.8 (43.7)	2.9 (25.2)	5.1 (17.2)	1.5 (8.7)	9.5 (106.7)
Total	374								
				Subsoil					
Ferralic Cambisol	19	62.7 (12.3)	25.1 (27.6)	12.2 (32.6)	0.6 (31.0)	2.5 (25.2)	4.7 (16.9)	1.6 (7.0)	27.4 (142.1)
Ferric Acrisol	111	65.4 (9.9)	23.0 (21.4)	11.6 (31.3)	0.6 (32.6)	2.4 (29.9)	4.5 (15.1)	1.6 (5.2)	8.6 (292.9)
Haplic Acrisol	132	70.4 (11.1)	19.6 (27.0)	10.0 (47.4)	0.6 (46.7)	2.8 (26.9)	4.8 (16.2)	1.6 (6.8)	5.7 (226.2)
Gleyic Acrisol	70	72.5 (13.7)	20.6 (32.1)	6.8 (76.3)	0.8 (37.5)	2.8 (31.0)	5.0 (17.0)	1.6 (8.8)	4.3 (167.5)
Gleyic Fluvisol	17	62.0 (21.2)	24.3 (37.2)	13.7 (40.7)	0.6 (37.9)	2.9 (20.6)	4.7 (23.5)	1.6 (9.0)	1.8 (198.9)
Total	349								

[a]CV: Coefficient of variation; [b]CEC (cmol(+)kg^{-1}); [c]BD: bulk density

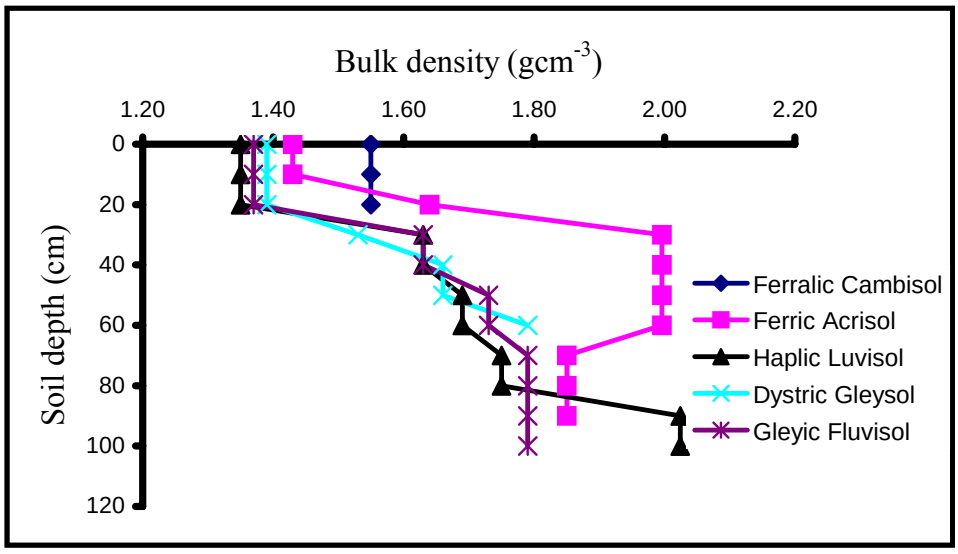


Figure 5.6. Bulk density at different soil depths for the representative soil profiles at Ejura

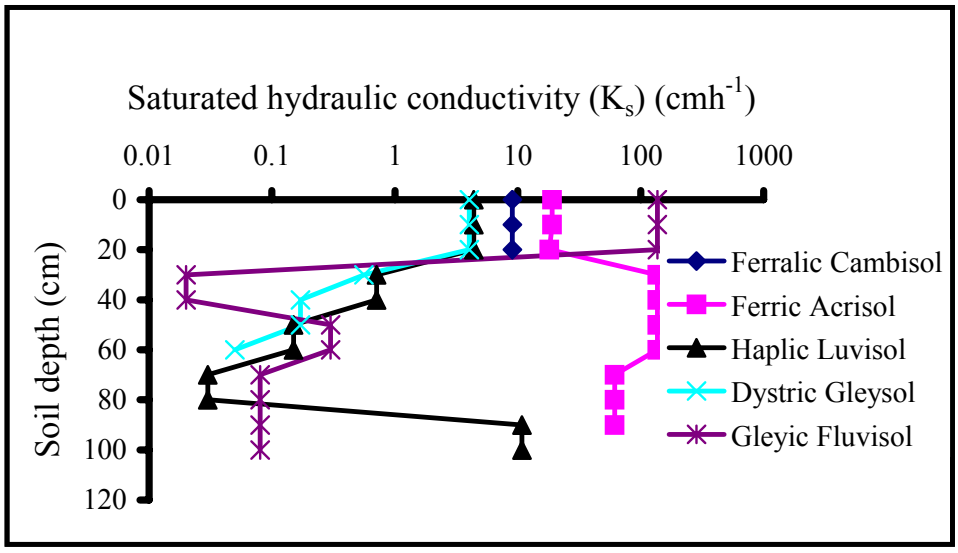


Figure 5.7. Saturated hydraulic conductivity at different soil depths for the representative soil profiles at Ejura

Figures 5.6 and 5.7 illustrate the bulk density and K_s, respectively, as these vary with soil depth at the site. The Ferralic Cambisol is a shallow soil with a depth of about 20 cm, after which there is an ironpan (Figures 5.5 and 5.6). The remaining soil types

have increasing bulk density with soil depth, with Ferric Acrisol showing a sharp increase as a result of the concretional layer followed by a decline (Figure 5.6).

The Haplic Acrisol, Gleyic Acrisol, and Gleyic Fluvisol soils have decreasing K_s with soil depth, corresponding to the increasing bulk density. However, the Ferric Acrisol has an increasing K_s with soil depth as a result of high porosity associated with loose iron concretions.

The CVs for the soil properties were high (15 % or more) for most parameters except for topsoil sand content, pH, and bulk density and subsoil sand content and bulk density. The K_s has the highest within soil type variation, with values ranging from 97 % to 293 %. This high variation highlights the heterogeneous nature of the soils, and therefore, the difficulty in classifying them meaningfully into soil types.

5.1.4 Comparison of distribution of soil types at the Tamale and Ejura sites

In general, the distribution of soil types follows the catena, with variation of soil morphological properties from slope summit through the mid-slope to the valley bottom. However, the variation of soil along the slope is more distinct at the Ejura site due to the steep nature of the terrain at that site.

At Tamale, the clay content in the subsoil is a higher than in the topsoil. This is mainly due to the higher clay content in the parent material at that site and its subsequent translocation into the subsoil over time. The Ejura soils are sandier throughout the top and subsoils.

Distinct soil types are found to dominate the soils of the uplands and lowlands. For instance, at both sites the soils in the lowlands and valley bottoms are observed to have more fine particles than the soils in the uplands and mid-slopes. Also, at the Tamale site, most of the soil types have many gravel concretions in both soil depths. These concretions are more dominant in soil types such as Haplic Luvisol, Lithic Leptosol, Ferric Acrisol and Plinthic Acrisol found on uplands and Dystric Plinthosol, and Eutric Plinthosol found on mid-slopes.

The dominant soil processes seen to influence soil formation and therefore soil properties at the two sites are vertical and lateral soil translocation and leaching. This has resulted in different soil types across the soil catena. At the Tamale site, soil

plinthization is also a very common process, which has resulted in a large amount of plinthic (concretional) material.

5.2 Distribution of soil properties for different soil types

The focus here is to determine the variation among the mapped soil types from section 5.1 in terms of soil properties. The similarities and/or differences among soil physical properties for specific soil types at the two sites were determined using non-parametric analysis.

The analysis was done using non-parametric median test (SPSS Inc., 1999) in preference to ANOVA because of the non-normally distributed nature of some of the soil properties and the limited sample size for some of the soil types. See Appendices 4 and 5 for Kolmorgorov-Smirnov normality test statistic and significance values, and Kruskal-Wallis mean rank for soil properties of the different soil types. These show that most of the parameters are significantly different from normal distribution ($p < 0.05$) and have wide differences based on the ranking. It must be noted that the usual assumption of sharp boundaries among different soil types is not true in reality, thus making it difficult to discern variation of soil properties across soil types. The median test is used to determine the existence of differences among the different soil types for a particular soil property at a given site. The pairwise analysis is used to determine which pairs of soil types are significantly different from one another for a given soil property.

The variation of soil-landscape patterns is the result of a combination of long- and short-term pedogeomorphic processes in conjunction with human activities. Since the two areas considered are of relatively small size, it is assumed that variation in soil is due mainly to differences in slope and elevation as geology is uniform and there are no changes in climate. The variation is usually the result of movement of soil downslope and longer water saturation regimes in the foot and toeslopes as compared to the backslope, shoulder and interfluve.

5.2.1 Comparison of soil properties for different soil types at the Tamale site

Table 5.5 presents the median, percentage of values greater than the parameter median (P), group median test analysis (Chi-Square and significance level) and pair-wise median test for the different soil properties at the Tamale site.

The median test resulted in a highest Chi-Square of 42.8 for subsoil K_s and a lowest value of 5.46 for clay content in the topsoil. There were significant differences among sand content, silt content, pH and bulk density for the topsoil, and silt content, bulk density and K_s for the subsoil at the Tamale site.

In the pair-wise comparison analysis, Eutric Plinthosol is not considered because of its very small data size (4 for the topsoil and 2 for the subsoil). The pair-wise median test significantly differentiates ($p < 0.05$) between the topsoil sand and silt content and subsoil silt content of the lowland soils (Dystric Gleysol and Eutric Gleysol) and the other soil types at mid-slope (Dystric Plinthosol) and upland soils (Haplic Luvisol, Lithic Leptosol, Ferric Acrisol and Plinthic Acrisol). These differences may be the result of the movement of fine soil particles downslope leaving the large particles behind, thus resulting in the mid-slope and upland soils being sandier compared to flat and valley bottom soils. The significant difference in topsoil pH is the result of differences in soil leaching resulting from differences in position of the soil type as well as land use intensity. The pattern of variation observed for sand, silt and clay content is similar to that observed by Nizeyimana and Bicki (1992), with highest sand content and lowest clay content found on the steeper slope (backslope) on hillslopes in Rwanda.

The difference in bulk density for both top and subsoil is between Plinthic Acrisol and the other soil types. This may be the result of the high concretional content of Plinthic Acrisol compared to the other soil types (see Table 5.2).

The lack of difference in K_s for the topsoil in Tamale may be due to the impact of soil management. In the subsoil, the high K_s of the Plinthic Acrisol compared to the other soil types is due to its high amount of loose iron concretion, which results in higher pore size. The higher amount of fine soil particles in the Dystric Gleysol and Eutric Gleysol also contributes to the low K_s in these soils at the subsoil level. The comparatively high K_s for subsoil of Haplic Luvisol is questionable because of its small sample size.

Table 5.5a. Soil properties classified using median test analysis based on soil types for topsoil (0-15 cm) at Tamale site

[a]Soil type	n	Sand (%) Median (P)[b]	Silt (%) Median (P)	Clay (%) Median (P)	Carbon (%) Median (P)	CEC[e] Median (P)	pH Median (P)	BD (g cm^{-3}) Median (P)	K_s (cm h^{-1}) Median (P)
Haplic Luvisol (1)	5	68.7 (60.0)	29.7 (40.0)	3.0 (40.0)	0.2 (40.0)	2.0 (60.0)	4.9 (60.0)	1.5 (40.0)	2.0 (60.0)
Lithic Leptosol (2)	9	61.6 (77.8)	33.4 (22.2)	4.1 (22.2)	0.5 (55.6)	1.7 (33.3)	5.2 (88.9)	1.5 (77.8)	3.7 (88.9)
Ferric Acrisol (3)	46	57.6 (52.2)	36.1 (43.5)	6.0 (52.2)	0.3 (45.6)	2.3 (67.4)	5.1 (73.9)	1.5 (50.0)	0.9 (47.8)
Plinthic Acrisol (4)	54	62.9 (72.2)	28.2 (27.8)	5.3 (46.3)	0.4 (50.0)	1.6 (46.3)	4.9 (61.1)	1.6 (72.2)	1.9 (59.3)
Dystric Plinthosol (5)	52	56.8 (50.0)	38.1 (55.8)	6.3 (53.8)	0.4 (53.8)	1.8 (50.0)	4.7 (38.5)	1.5 (42.3)	0.9 (46.1)
Eutric Plinthosol (6)	4	63.0 (75.0)	30.2 (25.0)	4.7 (50.0)	0.2 (0.0)	1.7 (50.0)	4.6 (25.0)	1.4 (25.0)	0.8 (50.0)
Eutric Gleysol (7)	65	53.3 (35.4)	41.6 (63.1)	5.6 (47.7)	0.4 (50.8)	1.7 (43.1)	4.7 (29.2)	1.5 (41.5)	0.9 (38.5)
Dystric Gleysol (8)	15	41.6 (0.0)	49.0 (100)	9.4 (66.7)	0.6 (60.0)	1.6 (46.7)	4.6 (26.7)	1.4 (20.0)	0.9 (40.0)
Total	250	56.9	37.2	5.6	0.41	1.8	4.8	1.5	1.3
Median test		35.3[c] (0.00)[d]	35.6 (0.00)	5.46 (0.60)	5.58 (0.59)	8.4 (0.30)	40.0 (0.00)	23.1 (0.00)	11.8 (0.11)
Pair-wise median test (significantly different at $p<0.05$)		1 ≠ 8 2 ≠ 8 3 ≠ 8 4 ≠ 5,7,8 5 ≠ 7,8 7 ≠ 3,5,8 8 ≠1,2,3,4,5,7	2 ≠ 7,8 3 ≠ 7,8 4 ≠ 4,5,7,8 5 ≠ 3,8 7 ≠ 2,3,4 8 ≠ 2,3,4,5,7				2 ≠ 3,5,7,8 3 ≠ 5,7,8 4 ≠ 7 5 ≠ 2,4 7 ≠ 2,3,4 8 ≠ 2,4	2 ≠ 8 3 ≠ 4 4 ≠ 3,5,7,8 5 ≠ 3 7 ≠ 3 8 ≠ 3	

[a]Soil type: (1) to (8) as in table; [b]P: percentage of values greater than median; [c]Chi-Square; [d]Probability; [e]CEC (cmol(+)kg^{-1})

Table 5.5b. Soil properties classified using median test analysis based on soil types for subsoil (30-45 cm) at Tamale site

[a]Soil type	n	Sand (%) Median (P)[b]	Silt (%) Median (P)	Clay (%) Median (P)	Carbon (%) Median (P)	CEC[e] Median (P)	pH Median (P)	BD (g cm^{-3}) Median (P)	K_s (cm h^{-1}) Median (P)
Haplic Luvisol (1)	5	39.1 (40.0)	27.6 (20.0)	28.5 (60.0)	0.3 (40.0)	1.8 (40.0)	4.8 (60.0)	1.5 (60.0)	1.6 (100)
Lithic Leptosol (2)	4	40.9 (50.0)	35.6 (100.0)	16.7 (50.0)	0.5 (75.0)	1.5 (25.0)	4.5 (50.0)	1.5 (50.0)	0.3 (25.0)
Ferric Acrisol (3)	36	48.0 (58.3)	28.3 (38.9)	20.4 (47.2)	0.3 (41.7)	2.3 (55.6)	4.6 (66.7)	1.5 (33.3)	0.8 (55.6)
Plinthic Acrisol (4)	41	46.8 (56.1)	29.1 (29.3)	23.4 (56.1)	0.4 (61.0)	1.9 (48.8)	4.5 (53.7)	1.6 (70.7)	6.1 (85.4)
Dystric Plinthosol (5)	43	38.4 (34.9)	32.6 (51.2)	27.2 (62.8)	0.4 (58.1)	2.3 (60.5)	4.5 (48.8)	1.5 (44.2)	0.5 (39.5)
Eutric Plinthosol (6)	2	54.9 (50.0)	23.8 (0.0)	21.3 (50.0)	0.3 (0.0)	1.2 (0.0)	4.3 (50.0)	1.2 (0.0)	0.6 (50.0)
Eutric Gleysol (7)	61	43.3 (52.5)	35.7 (65.6)	17.6 (36.1)	0.3 (44.3)	2.0 (47.5)	4.4 (39.3)	1.5 (54.1)	0.4 (29.5)
Dystric Gleysol (8)	13	36.4 (46.1)	36.6 (69.2)	23.4 (53.9)	0.3 (38.5)	1.4 (30.8)	4.4 (30.8)	1.5 (30.8)	0.2 (23.1)
Total	205	41.6	32.3	21.4	0.4	2.0	4.5	1.5	0.6
Median test		6.0[c] (0.54)[d]	24.5 (0.00)	8.5 (0.29)	8.8 (0.27)	7.6 (0.37)	9.1 (0.25)	16.2 (0.02)	42.8 (0.00)
Pair-wise median test (significantly different at $p<0.05$)			1 ≠ 2,7,8 2 ≠ 3 3 ≠ 7,8 4 ≠ 7,8					3 ≠ 4,7 4 ≠ 3,5,7,8 5 ≠ 3	1 ≠ 5,7,8 3 ≠ 4 4 ≠ 3,5,7,8 5 ≠ 1,3
			7 ≠ 1,3,4 8 ≠ 1,3,4						7 ≠ 1,3 8 ≠ 1,3

[a]Soil type: (1) to (8) as in table; [b]P: percentage of values greater than median; [c]Chi-Square; [d]Probability; [e]CEC (cmol(+)kg^{-1})

5.2.2 Comparison of soil properties for different soil types at the Ejura site

Table 5.6 presents the median, percentage of values greater than the parameter median (P), group median test analysis (Chi-Square and significance level) and pair-wise median test for the different soil properties at the Ejura site.

The median test resulted in varying Chi-Squares for the different parameters. The analysis gave the highest Chi-Square of 51.7 for the topsoil organic carbon content and the lowest of 1.9 for the topsoil bulk density (Tables 5.6). For all parameters except organic carbon content, the subsoil has higher Chi-Squares than the topsoil. Differentiating soil properties in the subsoil thus is easier than for the topsoil, due to the overriding effect of management practices that lead to mixing of the topsoil. Significant differences were observed among topsoil sand, silt, clay and organic carbon content and among subsoil sand, silt, clay and organic carbon content and soil pH.

Based on the pair-wise median test analysis, the topsoil of Gleyic Fluvisol in the lowland (accumulation) area had a lower sand content than the Acrisols found on the steep slopes of the upland and mid-slope. The Gleyic Fluvisol also had a very high clay content compared to the other soil types due to the washing of fine particles from the soils on the steep slopes to the lowland (Table 5.6a). The Ferrallic Cambisol has a comparable sand content to the Fluvisol because it is found on the more or less flat summit of the topography that experiences minimal soil translocation. Haplic Acrisols had the lowest silt content in the topsoil, which was significantly lower than that for Ferralic Cambisol and Ferric Acrisol on the summit-upper slope, due to less washing away of fine particles. Gleyic Fluvisol in the lowland area had higher silt levels due to the reception of fine particles. Much of the area covered by Gleyic Acrisol and Gleyic Fluvisol were under forest or fallow. This may have given rise to the higher carbon content in those soil types.

In the subsoil, sand and clay contents of Gleyic Fluvisols in the lowland area were comparable to those at the summit and upper slopes (Ferralic Cambisol and Ferric Acrisol), but significantly lower than those at mid-slope (eluviation) area (Haplic and Gleyic Acrisol) (Table 5.6b). Haplic Acrisols on the mid-slope have significantly lower silt content compared to the other soil types. This trend is due to the washing of fine soil particles from the steep slopes to the lowland and valley bottoms. The summits to upper slopes are not affected by this process. The subsoil carbon content for the Gleyic

Table 5.6a. Soil properties classified using median test analysis based on soil types for topsoil (0-15 cm) at the Ejura site

[a]Soil type	n	Sand (%) Median (P)[b]	Silt (%) Median (P)	Clay (%) Median (P)	Carbon (%) Median (P)	CEC[e] Median (P)	pH Median (P)	BD (g cm^{-3}) Median (P)	K_s (cm h^{-1}) Median (P)
Ferralic Cambisol (1)	27	72.4 (37.0)	23.2 (59.2)	4.0 (55.5)	0.6 (48.1)	3.1 (63.0)	5.2 (33.3)	1.4 (55.5)	5.0 (40.7)
Ferric Acrisol (2)	122	71.1 (39.3)	25.1 (62.3)	4.1 (58.2)	0.6 (35.2)	2.9 (53.3)	5.2 (47.5)	1.4 (49.2)	8.0 (50.8)
Haplic Acrisol (3)	134	75.1 (62.7)	21.3 (36.6)	3.5 (42.5)	0.6 (42.5)	2.7 (44.0)	5.2 (47.8)	1.4 (51.5)	7.5 (48.5)
Gleyic Acrisol (4)	74	74.3 (58.1)	22.4 (44.6)	3.0 (36.5)	1.1 (85.1)	2.7 (47.3)	5.4 (62.2)	1.4 (44.6)	9.4 (55.4)
Gleyic Fluvisol (5)	17	68.8 (11.8)	23.2 (58.8)	6.6 (100.0)	0.7 (64.7)	3.0 (64.7)	5.1 (41.2)	1.5 (58.8)	5.3 (41.2)
Total	374	72.8	23.1	3.8	0.7	2.8	5.2	1.4	7.7
Median test		27.9[c] (0.00)[d]	19.3 (0.00)	29.0 (0.00)	51.7 (0.00)	5.9 (0.20)	8.4 (0.08)	1.9 (0.76)	2.5 (0.65)
Pair-wise median test[e] (significantly different at $p<0.05$)		1 ≠ 3 2 ≠ 3,4,5 3 ≠ 1,2,5 4 ≠ 2,5 5 ≠ 2,3,4	1 ≠ 3,4 2 ≠ 3,4 3 ≠ 1,2,5 4 ≠ 1,2 5 ≠ 3	1 ≠ 4,5 2 ≠ 4,5 3 ≠ 4,5 4 ≠ 1,2,3,5 5 ≠ 1,2,3,4	1 ≠ 2 2 ≠ 1,4 3 ≠ 4 4 ≠ 2,3				

[a]Soil type: (1) to (5) as in table; [b]P: percentage of values greater than median; [c]Chi-Square; [d]Probability; [e]CEC (cmol(+)kg^{-1})

Table 5.6b. Soil properties classified using median test analysis based on soil types for subsoil (30-45 cm) at the Ejura site

[a]Soil type	n	Sand (%) Median (P)[b]	Silt (%) Median (P)	Clay (%) Median (P)	Carbon (%) Median (P)	CEC[e] Median (P)	pH Median (P)	BD (g cm^{-3}) Median (P)	K_s (cm h^{-1}) Median (P)
Ferralic Cambisol (1)	19	61.3 (26.3)	23.1 (79.0)	12.3 (63.2)	0.6 (52.6)	2.8 (52.6)	4.4 (36.8)	1.6 (63.2)	3.5 (58.8)
Ferric Acrisol (2)	111	63.4 (30.6)	24.1 (65.8)	11.9 (59.5)	0.6 (46.8)	2.5 (37.8)	4.3 (36.9)	1.6 (55.9)	1.3 (44.0)
Haplic Acrisol (3)	132	70.0 (63.6)	19.0 (31.8)	10.0 (47.7)	0.5 (38.6)	2.8 (55.3)	4.6 (54.5)	1.6 (42.4)	2.3 (56.6)
Gleyic Acrisol (4)	70	72.8 (67.1)	21.1 (45.7)	5.8 (20.0)	0.8 (75.7)	2.8 (57.1)	4.9 (65.7)	1.6 (47.1)	1.6 (48.5)
Gleyic Fluvisol (5)	17	62.3 (23.5)	23.6 (70.6)	12.8 (76.5)	0.6 (47.1)	2.7 (52.9)	4.2 (35.3)	1.6 (52.9)	0.2 (31.2)
Total	349	67.4	21.2	10.6	0.6	2.7	4.5	1.6	1.7
Median test		43.7[c] (0.00)[d]	38.2 (0.00)	35.1 (0.00)	25.9 (0.00)	9.6 (0.05)	18.3 (0.00)	6.1 (0.19)	6.6 (0.16)
Pair-wise median test[e] (significantly different at $p<0.05$)		1 ≠ 3,4 2 ≠ 3,4 3 ≠ 1,2,5 4 ≠ 1,2,5 5 ≠ 3,4	1 ≠ 3,4 2 ≠ 3,4 3 ≠ 1,2,4,5 4 ≠ 1,2,3 5 ≠ 3	1 ≠ 4 2 ≠ 3,4 3 ≠ 2,4,5 4 ≠ 1,2,3,5 5 ≠ 3,4	1 ≠ 4 2 ≠ 4 3 ≠ 4 4 ≠ 1,2,3,5 5 ≠ 4		2 ≠ 3,4 3 ≠ 2 4 ≠ 2		

[a]Soil type: (1) to (5) as in table; [b]P: percentage of values greater than median; [c]Chi-Square; [d]Probability; [e]CEC ($cmol(+)kg^{-1}$)

Acrisol was significantly higher than that of the other soil types as they are mostly in uncultivated areas (Brady and Weil, 1996). The subsoil of Gleyic Fluvisol, though mostly in uncultivated condition, had lower carbon content compared to the Gleyic Acrisol, due to the fact that the area is mostly under saturated conditions. The high pHs for Haplic and Gleyic Acrisol are due to the soils being in the driest parts (steep mid-slopes) of the topography. Except for the subsoil of Ferralic Cambisol, the influence of gravel concretions in the Ejura soils is very minimal at the soil depths considered.

5.2.3 Comparison of variation in soil properties for different soil types at the Tamale and Ejura sites

The analysis resulted in significant differences among the soil types for the soil particle size distribution (sand, silt and clay content) and pH. In general, the level of significance for the subsoil was higher than for the topsoil. This may be due to the minimal disturbance of the subsoil through land management activities. The differences for the Ejura soil properties were higher than at the Tamale site. This is mainly due to the steeper slope at the Ejura site, influencing runoff and soil movement. The differences observed follow the pattern of lateral soil translocation, as clay content is higher and sand content is lower on the lowlands compared to the steep mid-slopes and uplands.

Saturated hydraulic conductivity in most cases showed no differences between the different soil types. The lack of differentiation for K_s is due to the high variability of this property.

5.3 Land use type (LUT)

The term land use type in this study is defined in the context of agriculture and refers to a crop, crop combination or cropping system with a specified technical and socio-economic setting (FAO, 1983). The land use type and method of land preparation are identified by the cultivated crop and land preparation method at the time of sampling. The objective here is to determine the variation among the different LUT in terms of soil properties and their relationship to one another and to identify the main LUTs at the two sites and their possible relationship to the different soil properties. This is in view of the fact that, in the long term, land management has an influence on the soil

properties. The similarities and/or differences among soil physical properties for specific LUT was determined using non-parametric analysis for the same reason as stated in 5.2.

Table 5.7 presents the percentage coverage of the different crops with their botanical names as observed at the sampling points of the two sites. The different crops were grouped based on the crop types for ease of analysis, interpretation and visualization as shown in Figures 5.8 - 5.10. Table 5.7 provides the percentage coverage for uncultivated land and land covered by intercrops or mixed crops. No distinction is made between intercrop and mixed cropping and therefore the two terms are used interchangeably.

Table 5.7. Land use types (crop) with crop coverage (%) at Tamale (2001 cropping season) and Ejura (2002 first cropping season)

Crop (LUT) group	Crop or land use type	Botanical name	Coverage % Tamale	Ejura
Legume	Cowpea	*Vigna unguiculata*	1.6	12.3
	Groundnut	*Arachis hypogaea*	8.0	8.8
	Soybean	*Glycine max*	0.4	-
Upland cereal	Maize	*Zea mays*	8.8	5.1
	Sorghum	*Sorghum bicolor*		0.3
Vegetable	Pepper	*Capsicum spp.*	7.2	
	Okra	*Abelmoschus esculentus*	1.2	0.5
	Cotton	*Gossypium spp.*	1.6	-
Rice	Rice	*Oryza spp.*	11.6	0.3
Root and tuber	Cassava	*Manihot esculenta*	4.8	3.5
	Yam	*Dioscorea spp.*	5.2	6.7
Tree crop	Cashew	*Anacardium occidentale*	-	2.9
Fallow	Uncultivated (fallow, grazing land and forest)		49.6	59.7
	Intercrop with maize		23.0	4.6
	Intercrop with cowpea		1.6	6.6
	Total Mixed/intercrop		29.4	8.6

It was observed that in the 2001 cropping season about 50 % of the surveyed area at Tamale was uncultivated, and consisted mainly of grass and shrub vegetation. Either the area is unsuitable for cultivation and therefore left for grazing purposes or exhausted land, intentionally left to fallow. At the Ejura site, about 60 % of the sampled

area was uncultivated in the first growing season of 2002. The uncultivated land at the Ejura site consists mainly of forest left along the stream, and fallow land.

The four most important crops at the Tamale site in decreasing order of coverage were rice, maize, groundnut, and yam. At the Ejura site, the most important crops were legumes (cowpea and groundnut) covering about 52 % of the cultivated area, followed by yam and maize. Also observed at the Ejura site are cashew plantations, which are of recent development (i.e. less than 10 years). At the Tamale site, trees such as shea nut (*Butryospermum parkii*) and dawadawa (*Parkia clappertoniana*) are scattered across the site and were therefore not considered.

At the Tamale site, 29.4 % of cultivated land involves intercropping as compared to about 8.6 % at the Ejura site. A greater percentage of intercropping at the Tamale site involves maize. The higher rate of intercropping at the Tamale site is an indication of the higher pressure on land at that site as intercropping or mixed farming is mainly used as a management tool on low fertility land. The numerous crops cultivated at the Tamale site compared to that of Ejura site also confirm the pressure on land at the Tamale site. The pressure on land is best seen from the larger percentage of land under cultivation and the need to make maximum use of the single cropping season experienced at the Tamale site.

Figures 5.8 and 5.9 illustrate the LUT for the different crop groups based on different soil types at the Tamale and Ejura sites. The figures show a more even utilization of all soil types at Ejura compared to Tamale. At Tamale it is evident that most of the crop groups can be found on four or more soil types with the legumes seen on all soil types except Dystric Gleysol due to its water-logging characteristics. At the Ejura site, legumes and root and tuber crops can be found on all the soil types, thus indicating more suitable soil conditions for cultivation.

Figure 5.10 shows the use of three different land preparation methods (bullock ploughing, tractor ploughing and manual (hoeing)) for cultivation. The most widely used land preparation method for all crops at the two sites is manual land clearing using cutlass and/or hoe, thus emphasizing the subsistence nature of farming at both sites (Figure 5.10). The second most important land preparation method is tractor ploughing, which is also used for all crops except for root and tuber crops. The use of bullock ploughing is limited to Tamale (Figure 5.10).

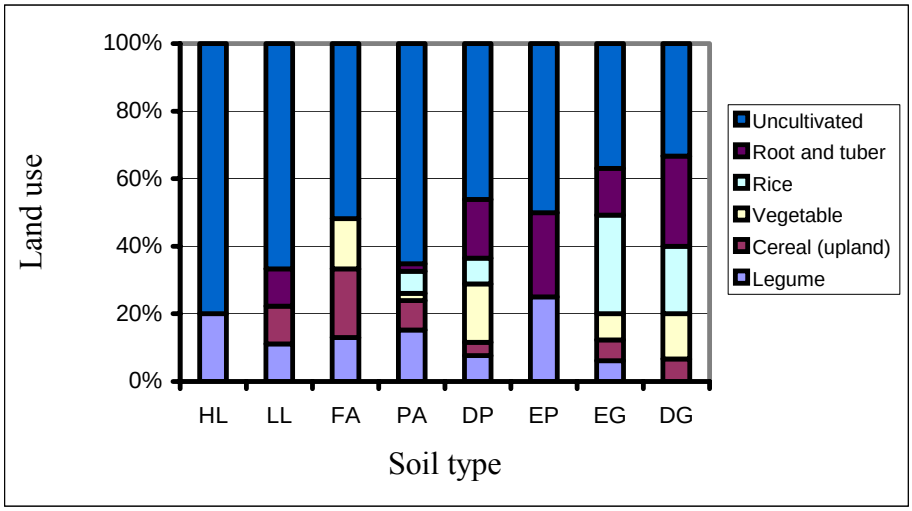


Figure 5.8. Percentage land use type on different soil types (Haplic Luvisol (HL), Lithic Luvisol (LL), Ferric Acrisol (FA), Plinthic Acrisol (PA), Dystric Plinthosol (DP), Eutric Plinthosol (EP), Eutric Gleysol (EG) and Dystric Gleysol (DG)) at the Tamale site

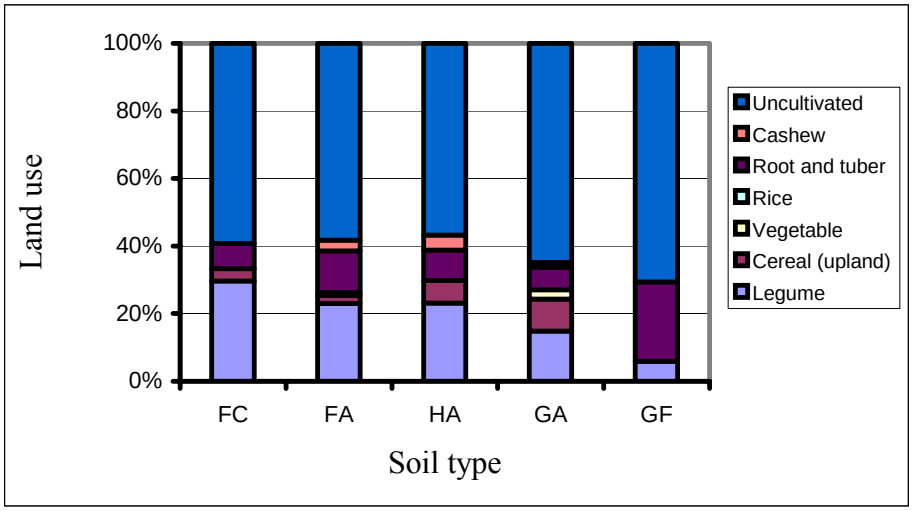


Figure 5.9. Percentage land use type on different soil types (Ferralic Cambisol (FC), Ferric Acrisol (FA), Haplic Acrisol (HA), Gleyic Acrisol (GA) and Gleyic Fluvisol (GF)) at the Ejura site

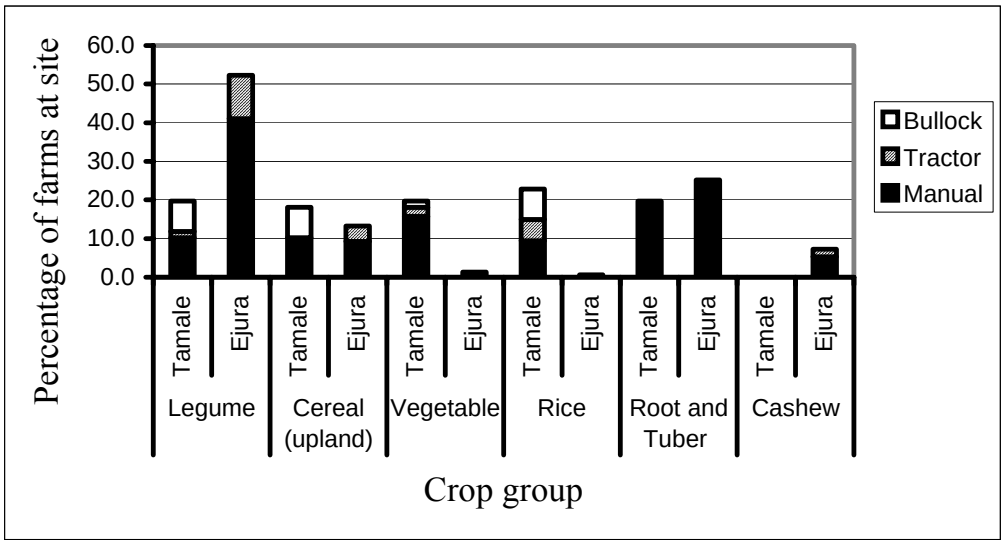


Figure 5.10. Land preparation methods (bullock, tractor and manual) for different crop groups at Tamale and Ejura pilot sites

5.3.1 Distribution of soil properties for different land use types

Though this single season LUT data may not be enough to check for any relationship between LUT and soil properties, it is done to test the method for when long term data is available. The analysis was carried out to determine the differences in the aforementioned soil properties among the different LUT groups listed in Table 5.7. The results of the median test analysis and pairwise comparison for sand, silt, clay and K_s for different land use types at the topsoil and subsoil are presented in Table 5.8 for Tamale and Table 5.9 for Ejura. Particle size distribution parameters were used because of the significant differences observed when they were used for the different soil types as explained in section 5.2 and saturated hydraulic conductivity is considered because it is the dependent variable for further analysis as described in Chapters 6-8.

At the Tamale site, a significant Chi-square for different LUT groups was observed for sand, silt and clay content in the topsoil and sand in the subsoil. It is highest (21.6) for the topsoil silt content and lowest (4.2) for the topsoil K_s (Table 5.8). This suggests the existence of differences in soil properties among LUT groups. In the pairwise comparison, fallow LUT is not considered as it represents vegetation of varying age.

For the Tamale site, the topsoils on which rice and root and tuber crops were planted had significantly lower sand and higher silt content than the soils on which legumes and upland cereals were cultivated. Also topsoils on which vegetables were cultivated had a significantly higher clay content (median = 6.0 %) compared to that of legumes (median = 3.8 %) (Table 5.8).

The subsoils of the Tamale site on which legumes were planted had significantly lower silt content (median = 26.1 %) than those with upland cereals, vegetables, rice and root and tubers crops. Vegetables were planted in soils with higher clay content compared to the other LUT. Subsoils on which legumes were cultivated showed a significantly higher K_s than those with rice, and root and tuber crops.

It is evident from the analysis that at the Tamale site, legumes and upland cereals were planted on soils of high sand and low silt content in the topsoil and subsoil, and thus on soils with good infiltration characteristics in the subsoil. Therefore, the LUT can be used to distinguish areas of uniform sand, silt, and clay content in the topsoil and subsoil at the Tamale site.

At the Ejura site, significant Chi-squares ($p < 0.05$) were obtained with all the investigated parameters for the topsoil, but only with sand and silt content for the subsoil. The highest Chi-square of 21.7 was obtained for topsoil sand and the lowest value of 9.6 for K_s in the subsoil (Table 5.9).

In the pairwise comparison analysis the fallow system was not considered for the same reason as stated earlier for the Tamale site. The vegetables and rice LUT at Ejura were left out because of the small number of data (≤ 2).

At Ejura, legumes, upland cereals and tree crops were cultivated on topsoil of high sand content compared to that with root and tuber crops. The legumes and upland cereals were planted on soils with lower silt content compared to those used for root and tuber crops. Legume were observed to be on topsoils with lower clay content (median = 23.0 %) than those with root and tuber crops (median = 25.2 %) (Table 5.10).

The sand content for subsoils on which tree crops and upland cereals were grown was significantly higher than that used for legumes, and even more so for those soils used for root and tuber crops. Subsoils under tree crops (median = 16.0 %) and upland cereals (median = 19.0 %) had a significantly lower silt content than those under

Table 5.8. Sand, silt, and clay content and saturated hydraulic conductivity classified using non-parametric median test based on land use type for topsoil (0-15 cm) and subsoil (30-45 cm) at Tamale site

	Topsoil median values					Subsoil median values				
[e]Land use type group	n	Sand (%)	Silt (%)	Clay (%)	K_s (cm h^{-1})	n	Sand (%)	Silt (%)	Clay (%)	K_s (cm h^{-1})
Fallow (1)	123	56.4[a] (48.8)[b]	37.4 (50.4)	6.2 (53.7)	1.3 (50.4)	98	41.0 (44.9)	32.2 (49.0)	23.2 (53.1)	1.0 (57.1)
Legume (2)	25	70.4 (76.0)	28.1 (20.0)	3.8 (20.0)	1.8 (56.0)	21	60.9 (81.0)	26.1 (14.3)	11.0 (19.0)	1.4 (66.7)
Upland cereal (3)	23	67.3 (69.6)	28.2 (30.4)	4.8 (34.8)	2.1 (60.9)	18	49.0 (55.6)	33.6 (61.1)	17.4 (44.4)	1.4 (55.6)
Vegetable (4)	25	56.9 (52.0)	37.8 (52.0)	6.0 (60.0)	0.8 (40.0)	19	38.7 (42.1)	31.8 (47.4)	29.4 (73.7)	0.6 (47.4)
Rice (5)	29	50.7 (27.6)	43.6 (72.4)	6.3 (51.7)	0.8 (41.4)	26	42.7 (50.0)	37.1 (69.2)	19.4 (42.3)	0.2 (19.2)
Root and tuber (6)	25	49.1 (36.0)	39.3 (68.0)	7.6 (60.0)	0.9 (40.0)	23	41.1 (43.5)	32.7 (56.5)	22.6 (56.5)	0.4 (26.1)
Total	250	56.9	37.2	5.6	1.26	205	41.6	32.3	21.4	0.6
Median test		18.2[c] (0.00)[d]	21.6 (0.00)	13.8 (0.02)	4.2 (0.52)		10.1 (0.07)	15.9 (0.01)	13.9 (0.02)	19.6 (0.00)
Pair-wise median		1 ≠ 2,3	1 ≠ 2	1 ≠ 2				1 ≠ 2	1 ≠ 2	1 ≠ 5,6
test[c] (significantly		2 ≠ 1,5,6	2 ≠ 1,5,6	2 ≠ 1,4,6				2 ≠ 1,3,5,6	2 ≠ 1,4	2 ≠ 5,6
different at $p<0.05$)		3 ≠ 1,5,6	3 ≠ 5,6					3 ≠ 2	3 ≠ 4	
				4 ≠ 2				4 ≠ 2,5	4 ≠ 2,3,5	4 ≠ 1,2,5
		5 ≠ 2,3	5 ≠ 2,3						5 ≠ 4	5 ≠ 4
		6 ≠ 2,3	6 ≠ 2,3	6 ≠ 2				6 ≠ 2		6 ≠ 1,2

[a]Median value, [b]percentage of values greater than median; [c]Chi-Square; [d]Probability; [e]Land use type: (1) to (6) as in table

Table 5.9. Sand, silt, and clay content and saturated hydraulic conductivity classified using non-parametric median test based on land use type for topsoil (0-15 cm) and subsoil (30-45 cm) at Ejura site

	Topsoil median values					Subsoil median values				
[e]Land use type group	n	Sand (%)	Silt (%)	Clay (%)	K_s (cm h^{-1})	n	Sand (%)	Silt (%)	Clay (%)	K_s (cm h^{-1})
Fallow (1)	233	73.2[a] (50.6)[b]	3.6 (43.8)	23.0 (44.2)	10.7 (54.9)	203	67.4 (50.7)	21.7 (53.2)	10.2 (50.7)	2.0 (53.2)
Legume (2)	79	73.2 (51.9)	4.1 (57.0)	23.0 (48.1)	5.5 (40.5)	71	66.6 (47.9)	21.0 (46.5)	10.6 (46.5)	1.4 (45.1)
Upland cereal (3)	20	74.6 (55.0)	3.1 (40.0)	23.1 (50.0)	3.6 (30.0)	18	72.7 (77.8)	19.0 (33.3)	9.5 (44.4)	1.9 (50.0)
Vegetable (4)	2	77.4 (50.0)	1.4 (0.0)	21.1 (50.0)	6.0 (0.0)	2	76.4 (100.0)	20.1 (50.0)	3.5 (0.0)	0.6 (0.0)
Rice (5)	1					1				
Root and tuber (6)	38	69.2 (18.4)	5.3 (71.1)	25.2 (73.7)	5.3 (36.8)	35	65.4 (37.1)	23.7 (71.4)	10.9 (57.1)	0.9 (37.1)
Tree crop (7)	11	78.3 (81.8)	2.8 (36.4)	18.6 (27.3)	5.4 (45.5)	9	74.0 (88.9)	16.0 (0.0)	9.7 (33.3)	4.8 (77.8)
Total	374	72.8	3.8	23.1	7.7	349	67.4	21.2	10.6	1.7
Median test		21.7[c] (0.00)[d]	14.5 (0.02)	13.1 (0.04)	16.6 (0.01)		14.8 (0.02)	20.0 (0.00)	5.7 (0.46)	9.6 (0.14)
Pair-wise median		1 ≠ 6	1 ≠ 6	1 ≠ 6	1 ≠ 2,3,6		1 ≠ 3,6,7	1 ≠ 7		
test[c] (significantly		2 ≠ 6	2 ≠ 6	2 ≠ 6	2 ≠1		2 ≠ 3,6	2 ≠ 7		
different at p<0.05)		3 ≠ 6	3 ≠ 6		3 ≠1		3 ≠ 1,2,6	3 ≠ 6		
		6 ≠ 1,2,3,7	6 ≠ 1,2,3	6 ≠ 1,2	6 ≠ 1		6 ≠ 1,2,3,7	6 ≠ 3,7		
		7 ≠ 6					7 ≠ 1,6	7 ≠ 1,2,6		

[a]Median value, [b]percentage of values greater than median; [c]Chi-Square; [d]Probability; [e]Land use type: (1) to (6) as in table

root and tuber crops (median = 23.7 %). Also, the tree crops were planted on subsoils of lower silt content compared to the legumes.

In general, at the Ejura site, root and tuber crops were planted on soils of lower sand content and higher clay content compared to the other LUT. This implies that also at the Ejura site, LUT can be used to differentiate soils of similar particle size distribution at both soil depths.

5.4 Conclusion

The soils at the two sites follow a catenary development. The soils at the Ejura site have a more distinct catenary developed compared to that at Tamale, which may be due to the steeper terrain at the Ejura site. Although, the soils at each site were put into eight soil type at Tamale site and five soil types at Ejura site, three main categories of soils emerged from the comparison of the soil properties. These are (1) soils on the uplands (summits and upper slopes) that are more stable with minimal soil translocation effect, (2) mid-slope soils on steep slopes that are influenced by translocation due to the slope and the illuviation clay, and (3) lowland and valley bottom soils that receive soil material. The main mechanisms influencing soil formation are lateral and vertical soil translocation at both sites and plinthization at the Tamale site.

At both sites, it was observed that the soil properties vary in terms of particle size distribution (sand, silt, and clay content) and pH, mainly as a result of soil translocation and leaching. In general, the clay content of the topsoil (0-15 cm) and subsoil (30-45 cm) for the lowland soil was higher than that at the mid-slopes as the fine particles are washed downslope leaving behind the larger particles. Differences in soil pH also resulted from the differential leaching associated with various topographical positions.

The parent material has given rise to a higher sand content at Ejura compared to Tamale. In addition, the most important factors influencing soil type are slope gradient and soil water saturation regime. The parameters K_s, bulk density and CEC in most cases do not vary significantly between soil types for the different soil levels due to the highly heterogeneous nature of the soils within the soil types, which are aggregated as a result of land management – most especially for K_s.

The lowlands and valleys at Ejura are less intensively used for cultivation, due probably to ample land availability, favorable rainfall regime and the two cropping season per year in the area. At the Tamale site, the lowland soils (Eutric and Dystric Gleysol) are the most intensively used, mainly for cultivation of rice and also for other crops such as cassava, pepper and maize due to the availability of soil moisture for longer periods during the growing season (Hanna et al., 1982). Due to sandier conditions and steeper slopes at Ejura, the soils in this area are more liable to nutrient losses through leaching and erosion; however, the high land use intensity at Tamale puts that area at far greater risk to these soil degradation processes.

The significant difference in particle size distribution based on LUT suggests that the choice of land for specific crops takes indirectly into account the soil properties. The smaller Chi-square for particle size distribution based on LUT compared to soil types implies that the latter can give better indication of these soil properties. Thus, LUT can be used for quick distinction of areas of different sand, silt and clay content as this approach is less demanding in terms of time and labor.

6. RELATIONSHIP BETWEEN SATURATED HYDRAULIC CONDUCTIVITY AND SOIL PROPERTIES, SOIL TYPE, LAND USE TYPE AND TERRAIN PARAMETERS

This study set out to identify the relationship between K_s and selected soil properties, soil type, land use type and terrain parameters using stepwise multiple regression (SMR) and generalized linear model (GLM), with data from the Tamale and Ejura sites in the Volta Basin of Ghana. The intention was to identify the key parameters for determining K_s and the relevance of the different parameter groups (soil properties, soil type, land use type and terrain attributes) through the use of the two statistical models.

The SMR and GLM were used to model for K_s using varying data sets (Tamale and Ejura data and combined data from both sites). These methods were selected because of their unique qualities such as the inclusion of input variables with good explanatory power in SMRs and the ability of GLMs to utilize interaction terms. One common advantage of these techniques is the ability to use both categorical and continuous parameters as input variables.

The chapter is divided into four main sections. The first section outlines the methodology used. The second and third sections present the results and discussion of each statistical model used in estimating K_s. The final section presents the conclusion.

6.1 Method of analysis

The importance of each parameter is evaluated based on its coefficient and significance level in the model, R^2 change for the SMR, and the effect size measure (Eta) in the case of GLM. The R^2 change gives the percentage of variance in the dependent variable (K_s) explained by the independent variable. For each statistical model, the adjusted coefficient of determination (R^2) and F-statistic is used as a measure of model performance.

The stepwise multiple regression is a sequential approach to variable selection, and was used because it allows the inclusion of input variables that better explain the response, leaving out parameters that are statistically insignificant or of low explanatory power due to the inclusion of other parameters (Hair et al., 1998). Stepwise regression analysis was done using the data listed in Table 6.1 with variable selection method,

entering and removal of parameters at $p < 0.05$ and $p < 0.1$, respectively, using SPSS 10.0 (SPSS Inc., 1999). The parameters used for the SMR and GLM are presented in Table 6.1 based on the different parameter grouping.

Table 6.1 Parameter groups used for SMR and GLM analyses

Location	*Soil properties*	*Terrain attributes*	*Soil type*	*Land use type group*
Site	Sand	Profile curvature	Haplic Luvisol	Legume
Soil Depth	Clay	Plan curvature	Lithic Leptosol	Upland cereal
	Silt	Curvature	Ferric Acrisol	Vegetable
	CEC	Elevation	Plinthic Acrisol	Rice
	Organic carbon	Wetness index	Dystric Plinthosol	Root and tuber
	Bulk density	Upslope contribution area	Eutric Plinthosol	Tree crop
	pH	Stream power index	Eutric Gleysol	Manual
	Gravel concretion	Slope gradient	Dystric Gleysol	Tractor
	Sabangular blocky structure	LS factor	Ferralic Cambisol	Bullock
	Grainular structure	Aspect	Ferric Acrisol	
	Weak structure		Haplic Acrisol	
	Moderately strong structure		Gleyic Acrisol	
	Strong structure		Gleyic Fluvisol	
	Fine structural size			
	Medium structural size			
	Course and medium structural size			

GLM differs from the well known multiple regressions in two main respects. First, the distribution of the dependent or response variable does not have to be continuous. It also allows categorical or nominal variables as input variables by recoding them into a number of dichotomous variables (Park and Vlek, 2002). Secondly, unlike multiple regressions, which are intrinsically univariate methods, GLM allows linear combinations of multiple independent variables (Park and Vlek, 2002).

This is a great advantage for this study, because it can take into account not only the relationships of the independent variables with the dependent variables, but also the relationships among the multiple independent variables. On the other hand, it may be a disadvantage because an identification of the 'best set' of independent variables may be less meaningful due to the possible increase of multi-colinearity. The GLM was carried out using SPSS 10.0 (SPSS Inc., 1999) and limiting the number of interactions to 2-way. The contribution of significant main and interaction terms were noted

6.2 Stepwise multiple regression (SMR) analysis

The result of the stepwise regression analysis using data from the Tamale site only is presented in Table 6.2. An R^2 of 0.58 was obtained with a significant F-statistic of 41.46. Thus only 58 % of the variance in K_s could be explained. The table presents the parameters in order of decreasing importance, with sand being the most important and the LS factor the least. The parameters CEC, aspect, rice, bulk density, silt, clay, upland cereal and LS factor have negative coefficients and therefore decrease with increasing K_s. The reverse is true for sand, soil depth and type as well as for organic carbon. The most important parameters at the Tamale site are those for soil properties and soil depth.

Table 6.2. Stepwise regression coefficients, standard error, significance level and R^2 change for K_s using data from Tamale site

Parameter	Coefficient	Standard error	Significance level	R^2 change
Constant	1.084	0.219	0.00	
Sand	0.167	0.056	0.00	0.393
Soil depth (topsoil = 1)	0.063	0.017	0.00	0.073
CEC	-0.106	0.027	0.00	0.047
Aspect	-0.039	0.017	0.02	0.015
Rice	-0.033	0.015	0.03	0.010
Bulk density	-0.207	0.043	0.00	0.008
Plinthic Acrisol	0.060	0.014	0.00	0.014
Lithic Leptosol	0.088	0.033	0.01	0.008
Organic carbon	0.088	0.034	0.01	0.005
Silt	-0.624	0.206	0.00	0.005
Clay	-0.113	0.049	0.02	0.005
Upland cereal	-0.038	0.018	0.03	0.005
LS factor	-0.083	0.042	0.05	0.004
F- statistic (N)	41.46 (386)			
R^2	0.58			

Table 6.3 presents the SMR results when Ejura data was used. It shows that the most important parameters are again sand and soil depth but in reverse order, followed by the other soil properties data. About 80 % of the total variance was explained by soil depth. There was a strong R^2 of 0.90 with a significantly high F-statistic of 590.92.

Table 6.3. Stepwise regression coefficients, standard error, significance level and R^2 change for K_s using data from Ejura site

Parameter	Coefficient	Standard error	Significance level	R^2 change
Constant	0.662	0.035	0.00	
Soil Depth (topsoil = 1)	0.130	0.010	0.00	0.804
Sand	0.216	0.026	0.00	0.056
Bulk density	-0.220	0.017	0.00	0.033
Silt	-0.077	0.026	0.00	0.002
Tractor	-0.026	0.008	0.00	0.002
pH	0.043	0.014	0.00	0.001
Aspect	-0.025	0.009	0.01	0.001
Rice	-0.086	0.034	0.01	0.001
CEC	-0.060	0.024	0.01	0.001
Fallow	0.008	0.004	0.05	0.001
F- statistic (N)	590.92 (661)			
R^2	0.90			

Presented in Table 6.4 are the SMR results using combined data from the Tamale and Ejura sites. A strong R^2 (0.89) with highly significant F-value of 531.56 was obtained. The most important parameters are those of soil properties and location (site and soil depth).

The R^2 for the different data groups obtained in the SMR analysis is illustrated in Figure 6.1. It shows that the most important data groups are location (site and soil depth) and the soil properties (sand, silt, and clay content and bulk density). The higher explanatory power for soil depth is mainly due to the difference between the more disturbed topsoil and the less disturbed subsoil. The effect of site on K_s may be the result of differences in parent material and land management practices at the two sites. This high effect of site on K_s explains why it is difficult estimate to K_s using a model calibrated with data from another site.

Table 6.4. Stepwise regression coefficients, standard error, significance level and R^2 change for K_s using combined data from Tamale and Ejura sites

Parameter	Coefficient	Standard error	Significance level	R^2 change
Constant	0.759	0.028	0.00	
Silt	-0.141	0.026	0.00	0.699
Soil Depth (topsoil = 1)	0.095	0.007	0.00	0.118
Sand	0.241	0.018	0.00	0.019
Bulk density	-0.281	0.016	0.00	0.020
Site (Tamale = 1)	-0.186	0.013	0.00	0.019
Plinthic Acrisol	0.068	0.011	0.00	0.005
CEC	-0.105	0.016	0.00	0.003
Lithic Leptosol	0.093	0.025	0.00	0.002
Rice	-0.028	0.011	0.01	0.001
Concretion (present = 1)	0.027	0.011	0.01	0.001
Aspect	-0.024	0.009	0.01	0.001
Haplic Luvisol	0.064	0.024	0.01	0.001
Ferric Acrisol	0.025	0.010	0.01	0.001
Upland cereal	-0.022	0.009	0.02	0.001
Organic carbon	0.041	0.016	0.01	0.001
Stream power index	-0.030	0.014	0.03	0.000
F- statistic (N)	531.56 (1046)			
R^2	0.89			

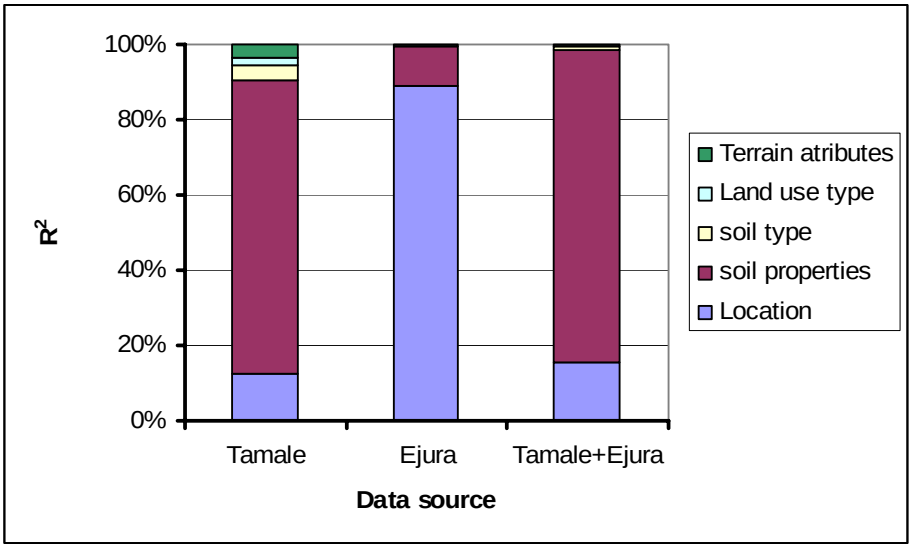


Figure 6.1. Comparison of variation of K_s explained by different data groups from Tamale, Ejura and combined Tamale and Ejura data in SMR

6.3 General Linear Model (GLM) analysis

In the GLM analysis, three comparisons were carried out using data from Tamale, Ejura and combined data from Tamale and Ejura. For each analysis a table (Tables 6.5-6.7) is presented giving the coefficient, standard error, significance level and measure of size effect (Eta) for the constant or intercept, main terms (effect) and the interaction terms. The Eta gives the measure of association between the main or interaction term and the dependent variable, K_s.

Table 6.5 presents the GLM analysis using data from Tamale site. It gives a strong R^2 of 0.64 and F-value of 50.05 compared to an R^2 of 0.58 and F-value of 41.46 when the SMR was used. The improvement in R^2 is mainly due to the inclusion of interaction terms. The most important parameters are soil properties followed by soil depth. In contrast to the SMR analysis, the soil type and land use type show some association with K_s, with a reduced effect of soil properties and location as main terms.

Table 6.5. GLM regression coefficients, standard error, significance level and effect size measure (Eta) for K_s using data from Tamale site

Parameter	Coefficient	Standard error	Significance level	Eta square
Constant	-0.506	0.372	0.18	0.005
Silt	1.005	0.404	0.01	0.016
Bulk density	3.385	0.715	0.00	0.057
Ls factor	0.210	0.132	0.11	0.007
Soil depth (topsoil = 1)	0.074	0.016	0.00	0.054
Lithic Leptosol	0.099	0.031	0.00	0.027
Plinthic Acrisol	0.053	0.013	0.00	0.040
Upland cereal	-0.036	0.017	0.03	0.012
Rice	-0.029	0.014	0.04	0.011
Sand*clay	0.370	0.068	0.00	0.073
Clay*bulk density	-0.474	0.056	0.00	0.163
Silt*bulk density	-3.582	0.774	0.00	0.054
Bulk density*LS factor	-0.670	0.275	0.02	0.016
pH*Aspect	-0.107	0.043	0.01	0.016
F- statistic (N)	50.05 (386)			
R^2	0.64			

Sand*clay: interaction between sand and clay

Table 6.6 presents the GLM analysis with data from Ejura were used. The F-value of 176.59, though significant, is lower than that for the SMR analysis. The R^2 was strong with a value of 0.93. For Ejura, the dominant parameters are the interaction terms

explaining about 62 % of the association between K_s and the independent variables. The remaining variation is mainly explained by the soil depth followed by soil properties with terrain attributes having only a minimal effect (Figure 6.2).

The results for the GLM analysis for the combined Tamale and Ejura data are presented in Table 6.7. An R^2 of 0.91 was obtained with a higher F statistic of 253.77. The Eta as a measure of the association between K_s and the different independent parameters reflecting the interaction terms is considerably reduced from 61 % for the Ejura site and 58 % for the Tamale to about 32 % for the combined data set (Figure 6.2). This is still high compared to the other data groups main terms. It is evident that the interaction terms are critical for estimating K_s at the individual sites but their importance reduces if data from multiple sites are used. Also shown in Figure 6.2 is the importance of location and soil properties data in estimating K_s.

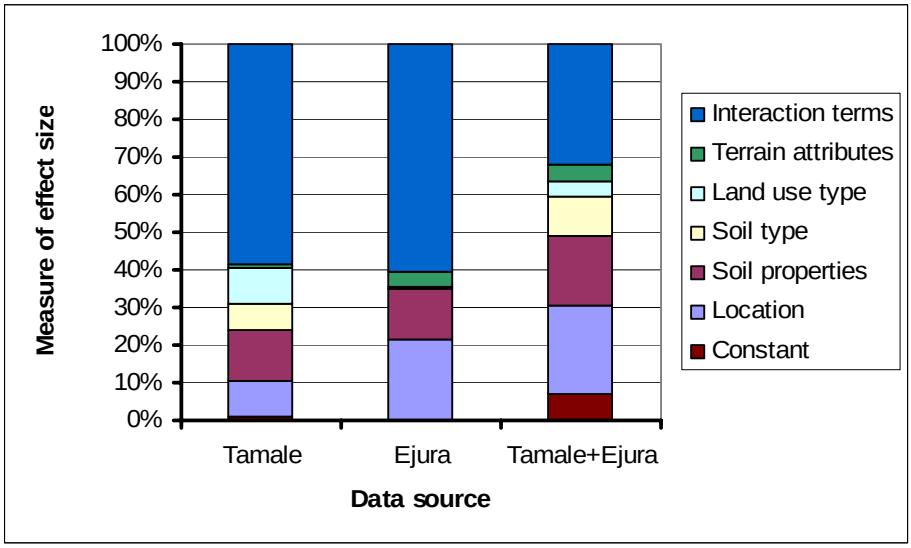


Figure 6.2. Comparison of measure of size effect (Eta) for K_s in terms of different data groups from Tamale, Ejura and combined Tamale and Ejura data in GLM

Of the main terms, the location data (site and soil depth), soil properties (sand, clay, silt, content and bulk density) remain important. The influence of soil type on K_s increases if sites are pooled. The importance of site may be explained by the variation in parent material and differences in land management. The influence of soil depth is mainly the result of differences in soil disturbance of the topsoil and subsoil. The effect

Table 6.6. GLM regression coefficients, standard error, significance level and effect size measure (Eta) for K_s using data from Ejura site

Parameter	Coefficient	Standard error	p-level	Eta square
Constant	-0.071	0.322	0.83	0.000
Sand	0.566	0.269	0.04	0.007
clay	2.432	0.430	0.00	0.049
Silt	1.054	0.304	0.00	0.019
CEC	-2.349	0.491	0.00	0.036
Organic carbon	0.661	0.278	0.02	0.009
Elevation	1.389	0.279	0.00	0.039
Aspect	-0.602	0.315	0.06	0.006
Soil Depth (topsoil = 1)	0.162	0.011	0.00	0.256
Fine structural size	-0.062	0.022	0.00	0.013
Medium structural size	-0.061	0.022	0.00	0.013
Course and medium structural size	-0.072	0.025	0.00	0.013
Rice	-0.076	0.031	0.01	0.010
Sand*clay	-1.637	0.265	0.00	0.058
Sand* organic carbon	0.834	0.274	0.00	0.015
Clay*bulk density	-0.422	0.102	0.00	0.027
Clay*pH	0.480	0.162	0.00	0.014
Clay*elevation	-0.577	0.330	0.08	0.005
CEC*organic carbon	-0.816	0.247	0.00	0.017
CEC*bulk density	0.467	0.111	0.00	0.028
CEC*pH	0.423	0.122	0.00	0.019
Organic carbon*pH	-0.312	0.119	0.01	0.011
Organic carbon*plan curvature	-0.112	0.030	0.00	0.023
Organic carbon*elevation	-1.150	0.193	0.00	0.054
Bulk density*elevation	-0.228	0.082	0.01	0.012
Sand*aspect	0.601	0.205	0.00	0.014
Clay*wetness index	-3.417	0.723	0.00	0.035
Clay*upslope contributing area	3.828	0.888	0.00	0.029
Clay*stream power index	-2.192	0.536	0.00	0.026
Clay*aspect	0.724	0.285	0.01	0.010
Silt*wetness index	-1.993	0.651	0.00	0.015
Silt*upslope contributing area	2.504	0.724	0.00	0.019
Silt*stream power index	-1.512	0.417	0.00	0.021
Silt*slope	-0.749	0.190	0.00	0.025
Silt*LS factor	0.655	0.119	0.00	0.047
Bulk density*stream power index	-0.162	0.069	0.02	0.009
CEC*wetness index	5.138	0.930	0.00	0.047
CEC*upslope contributing area	-6.376	1.164	0.00	0.046
CEC*stream power index	3.619	0.709	0.00	0.041
Organic carbon*aspect	-0.272	0.103	0.01	0.011
pH*wetness index	-0.475	0.257	0.06	0.006
pH*upslope contributing area	0.750	0.299	0.01	0.010
pH*stream power index	-0.430	0.179	0.02	0.009
pH*aspect	-0.161	0.064	0.01	0.010
Elevation*aspect	-0.273	0.097	0.01	0.013
F- statistic (N)	176.59 (661)			
R^2	0.93			

Sand*clay: interaction between sand and clay

Table 6.7. GLM regression coefficients, standard error, significance level and effect size measure (Eta) for K_s using combined data from Tamale and Ejura

Parameter	Coefficient	Standard error	Significance level	Eta square
Constant	2.093	0.272	0.00	0.056
Sand	-0.914	0.184	0.00	0.024
clay	-1.089	0.228	0.00	0.022
Silt	-1.371	0.260	0.00	0.027
CEC	0.198	0.057	0.00	0.012
Organic carbon	0.758	0.238	0.00	0.010
Bulk density	-0.688	0.149	0.00	0.021
pH	-1.275	0.246	0.00	0.026
Curvature	0.069	0.040	0.09	0.003
Stream power index	-0.128	0.030	0.00	0.017
Slope	-0.068	0.037	0.07	0.003
LS factor	0.201	0.051	0.00	0.015
Site (Tamale = 1)	-0.161	0.021	0.00	0.054
Soil Depth (topsoil = 1)	0.127	0.010	0.00	0.127
Course and medium structural size	-0.019	0.010	0.06	0.003
Haplic Luvisol	0.050	0.023	0.03	0.005
Lithic Leptosol	0.125	0.023	0.00	0.028
Ferric Acrisol	0.027	0.010	0.01	0.007
Plinthic Acrisol	0.070	0.010	0.00	0.043
Root and tuber	0.015	0.008	0.05	0.004
Manual	-0.018	0.005	0.00	0.013
Tractor	-0.033	0.009	0.00	0.013
Sand*clay	0.556	0.090	0.00	0.037
Sand*silt	0.668	0.132	0.00	0.025
Sand*CEC	-0.202	0.069	0.00	0.009
Sand* organic carbon	-0.443	0.156	0.00	0.008
Sand*bulk density	0.426	0.101	0.00	0.017
Sand*pH	0.934	0.155	0.00	0.035
Clay*silt	0.894	0.205	0.00	0.019
Clay*organic carbon	-0.428	0.130	0.00	0.011
Clay*pH	0.620	0.153	0.00	0.016
Silt*organic carbon	-0.478	0.170	0.01	0.008
Silt*pH	0.504	0.136	0.00	0.013
CEC*bulk density	-0.291	0.090	0.00	0.010
Organic carbon*bulk density	0.240	0.086	0.01	0.008
Organic carbon*plan curvature	-0.126	0.058	0.03	0.005
Sand*aspect	-0.086	0.027	0.00	0.010
Silt*bulk density	0.330	0.098	0.00	0.011
Bulk density*LS factor	-0.125	0.057	0.03	0.005
Bulk density*Aspect	0.049	0.029	0.09	0.003
F- statistic (N)	253.77 (1047)			
R^2	0.91			

Sand*clay: interaction between sand and clay

of land use type and terrain attributes remains low. The low effect of land use type on K_s may be the result of using current land use type rather than the past land use type that may have influence the soils' development. Though the influence of soil type, land use type and terrain on K_s is rather low compared to that of location and soil properties, their inclusion in estimation models does improve the estimation of K_s.

6.4 Conclusion

Stepwise multiple regressions and a generalized linear model were used to identify the key data groups (location, soil properties, soil type, land use type and terrain) as independent variables for estimating K_s. For this purpose three data sets were considered, i.e. data from the Tamale site, Ejura site and combined data from the two sites.

Two important data groups were observed for estimating K_s: location (site and soil depth) and soil properties (sand silt, and clay content and bulk density). The importance of location data may be due to their ability to capture differences in soil parent material, management practices and soil disturbance. The parameters of particle size distribution (sand, silt and clay content) and bulk density were found to be very important in estimating K_s as these directly influence soil pore size and distribution, which in turn determine how water flows through the soil and therefore K_s.

The level of influence of soil type, land use type and terrain in estimating K_s was observed to be low. The application of historical (past) land use type may influence K_s more strongly than current land use type used in this study. However, the inclusion of these parameters in the estimation of K_s may improve the result, though they can not be relied upon as the main independent variables for such estimation. With the improvement in high resolution image acquisition, these parameters are however fairly easy to obtain compared to data on soil properties, and may have a better influence on K_s when a larger area (such as regional scale) is considered.

Through the GLM analysis it was observed that interaction terms can play a key role in K_s estimation. Therefore, the inclusion of interaction terms in K_s models will improve their performance.

7 EVALUATION OF PEDO-TRANSFER FUNCTIONS

Increasing use of mathematical models in many fields of study (such as ecology, hydrology and climatology) has led to a growing need for accurate soil hydraulic properties (Rawls et al., 1998). However, widespread measurement of soil hydraulic properties is not only time consuming but also cost prohibitive (Schaap et al., 1999). Moreover, saturated hydraulic conductivity (K_s), the hydraulic parameter of interest in this study, is highly variable and numerous laboratory and field methods exist for determining it; however, none is satisfactory for all soil types under all conditions.

As a result, a great deal of work has been carried out in the past using soil morphological data as a surrogate for estimating hydraulic properties (Bouma and van Lenen, 1987; Bouma, 1989 and Rawls et al., 1992). It is reasonable, therefore, to seek useful relations between soil hydraulic properties and easily obtainable or existing information such as texture for areas in which the soils have similar mineralogy and genesis (Anderson and Bouma, 1977; Bouma et al., 1979; King and Franzmeier, 1981 and McKeague et al., 1982). The pedotransfer function (PTF) approach, involves the use of different input data and model structure to estimate hydraulic properties of un-sampled areas (Williams et al., 1992 and Schaap et al., 1998). Existing PTFs, using soils of our area of interest (Volta Basin), need to be evaluated before their adoption. Most of the work done in this field involves soils from the "Temperate World" and, as elaborated in Chapter 5, in the Volta Basin even soils within a kilometer are highly variable.

In the following, selected PTFs using soil data taken from the Tamale and Ejura pilot sites in the Volta Basin of Ghana are evaluated with the aim of identifying a suitable method for estimating K_s at the pilot sites and possibly in the Volta Basin as a whole.

The chapter is divided into five parts, providing an overview of K_s, an overview of PTFs, methods used in estimating and comparing PTFs, comparison of the different PTFs models, and a conclusion regarding this approach.

7.1 Saturated hydraulic conductivity

Saturated hydraulic conductivity is the measure of the soil's ability to transmit water under saturated conditions (Klute and Dickson, 1986). Saturated hydraulic conductivity is a very crucial parameter used in determining infiltration, irrigation practice, drainage design, runoff, groundwater recharge and in simulating leaching and other agricultural and hydrological processes. Knowledge of a soil's hydraulic properties is of major importance in modern agriculture, since these properties influence plant growth, soil aeration, soil temperature, drainage, irrigation, and trafficability.

Flow of water through the soil has to go through irregularly shaped, tortuous, and intricately interconnected pores that are limited by constrictions with occasional "dead-end" spaces (Hillel, 1998). Consequently, the fluid velocity varies from point to point even along the same passage, hence making it impossible to describe flow through soil at the microscopic level. Therefore the flow of any fluid through a soil is considered over a given volume of soil with the flow averaged over the cross-section of the soil, thus yielding a macroscopic flow-velocity. This approach has the implicit assumption that the soil volume taken is sufficiently large relative to the pore sizes and microscopic heterogeneities to permit the averaging of velocity over the cross-section (Hillel, 1998).

The flow of a fluid through a soil column with the flux density or specific discharge is directly proportional to the change in hydraulic head ΔH, and inversely proportional to the length of the column L with a constant of proportionality K_s (saturated hydraulic conductivity) given by Darcy's law:

$$q = -K_s \Delta H / L \qquad [7.1]$$

Saturated hydraulic conductivity is affected mainly by soil texture (Tieje and Hennings, 1996 and Bloemen, 1980), structure, and porosity. It is a function of particle size distribution, pore size distribution, pore continuity and configuration, and bulk density. Saturated hydraulic conductivity is a highly variable parameter that depends to a large extent on the measuring method, scale of measurement (field or laboratory, size of the measurement volume) and spatial variability (Tietje and Richter, 1992).

7.2 Pedo-Transfer Functions (PTFs)

Many studies in the past have investigated the PTF approach, which serves to translate through empirical regression or functional relationships the basic information obtained through soil survey (such as texture, organic matter content and bulk density) into parameters (such as hydraulic conductivity and moisture retention parameters) with broader application (Bouma and van Lanen, 1987; Bouma, 1989; McKeague et al., 1982; McKenzie MacLeod, 1989; Williams et al., 1992 and McKenzie et al., 1991).

A variety of PTFs with different mathematical concepts, estimation properties and input data requirements have been developed in the past. This may be empirically based (Mualem, 1976; van Genuchten, 1980) or physically based (Campbell, 1985; Brutsaert, 1967; Arya and Paris, 1981; Haverkamp and Parlange, 1986; and Rieu and Sposito, 1991) with the vast majority being empirically based on relatively simple linear regression equations.

The empirically based PTF models for estimating saturated hydraulic conductivity generally include the use of multiple linear and nonlinear regression approaches such as those derived by Cosby et al. (1984), Brakensiek et al. (1984), Saxon et al. (1986), Vereecken et al. (1990) and Pucket et al. (1985). These methods are further described in Section 7.3. Input parameters in this approach are mainly sand, silt, clay, organic matter content and bulk density.

The physically based PTF uses the functional relationship between the basic properties of the soil and saturated hydraulic conductivity. Examples of this approach are the models by Campbell (1985), Bloemen (1980), Rawls et al. (1993), Ahuja et al. (1984), and Brutsaert (1967). The equations for these models are provided in Section 7.3.

Many PTFs have been developed, but due to differences in data distribution and high soil variability (Schaap and Leij, 1998 and Teitje and Richter, 1992) all need to be calibrated on existing data before true estimations can be made. Inherent uncertainty exists in PTF estimations due to high variation of soil properties. Thus, PTFs developed on the basis of one population of soil data cannot automatically estimate K_s for another set of soil data, which often require entirely different PTFs (Schaap and Leij, 1998).

PTFs may be site specific or applicable to a particular range of soil types for which they were developed, and therefore applicable to a given range of soil particle

sizes (Espino et al., 1995). Inaccuracies in K_s-values using PTFs are mainly due to the inherent variability of the saturated hydraulic conductivity as a result of spatial variability, measurement and sampling errors (Tietje and Hennings, 1996). Therefore PTFs based on different K_s determination approaches applied to the same data set are likely to yield different results.

7.2.1 Review of selected PTF analysis results

According to Pucket et al. (1985), K_s exhibits extreme variability and is difficult to estimate. Their K_s values ranged from 3.0×10^{-4} to 14.18 cmh^{-1}. Fine sand and sandy clay percentages are highly correlated with K_s. The authors indicated that non-linear regression procedure gives the best estimates for K_s. Deviations were observed for soils with porosity greater than 0.42. Their analysis indicated that, in general, soils with 20% or more clay and lacking macropores might have K_s lower than 7.2 cmh^{-1}.

It is possible to calculate conductivity functions of any soil, provided that an adequate analysis of granular composition and humus or organic matter content analysis is available (Bloemen, 1980). Bloemen (1980) found no significant difference between K_s values for horizontally and vertically sampled soils.

Rawls et al. (1992), using a data set of 52 observations, estimated K_s with R^2 of 0.42 using the PTF by Brutsaert (1967). They recommended the use of Ahuja et al (1984) method for estimating K_s for a wide range of soils, especially for soils with less than 65 % sand and less than 40 % clay.

Tietje and Hennings (1996) evaluated the performance of six PTFs including Brakensiek et al. (1984), Saxton et al. (1986), Vereecken et al. (1990), Cosby et al. (1984), Campbell (1985), and Bloemen (1980), using 1161 soil samples from the north-western part of Germany. They determined K_s using the falling head method with horizontally oriented soil core volume of 250 cm^3. The geometric standard deviation of error ratio (GSDER) for the different methods evaluated by Tietje and Hennings (1996) differs only slightly between most of the PTFs, with the Bloemen (1980) PTF yielding the largest deviations between the estimated and measured K_s. The GSDER ranged from about 4 to 20. The method by Cosby et al. (1984) performed best for most of the investigated textural classes, based on its low geometric mean error ratio (GMER) (Tietje and Hennings, 1996).

Schaap et al. (1998) in comparing the performance of different published PTFs using 620 soils observed for R^2 = 0.30 (Cosby et al., 1984), R^2 = 0.43 (Saxton et al., 1986), R^2 = 0.22 (Vereecken et al., 1990), R^2 = 0.42 (Brakensiek et al., 1984) and R^2 = 0.54 (Ahuja et al., 1984). The R^2 obtained using Cosby and Saxton methods found by Schaap et al (1998) were not as good as those reported by the researchers themselves (0.84 by Cosby et al. (1984), and 0.95 by Saxton et al. (1986)).

Grouping PTFs by horizons improves estimates, especially in the subsoil (Pachepsky et al. 1996). This is because the relationship between the input and output parameters in the subsoil are less affected by land management practices.

7.3 PTFs models and estimation error for saturated hydraulic conductivity

For the purpose of identifying suitable PTFs for estimating K_s in the Volta Basin of Ghana, different PTFs were used to estimate K_s. The resulting estimated saturated hydraulic conductivity (K_{se}) was compared to the measured saturated hydraulic conductivity (K_{sm}) using methods explained in section 7.3.1. Equations for the different methods are presented below:

Method A: Cosby et al. (1984)

$$K_s = 2.54 \cdot 10^{(-0.6+0.0126s-0.0064c)} \qquad [7.2]$$

Method B: Brakensiek et al. (1984)

$$K_s = \exp(x) \qquad [7.3]$$

where

$$x = 19.52348 \cdot \phi - 8.96847 - 0.028212 \cdot c + 0.00018107 \cdot s^2 - 0.0094125 \cdot c^2 - 8.3952 \cdot \phi^2 + 0.077718 \cdot s \cdot \phi - 0.00298 \cdot s^2 \cdot \phi^2 - 0.019492 \cdot c^2 \cdot \phi^2 + 0.0000173 \cdot s^2 \cdot c + 0.02733 \cdot c^2 \cdot \phi + 0.001434 \cdot s^2 \cdot \phi - 0.000035 \cdot c^2 \cdot s$$

Method C: Saxon et al. (1986)

$$K_s = \exp\left(\begin{array}{c} 12.012 - 7.55\cdot 10^{-2}\cdot s + \\ \dfrac{-3.895 + 3.671\cdot 10^{-2}\cdot s - 0.1103\cdot c + 8.7546\cdot 10^{-4}\cdot c^2}{0.332 - 7.251\cdot 10^{-4}\cdot s + 0.1276\cdot \log_{10}(c)} \end{array}\right) \qquad [7.4]$$

Method D: Vereecken et al. (1990)

$$K_s = 0.0417\cdot \exp(20.62 - 0.96\cdot \ln(c) - 0.66\cdot \ln(s) - 0.46\cdot \ln(m) - 8.43\cdot b_d) \qquad [7.5]$$

Method E: Jabro (1992)

$$K_s = \exp(9.56 - 0.81\cdot \log(u) - 1.09\cdot \log(c) - 4.64\cdot b_d) \qquad [7.6]$$

Method F: Pucket et al. (1985)

$$K_s = 11.336\cdot \exp^{-0.1975\cdot c} \qquad [7.7]$$

Method G: Campbell (1985)

$$K_s = 14.125\left(1.3/b_d\right)^{1.3b}\cdot \exp(-6.9\cdot c - 3.7\cdot u) \qquad [7.8]$$

$$b = GMPS^{-0.5} + 0.2\cdot GSD$$

Method H: Bloemen (1980)

$$K_s = 0.000833\cdot M_d^{1.93}\cdot f^{-0.74} \qquad [7.9]$$

The grain size distribution index (f) is calculated from the basic data for the cumulative grain size distribution curve as a succession of weight percentages measured at a number of size interval limits. The slope of a distribution curve between two of these limits can easily be calculated as the tangent to the abscissa of a straight line through the two data points, if the ordinate and abscissa have the same scale. Using a

dimensionless log scale on both axes, the size interval limits (S_i) are plotted on the abscissa and the cumulative weight percentages (P_i) on the ordinate. The mean grain size distribution index (f) between the cumulative weight percentages of the lower and upper size limit of the analysis (P_1 and P_n) would be:

$$f = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} (P_{i+1} - P_i)} \qquad [7.10]$$

where

$$f_i = (P_{i+1} - P_i) tg_i \qquad [7.11]$$

with the slope between two limits as

$$tg_i = \frac{\log\left(P_{i+1} / P_i\right)}{Log\left(S_{i+1} / S_i\right)} \qquad [7.12]$$

See Table 7.1 for a sample calculation of the grain size distribution index f from textural data of a soil with 37.4 % sand, 19.6 % silt and 43 % clay.

Table 7.1. Sample calculation of grain size distribution index (f) of clay texture (37.4 % sand, 19.6 % silt and 43 % clay)

i	S_i (µm)	P_i	$P_{i+1} - P_i$	$\log \frac{S_{i+1}}{Si}$	$\log \frac{P_{i+1}}{P_i}$	f_i eq. (6.10)
1	2	43				
2	50	62.6	19.6	1.398	0.163	2.287
3	2000	100.0	37.4	1.602	0.203	4.749
Total			57.0			7.036
		Eq. (6.13): $f = 7.036/57.0 = 0.123$				

Method I: Rawls et al. (1993)

$$K_s = 4.41 \cdot 10^7 \left(\frac{\phi^{4/3}}{d^2} \right) \cdot R_1^2 \qquad [7.13]$$

where $R_1 = 0.148 \diagup \varphi$

$d = 1.86(2-\lambda)^{5.34}$

Method J: Ahuja et al. (1984)

$$K_s = 1015(\phi - \theta_{33})^4 \qquad [7.14]$$

Method K: Brutsaert (1967)

$$K_s = 3600a \frac{\phi_e^{\ 2}}{\varphi_b^{\ 2}} \cdot \frac{\lambda^2}{(\lambda+1)(\lambda+2)} \qquad [7.15]$$

where

K_s = saturated hydraulic conductivity (cmh^{-1})
s = % sand (50μm -2000μm)
u = % silt (2μm-50μm)
c = % clay (<2μm)
b_d = bulk density (gcm^{-1})
GMPS = geometric mean particle size (mm) (Shirazi and Boersma, 1984)
GSD = geometric standard deviation (dimensionless)
m = organic matter content (%)
θ_{33} = moisture content at –33kPa
λ = pore size distribution index (Table 7.2)
φ_b = geometric mean bubbling pressure (cm) (Table 7.2)
ϕ = porosity ($1-b_d/2.65$)
ϕ_e = effective porosity (total porosity minus residual soil water at –33kPa) ($cmcm^{-1}$)
a = constant (86) (Rawls et al. 1982)
M_d = median grain size
f = grain size distribution index
S_i = cumulative grain size
P_i = cumulative grain weight

Table 7.2 presents the hydraulic parameters for different soil textural classes as given by Rawls et al., (1982).

Table 7.2. Hydraulic and grain properties for different soil textural classes

Soil texture	Bubbling pressure[1] cm	Pore size distribution index	Water content at –33kPa cm^3cm^{-3}	Residual water content cm^3cm^{-3}	Median grain size μm
Sand	7.26	0.694	0.091	0.020	
Loamy sand	8.69	0.553	0.125	0.035	465.50
Sandy loam	14.66	0.378	0.207	0.041	226.80
Loam	11.15	0.252	0.270	0.027	82.55
Silt loam	20.76	0.234	0.330	0.015	86.90
Sandy clay loam	28.08	0.319	0.255	0.068	149.70
Clay loam	25.89	0.242	0.318	0.075	35.55
Silty clay loam	32.56	0.177	0.366	0.040	
Sandy clay	29.17	0.233	0.339	0.109	
Silty clay	34.19	0.160	0.387	0.056	
Clay	37.30	0.165	0.396	0.090	18.95

[1]Geometric mean; Source: Rawls et al., (1982).

7.3.1 Methods of comparing models and estimating model errors

Model accuracy is determined by comparing estimated saturated hydraulic conductivity (K_{se}) to the measured saturated hydraulic conductivity (K_{sm}) in a linear regression for the whole data set and for different textural classes using the coefficient of determination (R^2) and ANOVA.

The errors involved in the estimation of K_{se} compared to K_{sm} are evaluated using the geometric mean error ratio (GMER) and geometric standard deviation of error ratio (GSDER). This is done using the error ratio ε, or estimated to measured saturated hydraulic conductivity. Quantitatively, GMER and GSDER are given as (Tieje and Hennings, 1996):

$$GMER = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln(\varepsilon_i)\right) \qquad [7.16]$$

with $\varepsilon = K_{se} \big/ K_{sm}$

where K_{se} and K_{sm} are estimated and measured K_s, respectively

$$GSDER = \exp\left[\left(\frac{1}{n-1}\sum_{i=1}^{n}\left[\ln(\varepsilon_i) - \ln(GMER)\right]^2\right)^{1/2}\right] \qquad [7.17]$$

7.4 Performance of different PTFs

Presented in Table 7.3 is the coefficient of determination (R^2) for different PTFs using data from Tamale, Ejura and the combined data from the two sites. The results show that for most of the PTFs, combining the data from the two sites gives a better R^2 than when data from only one site is used. This improvement may be due to the lager data range for K_s. Consequently, all further analyses in this section consider the combined data from the two sites.

Table 7.3. Coefficient of determination (R^2) for comparing estimated K_s (PTFs) to measured K_s using data from Tamale, Ejura and combined data

	Data source		
PTF	Tamale	Ejura	Tamale + Ejura
Cosby	0.25	0.14	0.31
Brakensiek	0.21	0.31	0.31
Saxon	0.19	0.16	0.26
Vereecken	0.09	0.30	0.19
Jabro	0.14	0.33	0.31
Pucket	0.18	0.17	0.23
Campbell	0.27	0.31	0.38
Bloemen	0.20	0.10	0.27
Rawls	0.22	0.27	0.34
Ahuja	0.18	0.35	0.32
Brutsaert	0.23	0.24	0.35
N	421	705	1126

7.4.1 Overall model fit

The regression analysis presented in Figure 7.1 illustrates the relationship between the measured saturated hydraulic conductivity (K_{sm}) and the estimated saturated hydraulic conductivity (K_{se}) from different PTFs using R^2 as the measure of model accuracy. For most of the PTFs, the distribution of points is scattered with varying patterns along the regression lines for different PTFs. Three main scatter point patterns can be observed.

The first group of patterns (Jabro and Vereecken) illustrates a cloud of points at a central location along the regression line. They have a poor distribution of the points along the regression line, resulting in low R^2. The second distribution pattern (Cosby, Brakensiek, Saxon, Pucket, Campbell and Ahuja), to which most of the models belong illustrates a bimodal point cloud along the regression line. It shows a larger and smaller

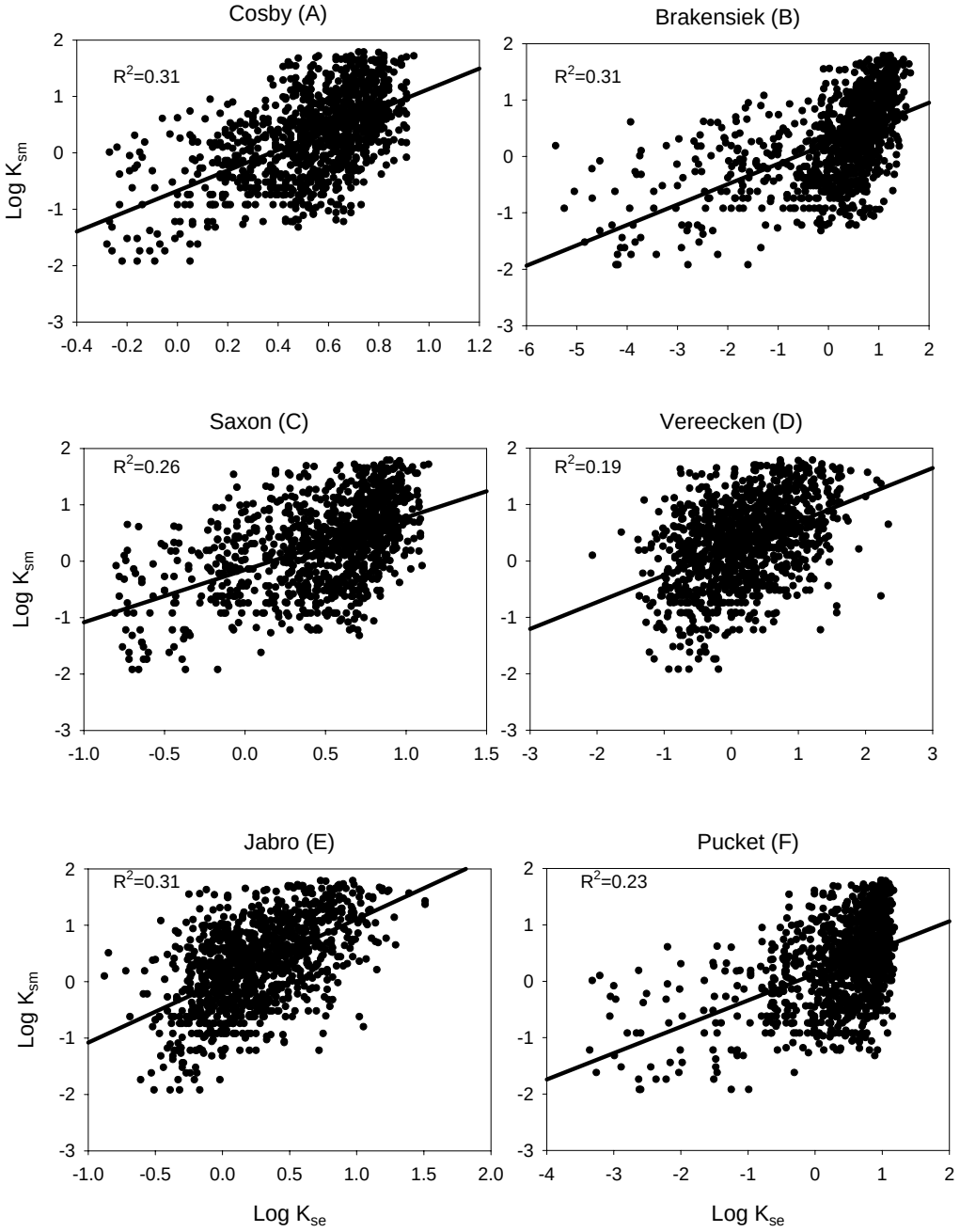


Figure 7.1a. Regression of measured saturated hydraulic conductivity (K_{sm}) to estimated saturated hydraulic conductivity (K_{se})

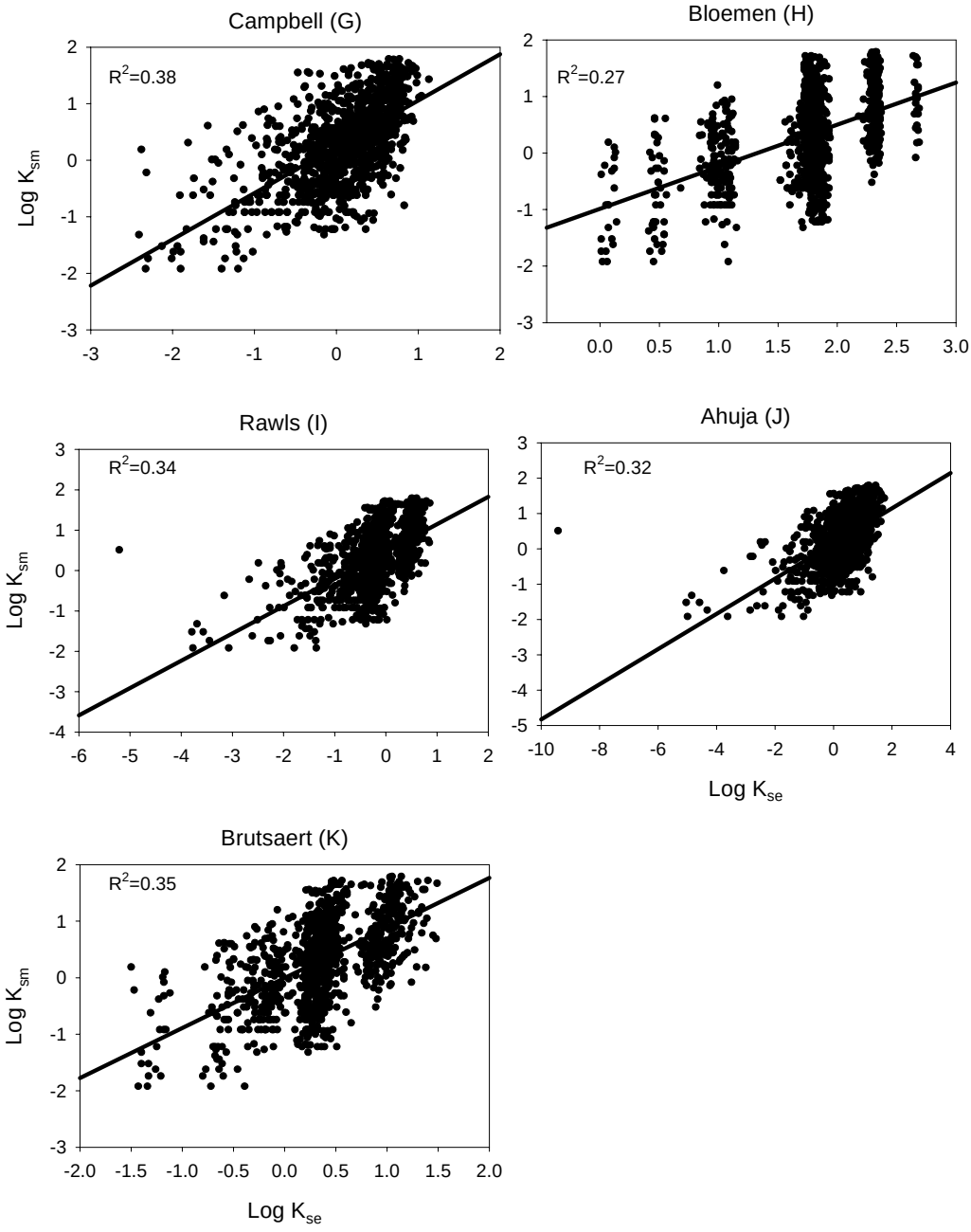


Figure 7.1b. Regression of measured saturated hydraulic conductivity (K_{sm}) to estimated saturated hydraulic conductivity (K_{se})

cloud of points around two locations along the regression line. The bimodal nature of the points negatively affects the resulting R^2. The third group of distribution pattern (Bloemen, Rawls and Brutsaert), shows classes or groups of scatter points produced by

the models. These patterns mainly reflect classes of soil texture, as these three models are highly dependent on soil texture.

The generally poor distribution of points resulted in low R^2 values in all PTFs. The R^2 observed ranged from 0.19 to 0.38. The four PTFs with highest R^2 in order of decreasing R^2 were Campbell (R^2=0.38), Brutsaert (R^2=0.35), Rawls (R^2=0.34), and Ahuja (R^2=0.32). The R^2 for Cosby (0.31), Saxton (0.26) and Vereecken (0.19) are comparable to that obtained by Schaap et al. (1998).

The different distribution patterns highlight the fact that overall R^2 cannot be used solely as a basis for selecting one model in preference to the other. In view of this, in the subsequent sections the different models are compared based on textural classes and errors involved in estimating K_s. The poor distribution patterns for the PTFs that lead to poor R^2 are not solely due to the models. Saturated hydraulic conductivity is difficult to measure and highly variable as described in Chapter 4. Considering the aforementioned problems associated with the K_s, the $R^2 > 0.25$ obtained for most of the models are appreciable; therefore some of the models can be selected for estimating K_s.

7.4.2 Model comparison for different soil textural classes

Table 7.4 presents the correlation between K_{sm} and K_{se} based on different PTFs for different soil textures. The R^2 is used to identify model(s) with a significant correlation at the 0.05-level between measured and estimated K_s.

The PTF methods of Vereecken, Jabro, Campbell, Rawls, Ahuja, and Brutsaert showed a significant positive correlation at the 0.05-level or better for either five or six textural groups. The remaining five methods showed a significant positive correlation at the 0.05-level or better for only three or less textural groups. Most of the PTFs had a significant positive correlation at the 0.05-level or better for loamy sand, sandy loam, clay loam and silt loam.

No model had a significant positive correlation for the sandy clay loam texture, due partly to the low sampling size. For clay texture, where only two PTFs (Rawls and Ahuja) had significant positive correlation. However, six PTFs (Vereecken, Jabro, Campbell, Rawls, Ahuja, and Brutsaert) showed a significant positive correlation at the 0.05-level for sand texture despite the low sample size. In the case of loam texture, three

PTFs (Campbell, Vereecken and Jabro) showed a significant positive correlation at the 0.05-level.

Table 7.4. Correlation coefficient of measured to estimated log K_s for different PTFs methods based on soil textural classes

Texture	[1]Loamy Sand	Sandy Loam	Clay Loam	Silt Loam	Sandy Clay L	Loam	Clay	Sand
n	(233)	(634)	(39)	(37)	(19)	(121)	(21)	(22)
PTFs								
Cosby	0.04	0.26**	0.05	0.32	-0.04	0.10	0.04	0.16
Brakensiek	0.46**	0.39**	0.28	0.50**	-0.07	0.13	-0.08	0.38
Saxon	0.08	0.20**	-0.02	0.35*	0.04	0.06	-0.40	0.11
Vereecken	0.38**	0.32**	0.57**	0.33*	0.28	0.18*	0.22	0.42*
Jabro	0.45**	0.39**	0.50*	0.42**	0.27	0.20*	0.31	0.47*
Pucket	0.09	0.16**	0.00	0.31	0.03	0.07	-0.31	0.19
Campbell	0.47**	0.41**	0.39*	0.52**	0.25	0.21*	0.31	0.43*
Bloemen	-0.04	-0.12	0.02	-0.17	0.03	-0.05	0.42	-0.25
Rawls	0.47**	0.36**	0.57**	0.53**	0.23	0.01	0.64**	0.44*
Ahuja	0.47**	0.36**	0.57**	0.52**	0.23	0.02	0.64**	0.44*
Brutsaert	0.47**	0.38**	0.55**	0.50**	0.28	0.18	0.26	0.43*

* Positive correlation is significant at the 0.05 level (2-tailed); ** positive correlation is significant at the 0.01 level (2-tailed); [1]Number of samples (n) in parenthesis

In Table 7.5, the mean log K_{se} is compared to log K_{sm} for different soil textural classes using ANOVA with the Bonferroni mean separation test. For specific soil textural groups, the difference in model value compared to the measured value is indicated with (*). The focus is to identify the method with the least number of (or highest number of not) significantly different estimated K_s-values from measured K_s for the different textural groups.

Table 7.5 shows that the Saxon, Ahuja, and Brutsaert methods had mean K_s values not significantly different from the measured K_s at the 0.05-level for either five or six textural classes out of the eight. The remaining methods showed significant differences between the K_{se} and K_{sm} in five or more textural classes. The Brutsaert method performed credibly well with means comparable to the measured data. However, its mean for coarse textured samples such as sand and sandy loam were significantly different from those of the measured data.

Based on the two comparison approaches, the models of Brutsaert and Ahuja show a significant correlation between K_{se} and K_{sm} for a high number of textural classes and also a high number K_{se}-values are not significantly different from K_{sm}-values.

Table 7.5. ANOVA comparison of mean log of estimated saturated hydraulic conductivity (K_{se}) to measured saturated hydraulic conductivity (K_{sm}) for different soil textural classes with standard error in ( )

Texture	Loamy Sand	Sandy Loam	Clay Loam	Silt Loam	Sandy Clay L	Loam	Clay	Sand
Ksm	0.83 (0.033)	0.32 (0.027)	-0.66 (0.121)	-0.25 (0.102)	-0.22 (0.103)	-0.16 (0.050)	-0.90 (0.157)	0.92 (0.112)
				PTFs (K_{se})				
Cosby	0.77 (0.003)	0.58* (0.004)	0.01* (0.014)	0.24* (0.015)	0.37* (0.020)	0.26* (0.008)	-0.19* (0.016)	0.90 (0.003)
Brakensiek	1.02* (0.018)	0.49* (0.020)	-2.92* (0.115)	-0.29 (0.068)	1.30* (0.302)	-0.93* (0.064)	-4.13* (0.149)	1.23* (0.045)
Saxon	0.86 (0.010)	0.53* (0.010)	-0.49 (0.018)	0.48* (0.034)	-0.22 (0.018)	0.13* (0.023)	-0.72 (0.011)	1.06 (0.010)
Vereecken	0.59* (0.042	0.10* (0.026)	-0.37* (0.064)	0.49* (0.083)	-0.82* (0.102)	-0.01 (0.044)	-0.44 (0.091)	0.58* (0.090)
Jabro	0.58* (0.022)	0.24* (0.013)	-0.18* (0.035)	0.23* (0.042)	-0.26 (0.059)	0.03 (0.023)	-0.18* (0.051)	0.65* (0.050)
Pucket	0.93* (0.010)	0.50* (0.015)	-1.62* (0.057)	0.49* (0.067)	-0.69 (0.041)	-0.19 (0.042)	-2.84* (0.066)	1.05 (0.016)
Campbell	0.59* (0.011)	0.15* (0.012)	-1.21* (0.076)	-0.27 (0.052)	-0.72 (0.080)	-0.44* (0.032)	-1.68* (0.107)	0.75 (0.021)
Bloemen	2.31* (0.002)	1.80* (0.002)	0.48* (0.006)	0.93* (0.012)	1.60* (0.008)	1.04* (0.006)	0.08* (0.009)	2.67* (0.003)
Rawls	0.44* (0.008)	-0.20* (0.007)	-1.42* (0.046)	-1.20* (0.036)	-1.31* (0.075)	-0.75* (0.042	-2.66* (0.144)	0.72 (0.017)
Ahuja	1.03* (0.020	0.38 (0.016)	-0.99 (0.099)	-0.68* (0.078)	-0.75 (0.166)	-0.21 (0.089)	-2.73* (0.297)	1.11 (0.045)
Brutsaert	0.98* (0.008)	0.35 (0.005)	-0.61 (0.014)	-0.34 (0.011)	-0.55 (0.031)	-0.14 (0.009)	-1.28 (0.024)	1.33* (0.019)
F-statistic	614.0	1007.3	173.8	107.6	46.9	135.1	116.7	127.5
N	233	634	39	37	19	121	21	22

*Mean K_{sp} is significantly different from K_{sm} at 0.05-level using Bonferroni test

7.4.3 Comparison of estimation error for different PTFs

Figure 7.2 depicts the error for estimating K_s using the various PTFs with GMER on the ordinate on a log scale. In general, the methods by Brutsaert, Saxon, Vereecken, Jabro and Campbell resulted in a smaller error in estimating K_s based on the GMER compared to the other methods. From Figure 7.2, it is evident that Brutsaert's method gave the lowest error in estimating K_s.

Figure 7.3 illustrates the standard deviation of the error ratio (GSDER) for the different PTFs with error bars indicating the standard deviation for the different textural classes. The methods of Brutsaert, Rawls, Campbell, and Jabro gave the lowest GSDERs with comparatively low standard errors for the different textural classes.

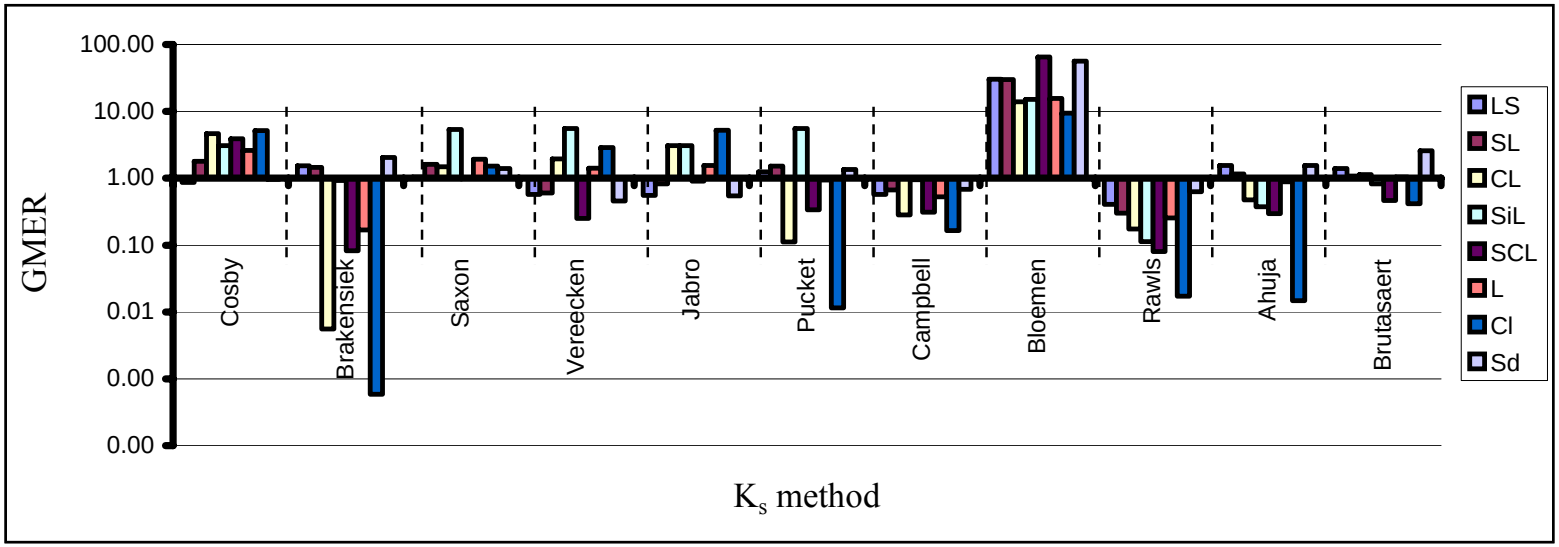


Figure 7.2. Geometric mean error ratio (GMER) for estimated saturated hydraulic conductivity using different PTFs based on different soil textural classes (loamy sand (LS), sandy loam (SL), clay loam (CL), silty loam (SiL), sandy clay loam (SCL), loam (L), clay (Cl), and sand (Sd))

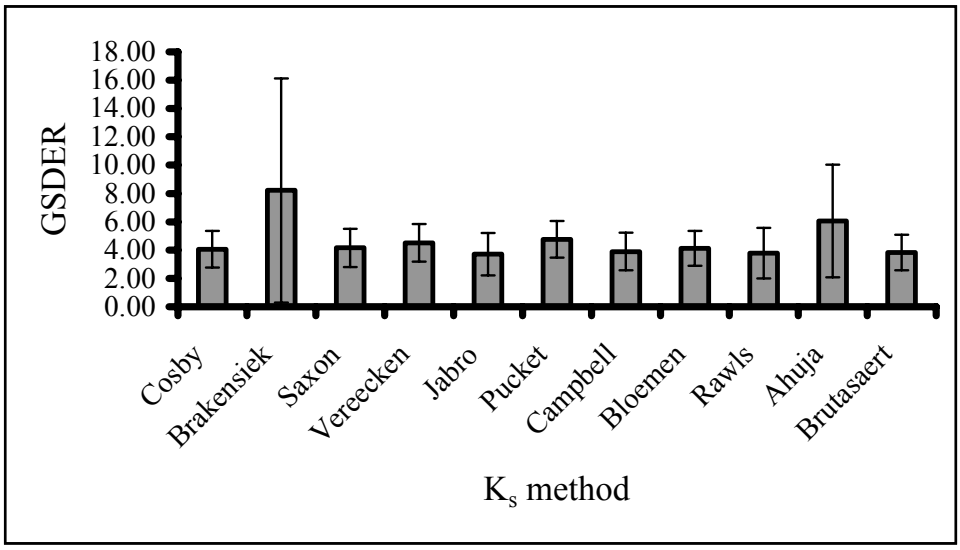


Figure 7.3. Geometric standard deviation of error ratio (GSDER) for different estimated K_s (PTFs) with error bars indicating the standard deviation for different soil textural classes

Based on the GMER and GSDER, the models of Brutsaert, Campbell and Jabro had the lowest estimation error compared to the others, illustrating that these models are more accurate in estimating measured K_s at the pilot sites.

7.5 General discussion of Pedo-Transfer Functions (PTFs)

As with all empirical studies, the results obtained are not universally applicable. Recalibration may improve the results but would not only require adaptation of the coefficients but also an evaluation of whether the input variables and expressions used are actually appropriate for the current data set. In view of this, a direct comparison of the current performance of the PTF models and published results of these PTFs was not carried out in depth. This is because such a comparison would not be fair, knowing that the published PTFs were calibrated on different data sets with possibly different data distribution. The PTF models were only evaluated on how well they perform on our independent data set, assuming that our data set quality is good. Also, their performances were evaluated in terms of their GMER and GSDER.

The models of Campbell (G) and Brutsaert (K) with R^2 of 0.38 and 0.35, respectively, were best in terms of correlating with the measured K_s. A correlation analysis in terms of soil textural groups also showed a better performance by the

Campbell, Brutsaert, Ahuja and Rawls models, yielding significant correlations at the 0.05-level or better for either five or six textural classes out of the eight classes (Table 7.3). The poor performance of the Bloemen method may be due to the inadequate number of particle size fractions (granular composition) – i.e. only three in this case (Bloemen, 1980).

In the ANOVA comparison of mean K_s for the different textural classes, the Brutsaert method correctly estimated six out of the eight textural classes and those of Ahuja (J) and Saxon (C) five out of the eight textural classes; the remaining methods estimated either three or less textural classes K_s (Table 7.4). The models of Campbell, Brutsaert, Ahuja and Rawls that directly considers the underlying physical processes showed the advantage of having a wider domain of applicability, as they are based on fundamental physical relationships. The other models' relationships are based historical data sets.

The GMER, which indicates the model with the lowest error (taken as one (1)) when the estimated is compared to the measured data, also showed that the models of Campbell and Brutsaert had the least deviation from one (Figure 7.2). These models also had the lowest GSDER with comparatively smaller standard deviations represented by the error bars (Figure 7.3).

In a nutshell, the models of Campbell and Brutsaert performed credibly well compared to the other methods. They resulted in appreciable R^2 and k_{se}-values not significantly different from the measured data for most textural classes. The two methods also yielded minimal errors when used for estimating K_s and may therefore be used for estimating K_s of soils in the Volta Basin of Ghana.

8 ARTIFICIAL NEURAL NETWORK (ANN) ESTIMATION OF SATURATED HYDRAULIC CONDUCTIVITY

Artificial Neural Networks (ANN) have become a common tool for modeling complex "input-output" dependencies. The resurgence in popularity of ANN is a worldwide phenomenon, which was started by the development of back-propagation for training multi-layer networks (Rumelhart et al., 1986). Many different types of neural networks are available with different structures (Bishop (1995), Ripley (1996) and Principe et al. (2000)).

In the past, neural network models have been used as a special class of PTFs using feed-forward back propagation or radial basis functions to approximate any continuous (non-linear) function (Hecht-Nielsen, 1990; Schaap and Bouten, 1996 and Pachepsky et al., 1996). Some of the set-backs in the use of ANN are the issue of data form or distribution, sensitivity and amount required to make a good estimate of the parameter of concern. Also, most ANN in the past have relied solely on soil physical properties – such as sand, silt and clay content, and bulk density – for estimating K_s.

In view of these, a three-set objective is proposed to investigate the use of ANN for soils in the Volta Basin of Ghana. These objectives are (1) identify data form or distribution, sensitive parameters among soil and terrain parameters, and possible data size that may be good for estimating K_s, (2) model K_s using soil and/or terrain parameters, and (3) estimate K_s for sites different from those of the training data.

This chapter is divided into four sections. The first section gives an overview of ANN, outlining the principles, data conditions required and some results of K_s estimation using ANN. The second section explains the ANN procedure used in this study. The third presents results and discussion. Finally, the last section presents the concluding remarks.

8.1 Overview of ANN

Studies in which ANNs have been examined from a statistical perspective indicate that ANN models with certain geometries, connectivities and internal parameters are either equivalent or close to existing statistical models (Cheng and Titterington, 1994; Hill et al., 1994; Sarle, 1994 and White, 1989). However, they are extremely valuable as they

are flexible, as by simply altering the transfer function or architecture one can vary model complexity; they can also be easily extended from univariate to multivariate cases incorporating non-linearities effortlessly. In neural networks, the concern is primarily one of estimation or prediction accuracy and methods that work, whereas the main objective of statisticians is to develop a universal methodology and to achieve statistical optimality (Breiman, 1994 and Tibshirani, 1994).

8.1.1 Principles of ANN

An artificial neural network consists of many interconnected simple computational elements called nodes or neurons. A neuron has multiple inputs and a single output as illustrated in Figure 8.1. Within a neuron, each input is weighted and combined (also suitably biased) to produce a single value, (*I*):

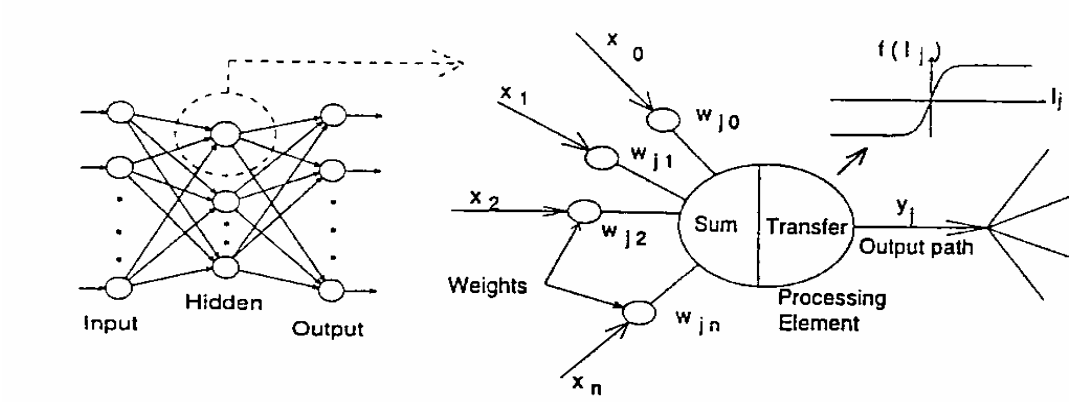


Figure 8.1. Operation of a typical feed forward artificial neural network (Source: Maier and Dandy, 2001)

$$I_j = \sum_{i=1}^{n} w_{ji} x_i + b \qquad [8.1]$$

where n is the number of inputs, x_i represents input values, and w_{ji} and b are the weights and bias associated with the neuron. The weighted input *I* is then operated on by an activation or transfer function *f*. The output value of a neuron, y, takes the form:

$$y_j = f(I_j) \qquad [8.2]$$

In ANN, one always works with a training data set and tests the performance on an independent test data set. Training of neural network entails (1) calculating output sets from the input sets, (2) comparing measured and estimated outputs, and (3) adjusting the weights and bias in the transfer function for each neuron to decrease the difference between measured and estimated values. The mean square sum of the difference between measured and estimated values serves as a measure of the goodness-of-fit.

The transfer function in the output layer has an influence on the range to which the data should be scaled (Maier and Dandy, 1999). In general, learning rate, momentum factors, network architecture, stopping criterion, and transfer function in a feed-forward ANN affect the development and results of training with a certain data set.

Learning extracts the required information from the input data with the help of the output. This is done through the selection of initial weights, learning rates, search algorithms, and stop criteria. The learning is a stochastic process that depends not only on the learning parameters, but also on the initial conditions (data quality). Stochasticity of adaptation means that there are many possible end points and therefore the need for multiple runs and results averaged.

The momentum learning used in this study is an improvement to the straight gradient-descent search in the sense that a memory term (the past increment to the weight) is used to speed up and stabilize convergence. The weights in this method are changed proportionally, based on how much they were updated in the previous iteration. For further reading on different types of algorithms see Golden (1996), Principe et al. (2000), and Yu Chen (1997).

How and when to stop the learning machine when it has learnt the task is one of the major issues in ANN. The selection of an acceptable output mean square error (MSE) level is useful in stopping the training. However, it does not address the problem of generalization, performance on data that do not belong to the training set. One approach to solving this problem is to stop the training at the point of maximum generalization and avoid overtraining (Vapnik, 1995) using early stopping, or stopping with cross-validation (smallest error in the validation set).

8.1.2 Data size, conditions and performance of ANN

The size of the training set directly influences the performance of any regression trained nonparametrically (e.g. neural networks). This class of learning machines requires a lot of data for appropriate training, because there are no *a priori* assumptions about the data.

It is generally accepted that ANNs are unable to extrapolate beyond the range of the data used for training (Minns and Hall, (1996). Consequently, it is unlikely that ANNs can account for deterministic components in the data (such as trends in the mean or variance, seasonal and cyclic components in time series). Methods of removing and dealing with trends in ANN are well covered in the literature (see Chng et al., 1996 and Khotanzad et al., 1997).

In contrast to earlier suggestion, the probability distribution of the input data needs to be known (Burke and Ignizo, 1992). Since the mean square error function is generally used to optimize the connection weights in ANN models, data needs to be normally distributed in order to obtain optimal results (Fortin et al., 1997). This needs to be confirmed empirically (i.e. using a data set).

Generally, different variables span different ranges. In order to ensure that all variables receive equal attention during the training process, they should be transformed to uniform ranges that are commensurate with the limits of the activation functions in the output layer (Masters, 1993).

ANNs belong to the class of data driven modeling approaches, which have the ability to determine which model inputs are critical, so there is no need for "…*a priori* rationalization about relationships between variables..." (Lachtermacher and Fuller, 1994). In situations with a large number of input variables, sensitivity analysis can be used to determine the relative significance of each input variable in the trained network (Maier, 1995 and Faraway and Chatfield, 1998).

There are two basic approaches that deal with the size of a learning machine (Hertz et al., 1991). By heuristics, we either start with a small machine and increase its size (growing method), or we start with a large machine and decrease its size by pruning unimportant components (pruning method). Pruning reduces the size of the learning machine by eliminating either inputs or processing elements.

8.1.3 Review of artificial neural network PTFs

Schaap et al. (1998), using neural networks in a hierarchical approach with 620 soil samples, developed models for estimating saturated hydraulic conductivity. They showed that the estimation of saturated hydraulic conductivity improved from $R^2 = 0.42$ to $R^2 = 0.57$ with increasing input data, such as the inclusion of soil water retention data.

Evaluating three data sets and the combined data in an ANN Schaap and Leij (1998) observed that the root mean square error (RMSE) of log (K_s) for the training data varied from 0.4 to 0.7 log (cm day^{-1}) and that testing with a different data set varied from 0.6 to 0.9 log (cm day^{-1}). The R^2 for one data set varied from 0.28 (testing with different data set) to 0.44 (training data set) compared to 0.55 (testing with different data set) to 0.67 (training data set) for another set of data. The uncertainty in estimating K_s was one-half to one order of magnitude and was smallest when the ANN calibration was done with all available data. Based on these results they concluded that the performance of PTFs may depend strongly on the data used in calibrating and testing the ANN. Testing using different data sets from that of the training data yields inferior results compared to the situation where the training and testing data belong to the same data set (Schaap and Leij, 1998). Furthermore, the PTFs based on the combination of the three data sets did not yield systematic errors in their estimation, thus indicating that PTFs are estimation methods that depend strongly on the quality of the calibration data (Schaap and Leij, 1998).

8.2 ANN Procedure

The ANN structure used in this study is the Multi-Layer Perceptron (MLP), which is the most commonly used neural network structure in ecological modeling and soil science (Schulze et al., 2000; Dawson and Wilby, 2001). The non-linear hyperbolic tangent transfer function was used to introduce non-linearity during training or calibration. For parsimony and ease of data interpretation and based on recommendations by Principe et al. (2000), one hidden layer was adopted. The number of neurons or processing elements (PEs) was determined by trying with different numbers. Four PEs were adopted for all analyses. Furthermore, through a number of trials using the momentum-learning rule, a step size of 1.0 and 0.01 was adopted for the hidden and output layers, respectively. A momentum value of 0.5 was used.

The maximum epoch was set at 1000, but an early stopping of learning was used for training by a cross-validation procedure using 10% of total training data set. Each analysis was carried out using five runs on different realizations of the data set, simulated by repeating 10 times with different randomization of the data. The best performance is that with lowest normalize mean square error (NMSE) (Equation 8.3) and highest coefficient of determination (R^2). The variability of the randomized estimations will be quantified by standard deviations (assuming the estimated parameters had approximately normal distribution).

$$NMSE = \frac{Mean\,square\,error}{Variance} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{S^2} \qquad [8.3]$$

where x_i and $\hat{x}_i$ are the measured and estimated parameters, respectively.

A sensitivity analysis (Equation 8.4) was carried to determine which input data have most influence on the output data, alternatively that is how changes of an input variable affect the output variable. The data in different formats (such as: raw (R), normalize (NORM), z-score transform (ZS), or min-max transform with the range of 0-1 (MM) for all continuous data and all categorical data coded as presence (1) or absence (0), were evaluated for their potential in estimation of the output, with the stability evaluated based on the standard deviation from 10 randomized data sets. Also, the optimum net for sensitive input parameters and data size were determined.

$$Sensitivity = \frac{\%\,change\,in\,\text{output}}{\%\,change\,in\,\text{input}} \times 100 \qquad [8.4]$$

For the data size requirement, a test data set (126 data points) was set-aside after each randomization for testing and the same test data set used for the different data sizes (200, 400, 600, 800, and 1000). The standard deviations provide information about the variability of the R^2 and NMSE among estimations of 10 randomization models. The ANN performance criteria adopted in this study is based on model accuracy or ability (NMSE and R^2 of training data set) and generalization or estimation ability using NMSE and R^2 for the testing data set. The models were implemented using the commercially available software package NeuroSolutions 4.0 (Principe et al., 2000).

8.3 Artificial Neural Network (ANN) model

This section is divided into three sub-sections, each addressing one of the objectives outlined in the introduction to this chapter.

8.3.1 ANN sensitivity and data size analysis

To understand the effect of input and output data distribution or form on ANN models, raw, normalized, z-score and minimum-maximum (0-1) transform data were used. The results presented in Figure 8.2 show the sensitivity of the soil input parameters in different data forms based on ANN performance. The raw, normalize, and z-score data exhibit high standard errors, which indicates instability. Also, contrary to expectation, the sensitivity of the categorical input parameters is comparatively higher than that of continuous (scale) parameters, except when minimum-maximum (0-1) data is used.

To determine the importance of the individual input parameters in modeling K_s, a sensitivity analysis was carried out. Table 8.1 presents the results of the sensitivity analysis from ANN across the different sites and soil depths with their standard deviation. The results show that the ten most sensitive parameters marked (*) are bulk density, silt content, sand content, clay content, CEC, LS factor, organic carbon, stream power index, aspect and profile curvature – in order of decreasing sensitivity. For detailed sensitivity data see Appendix 6.

To establish the effect of data distribution on ANN output, the coefficient of determination (R^2) and normalized mean square error (NMSE) for the different data distributions were compared using ANOVA. Presented in Table 8.2 are the R^2 and NMSE for training and testing for four different data forms (raw, normalize, z-score and minimum-maximum (0-1)). The analysis was done using only the ten most sensitive parameters listed, which, in order of decreasing sensitivity, are bulk density, sand content, site (Tamale), gravel and/or concretion, soil sampling depth (topsoil), soil structural grade (strong), structural type (sub-angular blocky), clay content, silt content, and structural size (course). The results show that the raw data has significantly lower R^2 and higher NMSE compared to the other data forms. For these reasons and those stated earlier, the minimum-maximum data form with the range of 0-1 is selected for use in further analysis.

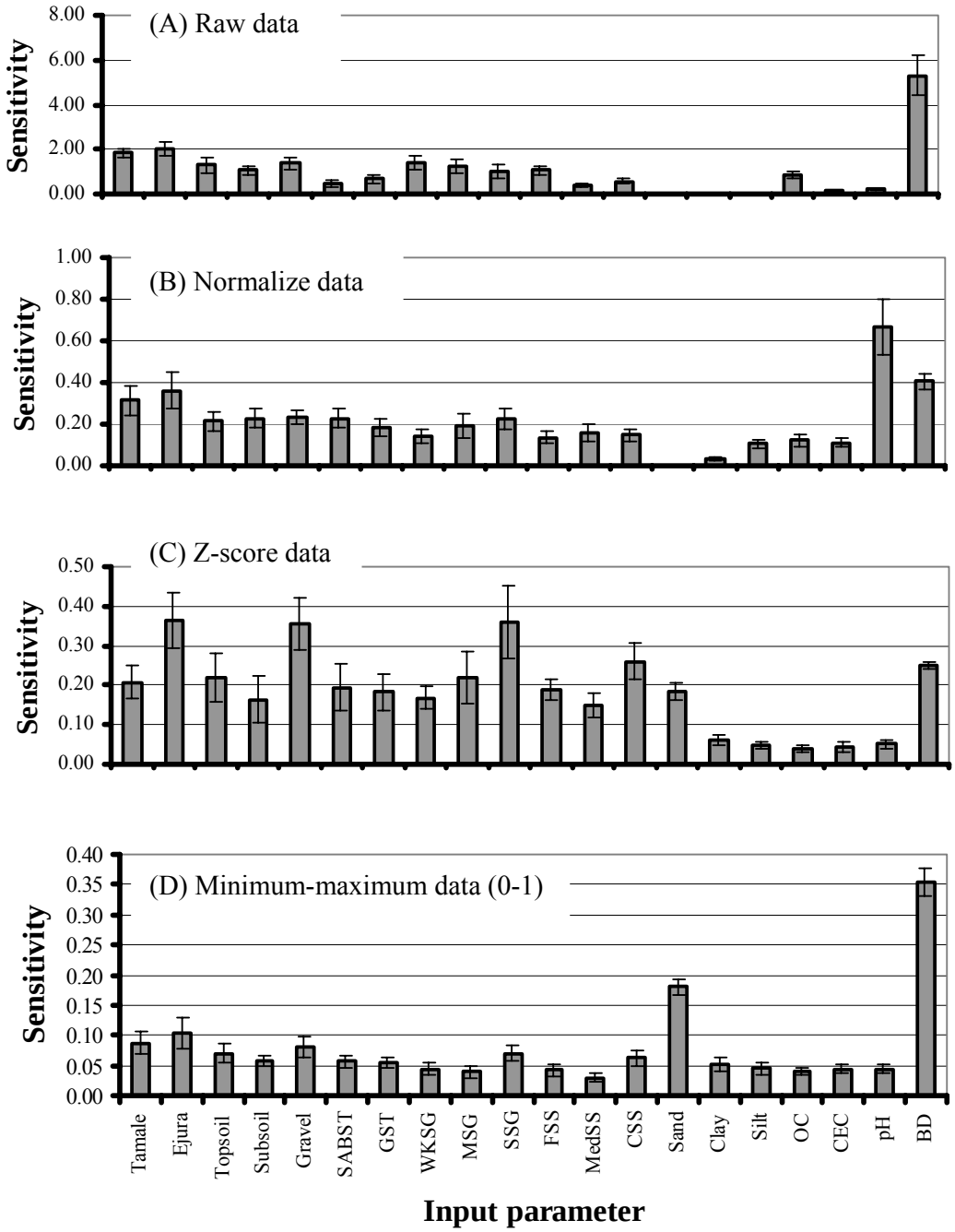


Figure 8.2. Sensitivity of different soil input parameters for different data forms (A) raw (B) normalize (C) z-score and (D) minimum-maximum (0-1) data for estimating saturated hydraulic conductivity in ANN (see Table 7.1 for complete list of parameters)

Table 8.1. Mean sensitivity of K_s estimated with all input parameters at the two sites and sampling depths and their standard deviation using ANN

Parameter	Mean sensitivity across site and depth	Standard deviation of sensitivity
Profile curvature*	0.058	0.054
Plan curvature	0.056	0.038
Curvature	0.046	0.046
Elevation	0.047	0.038
Wetness index	0.051	0.057
Upslope contribution area	0.044	0.035
Stream power index*	0.065	0.046
Slope gradient	0.053	0.046
LS factor*	0.072	0.067
Aspect*	0.059	0.036
Sand*	0.121	0.079
Clay*	0.099	0.087
Silt*	0.145	0.083
CEC*	0.088	0.061
Organic carbon*	0.069	0.036
Bulk density (BD)*	0.184	0.116
pH	0.041	0.033
Gravel/concretional presence	0.047	0.035
Sabangular blocky structure (SABST)	0.045	0.028
Grainular structure (GST)	0.041	0.044
Weak structure (WKSG)	0.035	0.023
Moderately strong structure	0.039	0.025
Strong structure (SSG)	0.058	0.048
Fine structural size (FSS)	0.052	0.039
Medium structural size (MedSS)	0.042	0.033
Course and medium structural size (CSS)	0.053	0.047

* Selected parameters for further analysis; See Appendix 4 for detailed sensitivity data

Figures 8.3(A) and 8.3(B) illustrate the effect of an increasing number of input parameters on the variation of R^2 and NMSE for estimating K_s for training and testing data. The input parameters were selected based in order of their sensitivity. Figure 8.3(A) depicts a rapid improvement in R^2 for both training and testing data for the first two most sensitive parameters; the increase then becomes gradual, with the training data maintaining a plateau after about eight input parameters, whereas for the testing data, R^2 declines with additional input parameters. The NMSE shows an opposite trend to that of the R^2 with a more or less constant value for the training data and a gradually increasing NMSE for the test data (Figure 8.3(B)). This illustrates the need to use only the most sensitive input parameters for estimation purposes.

Table 8.2. Comparison of ANN coefficient of determination (R^2) and normalized mean square error (NMSE) for training and testing data sets for saturated hydraulic conductivity based on different data forms using ANN

	Training data		Testing data	
Data form	Mean	Standard error	Mean	Standard error
	Coefficient of determination (R^2)			
Raw data	0.36b	0.021	0.29b	0.037
Normalize data	0.49a	0.015	0.40a	0.028
Z-score data	0.52a	0.011	0.40a	0.025
Min.-max data	0.51a	0.009	0.41a	0.026
F	15.20***		2.69**	
	Normalized Mean Square Error (NMSE)			
Raw data	0.64b	0.023	0.77b	0.052
Normalize data	0.51a	0.016	0.62a	0.032
Z-score data	0.48a	0.011	0.62a	0.030
Min.-max data	0.49a	0.010	0.60a	0.029
F	22.05***		4.62**	

***p < 0.01, **p < 0.05 and *p < 0.10; Mean values with the same letters in a given column are not significantly different at p < 0.05 using Bonferroni mean separation test

Figure 8.4 shows variations in R^2 and NMSE for training and testing data sets as training data size is varied. According to the figures (see also Equations 8.5 – 8.8), as the training data size increases, the R^2 and NMSE for the training data linearly increase and decrease, respectively. This is an indication of the increasing ability to train the input data as the size of the input data is increased. However, in the case of the testing data, the R^2 and NMSE increase and decrease, respectively, at a decreasing rate following a natural log function. The observation shows that with additional data (> 1000), one can expect an increase in model performance. Thus, with the maximum data size used, the model is still not able to capture all the relationships between the input parameters and K_s. The trend depicted by the testing data is an indication that with increasing training data it is possible to improve on the generalization ability of the ANN, but after a certain maximum training-data-set size there will be no further increase in the ability to generalize or estimate.

$$R^2 = 0.00008x + 0.417 \qquad R^2 = 1.00\,(Training\ data) \qquad [8.5]$$

$$NMSE = -0.00008x + 0.5856 \qquad R^2 = 1.00\,(Training\ data) \qquad [8.6]$$

$$R^2 = 0.0486\ln(x) + 0.128 \qquad R^2 = 0.92\,(Testing\ data) \qquad [8.7]$$

$$NMSE = -0.0649\ln(x) + 0.9865 \qquad R^2 = 0.91\,(Testing\ data) \qquad [8.8]$$

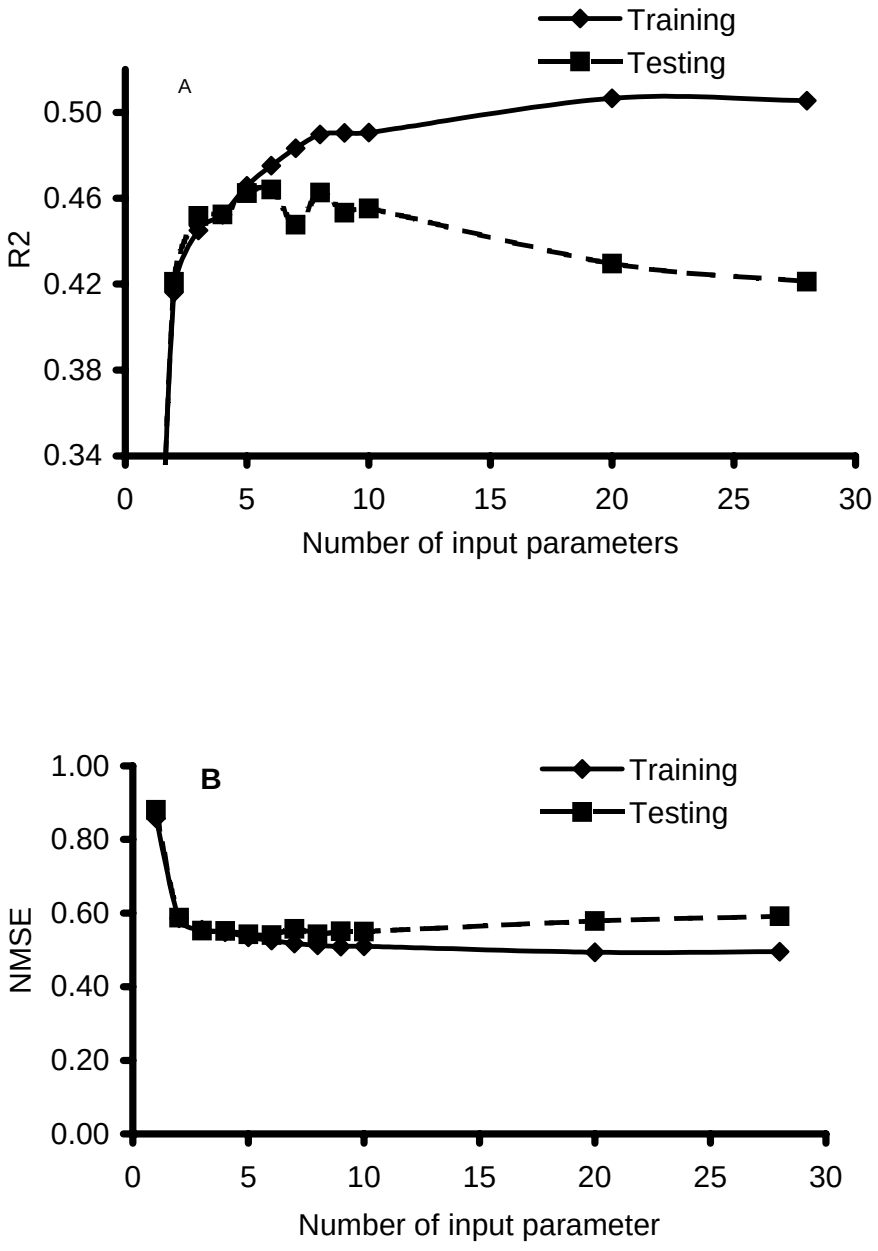


Figure 8.3. Variation in saturated hydraulic conductivity estimation with increasing number of input parameters for (A) coefficient of determination (R^2) and (B) normalize mean square error (NMSE) in ANN

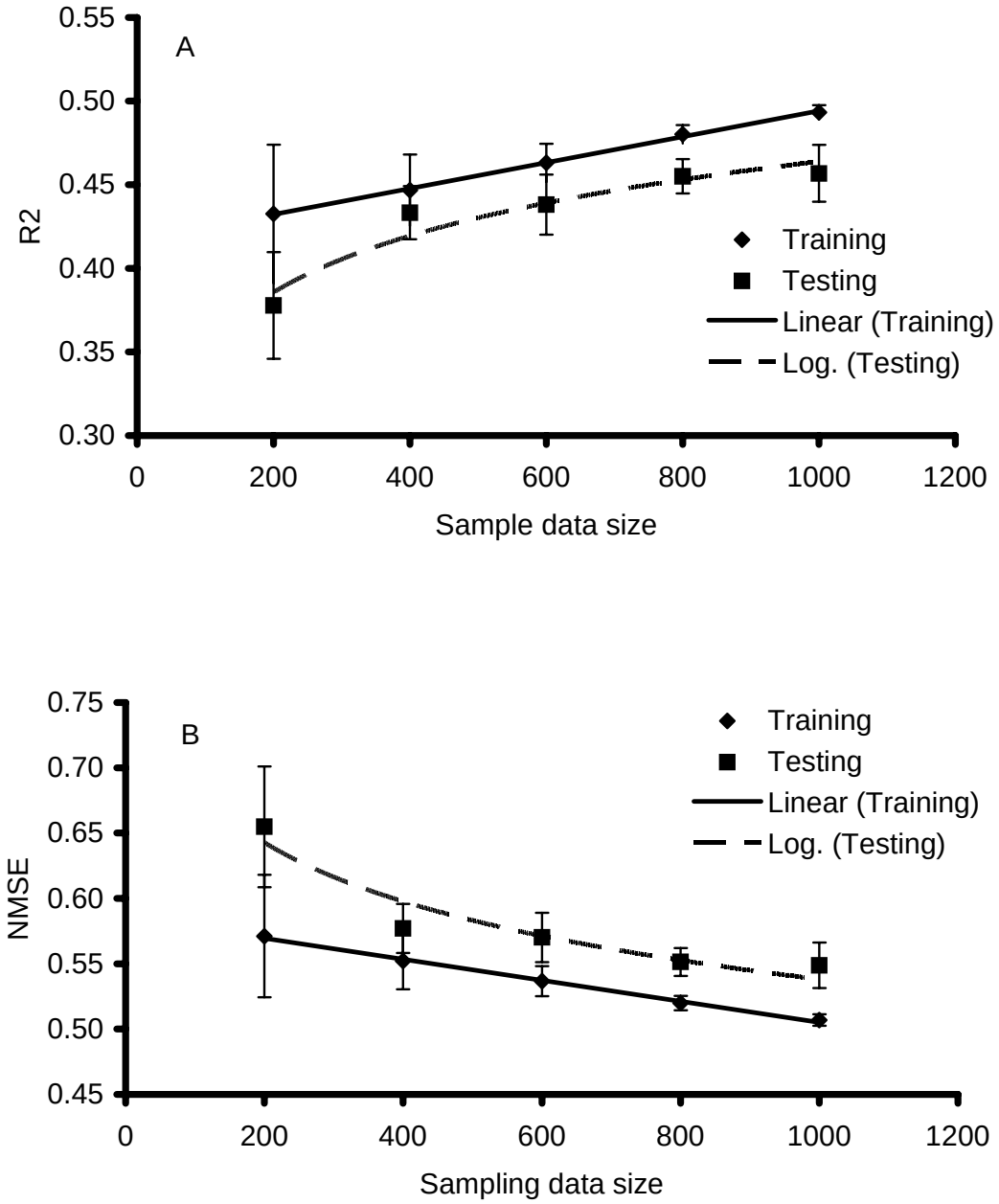


Figure 8.4. Training data size effect on (A) coefficient of determination (R^2) and (B) normalized mean square error (NMSE) for training and testing data for estimating saturated hydraulic conductivity using combined data from Ejura and Tamale in ANN

8.3.2 Artificial Neural Network (ANN) modeling with soil and terrain parameters

To investigate the importance of different data groups and the effect of soil and terrain data on R^2 and NMSE, four parameter groups were considered. Table 7.3 presents the ANOVA for the R^2 and NMSE with their standard error for training and testing data sets for the different groups of input data: (A) all parameters listed in Table 8.1, (B) ten most sensitive parameters as marked (*) in Table 8.1, (C) six most sensitive soil parameters (bulk density, silt content, sand content clay content, CEC and organic carbon content), and (D) only terrain attributes.

Table 8. 3. Coefficient of determination (R^2) and normalized mean square error (NMSE) for K_s using different data groups for all sites and sampling depths with standard error in ( )

Different input data	Mean value for training data	Mean value for test data
Coefficient of determination (R^2)		
All parameters used as input parameters (A)	0.60a (0.019)	0.47a (0.020)
Ten (10) most sensitive parameters used as input parameters (B)	0.58a (0.015)	0.51a (0.022)
Using six (6) sensitive continuous soil parameters (C)	0.56a (0.018)	0.50a (0.021)
Using only terrain parameters (D)	0.15b (0.023)	0.07b (0.013)
F-statistic (Significance)	85.1 (0.00)	92.0 (0.00)
Normalized Mean Square Error (NMSE)		
All parameters used as input parameters (A)	0.40a (0.019)	0.59a (0.026)
Ten (10) most sensitive parameters used as input parameters (B)	0.41a (0.016)	0.54a (0.025)
Using six (6) sensitive continuous soil parameters (C)	0.44a (0.019)	0.55a (0.026)
Using only terrain parameters (D)	0.87b (0.027)	1.04 b (0.031)
F-statistic (Significance)	118.9 (0.00)	79.0 (0.00)

Mean values with same letter are not significantly different at the 0.05 level using Bonferroni mean separation test

Using only terrain parameters (D) gives an R^2 that is significantly lower and an NMSE that is significantly higher than for the other three parameter groups for both training and testing data sets (Table 8.3). Using all input parameters (A), the ten most sensitive (B) and the six most sensitive soil parameters (C), R^2 are not significantly different in terms of training and testing data, suggesting that a minimum parameter set can be defined.

In investigating the effect of data source on R^2 and NMSE, an ANOVA was carried out for data from the two sites and soil depths (topsoil (0-15 cm) and subsoil

(30-45 cm)). Tables 8.4a and 8.4b give the R^2 and NMSE, respectively, for the training and testing data sets for the different input parameter groups (A to D).

Table 8.4a. Comparison of ANN coefficient of determination (R^2) with the standard error of K_s for training and testing data at different sites and sampling depths

	Training data		Testing data	
	Mean R^2	Std. error	Mean R^2	Std. error
	All parameters (A)			
Ejura topsoil	0.67a	0.021	0.56a	0.033
Ejura subsoil	0.58ab	0.024	0.42bc	0.021
Tamale topsoil	0.53b	0.038	0.36c	0.030
Tamale subsoil	0.60ab	0.051	0.753ab	0.035
F-statistic	3.09**		9.34***	
	Ten (10) most sensitive parameters (B)			
Ejura topsoil	0.66a	0.020	0.59	0.034
Ejura subsoil	0.51b	0.031	0.44	0.037
Tamale topsoil	0.60ab	0.024	0.44	0.046
Tamale subsoil	0.59ab	0.031	0.56	0.037
F-statistic	5.42***		4.06**	
	Six (6) sensitive continuous soil parameters (C)			
Ejura topsoil	0.66a	0.015	0.61a	0.031
Ejura subsoil	0.52bc	0.021	0.46bc	0.033
Tamale topsoil	0.45c	0.030	0.38c	0.044
Tamale subsoil	0.61ab	0.028	0.54ab	0.019
F-statistic	12.36***		9.32***	
	Only terrain parameters (D)			
Ejura topsoil	0.15ab	0.011	0.06	0.023
Ejura subsoil	0.05b	0.014	0.03	0.015
Tamale topsoil	0.10b	0.042	0.06	0.028
Tamale subsoil	0.28a	0.066	0.11	0.034
F-statistic	6.22***		1.64	

***p < 0.01, **p < 0.05 and *p < 0.10; Mean values with same letter in the same column are not significantly different at p < 0.05 using Bonferroni mean separation test

Focusing on the sensitive input parameter groups B and C, there were differences between the training and testing ability of the data from the different sites and soil depths. Tables 8.4a and 8.4b show that the training for Ejura topsoil is significantly better than that for Ejura subsoil with higher R^2 and lower NMSE values. These results are comparable to those obtained by Schaap and Leij (1998) (0.44-0.67 for training data set and 0.28-0.55 for testing with different data sets).

The generally low R^2 obtained for the parameter groups A and C for the Tamale topsoil compared to that of Ejura site underlines the fact that it is more difficult to

estimate saturated hydraulic conductivity for highly disturbed land areas such as those at Tamale.

Table 8.4b. Comparison of ANN normalized mean square error (NMSE) with standard error of K_s using training and testing data at different sites and sampling depths

	Training data		Testing data	
	Mean NMSE	Std. error	Mean NMSE	Std. error
		All parameters (A)		
Ejura topsoil	0.33 a	0.021	0.51	0.059
Ejura subsoil	0.42 ab	0.024	0.62	0.039
Tamale topsoil	0.48 b	0.038	0.69	0.039
Tamale subsoil	0.38 ab	0.052	0.53	0.057
F-statistic	2.93*		2.98**	
		Ten (10) most sensitive parameters (B)		
Ejura topsoil	0.34 a	0.021	0.44	0.041
Ejura subsoil	0.49 b	0.033	0.61	0.054
Tamale topsoil	0.40 ab	0.024	0.60	0.049
Tamale subsoil	0.43 ab	0.033	0.51	0.045
F-statistic	5.14***		2.72*	
		Six (6) sensitive continuous soil parameter (C)		
Ejura topsoil	0.34 a	0.015	0.42 a	0.040
Ejura subsoil	0.48 bc	0.032	0.57 ab	0.041
Tamale topsoil	0.56 c	0.033	0.68 b	0.058
Tamale subsoil	0.40 ab	0.032	0.52 ab	0.037
F-statistic	10.0***		5.80***	
		Only terrain parameter (D)		
Ejura topsoil	0.85 ab	0.011	1.05	0.052
Ejura subsoil	0.97 b	0.023	1.03	0.035
Tamale topsoil	0.92 ab	0.050	1.06	0.062
Tamale subsoil	0.75 a	0.079	1.04	0.097
F-statistic	4.01**		0.05	

***$p < 0.01$, **$p < 0.05$ and *$p < 0.10$; Mean values with same letter in the same column are not significantly different at $p < 0.05$ using Bonferroni mean separation test

Mean separation was done using the Bonferroni test. When only soil data (C) is used in the model, the K_s for Ejura topsoil and Tamale subsoil training samples were estimated better than those of the Tamale topsoil. For testing data, when C was used , K_s values for Ejura topsoil and Tamale subsoil were estimated significantly better than for Tamale topsoil, whereas K_s for Ejura topsoil was estimated better than for Ejura subsoil in terms of R^2. The inclusion of terrain parameters as in the data groups A and B

improves the R^2 and NMSE for Tamale topsoil to values comparable to those of Ejura topsoil for both training and testing data (Tables 8.4a and 8.4b).

To illustrate the relationship between measured and estimated data, a line plot of measured saturated hydraulic conductivity (K_{sm}) and estimated saturated hydraulic

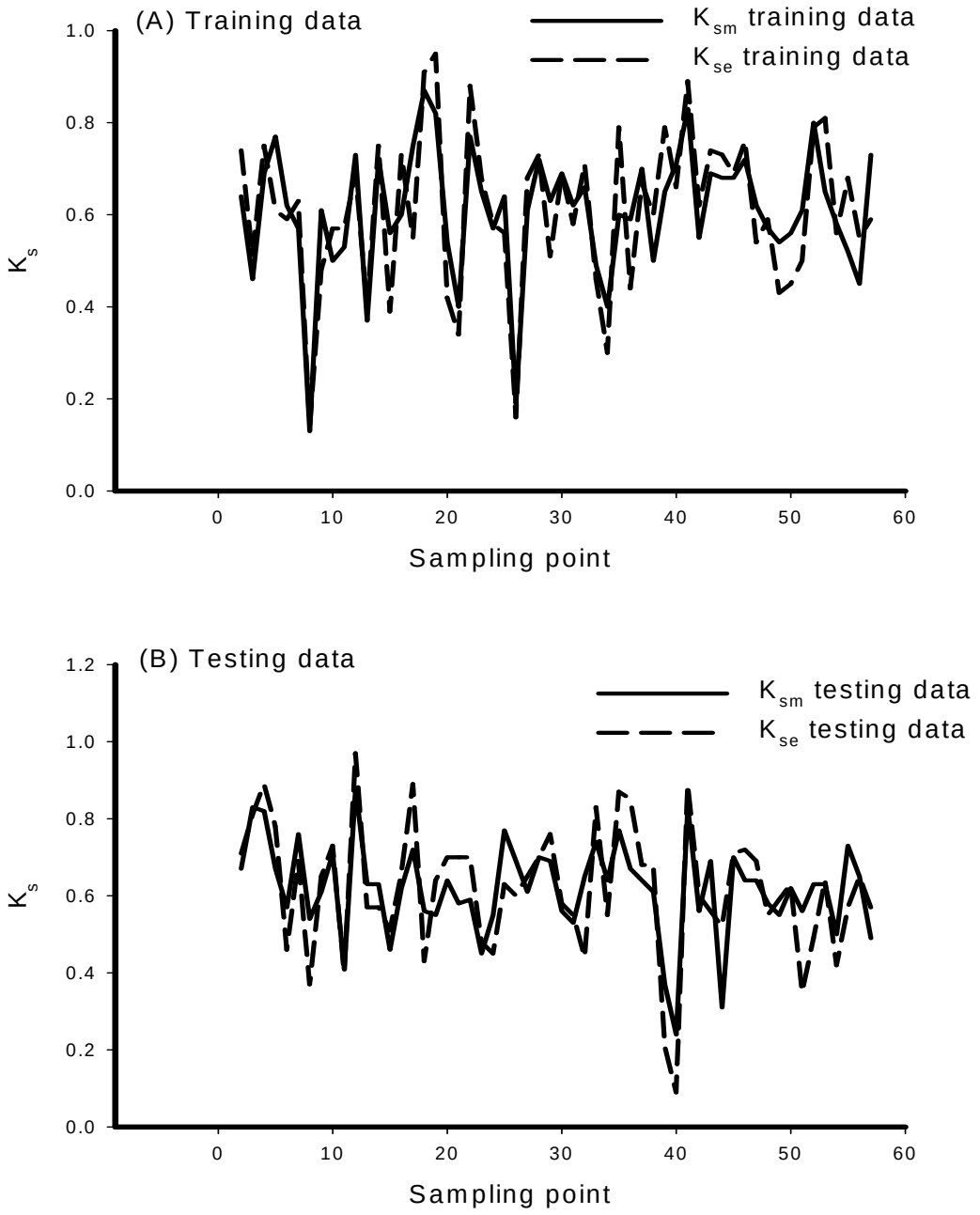


Figure 8.5. Measured compared to estimated saturated hydraulic conductivity (K_s) for (A) training and (B) testing data for a given randomization of Ejura topsoil data set in ANN

conductivity (K_{se}) for a given randomization of training and testing data was generated (Figure 8.5). For this particular Ejura topsoil data set, an R^2 of 0.69 (NMSE of 0.31) and 0.64 (NMSE of 0.36) was obtained for the training and testing data, respectively, using only the ten most sensitive input parameters (B). Figure 8.5 illustrates a very good relationship between K_{sm} and K_{se}.

From the results so far it can be said that with sufficiently sensitive data it is possible to use ANN to model K_s for soils of different sites and soil depths. However, using terrain data alone yields poor estimation of K_s, while the inclusion of sensitive terrain parameters in addition to soil properties improves the model and estimation performance of the model.

8.3.3 Estimating K_s using ANN

In order to evaluate the potential of using ANN to estimate K_s for sites very different and far removed from the particular site which data is used to develop or build the model, different model scenarios was considered. For this purpose, data from the Ejura site was used to build a model and the model tested with data from both sites and vice versa. Figures 8.6, 8.7, and 8.8 illustrate the behavior of this estimation potential.

Figure 8.6 illustrates the coefficient of determination (R^2), Figure 8.7 illustrates the NMSE for the different sites by soil depths used for estimating K_s when testing data is from the same or a different site as the training data. Shown on the graphs are the Bonferroni mean separation results using a, b, c. Also marked on the graphs are standard error bars. The two figures show a comparably high R^2 and low NMSE when testing data is from the same site as the training data. However, the R^2 for the topsoil at the two sites is high for situations when the training and testing data are from the same site and low when the testing data is from a site different from that of the training data, while R^2 and NMSE for subsoil at both sites were not significantly different whether the testing and training data are from same site or not. This indicates a more stable K_s estimation for subsoil than for the topsoil. This difference is mainly due to the higher influence of management practices on K_s at the topsoil compared to the subsoil.

Figure 8.8 illustrates the closer relationship between saturated hydraulic conductivity, measured (K_{sm}) and estimated (K_{se}) for the for the subsoil compared to the topsoil at different sites by soil depths using training data from a different site.

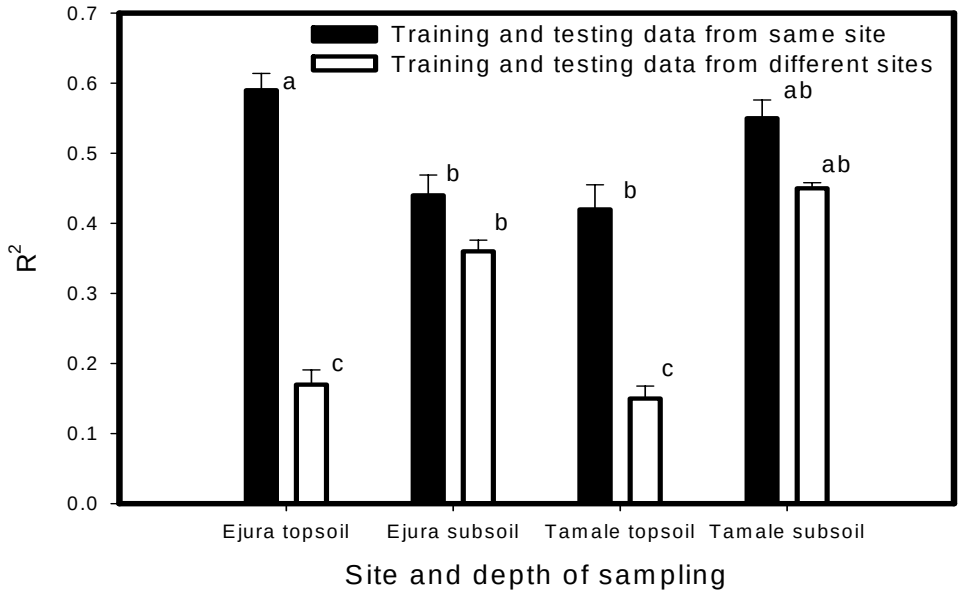


Figure 8.6. Comparison of coefficient of determination (R^2) for estimated saturated hydraulic conductivity for different testing data using training data from the same site and from a different site and indicating standard error bars in ANN

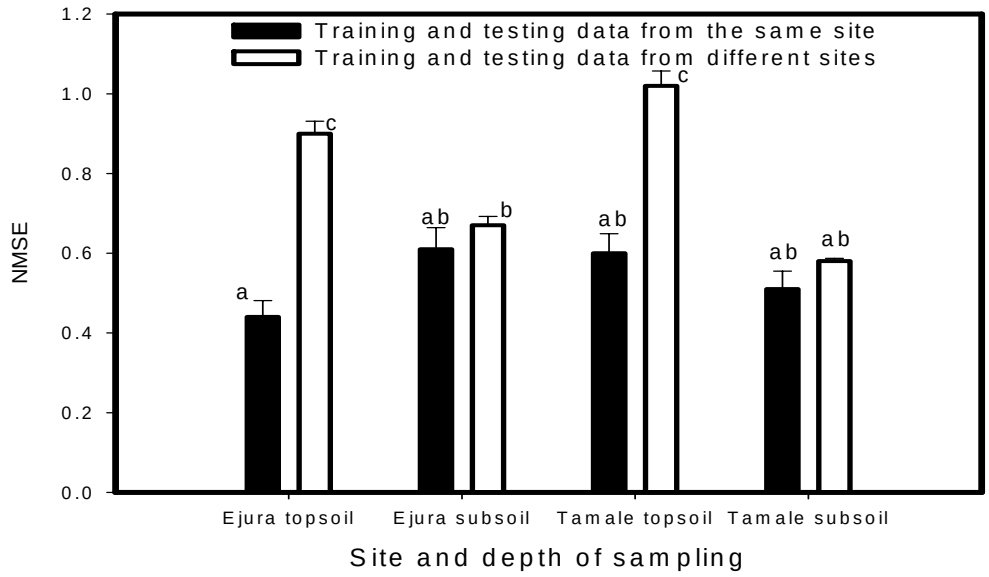


Figure 8.7. Comparison of normalized mean square error (NMSE) for estimated saturated hydraulic conductivity for different testing data using training data from the same site and different sites and indicating standard error bars in ANN

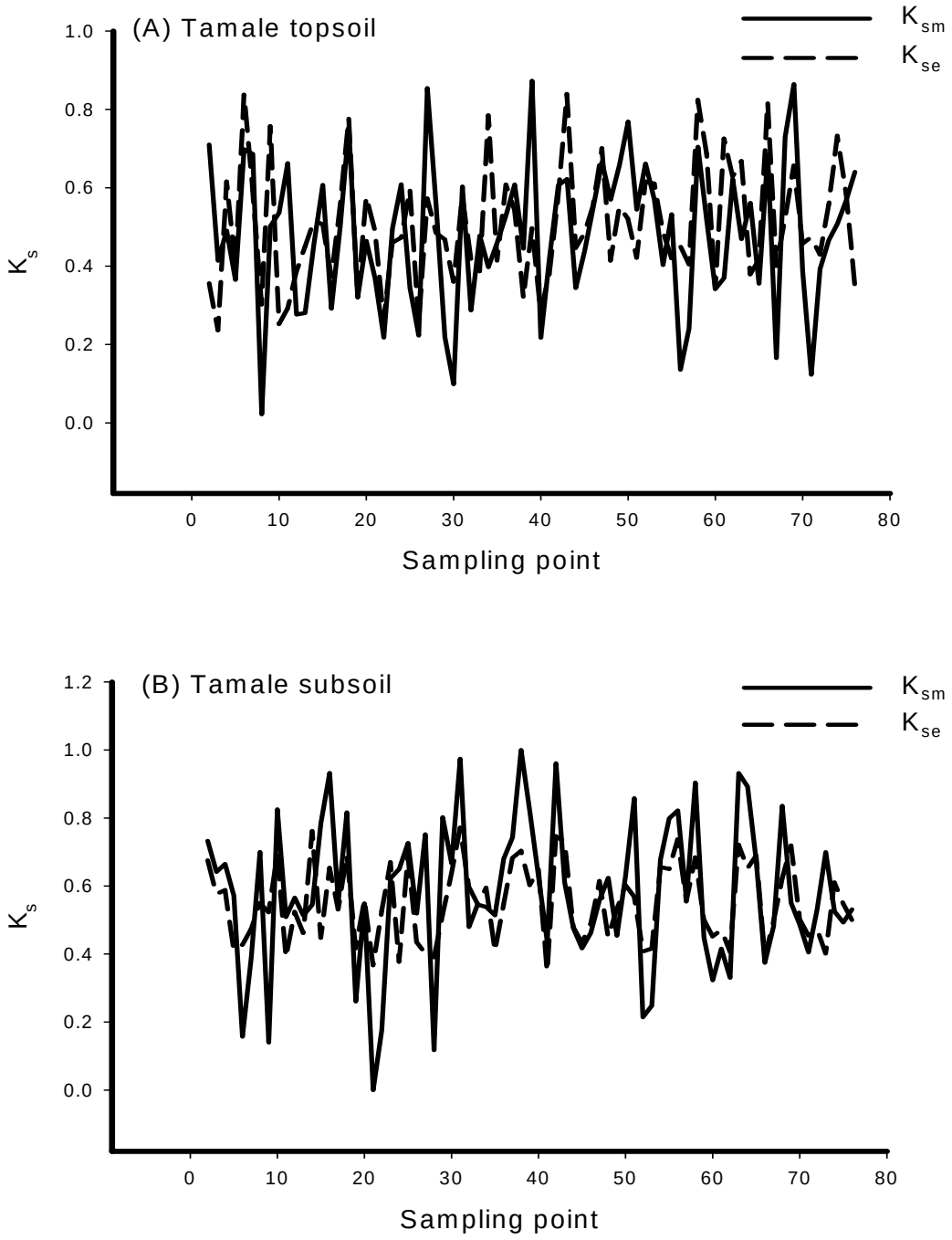


Figure 8.8a. Comparison of measured (K_{sm}) and estimated (K_{se}) saturated hydraulic conductivity for Tamale site – using training data from different site – in ANN with K_s transformed between 0 and 1

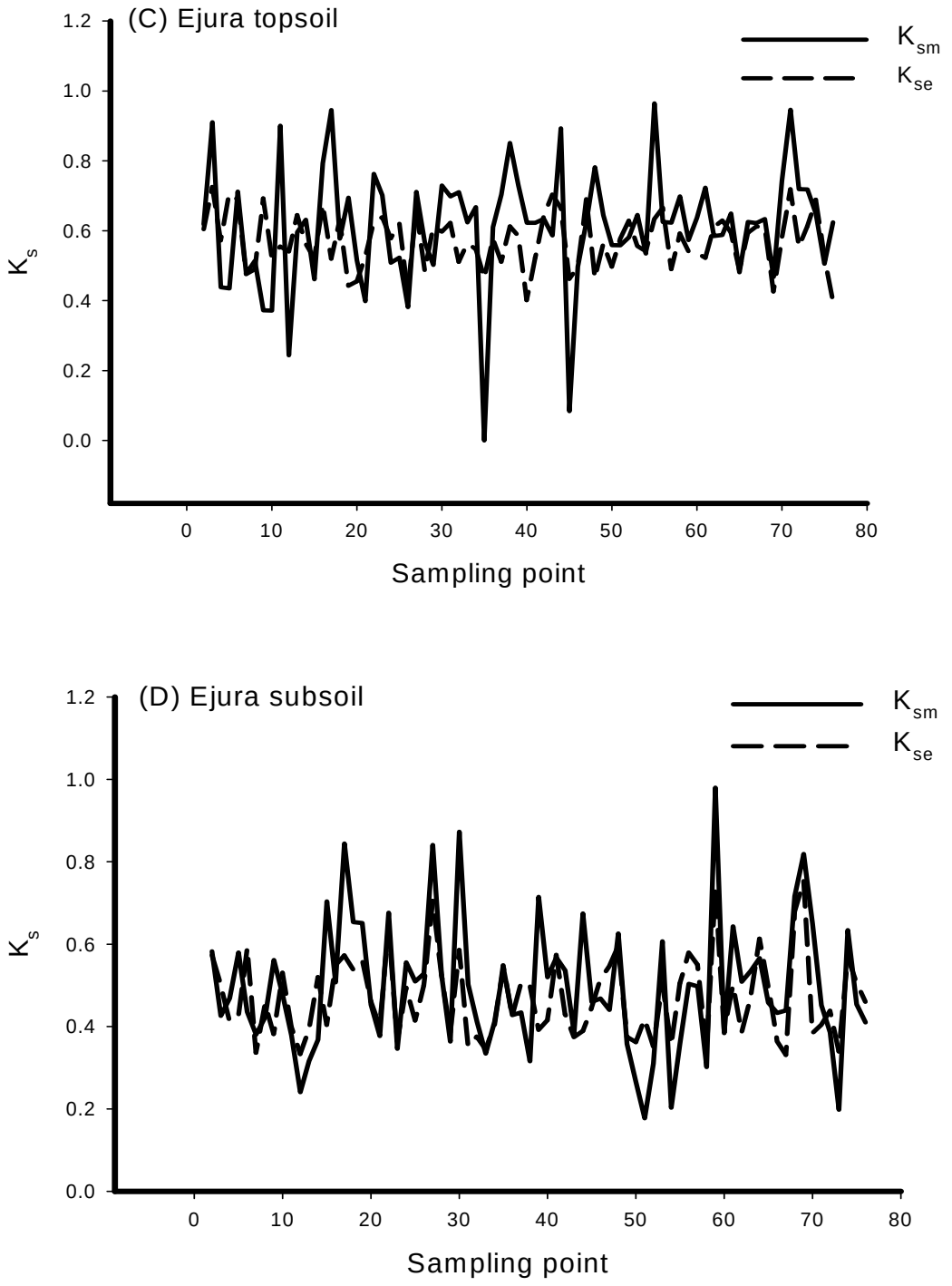


Figure 8.8b. Comparison of measured (K_{sm}) and estimated (K_{se}) saturated hydraulic conductivity for Ejura site – using training data from different site – in ANN with K_s transformed between 0 and 1

8.4 Conclusion

Based on the analysis, it was found that transforming the data into a maximum-minimum range of 0-1 yields the best ANN model performance. This is because with this transformation all data falls within the same range and therefore the undue influence of contrasting data ranges is minimized. Also, the use of sensitive data was found to be critical for modeling K_s with ANN, as too many insensitive input parameters reduce the generalization or estimation accuracy of the model. It was also evident that increasing the data size, even beyond the 1000 data set used for this analysis, may help in improving the model performance of the ANN.

ANN yields a strong R^2 if adequate sensitive data is used, suggesting good estimation of K_s. ANN can be used to estimate K_s using soil data with possible improvement in model performance when additional parameters from terrain attributes are included. The ANN models yield good results compared to the PTF models discussed in Chapter 7, if adequate data is available for training the ANN. Though the inclusion of terrain parameters can improve the estimation of K_s using ANN, it can not be relied upon solely for modeling K_s.

The comparatively better performance of ANN compared to the other methods and the fact that it does not require an *a priori* model makes it ideal for estimating K_s. However, it should be remembered that increasing the number of samples increases the estimation potential of ANN and also that the use of adequately sensitive data is important for model development (Figures 8.2 and 8.3).

For different data randomization, the standard deviations were in most cases 5 to 10% of the average R^2 and NMSE values. This implies that models with different performances are obtained when slightly different calibration data sets are used. Since cross-sectional data – compared to aggregate time series or replicated data - are used, an $R^2 \geq 0.5$ is high (Greene, 2000), knowing that the parameter being considered (K_s) is highly variable.

As shown in Figure 8.5 it is evident that, for estimation purposes, when using ANN it is important to use data from the same environment to do the training when topsoil is being considered. In the case of subsoil, where the land management effect is minimal, it is possible to estimate K_s for an area far away from the training data site (within the Volta Basin of Ghana) if the necessary input data exist.

9 GENERAL CONCLUSION AND RECOMMENDATIONS

Soil data is important for hydro-ecological and climatological modeling as it serves as the source and sink for heat and moisture to and from the atmosphere, as the natural source of nutrients for plant growth, and as the control medium for materials that enter the soil. Soil survey as a tool for collecting soil data has in the past been of tremendous value in natural resource management. However, standard soil surveys are expensive and are not designed to provide detail (high resolution) soil data, especially those relating to soil physical properties. Soil properties, especially the hydraulic properties, are spatially highly variable and measuring them is time-consuming and expensive. Therefore, efficient methods for estimating soil hydraulic properties are important. Pedotransfer functions (PTFs), are well-known methods by which hydraulic properties can be estimated through the use of easy to measure or widely available soil parameters. In recent times, artificial neural network models that utilize feed-forward back propagation to approximate any continuous (non-linear) function have been used to estimate hydraulic properties.

This study was carried out to characterize the spatial distribution of soil physical properties and to examine the variation among soil types and land use type in terms of these soil properties at the topo-scale level. Furthermore, it was to identify the most important soil, land management and terrain parameters and estimate the K_s using Stepwise Multiple Regression (SMR) and Generalized Linear Model (GLM) based on these parameters. Finally, it was to identify suitable existing pedotransfer function models and evaluate the potential of Artificial Neural Network (ANN) for estimating saturated hydraulic conductivity (K_s).

The study was carried out at two locations in the Volta Basin of Ghana near Tamale (9°28'N and 0°55'W) and Ejura (7°19'N and 1°16'W). The sites were selected to represent the diverse environmental condition in the basin across the country. The Volta Basin in Ghana covers about 42 % of Volta River catchment and about 69 % of the total 238,539 km² land area of the country. The Tamale and Ejura sites are located in the Guinea savannah and forest-savannah transitional zones, respectively. The Tamale area is gently undulating with broad valleys and isolated low-lying hills and inselbergs in the tropical continental or interior savanna climate, with a dry hot low latitude

climate (Aw) and receives about 1000 – 1200 mm of rainfall in a single rain season (April – October) with a mean annual temperature of 28°C. The Ejura area is one of the high elevation points within the Volta Basin with a tropical monsoon climate (Am), a bimodal mean annual rainfall of 1200 - 1300 mm with a high annual and monthly variability, no clear-cut beginning and end of the rains and a mean annual temperature of 26.6 °C. The geology found at both pilot sites is Voltaian sandstone made up mainly of sandstone, quartzite, shale, arkose and mudstone, covering the largest portion of the basin (about 45% of the country). The soils are predominantly Lixisols, Leptosols, Plinthosols, Acrisols and Luvisols based the FAO classification.

Soil and land use data were collected from an area of 6 km^2 at the Tamale site and 0.64 km^2 at the Ejura site using a regular grid interval. Data was collected on soil morphological properties using mini-pits, and disturbed and undisturbed soil samples were collected from 0 – 15 cm (topsoil) and 30 – 45 cm (subsoil) depths. The samples were analyzed for particle size distribution, pH, organic carbon, cation exchange capacity (CEC) bulk density, and K_s. Land use data was collected based on crop type and land preparation methods. Differential Global Positioning System (DGPS) equipment was used to generate point elevation measurements, mark sampled points and also to develop a DEM, which is used to generate terrain attributes.

Based on the fieldwork and the subsequent analysis the following conclusions can be drawn:

9.1 Soil properties, terrain attributes and their spatial distribution

Soil physical properties have high spatial variability (i.e. laterally) most especially the K_s at both soil depths at the two sites, irrespective of the terrain attributes in the locality. The soil textural conditions at the Ejura site are more uniform, while a sharp increase in clay content from the topsoil to the subsoil is typical for the soils at the Tamale site. This difference may be due to difference in parent material at the two sites, with the Tamale soils being of high inherent clay content with translocation down the soil profile over time. The very high variability of K_s underlines the difficulty in accurately measuring this soil property and therefore the need for suitable estimation methods. The least variable parameter was bulk density.

The Ejura site is characterized by steeper slopes compared to the Tamale site and the terrain attributes were more pronounced. This implies that the Ejura site is more prone to high runoff and erosion if the soils are exposed.

Higher spatial dependencies were observed for soil properties in the subsoil compared to the topsoil as a result of less soil disturbance in the subsoil. Furthermore, the spatial dependencies for soil properties at Tamale were generally higher than those at the Ejura site, reflecting the gentler slopes at the Tamale site.

The K_s was observed to have a good relationship with soil properties such as sand, silt and clay content and bulk density but the relationship with terrain attributes was poor. This implies that the use of soil properties to estimate K_s will probably be more successfully compared to terrain attributes.

9.2 Spatial distribution of soil types and land use types and their relation to soil properties

The soils at the two sites follow a catenary development. The soils at Ejura have a more distinct catenary development than at Tamale, which may be due to the steeper terrain around Ejura.

Soil types vary widely across the two pilot sites, with each type used for diverse crops and mixed/intercropping, especially at the Tamale pilot site. The identified soil types varied at each site based on the textural properties and pH of the topsoil (0-15 cm) and subsoil (30-45 cm). Although the soils at each site were put into eight Tamale and five (Ejura) groups (soil type), three main categories of soil emerged when comparing soil properties. These are the soils on the upland (. summits and upper slopes), which are more stable due to the minimal effect of soil translocation, the mid-slope soils on steep slopes that are influenced by soil movement due to the slope, and the soils of the lowland and valley bottoms that receive soil material. The main mechanisms influencing soil formation are lateral and vertical soil translocation at both sites and plinthization at the Tamale site.

At both soil depths, soil properties such as sand, silt, clay and pH vary for the different soil types across the landscape mainly as a result of soil translocation and leaching. Sand content is usually highest on the upper and mid slopes compared to the summit and lowland and valley bottom, while clay content is usually highest in the

lowland and valley bottom soil types. Beside the parent material that may have given rise to a higher sand content at Ejura than Tamale, the most important factors determining the soil type are slope gradient and the soil drainage regime.

The parameters K_s, bulk density, and CEC in most cases do not vary significantly between soil types at the different soil levels due to the highly heterogeneous nature of the soils as a result of land management; this results in high variation of these parameters within a given soil type – most especially of K_s.

Different crop groups are cultivated on soils with different textures, e.g. rice and root and tuber crops are cultivated on soils of high clay content. The significant difference in particle size distribution between land use types (LUT) suggests that the choice of land for specific crops is made taking indirectly into account the soil properties. The differences in soil texture in terms of LUT were smaller than those of soil type, meaning the latter can give better distinction of these soil properties. However, LUT can be used for quick distinction of areas of different texture, as LUT mapping is less time and labor consuming.

9.3 Saturated hydraulic conductivity relationship with soil properties, soil type, land use type and terrain parameters

To identify the key data groups (location, soil properties, soil type, land use type and terrain attributes) as independent variables for estimating K_s, stepwise multiple regression and "generalized linear model" (GLM) were used. Based on three data sets (i.e. data from the Tamale site, the Ejura site and the combined data from the two sites) it was found that there are two important data groups for estimating K_s, which are location (i.e. site and soil depth) and soil properties (sand, silt, clay and bulk density).

The importance of location data may be due their ability to capture differences in soil parent material, management practices and soil disturbance. The textural properties (sand, silt and clay content) and bulk density were found to be very important in estimating K_s as these parameters directly influence soil pore size and distribution which in turn determine how water flows through the soil and therefore K_s.

Soil type, land use type and terrain attributes had little effect on estimating K_s. They cannot be relied upon as the sole independent parameters for estimating K_s, but can be useful as additional data to improve the results. These parameters are fairly easy

to obtain with the advent of high resolution image acquisition compared to data on soil properties that require field sampling work.

Through the GLM analysis it was shown that the interaction terms are very important for modeling K_s. Therefore, the inclusion of interaction terms in K_s models may improve their performance.

9.4 Pedo-Tansfer Functions (PTFs)

To evaluate the performance of PTFs for estimating saturated hydraulic conductivity in the Volta Basin, eleven PTFs were evaluated using coefficient of determination, ANOVA, geometric mean error ratio (GMER) and geometric standard deviation of error ratio (GSDER). PTFs proposed by Campbell (1985) and Brutsaert (1967) can be used for estimating saturated hydraulic conductivity in the basin with an R^2 of 0.38 (Campbell) and 0.35 (Brutsaert) with possible improvement when specific textural classes are considered. These two models performed credibly well compared to the other nine models considered. The R^2 are high, given that saturated hydraulic conductivity is a highly variable soil parameter and its accuracy is very often determined in terms of order of magnitude. These two PTFs also gave the least errors based on their GMER and GSDER.

It is evident from the results that, as with all empirical studies, the results obtained are not universal. Recalibration may improve the results, but would not only require adapting of the coefficients but also evaluating whether the input variables or expressions used are actually appropriate for the current data set. In view of this, a direct comparison of the PTF models and published PTFs should be done with caution. This is because such a comparison would not be objective, given that the eleven PTFs were established on different data sets, with possibly different data characteristics. The PTF models were only evaluated on how well they perform based on the Tamale and Ejura data set, based on the assumption that the data set quality is fairly good. Also their performances were evaluated in terms of their GMER and GSDER.

The models of Campbell, Brutsaert, Ahuja and Rawls that directly consider the underlying physical processes, portrayed their advantage of having a wider domain of applicability, as they are based on fundamental physical relationships. The other investigated models are based on historical data sets.

9.5 Artificial neural network (ANN)

The use of ANN in estimating saturated hydraulic conductivity turned out to be comparatively better than the use of the PTFs considered in this study. The ANN used can estimate K_s provided adequate sensitive data are available. Though the inclusion of terrain parameters can improve the estimation of K_s in using ANN, they cannot be relied upon solely for modeling K_s. The better performance of ANN compared to the PTF methods and the fact that it does not require an *a priori* model makes it ideal for estimating K_s.

In order to achieve a satisfactory model using ANN it is good to do an extensive data pre-processing (in this case: normalize maximum-minimum (0-1) data transformation) as required for statistical models, contrary to the popularly held view of 'no need for data pre-processing in ANN'. Data transformation reduces undue influence due to contrasting data ranges. Also, the use of adequately sensitive parameters was found to be critical for modeling K_s using ANN, as too many insensitive input parameters reduce the generalization or estimation accuracy of the model. A large data set (> 1000) is required for good estimation results for K_s. Increasing the number of samples increases the estimation potential of ANN.

It was found that for estimating K_s using ANN, it is important to use data from the same environment to do the training when the topsoil is being considered. In the case of subsoil, the land management effect is minimal; it was possible to estimate K_s for an area far away from the training data site within the basin.

9.6 Summary of conclusion and the way forward

It is evident from this study that, though K_s is highly spatially variable, it can be estimated using suitable PTFs, such as those developed by Brutsaert (1967) and Campbell (1985) for soils in the Volta Basin of Ghana. Furthermore, when adequate sensitive data are available, then ANN may offer a better K_s estimation approach compared to the PTFs.

On-going studies by the Glowa-Volta project in the Volta Basin of West Africa, with the goal of developing a Decision Support System (DSS) for sustainable land and water use in the basin will require a good approximation of K_s and other soil data for initialization of the model. For example, K_s is one of the most sensitive parameters in

the land surface model of the MM5 for modeling climate change, which is one of the key components of the DSS.

It should be noted that this study was only carried out on a pilot scale level with small grid size compared to a basin-wide scale with a course resolution of 3 km × 3 km required by the MM5. This calls for generating a K_s map for the basin, which can be done based on the knowledge acquired through this study. The way forward will be to use the existing soil map of the basin as the first approximation of K_s. Based on the observation that particle size distribution (sand, silt and clay content) varies from the upland to the lowland and the fact that it is the most important input parameter for K_s estimation, terrain attributes generated from DEM can be used to improve on the quality of the said K_s map. Also, the use of land use intensity map may help improve on any resulting K_s map as the topsoil is strongly influenced by land management practices.

9.7 Recommendations

For further studies, the following should be considered:

- For the purposes of estimating soil input parameters for hydro-climatological models, it is advisable to focus on particle size distribution (texture) as this can be used to generate almost all required soil input parameters for such models. The use of historical or past land use data rather than the current land use data applied in this study may provide a better understanding of land use influence in the estimation of saturated hydraulic conductivity. Also, the use of an object-oriented approach, through which form elements are considered rather than the grid point terrain attributes used in this study, may improve the use of terrain information in estimating K_s.
- Further studies should address the issue of up-scaling of the knowledge acquired in this study as environmental and hydro-climatological models will require K_s input data based on the model grid (e.g. 3 km × 3 km or 9 km × 9 km). They should also look at identifying appropriate aggregations schemes (such as the use of arithmetic, geometric, harmonic averaging and wavelet analysis) for the different input soil parameters required for GCM and other environmental models, as changing the scale of observation affects research results and their interpretation.

- For purposes of generating elevation data using differential GPS to develop DEM, data collection in shaded areas (e.g. forest areas) must be done with care as these data can reduce the accuracy.

REFERENCES

Adu SV (1969) Soils of the Navrongo-Bawku area, upper region, Ghana. Memoir no. 5, Soil Research Institute, Kumasi Ghana

Adu SV (1995) Soils of the Bole-Bamboi area: Northern regions, Ghana. Soil Research Institute (Council for Scientific and Industrial Research). Memoir no. 14: 16-52

Adu SV and Mensah-Ansah JA (1995). Soils of the Afram basin: Ashanti and Eastern regions, Ghana. Soil Research Institute (Council for Scientific and Industrial Research). Memoir no. 12:17-53

Ahuja LA, Naney JW, Green RE and Nielsen DR (1984) Macroporosity to characterize spatial variability of hydraulic conductivity and effects of land measurement. Soil Sci. Soc. Am. J. 48: 699-702

Anderson JL and Bouma J (1977) Water movement through pedal soils: I. Saturated flow. Soil Science Society of America Journal 43: 413-418

Arya LM and Paris (JF) 1981 A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. Soil Science Society of America Journal 45: 1023-1030

Ashtech (1998a) Ashtech manual. Ashtech Inc.

Ashtech (1998b) Locus processor user's guide. pp 1-180. Ashtech Inc

Asiamah RD (2002) Plinthite development in upland agriculture soils in Ghana, Paper presented at 17th WCSS, 14-21 August 2002, Paper no. 1448: 1-12

Bates DA (1962a) Geology. In: Wills JB (ed) Agriculture and land use in Ghana. Ghana Ministry and Food and Agriculture, pp 51-61. Oxford University press, London

Bates DA (1962b) Rural water supplies. In: Wills JB (ed) Agriculture and land use in Ghana. Ghana Ministry and Food and Agriculture, pp 62-76. Oxford University press, London

Bauer J, Rohdenburg, H and Bork HR (1985) Ein Digitales Reliefmodell als Vorraussetzung fuer ein deterministisches Modell der Wasser- und Stoff-Fluesse, In: H.R. Bork and H. Rohdenburg (Editors), Parameteraufbereitung fuer deterministische Gebiets-Wassermodelle, Grundlagenarbeiten zu Analyse von Agrar-Oekosystemen. Landschaftsgenese und Landschaftsoekologie H10:1-15

Beckett PHT and Webster R (1971) Soil variability: a review. Soils and Fertilizers 34: 1-15.

Bishop CM (1995) Neural Networks for Pattern Recognition. Oxford University Press, Oxford

Blake GR and Hartge KH (1986) Bulk Density. In: A. Klute (ed) Methods of Soil Analysis, Part 1, 2nd Agron. Monogr. 9, ASA, pp 363-375. Madison, WI

Bloemen GW (1980) Calculation of hydraulic conductivities of soils from texture and organic matter content. Zeitschrift für Pflanzenernährung und Bodenkunde, 143: 581-605

Bonsu M (1992) A study of texture-based equation for estimating the saturated hydraulic conductivity of an alfisol in the sudan savannah ecological zone, Ghana. Hydrological Sciences 37(6) 599-606

Bouma J (1989) Land qualities in space and time. In: Bouma J and Bregt AK (eds) Land qualities in space and time, pp 3-13. PUDOC, Wageningen, the Netherlands

Bouma J and van Lanen JAJ (1987) Transfer functions and threshold values: From soil characteristics to land qualities. In: Beek KJ, Burrough PA and McCormack DE (eds) Proc. ISSS/SSSA workshop on Quantified land evaluation, International Institute or Aerospace Surv. Earth Sci. ITC publ. 6, pp 106-111. Enschede, the Netherlands

Bouma J Jongerius A and Schoonderbeek D (1979) Calculation of saturated hydraulic conductivity of some pedal clay soils using micromorphometric data. Soil Sci. Soc. Am. J. 43: 261-264

Bouyoucos GJ (1962) Hydrometer method improved for making particle size analysis of soil. Agron. J. 54: 464-465

Bowen HD (1981) Alleviating mechanical impedance. In: Arkin GF and Taylor HM (eds.) Modifying the root environment to reduce crop stress, ASAE, pp 21-57 St. Joseph, MI

Brady NC and Weil RR (1996) The nature and properties of soil. 12ed Prentice Hall, New Jersey, USA

Brakensiek DL Rawls WJ and Stephenson GR (1984) Modifying SCS hydrologic soil groups and curve numbers for rangeland soils. ASAE Paper No. PNR-84-203. St. Joseph, MI.

Breiman L (1994) comment on "Neural networks: A review from a statistical perspective" by Cheng B and Titterington DM, Stat. Sci. 9(1): 38-42

Brutsaert W (1967) Some methods of calculating unsaturated permeability. Trans. ASAE 10:400-404

Buol SW Holes FD and Macracken RJ (1989) Soil Genesis and Classification. pp 446 Iowa State Univ. Press, Ames.

Burke LI and Ignizio JP (1992) Neural networks and operations research: an overview. Computer Oper. Research, 19(3/4): 179-189

Burrough PA (1986) Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford, pp 193

Burrough PA, van Gaans PFM and MacMillan RA (2000) High-resolution landform classification using fuzzy k-means. Journal of Fuzzy Sets and Systems 113:37-52

Burrough PA, Wilson JP, van Gaans PFM and Hansen AJ (2001) Fuzzy k-means classification of digital elevation models as an aid to forest mapping in the Greater Yellowstone Area, USA. Landscape Ecology of Fuzzy Sets and Systems 16:523-546

Burt TP and Butcher DP (1986) Development of topographic indices for use in semi-distributed hillslope runoff models. In: Baltenau, D. and Slaymaker, O. (eds) Geomorphology and Land Management, Zeits. Geomorph. Suppl. Band 58: 1-19

Campbell GS (1985) Soil physics with BASIC. Development in soil science, Vol. 14. Elsevier, New York

Chen F and Dudhia J (2001) coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part 1: Model implementation and sensitivity. American Meteorology Society 129: 569-585

Cheng B and Titterington DM (1994) Neural networks: A review from a statistical perspective. Stat. Sci. 9(1) 2-54

Chhabra R, Pleysier J and Cremmers A (1975) The measurement of the cation exchange capacity and exchangeable cations in soils: A new method. In: Proceedings of the 5th International Clay Conference, pp 439-449. Applied Publishing, Willmette, Illinois

Chng ES, Chen S and Mulgrew B (1996) Gradient radial basis function networks for nonlinear and nonstationary time series prediction. IEEE Transactions on Neural-Networks 7(1): 191-194

Clapp RB and Hornberger GM (1978) Empirical equation for some soil hydraulic properties, Water Resour. Res. 14: 601-604

Conacher AJ and Darlrymple JB (1977) 'The nine-unit landsurface model': An approach to pedogeomorphic research. Geoderma, 18: 1-153.

Conrad O (2001) DEM-Software for digital elevation model. Ph.D. Thesis (in German), University of Goettingen, Germany. Available at http://www.geogr.uni-goettingen.de/pg/saga/digem/download/diplom1.pdf.

Cook SE, Corner RJ, Grealish G, Gessler PE and Chartes CJ (1996) A rule-based system to map soil properties. Soil Sci. Soc. Am. J., 60:1893-1900.

Cosby BJ, Hornberger GM, Clapp RB and Gin TR (1984) A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour. Res. 20: 682-690

Costa-Cabral MC and Burges SJ (1994) Digital elevation model networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal area. Water Resources Research 30: 1681-1692

Dangler EW, El-Swaify SA and Barnett PA (1975) Erosion losses from Hawaii soils under simulated rainfall. Research Bulletin 181. Hawaii Agricultural Experiment Station, University of Hawaii, pp 63

Dawson CW and Wilby RL (2001) Hydrological modeling using artificial neural networks. Progress in Physical Geography 25: 80-108

Desmet PJJ and Govers G (1996) Comparison of routing algorithms for digital elevation models and their implications for predicting ephemeral gullies. International Journal of Geographical Information Systems 10: 311-331

Dickson, K.B. and Benneh, G. 1988. A new geography of Ghana. London: Longman.

Dikau R (1989) The application of a digital relief model to landform analysis in geomorphology. In: J. Raper J (ed) Three Dimensional Applications in Geographic Information Systems, pp 51-77. Taylor and Francis, New York

Ek M and Cuenca RH (1994) Variation in soil parameters: implications for modeling surface fluxes and atmospheric boundary-layer development. Boundary-Layer Meteorology 70: 369-383

Elliot WJ, Kohl KD and Laflen JM (1988) Methods of collecting WEPP soil erodibility data. Paper No. MCR 88-138. ASAE, St. Joseph, MI

Ellison WD (1947) Soil erosion studies. Part 1, Agricultural Engineering 28 (4): 145-146

Espino A, Mallants, Vanclooster M and Feyen J (1995) Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties. Agricultural Water Management 29:235-253

Fairfield J and Leymarie P (1991) Drainage networks from grid digital elevation models. Water Resources Research 27: 709-717, 2809.

FAO (1983) Guidelines: Land evaluation for rain-fed agriculture. Soils Bulletin 52. FAO, Rome

FAO (1988) Soil Map of the World, Revised Legend, Reprinted with corrections. World Soil Resources Report 60, FAO, Rome

FAO (1990) Guidelines for Soil Description, Third edition (revised) Soil Resources, Management and Conservation Service, pp 14-66. Land and Water Development Division. FAO, Rome

Faraway J and Chatfield C (1998) Time series forecasting with neural networks: a comparative study using the airline data. Appl. Stat. 47(2): 203-209

Fortin V, Ouarda TBMJ and Bobee B (1997) Comment on "The use of artificial neural networks for the prediction of water quality parameters" by Maier HR and Dandy GC, Water Resources Research 33(10): 2423-2424

Foth HD and Ellis BG (1997) Soil fertility. Second edition, pp 290. CRC press Boca Raton, Florida

Freeman GT (1991) Calculating catchment area with divergent flow based on regular grid. Computers and Geosciences 17: 413-422

Gaur A, Horton R, Jaynes DB, Lee J and Al-Jabri SA (2003) Using surface time domain reflectometry measurements to estimate subsurface chemical movement. Vadose Zone Journal 2(4): 539-543

Gerakis A and Baer B (1999) A computer program for soil textural classification. Soil Science Society of America Journal 63:807-808

Gessler PE, Moore ID, McKenzie NJ and Ryan PJ (1995) Soil-landscape modeling and spatial prediction of soil attributes. Geographical Information Systems 9: 421-431

Giles (1998) Geomorphological signatures: classification of aggregated slope unit objects from digital elevation and remote sensing data, Earth Surface Processes and Landforms 20: 582-94

Glowa-Volta (1999) Sustainable water use under changing land use, rainfall reliability and water demands in the Volta Basin. ZEF, Bonn, http://www.glowa-volta.de/general/pubs/proposal.pdf

Golden RM (1996) Mathematical methods for neural network analysis and design. MIT press, Cambridge, Mass., pp 419

Goldensoftware Inc., (1999) Surfer 7.0. Goldensoftware Inc.

Grayson RB, Blöschl G, Barling RD and Moore ID (1992) Process, scale, and constraints to hydrological modeling in GIS. In: Kovar K and Nachtnebel (eds) Application of Geographic Information Systems in Hydrology and Water Resources: Proceedings of the HydroGIS 93 Conference held in Vienna, April 1993, Wallingford, UK: International Association of Hydrological Sciences 211: 83-92

Greene WH (2000). Econometric analysis. Forth edition. Prentice Hall International, Inc. Upper Saddle River, New Jersey, pp 241

Grubaugh K (2003) Geology of Ghana. Presented at SME Annual Meeting, February, 26-28, 2001, Denever, Co. http://me.smenet.org/200301/pdf/min0301_39.pdf

GSS (2002a). 2000 Population and Housing Census. Special report on 20 largest localities. Ghana Statistical Service. March, 2002. Medialite Co. Ltd.

GSS (2002b) 2000 Population and Housing Census. Summary report of final results. Ghana Statistical Service. March, 2002. Medialite Co. Ltd.

Hair JF, Anderson RE, Tatham RL and Black WC (1998) Multivariate data analysis. Prentice Hall, Upper Saddle River, New Jersey, pp 147, 178

Hamilton LC (1990) Modern data analysis. A first course in applied statistics, pp 684

Hanna AY, Harlan PW and Lewis DT (1982) Soil available water as influenced by landscape position and aspect. Agron. J. 74: 999-1004

Hatano R and Booltink HWG (1992) Using fractal dimensions of stained flow patterns in a clay soil to predict bypass flow. Journal of Hydrology 135: 121-131

Haufe HK (1989) Site characterization of the West Ridge Upland Experimental Area (NAES). In: Nyankpala Agricultural Experimental Station Research Report no. 6, Annual Report 1987/88, NAES, pp139-154

Haverkamp R and Parlange JY 1986. Predicting the water retention curve from particle size distribution:1. Sandy soils without organic matter. Soil Science 142(6): 325-339

Hecht-Nielsen R (1990) Neurocomputing . Addison-Wesley Publ. Co., Reading, MA

Hertz J, Krogh A and Palmer R (1991) Introduction to the Theory of Neural Computation, Addison-Wesley, Reading, MA

Hill T, Marquez L, O'Connor M and Remus W (1994) Artificial neural network models for forecasting and decision making. International Journal of Forecasting 10: 5-15

Hillel D (1998) Environmental soil physics. Academy Press, New York. pp 19-123, 173-201

Hudson BD (1992) The soil survey as paradigm-based science. Soil Science Society of American Journal 56: 836-841

Hutchinson MF and Gallant JC (2000) Digital elevation models and representation of terrain shape. In: Wilson JP and Gallant J.C (eds) Terrain Analysis: Principles and Application, pp 29-50. John Wiley and Sons

Isaak EH and Srivastava (1989) An introduction to applied geostatistics. pp 140-183. Oxford University Press, Oxford

Jabro JD (1992) Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data. American Society of Agricultural Engineers 35:557-560

Jenny H (1941) Factors of soil formation - a system of quantitative pedology. McGraw-Hill, New York, N.Y, pp 281

Jenny H (1980) The Soil Resource: Origin and Behavior, Springer-Verlag, New York, pp 377

Kar S, Varade SB, Subramanyan TK and Ghildyal BP (1976) Soil physical conditions affecting rice root growth: bulk density and submerged soil temperature regime effects, Agronomy journal 68:23-26

Kasei CN (1993) A synopsis on the climate of the north of Ghana. In: proceedings: Workshop on improving farming systems in the interior savanna of Ghana. Nyankpala Agricultural Experimental Station (NAES), Nyankpala, Ghana

Khotanzad A, Afkhami-Rohani R, Lu T-L, Abaye A, Davis M and Maratukulam DJ (1997) ANNSTLF – A neural-network-based electric load forecasting system. IEEE Transactions on Neural-Networks 8(4): 835-846

King JJ and Franzmeier DP (1981) Estimation of saturated hydraulic conductivity from soil morphology and genetic information. Soil Sci. Soc. Am. J. 45: 1153-1156.

Klute A and Dirksen C (1986) Hydraulic conductivity and diffusivity: Laboratory methods. In: Klute A (ed) Methods of Soil Analysis, Part 1, pp 687-734. Madison, WI: Am. Soc. Agron.

Kohler A, Abbaspour KC, Fritsch M and Schuhn R (2003) Using simple bucket models to analyze solute export to subsurface drains by preferential flow. Vadose Zone Journal 2(1): 68-75

Kreznor WR, Olson KR, Banwart, WL and Johnson DL (1989) Soil, landscape and erosion relationships in a northwest Illinois watershed. Soil Science Society of America journal 53:1763-1771

Lachtermacher G and Fuller JD (1994) Backpropagation in hydrological time series forecasting, In: Hipel KW, McLeod AI, Panu US and Singh VP (eds) Backpropagation in hydrological time series forecasting, pp 229-242. Kluwer Academic Publishers, Dordrecht

Lagacherie P, Andrieux P and Bouzigues R (1996) Fuzziness and uncertainty of soil boundaries from reality to coding in GIS. In: Burrough PA and Frank AU (Eds), Geographic objects with intermediate boundaries, pp 275-286 Taylor and Francis, London

Lal R, Sobecki TM, Iivari T, Kimble JM (2003) Soil degradation in the United States: extent, severity and trends, pp 3-21. CRC Press LLC, Boca Raton, Florida

Landon JR (1991) A booker tropical soil manual: A handbook for soil survey and agricultural land evaluation in the tropics and subtropics. Booker Tate, Thame, Oxon, UK, pp 58-125

Maier HR (1995) Use of artificial neural networks for modeling univariate water quality time series, PhD. Thesis, Department of Civil and Environmental Engineering, The University of Adelaide

Maier HR and Dandy GC (1999) Empirical comparison of various methods for training feed-forward neural networks for salinity forecasting. Water Resources Research, vol 35(8): 2591-2596

Maier HR and Dandy GC (2001) Neural network based modeling of environmental variables: a systematic approach, Mathematical and Computer Modeling 33: 669-682

Masters T (1993) Practical neural network recipes in C++, Academic press, San Diego

Matheron G (1971) The Theory of Regionalized Variables and Its Applications, Ecole de Mines, Fontainbleau, France

Mathsoft Inc. (1999) S-Plus 2000 Professional Release 2. MathSoft Inc

McBratney AB, Odeh IOA, Bishop TFA, Dunbar MS and Shatar TM (2000) An overview of pedometric techniques for use in soil survey. Geoderma 97: 293-327

McKeague JA, Wang C and Topp GC (1982) Estimating saturated hydraulic conductivity from soil morphology. Soil Sci. Soc. Am. J. 46: 1239-1244

Mckenzie NJ and MacLeod DA (1989) Relationships between soil morphology and soil properties relevant to irrigated and dryland agriculture. Aust. J. Soil Res. 27:235-258

McKenzie NJ, Gessler PE, Ryan PJ and O'Connell (2000) The role of terrain analysis in soil mapping. In: Wilson JP and. Gallant JC (eds) Terrain Analysis: Principles and Application, pp 245-265. John Wiley and Sons

McKenzie NJ, Jacquie DW, Ashton LJ and Cresswell HP (2000) Estimation of soil properties using the Atlas of Australian soils, CSIRO Land and Water, Technical Report 11/00 (http://www.clw.csiro.au/publications/technical)

Mckenzie NJ, Smettem KRJ and Ringrose-Voase AJ (1991) Evaluation of methods for inferring air and water properties of soils from field morphology. Aust. J. Soil Res. 29:587-602.

Minns AW and Hall MJ (1996) Artificial neural network as rainfall-runoff models. Hydrological Science Journal 41(3): 399-417

Mitasova H, Hofierka J, Zlocha M and Iverson LR (1996) Modeling topographic potential for erosion and deposition using GIS. International Journal of Geographical Information Systems 10: 629-641

MoFA (1991) Agriculture in Ghana: facts and figures. PPME, Ministry of Food and Agriculture (MoFA), pp 1-30.

Montgomery DR and Dietriech WE (1995) Hydrologic processes in low-gradient source area. Water Resources Research 31:1-10

Moore ID and Hutchinson MF (1991) Spatial extension of hydrologic process modeling, In: Proceedings of the International Hydrology and Water Resources Symposium, Perth, 2-4 October 1991, pp 803-808. Canberra: Institute of Australian Engineers

Moore ID, Gessler PE, Nielsen GA and Peterson GA. (1993b) Terrain analysis for soil specific management. In "Soil Specific Crop Management" (P. C. Robert, R. H. Rust, and W. E. Larson (eds.), pp. 27-55. ASA, CSSA, and SSSA, Madison, WI

Moore ID, Gessler PE, Nielson GA and Peterson GA (1993a) Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57:443-452

Moore ID, Grayson RB and Ladson AR (1991) Digital terrain modeling: A review of hydrological, geomorphological and biological applications. Hydrol. Processes 5:3-30

Moore ID, Lewis A and Gallant JC (1993c) Terrain attributes: estimation methods and scale effects. In: Jakeman AJ, Beck MB and McAleer MJ (eds) Modeling change in environmental systems, pp 189-214. New York, Wiley

Moore ID, O'Loughlin EM and Burch GL (1988). A contour-based topographic model for hydrological and ecological applications. Earth Surface Processes and Landforms 13: 305-320

Mualem Y (1976) A new models for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12:513-522

Müller-Sämann KM and Kotschi J (1994) Sustaining growth: soil fertility management in tropical smallholders. pp 81- 138. Margraf Verlag, Weikersheim, Germany.

Nelson PN, Dictor M-C and Soulas (1994) availability of organic carbon in soluble and particle-size fractions from a soil profile. Soil Biol. Biochem. 26(11): 1549-1555

Nizeyimana E and Bicki TJ (1992) Soil and landscape relations in the north central region of Rwanda, East-Central Africa. Soil Science 153(3): 225-236

O'Callaghan JF and Mark DM (1984) The extraction of drainage networks from digital elevation data. Computer Vision, Graphics and Image Processing 28: 323-344

Odeh IOA, Chittleborough DJ and McBratney AB (1991) Elucidation of soil-landform interrelationships by canonical ordination analysis. Geoderma 49: 1-32

Ofori-Sarpong E (1985) The nature of rainfall and soil moisture in the northeastern part of Ghana during the 1975-1977 drought. Geografiska Annaler 67A 3-4, pp 177-186

Pachepsky YA and Rawls WJ (2003) Soil structure and pedotransfer functions, European Journal of Soil Science 54(3): 443-452

Pachepsky, YaA, Timlin, D and Varallyay G (1996) Artificial neural networks to estimate soil water retention from easily measurably data. Soil Sci. Soc. Am. J. 60: 727-773

Panuska JC Moore ID and Kramer L A (1991) Terrain analysis: integration into the agriculture non-point source (AGNPS) pollution model. Journal of Soil and Water Conservation 46: 59-64

Park SJ and Vlek PLG (2002). Environmental correlation of three-dimensional spatial soil variability: A comparison of three adaptive techniques. Geoderma 109: 117-140.

Park SJ McSweeney K and Lowery B (2001) Identification of the spatial distribution of soils using a process-based terrain characterization. Geoderma 103: 249-272

Park SJ, Ruecker GR, Agyare WA, Akramhanov A and Vlek PLG (2003) Sensitivity of spatial resolution and flow routing algorithm to predict the spatial distribution of soil properties, Geoderma (in review)

Pleysier JL and Juo ASR (1980) A single extraction method using silver-thiourea for measuring exchangeable cations and effective CEC in soils with variable charges, Soil Sci. 129: 205-211

Principe JC, Euliano NR and Lefebvre WC (2000) Neural and adaptive systems: Fundamentals through simulations. John Wiley & Sons, New York, pp 173-222, 441-442

Pucket, WE, Dane JH and Hajek BF (1985) Physical and mineralogical data to determine soil hydraulic properties. Soil sci. Soc. Am. J. 49: 831-836

Quinn PF, Beven KJ and Lamb R (1995) The ln(Ãtaɳβ) index: how to calculate it and how to use it within TOPMODEL framework. Hydrological Processes 9: 161-182.

Quinn PF, Beven KJ, Chevallier P and Planchon O (1991) The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrological Processes 5: 59-79

Rawls WJ and Brakensiek DL (1982) Estimating soil water retention from soil water properties, Trans. ASAE 108(IR2): 166-171

Rawls WJ, Ahuja LR and Brakensiek (1992) Estimating soil hydraulic properties from soil data In: Van Genuchten MTh, Leij FJ and Lunds LJ (eds) Proc. Int. Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, October 11-13, 1989, pp 329-340. University of California, Riverside, CA

Rawls WJ, Brakensiek DL and Logsdon SD (1993) Predicting saturated hydraulic conductivity utilizing fractal principles, Soil Science Society of America Journal 57: 1193-1197

Rawls WJ, Brakensiek DL and Saxton KE (1982) Estimation of soil water properties. Transactions of ASAE 25: 1316-1320

Rawls WJ, Gimenez D and Grossmann R (1998) Use of soil texture, bulk density, and slope of water retention curve to predict saturated hydraulic conductivity, Transactions of ASAE, 41(4): 983-988

Reu M and Sposito G (1991) Fractal fragmentation, soil porosity, and soil water properties: I theory, II applications. Soil Science Society of America Journal 55: 1231-1244

Ripley BD (1996) Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge

Ruhe RV and Walker PH (1968) Hillslope models and soil formation. II: Open systems. Transactions of the Ninth Congress of the International Soil Science Society, pp 551-560. Adelaide, Australia

Rumelhart DE, Hinton GE and Williams RJ (1986) Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL (eds) Parallel Distributed Processing, pp 318-362. MIT Press

Sarle WS (1994) Neural network and statistical models, Nineteenth Annual SAS Users Group International Conference, pp1538-1550

Saxton KE, Rawls WJ, Romberger JS and R. I. Papendick. (1986) Estimating generalized soil watercharacteristics from texture. Soil Sci. Soc. Am. J. 50: 1031-1035

Schaap MG and Bouten W (1996) Modeling water retention curves of sandy soils using neural networks. Water Resources Research 32(10): 3033-3040

Schaap MG and Leij FJ (1998) Database-related accuracy and uncertainty of pedotransfer functions. Soil Science 163(10): 765-779

Schaap MG, Leij FL and van Genuchten MTh (1998) Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J. 62: 847-855.

Schaap MG, Leij FL and van Genuchten MTh (1999) A bootstrap-neural network approach to predict soil hydraulic parameters. In: Van Genuchten MTh, Leij FJ and Wu L (eds) Proc. Int. Workshop on characterization and measurement of the hydraulic properties of unsaturated porous media, pp 1237-1250. University of California, Riverside, CA

Schulze R (2000) Transcending scales of space and time in impact studies of climate and climate change on agrohydrological responses. Agriculture, Ecosystems and Environment 82:185-212

Searle PL (1984) the single-extraction silver thiourea method for measuring the cation exchange capacity of soil: some preliminary comments. New Zealand soil News 32(4): 133-136

Shirazi MA and Boersma L (1984) A unifying quantitative analysis of soil texture. Soil Science Society of America 48:142-147

Smith GK (1962) Report on soil and agricultural survey of Sene-Obosum river basins, Ghana: East Brong-Ahafo and Ashanti regions. pp 29-98. United States Agency for International Development mission to Ghana

Speight JG (1974) A parametric approach to landform regions. In: Brown EH and Waters RS (eds) Progress in Geomorphology, pp 213-230. Alden, London

SPSS Inc. (1999) SPSS base 10.0 for Windows User's Guide. SPSS Inc., Chicago IL

Tarboton DG (1997) A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research 33: 309-319

Taylor CJ (1952) The vegetation zones of the Gold Coast, Govt. Printer Forestry Dept. Bull. No. 4. Accra

Thomas, GW (1996) Soil pH and soil acidity. In: Sparks DL (ed) methods of soil analysis, part 3 chemical methods, pp 475-490. Soil Sci. Soc. of Amer., Madison, Wisconsin

Tibshirani R (1994) Comment on 'Neural networks: A review from statistical perspective' by Cheng B and Titterington (initials), Statistical Science 9(1): 48-49

Tietje O and Hennings V (1996) Accuracy of the saturated hydraulic conductivity prediction by pedotransfer functions compared to the variability within FAO textural classes. Geoderma 69: 71-84

Tietje O and Richter O (1992) Stochastic modeling of the unsaturated water flow using auto-correlated spatially variable hydraulic parameters. Modeling Geo-Biosphere Processes 1(2): 163-183

UNEP (2003) Africa environment outlook. http://www.unep.org/aeo/166.htm, 29th December 2003. pp 1-2

Van de Giesen N, Andreini M, van Edig A and Vlek PLG (2001) Competition for water resources of the Volta Basin, Regional Management of Water Resources (Proceedings of symposium held during the sixth IAHS Scientific >Assembly at Maastrict, The Netherlands, july 2001), pp 199-205 IAHS publ. no. 268

van Genuchten MTh (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 4: 892-898

Vapnik V (1995) The Nature of Statistical Learning Theory, Springer Verlag, New York

Vereecken H, Maes J Feyen J (1990) Estimating unsaturated hydraulic conductivity from easily measured soil properties. Soil Science Society of America journal 149: 1-12

Vieux BE 1993 DEM aggregation and smoothing effects on surface runoff modeling. Journal of Computing in Civil Engineering 7:310-338

Vlek PLG (1993) Strategies for sustaining agriculture in sub-Saharan Africa: the fertilizer technology issue. In Technologies for sustainable agriculture in the tropics (eds) ASA, CSSA, SSSA, pp265-277. Madison, WI, US

Vlek PLG, Kühne RF and Denich M (1997) Nutrient resources for crop production in the tropics. Phil. Trans. R. Soc. Lond. B. 352: 975-985

Walker HO (1962) Weather and climate. In: Wills JB (ed) Agriculture and land use in Ghana. Ghana Ministry and Food and Agriculture, pp 3-50. Oxford University press, London

Warrick AW and Nielson DR (1980) Spatial variability of soil physical properties in the field. In: Hillel D (ed) Applications of soil physics, pp 319-344. Academic press, New York

Western AW and Blöschl G (1999) On the spatial scaling of soil moisture. Journal of Hydrology 217: 203-224

White H (1989) Learning in artificial neural networks: A statistical perspective, Neural Comput. 1: 425-464

Wilding LP (1984) Spatial variability: its documentation, accommodation and implication to soil surveys. In: Nielsen DR and Bouma J (eds) Soil Spatial Variability, pp 166–193. Pudoc, Wageningen

Wilding LP and Drees LR (1983) Spatial variability and pedology. In: Wilding LP, Smeck NE and Hall GF (eds) Pedogenesis and soil taxonomy I: concept and interaction, pp 83-116, Elsevier Science Publishers, Amsterdam, Netherlands

Williams RD, Ahuja LR and Naney (1992) Comparison of methods to estimate soil water characteristics from soil texture, bulk density, and limited data. Soil Science 153(3) 172-184

Wilson JP and Gallant JC (2000) Digital Terrain Analysis. In: Wilson JP and. Gallant JC (eds) Terrain Analysis: Principles and Application, pp 1-27. John Wiley and Sons

Wilson JP, Repetto PL and Snyder RD (2000) Effect of data source, grid resolution, and flow routing method on computed topographic attribute. In Wilson JP and. Gallant JC (eds) Terrain Analysis: Principles and Application, pp 133-161. John Wiley and Sons

Wilson MF, Henderson-Sellers A, Dickinson RE and Kennedy PJ (1987) Sensitivity of the Biosphere-Atmosphere Transfer Scheme (BATS) to the inclusion of variable soil characteristics. J. Clim. Appl. Meteorol.26:18-27

Wolock DM and McCabe GJ (1995) Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. Water Resources Research 31:1315-1324

Young A (1976) Tropical soils and survey. pp 467. Cambridge, Cambridge University Press

Young, F.J. and R. D. Hammer. 2000. Defining geographic soil bodies by landscape position, soil taxonomy, and cluster analysis. Soil Sci. Soc. Am. J. 64:989-998

Yu XH and Chen GA (1997) Efficient backpropagation learning using optimal learning rate and momentum. Neural networks 10:517.527

Zevenbergen LW and Thorne CR (1987) Quantitative analysis of land surface topography. Earth Surface Processes and Landforms 12: 47-56

Zhang WH and Montgomery DR (1994) Digital elevation model grid size, landscape representation, and hydrologic simulations. Water Resources Research 30: 1019-1028

Zhu J and Mohanty BP (2002) Spatial averaging of van Genuchten hydraulic parameters for steady-state flow in heterogeneous soils. Vadose Zone Journal 1: 261-272

Zhu, A.-X., L. Band, R. Vertessy and B. Dutton. 1997. Derivation of soil properties using soil land inference model (SoLIM). Soil Sci. Soc. Am. J., 61:523-533

APPENDICES

Appendix 1. Soil series and their classification (FAO)

Soil series	*FAO soil classification*
Tamale site	
Chanalili	Gleyic Plinthosol
Kpelesawgu	Dystric plinthosol
Kumayili	Ferric Acrisol
Lima	Eutric Gleysol
Nyankpala	Plinthic Acrisol
Tingoli	Haplic Luvisol
Volta	Dystric Gleysol
Wenchi	Lithic Leptosol
Ejura site	
Amantin	Haplic Luvisol
Denteso	Dystric Gleysol
Samari	Ferralic Cambisol
Sene	Gleyic Fluvisol
Techiman	Ferric Acrisol

Appendix 2a. Profile description for Ferric Acrisol (TP 1) (Kumayili series)

- 0-10 (Ap) reddish brown loamy sand with weak fine granular structure and clear and smooth boundary
- 10-29 (B) yellowish red loamy sand with weak fine granular structure and clear and smooth boundary
- 29-55 (Bt1) yellowish red loamy sand with weak fine and medium granular structure and gradual and smooth boundary
- 55-82 (Bt2) yellowish red sandy loam with weak fine and medium granular structure and abrupt and smooth boundary
- 82-132 (BtC) dark red sandy clay loam with massive structure

Analytical data for Ferric Acrisol

Horizon	Ap	B	Bt1	Bt2	BtC
pH	0.01	0.26	5.37	4.72	4.67
%SAND	78.88	80.88	77.88	71.88	52.88
%CLAY	1.24	1.24	3.24	5.24	32.24
%SILT	19.88	17.88	18.88	22.88	14.88
%C	1.95	1.10	0.20	0.21	0.10
CECs	3.63	2.31	2.00	2.56	3.94
%B SAT	48.40	30.81	22.37	19.03	74.61
CECc	293.01	185.92	61.62	48.91	12.22

Appendix 2b. Profile description for Dystric Plinthosol (TP2) (Kpelesawgu series)

- 0-9 (Ap) brown loam with moderate fine and medium granular structure and a clear and smooth boundary
- 9-23 (BA) yellowish brown clay with moderate fine and medium subangular blocky structure and a clear and smooth boundary
- 23-40 (Bt1) brown clay with moderate medium subangular blocky structure and a clear and smooth boundary
- 40-54 (Bt2) yellowish brown clay loam with moderate medium and course subangular blocky structure and a clear and smooth boundary
- 54-75 (BC) reddish brown clay loam with moderate medium subangular blocky structure and a clear and smooth boundary
- 75-120 (CB) reddish brown clay loam with weak medium subangular blocky structure

Analytical data for Dystric Plinthosol

Horizon	Ap	BA	Bt1	Bt2	BC	CB
pH	4.85	4.53	4.41	5.28	6.04	6.27
%SAND	66.88	42.88	18.88	30.88	26.88	20.88
%CLAY	2.24	20.24	48.24	40.24	34.24	28.24
%SILT	30.88	36.88	32.88	28.88	38.88	50.88
%C	0.36	0.62	0.75	0.35	0.18	0.07
CECs	3.09	5.46	10.64	10.13	10.46	10.69
%B SAT	29.69	35.86	56.28	80.75	86.62	82.18
CECc	138.11	26.96	22.05	25.18	30.55	37.85

Appendix 2c. Profile description for Eutric Gleysol (TP3) (Lima series)

- 0-10 (Ap) dark yellowish brown sandy loam with moderate fine granular structure and clear and smooth boundary
- 10-29 (AB) brown loam with weak fine and medium subangular blocky structure and clear and smooth boundary
- 29-54 (Btg1) yellowish brown loam with weak fine and medium subangular blocky structure, olive yellow common faint mottles and clear and smooth boundary
- 54-82 (Btg2) light olive brown loam with moderate medium subangular blocky structure, olive yellow common distinct mottles and clear and smooth boundary
- 82-107 (Btg3) dark yellowish brown clay loam with weak fine and medium subangualar blocky structure, olive yellow few faint mottles and an abrupt and smooth boundary
- 107-150 (BtC) brown clay loam with weak fine and medium subangular blocky structure, many iron MnO2 concretions, rock intrusions and absence of roots

Analytical data for Eutric Gleysol

Horizon	*Ap*	*ABg*	*Btg1*	*Btg2*	*Btg3*	*BtC*
pH	4.71	4.18	4.21	4.48	5.86	6.56
%SAND	60.88	38.88	36.88	30.88	26.88	22.88
%CLAY	8.24	12.24	18.24	24.24	34.24	50.24
%SILT	30.88	48.88	44.88	44.88	38.88	26.88
%C	0.72	0.20	0.18	0.12	0.10	0.12
CECs	4.19	8.06	9.79	8.56	7.33	9.17
%B SAT	43.49	21.25	28.51	49.77	78.45	88.93
CECc	50.90	65.88	53.68	35.32	21.41	18.25

Appendix 3a. Profile description for Lithic Leptosol (EP1) (Wenchi series)

- 0-20 (Ap) brown loamy sand with granular and abrupt and smooth boundary
- hard pan

Analytical data for Lithic Leptosol

Horizon	Ap
pH($CaCl_2$)	5.25
%SAND	77.36
%SILT	17.60
%CLAY	4.72
%C	0.58
CECs	2.93
Base Sat	38.55
CECc	62.06

Appendix 3b. Profile description for Ferric Acrisol (EP2) (Techiman series)

- 0-17 (Ap) dark grayish brown sandy loam with granular structure and clear and smooth boundary
- 17-29 (B) yellowish brown sandy loam with fine and medium sub-angular blocky structure and abrupt and smooth boundary
- 29-61 (Bcs) yellowish brown sandy loam with dominant round concretion in granular structure with an abrupt boundary
- 61-90 (Bv) yellowish brown loam with dominant massive angular concretions structure

Analytical data for Ferric Acrisol

Horizon	Ap	B	Bcs	Bv
pH($CaCl_2$)	3.65	4.07	3.77	3.76
%SAND	71.24	63.24	59.24	51.24
%SILT	24.28	25.28	21.28	27.28
%CLAY	3.76	10.76	18.76	20.76
%C	0.60	0.42	0.56	0.59
CECs	1.89	3.71	5.60	6.03
Base Sat	68.33	19.19	17.90	18.68
CECc	50.39	34.50	29.87	29.02

Appendix 3c. Profile description for Haplic Luvisol (EP3) (Amantin series)

- 0-27 (Ap) brown sandy loam with granular structure and gradual and smooth boundary
- 27-40 (AB) yellowish brown sandy loam with granular structure and clear and smooth boundary
- 40-61 (Bt1) dark yellowish brown sandy loam with weak fine and medium sub-angular blocky structure and clear and smooth boundary
- 61-89 (Bt2) yellowish brown sandy loam with fine distinct strong brown mottles, weak fine and medium sub-angular blocky structure and abrupt and smooth boundary
- 89-106 (Btcs) light yellowish brown sandy clay loam dominated by round medium sized concretions with a massive structure

Analytical data for Haplic Luvisol

Horizon	Ap	AB	Bt1	Bt2	Btcs
pH($CaCl_2$)	4.27	3.89	3.55	3.56	3.57
%SAND	73.24	67.24	61.24	61.24	53.24
%SILT	23.28	23.28	23.28	19.28	23.28
%CLAY	2.76	8.76	14.76	18.76	22.76
%C	0.62	0.50	0.48	0.43	0.51
CECs	1.48	2.86	5.25	5.40	6.30
Base Sat	45.97	15.96	8.52	9.28	22.25
CECc	53.65	32.60	35.55	28.79	27.69

Appendix 3d. Profile description for Dystric Gleysol (EP4) (Denteso series)

- 0-21 (Ap) dark grayish brown loamy sand with granular structure and a gradual and smooth boundary
- 21-31 (AB) brown sandy loam with few faint strong brown mottles, weak fine and medium sub-angular blocky structure and clear and smooth boundary
- 31-51 (B1) yellowish brown sandy loam with common faint strong brown mottles, weak fine and medium sub-angular blocky structure and distinct and smooth boundary
- 51-61 (B2) yellowish brown loamy sand with many distinct strong brown mottles and granular structure
- water was coming into the pit so it could not be dig further

Analytical data for Dystric Gleysol

Horizon	Ap	AB	B1	B2
Ph(CaCl2)	5.28	4.38	4.12	4.28
%SAND	73.24	71.24	73.24	81.24
%SILT	25.28	24.28	23.28	17.28
%CLAY	0.76	3.76	2.76	0.76
%C	0.72	0.46	0.11	0.30
CECs	2.30	2.20	1.07	0.87
Base Sat	47.92	18.27	25.26	31.09
CECc	303.17	58.58	38.78	114.57

Appendix 3e. Profile description for Ferric Luvisol (EP5) (Sene series)

- 0-21 (Ap) dark grayish brown sandy loam with weak fine and medium sub-angular blocky structure and gradual and smooth boundary
- 21-46 (B1) brown sandy loam with weak fine and medium sub-angular blocky structure and clear and smooth boundary
- 46-67 (B2) pale brown sand loam with very few fine and round concretion, moderate fine and medium sub-angular blocky structure and gradual and smooth boundary
- 67-103 (B3) pale brown sandy clay loam with few fine and round concretion, moderate medium and coarse sub-angular blocky structure and abrupt and smooth boundary
- hard pan

Analytical data for Ferric Luvisol

Hor.sym	Ap	B1	B2	B3
Ph(CaCl2)	5.29	5.08	4.52	4.12
%SAND	69.36	65.36	69.36	58.36
%SILT	28.60	23.60	11.60	18.60
%CLAY	1.72	10.72	18.72	22.72
%C	0.65	0.82	0.27	0.44
CECs	2.22	2.74	3.97	5.22
Base Sat	50.54	59.88	57.16	46.38
CECc	129.30	25.57	21.20	22.99

Appendix 4a. Kolmogorov-Smirnov normality test statistic and significance, and Kruskal-Wallis mean rank for different soil parameters at the Tamale site

Soil parameter	Soil type[a]	Topsoil			Subsoil		
		KS[b] normality		KW[c] rank	KS normality		KW rank
		Statistics	p		Statistics	p	
Sand content	HL	0.25	0.20	160.30	0.22	0.20	100.80
	LL	0.14	0.20	161.56	0.30		104.63
	FA	0.08	0.20	128.90	0.13	0.13	109.43
	PA	0.11	0.09	161.08	0.09	0.20	114.02
	DP	0.11	0.10	116.57	0.10	0.20	83.80
	EP	0.22		159.50	0.26		141.00
	EG	0.06	0.20	109.10	0.10	0.18	107.61
	DG	0.09	0.20	46.70	0.18	0.20	86.77
Silt content	HL	0.31	0.20	95.50	0.21	0.20	55.60
	LL	0.33	0.20	86.67	0.42		143.75
	FA	0.15	0.20	119.15	0.13	0.11	90.04
	PA	0.20	0.08	84.01	0.11	0.20	79.10
	DP	0.13	0.20	134.85	0.05	0.20	104.83
	EP	0.23		100.13	0.26		30.25
	EG	0.14	0.06	149.43	0.07	0.20	121.71
	DG	0.19	0.20	198.30	0.17	0.20	137.31
Clay content	HL	0.18	0.12	95.80	0.22	0.20	125.20
	LL	0.15	0.01	86.39	0.24		76.63
	FA	0.07	0.01	129.13	0.12	0.20	99.64
	PA	0.11	0.00	122.70	0.07	0.20	110.74
	DP	0.07	0.02	130.21	0.08	0.20	122.99
	EP	0.27		92.88	0.26		90.00
	EG	0.11	0.00	121.49	0.12	0.04	85.80
	DG	0.18	0.15	167.53	0.15	0.20	104.08
Organic carbon	HL	0.35	0.05	113.50	0.26	0.20	109.50
	LL	0.19	0.20	151.50	0.30		146.25
	FA	0.20	0.00	116.48	0.13	0.16	95.92
	PA	0.14	0.02	120.90	0.09	0.20	113.28
	DP	0.11	0.10	129.35	0.13	0.06	114.45
	EP	0.16		55.75	0.26		61.00
	EG	0.09	0.20	129.02	0.16	0.00	92.66
	DG	0.14	0.20	148.17	0.17	0.20	91.50

[a]Haplic Luvisol (HL) Lithic Leptosol (LL), Ferric Acrisol (FA), Plinthic Acrisol (PA), Dystric Plinthosol (DP), Eutric Plinthosol (EP), Eutric Gleysol (EG), and Dystric Gleysol (DG); [b]KS: Kolmogorov-Smirnov normality test ($p<0.05$ data significant from normal distribution); [c]KW: Kruskal-Wallis mean rank

Appendix 4b. Kolmogorov-Smirnov normality test statistic and significance, and Kruskal-Wallis mean rank for different soil parameters at the Tamale site

Soil parameter	Soil type[a]	Topsoil			Subsoil		
		KS[b] normality		KW[c] rank	KS normality		KW rank
		Statistics	p		Statistics	p	
CEC	HL	0.31	0.14	112.20	0.26	0.20	97.80
	LL	0.32	0.01	115.89	0.27		65.25
	FA	0.13	0.07	149.54	0.16	0.02	111.14
	PA	0.20	0.00	124.93	0.16	0.01	101.44
	DP	0.15	0.01	125.37	0.13	0.09	114.44
	EP	0.22		102.75	0.26		39.50
	EG	0.20	0.00	119.29	0.16	0.00	102.56
	DG	0.22	0.04	97.47	0.19	0.20	73.00
pH	HL	0.17	0.20	129.50	0.20	0.20	119.70
	LL	0.30	0.02	179.22	0.26		90.25
	FA	0.06	0.20	167.66	0.16	0.03	127.85
	PA	0.09	0.20	143.25	0.14	0.03	104.54
	DP	0.11	0.17	114.36	0.13	0.05	107.92
	EP	0.24		100.50	0.26		75.75
	EG	0.17	0.00	92.85	0.15	0.00	88.66
	DG	0.14	0.20	85.50	0.19	0.20	82.04
Bulk density	HL	0.34	0.06	132.30	0.24	0.20	126.00
	LL	0.17	0.20	166.06	0.22		87.88
	FA	0.09	0.20	121.84	0.12	0.20	86.43
	PA	0.10	0.20	175.21	0.09	0.20	130.87
	DP	0.08	0.20	116.62	0.11	0.20	97.17
	EP	0.24		86.13	0.26		4.00
	EG	0.09	0.20	100.51	0.12	0.04	104.12
	DG	0.11	0.20	80.77	0.08	0.20	86.04
K_s	HL	0.23	0.20	140.60	0.34	0.05	147.30
	LL	0.27	0.06	193.67	0.43		81.75
	FA	0.24	0.00	120.71	0.26	0.00	105.03
	PA	0.20	0.00	147.80	0.22	0.00	154.70
	DP	0.24	0.00	115.39	0.34	0.00	92.41
	EP	0.26		101.00	0.26		102.25
	EG	0.21	0.00	113.30	0.33	0.00	81.48
	DG	0.16	0.20	108.43	0.28	0.01	60.00

[a]Haplic Luvisol (HL) Lithic Leptosol (LL), Ferric Acrisol (FA), Plinthic Acrisol (PA), Dystric Plinthosol (DP), Eutric Plinthosol (EP), Eutric Gleysol (EG), and Dystric Gleysol (DG); [b]KS: Kolmogorov-Smirnov normality test ($p<0.05$ data significant from normal distribution); [c]KW: Kruskal-Wallis mean rank

Appendix 5. Kolmogorov-Smirnov normality test statistic and significance, and Kruskal-Wallis mean rank for different soil data at the Ejura site

Soil parameter	Soil type[a]	Topsoil			Subsoil		
		KS[b] normality		KW[c] rank	KS normality		KW rank
		Statistics	p		Statistics	p	
Sand	FC	0.16	0.09	159.57	0.15	0.20	108.16
	FA	0.05	0.20	165.93	0.13	0.00	138.81
	HA	0.08	0.04	218.81	0.07	0.20	201.09
	GA	0.09	0.20	196.33	0.08	0.20	216.31
	GF	0.22	0.03	101.41	0.19	0.14	113.29
Silt	FC	0.15	0.10	215.17	0.20	0.06	224.37
	FA	0.09	0.02	214.48	0.13	0.00	206.45
	HA	0.05	0.20	153.35	0.06	0.20	141.56
	GA	0.08	0.20	184.24	0.11	0.06	165.68
	GF	0.20	0.07	233.38	0.22	0.04	212.53
Clay	FC	0.15	0.11	203.70	0.17	0.20	218.63
	FA	0.10	0.00	199.05	0.06	0.20	206.49
	HA	0.15	0.00	176.29	0.07	0.20	172.02
	GA	0.18	0.00	154.89	0.16	0.00	102.76
	GF	0.28	0.00	309.21	0.14	0.20	241.24
Organic carbon	FC	0.17	0.04	181.94	0.18	0.17	169.45
	FA	0.09	0.02	151.59	0.09	0.02	163.23
	HA	0.12	0.00	168.09	0.16	0.00	154.15
	GA	0.09	0.18	277.02	0.04	0.20	236.81
	GF	0.17	0.18	217.32	0.16	0.20	165.41
CEC	FC	0.25	0.00	194.06	0.15	0.20	158.39
	FA	0.07	0.20	195.73	0.06	0.20	142.54
	HA	0.07	0.07	174.79	0.08	0.03	190.70
	GA	0.06	0.20	190.03	0.09	0.20	192.72
	GF	0.17	0.20	207.24	0.17	0.20	210.65
pH	FC	0.18	0.03	174.43	0.22	0.03	157.13
	FA	0.07	0.20	178.80	0.18	0.00	152.46
	HA	0.10	0.00	186.76	0.11	0.00	181.40
	GA	0.07	0.20	216.82	0.12	0.03	210.65
	GF	0.15	0.20	148.94	0.26	0.01	145.62
Bulk density	FC	0.12	0.20	205.35	0.17	0.20	203.55
	FA	0.06	0.20	183.82	0.08	0.08	189.64
	HA	0.04	0.20	191.57	0.12	0.00	162.83
	GA	0.05	0.20	171.54	0.09	0.20	164.04
	GF	0.17	0.20	222.94	0.14	0.20	187.18
Ks	FC	0.21	0.00	162.96	0.28	0.00	223.71
	FA	0.18	0.00	184.84	0.37	0.00	160.62
	HA	0.17	0.00	185.13	0.33	0.00	181.74
	GA	0.21	0.00	211.54	0.28	0.00	164.51
	GF	0.20	0.06	159.62	0.35	0.00	105.47

[a]Soil type Ferralic Cambisol (FC), Ferric Acrisol (FA), Haplic Acrisol (HA), Gleyic Acrisol (GA), and Gleyic Fluvisol; [b]KS: Kolmogorov-Smirnov normality test ($p<0.05$ data significant from normal distribution); [c]KW: Kruskal-Wallis mean rank

Appendix 6. Sensitivity of input data at two sites and sampling depths with their standard deviation in ( ) from ANN

Input parameters	Ejura topsoil	Ejura subsoil	Tamale topsoil	Tamale subsoil
Profile curvature*	0.096 (0.054)	0.059 (0.050)	0.039 (0.051)	0.039 (0.044)
Plan curvature	0.057 (0.043)	0.074 (0.044)	0.046 (0.035)	0.048 (0.026)
Curvature	0.054 (0.047)	0.018 (0.008)	0.068 (0.049)	0.044 (0.054)
Elevation	0.026 (0.025)	0.050 (0.040)	0.036 (0.030)	0.078 (0.039)
Wetness index	0.056 (0.078)	0.050 (0.036)	0.059 (0.061)	0.039 (0.050)
Upslope area	0.043 (0.029)	0.056 (0.043)	0.033 (0.028)	0.044 (0.040)
Stream power index*	0.057 (0.059)	0.053 (0.044)	0.083 (0.034)	0.066 (0.042)
Slope gradient	0.083 (0.065)	0.054 (0.028)	0.032 (0.021)	0.041 (0.045)
LS factor*	0.129 (0.095)	0.072 (0.048)	0.049 (0.040)	0.038 (0.033)
Aspect*	0.079 (0.029)	0.036 (0.013)	0.052 (0.039)	0.068 (0.045)
Sand*	0.209 (0.085)	0.089 (0.038)	0.086 (0.053)	0.099 (0.061)
Clay*	0.055 (0.038)	0.135 (0.075)	0.033 (0.026)	0.174 (0.103)
Silt*	0.162 (0.046)	0.104 (0.059)	0.201 (0.078)	0.111 (0.107)
CEC*	0.034 (0.028)	0.127 (0.059)	0.119 (0.035)	0.072 (0.067)
Organic carbon*	0.071 (0.041)	0.071 (0.035)	0.053 (0.021)	0.082 (0.042)
Bulk density*	0.351 (0.047)	0.135 (0.055)	0.121 (0.067)	0.129 (0.086)
pH	0.033 (0.021)	0.038 (0.021)	0.027 (024)	0.067 (0.046)
Gravel/concretion	0.054 (0.042)	0.000 (0.000)	0.061 (0.031)	0.027 (0.023)
Sabangular blocky structure	0.048 (0.033)	0.050 (0.034)	0.041 (0.027)	0.041(0.017)
Grainular structure	0.049 (0.049)	0.051 (0.064)	0.024 (0.019)	0.042 (0.030)
Weak structure	0.034 (0.023)	0.044 (0.028)	0.025(0.019)	0.037 (0.018)
Mod. strong structure	0.033 (0.027)	0.047 (0.028)	0.035 (0.022)	0.042 (0.023)
Strong structure	0.058 (0.050)	0.065 (0.051)	0.000 (0.081)	0.041 (0.031)
Fine structural size	0.085 (0.049)	0.047 (0.033)	0.038 (0.032)	0.040 (0.022)
Medium structural size	0.046 (0.035)	0.054 (0.036)	0.030 (0.027)	0.038 (0.032)
Course structural size	0.052 (0.042)	0.085 (0.60)	0.043 (0.043)	0.031 (0.020)

* Parameters selected for further analysis

ACKNOWLEDGEMENT

I wish to express my gratitude to the German Government through its Ministry of Education, Science, and Technology (BMBF) for its financial support, and to the gracious staff of ZEFc, GLOWA project and ZEF Doctorial Secretariat for their technical and administrative support. This gratitude also goes to the management and staff of SARI, my research home since 1996, for their technical support and to the mother institute CSIR for the grant of study leave.

On a more personal level, I wish to say a special thanks to Prof. Dr. Paul L.G. Vlek for his able guidance, invaluable advice and thoughtful comments. The support of my tutor Dr. Soojin Park whose, comments did much to improve this write-up, is gratefully acknowledged. Special thanks also to Prof. Dr. Richard Dikau for his interest and willingness to take time to read the write-up and give comments and advice. The comments of Prof. A. Skowronek on parts of this write-up are appreciated.

Thanks to Prof. M. Bonsu (UCC), Messer Boasiako Ohene Antwi, Dela C. Dedzoe, James K. Seneya, and Owusu Dwomo (all of SRI), to Alhassan I. Zakaria and the staff of Soil Chemistry Laboratory at SARI for their contribution during the field work. Also my heartfelt thanks go to my numerous friends, colleagues and teachers who have played a part in this work or before in my academic carrier.

Finally, my profound gratitude goes to my family. To my mother Grace Asiedu and grandmother Ama Oye, through whose toil and sweat I had my basic, secondary and first-degree education: I say thank you very much for what you gave me. To my brothers Alex and George, my sister Mercy and her husband Ernest: thanks for your prayer and support. To my son Kelvin I wish to say a big thank you for being a good boy during my absence from home. Last but not least, my heart-felt and deepest gratitude goes to my wife Juliana for being a continuous source of inspiration, encouragement, prayer and understanding before and during the execution of this study.