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

Zentrum für Entwicklungsforschung
Center for Development Research
University of Bonn

ZEF Bonn

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

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2004
Zugl.: Bonn, Univ., Diss., 2004
ISBN 3-86537-100-0

D 98

1. Referent : Prof. Dr. Paul L.G. Vlek

2. Referent : Prof. Dr. Richard Dikau

Tag der Promotion: 19. April 2004

Angefertigt mit Genehmigung der Hohen Landwirtschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

⊕ CUVILLIER VERLAG, Göttingen 2004
Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2004
Gedruckt auf säurefreiem Papier

ISBN 3-86537-100-0

To my mother (Grace Asiedu) and my wife (Juliana Agyare).

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) at the Tamale site. 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 the 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 correlation and deviation 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 with little or no use of fertilizer leads to soil nutrient mining 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 modeling 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 properties

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 low scale of 1:250,000 and provides 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 also have 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 % 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 Δ 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 pedo-transfer 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.