

# Estimating Soil Bulk Density and Total Nitrogen from Catchment Attributes in Northern Ethiopia

**Gebreyesus Brhane Tesfahunegn**

College of Agriculture, Aksum University-Shire Campus, P.O. Box 314, Shire, Ethiopia  
(gebre042001@yahoo.com)

## ABSTRACT

Even though data on soil bulk density (BD) and total nitrogen (TN) are essential for planning modern farming techniques, their data availability is limited for many applications in the developing world. This study is designed to estimate BD and TN from soil properties, land-use systems, soil types and landforms in the Mai-Negus catchment, northern Ethiopia using stepwise multiple regression (SMR) and generalized linear model (GLM) analyses. Different soil properties and other catchment attributes were collected following a standard procedure. The SMR analysis showed overall model coefficients of determination ( $R^2$ ) of 0.91 and 0.89 with significant F-statistics for the relationships of BD and TN, respectively, with soil properties, land-use and landforms. In addition, the GLM analysis resulted in an overall  $R^2$  of 0.92 with significant F-statistic for BD, and an  $R^2$  of 0.94 with significant F-statistic for TN. The model coefficients of both analyses for the dependent variables showed higher for organic carbon (OC) as compared to the other variables even though higher values were found from GLM. This study thus confirmed that practices which improve OC can strongly influence the variation of both dependent variables. This study suggested that BD and TN should be estimated based on the relationships explained by the different techniques (analysis) in similar conditions in order to improve data availability; however, the GLM is preferable as it considers the effect of the interaction terms.

**Keywords:** Bulk density, Total nitrogen, Stepwise multiple regression, Generalized linear model, Mai-Negus catchment.

## 1. INTRODUCTION

The need to achieve sustainable use of soil resource has been an increasing concern to decision and policy makers. This is mainly the concern of many developing countries like Ethiopia, because soil degradation such as soil nutrient depletion and physical degradation have alarmingly increased and become serious threats to agricultural productivity (Fassil Kebede and Yamoah, 2009). Previous studies have showed that low soil nutrients are among the most crop production limiting factors in Ethiopia (e.g., Kamara and Haque, 1988; Chikowo et al., 2010). The positive feedback dynamics between growing population, land-cover and climate change have led to a rapid loss in the capacity of soils to deliver essential nutrients such as nitrogen (Gebreyesus Brhane Tesfahunegn et al., 2011). Other studies also reported that most crop and soil management practices and topography of tropical ecosystems (e.g., Ethiopia) could cause significant modifications in soil properties such as nitrogen and bulk density as such factors influence erosion and nutrient losses (Lal, 1996; Shepherd et al., 2000; Agoumé and Birang, 2009). However, literature showed that modifications of soil

biological and chemical properties are more rapid than soil physical properties (Schipper and Sparling, 2000; Birang et al., 2003; Agoumé and Birang, 2009).

Strong relationships between plant communities and soils, and landscapes are documented by various studies (e.g., Wondzell et al., 1996). Such studies suggest that the patterns of plant communities vary as a result of the influence of soil properties on soil water and nutrient availability across landscapes. However, site specific understanding on the relationships of soil nutrient (e.g., total nitrogen, TN) and soil physical parameter (e.g., bulk density, BD) with catchment attributes (soil properties, landform, land-use systems, and soil types) demands research attention in fragile ecosystems in order to highlight the main influencing parameters. To do so, there are many cases in which it is desirable to develop empirical relationships among some soil physical and chemical properties and with other catchment attributes (Rashidi and Seilsepour, 2009; Seybold et al., 2009). This is based on the fact that BD and TN are often determined using laborious and time consuming methods but these soil parameters can be estimated on the basis of easily available parameters (Périé and Ouimet, 2008). In addition, dense sampling is required to adequately characterize the spatial variability of soil BD and TN, which may not be economically affordable (Rashidi and Seilsepour, 2009). It may be thus more suitable and economically feasible if a method which uses easily available soil data and other catchment attributes that influence soil conditions in a catchment is developed to serve as a proxy for estimating soil parameters such as BD and TN (Heuscher et al., 2005; Rashidi and Seilsepour, 2009; Périé and Ouimet, 2008).

Literature showed that regression models have been developed to predict soil parameters such as BD and TN from soil physical and chemical datasets (Akgül and Özdemir, 1996; Heuscher et al., 2005). A linear regression model for predicting soil TN from soil organic carbon (OC) by Rashidi and Seilsepour (2009) and logarithmic regression for BD estimation by Prévost (2004) and Périé and Ouimet (2008) were obtained. According to Heuscher et al. (2005), stepwise multiple regressions indicated that OC was the strongest contributor to BD prediction. However, such estimations of TN and BD were as a function of soil OC only or using some soil properties, indicating that there is a limitation in the past studies as mainly focused on specific or limited soil datasets. Consequently, there is lack of published results that demonstrated the relationship of BD and TN with other soil properties and catchment attributes such as soil types, land-use types, and landforms in the existing literature (Wagner et al., 1994; Akgül and Özdemir, 1996; Heuscher et al., 2005; Rashidi and Seilsepour, 2009). In general, the importance of soil nutrients such as TN for plant production is clearly demonstrated by experts and farmers in Ethiopia conditions. For example, Fassil Kebede and

Yamoah (2009); and Rashidi and Seilsepour (2009) have reported that nitrogen (N) is among the soil macronutrients that often limit plant growth. Despite of this fact, presently, there is no generally adopted and completely reliable and easy method for prediction of nitrogen in soil. Knowledge of soil BD is also an essential indicator for soil management, as BD shows soil compaction and structural degradation. In addition, soil BD is often used in models, characterizing field conditions, estimate soil porosity, and convert volumetric measurements (Reinsch and Grossman, 1995; Arshad et al., 1996). Other studies showed that soil BD is a basic soil property which influences other soil physical and chemical properties and catchment characteristics (e.g., Arshad et al., 1996; Périé and Ouimet (2008)). Site specific knowledge of soil BD and TN are essential for planning modern farming techniques and thus for sustainable management of soil resources (Akgül and Özdemir, 1996).

Despite the above facts, soil BD and TN are dynamic soil property which vary with the structural condition of the soil as this can be altered by cultivation, trampling by animals, land-use types, erosion-deposition processes, and weather condition such as raindrop impact (Reinsch and Grossman, 1995; Akgül and Özdemir, 1996; Arshad et al., 1996). This makes measuring and/or estimating BD and TN difficult, time consuming and expensive at catchment-scale. As a result of this, many studies use single BD value in most cases while estimating for other locations. Similarly, direct laboratory measurement of soil TN is generally impractical to get data for most applications. It is therefore essential to explore methods that are required for predicting soil BD and TN from more easily and routinely measured soil properties and other catchment attributes.

Though few studies illustrated the development of BD and TN prediction equations using a range of soil data (e.g., Arshad et al., 1996), little is documented on the relationships of soil BD and TN with soil properties and other catchment attributes such as land-use systems, soil types, and landforms (small units of the landscape that possess similar slope, flow and deposition). However, acquiring scientific information on their relationships of a combined and separate effect on BD and TN is required so as to guide for decision makers in the choice of appropriate cropping systems and suitable land-use and soil management practices in a catchment (Aruleba and Ajayi, 2011).

The objective of this study is to examine the relationships of soil BD and TN with selected soil properties, land-use systems, soil types, and landforms in Mai-Negus catchment, northern Ethiopia, using stepwise multiple regression (SMR) and generalized linear model (GLM). The target of this study is to identify the main and interaction terms of soil parameters and other catchment attributes, and assesses the relevance of the data groups (soil properties, soil

type, land-use systems and landform) in predicting BD and TN using the two statistical models. Such prediction models improve the availability of soil BD and TN data for researchers and development workers and their interpretations can be applied for decision-making processes such as soil-plant management planning in similar conditions.

## 2. MATERIALS AND METHODS

### 2.1. Study Site

This study was conducted in the Mai-Negus catchment of Tigray region ( $12^{\circ}00' - 15^{\circ}00'$  N latitude and  $36^{\circ}30' - 41^{\circ}30'$  E longitude), northern Ethiopia (Fig 1). The catchment has an area of 1240 ha and altitude ranging from 2060 to 2650 m above sea level. The study catchment has a mean annual temperature of  $22^{\circ}\text{C}$  and precipitation of 700 mm (Aksum Meteorological Station, 2009). Most of the rainfall ( $> 70\%$ ) occurs within July and August. Land use is predominantly arable, with teff (*Eragrostis tef*) being the major crop along with different proportions of pasture land and scattered patches of trees, bushes and shrubs (Gebreyesus Brhane Tesfahunegn, 2011). The major rock types are lava pyroclastic and meta-volcanic. Soils are mainly Leptosols on the very steep positions, Cambisols on middle to steep slopes and Vertisols on the flat areas (Gebreyesus Brhane Tesfahunegn, 2011).

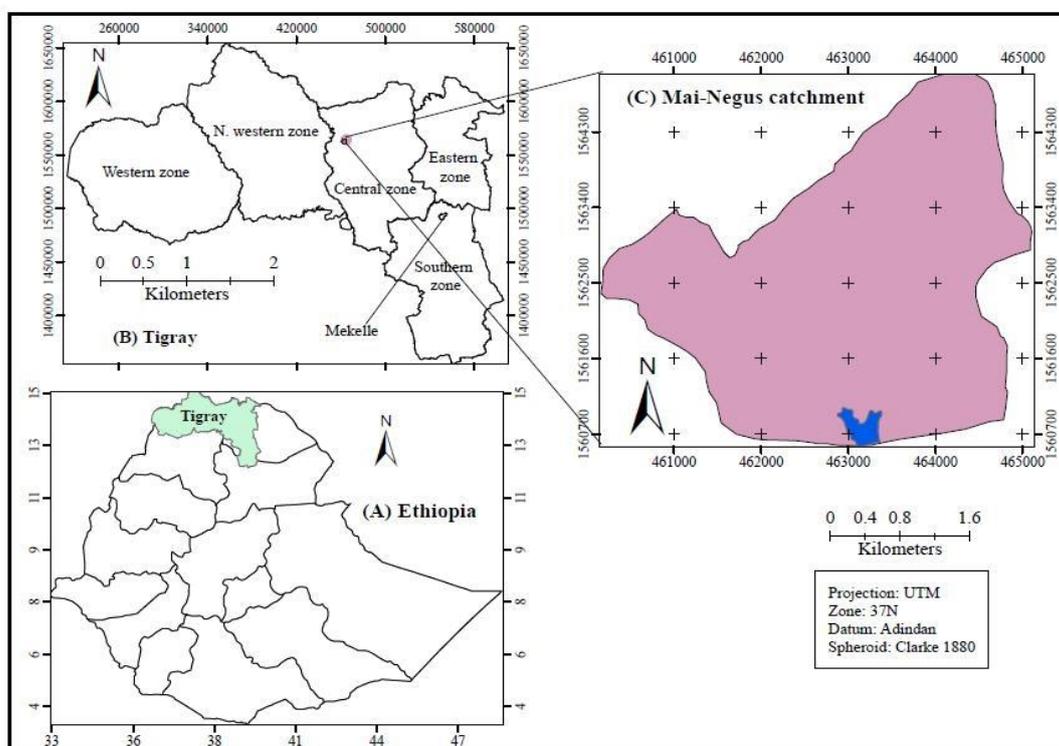


Figure 1. Study area: Ethiopia (A), Tigray (B) and Mai-Negus catchment (C). The blue color shaded area is the reservoir.

## 2.2. Generating Catchment Landforms, Elevation, Soil Types, and Land-Use Types

In order to develop the different catchment attributes, field reconnaissance surveys and informal group discussions were executed with a team consisting of the author, two development agents and six farmers who are knowledgeable about the study catchment. Data were collected from June to December 2009 using the research framework which is given in figure 2. The detail data collection and generating procedures of each catchment attribute is given below.

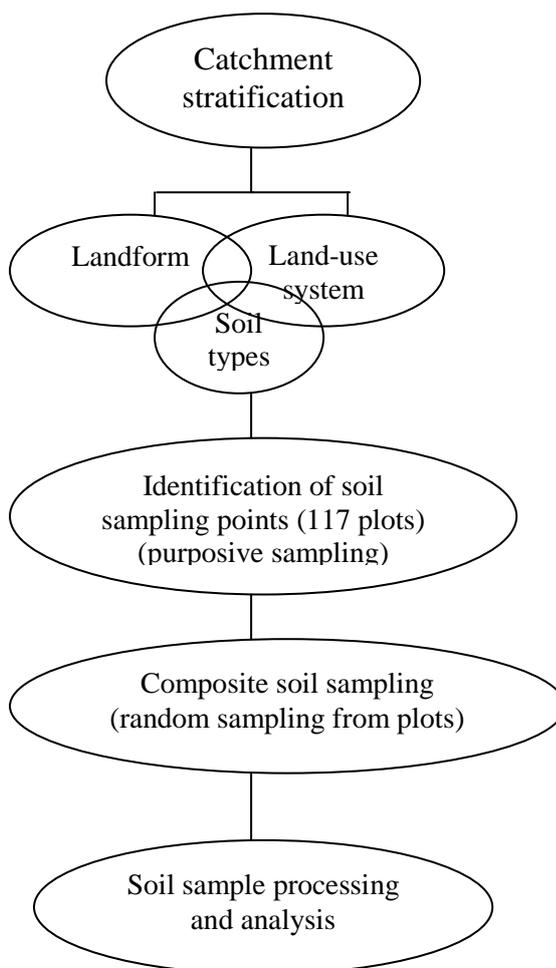


Figure 2. Research framework used during this study. Note that the intersections of attributes (landform, land-use, soil types) indicate that their overlap in the catchment.

The landforms in the catchment (Fig 3A) were classified using data from field and topographic map of the area. Considering elevation, slope, and geomorphologic character (surface and subsurface flows, alluvial and colluvial deposition), the catchment was classified into six main landforms in ArcGIS (Fig 3A). These are the valley (19% of the catchment area), plateau (8%), rolling hill (9%), central ridge (27%), escarpment (29%), and mountainous (6%). The reservoir was considered as a separate landform, which covered about 2% of the catchment area. The elevation pattern of the catchment is shown in figure

3B. The elevation map of the study catchment was developed in ArcGIS from topographic map (scale 1:50,000) obtained from Ethiopian Mapping Agency (EMA, 1997). The topographic map was scanned, geo-referenced using ground control coordinate points and contours and spot heights were digitized and tagged with elevation values in a GIS environment.

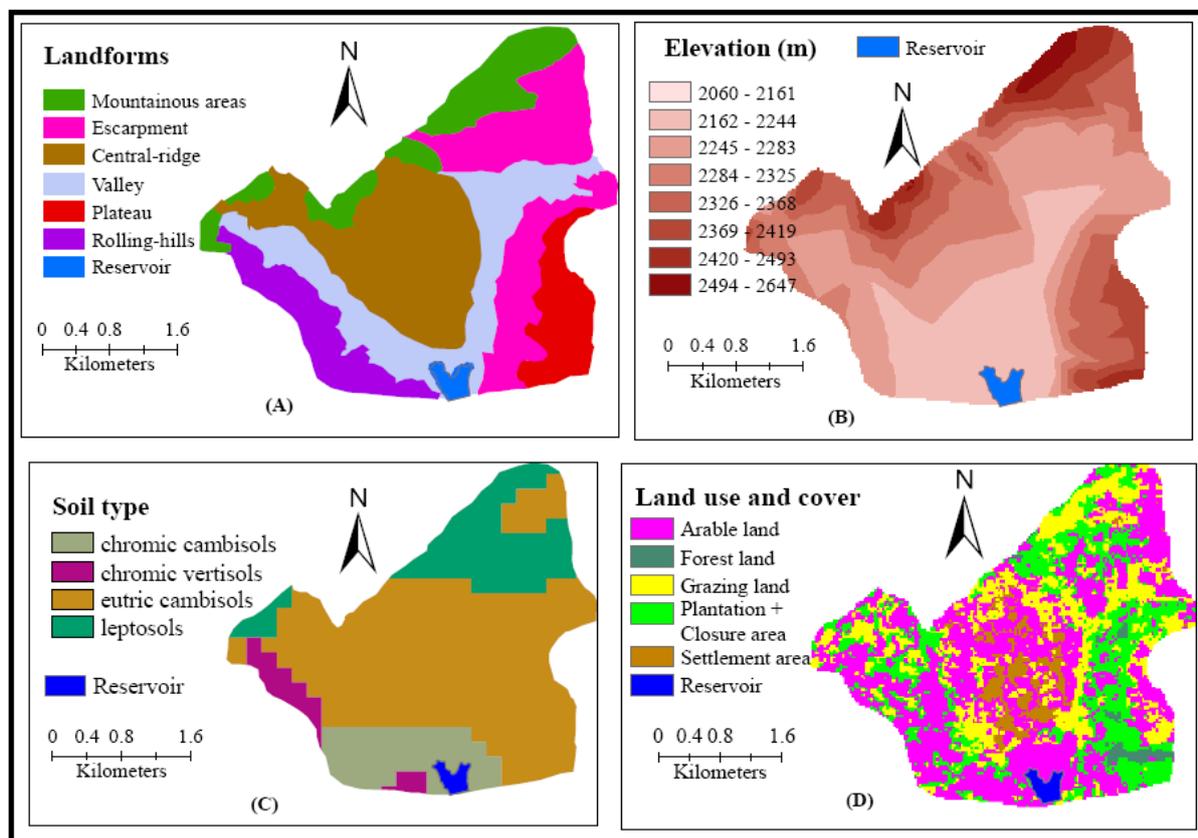


Figure 3. Landform (A); digital elevation model (B); major soil types (C) and land-use and land-cover (D) of the Mai-Negus catchment, northern Ethiopia

The soil data such as available water content, soil texture, bulk density, hydraulic conductivity were obtained from the NEDECO database (NEDECO, 1998). Additional soil physical, chemical and morphological properties were determined on-site to supplement the data gap in the database. According to the FAO-UNESCO (1974) Soil Classification System, the main soil types in the study catchment are Eutric Cambisols (67%), Chromic Cambisols (13%), Leptosols (15%) and Chromic Vertisols (5%). Chromic Cambisols and Vertisols occupy almost the flat areas, Eutric Cambisols the undulating plains and rolling land and Leptosols steep to very steep lands (Fig 3C).

Landsat image of November 2007 supplemented by field survey was used to derive land-use and land-cover (LULC) types. The LULC image was rubber sheeted to match ground control locations, and the area containing only the Mai-Negus catchment was extracted from the full

scene. A root mean square error less than 5 m was achieved while geo-referenced. Ground truthing was conducted in the dry season between mid-September and November 2009 using a Geographic Positioning System (GPS) (*Garmin III*). The GPS points were used to geo-reference the image and as training samples for supervised classification of the Landsat image 30 m x 30 m resolution. Sampling points ranging from 10 to 15 were selected for each land-use class. Maximum likelihood classifier was then applied to produce six LULC classes (Fig. 3D). Overall accuracy of the land use classification was satisfactory in which about 85.6% was correctly classified. Because of the relatively similar reflectance between grazing areas and cultivated fields, different enhancement and transformation techniques and ground truth data were used to enhance separating of cover types and delineate training areas.

About 55% of the land area in the catchment was classified as arable land, 21% for grazing, and 14% as enclosure. Dense bush and woodland with mixed forest accounted for about 2% of the catchment. The rest of the land was miscellaneous such as settlement, marginal area and reservoir (8%). A farmland after harvest is used as a grazing land to use grasses in the field and around the boarders for short time (not more than an hour). Grasses in an enclosure (landscape under rehabilitation) was used for livestock by cut and carrying system. Part of the areas covered by bush and wood land were also opened for grazing but these are dominated by unpalatable species and in many cases due to their shading effects grasses are rarely grown. The data on cropping systems and soil management practices were collected through informal group discussions with local farmers and development agents in the study catchment.

### **2.3. Soil Sampling and Analysis**

During soil sample collection, the locations of the soil sampling points considered the spatial distribution of the different landforms, land-use and land-cover and soil types in the catchment as presented in table 1. A total of 117 soil samples were collected and analysed from the sampling points (plots) with an area ranged 150 to 300 m<sup>2</sup> in the catchment. The number of soil samples varied with size of the sampling point in the landforms, land-use and soil types. A sampling plot was located considering its homogeneity in hydrological condition on the basis of researcher's field observations. From each soil sampling points composite samples of 5 to 8 at 0-20 cm soil depth were collected since this depth is where most changes are expected to occur due to natural and anthropogenic activities. The soil samples were higher in the sampling points that possessed larger plots. Disturbed soil samples from each sampling point were air-dried, pooled, homogenized and sieved to pass through 2-mm sieve.

Table 1. Soil parameters and other catchment attributes used for SMR and GLM analyses under each data group in the Mai-Negus catchment, northern Ethiopia.

<i>Soil properties (n = 117)</i>	<i>Soil types</i>	<i>Land-use types</i>	<i>Landforms</i>
Sand	Chromic Cambisols (n = 34)	Forest land (n = 12)	Rolling-hills (n = 15)
Clay	Chromic Vertisols (n = 13)	Plantation protected	Mountainous (n = 15)
Silt	Eutric Cambisols (n = 54)	Land (n = 16)	
Bulk density	Leptosols (n = 16)	Grazed land (n = 13)	Central-ridge (n = 28)
pH		Teff ( <i>Eragrostis tef</i> )-	Valley(n = 24)
EC		Faba bean ( <i>Vicia faba</i> )	Plateau (n = 12)
Exchangeable Na		rotation (n = 22)	Escarpment (n = 17)
Exchangeable K		Teff-wheat ( <i>Triticum</i>	Reservoir (n = 6)
Exchangeable Ca		<i>vulgare</i> )/ Barley	
Exchangeable		( <i>Hordeum vulgare</i> )	
Mg Sum of basic cations		rotation (n = 17)	
CEC		Teff mono-cropping	
Organic carbon		(n = 13)	
Total nitrogen		Maize ( <i>Zea mays</i> )	
Available phosphorus		mono-cropping (n = 13)	
Total phosphorus		Uncultivated marginal land (n = 11)	
ESP			
Base saturation percentage			
Fe & Zn			

SMR, stepwise multiple regression; GLM, generalized linear model; n is number of samples which were taken by considering the area coverage of the factors in the catchment.

Soil samples were determined for soil texture using the Bouyoucos hydrometer method (Gee and Bauder, 1986), soil bulk density (BD) by the core method (Blake and Hartage, 1986), electrical conductivity (EC) by an EC meter in a 1:2.5 soil to water suspension (Rhoades, 1982a), soil pH by suspending the soil solution in a 1:2.5 soil to water ratio using a pH-meter and a combined glass electrode (Thomas, 1996). Soil organic carbon (OC) was determined by the Walkley-Black method (Bremmer and Mulvaney, 1982), available phosphorus (Pav) by Olsen (Olsen and Sommers, 1982), total phosphorus (TP) extracted by HClO<sub>4</sub> digestion determined calorimetrically (Jackson, 1964) and total nitrogen (TN) by the Kjeldhal Digestion method (Anderson and Ingram, 1993). Cation exchange capacity (CEC) was determined by ammonium acetate extraction buffered at pH 7 (Rhoades, 1982b). Exchangeable bases (Ca, Mg, K, Na) were analyzed after extraction using 1M ammonium acetate at pH 7.0. Readings for Ca and Mg in the extracts were determined using an atomic absorption spectrophotometer, while Na and K were determined by flame photometry (Black et al., 1965). Exchangeable sodium percentage (ESP) was calculated by dividing exchangeable Na<sup>+</sup> by CEC. Base saturation percentage (BSP) was calculated by dividing the

sum of base-forming cations by CEC (Coyne and Thompson, 2006). Iron and Zinc were determined by the method described in Baruah and Barthakur (1999) using 0.005M diethylene triamine pentaacetic acid extraction.

#### **2.4. Data Analysis**

Soil properties, land-use systems, soil types, and landforms were used as predictors to assess their relationships with soil BD and TN using SPSS 18.0 software (SPSS, 2011). To do so, all soil types, land-use systems and landform types (Table 1) were coded as categorical (dichotomous) variable using 1= presence and 0 = absence, for indicating their presence and absence in each sampling point. For example, in the sampling point one the soil type was Vertisols, land-use system of teff-pulse rotation and a rolling hills landform, in which these variables were coded as 1 whereas the other soil types, land-use and landforms were absented in sampling point one and consequently, they were coded as 0. The same procedure holds true for the other sampling points in the catchment. The continuous value of the soil properties determined from each sampling point was entered into the corresponding independent soil variables. After this, the stepwise multiple regression (SMR) was used to test the strength of the relationships of BD and TN with all the independent variables. The generalized linear model (GLM) was also applied to examine for the relationships of soil BD and TN with the main and interaction terms of the different parameters. These two methods were selected because of their unique qualities such as the inclusion of input variables with good explanatory power in SMR and the ability of GLM to utilize interaction terms (Agyare, 2004). One common advantage of these techniques is the ability to use both categorical and continuous parameters as input variables. 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 non-significant or of low explanatory power due to the inclusion of other parameters (Hair et al., 1998). The GLM differs from the well-known multiple regressions in many respects. For instance, the distribution of the independent or response variable does not have necessary to be continuous for GLM. Meaning, GLM model allows categorical or nominal variables as input variables by recoding them into a number of dichotomous variables (Agyare, 2004). The parameters and their data group used for the SMR and GLM analysis are presented in table 1.

When soil BD and TN were used as the dependent variable in the analysis, the importance of each independent parameter was evaluated based on the size of the model coefficient, significance level, and coefficient of determination ( $R^2$ ) of each statistical model. Besides to this, the effect size measure (Eta) in the case of GLM was used as a measure of model

performance i.e., measures the association between the main or interaction term and the dependent variable. Only data elements that contributed significantly ( $P \leq 0.05$ ) to predicting soil BD or TN were presented in this study. The  $R^2$  change gives the percentage of variance in the dependent variables (BD and TN) explained by the independent variables. For each statistical model, the  $R^2$  and F-statistic is used as a measure of model performance. The stepwise regression analysis was done using the data lists with variable selection method, i.e., entering and removal of parameters at  $P \leq 0.05$  and  $P \geq 0.1$ , respectively, using SPSS 18.0 (SPSS, 2011). For GLM analysis, the number of interactions is limited to 2-way due to the possible increase of multi-collinearity.

### 3. RESULTS AND DISCUSSION

#### 3.1. Estimation of Soil Bulk Density Using SMR Analysis

The result of the stepwise multiple regression (SMR) analysis for soil bulk density (BD) as the dependent variable is shown in table 2. An overall model coefficient of determination ( $R^2$ ) of 0.91 with a significantly high F-statistic of 176.78 was obtained for the relationship of soil BD with the soil properties, land-use system and landforms in the catchment. This indicates that about 91% of the variance in BD can be explained by the independent variables. However, there was no significant ( $P > 0.05$ ) relationship between BD and soil types of nominal data group, and some soil properties (e.g., Ca, Mg, Na) (data not shown). The non-significant variables were not included in the SMR analysis result due to their low relative contributions.

Table 2. Stepwise regression coefficients, standard error, significance level and  $R^2$  change for soil bulk density estimation in the Mai-Negus catchment, northern Ethiopia.

<i>Parameter</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>Significance level</i>	<i>R<sup>2</sup> change</i>
Constant	0.686	0.024	0.00	
Organic carbon	-0.256	0.012	0.00	0.710
Sand	0.211	0.015	0.00	0.051
Silt	-0.135	0.027	0.00	0.033
CEC	-0.022	0.008	0.02	0.006
Pav	-0.031	0.010	0.02	0.004
Reservoir	-0.205	0.033	0.00	0.033
Central-ridge	0.041	0.017	0.01	0.014
Mountainous	0.064	0.016	0.01	0.013
Forest land	-0.159	0.038	0.01	0.016
Plantation protected land	-0.163	0.042	0.001	0.017
Marginal land	0.055	0.019	0.03	0.013
F-statistic (n)	176.78 (117)			
Overall Model $R^2$	0.91			

$R^2$ , coefficient of determination; CEC, cation exchange capacity; n, number of samples included in the analysis.

The most important parameters that influence the estimation of BD were identified as OC, followed by sand content, the reservoir landform, plantation protected land and forest land-use systems. This is because these variables showed higher SMR coefficients than the others. In line to this, previous reports indicated that sand content and OC were among the most important soil property that affected BD in which BD decreases with increasing OC and decreasing sand content (e.g., Bauer and Black, 1992; Wagner et al., 1994). Such findings thus suggested that soil BD can be estimated using soil texture parameters along with OC values. However, Heuscher et al. (2005) reported using stepwise multiple regression that OC was the strongest contributor to BD prediction as compared to other soil properties.

The soil parameters such as CEC and Pav showed lower relationships with soil BD as compared to the others, even though their relationships were influenced significantly (Table 2). The implication is that while estimating BD, the parameters that had larger regression coefficients are preferable to those with lower values. The variables such as OC, silt, CEC, Pav, reservoir, forest land, plantation protected land-use system that had negative regression coefficients indicated an inverse relationship with BD. On the other hand, the variables such as sand content, marginal land-use system, central-ridge and mountainous landforms showed positive coefficients for their relationships with BD. In line to this, other studies reported that positive coefficient indicates that the dependent variable increases as the corresponding independent variables increase and vice versa for negative coefficient (e.g., Rawls, 1983; Federer et al., 1993; Manrique et al., 1993; Neupane et al., 2002).

The highest variance ( $R^2$ ) while estimation soil BD was explained by organic carbon only for 71%, and the lowest was by Pav (0.4%) followed by CEC (0.6%). According to Hamilton (1990), strong and weak values of  $R^2$  are defined between  $\pm (0.64-1.0)$  and  $\pm (0.04-0.25)$ , respectively. Generally, considering the different data source groups (soil properties, landform and land-use systems) as independent variables, the  $R^2$  obtained using the SMR analysis to estimate soil BD is illustrated in figure 4. The  $R^2$  for soil properties as the data source group in estimating BD was explained by the variance of 80%. This was followed by the data source group of the landform for 6% and land-use systems for 5%, in explaining the variability of soil BD. Those values indicated that BD variability in the study catchment can be explained mainly by the soil properties data group. In light of the above model results, other studies prove that stepwise multiple regression (SMR) is more efficient than the full model regression to determine predictive equation for yield and yield components (e.g. Naser and Leilah, 1993; Mohamed, 1999).

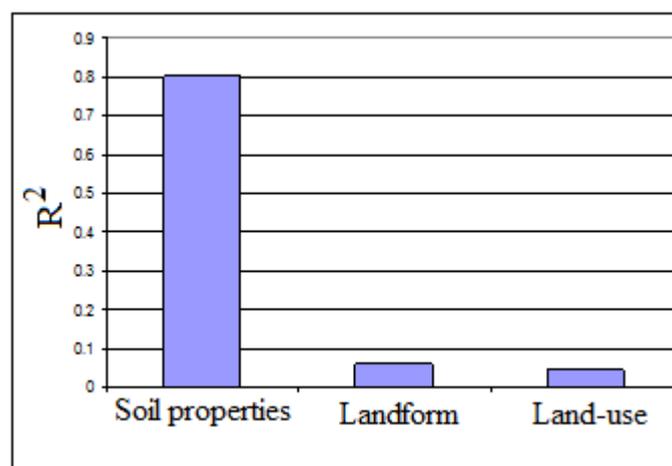


Figure 4. Comparison of coefficient of determination ( $R^2$ ) for soil bulk density (BD) as explained by different data source groups in the Mai-Negus catchment, northern Ethiopia.

Table 3. Stepwise regression coefficients, standard error, significance level and  $R^2$  change for soil total nitrogen estimation in the Mai-Negus catchment, northern Ethiopia.

<i>Parameter</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>Significance level</i>	<i>R<sup>2</sup> change</i>
Constant	0.176	0.021	0.000	
Organic carbon	0.067	0.004	0.000	0.690
Silt	-0.025	0.001	0.010	0.018
Pav	0.042	0.017	0.020	0.011
Bulk density	-0.009	0.009	0.000	0.006
CEC	0.010	0.003	0.004	0.001
Clay	0.036	0.005	0.001	0.009
Reservoir landform	0.027	0.003	0.000	0.045
Central-ridge landform	-0.021	0.011	0.010	0.012
Mountainous landform	-0.009	0.002	0.003	0.001
Forest land	0.013	0.010	0.000	0.039
Marginal land	-0.014	0.013	0.010	0.002
Teff mono-cropping	0.011	0.008	0.010	0.003
Maize mono-cropping	0.013	0.010	0.000	0.009
Plantation protected land	0.046	0.014	0.002	0.047
F-statistic (n)	162.57 (117)			
Overall Model R <sup>2</sup>	0.89			

$R^2$ , coefficient of determination; CEC, cation exchange capacity; n, number of samples included in the analysis.

### 3.2. Estimation of Soil Total Nitrogen Using SMR Analysis

The result of stepwise multiple regression (SMR) analysis between the dependent variable (TN) and independent variables (catchment attributes) is presented in table 3. The parameters in this table were significant predictors of TN for an overall  $R^2$  of 0.89 with a significant F-

statistic of 162.57. However, the coefficient and  $R^2$  of each variables showed that soil OC followed by the plantation protected land-use system and reservoir landform have significant contribution for the regression model when compared with the others (Table 3). The better estimation of TN by OC could be associated with the organic source of nitrogen in the study catchment and this is consistent with the finding reported using linear regression in Rashidi and Seilsepour (2009); and Prévost (2004). The SMR model coefficient and  $R^2$  of the forest land-use system was also significantly predicted TN better than the other variables. The lowest model coefficient (-0.009) was found for the relationships between TN with BD and the mountainous landform.



Figure 5. Comparison of coefficient of determination ( $R^2$ ) for soil total nitrogen (TN) explained by different data source groups in the Mai-Negus catchment, northern Ethiopia.

The total variance explained for TN by the different data source groups is shown in figure 5. This figure indicates that soil properties followed by land-use systems and landform can better predict TN compared to the other landscape attributes. The soil properties (OC, silt, clay, BD, Pav, CEC) could predict 74% of the variation in soil TN, even though OC only accounted for the largest part (about 69%). Such  $R^2$  results are rated as strong relationship according to the rates described by Hamilton (1990). The land-use system data group (plantation protected land, forest land, maize mono-cropping, teff mono-cropping, marginal land-use systems) explained the variability in the prediction of TN for 10%. However, the larger portion (8.6%) of TN variability was explained by both plantation protected and forest land-use systems. The landform data source group (reservoir, central-ridge and mountainous landforms) explained for the variability of TN for about 5.8%. From this value, the reservoir accounted for explaining the variability in TN for about 3.9%. This study indicated that land-

use and landform are not good enough to be useful to predict soil TN as compared to the soil properties data group.

Table 4. GLM regression coefficients, standard error, significance level and effect size measure (Eta) for soil BD estimation in the Mai-Negus catchment, northern Ethiopia.

<i>Parameter</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>Significance level</i>	<i>Eta square</i>
Constant	0.735	0.241	0.001	0.002
OC	-0.314	0.102	0.001	0.114
Silt	-0.155	0.052	0.003	0.003
Sand	0.126	0.065	0.004	0.025
Forest land	-0.094	0.020	0.005	0.023
Plantation protected land	-0.053	0.017	0.002	0.019
Reservoir	-0.202	0.070	0.005	0.033
Central-ridge	0.073	0.013	0.014	0.015
Mountainous	0.051	0.010	0.020	0.011
Chromic Cambisols	-0.025	0.008	0.030	0.006
Reservoir × OC	-0.421	0.135	0.001	0.160
OC × clay	-0.365	0.112	0.001	0.125
OC × silt	-0.226	0.127	0.006	0.136
Sand × clay	0.263	0.105	0.002	0.035
Valley × silt	-0.231	0.096	0.003	0.037
Forest land × OC	-0.400	0.127	0.005	0.112
Protected land × OC				
Marginal land × sand	-0.137	0.001	0.012	0.045
Teff mono-cropping × sand	0.205	0.130	0.003	0.029
F-statistic (n)	187.45 (117)			
Overall Model R <sup>2</sup>	0.92			

*BD, soil bulk density; x, interaction between parameters; OC, organic carbon; n, number of samples included in the analysis; R<sup>2</sup>, coefficient of determination.*

### 3.3. Estimation Soil Bulk Density Using General Linear Model (GLM)

The GLM analysis result for the relationship of soil BD with the independent variables is presented in table 4. The result shows the coefficient, standard error, significance level and measure of size effect (Eta) for the constant or intercept, the main and interaction effects of the independent variables. The Eta gives the measure of the association between the main or interaction term and the dependent variable, in this case soil BD. The GLM analysis resulted in an overall R<sup>2</sup> of 0.93 with F-statistic of 187.45. The main term effect among the parameters in the GLM analysis showed the highest model coefficient and Eta square for OC followed by the reservoir landform. However, the interaction term effect of OC with the other parameters on the coefficients and Eta square was higher when compared to the main terms

effect on BD prediction as this is shown by the GLM analysis (Table 4). In contrast to this study, Agyare (2004) has reported that OC was found to be least important in estimating hydraulic conductivity, even though this directly influences soil pore size and distribution. The least important main term effect in the prediction of soil BD using the GLM analysis was explained by the soil type.

The dominant parameters that influenced soil BD prediction were associated with the interaction terms, explaining about 68% of the variance compared to the other data groups (Fig 6). The remaining variation in BD was mainly explained by the soil properties (15.2%), landforms (8.2%) followed land-use systems (3.2%) and soil type (0.60%) data groups. This result indicates that the interaction terms in the GLM analysis are better in estimating soil BD in the study catchment conditions.

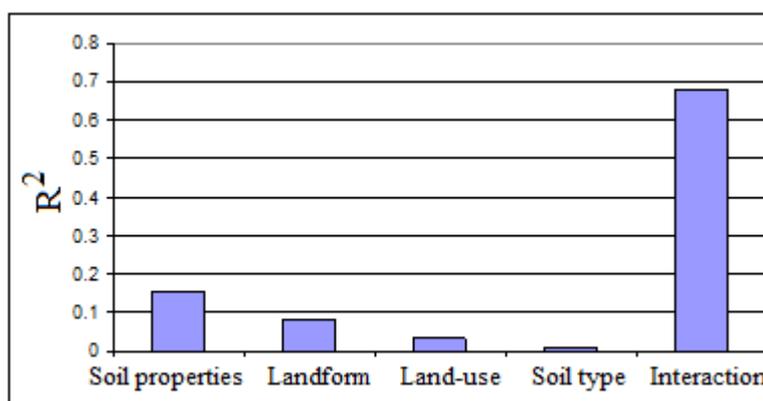


Figure 6. Comparison of measure of size effect (Eta) for soil BD estimation in-terms of different data groups using GLM analysis for Mai-Negus catchment, northern Ethiopia.

### 3.4. Estimation of Soil Total Nitrogen Based on GLM

The coefficient, standard error, significance level and measure of size effect (Eta) for the constant or intercept, the main and interaction effects of the independent variables on the dependent variable of TN using the GLM analysis is presented in table 5. The overall model R<sup>2</sup> of the GLM analysis was 0.94 with F-statistic of 193.05 for TN. Among the main terms, OC (11.7%) showed significantly higher model coefficient and Eta square followed by the plantation protected land-use system (1.6%) in explaining TN variability. The minimum coefficient and Eta square of the main term effect was obtained due to soil pH even though this was significantly influenced soil TN variation. Of the interaction terms, the OC × plantation protected land system followed by OC × forest land system showed significantly higher model coefficient and Eta square than the others. The OC × reservoir interaction term also contributed better for TN estimation. This indicates that the soil TN variability can be

critically influenced by the OC, land-use system and landform in the catchment. This result is consistent with the report in Aruleba and Ajayi (2011); and Rashidi and Seilsepour (2009).

Table 5. GLM regression coefficients, standard error, significance level and effect size measure (Eta) for soil total nitrogen estimation in the Mai-Negus catchment, northern Ethiopia.

<i>Parameter</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>Significance level</i>	<i>Eta square<sup>a</sup></i>
Constant	0.159	0.116	0.008	0.001
Organic carbon	0.456	0.103	0.001	0.117
Sand	-0.113	0.083	0.007	0.010
Silt	0.235	0.091	0.008	0.013
Clay	0.204	0.108	0.010	0.015
Bulk density	-0.200	0.132	0.006	0.002
CEC	0.138	0.034	0.026	0.001
pH	0.036	0.010	0.009	0.001
Pav	0.143	0.046	0.00	0.002
Forest land	0.107	0.049	0.015	0.011
Teff-pulse rotation land	0.083	0.007	0.003	0.001
Plantation protected land	0.110	0.020	0.002	0.016
Marginal land	-0.062	0.010	0.020	0.001
Reservoir	0.146	0.057	0.021	0.012
Escarpment	0.105	0.031	0.013	0.005
Central-ridge	-0.095	0.024	0.010	0.002
Mountainous	-0.053	0.017	0.030	0.002
OC × BD	-0.586	0.223	0.004	0.126
BD × clay	-0.135	0.047	0.009	0.018
BD × silt	-0.413	0.210	0.006	0.017
Sand × silt	-0.031	0.005	0.008	0.022
BD × sand	0.205	0.164	0.005	0.015
OC × Plantation protected land	0.521	0.187	0.001	0.137
Maize mono-cropping × silt	0.046	0.006	0.003	0.018
OC × teff-pulse rotation	0.104	0.053	0.001	0.121
BD × reservoir	0.472	0.324	0.005	0.016
OC × clay	0.372	0.145	0.001	0.120
OC × Forest land	0.503	0.203	0.001	0.131
F-statistic (n)	193.05 (117)			
Overall Model R <sup>2</sup>	0.94			

*x*, interaction between parameters; OC, organic carbon; BD, bulk density; CEC, cation exchange capacity, N, number of samples included in the analysis; R<sup>2</sup>, coefficient of determination.

<sup>a</sup>Eta square is analogous to  $r^2$  in correlation and regression. Eta-square provides the most accurate representation of continuous as well as categorical data, and is well known and easy to interpret. It is a descriptive statistic that quantifies the strength of the relationship between group membership and the variable you measured.

The remaining main and interaction terms also significantly explained the estimation of TN in the study catchment (Table 5). The interaction terms accounted for overall Eta square of 0.74, which means that this is explained for 74% of the variation in soil TN. This is followed by the soil properties, land-use and landform in descending order as it is shown in figure 7.

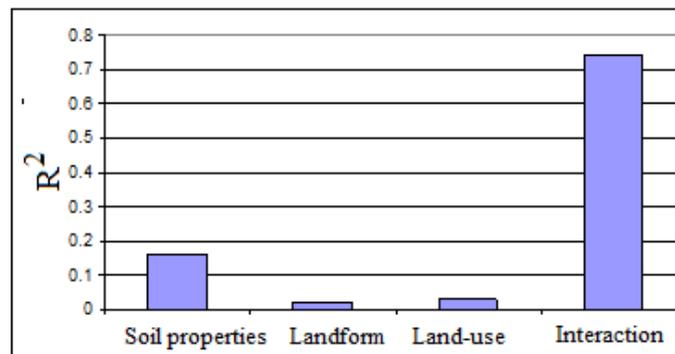


Figure 7. Comparison of measure of size effect (Eta) for soil TN estimation in-terms of the different data groups using GLM analysis in the Mai-Negus catchment, northern Ethiopia.

### 3.5. Synthesis of SMR and GLM Analysis Results

Currently, literature shows that there is no comprehensive model for estimating soil parameters which are not easily available, using basic soil properties obtained in soil survey and catchment attributes collected directly from a field (Seybold et al., 2009). To contribute towards this information gap, this study examined the variables that best explain for variation in soil BD and TN using different analysis methods. In this study, the overall  $R^2$  found from SMR and GLM analysis employed to predict soil BD and TN based on key landscape attributes is greater than 0.60. This is in agreement with the report by Seybold et al. (2009) which stated that when there are no measured data, regression equations developed with  $R^2 > 0.60$  are useful and can be used in estimating soil properties. Thus, the relationships of soil BD and TN with catchment attributes developed in this study can be used to estimate BD and TN in similar conditions. In general, the overall model  $R^2$  (0.93-0.94) for GLM analysis are higher than that of SMR (0.89-0.91) for both dependent variables (Tables 2-5). This could be attributed to the inclusion of interaction terms in the GLM, while SMR does not explicitly consider that interaction effect. Similarly, Agyare (2004) reported that the interaction terms of soil properties with catchment attributes are critical for estimating hydraulic conductivity.

In addition, the soil type data group had no significant relationship with BD and TN in the SMR analysis whereas the reverse is observed in the GLM analysis. The overall  $R^2$  due to the independent parameters on TN (0.94) for the GLM analysis is greater than that of soil BD (0.92). However, the  $R^2$  is higher in the SMR analysis for soil BD (0.91) than TN (0.89). The

improvement in  $R^2$  may be associated with the inclusion of parameters more suitable for the dependent variable of soil BD than TN in the SMR analysis.

The Eta, which is a measure of the association between the two dependent and the different independent variables in reflecting the interaction terms is considerably reduced from 74% for soil TN to 68% for soil BD variance (Figs 6 &7). This holds true for the Eta of the soil properties data group with variance of 16% for TN and 15% for BD. Of the independent variables, OC contributed better in predicting the variability of both BD and TN. The  $R^2$  for the relationship of BD and OC using SMR analysis was 0.71, whereas it was 0.69 with TN. However, the variance contributed by OC as main term effect for the variability of BD and TN in the GLM analysis were 0.114 and 0.118, respectively, which are lower than that of the SMR. This shows that linear regression model may be better suited for predicting soil TN from soil OC as suggested by Rashidi and Seilsepour (2009). Despite this fact, OC contributed higher in both methods as compared to the other independent variables used. The model coefficients found from both analysis methods for both dependent variables were higher for OC as compared to the other variables. This study indicates that a unit increase in organic matter can be caused a relatively larger decrease in soil bulk density but larger increase in TN, which is consistent with the findings reported in Federer et al. (1993); and Akgül and Özdemir (1996). Thus, management and land-use practices that improve OC should be considered in determining representative fields as OC is useful in estimating soil BD and TN for soil survey when there are no measured data available.

#### 4. CONCLUSION

In this study, the potentials of stepwise multiple regressions (SMR) and generalized linear models (GLM) were used to predict the variation of soil bulk density (BD) and total nitrogen (TN) based on selected environmental data group (soil properties, land-use, landform, soil type) at a catchment-scale in northern Ethiopia. The study also attempted to identify the most useful environmental variables that predict the BD and TN. The results of SMR and GLM analysis showed strong relationships ( $R^2 > 0.89$ ) between the two soil properties (soil BD and TN) and catchment and soil attributes. Both methods indicated that soil properties are best suited to predict BD and TN. Among the different soil properties, organic carbon (OC) accounted for the largest share of the variation in BD and TN. The second landscape data group that best predicted BD was land-use type followed by landform for BD. However, the Eta as a measure of association in the GLM analysis showed higher model coefficient and Eta square for the interaction terms of OC  $\times$  plantation protected land system, OC  $\times$  forest land

system, OC  $\times$  reservoir with BD and TN. However, soil TN (68%) is better explained by the interaction terms as compared to BD (74%), even though both are significantly related to the independent variables. Generally, a higher overall  $R^2$  of GLM for both dependent variables (BD and TN) as compared to that of SMR indicated that inclusion of the interaction terms in the GLM analysis improved the variance values. However, the  $R^2$  of OC in the SMR analysis for BD estimation was 0.71 whereas for TN it was 0.69. These values are higher than the Eta square of OC in the GLM analysis, indicating that SRM is preferable for estimating BD and TN variations using the main term. This study suggests that soil BD and TN should be estimated based on the soil properties such as organic carbon and other catchment attributes using appropriate statistical techniques such as the GLM which is more suitable to consider the effects of interaction terms in the study catchment conditions in order to get data for most applications.

## 5. ACKNOWLEDGEMENTS

The author gratefully acknowledge the financial support by DAAD/ GTZ (Germany) through the Center for Development Research (ZEF), University of Bonn (Germany), and the support of Aksum University (Ethiopia) during the field work. The author deeply appreciates the assistance offered by the local farmers and extension agents during the field study. The author is also grateful to the anonymous reviewers for their comments, suggestions and corrections of this manuscript.

## 6. REFERENCE

- Agoumé, V & Birang, A.M. 2009. Impact of land-use systems on some physical and chemical soil properties of an Oxisol in the humid forest zone of Southern Cameroon. *Tropicultura*, **27**: 15-20.
- Agyare, W.A 2004. Soil characterization and modeling of spatial distribution of saturated hydraulic conductivity at two sites in the Volta Basin of Ghana. Pub. PhD Thesis, Center for Development Research, University of Bonn, Germany.
- Akgül, M & Özdemir, N. 1996. Regression models for predicting bulk density from measured soil properties. *Tropical Journal of Agriculture and Forest*, **20**: 407-413.
- Aksum Meteorological Station, 2009. Rainfall and temperature data for Aksum area. Aksum, Ethiopia.
- Anderson, J.M & Ingram, J.S.I. 1993. Tropical soil biology and fertility. A Handbook of methods. CAB International, Wallingford.

- Arshad, M.A., Lowery, B & Grossman, B. 1996. Physical tests for monitoring soil quality. In: W. Doran, & A.J. Jones, (eds.), *Methods for Assessing Soil Quality*, Spec. Publ., vol. 49, *Soil Science Society of America*, Madison, WI, 123-142pp.
- Aruleba, J.O & Ajayi, A.S. 2011. Use of models in assessing the impact of cropping system, land types and suitability on land degradation in Southwestern Nigeria. *African Journal of Agricultural Research*, **6**: 1090-1096.
- Baruah, T.C & Barthakur, H.P. 1999. A text book of soil analysis. Vikas Publishing House Pvt. Ltd., New Delhi.
- Bauer, A & Black, A.L. 1992. Organic carbon effects on available water capacity of three soil textural groups. *Soil Science Society of America Journal*, **56**: 248-254.
- Birang, A.M., Hauser, S., Brussaard, L & Norgrove, L. 2003. Earthworm surface-casting activity on slash-and-burn cropped land and in undisturbed *Chromolaena odorata* and young forest fallow in south Cameroon. *Pedobiologia*, **47**: 811-818.
- Black, C.A., Evans, D.D., White, J.L., Ensminger, L.E & Clark, F.E. (eds.) 1965. *Methods of Soil Analysis, Part 1. Physical and Mineralogical Properties, Including Statistics of Measurement and Sampling. America Society of Agronomy*, Madison, WI.
- Blake, G.R & Hartge, K.H. 1986. Bulk Density. In: A. Klute (ed.), *Methods of Soil Analysis, Part 1, 2nd edition, Agronomy Monograph 9, America Society of Agronomy*, Madison, WI, pp.363-375.
- Bremner, J.M & Mulvaney, C.S. 1982. Nitrogen total. In: A.L. Page, (ed.), *Method of Soil Analysis, Part 2. Chemical and Microbiological Properties, Agronomy Monograph 9, America Society of Agronomy*, Madison, WI, pp.595-624.
- Chikowo, R., Corbeels, M., Mapfumo, P., Tittonell, P., Vanlauwe, B & Giller, K.E. 2010. Nitrogen and phosphorus capture and recovery efficiencies, and crop responses to a range of soil fertility management strategies in sub-Saharan Africa. *Nutrient Cycling in Agroecosystems*, **88**: 59-77.
- Coyne, M.S & Thompson, J.A. 2006. *Math for Soil Scientists*. Thomson Delmar Learning, Clifton Park, New York.
- Ethiopian Mapping Agency (EMA). 1997. Ethiopia 1:50000 topographic maps. Aksum Sheet.
- FAO-UNESCO. 1974. *Soil map of the world. volume 1, Legend*, FAO, Rome.
- Fassil Kebede & Yamoah, C. 2009. Soil fertility status and Nufass fertilizer recommendation of Typic Haplusterts in the Northern Highlands of Ethiopia. *World Applied Science Journal*, **6**: 1473-1480.

- Federer, C.A., D.E. Turcotte & Smith, C.T. 1993. The organic fraction-Bulk-density relationship and the expression of nutrient content in forest soils. *Canadian Journal of Forest Research*, **23**: 1026-1032.
- Gebreyesus Brhane, Lulseged Tamene & Vlek, P.L.G. 2011. Evaluation of soil quality identified by local farmers in Mai-Negus catchment, northern Ethiopia. *Geoderma*, **163**: 209-218.
- Gee, G.W & Bauder, J.W. 1986. Particle-size analysis. In: A. Klute, (ed.), *Methods of Soil Analysis, Part 1, 2nd ed. America Society of Agronomy, Soil Science Society of America*, Madison, WI, pp.383-411.
- Hair, J.F., Anderson, R.E., Tatham, R.L & Black, W.C. 1998. *Multivariate data analysis*. Prentice Hall, Upper Saddle River, New Jersey.
- Hamilton, L.C. 1990. *Modern data analysis. A first course in applied statistics*, 684 pp.
- Heuscher, S.A., Brandt, C.C & Jardine, P.M. 2005. Using soil physical and chemical properties to estimate bulk density. *Soil Science Soc., of America Journal*, **69**: 51-56.
- Jackson, M.L. 1964. *Soil chemical analysis*. Englewood Cliffs, Prentice Hall, New York.
- Kamara, C.S & Haque, I. 1988. Characteristics of Vertisols at ILCA research and outreach sites in Ethiopia. Revised version. Plant Science Division Working Document B5. ILCA (International Livestock Centre for Africa), Addis Ababa, Ethiopia.
- Lal, R. 1996. Deforestation and land-use effects on soil degradation and rehabilitation in Western Nigeria. I. Soil physical and hydrological properties. *Land Degradation & Development*, **7**: 19-45.
- Manrique, L.A & Jones, C.A. 1991. Bulk-density of soils in relation to soil physical and chemical properties. *Soil Science Society of America Journal*, **55**: 476-481.
- Mohamed, N.A. 1999. Some statistical procedures for evaluation of the relative contribution for yield components in wheat. *Zagazig J. of Agricultural Research*, **26**: 281-290.
- Naser, S.M & Leilah, A.A. 1993. Integrated analysis of the relative contribution for some variables in sugar beet using some statistical techniques. *Bulletin no.* **44**: 253-266.
- Netherlands Engineering Consultants (NEDECO). 1998. Tekeze river basin integrated master plan development project. Natural Resource, Soils and Terrain, V. xiv, Addis Ababa.
- Neupane, R.P., Sharma, K.R & Thapa, G.B. 2002. Adoption of agroforestry in the hills of Nepal: A logistic regression analysis. *Agricultural Systems*, **72**: 177-96.
- Olsen, S.R & Sommers, L.E. 1982. Phosphorous. In: A.L. Page, (ed.), *Method of Soil Analysis, Part 2. Chemical and Microbiological Properties*. Agron. Monogr. 9, *America Society of Agronomy*, Wisconsin, WI, pp.403-430.

- Périé, C & Ouimet, R. 2008. Organic carbon, organic matter and bulk density relationships in boreal forest soils. *Canadian Journal of Soil Science*, **88**: 315-325.
- Prévost, M. 2004. Predicting soil properties from organic matter content following mechanical site preparation of forest soils. *Soil Science Society of America Journal*, **68**: 943-949.
- Rashidi, M & Seilsepour, M. 2009. Modeling of soil total nitrogen based on soil organic carbon. *Journal of Agricultural and Biological Science*, **4**: 1-5.
- Rawls, W.J. 1983. Estimating soil bulk-density from particle-size analysis and organic matter content. *Soil Science*, **135**: 123-125
- Reinsch, T.G & Grossman, R.B. 1995. A method to predict bulk density of tilled Ap horizons. *Soil and Tillage Research*, **34**: 95-104.
- Rhoades, J.D. 1982a. Soluble salts. In: A.L. Page, (ed.), *Methods of Soil Analysis V.2, America Soc., of Agronomy, Soil Science Soc., of America*, Madison, WI, pp.167-179.
- Rhoades, J.D. 1982b. Cation exchange capacity. In: A.L. Page, R.H. Miller & D.R. Keeney, (eds.), *Methods of Soil Analysis, Agronomy Monograph 9, America Society of Agronomy*, Madison, WI, **9**: pp.149-157.
- Schipper, L.A & Sparling, G.P. 2000. Performance of soil condition indicators across taxonomic groups and land uses. *Soil Science Soc., of America Journal*, **64**: 300-311.
- Seybold, C.A., Finnell, P.R & Elrashidj, M.A. 2009. Estimating total acidity from soil properties using linear models. *Soil Science*, **174**: 88-93.
- Shepherd, G., Bureh, R.J & Gregory, P.J. 2000. Land use affects the distribution of soil inorganic nitrogen in smallholder production systems in Kenya. *Biology and Fertility of Soils*, **31**: 348-355.
- SPSS. 2011. Statistical package for social sciences. Release 18.0, SPSS Inc.
- Thomas, G.W. 1996. Soil pH and soil acidity. In: D.L. Sparks (ed.), *Methods of Soil Analysis, Part 3, Chemical Methods, Soil Science Society of America*, Madison, WI, pp.475-490.
- Wagner, L.E., Ambe, N.M & Ding, D. 1994. Estimating a Proctor density curve from intrinsic soil properties. *American Society of Agricultural and Biological Engineers*, **37**: 1121-1125.
- Wondzell, S.M., Cunningham, G.L & Bachelet, D. 1996. Relationships between landforms, geomorphic processes, and plant communities on a watershed in the northern Chihuahuan Desert. *Landscape Ecology*, **11**: 351-362.