APPLICATION OF REGRESSION TREE IN MODELLING AND MAPPING OF CATION EXCHANGE CAPACITY OF SOILS IN AKWA IBOM STATE, NIGERIA

***Akpan[1], U. S.; Uwah[2] I.D. and Nkannga[1], N.A.**

[1]Department of Soil Science and Land Resources Management, University of Uyo , Akwa Ibom State, Nigeria

[2]Department of Agricultural Technology, Akwa Ibom State College of Technology, Nung Ukim Ikono, Akwa Ibom State, Nigeria

*: E-mail: **udyakpa2k2@yahoo.com**. 08023783311

ABSTRACT

Regression tree was used in modelling and mapping cation exchange capacity of soils in Akwa Ibom State, Nigeria.The aim was to provide an alternative techinque of estimating ECEC from more readily available soil data and map the distribution for site-specific soil management. The study area (Akwa Ibom State) was grouped into four major mapping units based on parent materials, namely: coastal plain sand, sandstone, shale and beach ridge sand. Each parent material (major mapping unit) was subdivided into four commonly practice landuse types/soil management systems namely homestead or compound farmland, oil palm plantation, secondary forest of 3 years and above and cultivated farmland. In each landuse type/soil management system, mini- soil profile pit was dug to a depth of 100cm at representative location. Soil samples were collected from designated depths of 0-20, 20-60 and 60-100 cm. A total of 144 samples were generated for laboratry analysis. The study revealed that ECEC can be predicted using soil organic carbon, clay, silt and soil pH in the study area. The results of independent variable importance to the model showed that organic carbon was the most significant predictor of ECEC with 22.1 % contribution, followed by clay with 17.3 %, followed by silt with 8.5% while soil pH was the least predictor of ECEC with 0.8% contribution in the study area. Based on the model, organic carbon content predicted ECEC of 32.4 cmol/kg in sandstone soil while organic carbon in combination with clay predicted ECEC of 40.24 cmol/kg in soils developed from shale parent material. In coastal plain sand soils, organic carbon in combination with clay and silt predicted ECEC of 24.3 cmol/kg. In beach ridge sand soils, organic carbon in combination with clay and silt predicted ECEC of 17.8 cmol/kg. The model showed that organic carbon content was the only significant predictor of ECEC in sandstone soils while organic carbon in combination with clay made significant prediction of ECEC in shale parent material. In coastal plain sand and beach ridge sand soils, Organic carbon in combination with clay and silt made significant prediction of ECEC. In the application of the model, independent variables included in the final model and measured in the same unit should be used.

Keywords: modelling, regression tree, CEC, soil of Akwa Ibom State.

Introduction

Cation Exchange Capacity (CEC) is one of the most important soil properties that is required in soil databases. It is a good indicator of soil fertility, crop growth and pollutant transport and is used as an input in soil and environmental models (Manrique *et al*., 1991; Keller *et al*., 2001). Cation Exchange Capacity is the total exchangeable cations that a soil can hold at a specified pH. It measures the potential capacity of soil to hold and exchange cations at a specified pH, measures soil's capability to store and filter chemicals, buffer soil chemical properties against changes, influenced structural stability, nutrient availability, soil pH and the soil's reaction to fertilisers and other ameliorants. In acid soil, (pH <5.5), the sum of Ca^{2+}, Mg^{2+}, K^{+} and Na^{+} is often less than the exchange capacity of the soil, the remainder is filled by Al^{3+}, H^{+} and Mn (Rengasamy and Churchman, 1999).

Soil components known to contribute to CEC are clay and organic matter, and to a lesser extent, silt (Manrique *et al*., 1991). Both of these are negatively charged particles in the soil that attract the positively charged (cations) nutrients. In clays, the negative charge is due to an excess of oxygen atoms in the crystal clay structure (permanent charge). The types of clay such as the swelling clays (smectite and montmorillonite) have a higher CEC than the non- swelling type clays (kaolinite). Soil organic matter on the other hand, does contribute to the CEC but in a complex way. The contribution depends on the type of organic matter and the inherent soil chemical environment (Charman and Murphy, 2007). According to Oades *et al*. (1989), carboxyl groups of organic matter are the major source of the negative charge that contributes to CEC. The humification of soil organic matter gives rise to a wide range of compounds such as humin, humic acids (HAs), fulvic acids (FAs) and derivative of amino acid and phenolic acid which have the general form of R-CH_3-COOH and R-CH_2-COOH (Stevenson, 1982). As the pH of the soil increases from 5.0 to 7.0, the hydroxyl groups become disassociated or ionized to form negatively charged (R*COO) sites which are available for cation exchange. Carboxyl group in humus ionise or disassociate mainly in the acid part of the pH scale while phenolic hydroxyl groups disassociate at pH above 6.0. When the pH falls below 5.0, many of the compounds of organic matter are not dissociated and so do not displays a negative charge (Stevenson, 1982). Partitt *et al*. (1995) found that most of the CEC attributed to topsoil organic matter was from carboxyl functional groups and a critical limite of 2% of soil organic carbon was necessary for soil organic matter to have an appreciable effect on CEC.

Although CEC can be measured directly, its measurement is difficult, time consuming and expensive. Pedotransfer Functions (PTFs) provide an alternative mean of estimating CEC from more readily available soil data. The term pedotransfer function was coined by Bouma (1989) as translating data we have in to what we need. In recent years, several researchers tried to estimate CEC from basic physical and chemical soil properties (Bell and McBratney *et al.*, 2002). In many of the models, CEC is assumed to be a linear function of soil organic matter and clay content and therefore used linear regression in the modelling. Drake and Motto (1982) reported that greater than 50% of the variation in CEC was explained by the variation in clay and organic C content in several New Jersey soils. Krogh *et al.* (2000) reported 90% variation in CEC due to variation in silt, clay, organic carbon and soil pH. For valid application of linear regression in modelling, certain assumptions must be satisfied. Some of the assumption includes, the relationship between dependent variable and independent variables must be linear; independent variables must be linearly independent of each other otherwise multicollinearity is present, the distribution must be normally distributed (parametric) etc. Some relationships between dependent variable and independent variables may not be linear and the distribution may not be normally distributed (non-parametric). To overcome these limitations, some techinques such as classification and regression tree can be resorted to without undergoing data transformation and principal component analysis.

Classification and regresson tree (CART) is one of the popular machine learning and data mining algorithms (Huang *et al*., 2010). Machine learning and data mining is a field of computer science that gives computers the ability to learn based on the inherent characteristics of data, without being explicitly programmed, uncovering patterns and structures in (large) data sets and deriving predictive relationships (Diplaris *et al*., 2006). Classification and regression tree is a systematic approach used in building classification and regression models from an input data set. The model or tree like- structure consists of nodes and leaves. The nodes are spltting points while the leaves are the terminal nodes. The objective of building a classification and regression tree is to find a model or tree-like structure which recursively partitions the learning data set into predefined class (classification tree) or mean value (regression tree). It is most suited for predicting or describing data set with binary or nominal catorgories. The main characteristic of classification tree is that the feature space (i.e., the space spanned by all predictor variables) is recursively partitioned into a set of rectangular areas (Brieman *et al*., 1984). Recursive partitioning divides up the 'p' dimensional space of the 'x' variables into non-overlapping rectangles. This division is accomplished recursively. The partition is created such that observations with similar response values are grouped into rectangular areas. Each rectangle is as homogenous or as 'pure' as possible (Brieman *et al*., 1984). Pure here refers to area containing points that belong to just one class which can be displayed either as a tree, or as a rectangular partition of the feature space. The prediction at the leaf is the mean value of data points or the model class. In selecting the splitting variable and cutpoint, classification tree follows the approach of impurity reduction. Each split in the tree-building process results in daughter nodes that are more pure than the parent node in the sense that groups of subjects with a majority for either response class are isolated. The impurity reduction achieved by a split is measured by the difference between the impurity in the parent node and the average impurity in the two daughter nodes. The most common methods of measuring impurity in classification and regression tree are the Shannon entropy and the Gini index. Gini Index and Shannon entropy are used to quantify the impurity in each node. These entropy measures have in common that they reach their minimum for perfectly pure nodes with the relative frequency of one response class being zero and their maximum for an equal mixture with the same relative frequencies for both response classes.

According to Wagner *et al*. (2001), models developed for one region may not give adequate estimates for a different region. When soils are grouped by similarities in origin or properties, accuracy of predictive models are improved (Pachepsky and Rawls, 1999). In Akwa Ibom State, very limited work has been carried out to estimate or model soil CEC from easy available soil physical and chemical properties using regression tree. Therefore, this study was conducted to model and mapped cation exchange capacity of soils of Akwa Ibom State, for site-specific soil management using regression tree.

MATERIALS AND METHODS

The study area

The study was conducted in Akwa Ibom State, located in south eastern Nigeria. It lies between latitudes $4^0$30' and $5^0$30'N and longitudes 7^0 30' and $8^0$20'E, and underlain mainly by coastal plain sands, beach ridge sands, sandstone / shale and alluvial deposits parent materials. The climate is humid tropical, annual rainfall ranges from more than 3000 mm along the coast to about 2250 mm at the extreme north, with 1 – 3 dry months in the year. Mean annual temperature varies between 26 and 28^0C, while relative humidity varies between 75 – 80 %. The original natural vegetation which comprised lowland rainforest, mangrove forest and coastal vegetation, has given way to a mosaic farmland, riparian forest and oil palm forest (Petters *et al.*, 1989).

Field work

The study area was grouped into four major mapping units based on the parent materials, namely: coastal plain sand, sandstone, shale and beach ridge sand. Each parent material (major mapping unit) was subdivided into four commonly practice landuse types/soil management sysytems namely homestead or compound farmland, oil palm plantation, secondary forest of 3 years and above and cultivated farmland. In each landuse type/soil management system, mini- soil profile pit was dug to a depth of 100cm at representative location. Soil samples were collected from designated depths of 0-20, 20-60 and 60-100 cm. A total of 144 samples (3 replications x 4 landuse types x 4 parent materials x 3 depths =1 44 samples) were generated for laboratry analysis.

Laboratory analysis

The following analyses were carried out using appropriate standard procedures: Particle size analysis was carried out using the bouyoucos hydrometer method as described by Udo *et al.* (2009). Soil pH was determined in water using a 1:2.5 soil to water suspension and the soil pH was read using a glass electrode. Organic carbon was determined by the dichromate wet-oxidation method as described by Nelson and Sommers (1996). The value was multiplied by 1.732 to obtain organic matter content. Exchangeable bases: Ca, Mg, Na, K. were extracted using normal ammonium acetate (Thomas, 1986). The exchangeable K and Na were determined by flame photometer while Ca and Mg were determined using atomic absorption spectrometer. Effective cation exchange capacity (ECEC) was determined by summing up exchangeable cations and exchangeable acidity.

RESULTS AND DISCUSSION

1. Soil properties of the study area

The mean, minimum, maximum and variance of soil properties of the study area are presented in Table 1. In soils developed from beach ridge sand parent material, the mean sand fraction was 84.85 %, silt fraction was 7.17 % while clay fraction was 11.03 %. In soils developed from coastal plain sand parent material, mean sand fraction was 75.45 %, silt was 6.17 % while clay fraction was 16.42 %. In sandstone soils, sand fraction was 75.78 %, silt was 4.33 % while clay fraction was 19.67 %. In soils developed from alluvium parent material, mean sand fraction was 51.50 %, silt was 14.67 % while clay fraction was 33.83 %. Based on the USDA textural classes, the soil texture of beach ridge sand ranged from sand in the surface soil to sandy clay loam in the subsurface soil. In soils developed from coastal plain sand and sandstone parent materials, soil texture ranged from sand in the surface to sandy clay in the subsoil. In soils developed from alluvium, soil texture ranged from loamy sand in the surface to clay in the subsurface soil. The variation in soil texture among the parent materials could be attributed to influence of parent material (coastal plain sand, beach ridge sand, sandstone /shale) (Soil Survey Staff, 2006).

The mean soil pH in water of beach ridge sand soils was 5.3 , coastal plain sand soils was 5.5, sandstone soils was 5.2 while alluvium soils was 5.7. The mean hydrogen ion concentration (pH) indicated that beach ridge sand soils, coastal plain sand and sandstone soils were strongly acid while allvium was moderately acid. The general acidity in the study area could be attributed to high rainfall in the area and as well as high agricultural activities. In humid environment, soil pH decreases over time in a process called acidification due to leaching caused by high amount of rainfall. Also, the coarse texture soil in the soil surface with low buffering capacity could be responsible for the soil acidity. Soils with high clay and organic matter content are more able to resist a drop or rise in pH (have greater buffering capacity) (Whitebread *et al*., 1998).

The mean soil organic carbon of beach ridge sand soils was 3.7 %, coastal plain sand soils was 4.5 %, sandstone soils was 0.7 % while alluvium soil was 2.0 %. The values indicated that organic carbon was very high in all the parent materials in the study area execpt in sandstone soils (Enwezor *et al*., 1989). The low organic carbon content in sandstone soils than others could be attributed to low organic matter inputs coupled by reduced physical protection of SOC as a result of tillage and increased oxidation of soil organic matter (John *et al*., 2005).

The mean available P of beach ridge sand soils was 12.0 mg/kg, coastal plain sand soils was 24.1 mg/kg, sandstone soils was 10.3 mg/kg while alluvium soil was 15.0 mg/kg. The mean available P was high in soils developed from coastal plain sand, and moderate in all other parent materials. The high P content in coastal plain sand soils could be attributed to high organic matter content in the soil (John *et al*., 2005).

The mean exchangeable Ca of beach ridge sand soils was 3.7 cmol/kg, coastal plain sand soils was 3.6 cmol/kg, sandstone soils was 4.3 cmol/kg while alluvium soils was 3.3 cmol/kg. The mean exchangeable Mg of beach ridge sand soils was 0.8 cmol/kg, coastal plain sand soils was 1.6 cmol/kg, sandstone soils was 2.0 cmol/kg while alluvium soils was 1.4 cmol/kg.. The mean exchangeable Na of beach ridge sand soils was 1.0 cmol/kg, coastal plain sand soils was 0.06 cmol/kg, sandstone soils was 0.06 cmol/kg while alluvium soils was 0.1 cmol/kg. The mean exchangeable K of beach ridge sand soils was 0.14 cmol/kg, coastal plain sand soils was 0.19 cmol/kg, sandstone soils was 0.06 cmol/kg while alluvium soils was 0.17 cmol/kg. Mean exchangeable Ca and K were low in the study area. Mean exchangeable Mg was low in beach ridge sand soils, moderate in coastal plain sand, sandstone and alluvium soils. Mean exchangeable Na was high in beach ridge sand soils and low in all other parent materials in the study area. The variation is due to difference in parent materials and soil organic matter (John *et al*., 2005).

The mean cation exchange capacity (CEC) of beach ridge sand soils was 6.9 cmol/kg, coastal plain sand soils was 9.9 cmol/kg, sandstone soils was 11.4 cmol/kg while alluvium soils was 8.9 cmol/kg. The CEC of the study area was low. The low CEC of the study area could be attributed to the type and quantity of clay and low organic matter content. Soils low in clay (< 15 %) are much more dependent on soil organic matter to provide CEC and biotic processes dominate, whereas soils high in clay (>35 %), biotic processes are minimal (Oades, 1993).

Table 1: Range and mean of soil properties in the study area

Soil property		**Parent material**			
		Beach ridge sand	**Coastal plain sand**	**Sandstone**	**Alluvium**
Sand (%)	Mean	84.85	75.45	75.78	51.50
	Mini	69.20	18.00	59.00	27.00
	Max	95.20	90.60	90.00	75.00
	Variance	67.1	194.5	124.2	547.1
Silt (%)	Mean	7.17	6.17	4.33	14.67
	Mini	1.40	2.00	1.00	12.00
	Max	22.10	14.00	8.00	20.00
	Variance	28.6	16.9	6.3	11.9
Clay (%)	Mean	11.03	16.42	19.67	33.83
	Mini	0.80	3.00	6.00	12.00
	Max	41.40	39.00	38.00	60.00
	Variance	90.5	125.6	141.7	451.8
pH(water)	Mean	5.3	5.5	5.2	5.7
	Mini	4.0	4.2	4.6	5.1
	Max	6.9	7.5	5.9	6.2
	Variance	0.5	0.7	0.2	0.2
OC (%)	Mean	3.7	4.5	0.7	2.0
	Mini	0.2	0.0	0.2	1.3
	Max	15.9	19.2	1.4	2.4
	Variance	29.6	37.4	0.2	0.2
AvP (mg/kg)	Mean	12.0	24.1	10.3	15.0
	Mini	1.2	4.7	4.1	8.1
	Max	56.0	71.9	18.4	34.0
	Variance	220.1	375.3	33.1	92.1
Ca (cmol/kg)	Mean	3.7	3.6	3.6	3.3
	Mini	0.2	0.9	0.9	1.6
	Max	14.4	7.9	7.9	5.1
	Variance	18.9	3.2	3.2	2.0

Mg (cmol/kg)	Mean	0.8	1.6	2.0	1.4
	Mini	0.3	0.8	1.2	1.2
	Max	2.4	2.8	3.1	1.8
	Variance	0.5	0.2	0.4	0.006
Na (cmol/kg)	Mean	1.0	0.06	0.06	0.10
	Mini	0.04	0.03	0.04	0.04
	Max	3.5	0.2	0.1	0.21
	Variance	0.7	0.001	0.001	0.005
K (cmol/kg)	Mean	0.14	0.19	0.06	0.2
	Mini	0.03	0.04	0.03	0.008
	Max	1.8	0.6	0.11	0.3
	Variance	0.1	0.03	0.001	0.007
CEC (cmol/kg)	Mean	6.9	9.9	11.4	8.9
	Mini	1.6	5.2	9.6	5.0
	Max	17.4	15.5	17.3	14.4
	Variance	24.6	7.3	5.4	16.4

2: Method of growing tree and Specifications

The summary of the tree growing method is presented in Table 2: The method adopted for this analysis was classification and regression tree (CRT). The independent variables used in the study were clay, organic carbon, silt and soil pH. Specifications were five maximum tree depth, thirty minimum cases in parent node and five minimum cases in child node. Method of validation was cross validation. After analysis, the results showed that all the selected variables contributed in the prediction of ECEC, final number of nodes were 9 with five terminal nodes and four tree depths.

Table 2: Tree growing method and Specifications

Specifications	Growing Method	CRT
	Dependent Variable	CEC
	Independent Variables	OC, pH, silt, clay
	Validation	Cross Validation
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	30
	Minimum Cases in Child Node	5
Results	Independent Variables Included	clay, silt, OC, pH
	Number of Nodes	9
	Number of Terminal Nodes	5
	Depth	4

3: Independent variables importance

The independent variable importance is presented in Table 3 and the barchart in Figure 1. The model summary indicated that all the independent variables selected for analysis made significant contribution to prediction of CEC except soil pH with insignificant contribution. Among the variables selected, clay was the most significant predictor of CEC with 5.7 % contribution, followed by organic carbon with 1.3 %, followed by silt with 1.2% while soil pH with 0.6% was the least predictor of CEC in the study area. The trend was as followed: clay fraction > organic carbon > silt fraction > pH. The high contribution of clay to CEC than organic matter and silt fractions in the study area could be attributed to charge development on clay. Although, kaolinite is the dominant clay mineral type in the study area, which is non-expanding with low charge development; the total number of charge developed on clay particles was more than that developed on organic matter due to lower quantity. Clay fraction has more surface area than silt fraction, which account for clay having more charge development and higher contribution to CEC than silt fraction in the study area.

Table 3: Independent Variable Importance

Independent Variable	Importance	Normalized Importance
clay	5.788	100.0%
OC	1.294	22.4%
silt	1.240	21.4%
pH	.561	9.7%

Growing Method: CRT

Dependent Variable: CEC

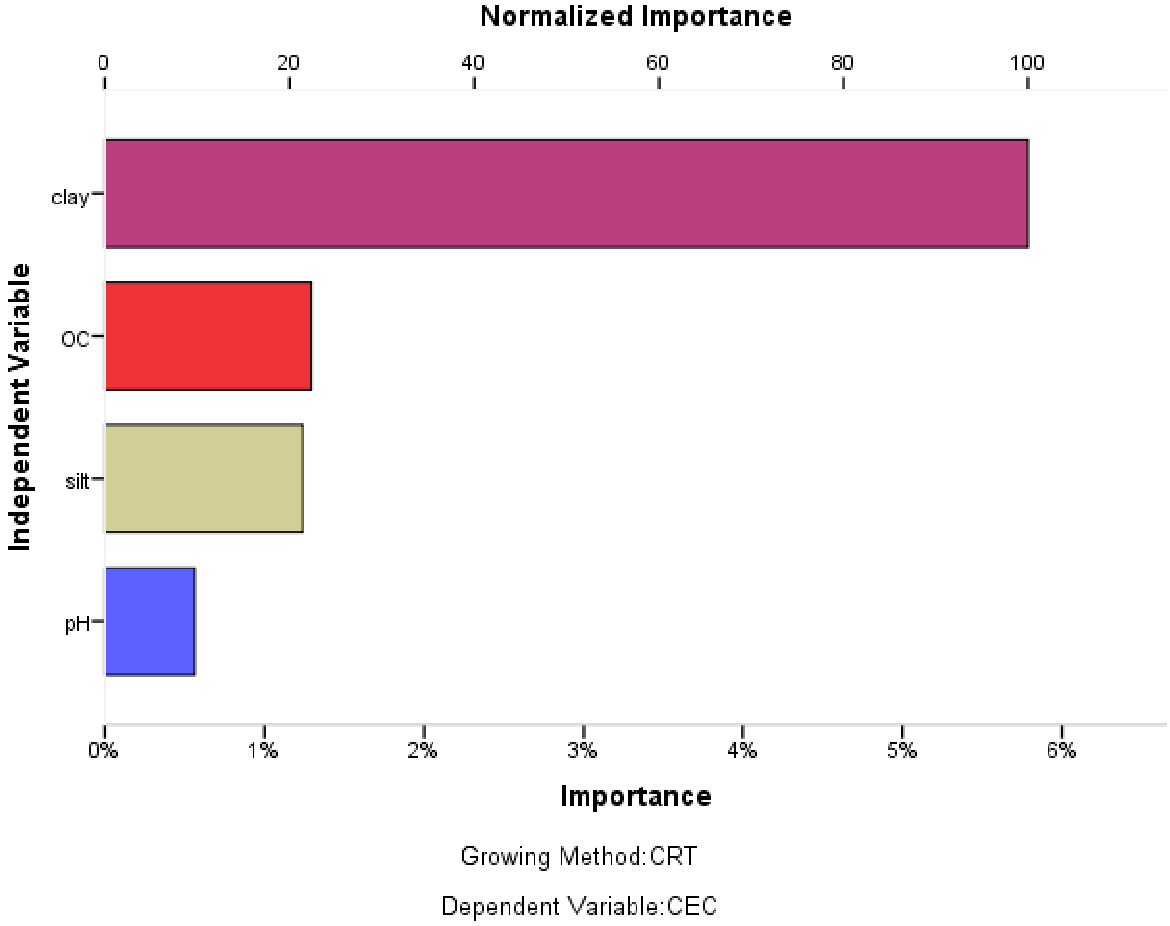

Fig1: **Independent Variable Importance**

3: PREDICTION MODEL OF CEC USING REGRESSION TREE

The regression tree model of CEC of the study area is presented in Table 4 and Figure 2. The prediction model showed that clay fraction less than or equal to 2.2 % predicted CEC of 1.97 cmol/kg (node 1) in the study area. Clay fraction of greater than 2.2 % combined with silt of less than or equal to 3.2 % to predict CEC of 11. 5 cmol/kg (node 3).Organic carbon of less than or equal to 4.3 %, combined with clay of less than or equal to 5.15% predicted a mean ECEC of 40.2 cmol/kg (node 3). Organic carbon content of less than or equal to 4.3% combined with clay of greater than 5.12 % and silt less than or equal to 3.75 % predicted a mean ECEC of 17.8 cmol/kg (node 5). Organic carbon content of less than or equal to 4.3%

combined with clay of greater than 5.12 % and silt of greater than 3.75 % predicted a mean ECEC of 24.3 cmol/kg (node 6). The organic carbon being the significant prediction of ECEC in the study area within 100 cm soil depth could be attributed to coarse texture soil of the study area with coastal plain sand, beach ridge sand, sandstone / shale as parent materials. Soils low in clay (< 15 %) are much more dependent on soil organic matter to provide CEC and biotic processes dominate, whereas soils high in clay (>35 %), biotic processes are minimal (Oades, 1993).

Table 4: Classification tree model

Tree Table

Node	Mean	Std. Deviation	N	Percent	Predicted Mean	Parent Node	Primary Independent Variable		
							Variable	Improvement	Split Values
0	8.8881	4.02052	75	100.0%	8.8881				
1	1.9729	.25934	7	9.3%	1.9729	0	clay	4.923	<= 2.200
2	9.6000	3.51239	68	90.7%	9.6000	0	clay	4.923	> 2.200
3	11.4784	3.65150	19	25.3%	11.4784	2	silt	1.240	<= 3.200
4	8.8716	3.20654	49	65.3%	8.8716	2	silt	1.240	> 3.200
5	5.4243	2.57305	7	9.3%	5.4243	4	OC	1.294	<= .676
6	9.4462	2.94980	42	56.0%	9.4462	4	OC	1.294	> .676
7	9.8662	2.84109	37	49.3%	9.8662	6	clay	.731	<= 22.530
8	6.3380	1.68351	5	6.7%	6.3380	6	clay	.731	> 22.530

Growing Method: CRT

Dependent Variable: CEC

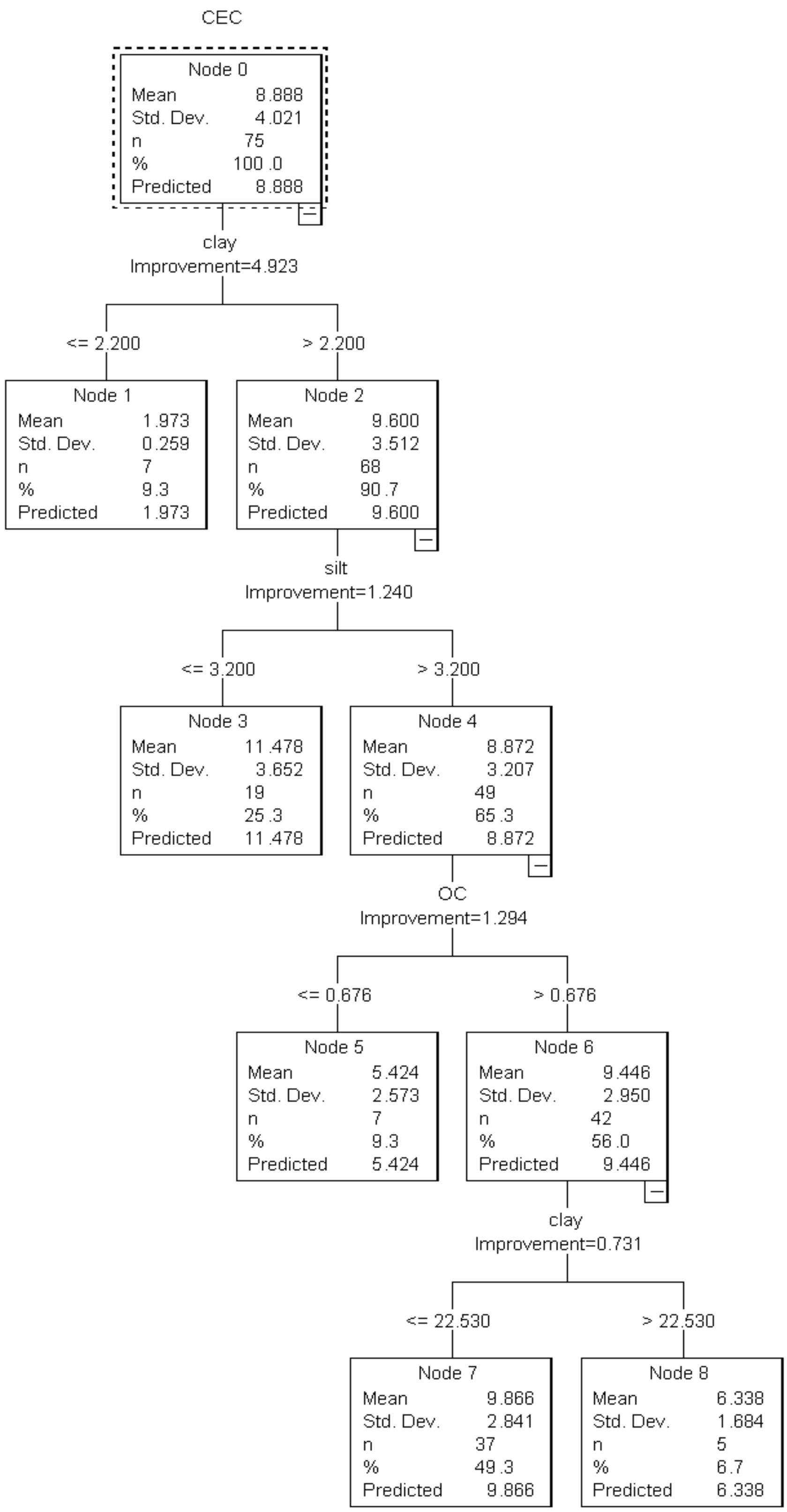

Fig.2: Graphical presentation of tree model

4.0: MAPPING OF ECEC USING REGRESSION TREE

Partitioning of the data space is shown in Figure 3 while Figure 4 shows the predicted map of ECEC of the study area. Node 2 with mean ECEC of 32.4 cmol/kg is located within the sandstone parent material while node 3 with mean ECEC of 40.2 cmol/kg is located within the shale parent material. Node 5 with mean ECEC of 17.8 cmol/kg is located within the beach ridge sand parent material while node 6 with mean ECEC of 24.3 cmol/kg is located within the coastal plain sand parent material. The model showed that organic carbon content was the only significant predictor of ECEC in sandstone soil while organic carbon in combination with clay made significant prediction of ECEC in shale parent material. In coastal plain sand and beach ridge sand soils, Organic carbon in combination with clay and silt made significant prediction of ECEC.

<table>
<tr><td colspan="2">Node 2:
Mean ECEC = 32.4 cmol/kg
.> 4.3 % organic carbon</td></tr>
<tr><td rowspan="2">Node 3:

Mean ECEC=40.2 cmol/kg

<= 4.3 % organic carbon, <= 5.15 % clay</td><td>Node 5: Mean ECEC = 17.8 cmol/kg
. <= 4.3 % organic carbon, <= 5.15 % clay
< 3.75 % silt</td></tr>
<tr><td>Node 6: Mean ECEC = 24.3 cmol/kg
. <= 4.3 % organic carbon, > 5.15 % clay,
> 3.75 % silt</td></tr>
</table>

Fig. 3: Partition of the data space

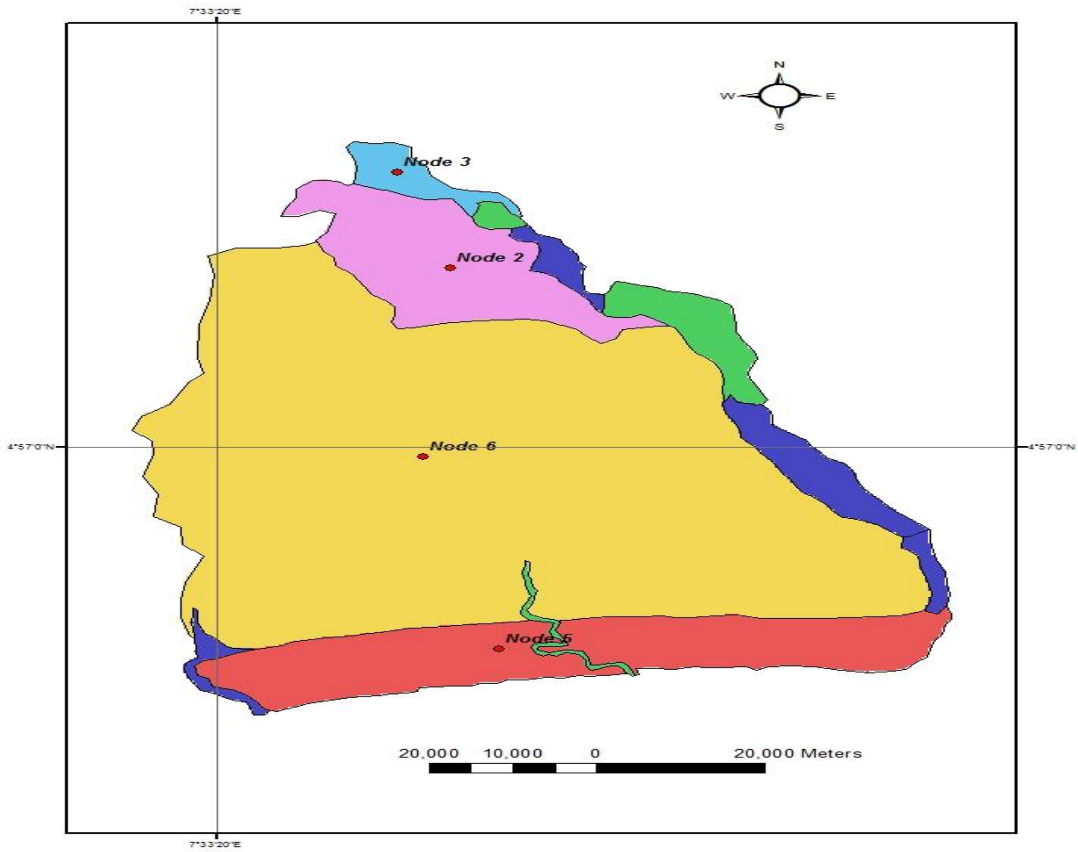

Fig.4: Predicted map of ECEC in the study area

5.0: EVALUATION OF THE MODEL

A plot of measured values on tha Y-axis and pedicted values in tha X-axis is shown in Fig. 5. The r^2 of the plot of measured ECEC values on tha Y-axis and pedicted values in tha X-axis was 1.0, indicating that the predicted values at the node were exact or very close to the actual values. This shows hat the model is a good model and provides information. But the risk estimate shown in Table 5, which is a measure of the within-node variance (unexplained variance) which is an indicator of model performance showed that the proportion of the variance explain by the model is 29%. The computations were as follows: total variance = within-node (error) variance (unexplained variance) + between node variance (explained variance). The within-node variance is the risk estimate value which is 111.87 (Table 5). Total variance is the variance for the dependent variable before consideration of any independent variables, which is the variance at the root node. The standard deviation displayed at the root node is 12.510; so the total variance is the square of 12.510 = 156.50. The proportion of variance due to error (unexplained variance) = 111.87/156.50 = 0.71. The proportion of the variance explains by the model 1-0.71 = 0.29 or 29%. The low proportion of variance explain by the model could be attributed to the use of ECEC instead of CEC for computation. In this study ECEC was computed from the summing up of exchangeable cations and exchangeable acidity instead of direct determination of CEC.

Table 5: Miscalculation Risk

Method	Estimate	Std. Error
Resubstitution	7.761	1.555
Cross-Validation	15.178	2.687

Growing Method: CRT

Dependent Variable: CEC

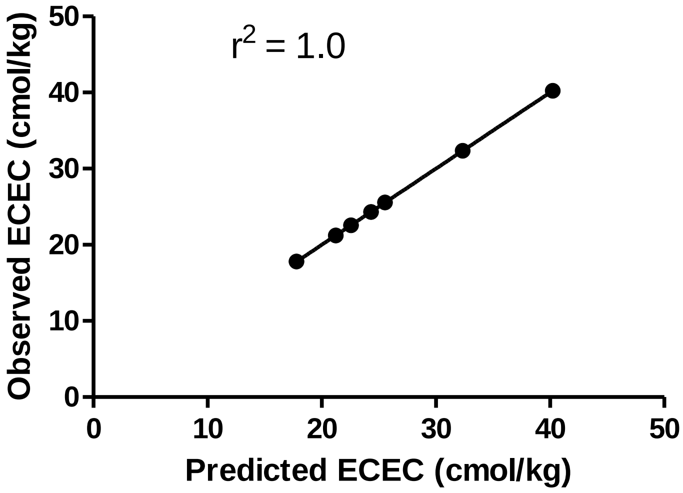

Fig. 5: Plot of measured mean ECEC vs predicted mean ECEC

6.0: Surrogates

The surrogate table (Table 6) indicates how surrgates were used in the model. For any missing value of organic carbon, soil pH was used as the surrogate predictor although it has a fairly low association value of only 0.04. For any missing value of clay, silt was used as the surrogate predictor, since this variable has a fairly moderate association of 0.4 with clay. Also clay was used as surrogate predictor for silt, and organic carbon for pH.

Table 6: Surrogates

Parent Node	Independent Variable		Improvement	Association
0	Primary	clay	4.923	
2	Primary	silt	1.240	
	Surrogate	clay	.134	.211
		OC	.000	.211
4	Primary	OC	1.294	
	Surrogate	pH	.561	.286
6	Primary	clay	.731	

Growing Method: CRT

Dependent Variable: CEC

Conclusion

The study revealed that ECEC can be predicted using soil organic carbon, clay, silt and soil pH in the study area. The results of independent variable importance to the model showed that organic carbon was the most significant predictor of ECEC, followed by clay, followed by silt while soil pH was the least predictor of ECEC in the study area. Based on the model, organic carbon content was the only significant predictors of ECEC in sandstone soil while organic carbon in combination with clay were the predictors of ECEC in soils developed from shale parent material. In coastal plain sand and beach ridge sand soils, organic carbon in combination with clay and silt were the predictors of ECEC. The model showed that organic carbon content was the only significant predictor of ECEC within 100 cm soil depth in the study area due to coarse texture of the soil (less than 15 % clay within 100cm soil depth). In the application of the model, independent variables included in the final model and measured in the same unit should be used.

REFERENCES

Bell, M.A., and Van Keulen J. 1995. Soil predotransfer functions for Mexican soils. Soil Sci. Soc. Am. J. 59:865-871.

Breiman, L.; Friedman, J.; Olshen, R. And Stone, C. 1984. Classification and Regression Trees. Belmont, CA: Wadsworth Int. Group.

Charman P. and Murphy, B. 2007. Soils: The Properties and Management. Oxford University Press, South Melbourne, Victoria. 3205pp.

Diplaris, S.; Symeonidis, A.L.; Mitas, P.A.; Banos, G. and Abas Z.2006. A decision based alarming system for the validation of national genetic evaluation. Journal of computer and electronics in agriculture 52: 21-35

Drake, E. H., and H. L. Motto. 1982. An analysis of the effect of clay and organic matter

content on the cation exchange capacity of New Jersey soils. Soil. Sci. 133:281-
288.

Enwenzor, W. O., Ochiri, A. C., Opuwaribo, E. E., Udo, E.J. 1980. A review of fertilizer use on crops in southern zone of Nigeria. In: *Literature Review on Soil Fertility Investigation in Nigeria*. FMANR. Lagos.6.

Huang, J.; Fang, H and Fan, X. 2010. Decision forest for classification of gene expression data. Computers in Biology and Medicine, Vol. 36, Issue 8, 698-704

John, B., Yamashita, T., Ludwig, B., and Flessa, H. 2005. Storage of organic carbon in aggregate and density fractions of soils under different types of land use. Geoderma. 128: 63-79

Keller, A., B. Von steiger, S.T. Van der Zee, and R. Schulin. 2001. A stochastic empirical model for regional heavy metal balances in agroecosystems. J.E.Q, 30: 1976-1989.

Krogh, L., B.M. Henrik, H.G. Mogens, 2000. Cation- exchange capacity pedotransfer functions for Danish soils. Acta Agriculturae Scandinavica, Section B- Plant Soil Science, 50; 1-12.

Manrique, L. A., C. A. Jones and P. T. Dyke. 1991. Predicting cation exchange capacity from
soil physical and chemical properties. Soil Sci. Soc. Am. J. 55:787-794.

Pachepsky, Y.A. and W.J. Rawls. 1999. Accuracy and reliability of pedotransfer functions as
affected by grouping soils. Soil Sci. Soc. Am. J, 63: 1748-1757.

Parfitt, R.L., D.J. Giltrap, and J.S. Whitton. 1995. Contribution of organic matter and clay minerals to the cation exchange capacity of soils. Commun. Soil Sci. Plant Anal. 26:1343–1355.

Rengasamy, P. and Churchman, G.J. 1999. Cation Exchange Capacity, Exchangeable Cations
and Sodicity. In Soil Analysis: an Interpretation Manual. Csiro Publishing. Collingwood, Vic. 147-157

Stevenson, F.J. 1982. Humus Chemistry- Genesis, Composition and Reactions. John and Wiley and Sons, New York, pp 158

Soil Survey Staff. (2006). Soil Taxonomy. A Basic System of Soil Classification for Making and Interpreting Soil Surveys. USDA Agric. Handbook. No.436, second edition. U.S.Govt. Printing Office, Washington, D.C. 871.

Solomon, D., Fritzsche, F., Tekalign, M., Lehmaann, J.& Zech, W. 2002. Soil organic matter composition in the subhumid Ethiopian highlands as influenced by deforestation and agricutural management. *Soil Sci. Soc. Am. J*. 66, 68-82

Thomas G. W. 1982. Exchangeable cation. Methods of Soil Analysis Part 2, Agronomy, Monography 9 (2nd edition). SSSA Madison, Wisconsin and USA, pp 159-165

Nelson, D.W., Sommers, L. E. 1996. Total organic carbon and organic matter. In: *Methods of Soil Analysis, 2nd Ed.* D. L.Spans; A.L. Page; P.A. Helmke; R.H. Leoppert; P.N. Soltanpour; M.A. Tabatabai & C.I. Johnson, Eds. American Society of Agronomy: Madison, WI, US.A pp. 961-1010.

Oades, J.M. 1993. The role of soil biology in the formation, stabilization and degradation of soil structure. *Geoderma*, 56, 377-400.

Petters, S.W.; Usoro, E.J.; Udo, E. J.; Obot, U. W. & Okpon, S.N. 1989. Akwa Ibom State Physical Background, Soils and Land Use Ecological Problems. Technical Report of the Task Force on Soils and L anduse. Govt. Printer Uyo, pp 602.

Udo, E. J., Ibia, T.O., Ogunwale, J.A., Ano, A.A. & Esu I. E. 2009. Manual of Soil, Plant and Water Analysis. Sibon Books Limited, Lagos, pp 82-92

Whitebread, A.M., Lefroy, R.D.B. & Blair, G.J. 1998 . A survey of the impact of cropping on soil physical and chemical properties New South Wales. *Aus.J. Soil Res. 36,669-681*.

Wagner, B., V.R. Tranawski, V. Hennings, U. Muller, G. Wessolek,.and R. Plagge. 2001. Evaluation of pedotransfer functions for unsaturated soil hydraulic conductivity using an independent data set. Geoderma, 102: 275-279.