DOI: https://dx.doi.org/10.4314/gjass.v23i1.7

GLOBAL JOURNAL OF AGRICULTURAL SCIENCES VOL. 23, 2024: 71-83
COPYRIGHT© BACHUDO SCIENCE CO. LTD PRINTED IN NIGERIA ISSN 1596-2903 e-ISSN: 2992 – 4499
www.globaljournalseries.com.ng, Email: globaljournalseries@gmail.com

ASSESSMENT OF SOIL QUALITY FOR SUSTAINABLE LAND MANAGEMENT USING MACHINE LEARNING AND DIGITAL SOIL MAPPING TECHNIQUES IN OBUDU CATTLE RANCH, NIGERIA

AFU, S.M., OLIM, D.M., AFANGIDE, A.I., EDIENE, V.F., AKPAMA, A.I. AND BISONG, S.B.
E-mail: [2]olimdenis001@gmail.com

(Received 11 April 2024; Revision Accepted 23 May 2024)

ABSTRACT

Soil quality assessment is essential to know variation in nutrient concentrations within landscape for sustainable soil management. This study assessed soil quality in Obudu cattle ranch using machine learning and digital techniques. A total of 60 composite soil samples (0–30 cm depth) were collected at intervals of 200–500 m and selected soil physicochemical properties were determined. Digital elevation model (DEM) and Sentinel-2 satellite imageries were obtained, processed and applied for modelling. Soil quality was measured using total dataset (TDS) and minimum dataset (MDS). Linear (L) and non-linear (NL) scoring functions were applied, yielding four indices, MDS_L, MDS_NL, TDS_L and TDS_NL. Sixteen soil quality indicators (SQI) were used as TDS and were further screened for MDS using principal component analysis (PCA). Multiple linear regression was used to predict soil quality index in unsampled locations. The result showed that the soils were sandy loam, loam and sandy clay loam in texture. pH ranged from strongly acidic to slightly acidic. Soil organic carbon, CEC and base saturation were high while available P, exchangeable cations, exchangeable acidity as well as ECEC were low. The mean estimated soil quality for MDS_L, MDS_NL, TDS_L and TDS_NL were 0.415, 0.51, 0.42, and 0.49 respectively. MDS_NL model was the best model in predicting soil quality index in the area. All the models showed almost similar spatial distribution, with a high soil quality region mostly found in the southwestern part while low soil quality areas were located mostly in the central part and northwestern part of Obudu Mountain Resort. The soil quality prediction class showed moderate class (class III) to be the dominant class covering greater part of the area with MDS_NL model. The predictive maps derived from this study should serve as a guide in the establishment of regionalized soil nutrient management programmes.

KEYWORDS: soil quality, dataset, land management decisions, Obudu mountain

INTRODUCTION

The Obudu Mountain Resort, located in Cross River State, Nigeria, is a popular tourist destination known for its scenic beauty and rich biodiversity. However, the region's soil quality and land management practices have not been comprehensively studied.

Knowing soil quality is essential when considering land degradation assessment, soil management, crop production and food security in Obudu Mountain Resort. Traditional soil assessment methods are often time-consuming, labour-intensive and costly, making them impractical for large-scale applications.

Afu, S.M., Department of Soil Science, Faculty of Agriculture, University of Calabar, Nigeria
Olim, D.M., Department of Soil Science, Faculty of Agriculture, University of Calabar, Nigeria
Afangide, A.I., Department Soil Science & Technology, Federal University of Technology, Owerri, Nigeria
Ediene, V.F., Department of Soil Science, Faculty of Agriculture, University of Calabar, Nigeria
Akpama, A.I., Department of Soil Science, Faculty of Agriculture, University of Calabar, Nigeria
Bisong, S.B., Department of Soil Science, Faculty of Agriculture, University of Calabar, Nigeria

However, recent advancements in machine learning and digital soil mapping techniques have provided promising alternatives for assessing soil quality in a more efficient and cost-effective manner (Chakraborty et al., 2019).

A decrease in soil quality can disrupt essential soil functions and potentially impact crop production and food security. However, evaluating soil quality is crucial for identifying areas with varying soil quality levels and assessing their suitability for agricultural land use, specifically for cultivated crops. This information is valuable for farmers, land managers, and policymakers to make informed decisions regarding sustainable land management practices as it can provide insights into potential soil degradation and nutrient depletion. Various models, such as fuzzy set techniques (Rezaee et al., 2020), Nemoro soil quality index, simple additive soil quality index (Mukherjee and Lal, 2014), weighted additive soil quality index (Vasu et al., 2016), among others have been developed and utilized to estimate soil quality index. It is important to note that unlike air and water, soil is not directly consumed making it challenging to estimate its quality (Debi et al., 2019). Additionally, soil quality is not determined by a single factor but rather the integration of physical, chemical and biological factors which are considered in soil quality index quantification (Shekhovtseva and Maltseva, 2015). Therefore, an initial approach to evaluating soil quality in cultivated land involves estimating the soil quality index using indicators that are sensitive to changes in soil management practices. Soil quality indices are models that provide numerical data on the soil's capacity to perform one or more functions (Asensio et al., 2013). The selection of soil quality indicators such as physical, chemical and biological properties, depends on their sensitivity to influencing soil function and the available financial resources. In this study, soil physical parameters, along with selected chemical properties were chosen as proxies to better understand soil quality in sub-Saharan Africa.

Evaluating soil quality typically requires a large number of sampling points, and conducting conventional soil surveys to quantify soil quality over extensive areas can be laborious, expensive and time-consuming. To overcome these challenges, the digital soil mapping (DSM) technique offers an alternative approach by utilizing soil properties, remotely-sensed data, digital elevation models (DEMs), micro-climatic data, land use and cover data and geological data as covariates or ancillary variables, aided by geo-statistics and machine learning to predict soil quality index or soil quality classes for unsampled areas (Nabiollahi et al., 2018; John et al., 2021; Zeraatpisheh et al., 2020). Machine learning and geostatistical models have already been employed to predict soil properties in sub-Saharan Africa, particularly in Nigeria (Ogunwole et al., 2014). The application of these tools has expanded in West Africa including the Africa Soil Information Services project (AfSIS) (Hengl et al., 2017), where they were used to model the spatial distribution of selected soil nutrient indicators, albeit at a coarse scale. Many sub-Saharan African countries have since adopted similar methods to create detailed maps of soil nutrients at various scales. However, the use of machine learning and geostatistics to model soil quality remains relatively limited in global studies (Nabiollahi et al., 2018; Paul et al., 2020; Zeraatpisheh et al., 2020).

Soil quality of any given area is affected by several factors among which parent materials and environmental factors are said to exert the strongest influence in tropical soils. Remote-sensing-based variables and topographic data are identified as the major drivers of spatial variability in soil quality index. Despite the active engagement in food crop production, they have been no feasible studies conducted on this approach in sub-Saharan Africa. Nevertheless, there is a growing interest in applying machine learning and geostatistical techniques to generate detailed soil quality maps in the region aiming to monitor soil resources and support crop production in light of the threat posed by land degradation in African farming systems. Thus, information and knowledge about soil quality are crucial for guiding management and restorative measures, enabling farmers and land managers to make informed decisions. These insights can help prevent inappropriate land use and soil management which can lead to soil quality deterioration, reducing crop production, food security, economic growth and environmental health. In Nigeria and in Cross River State in particular, the hectarage of land committed to agricultural cultivation has recently doubled due to increasing demand for food. With this level of engagement in crop cultivation in the state, assessment of soil quality which is a necessary tool for sustainable and optimal crop production is lacking across the State. This is because evaluating and monitoring soil quality can reveal areas of land with high potential for degradation and provide guide for farmers, environmentalists and policy makers on formulating best soil management techniques. Therefore, the purpose of this study is to estimate the soil quality index to assist farmers, land managers and policymakers in making decisions related to sustainable cropland management in Cross River State, Nigeria.

2 MATERIALS AND METHODS

2.1 The study area and soil sample collection

The study area, Obudu cattle ranch, is located in Obanliku Local Government Area of Cross River State, between latitudes 6° 21′ N and 6° 24′ N and longitudes 9° 22′ E and 9° 25′ E (Figure 1), and has varied soil types, land use and cover and topography with elevation reaching 1654 m Above Sea Level (asl). The area has moist tropical humid climate with rainfall exceeding 2000 mm/annum, temperature range from 15 to 31.8 ^{0}C.

The vegetation is tropical rainforest and the major crops grown in the area are oil palm, cocoa, maize, banana, plantain, okra, groundnut and cocoyam. Sixty georeferenced soil samples were collected randomly to cover the entire study area at the distance of 200 m -500 m apart. At each sampling point, a composite of 3 samples (0-30 cm soil depth) was collected randomly within the grid area using soil auger, hand mixed, placed in labeled plastic bags and transported to the laboratory for physicochemical analysis.

2.2 Laboratory analysis

In the laboratory, the samples were processed using standard procedure for analysis, bagged, labeled and analyzed. Particle size was determined by Bouyoucos hydrometer method as outlined by Gee and Or (2002). pH was determined potentiometrically in a soil: water suspension (1:2.5) using the procedure outlined by Udo et al. (2009), while organic carbon was

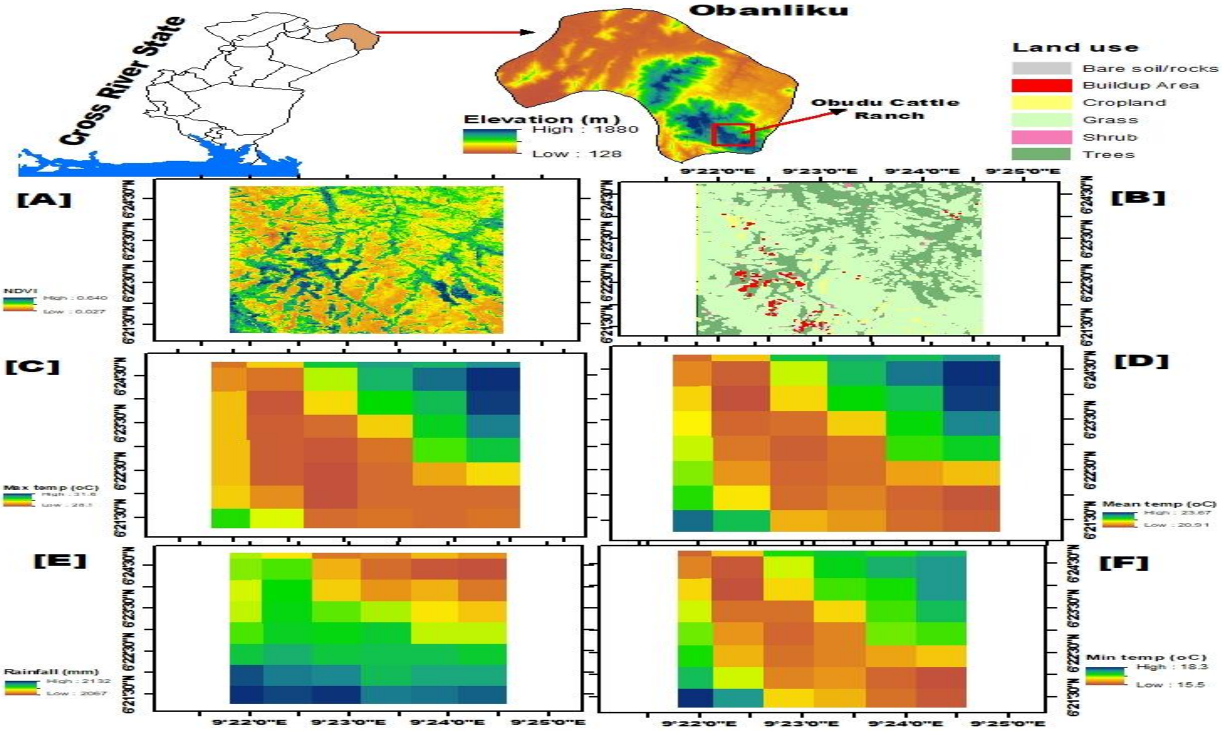


Fig. 1: Map of Cross River State showing location of the study area

determined by Walkley-Black wet oxidation method using acid dichromate ($K_2Cr_2O_7$) method outlined by Nelson and Sommers (1996). Total nitrogen was determined using modified micro-Kjeldhal method (Udo et al., 2009), while available phosphorus was determined using Bray P-1 according to the procedures of Kuo (1996). Exchangeable cations were determined using the extract obtained after leaching samples with one normal neutral ammonium acetate (1 N, NH_4OAc, pH 7.0) solution. Calcium and magnesium were determined by the EDTA titration method while potassium and sodium were estimated by flame photometer using the method of Udo et al. (2009). Exchangeable acidity was determined by titration using 0.1 N Na0H solution following procedure in Udo et al., (2009). CEC was determined using the procedure outlined by Udo et al. (2009) while effective cation exchange capacity (ECEC) and base saturation were calculated.

ENVIRONMENTAL DATA

Environmental covariates such as elevation, slope and aspect were derived from the digital elevation model (DEM) obtained from ASTER data at a spatial resolution of 30 m and were processed using the System for Automated Geoscientific Geographical Information System (SAGA-GIS) software terrain analysis toolbox. Cloud-free Sentinel-2 imageries were acquired from the European Space Agency's Copernicus Open Access Hub. Spectral indices such as land surface temperature (LST), normalized difference vegetation index (NDVI), soil-adjusted vegetation index (SAVI), canal network base level and normalized difference moisture index (NDMI) were estimated from the images. Interpolated soil properties; texture (clay), soil organic carbon (SOC) and pH were obtained by interpolating the soil database at a resolution of 30 m while mean rainfall, minimum temperature, maximum temperature and mean temperature, covering the study area, were obtained from the WorldClim database and processed using ArcGIS software. All maps were geo-referenced to the Universal Transverse Mercator (UTM) Zone 32 N coordinate system and the area of interest (AoI) for soil and environmental data was demarcated using the polygon feature of the study areas in ArcGIS software.

SOIL QUALITY COMPUTATION PROCEDURE

Computation of soil quality index (SQI) in this study was done in three steps, including, selection of soil quality (SQ) indicators using both total dataset (TDS) and minimum dataset (MDS), indicator transformation and indicator integration into overall index. Principal components analysis (PCA) was used as a method of MDS selection. In using PCA, only the principal components (PCs) with eigenvalues ≥1 and which explained at least 5 % variation of the data were retained for interpretation. Andrews and Carroll (2001) suggested that indicators with weighted absolute values within 10 % of the highest indicator value for each PC should be selected as the MDS. This criterion only accounts for the loading of variable to a single PC and does not provide information for the variable on a multi-dimensional space, hence norm values as suggested by Yemefack et al. (2006) were used for grouping and selection of variables as MDS. The norm value is obtained using:

$$N_{ik} = \sqrt{\sum_{i=1}^{k}\left(u_{ik}^{2}\lambda_{ik}\right)} \qquad 1$$

where,

N_{ik} = load for i^{th} soil property on PCs with eigenvalues ≥1,

u_{ik} = load for the i^{th} soil property on the principal component of k,

λ_{ik} = the eigenvalue of the i^{th} soil property on the principal component of k

In indicator transformation each soil indicator has different units, they were transformed and normalized into a unitless score between 0 and 1 using both linear and non-linear scoring approach before final integration into overall soil quality index (SQI). Three established soil scoring functions (SSFs) based on if it has a “negative” or “positive” relationship with soil quality, or if it is positively or negatively related within an “optimum range” (Li et al., 2018) were used. The linear and non-linear functions (Equations 2 to 4) used are presented below. Equations (2) and (3) are linear functions and correspond to “less is better” function (L), and “more is better” function (M), respectively, while equation (4) is a non-linear (NL) model:

$$S_i = 1 - \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) \qquad 2$$

$$S_i = \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) \qquad 3$$

$$S_{NL} = \frac{1}{\left[1 + \left(x / x_o\right)^{b}\right]} \qquad 4$$

were

x = measured value of the soil quality indicator, x_{min} = minimum value of soil quality indicator, x_{max} = maximum value of soil quality indicator, b is the slope assumed to be -2.5 for more is better and +2.5 for less is better, and x_o is the mean value of soil variable.

For indicator integration into overall index the soil quality index (SQI) of each sampling site was calculated using the integrated quality index (IQI) (Equation 5)

$$\mathrm{IQI} = \sum_{i=1}^{n} W_i S_i \qquad 5$$

where, IQI is the weighted additive soil quality index, n is the number of selected soil properties, S_i is the score of soil properties, i, W_i is the assigned weight of each soil property for the TDS and MDS based on communality of Principal Component Analysis (PCA).

The weights were computed as:

$$W_i = \frac{C_i}{\sum_{i=1}^{n} C_i} \qquad 6$$

where, ith is the soil variable, and C_i is the communality value of a soil variable, ith

Soil quality index validation: The performance of the proposed SQ models were validated using sensitivity index (SI)(Equation7)

$$\mathrm{SI} = \frac{\mathrm{SQI}_{\max}}{\mathrm{SQI}_{\min}} \quad 7$$

where, SQI_{max} = maximum soil quality index value and SQI_{min} = minimum soil quality index value

Statistical Analysis: Descriptive statistics, including mean, minimum, maximum, standard deviation, coefficient of variation, skewness and kurtosis were used to describe the data distributions. Multiple linear regression (MLR) was employed to model the relationship between SQI and selected environmental predictors. Thus, the SQI of interest is predicted using:

$$\hat{y}_{(i)} = \hat{\beta}_0 + \sum_{k=1}^{k} \hat{\beta}_k X_{k(i)} \quad 8$$

where, ŷ(i) is the predicted SQI at location i, $\hat{\beta}_0$ the estimated intercept, $\hat{\beta}_k$ the estimated regression coefficient for predictor k and $X_{k(i)}$ the value for the k^{th} predictor at location i.

Model evaluation

Coefficient of determination (R^2), root mean square error (RMSE), and Lin's concordance correlation coefficient (CCC) were used to compare and select the best soil quality model for prediction of SQI (Equations 9,10 and 11).

$$\mathrm{R}^2 = 1 - \frac{\sum_i^n \left(Z_{oi} - Z_{pi}\right)^2}{\sum_i^n \left(Z_{oi} - \bar{Z}_{pi}\right)^2} \quad 9$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Z_{pi} - Z_{oi}\right)^2} \quad 10$$

$$CCC = \frac{2\mathrm{r}\sigma_o\sigma_p}{\sigma_o^2 + \sigma_p^2 + (\bar{\mathrm{Z}}_p \_ \bar{\mathrm{Z}}_o)} \quad 11$$

where, Z_{pi}= predicted values, Z_{oi} = observed values, n = the size of the observations, for the i-th term observation, $\bar{Z}_p$ = average of the predicted values, $\bar{Z}_o$ = average of the predicted values, CCC = Lin's concordance correlation coefficient, σ_o^2 and σ_p^2 are the variances of the predicted and observed values, and r is the Pearson correlation coefficient between the predicted and observed values.

RESULTS AND DISCUSSION

Physicochemical properties

The soil physicochemical properties used in this study to estimate soil quality index are given in Table 1. The pH varied from strong to slightly acidic with values ranging from 5 to 6.6, with a mean of 5.44 . At a pH of 5.44, most of the important soil nutrient elements needed by plants may be fixed (Upadhyay and Raghubanshi 2020), since most nutrient elements are available at near-neutral pH. Soil organic carbon was generally high since there may be other low values between 2.6 to 44.7 mg/kg, while total nitrogen ranged from 0.1 to 3.8 mg/kg. Available phosphorus ranged from 2 to 27 mg/kg with a mean of 4.3 mg/kg. Ca, Mg, K and Na had means of 1.989 cmol/kg, 0.777 cmol/kg, 0.094 cmol/kg and 0.074 cmol/kg, respectively. Sand was the most dominant among the soil separates with mean value of 589 %, while silt and clay had means of 266 % and 145 % (Table 1). However, the values of coefficient of variation (CV) were all < 15 % meaning that the soil properties were homogeneously distributed in the study area (Essington, 2006). All the soil properties were low except base saturation (high > 60 %), OC (high >2%) and CEC (high >25 cmol/kg) (Landon, 1991)

TABLE 1: Summary Statistics for Studied Soil Quality Indicators

Soilquality indicators	Minimum	Maximum	Arithmetic Mean	Standard Deviation	CV (%)	Skewness	Kurtosis
pH	5	6.6	5.443	0.438	0.08	1.329	0.399
OC (mg/kg)	2.6	44.7	25.258	9.597	0.38	-0.155	-0.441
TN (mg/kg)	0.1	3.8	2.167	0.826	0.381	-0.252	-0.261
Av.P (mg/kg)	2	27	4.3	3.309	0.77	5.725	38.586
Ca (cmol/kg)	0.8	7.6	1.989	1.302	0.655	2.812	8.157
Mg (cmol/kg)	0.2	2.8	0.777	0.593	0.763	1.88	3.217
K (cmol/kg)	0.07	0.13	0.094	0.011	0.118	0.765	1.329
Na (cmol/kg)	0.05	0.11	0.074	0.012	0.161	0.621	0.341
Al (cmol/kg)	0	1.12	0.397	0.313	0.789	0.12	-1.004
H (cmol/kg)	0.08	1.16	0.584	0.165	0.283	0.586	2.429
ECEC (cmol/kg)	2.3	10.68	3.926	1.635	0.416	2.476	6.638
BS (%)	6.8	96	70.713	14.374	0.203	-1.196	5.328
CEC (cmol/kg)	15	44	25.9	6.701	0.259	0.66	-0.026
Sand (g/kg)	492	732	589	64.711	0.11	0.492	-0.512
Silt (g/kg)	160	370	266	53.526	0.201	0.101	-0.703
Clay (g/kg)	68	268	145	35.668	0.246	0.623	1.532

CV = Coefficient of Variation

TOTAL DATASET (TDS) AND MINIMUM DATASET (MDS) SELECTION

The soil properties presented in Table 1 were used as total dataset in this study. The use of total dataset in studying soil quality has been reported to cause change in soil functions (Andrews et al., 2004). Physicochemical properties have been used to study soil quality by many researchers (Isong et al., 2022; Fathizad et al., 2020; Choudhury and Mandal, 2021). The indicators were screened for the MDS using correlation and principal component analysis (PCA) with varimax rotation. The correlation result presented in Fig. 2, shows that strong positive and significant correlations were observed between pH and BS ($r = 0.66$; $p<0.01$), pH and ECEC ($r = 0.62$; $p<0.01$), pH and Na ($r = 0.53$; $p<0.01$), pH and K ($r = 0.55$; $p<0.01$), pH and Mg ($r = 0.47$; $p<0.01$) and, pH and Ca ($r = 0.26$). Similarly, negative and significant correlations were observed between pH and OC ($r = -0.39$; $p<0.01$), pH and TN ($r = -0.38$; $p<0.01$), pH and Al ($r = -0.69$; $p<0.01$), and pH and CEC ($r = -0.29$; $p<0.01$). The observed relationships from the correlation analysis were indicative of the intricate connections among the various soil properties which could hardly be observed when using raw data obtained directly from laboratory analysis.

In this study, the studied soil indicators were grouped into six PCs, as they had eigenvalues >1, each explaining at least 5 % of the data variation and accounting for 84.43 % of the total variance in the dataset (Table 2). In Group 1, ECEC had the highest norm value (2.19), and no other soil indicator had a norm value falling within the scope of 90 % of the highest value, hence ECEC was selected as MDS for group 1 and other indicators were eliminated. Similarly, in Group 2, OC had the highest norm value (1.6) which exceeded those of other variables and was selected for group 2. In group 3, phosphorus was selected, in group 4 H was selected. In group 5, sand and silt were selected while in group 5 clay was selected (Table 2).

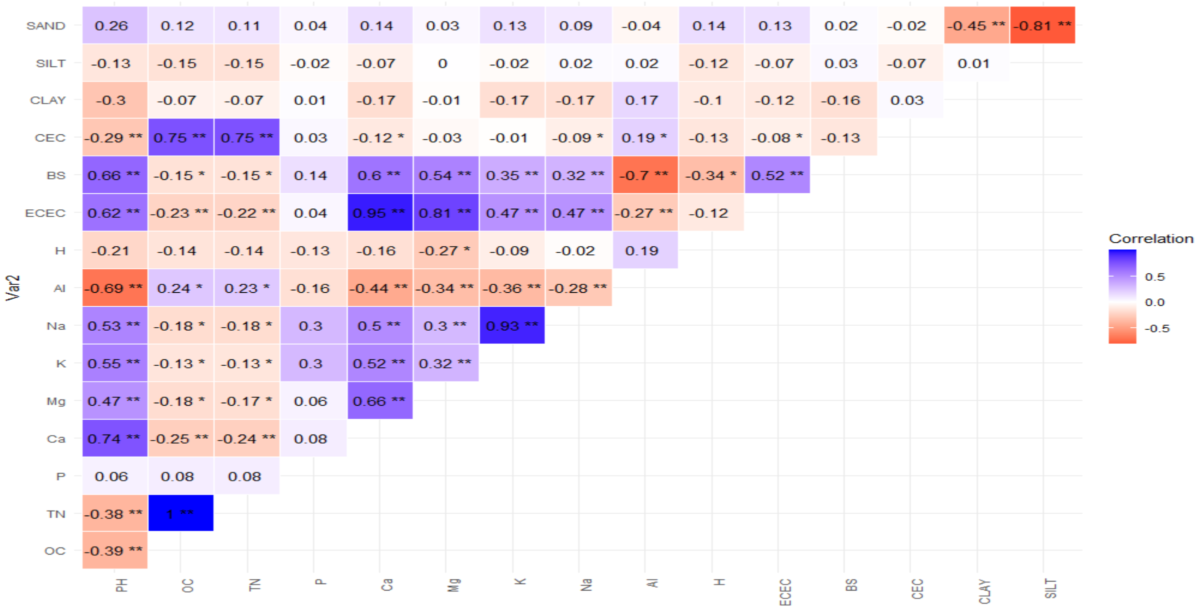


Fig. 2 Correlation between soil quality indicators
Note: ** and * mean significant at 1 % and 5 % level of probability

SOIL QUALITY INDICES

The results presented in Table 3 are the weight values of TDS and MDS. The weight value for TDS showed that ECEC (0.0706) had the highest weight, while H (0.0446) had the lowest value. Similarly, for MDS, sand (0.209) had the highest weight while clay (0.067) had the lowest value. The soil properties are indicators of soil texture, nutrient and soil acidity and could play an important role in assessing soil quality of an area. The screened indicators for both TDS and MDS were scored using linear and non-linear scoring functions (Table 3). The results based on sensitivity index showed that MDS_L was the most sensitive index with a value of 2.76 for evaluation of soil quality index in the study area, while MDS_NL had the least sensitivity index value of 1.73 (Table 3). The MDS_L selected through validation using sensitivity analysis was further correlated with soil organic matter (OM) (Fig. 3) to check its scientific credibility. The result showed strong positive correlation between MDS_L and OM ($r = 0.279$, $p < 0.05$). This implies that soil quality in the area was sensitive to change in OM content. This means that MDS_L can be used to monitor soil quality and crop yield in the study area.

TABLE 2: Results of the Principal Component Analysis

									TDS		MDS	
Indicators	PC1	PC2	PC3	PC4	PC5	PC6	Norm value	G	Com	weight	Com	weight
pH	0.517	-0.335	0.282	0.515	0.173	0.311	1.539303	1	0.852	0.063064		
OC	-0.159	**0.956**	-0.04	-0.014	0.096	0.038	1.601293	2	0.951	0.070392	0.682	0.152573
TN	-0.151	**0.956**	-0.043	-0.008	0.094	0.04	1.597226	2	0.948	0.07017		
P	-0.205	0.092	**0.665**	0.281	0.108	-0.288	1.149147	3	0.666	0.049297	0.422	0.094407
Ca	**0.854**	-0.132	0.235	0.245	0.074	0.125	2.022192	1	0.883	0.065359		
Mg	**0.834**	-0.026	0.036	0.259	-0.01	-0.126	1.940686	1	0.78	0.057735		
K	0.375	-0.042	**0.843**	0.082	-0.003	0.168	1.473568	3	0.888	0.065729		
Na	0.363	-0.104	**0.857**	0.005	-0.039	0.175	1.479915	3	0.909	0.067283		
Al	-0.172	0.235	-0.191	**-0.775**	-0.041	-0.213	1.131041	4	0.769	0.056921		
H	-0.132	-0.256	0.033	**-0.64**	0.152	0.293	0.98403	4	0.603	0.044634	0.639	0.142953
ECEC	**0.95**	-0.085	0.179	0.088	0.062	0.036	2.19868	1	0.954	0.070614	0.762	0.17047
BS	0.452	-0.09	0.121	**0.739**	-0.026	0.144	1.387149	4	0.794	0.058771		
CEC	0.041	**0.878**	0.018	**-0.088**	-0.038	-0.042	1.435799	2	0.783	0.057957		
Sand	0.063	0.046	0.057	-0.023	**0.906**	0.321	1.060406	5	0.934	0.069134	0.936	0.209396
Silt	-0.051	-0.078	0.014	0.033	**-0.962**	0.088	1.072052	5	0.942	0.069726	0.732	0.163758
Clay	-0.018	-0.065	-0.069	-0.098	-0.135	**-0.902**	0.94902	6	0.85	0.062916	0.298	0.066667
Eigenvalue	5.247	2.643	1.956	1.428	1.199	1.036						
Variance (%)	32.792	16.519	12.227	8.924	7.491	6.476						
cumulative Variance (%)	32.792	49.311	61.538	70.462	77.953	84.429						
KMO	0.505											
BTS	1167.7											
Df	120											
Sig	0.001											

PC = principal component; Bold PC loadings are considered highly weighted. Bold and Underlined factor loadings are selected as minimum data set (MDS); Com = communality; TDS = total dataset; MDS = minimum dataset. Grouping was based on norm values; Kaiser-Meyer-Olkin Measure of Sampling Adequacy; Bartlett's Test of Sphericity; df = degree of freedom; KMO = Kaiser-Meyer-Olkin; BTS = Bartlett’s test of Sphericity; G = Grouping

Table 3: Summary Statistics of Soil Quality Indices

	Minimum	Maximum	Mean	Std. Deviation	SI
MDS_L	0.21	0.58	0.3781	0.07421	2.76
MDS_NL	0.34	0.59	0.4475	0.05485	1.73
TDS_L	0.26	0.64	0.4078	0.08369	2.46
TDS_NL	0.34	0.61	0.4611	0.06817	1.79

SI = Sensitivity index

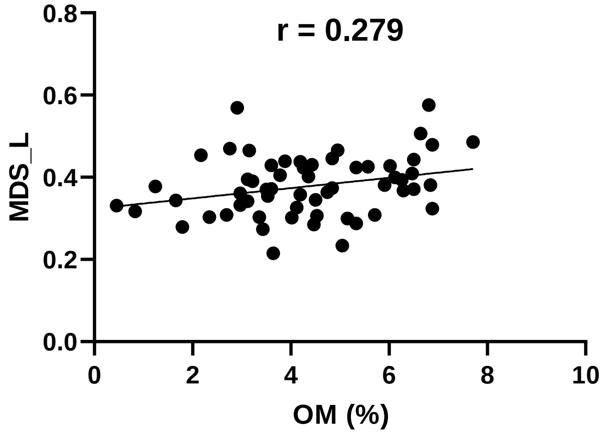
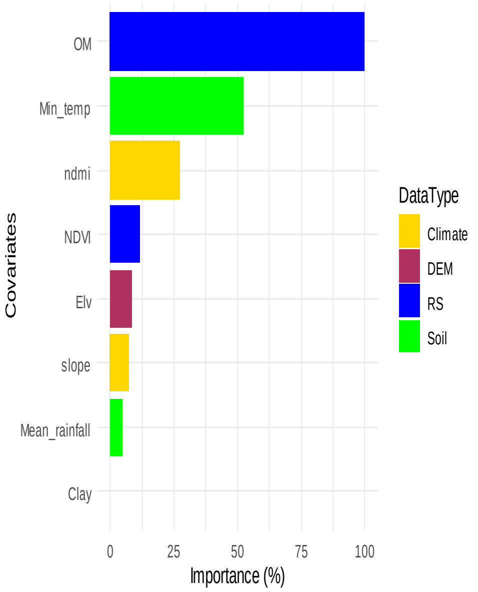


Fig. 3 Linear relationship between soil quality index (MDS_L) and organic matter

SELECTED COVARIATES FOR MODELLING

The variate inflation factor (VIF) was used in selecting covariates utilized in modeling soil quality in this study. MLR was used to quantified the effect of predictors on the soil quality index (Figs. 4-7). Organic matter (OM) was ranked first, with a relative importance of about 100% in both TDS and MDS of soil quality index. However, in MDS models, climate (minimum temperature) and remote sensing (RS) variables (NDVI and NDMI) were among the top four variables influencing soil quality index, whereas, in TDS models, clay, elevation and rainfall variables were among the top four variables influencing soil quality index in the study area. NDMI and NDVI were also the most important variables that detected variability of SOC in a study carried out by Falahatkar et al. (2016). In line with findings of this study, Paul et al. (2020), Fathizad et al. (2020) and Zeraatpisheh et al. (2019), reported organic matter, climate, NDVI and NDMI as factors affecting the prediction of soil quality index.



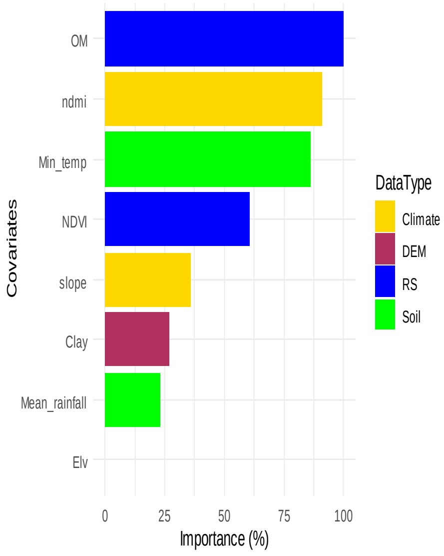
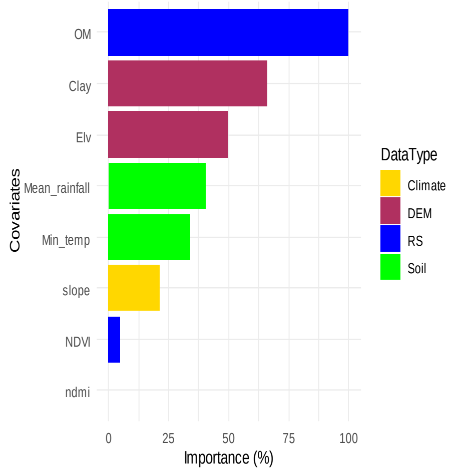


Fig. 4: Importance of variables in the MDS_L soil quality Fig. 5: Importance of variables in the MDS_NL soil quality index via MLR
index via MLR



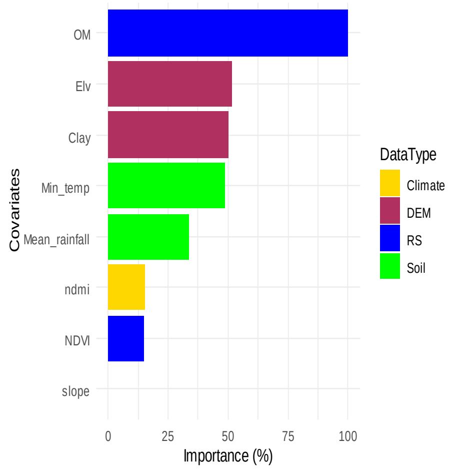


Fig. 6: Importance of variables in the TDS_L soil quality Fig. 7: Importance of variables in the TDS_NL soil quality index via MLR
index via MLR

SPATIAL PREDICTION OF SOIL QUALITY INDEX AND PERFORMANCE OF MACHINE LEARNING MODELS

The spatial distribution of soil quality index class predicted by MLR based on MDS_L, MDS_NL, TDS_L, and TDS_NL are shown in Fig. 8. The area had soil quality ranging from very low to very high. The results revealed that in all models, high soil quality index was found in the south-western part of the study area while low soil quality areas were located mostly in the central part and north-western part of the study area with moderate (class III) as the most dominant class in the study area (Fig 8). The values of RMSE, MSE, bias, R^2 and CCC in Table 4 show that soil quality models had different abilities to predict SQI at unsampled locations in the study area. TDS_NL (R^2 = 0.375) had high precision with CCC (0.395), implying good agreement with the 45º line. MDS_NL had lowest RMSE of 0.0466, MSE of 0.00217 and bias of -0.00614indicating that MDS_NL model predicted soil quality better than other models. MDS_L, TDS_L and TDS_NL models showed high tendency for either overestimation or underestimation (Figs. 9 a, b, c & d).

TABLE 4: Performance of Predictive Models in Predicting Soil Quality Index

Models	R^2	CCC	MSE	RMSE	Bias
MDS_L	0.080	0.183	0.00041	0.0644	-0.0144
MDS_NL	0.221	0.362	**0.0021**	**0.0466**	**-0.0061**
TDS_L	**0.342**	**0.369**	0.0049	0.0699	-0.0229
TDS_NL	0.376	0.395	0.0034	0.0581	-0.0126

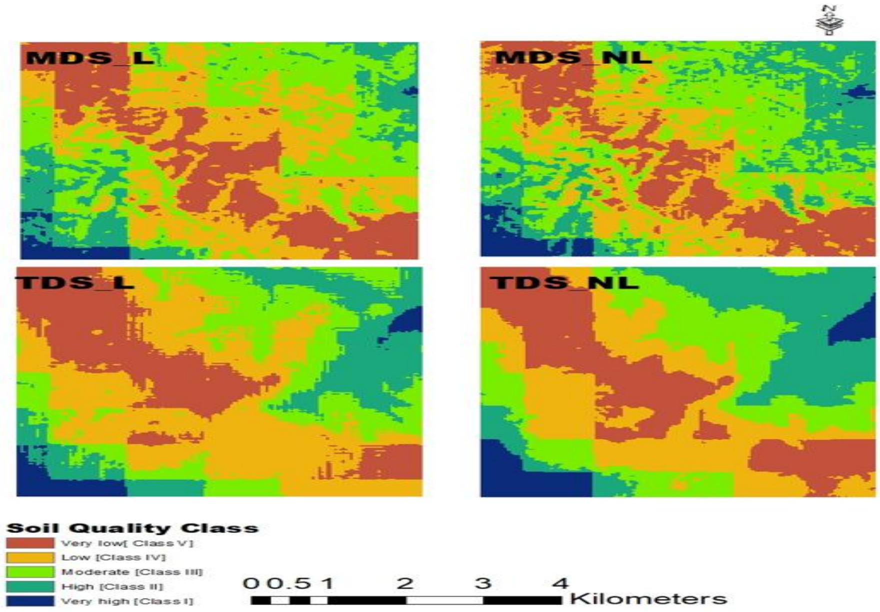


Fig. 8: Spatial distribution of soil quality index class predicted by MDS_L, MDS_NL, TDS_L, TDS_

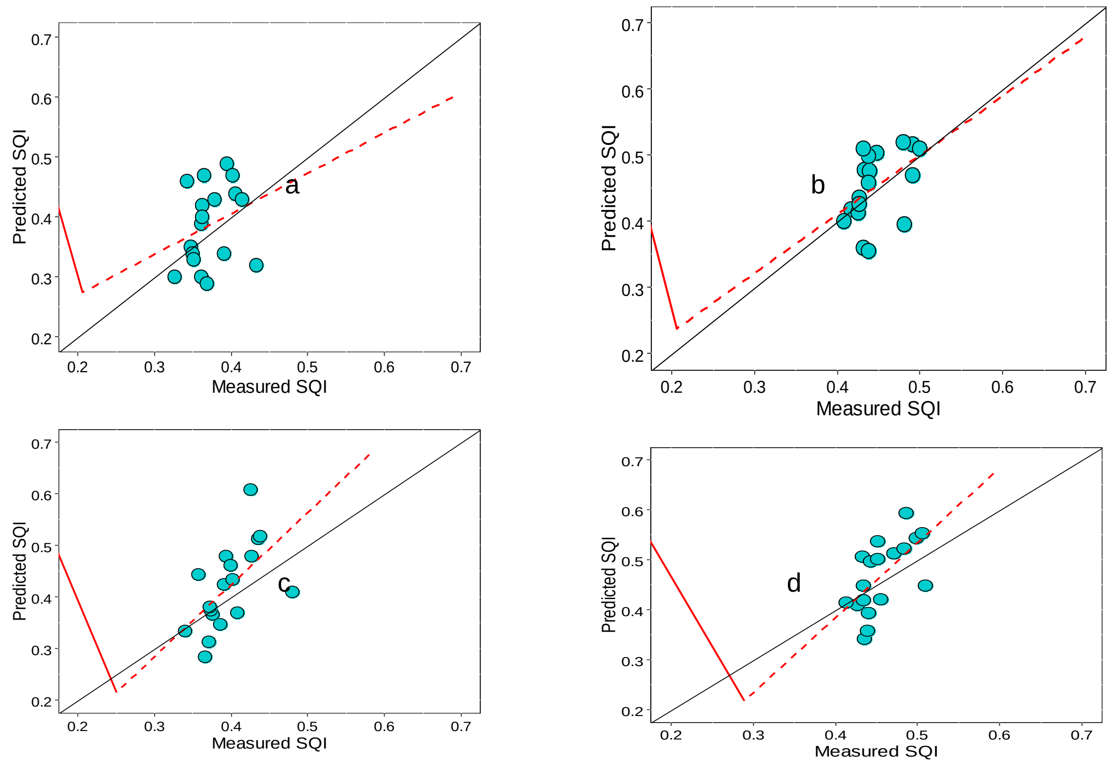


Fig. 9: Measured and predicted values of soil quality index using (a) TDS_NL, (b) MDS_NL, (c) TDS_L and (d) TDS_NL models

SUMMARY

This is the first study that has been conducted with the aim of predicting soil quality index or class in Obudu cattle ranch, Nigeria. Soil-environmental factors including NDVI, SAVI, clay, mean rainfall, temperature, elevation and slope were used as covariates while the models used were evaluated using bias, R^2, correlation CCC, MSE and RMSE. This study found that organic matter has the highest effect on the prediction of soil quality index in mountainous areas while slope, NDVI and mean rainfall were among the soil environmental factors with the least effect. Moderate soil quality class (class III) was the dominant class in greater part of the study area using MDS_NL model. MDS_NL model predicted soil quality better than other models. All the models showed almost similar spatial distribution of soil quality, with high soil quality found in the south-western part while low soil quality areas were located mostly in the central part and north-western. Organic manures/crop residues application, liming and fallow cropping system should be implemented as measures to increase soil quality from its current moderate state to high and very high for optimum productivity of the soils.

REFERENCES

Akpan, S., 2005. "Vegetation and Crop Farming Systems of Obanliku, Cross River State." Journal of Agriculture and Ecology in Tropical Regions 32, 15-25.

Andrews, S. S. and Carroll CR., 2001. Designing a soil quality assessment tool for sustainable agroecosystem management. Ecological Application 11, 1573–1585.

Andrews, S. S., Karlen, D. L., and Cambardella, C. A., 2002. The soil management assessment framework: A quantitative soil quality evaluation method. Soil Science Society of America Journal, 66(3), 1004-1019.

Asensio V, Guala SD, Vega FL, Covelo EF., 2013. A soil quality index for reclaimed mine soils. Environ Toxicol Chem, 32:2240–2248

Chakraborty, A., Bhattacharjee, P., Talukdar, S., and Bera, S., 2019. Soil quality assessment using machine learning algorithms: A review. Geoderma, 337, 536-551.

Choudhury BU, Mandal S., 2021. Indexing soil properties through constructing minimum datasets for soil quality assessment of surface and profile soils of Intermontane valley, Barak, North East India. Ecological Indicators, 123: 1-17.

Debi SR, Bhattacharjee S, Aka TD, Paul SC, Roy MC, Salam MS, Islam MS, Azady AR., 2019. Soil quality of cultivated land in urban and rural area on the basis of both minimum data set and expert opinion. Int J Hum Capital Urban Manag, 4(4):247–258

Falahatkar S, Hosseini MS, Ayoubi S, Mahiny AS., 2016. Predicting soil organic carbon density using auxiliary environmental variables in northern Iran. Archives of Agronomy and Soil Science, 62: 375-393.

Fathizad H, Ali M, Ardakani H, Heung B, Sodaiezadeh H, Rahmani A, Fathabadi A, Scholten T, Taghizadeh-Mehrjardi, R., 2020. Spatio-temporal dynamic of soil quality in the central Iranian desert modeled with machine learning and digital soil assessment techniques. Ecological Indicators 118:1-16.

Gee, W. G. and Or. D., 2002. Particle-Size Analysis. 255–293. In: Dane, J., and Topp, G.C. (eds.). Methods of Soil Analysis. Book Series: 5. Part 4. Soil Science Society of America. USA.

Hengl T, Leenaars JGB, Shepherd KD, Walsh MG, Heuvelink GBM, Mamo T, Tilahun H, Berkhout E, Cooper M, Fegraus E, Wheeler I, Kwabena N., 2017. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutrient Cycling Agroecosystem, 109:77–102. https:// doi. org/ 10. 1007/ s10705- 017- 9870-x

John, K., Agyeman, P.C., Kebonye, N.M., Isong, I.A., Ayito, E.O., Ofem, K.I. and Qin, C., 2021. Hybridization of cokriging and Gaussian process regression modelling techniques in mapping soil sulphur. Catena 206:105534

Li, Q., Li, Y., Lu, H., Huang, G., and Zhang, J., 2018. Irrigation scheduling using a modified support vector regression machine algorithm based on the FAO-56 model. Agricultural Water Management, 196, 71-79.

Mukherjee, A. and Lal, R., 2014. Comparison of soil quality index using three methods. PLoS ONE 9(8):e105981

Nabiollahi, K., Taghizadeh-Mehrjardi, R. and Eskandari, S., 2018. Assessing and monitoring the soil quality of forested and agricultural areas using soil quality indices and digital soil-mapping in a semi-arid environment. Arch Agron Soil Sci, 64:696–707

Nelson, D. W. and Sommers, L. E., 1996. Total carbon, organic carbon and organic matter. In D.L. Sparks (Eds.). Methods of soil analysis. Part 3. Chemical methods. SSSA Book Ser. 5. SSSA, Madison, WI. p. 961-1010.

Ogunwole, J.O., Obidike, E.O., Timm, L.C., Odunze, A.C., Gabriels, D.M., 2014. Assessment of spatial distribution of selected soil properties using geospatial statistical tools. Commun Soil Sci Plant Anal, 45:2182–2200.

Paul, G.C., Saha, S., and Ghosh, K. G., 2020. Assessing the soil quality of Bansloi river basin, eastern India using soil quality indices (SQIs) and Random Forest machine learning technique. Ecological Indicators, 118, 1-17.

Rezaee, L., Akbar, M.A., Naser, D. and Sepaskhah, A.R., 2020. Soil quality indices of paddy soils in Guilan province of northern Iran: spatial variability and their influential parameters. Ecol Indic, 117:106566

Shekhovtseva, O.G. and Mal'tseva, I.A., 2015. Physical, chemical and biological properties of soils in the city of Mariupol, Ukraine. Eurasian Soil Science 48:1393–1400

Udo, E J., T.O. Ibia, J.O. Ogunwale, A.O. Ano and I. Esu, 2009. Manual of Soil, Plant and Water Analysis. Sibon Books Ltd. Lagos.

Upadhyay, S. and Raghubanshi, A.S., 2020. Determinants of soil carbon dynamics in urban ecosystems. In Urban Ecology. 299-314.

Elsevier. Vasu, D., Singh, S.K., Ray, S.K., Duraisami, V.P., Tiwary, P., Chandran, P., Nimkar, A.M. and Anantwar SG., 2016. Soil quality index (SQI) as a tool to evaluate crop productivity in semi-arid Deccan Plateau, India. Geoderma, 282:70–79

Yemefack M, Jetten VG, Rossiter DG., 2006. Developing a minimum data set for characterizing soil dynamics under shifting cultivation systems. Soil Tillage Research, 86: 84–98.

Zeraatpisheh, M., Ayoubi, S., Sulieman, M. and Rodrigo-Comino, J., 2019. Determining the spatial distribution of soil properties using the environmental covariates and multivariate statistical analysis: a case study in semi-arid regions of Iran. Journal of Arid Land, 11(4): 551–566.

Zeraatpisheh, M., Bakhshandeh, E., Hosseini, M. and Alavi, S.M., 2020. Assessing the Effects of deforestation and intensive agriculture on the soil quality through digital soil mapping. Geoderma, 363:114139