JJEES

JJEES (2022) 13 (1): 26-36
ISSN 1995-6681

Jordan Journal of Earth and Environmental Sciences

Estimation of Surface Soil Particles Using Remote Sensing-based Data in Al-Ghab Plain, Syria

Alaa Khallouf [1,2*], Sameer Shamsham[2], Younes Idries[3]

General Commission for Scientific Agricultural Research (GCSAR), Damascus, Syria.
Department of Soil and Land Reclamation- Faculty of Agriculture- Al-Baath University, Homs, Syria.
General Organization of Remote Sensing (GORS), Damascus, Syria.

Received 17 October 2020, Accepted 2 July 2021

Abstract

This study focuses on digital soil mapping and its role to predict topsoil texture particle distribution. The study was carried out by collecting 146 surface soil samples (0-30 cm) from Al-Ghab plain, Hamah governorate, and 40 environmental covariates were derived from Landsat 8 OLI, and Digital Elevation Model (DEM). The surface soil particle models were obtained by Multiple Linear Regression (MLR) via R programming software, SAGA-GIS, and ArcGIS. Statistical analysis and results demonstrate no outlier in the data. Normal distribution for all covariates was examined, where skewness and kurtosis values were highly varied and ranged from 2.398 to 4.1090, and -1.772 to 20.1603 respectively, indicating that some predictors are highly skewed. Therefore, Tukey's ladder of power transformation was applied for all environmental predictors to be normally-distributed or close to normal distribution. All normal or close to normal data of environmental variables were used in MLR prediction models and cross-validation. Three MLR methods including forwarding selection, backward elimination, and stepwise selection were implemented for predictors selection. Relative importance also was calculated to estimate the contribution of each regressor in ML. The backward elimination method for the three soil texture particles sand, silt, and clay has the highest R^2 values (31.3, 28.6, 45.5%) respectively, and with the lowest values of mean absolute error (MAE) Root Mean stander error (RMSE), so it is the best prediction method for soil particles the study area.

Keywords: Backward elimination, Cross-validation, Forward selection, Multiple Linear Regression (MLR), Stepwise selection, Surface soil texture.

1. Introduction

Sustainable development as an overall goal usually deals with several issues at the regional and global levels involving land degradation, high food demand, water scarce, and climate change. Addressing these problems needs continuous and up-to-date soil information (McBratney, 2014; Zhang et al, 2013). Soil texture affects tightly many physical and chemical soil characteristics such as water and ventilation features, soil porosity, its fertility status, drainage characteristics, etc. (Makabe et al., 2009; Akpa et al., 2014). Soil texture and soil organic matter content are regularly used as the most significant parameters of hydrological characteristics and soil Cation Exchange Capacity (CEC) prediction (Wösten et al., 2001; Seybold et al., 2005). Besides, soil texture is a strong indicator to anticipate the soil erosion hazard conditions and eroded soil amount (Le Bissonnais 1996; Warrington et al., 2009). Therefore, accurate and reliable soil texture maps are crucial needs for hydrological and environmental modeling, land management practices, and environment protection, particularly when budget, labor, and time are limited (Zhao et al., 2009). The traditional survey process has been subjected to many limitations; firstly, the changes in environmental conditions are not easy to observe, mainly when the processing of many variables works concurrently; secondly, the entire steps must be repeated for any updating process, that leads to ineffective soil survey updates (Zhu et al., 2010). Remote sensing data was used for the first time to predict soil properties in the 1960s (McBratney et al., 2003). Digital soil mapping (DSM), which means the spatial prediction of soil properties in unobserved locations using statistical assumptions, have increasingly been applied recently since their early development at the beginning of the 19th century (Webster and Oliver, 2007). DSM relies on defining the relationship between soil attributes with some terrain characteristics which are usually derived from the Digital Elevation Model (DEM) such as Slope, Aspect, Curvature, Length of slope (LS), and others, in addition to derived remote sensing indices from satellite images. This approach was significantly used during the past three decades in predicting soil properties and has been achieved required rapidness and accuracy (Chen, et al., 2008; D'Acqui et al., 2010). Many studies examined the relation between soil reflection of the visible and near-infrared bands (400-2500 nm) and surface soil texture (Al-Abbas et al., 1972; Suliman and Post, 1988; Zhang et al., 1992; Sullivan et al., 2005). Liao et al. (2013) applied remote sensing data of Landsat7 ETM - digital number (DN) of six bands (Bands 1–5 and Band 7)- as covariates for estimation of surface soil particles distribution in the city of Pingdu lands, Shandong Province, China. Correlation analysis revealed that surface soil particles were

* Corresponding author e-mail: alaakhallouf@gmail.com

highly correlated with Landsat ETM data. Multiple Linear Regression (MLR) is one of the major statistical models that is commonly used to describe the relationship between environmental covariates and soil particles (Lagacherie et al., 2008; Mitran et al., 2018). Al-Ghab Plain is a highly valuable agricultural area in Syria, where various strategic crops such as wheat, sugar beets, cotton, tobacco, sesame, sunflower, soybeans, olives, and vegetables are grown. Therefore, land use planning procedures require beneficial and accurate soil data. Thus, the main objectives of this study are: (i) setting a digital map of the spatial distribution of sand, silt and, clay in the Al-Ghab Plain area and, (ii) fitting the best model for estimating the distribution of soil particles.

2. Material and Methods

2.1 Site description:

Al-Ghab Plain is located in the northwestern of Hamah governorate in the middle of Syria (Figure 1) and it extends between 36º 19` 12``36º 23` 45`` E and 35º 44` 40 ``- 35º 16` 34``N. It covers about 477 km^2 with altitude ranges between 146-176 m above sea level (asl). Al-Ghab plain is described as one of the Pliocene lakes. At the end of the Pliocene period, vigorous movement occurring along the already existing faults formed the actual landscape, although the movement continued during the Quaternary period. According to the existing geological map, a major part of the plain was covered by peat on the subsoil of gray plastic marl or white lime ooze. Gravels, sand, marls, and coquina are found in small parts in old riverbeds or lakeshore. White limestone and dolomite formations occur in the eastern border of the plain with basalts on the northeastern boundary (FAO, 1972). The Orontes river passes from south to north which is forming many terraces and irrigation channels. The climate of Al-Ghab plain belongs to the Mediterranean climate which is described as cold and rainy winters, hot and dry summers with two transitional seasons (spring and fall) (Moses, 1978). The soil climate is classified based on USDA as Xeric for soil moisture and Thermic for soil temperature regime (Ilaiwi, 1985).

2.2 Soil sampling and laboratory analyzing:

The study was conducted by collecting 146 surface soil samples (0-30cm) during 1-23/4/2020 (Figure 1), and their geographic locations were recorded by using the Global Positioning System (Garmin GPS, accuracy ± 5 m). The collected soil samples were air-dried, ground, and sieved through a 2mm sieve. The hydrometer method was used to estimate soil particle distribution (sand, silt, and clay as percentage) (Gee and Bauder, 1986).

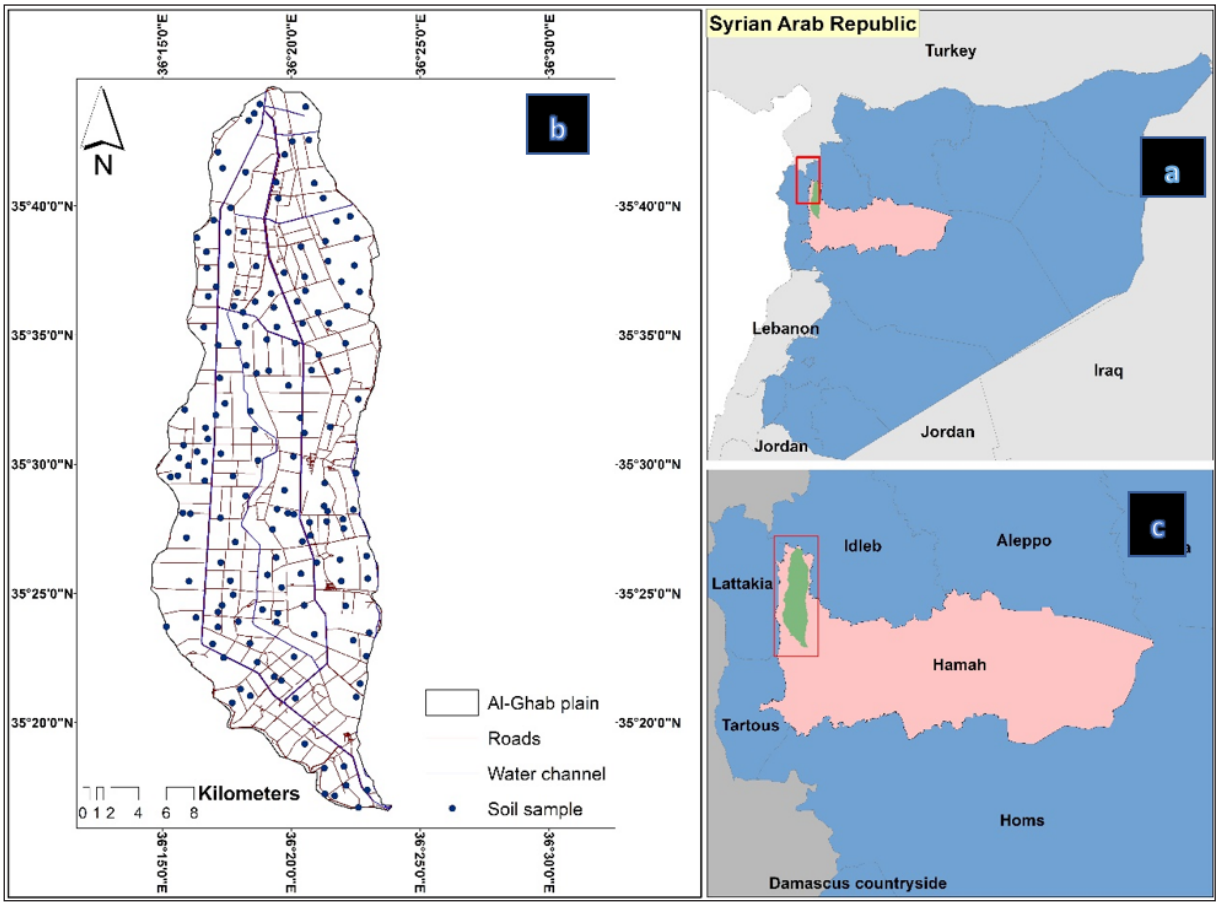

Figure 1. (a) Location of studied area to Syria, (b) soil samples distribution, (c) Location of the studied area to Hamah governorate

2.3. Remote sensing data and derived covariates:

To predict the relationship between surface soil distribution and the environmental predictors (variables) mainly topographic factors, hydrology, and vegetation cover indicators, in total 40 parameters were derived from Landsat8 OLI and Digital Elevation Model (DEM) but only 37 parameters were convenient for predicting soil texture (Table.1). Firstly, the satellite image was downloaded from https://earthexplorer.usgs.gov/, its date 18/4/2020 and re-projected, then bands reflectance was obtained.

The remote sensing imagery of Landsat8 OLI involves reflectance of bands (2to7), or some vegetation indices such as NDVI, SAVI, OSAVI, MSAVI, GS, etc…. Also, some transformations including principle component analysis PCA as PCA1 for reflectance bands 2,3,4; PCA2 for bands 5,6,7; and environmentally considered bands, finally tasseled cap transformation (brightness, greenness, and wetness). On the other hand, a 30 m-resolution DEM (Ministry of Economy, Trade, and Industry of Japan and the United States National Aeronautics and Space Administration, 2009), was sink-filled and SAGA-GIS was used to derive 13 environmental parameters (Table.1) using the basic terrain analysis, terrain-hydrology analysis, and morphometry analysis (Olaya, 2006).

Table 1. input environmental variables and derived covariates

Parameters	Description	Formula	Reference
Landsat8 band2	blue band (B)		
Landsat8 band3	green band (G)		
Landsat8 band4	red band (R)		
Landsat8 band5	near infra-red band (NIR)		
Landsat8 band6	short wave infra-red1 band (SWIR1)		
Landsat8 band7	short wave infra-red2 band (SWIR2)		
PCA1	principle component analysis	Derived from of bands 2,3,4	Frazier and Cheng, 1989
PCA2	principle component analysis	Derived from of bands 5,6,7	Frazier and Cheng, 1989
PCA3	principle component analysis	Derived from of bands 2,3,4,5,6,7	Frazier and Cheng, 1989
SR	Simple ratio	NIR/R	Malthus et al,1993
DVI	Difference Vegetation Index	NIR-R	Foody et al., 2001
OSAVI	Optimized Soil Adjusted Vegetation Index	(NIR -R) / (NIR + R + 0.16)	Nikolakopoulos, 2003
SAVI	Soil Adjusted Vegetation Index	[NIR-R)/(NIR+R+L)] × (1+L)	Pettorelli et al., 2005
MSAVI	Modified Soil Adjusted Vegetation Index	[NIR-R)/(NIR+R+L)] × (1+L)	Bannari et al 1995
TVI	Transformed Vegetation Index	(SWIR1-R)/(SWIR1+R)	Bannari et al 1995
MNDVI	Normalized Difference NIR/MIR Modified Normalized Difference Vegetation Index	(NIR-MIR)/(NIR+MIR)	Jürgens,1997
GVI	Green Vegetation Index	-0.29 (G) -0.56(R)+0.6(IR)+0.49(IR)	Leblon, 1993
OI	Simple Ratio Red/Blue Iron Oxide	R/B	Hewson et al.,2001
SRRed/NIR	Simple Ratio Red/NIR Ratio Vegetation-Index	R/NIR	Bannari et al 1995
SCRI	Surface clay index	SWIR2/SWRI1	Bannari et al 1995
SARVI2	Soil and Atmospherically Resistant Vegetation Index2	2.5(NIR-R)/(1+NIR+6R-7.5B)	Heute et al,1997
Wetness	The humid band derived from tasseled cap		Bahtti et al., 1991
Brightness	Brightness derived from tasseled cap		Bahtti et al., 1991
Greenness	Greenness band derived from tasseled cap		Bahtti et al., 1991
ASPECT	facing direction		Castro-Franco et al,2018
GENERALCUR	General curvature		Ließ et al,2012
LONCUR	Long curvature		Ließ et al,2012
LSFACTOR	Length of slope steepness		Castro-Franco et al,2018
PROFILECUR	Profile curvature		Ließ et al,2012
PLANCUR	Plan curvature		Ließ et al,2012
MRVBF	Multi-resolution Valley Bottom Flatness Index		Dobarco et al,2016
MRRTF	Multi-resolution of ridge top flatness index		Dobarco et al,2016
FLOWLINE	Flow line		Zevenbergen et al.,1987
MAXCUR	Maximum curvature		Ließ et al,2012
TWI	Topographic Wetness Index		Dobarco et al.,2016
SLOPE	Slope		Castro-Franco et al, 2018
DEM	Digital Elevation Model		Pinheiro et al, 2018

2.4. Descriptive statistics of environmental covariates and data transformation:

R programming software was used to calculate the criteria of descriptive statistics for both soil particles and environmental predictors. It includes the mean, minimum, maximum, standard deviation (SD), coefficient of variation (CV), skewness, and kurtosis. The CV% is usually used to explain the variance of each soil characteristic and predictors. The CV% was classified into: high variance (% CV> 35%), medium variance (15< %CV <35), low variance (% CV <15) (Wilding, 1985). The normal distribution test was applied for all variables by plotting histogram test companion package), skewness, and kurtosis were determined by using (moments package). For non-normal distributed variables, Tukey`s ladder of power transformations was applied (Mondejar and Tongco, 2019) by using the equation:

$$y=\begin{cases} x^{\lambda} \text{ if } \lambda>0 \\ \log x \text{ if } \lambda=0 \\ -(x^{\lambda}) \text{ if } \lambda<0 \end{cases} \quad \text{..... (1)}$$

Where λ is the power used to convert the factor closer to the normal distribution (Mangiafico, 2016; Scott, 2018). in this study λ values were (-2, -1, -1/2, 0, 1/2, 1, 2) (Tukey, 1977)

The transformation was performed for non-normal distributed variables for the Value of skewness to be close to zero and Kurtosis ranges between 1 to3.

2.5. Multiple Linear Regression predictive model:

Multiple Linear Regression predictive (MLR) has been widely utilized for estimating soil characteristics based on environmental covariates, where the correlation between predictors was examined. MLR equation is defined as:

$$Y`=\beta_0+\sum_{j=1}^{p} Xj\beta j \quad \text{..... (2)}$$

where: Y`: dependent variable, β_0: constant, Xj the matrix of the input independent covariates, and is the unknown coefficients for the involved predictors (McDonald, 2014).

Three methods of MLR involving Backward elimination, Forward Selection, and stepwise regression were applied in SAGA-GIS and a 95% confidence level was used for fitting the estimation model. All selected methods of MLR were subjected to Kolmogorov-Smirnov (Lilliefors) test.

2.6. Model fitting and validation:

The determination coefficient (R^2) is used as a function of goodness of fitting linear models for cross-validation and it is expressed as:

$$R^2=\frac{\sum_{i=1}^{n}(\hat{y}i-y`)^2}{\sum_{i=1}^{n}(yi-y`)^2} \quad \text{..... (3)}$$

The predictive model's performance was examined via cross-validation technique [5-fold method in R (Carte Package)] (Kuhn and Johnson, 2013), since the validation dataset was unavailable, in addition to the limited soil samples number (Ballabio et al., 2016). The accuracy of models was determined by calculating the coefficient of determination (R^2), the mean absolute error (MAE), and the root mean squared (RMSE) as following equations:

$$MAE=\frac{\sum|\hat{y}i-yi|}{n} \quad \text{..... (4)}$$

$$RMSE=\sqrt{\frac{\sum_{n=i}^{1}(yi-\hat{y}i)^2}{n}} \quad \text{..... (5)}$$

where $\hat{y}_i$: is the predicted value, y_i: is the observed value, y`: is the mean of observed values, and n:is the number of observed points.

2.7. Relative importance of environmental variables:

Relative importance is a topic that has seen a lot of interest in recent years, particularly in applied work. Relative importance (RI) is described as "the quantification of individual regressor's part to a multiple regression model" (Gromping, 2006). RI is typically implemented as part of the model-building procedure, e.g., forward variable selection or backward elimination, …etc. It is an alternative to the multiple regression techniques and it addresses the multi-collinearity problems and also helps to calculate the importance rank of variables. It helps to answer " Which variable is the most important and rank variables based on their contribution to R-Square". RI was obtained using R (Relaimpo package) and its values are calculated as percentages of 100% (Mondejar and Tongco, 2019).

2.8. Map generating:

Soil texture maps were generated based on the best MLR predicting model by using R (Automap package) and saved all maps as a raster (FAO,2018).

3. Results and discussion:

3.1. Descriptive statistics of environmental variables:

The former step in this study was examining data including descriptive statistics and transformation. Table.2 shows the summary statistics for untransformed data. The results showed high CV% and it ranges between 1.41 to 284.08%, with medium CV% for soil particles sand, silt, clay (21.5, 30.6, 18.3%) respectively, while CV% values were high for OSAVI, slope, LS factor, and Aspect (284.08, 245.46,104.29, and 100.63%) respectively. On the other hand, skewness, and kurtosis values range from 2.39 to 4.10; -1.77, and -20.16 respectively, these results imply that some predictors are highly skewed, and this is not appropriate to the normal assumption. It is obvious in Table.2 that the three soil particles were normally distributed, since skewness values for soil particles as a flowing: -0.154, 0.504, 0.762, and kurtosis values were -0.69, 0.768, 0.593 respectively, whereas results of some predictors were non-normally distributed based on histogram test and skewness and kurtosis values. Therefore, Tukey`s ladder of power transformation was applied to transform the environmental predictors into normally or close to a normal distribution as shown in Table. 3. The histogram, skewness, and kurtosis tests were applied for transformed data of non-normally distributed predictors.

Table 2. Descriptive statistics of environmental variables

Parameters	Minimum	Maximum	Mean	Standard Deviation	CV%	Skewness	Kurtosis
Sand %	22.0000	57.0000	38.874	7.9929	21.5	-0.154	-0.69
Silt%	10.0000	52.0000	27.361	10.1320	30.6	0.504	-0.77
Clay%	18.0000	62.0000	33.765	8.9702	18.3	0.762	0.59
Landsat8 band2	0.0033	0.0061	0.004	0.0005	2.24	0.995	2.21
Landsat8 band3	0.0166	0.0279	0.020	0.0020	4.47	0.703	1.12
Landsat8 band4	0.0325	0.0698	0.044	0.0069	8.31	0.791	1.20
Landsat8 band5	0.0317	0.1267	0.057	0.0156	12.49	0.966	2.35
Landsat8 band6	0.0350	0.1382	0.072	0.0222	14.90	0.098	-0.54
Landsat8 band7	0.0288	0.1061	0.061	0.0200	14.14	0.151	-0.97
PCA1	0.0282	0.0484	0.035	0.0036	6.00	0.736	1.12
PCA2	0.0567	0.2047	0.110	0.0319	17.86	0.120	-0.52
PCA3	0.0739	0.2277	0.129	0.0328	18.11	0.149	-0.43
SR	0.2491	0.2725	0.260	0.0037	6.08	-0.022	0.36
DVI	1314.44	1324.294	1319.367	2.5066	158.32	-0.056	-1.09
OSAVI	1244.56	1278.028	1261.294	8.0699	284.08	0.003	-0.91
SAVI	0.0000	0.5403	0.327	0.0904	30.07	-0.620	0.74
MSAVI	0.1058	0.6499	0.378	0.1059	32.54	-0.389	0.09
TVI	-0.0071	0.3780	0.218	0.1052	32.43	-0.622	-0.66
MNDVI	0.0000	0.3050	0.042	0.0568	23.83	1.983	5.25
GVI	0.0067	0.0948	0.040	0.0200	14.14	0.029	-0.77
NDVI	0.0000	0.3682	0.112	0.0698	26.42	0.722	1.24
MVI	0.0000	0.3502	0.141	0.1058	32.53	0.072	-1.18
ASPECT	0.0000	360.0000	118.055	118.4584	100.63	0.620	-0.87
GENERALCUR	0.0000	0.0089	0.001	0.0018	4.24	2.032	3.71
LONCUR	0	0.0066	0.001	0.0011	3.32	1.947	4.88
LSFACTOR	0	25.8273	10.95	11.502	104.29	0.229	-1.77
PROFILECUR	0	0.0034	0	0.0006	2.45	2.069	5.78
PLANCUR	0	0.1333	0.007	0.0204	14.28	4.109	20.16
MRVBF	5.5045	79.6704	57.786	20.4601	452.3	-0.834	-0.39
MRRTF	0.0872	8.9686	6.697	2.6469	162.69	-0.983	-0.19
FLOWLINE	0.6381	1.0385	0.968	0.0846	29.09	-1.768	3.13
MAXCUR	0.8736	1.1477	1.115	0.0818	28.60	-2.398	3.95
TWI	0.301	1.0404	0.523	0.1816	42.61	0.094	-0.84
SLOPE	0	0.9989	0.342	0.2706	245.46	-0.138	-1.2
DEM	0.3094	0.3152	0.311	0.001	3.16	0.711	1.14
OI	0.3179	0.8751	0.492	0.101	31.78	0.806	1.19
SRRed/NIR	0.4513	1.0142	0.655	0.1527	39.08	0.874	-0.29
VARIgreen	0	0.2188	0.067	0.0415	20.37	0.695	1.2
SCRI	0	0.4669	0.136	0.1104	33.23	0.898	0.38
SARVI2	0.285	0.2986	0.29	0.0018	4.24	1.277	5.84
Wetness	-0.06426	0.008739	-0.02631	0.018262	13.51	-0.075	-1.007
Brightness	0.087	0.2293	0.133	0.0275	16.58	0.313	0.18
Greenness	0	0.0006	0	0.0002	1.41	-0.352	-0.89

The results are shown in Table.3 reveals that all transformed data of predictors are normal or near to normal distribution, and values of skewness and kurtosis range -0.834 to 0.995 and -1.204 to 2.206 respectively. Consequently, all environmental predictors were used in MLR modeling with normal or close to normal distribution.

Table 3. The transformation data for some non-normally distributed using Tukey's method

soil variables	Trans formation-mode	Min	Max	Mean	Standard Deviation	CV%	Skew	Kurtosis
band3	Log(x)	0.017	0.028	0.02	0.08	3.65	0.703	1.115
band2	Log(x)	0.003	0.006	0	0.06	4.72	0.995	2.206
SR	X^2	0.249	0.273	0.26	0.07	6.23	-0.022	0.361
PCA1	Log	0.028	0.048	0.03	0.07	5.24	0.736	1.115
SAVI	Log(x)	0	0.54	0.33	0.09	12.25	-0.62	0.744
General curvature	Log(x)	-2.653	-2.051	-2.5	0.2	4.65	0.69	-1.036
Profile-curvature	-1/x	-3.654	-2.465	-2.96	0.16	12.58	-0.473	2.125
Plan curvature	Log(x)	-1.722	-0.875	-1.37	0.23	15.26	0.809	0.785
MRVBF	Log(x)	5.504	79.67	57.79	20.46	48.62	-0.834	-0.391
Max- curvature	-1/x	0.273	0.332	0.33	0.02	25.62	-0.407	2.01
TWI	Log(x)	0.301	1.04	0.52	0.18	3.69	0.094	-0.838
SLOPE	Log(x)	0	0.999	0.34	0.27	17.82	-0.138	-1.204
DEM	Log(x)	0.309	0.315	0.31	0.02	6.36	0.711	1.142
Greenness	Log(x)	0	0.001	0	0	10.28	-0.352	-0.892

3.2 Multiple linear regression modeling:

Table.3 shows the MLR results for the three methods that were chosen namely forward selection, backward elimination, and stepwise selection. All selected methods of MLR were subjected to Kolmogorov-Smirnov (Lilliefors) test. Homoscedasticity and normality test of residuals were determined to confirm the homogeneity of variance assumption (Heil and Schmidhalter, 2017), where the Lilliefors value is more than 0.05 that indicates enhancing in normality and stability of variance for both independent and dependent variables (Tsai et al., 2017), minimizing errors, and unbiased estimation (Zhang et al., 2013).

3.3 Cross-validation of MLR selected methods

Cross-validation of selected methods was carried out and R^2, MAE, RMSE, Homoscedasticity, and normality tests for residuals including skewness, kurtosis, and Lilliefors were calculated for estimating the accuracy of the model's performance. Homoscedasticity and normality tests of residuals were determined to confirm the homogeneity of variance assumption (Heil and Schmidhalter, 2017). The results in Table.4 show that the backward elimination method has achieved the best model performance in comparison with forwarding selection and stepwise methods, where the MAE, RMSE values were less than other methods in all soil particles predicted models. However, the backward elimination method has the highest R^2 for sand, silt, and clay (31.3, 28.6, 45.5%) respectively. Figure 2 illustrates the scatter plots of residuals and predicted values of homoscedasticity test, it is clear that the homoscedasticity was not broken by any all three MLR methods for soil texture particles, since linearity theory was met having no identifiable pattern and the residuals randomly scattered, in addition to the almost of residuals were symmetrically clustering towards the center of the plot. Graphically, Figure 3 explains clearly the test of normality of residuals for three MLR methods where most of the points are placed almost directly either below or above closely to the reference line with no evidence of outlier's presence (Mondejar and Tongco, 2019). It is also obvious that the most of points of the sand regression models were located very close to the reference line. These results were statistically confirmed in Table.4, where the regression models of sand particles were the least in both skewness and kurtosis values. Table 4 also shows that the distribution of the residuals was statistically near to normal distribution based on the Lilliefors normality test.

Table 4. Cross-validation of the selected methods of MLR

Soil particle	Regression model	Predictors number	MAE	RMSE	R^2	Residuals normality		
						skewness	kurtosis	Lilliefors
sand	Forward	2	5.9	7.11	20.2	-0.215	-0.815	0.96
	Backward	10	5.4	6.6	31.3	-0.192	-0.49	0.57
	Stepwise	2	5.9	7.11	20.2	-0.215	-0.851	0.96
silt	Forward	2	7.3	8.85	23.1	0.479	-0.428	0.97
	Backward	6	7.05	8.52	28.6	0.322	-0.365	0.7
	Stepwise	2	7.3	8.85	23.1	0.479	-0.428	0.97
clay	Forward	3	5.71	7.21	34.8	0.344	-0.089	0.055
	Backward	14	5.79	7.01	45.5	0.283	-0.281	0.073
	Stepwise	3	5.71	7.21	34.8	0.344	-0.089	0.055

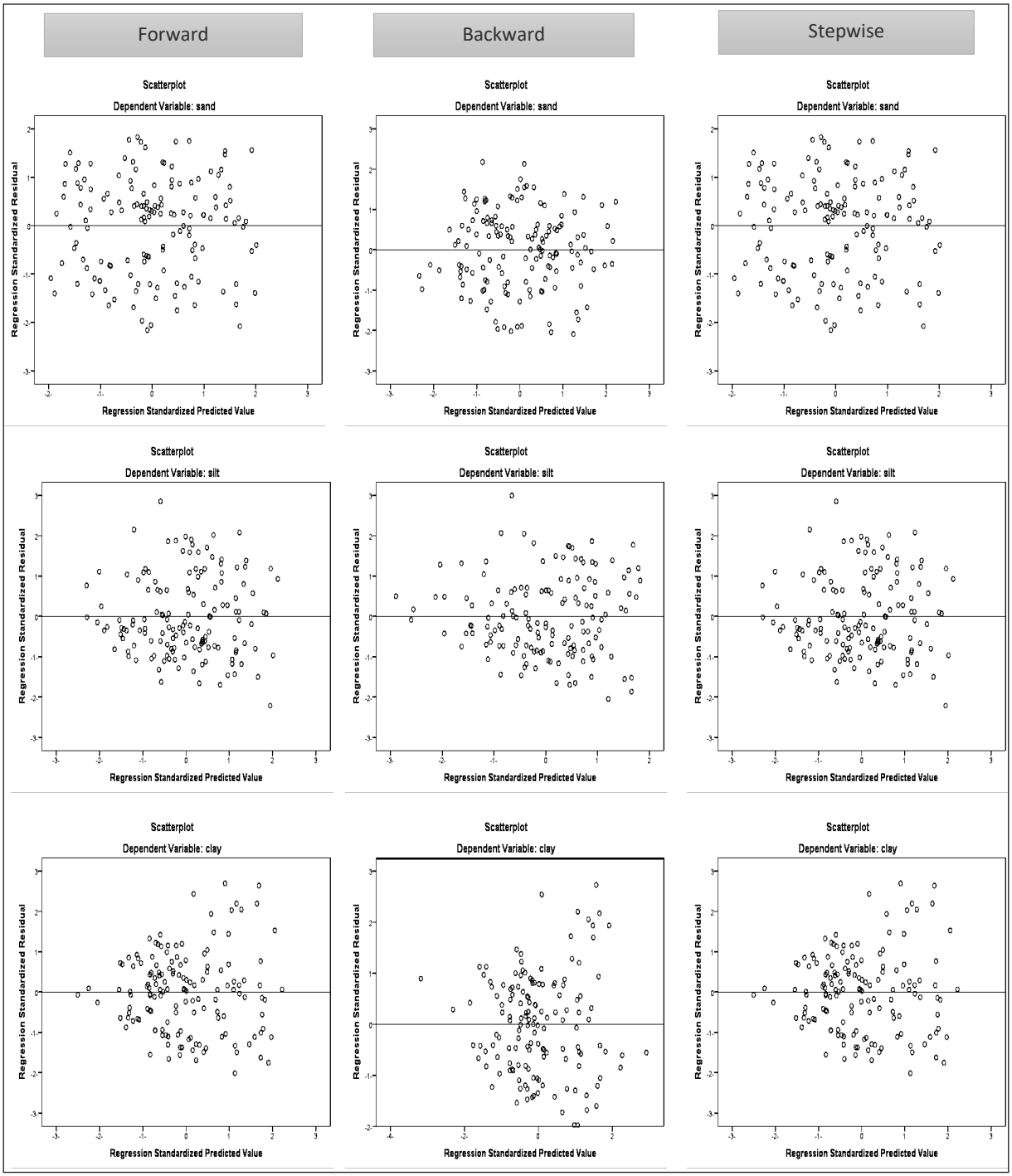

Figure 2. Residuals plots of MLR models based on selection methods

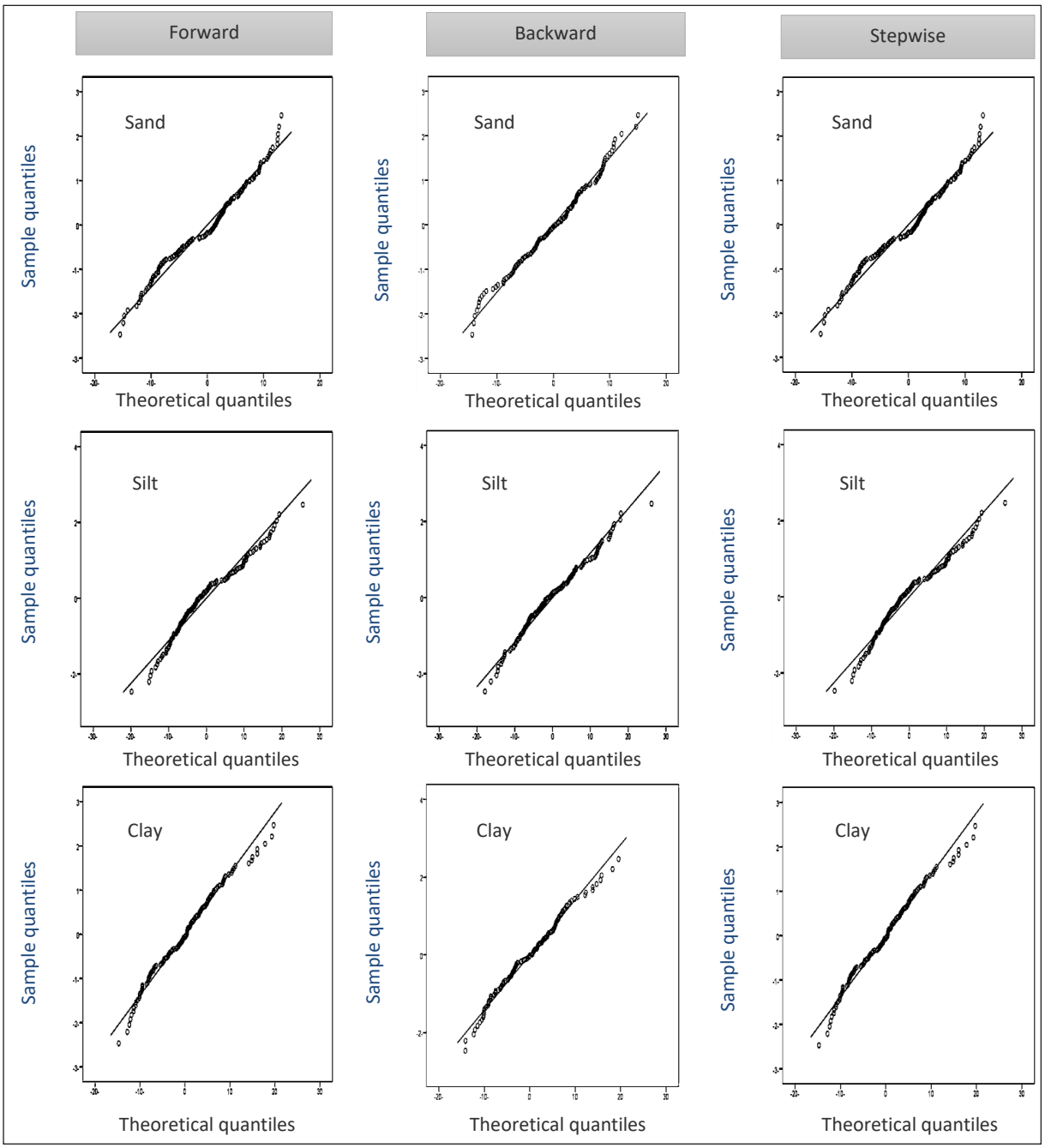

Figure 3. QQ plot of residuals for MLR selected methods

3.4 Relative importance of environmental covariates:

Figure 4 demonstrates that TWI, PCA1, SARVI, MAX-CUV, MRVBF, SCRI, OSAVI, MSAVI, Wetness, and NDVI were the most significant environmental covariates for predicting surface soil texture particles. In the sand predicting model, 10 predictors for the backward elimination method, since TWI, PCA1, SARVI, SRRED/NIR, and slope were the most important predictors with RI values 18.47, 15.4, 12.79, 11.32, and 9.6 % respectively. These results agreed with Mehrabi-Gohari et al. (2019) results since TWI was the most significant auxiliary variable to predict sand fraction distribution. Whereas Mondejar and Tongco (2019) found that stream proximity index (SPI) was the most crucial variable affected in sand fraction prediction. In Figure 4, the silt regression model is predicting silt values by using 6 predictors. Similarly, TWI, MAX-CURV, and MRVBF were the most critical variables for silt regression predicting models. In many studies, TWI and MrVBF were considered as a significant predictive variable in the prediction model of the silt fraction (Mehrabi-Gohari et al., 2019; Jafari et al., 2012), while Mosleh et al. (2016) reported steam power index, and plan curvature were the most auxiliary variables to silt prediction. For the clay predictive model, Figure 5 demonstrates that 14 environmental variables were considered in the predictive model since GSAVI, MSAVI, and DVI respectively were found the most significant variables of clay fraction prediction. These results agreed with Mehrabi-Gohari et al. (2019) results. Topography factors can influence soil physicochemical properties (soil depth, texture, and mineral contents), incoming solar radiation, precipitation and affect crop production. As an increased topography/elevation significantly increased soil

moisture, precipitation, soil significantly lower at the higher elevations, whereas bulk density, pH, and soil temperature were significantly lower at higher elevations (Abate and Kibret, 2016).

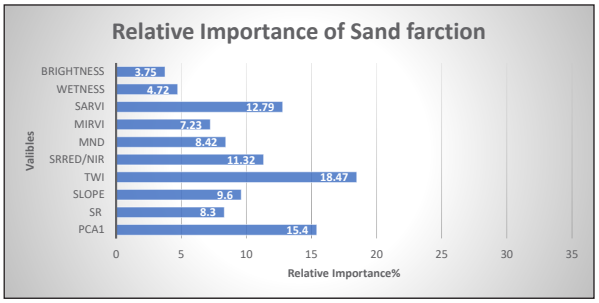

Figure 4. Relative importance of environmental covariates for soil texture prediction

3.5. Predictive soil texture maps:

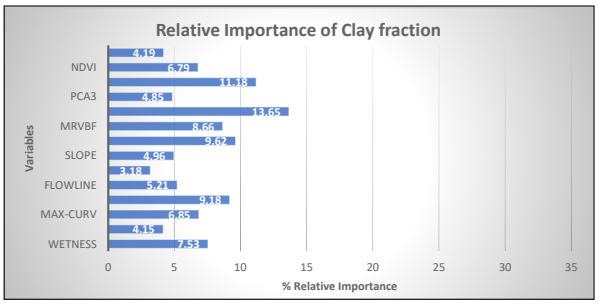

Figure 5. Relative importance of environmental covariates for soil texture prediction

Figures 6, 7, and 8 show the spatial distribution of sand, silt, and clay fraction based on the backward estimation predictive model. Sand content ranges from 20.7 to 59.6%, with a high concentration in the west northern part of Al-Ghab plain. Silt also ranges from 11.2 to 46.9% since most northern lands are high content of silt. The high content of clay (43.4-60.7%) is in the south and southeastern lands of Al-Ghab plain.

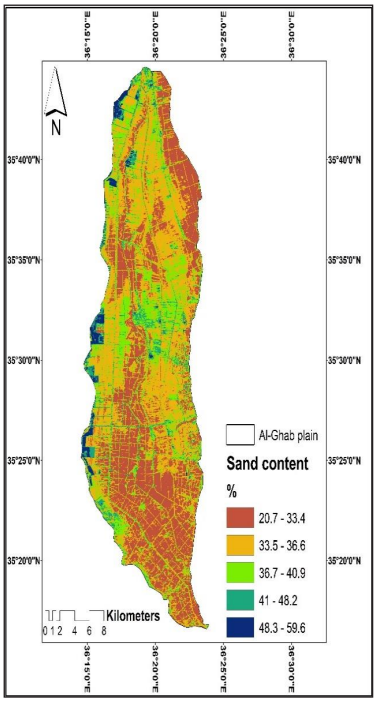

Figure 6. Sand spatial distribution

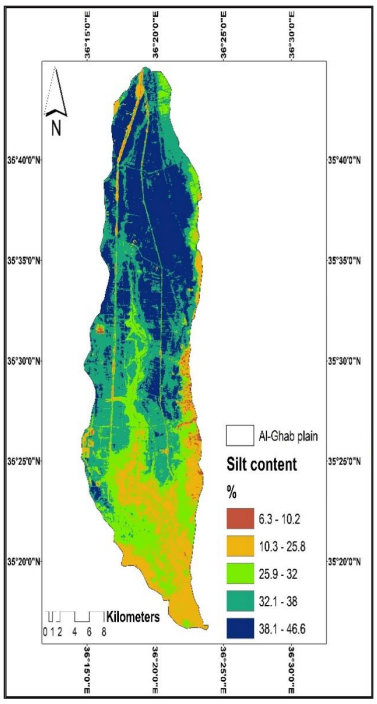

Figure 7. Silt spatial distribution

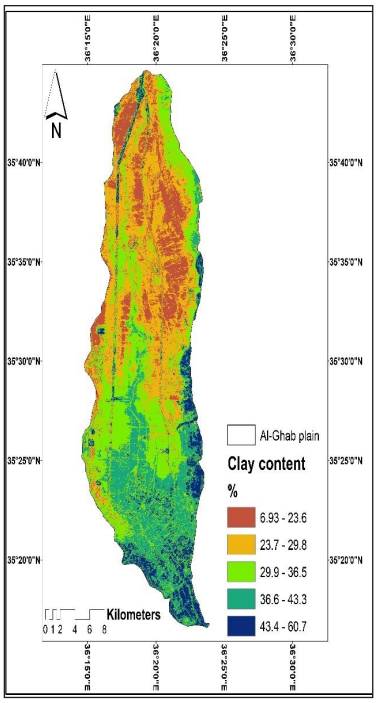

Figure 8. Clay spatial distribution

4. Conclusions:

Soil texture is one of the most crucial soil properties because it affects many soil functions involving the availability of water, nutrients, and many physical soil properties. In this study, 40 environmental RS-based covariates were derived from Landsat 8 OLI and DEM with a spatial resolution of 30 m. Multiple linear regression was used to determine the relationship between soil texture particles and environmental variables. The results revealed that backward elimination was the best predictive model for the three soil particles (sand, silt, and clay) with the highest R^2 values (31.3, 28.6, 45.5%) respectively, and with the lowest values of MAE and RMSE. On the other hand, TWI,

MRVBF, GSAVI, and MSAVI were the most significant predictors to predict topsoil texture fractions with a high relative importance percentage.

References

Abate, N, and Kibret, K (2016). Effects of Land Use, Soil Depth, and Topography on Soil Physicochemical Properties along the Toposequence at the Wadla Delanta Massif, Northcentral Highlands of Ethiopia. Environment and Pollution; 5(2): 57-71.

Akpa, S.I.C.; Odeh, I.O.A.; and Bishop, T.F.A. (2014). Digital mapping of soil particle-size fractions for Nigeria. Soil Sci. Soc. Am. J., 78: 1953–1966.

Al-Abbas, A.H., Swain, P.H., and Baumgardner M.F. (1972). Relating organic matter and clay content to the multispectral radiance of soils. Soil Sci, 114: 477–485.

Bahtti A.U., Mulla, D.J., and Frazier, B.E. (1991). Estimation of soil properties and wheat yields on complex eroded hills using geostatistics and thematic mapper images. Remote sens. Environment, 31: 181-191.

Ballabio, C., Panagos, P., and Monatanarella L. (2016). Mapping topsoil physical properties at a European scale using the LUCAS database. Geoderma, 261:110–23.

Bannari, A., Morin, D., Bonn, F., and. Huete, A. R. (1995). A review of vegetation indices.

Ben-Dor, E, Inbar, Y., and Chen, Y. (1997). The reflectance spectra of organic matter in the visible near-infrared and short-wave infrared region (400–2500 nm) during a controlled decomposition process. Remote Sens. Environ, 61: 1–15.

Ben-Dor, E, Irons, J., and Epema, G. (1999). Soil spectroscopy. In Manual of Remote Sensing, 3rd ed., Ed. Rencz A., John Wiley, and Sons, New York. 111–188 p.

Ben-Dor, E. and Banin, A. (1995). The near-infrared analysis is a rapid method to simultaneously evaluate several soil properties. Soil Sci. Soc. Am. J, 59: 364–372.

Ben-Dor, E., Goldshleger, N., Eshel, M., Mirablis, V., and Bason, U. (2008). Combined active and passive remote sensing methods for assessing soil salinity. In Remote Sensing of Soil Salinization: Impact on Land Management, Ed. Metternicht G, Zinck A, CRC Press, Boca Raton, FL, USA. 235–258.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne, M., and Reinsch, T.G. (2006). Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma, 132, 273–290.

Castro-Franco M, Domenech, M.B., Borda, M.R. and Costa, J.L. (2018). A spatial dataset of topsoil texture for the southern argentine pampas. Geoderma Reg, 12:18–27.

Chen, F., Kissel, D.E., West, L.T. and Adkins, W. (2000). Field-scale mapping of surface soil organic carbon using remotely sensed imagery. Soil Sci. Soc. Am. J 64: 746–753.

Chen, F., Kissel, D.E., West, L.T., Adkins, W., Rickman D., and Luvall, J.C. (2008). Mapping soil organic carbon concentration for multiple fields with image similarity analysis. Soil Sci. Soc. Am. J, 72: 186–193.

D'Acqui, L.P., Pucci, A., and Janik, L.J. (2010). Soil properties prediction of western Mediterranean islands with similar climatic environments employing mid-infrared diffuse reflectance spectroscopy. Eur. J. Soil Sci, 61: 865–876.

Dobarco, M.R., Orton, T.G., Arrouays, D., Lemercier, B., Paroissien, J.B. and Walter, C, et al. (2016). Prediction of soil texture using descriptive statistics and area-to-point kriging in region Centre (France). Geoderma Reg, 7:279–92.

FAO (Food and Agriculture Organization). (1972). report on the soil survey of Ghab Valley. Damascus, Syria.

FAO (Food and Agriculture Organization). (2018) Soil organic carbon mapping cookbook. Rome, Italy.

Foody, G.M., Cutler, M., Mcmorrow, J., Pelz, D., Tangki, H., Boyd, D.S. and Douglas, I. (2001). Mapping the biomass of Bornean tropical rain forest from remotely sensed data. J. Global Ecology Biogeography, 10: 379-387.

Frazier, B.E., and Cheng, Y. (1989). Remote sensing of soils in eastern Palouse region with LandSat thematic mapper, Remote sense. Environment, 28: 317-325.

Gawlik, B.M., Bo, F., Kettrup, A., and Muntau, H. (1999). Characterization of the second generation of European reference soils for sorption studies in the framework of chemical testing – Part I: Chemical composition and pedological properties. Sci. Total Environ, 229: 99–107.

Gee, G.W., and Bauder, J.W. (1986). Particle-size analysis. p. 383–411. In A. Klute (ed.) Methods of soil analysis. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.

Gromping, U. (2006). Relative importance for linear regression in R: the package relaimpo. J Stat Softw, 17:1–27.

Hassink, J. (1992). Effects of soil texture and structure on carbon and nitrogen mineralization in grassland soils. Biol. Fert. Soils, 14: 126–134.

Heil, K., and Schmidhalter, U. (2017). Improved evaluation of field experiments by accounting for inherent soil variability. Eur J Agron, 89:1–15.

Henderson, T.L., Szilagyi, A., Baumgardner, M.F., Chen, C.T., and Landgrebe, D.A. (1989). Spectral band selection for classification of soil organic matter content. Soil Sci. Soc. Am. J, 53: 1778–1784.

Hewson, R. D., Cudahy, T. J., and Huntington, J. F. (2001). Geologic and alteration mapping at Mt Fitton, South Australia, using ASTER satellite-borne data.

Huete, A. R., Liu, H. Q., Batchily, K, and Leeuwen, W.V. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS.

Hunt, J.r., Raymond, E., Daughtry, C. S., Eitel, T., Jan, U. H., and Long, D. S. (2011). Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index.

Ilawi, M. (1985). Soil map of Syria and Lebanon (ACSAD)Map.

Jafari, A.; Finke, P.A., Wauw, J.V., Ayoubi, S., and Khademi, H. (2012). Spatial prediction of USDA-great soil groups in the arid Zarand region, Iran: Comparing logistic regression approaches to predict diagnostic horizons and soil types. Eur. J. Soil Sci, 63: 284–298.

Jürgens, C. (1997).The modified normalized difference vegetation index (mNDVI) - a new index to determine frost damages in agriculture based on LANDSAT TM data.

Kuhn, M., and Johnson, K. (2013). Applied predictive modeling. 1st ed. New York: Springer.

Lagacherie, P., Baret, F., Feret, J.B, Madeira Netto, J., and Robbez- Masson, J.M. (2008). Estimation of soil clay and calcium carbonate using laboratory, field, and airborne hyperspectral measurements. Remote Sens. Environ, 112: 825–835.

Le Bissonnais, Y. (1996). Aggregate stability and assessment of crushability and erodibility: 1. Theory and methodology. Eur. J. Soil Sci, 47: 425–437.

Leblon, B. (1993). Soil and vegetation optical properties. In: Applications in Remote Sensing, Volume 4, The International Center for Remote Sensing Education. Wiley Press, New York, USA.

Liao KH, Xu SH, Wu JC, Zhu Q. (2013). Spatial estimation of surface soil texture using remote sensing data. Soil Sci Plant Nutr. 59:488–500.

Ließ, M., Glaser, B., and Huwe, B. (2012). Uncertainty in the spatial prediction of soil texture comparison of regression tree

and Random Forest models. Geoderma, 170:70–9.

Mangiafico, S.S. (2016). Summary and analysis of extension program evaluation in R, version 1.18.1. New Brunswick: Rutgers Cooperative Extension.

Makabe, S., Kakuda, K., Sasaki, Y., Ando, T., Fujii, H., and Ando, H. (2009). Relationship between mineral composition or soil texture and available silicon in alluvial paddy soils on the Shounai Plain, Japan. Soil Sci. Plant Nutr, 55: 300–308.

Malthus, T. J., Andrieu, B., Danson, F. M., Jaggard, K. W., and Steven, M. D. (1993). Candidate high spectral resolution infrared indices for crop cover. Remote Sensing of Environment.

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

McBratney, A.B., Santos, M.M., Minasny, B. (2003). On digital soil mapping. Geoderma, 117(1–2):3–52.

McBratney, A. (2014). Field DJ, Koch A. The dimensions of soil security. Geoderma, 213: 203–13.

McCarty, G.W., Reeves, G.B. III, Reeves, V.B., Follett, R.F., and Kimble, J.M. (2002). Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci. Soc. Am. J, 66: 640–646.

McDonald, J.H. (2014). Handbook of biological statistics. 3rd ed. Baltimore: Sparky House Publishing.

McKenzie, N.J., Ryan, P.J. 1999. Spatial prediction of soil properties using environmental correlation. Geoderma, 89(1-2): 67-94.

Mehrabi-Gohari, E., Matinfar, H. R., Jafari, A., Taghizadeh-Mehrjardi, R., and Triantafilis J. (2019). The Spatial Prediction of Soil Texture Fractions in Arid Regions of Iran. Soil Syst, 3: 65 doi:10.3390/soilsystems3040065.

Mitran, T., Mishra, U., Lal, R., Ravisankar, T., Sreenivas, K. (2018). Spatial distribution of soil carbon stocks in a semi-arid region of India. Geoderma Regional, 15:1–9

Mondejar, J.P., and Tongco A. F. (2019). Estimating topsoil texture fractions by digital soil mapping - a response to the long-outdated soil map in the Philippines. Sustainable Environment Research, 29:31, doi.org/10.1186/s42834-019-0032-5.

Moses, A. (1987). The climate of Syria, Damascus university publication 123. p

Mosleh, Z., Salehi, M. H., Jafari, A., Borujeni, I. E., and Mehnatkesh, A. (2016). The effectiveness of digital soil mapping to predict soil properties over low-relief areas. Environ Monit Assess 188:195. DOI 10.1007/s10661-016-5204-8.

Nikolakopoulos, K.G. (2003). Use of vegetation indexes with ASTER VNIR data for burnt areas detection in Western Peloponnese, Greece. IEEE Int. Geoscience and Remote Sensing Symp., September 21-25, Toulouse, France.

Oberthür, T., Dobermann, A., and Neue, H.U. (1996). How good is a reconnaissance soil map for agronomic purposes? Soil Use Manage, 12: 33–43.

Olaya, V. (2006). Basic Land-Surface Parameters. In: Hengl, T., Reuter, H.I. [Eds.]: Geomorphometry: Concepts, Software, Applications. Developments in Soil Science, Elsevier, 33: 141-169.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J. and Stenseth, N.C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. J. Trends Ecology Evolution, 9:503-510.

Pinheiro, H.S.K., de Carvalho, W., Chagas, C.D., dos Anjos, L.H.C., and Owens, P.R. (2018). Prediction of topsoil texture through regression trees and multiple linear regressions. Rev Bras Cienc Solo, 42:1–21.

Rawls, W.J, Brakensiek, D.L., and Saxton, K.E. (1982). Estimation of soil water properties. Trans. ASAE, 25: 1316–1320, 1328.

Scott D.W. (2018). Tukey ladder of powers. In: lane DM, editor. Introduction to statistics. Online Statistics Education: A Multimedia Course of Study. http:// onlinestatbook.com/Online_Statistics_Education.pdf. Accessed 21 Dec. 2018.

Seybold, C.A., Grossman, R.B., and Reinsch, T.G. (2005). Predicting cation exchange capacity for soil survey using linear models. Soil Sci. Soc. Am. J, 69: 856–863.

Sudduth, K.A., and Hummel, J.W. (1991). Evaluation of reflectance methods for soil organic matter sensing. Trans. ASAE, 34: 1900–1909.

Suliman, A.S., and Post, D.F. (1988). Relationship between soil spectral properties and sand, silt, and clay content of the soil on the University of Arizona Maricopa Agricultural Center. Proc. Hydrology and Water Resources in Arizona and the Southwest. J. Arizona Nevada Acad. Sci, 18: 61–65.

Sullivan, D.G., Shaw, J.N., and Rickman, D. (2005). IKONOS imagery to estimate surface soil property variability in two Alabama physiographies. Soil Sci. Soc. Am. J, 69: 1789–1798.

Tsai, AC, Liou, M., Simak, M., and Cheng, P.E. (2017). On hyperbolic transformations to normality. Comput Stat Data An, 115:250–66.

Tukey, J. W. (1977). Exploratory Data Analysis. Reading, Massachusetts: Addison-Wesley.

Warrington, D.N., Mamedov, A.I., Bhardwaj, A.K., and Levy, G.J. (2009). Primary particle size distribution of eroded material affected by the degree of aggregate slaking and seal development. Eur. J. Soil Sci, 60: 84–93.

Webster R, Oliver MA. (2007). Geostatistics for environmental scientists. Chichester: John Wiley & Sons, Ltd.

Wilding, L. P. (1985). Spatial Variability: its documentation, accommodation, and implication to soil surveys. In Nielsen, DR. Bouma, J. (Eds.). Soil Spatial Variability, Pudoc, Wageningen, Netherlands.

Wösten, J.H.M., Pachepsky, Y.A., and Rawls, W.J. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol., 251: 123–150.

Zevenbergen, L.W., and Thorne, C.R. (1987). Quantitative analysis of land surface topography. Earth Surface Processes and Landforms, 12: 47-56.

Zhang, R., Warrick, A.W., and Myers, D.E. (1992). Improvement of the prediction of soil particle size fractions using spectral properties. Geoderma, 52: 223–234.

Zhang, S.W., Shen, C.Y., Chen, X.Y., Ye, H.C., Huang, Y.F. and Lai, S. (2013). Spatial interpolation of soil texture using compositional kriging and regression kriging with consideration of the characteristics of compositional data and environment variables. J Integr Agr, 12:1673–83.

Zhao, Z.Y., Chow, T.L., Rees, H.W., Yang, Q. Xing Z.S., and Meng, F.R. (2009). Predict soil texture distributions using an artificial neural network model. Comput. Electron. Agr, 65: 36–48.

Zhu, Q., and Lin, H.S. (2010). Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. Pedosphere, 20: 594–606.