Journal of Zankoy Sulaimani
Part-A- (Pure and Applied Sciences)
http://jzs.univsul.edu.iq

Prediction of Crack Porosity from Other Easily Soil Properties Using Ridge Regression Analysis

Shadan H. Khurshid[1*]

1Department of Natural Resources, College of Agricultural Engineering Sciences, University of Sulaimani, Sulaimani,Kurdistan Region-Iraq.
**Corresponding author's e-mail: shadan.khurshid@univsul.edu.iq*

Article info

Original: 18 August 2019
Revised: 6 January 2020
Accepted: 30 January 2020
Published online: 20 June 2020

Key Words:
Ridge regression,
multicollinearity,
Soil crack porosity.

Abstract

The development of accurate, easy, and low-cost method to determine soil cracks porosity (SCP) is important in the evaluation of ecosystem to manage the hydrological, erosional, and geochemical cycles. Indeed, these procedures are cumbersome, time-and energy-consuming, and costly. Accordingly, intensive efforts are being made to formulate a high-performance model to estimate *SCP*.

Pedotransfer functions (PTFs) have often been developed using multiple linear regression models with no distinction about their involvement in multicollinearity. The paper was focused on ridge regression (RR) to overcome multicollinearity problems via regularizing the regression coefficients by imposing a penalty on their magnitudes. In the application of *RR,*choosing the ridge parameter (k) is important to control the amount of shrinkage of the regression coefficients. A total of 61 soil samples were analyzed from different land uses (forest, cropland, and pasture) inIraqi Kurdistan Region. Eighty-two percent of the soil sampleswere selected as the training set, and the rest of 11 soil samples (18%) were used as the testing set. Threemethods of modeling regression (simple linear regression SLR, multiple linear regression MLR, and ridge regression RR) were used to formulate accurate PTFs for predicting *SCP*. The results clearly showed multicollinearity problem (wrong sing and value of regression coefficient) in *SMR* and the most of *MLR*; therefore,they are not recommended to be used in predicting *SCP*. Soil crack porosity was positively correlated with each of clay (C), liquid limit (LL), plastic limit (PL), and percent of shrinkage limit (PLS). While, the most influential variable for predicting *SCP* is liquid limit (RMSE=126.194 $cm^2\ m^{-2}$; adj.R^2=0.884). The best model (RMSE=122.786 $cm^2\ m^{-2}$, adj. R^2=0.891) to predict *SCP* among all the models is formulated by *RR* technique when k=0.31395 which gave the lowest intercept value with very low *VIF* (0.582).

Introduction

Soil Cracks Porosity (SCP) is considered to be one measure of the soil quality [1], that makes the analysis of crack porosity and its management are very important. Soil Cracks Porosity is a complex phenomenon in materials like soils; it is a natural process that involves weathering and chemical changes. Desiccation cracking is the product of volumetric shrinking of clays due to soil moisture loss. Cracking initiates when tensile stresses generated by increasing suctions exceed the soil strength as a result of decreasing water content [2].

There are many factors that affect the development of soil cracking such as clay content, mineral composition, layer thickness and size, boundary conditions, wetting and drying cycles [3; 4; and 5]. Also, the soil cracking increases or decreases as result of the effect of grown crops, bulk density, groundwater and soil moisture characteristics of soils [6] particles size distribution, exchangeable cations [7] organic matter and cation exchange capacity [8; and 6] soil indices [9] clay content and liquid limit [10] liquid limit and plastic limit [1; and 11].

Cracking significantly affects soil performance, creating a zone of weakness in a soil mass and decrease the overall strength and stability of soil [12]. It causes altering in infiltration, runoff, evapotranspiration, and redistribution of water and chemicals. Consequently, crack contributes to the complex spatial and temporal variability of water redistribution in the landscape and creates challenges to the modeling of surface hydrology. Soil cracking exists in various geotechnical, agricultural and environmental applications such as irrigated land, tailing ponds for mining waste, landfill liners, earth embankment, reservoir beds [13]. These problems which created as a results of soil cracking can be reduced when the tendency of a soil to shrink and swell are known, that accurately requires both the knowledge of soil properties that affect swelling and shrinking, and the value of these parameters [10; and 14].To achieve this purpose a huge amount of data is required that is difficult to be obtained, especially in developing countries, due to the high cost of such studies [10; 15; and 16]. This situation causes impossibilities in using these soil properties in predicting soil crack parameters; also, caused difficulties or impossible sustainability of agricultural production with minimum environmental degradation in Vertisols [10; and 11] therefore, development of accurate, easy, and low cost ways is important to estimate soil cracks in the evaluation of soil quality for the purpose of improving sustainability of the soil system in areas where data is limit [10; 17; and 18].On the other hand,due to that, the crack exists in the soil interior, the change of outside condition leads the crack to open or close, therefore, the process of change is complex and direct or indirect observation is difficult. Until the present time, describing and analyzing the process of the crack developing cannot be accurate. Indeed, these procedures are cumbersome, time-and energy-consuming, and costly. Accordingly, intensive efforts are being made to develop predictive relationships [10]. Prediction of soil crack parameters from readily available soil properties such as clay, liquid limit, and organic matter has been proposed by many researchers [1; 7; 8]. Most of these models are individual regression or careless to the presence of multicollinearity among the data. Collinearity or multicollinearity was introduced as a concept by Frisch [19] reflecting the existence of near-linear relationships among the independent variables. For this purpose, the current study was carried out to overcome the effect of multicollinearity on the goodness of predictive model by using *RR*. It is a method for analyzing multiple regression data with high rate of multicollinearity problem [20]. Ridge regression is based on minimization of the usual least-squares criterion plus a penalty term. As such, it shrinks the coefficient estimates towards zero [21]. This introduces bias, but reduces the variance consequently; reduces the standard errors and expects that the net effect will give estimates that are more reliable.

Materials and Methods

Study areas and soil sampling:

Sixty-one soil samples were collected from different land usesnamely, forest, cropland, and pastureland in Iraqi Kurdistan Region, with soil samples taken to a depth of 30 cm. In other words, wide ranges in soil physical, chemical, geotechnical, and mineralogy properties were expected.The soils were air-dried, ground to pass through a 2mm sieve and stored in plastic containers until use. Also, undisturbed soil samples were taken from each soil using steel cores 5.6cm in diameter by 4cm long to determinetheir insitu bulk densities at the normal state of compaction. The cores were driven into the soil at or near field capacity with the aid of a core sampler.

Physical and chemical analyses:

Particle size distribution was determined by sieving and pipette according to the methods described by [22], Liquid limit was measured according to [23], while plastic limit and plasticity index were determined according to [24]. On the other hand, percent of linear shrinkage was measured according to the method outlined by [25]. Bulk density was measured by core method as described by [26]. calcium carbonate contents using acid neutralization method according to Richards as described in [27], total OM content applying Walkley and Black method (wet oxidation method) as described by [28], pH, and some soluble ions were determined according to the procedure outlined by [29].

Samples preparation

Predetermined portions of soil were packed in 183 metal pots each with an average diameter of 25cm and

23cm in height under a normal state of soil compaction.Before packing, soil moisture content of the soils were raised to nearly optimum water content by adding the required amount of water using a special sprayer. The samples were then thoroughly mixed by hands, kept in polyethylene bags and stored in a cool place for a period of 24hrs, to ensure soil water distribution.

The packing was performed with the aid of metal packer, and the objective was to achieve a normal state of soil bulk density. For this purpose, the optimum water content was determined according to the model described by [30]. Then the water was added to the pots to restore the soil water content to nearly field capacity. Later the prepared pots subjected to four cycles of wetting and drying, at the end of the last cycle of drying, the soil columns were again wetted and monitored daily to fix the water content at which cracks start to develop and at the water content at which the cracks reach a stabilized state. This was attained by periodical weighing of the cylinders together with visual observations on cracks.

Crack porosity

Soil crack porosity of sixty one soils was determined by applying Novak procedure thatis defined as crack area in a representative area of soil surface [34].

$P_c(0)=A_c(0)/A$

Where:

$P_c(0)$ = crack porosity on the soil surface.

$A_c(0)$ = crack area.

A = representative area of the soil surface.

Novak et al. (2000) [35] pointed out the $P_c(0)$ represents maximum crack porosity corresponding to zero soil water content, i.e., it represents the intercept of the linear relationship between crack porosity and gravimetric water content.

To calculate the crack area $A_c(0)$the basic characteristics (depth, width, and length) of soil cracks were measured assuming the triangular shape of cracks [31]. The average depth and width of crack were based on measurements made at several random points along the length of the cracks. The crack depth was measured by inserting a 2mm steel rod until it countered resistance to further penetration and the width at the same point as that of depth recording, was measured using vernier [32]. On the other hand, the crack length was measured by running a flexible twine along the crack and measuring the total length [33].

$SA = \sum 2\ CL$

$C=[(0.5w)^2 + d^2]^{1/2}$

Where:

SA = surface area (m^2).

w = the mean width of cracks (m).

d = the mean depth of crack (m).

L = the length of the crack (m).

C = the parameter based on (w) and (d).

Generating prediction model:

As stated earlier, the models were developed by using simple linear, multiple linear and ridge regression analyses. The coefficients of the multivariate were determined by using stepwise (forward and backward) procedures.To model the data in various land uses, a holdout method-cross validation was used for validation of the models; the soil samples were randomly split into 82% and 18% as training and validation set, which comprised as 50 and 11 soil samples, respectively. Clay, liquid limit, plastic limit, and percent of linear shrinkage were used as predictor variables. Descriptive analyses such as mean, minimum, maximum, standard deviation, and coefficient of variation were carried out using NCSS Version 12.0.2 (NCSS LLC, Utah, USA).

Simple Linear Regression Models (SLR)

Linear Regression points out to a set of techniques for fitting and describing the straight-line relationship between two variables. Individual regression models between *SCP* and each of (CL, LL, PL, and PLS) were formulated using a least-square difference (LSD) technique. The general form of *SRL* equation for estimating *SCP*:

$P_c = \beta_0 + \beta_1 X + \varepsilon$

where$P_c$is given in $cm^2\ cm^{-2}$, *X* is the individual predictor variable (CL, LL, PL, and PLS), and β_0 is the *Y*-intercept, β_1 is the slope, and ε is the error. (e.g., for the predicted P_c equation as a function of *CL*, the intercept β_0 corresponds to the value of *SCP* when the clay content is zero in a soil sample, and the slope β_1 is rate effect of clay content on P_c).

Multiple Linear Regression Models (MLR)

Pedotransfer functions have often been developed using MLR models [36]. It refers to a group of techniques for describing the straight-line relationships among two or more variables. Multiple regression models among *SCP* and CL, LL, PL, and PLS were formulated using a least-square difference (LSD) technique. The general form of *MRL* equation for estimating *SCP* from all properties is

$P_c = \beta_0 + \beta_1\, CL + \beta_2\, LL + \beta_3\, PL + \beta_4\, PLS + \varepsilon$

β_0 is the *Y*-intercept; β_1, β_2, β_3, β_4, are the coefficients to CL, LL, PL, and PLS respectively; ε is the error of the model.

Also,apedotransfer function was developed for predicting soil crack porosity based on stepwise (forward and backward) multiple regression procedure, is a technique of fitting regression models in which the choice of independent variables is conducted by an automatic procedure by XLSTAT (Version 2016.02).

Ridge regression

One of the alternative estimation technique for analyzing unknown parameters of the linear regression models is ridge regression (RR), and it belongs to the category of biased regression methods [37]. This technique decrease the variance of the parameter estimates based on increasing bias in the regression equation, This bias is entered with the ridge parameter, which determines the extent of the shrinkage of the least-squares estimates. The most difficult stepin *RR* is to determine the ridge parameter. The value of the ridge parameter that minimized the root mean squared error was chosen [38].

Multicollinearity diagnosis

Multicollinearity makes determining the contribution of each explanatory variable difficult because the effects of these variables are mixed. Coefficients may have the wrong sign or an illogical magnitude [39]. Various indicators have been used to detect multicollinearity problems. Draper and Smith (1998) [40] suggest the following: (1) check if some regression coefficients have the wrong sign, based on sign of singular regression coefficients; (2) check the correlations between all pairs of independent variables, to note if any is suddenly high; (3) check if deletion of a row or a column of the X matrix produces suddenly large changes in the fitted model; (4) check if predictors anticipated to be important based on prior knowledge have regression coefficients with small t-statistics; and (5) examine the variance inflation factors (VIF).

$$VIF = \frac{S^2(n-1)SD^2}{MSE}$$

Root mean square error (RMSE) and adjusted-R^2 (adj.R^2).

After generating the models by fitting with the observed data, it is necessary to know if the fitted model is valid. Often both *RMSE* and adj.R^2 are used to exam the validation of model.

1. $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(O_i - P_i\right)^2}$

2. $adj.R^2 = 1 - \left[\frac{(n-1)}{(n-k-1)}(1-R^2)\right]$

Where:
O_i=observed values; P_i=predicted values; n=number of observations; k=number of independent variables; R^2=coefficient of determination; *RMSE*=root mean square error;and $Adj.R^2$=Adjusted R-square.

Results and Discussion

Descriptive statistics regarding *SCP* in the study areas are shown in Tables 1 and 2 for training and validation sets, the minimum *SCP* values were 224 and 347 cm^2 m^{-2} in training and validation sets, respectively, while the maximum values were equal in both sets 1510 cm^2 m^{-2} (Tables 1 and 2). The level of dispersion as reflected by %CV are moderate class for all soil characteristics in different sets based on [41], except the *SCP* in validation set is high class (38.99%) which was more than 36%.

Table – 1: Some descriptive statistics for the training data

Variable	*Observations*	*Minimum*	*Maximum*	*Mean*	*Standard deviation*	*CV %*
SCP cm^2 m^{-2}	50	224	1510	1009.80	11.54	31.58
Clay %	50	11	57	38.42	10.59	30.04
LL%	50	19	58	42.82	5.44	24.72
PL%	50	11	31	23.64	4.09	22.99
PLS%	50	4	20	12.96	318.91	31.52

Table – 2: Some descriptive statistics for the test data

Variable	*Observations*	*Minimum*	*Maximum*	*Mean*	*Standard deviation*	*CV %*
SCP cm^2 m^{-2}	11	347	1510	951.76	11.22	38.99
Clay %	11	18	51	38.73	10.63	28.96
LL%	11	23	55	43.09	5.81	24.68
PL%	11	13	30	24.27	3.80	23.96
PLS%	11	7	18	13.36	371.10	28.45

Table 3 shows the values of correlation for the studied variables in the training data set. Soil crack porosity (SCP) was highly positive influenced by all studied variables (C, LL, PL, and PLS), and the highest correlation was appear with each of clay and *LL* as values r= 0.91 (Table 3). Also, the same table showed highly positive correlations among independent variables, and the highest correlation values were appeared among clay with each of *LL* and *PLS* (0.95). Hence, whenever the *r*-value between one independent variable and the rest is greater than or equal to 0.95, then multicollinearity will be present.

Table-3: The correlation matrix among studied variables in the training set.

Variables	*Clay*	*LL*	*PL*	*PLS*	*SCP*
Clay	1.00				
LL	0.95	1.00			
PL	0.88	0.90	1.00		
PLS	0.95	0.94	0.85	1.00	
CSP	0.91	0.91	0.81	0.89	1.00

The presences of multicollinearity problem (value of regression coefficient and its sign) were evident in *SMR* and the most of *MLR* (Table 4). Ridge Regression (RR) technique was used to overcome these problems.

Many values of k-factors in *RR* have been used in an attempt to reduce multicollinearity effect; from these values, there were selected just *k*=1, *k*=0.31395, *k*=0.05, and *k*=0, which mean without RR technique. By giving the maximum value of k=1, model no.18 was formulated to have the lowest value of *RMSE* and less multicollinearity effect with *VIF* value of 0.127 (very low value). By using this value of *k*, the intercept value of the model will be high (137.157, table 4) compared to other intercept values, resulted with lower *k* value. This high intercept value is part of the model error; that being appeared when the independent variables are zeroat time of predicting the dependent variable (herein, if values of *C*, *LL*, *PL*, and *PLS* are zero, the predicted value of *SCP* will take intercept value which is 137.157 $cm^2 m^{-2}$).

In case of giving the minimum value of k=0 (formulating model without using *RR* technique) the multicollinearity problem will be appeared clearly, the negative sign of *PL* coefficient in model no.20 instead of being positive as appeared on model no.3 (Table 4). It means that the individual effect of *PL* on *SCP* is positive. On the other indicators of multicollinearity problem when k=0, *VIF* value is raising to 15.489. (Table 4). Applying model no.20 will give the inaccurate predicting value of *SCP* despite low *RMSE* value and high adj.R^2, especially when the *PL* value is high, as its effect will be inversed on *SCP* value.

The best *k* value to overcome multicollinearity problem is *k*= 0.31395, giving the lowest intercept value (model no.19, table 4) among all other attempts to approaches zero value with very low *VIF* value (0.582). Model no.19 after validation gave the lowest *RMSE* value (122.786 $cm^2 m^{-2}$) compare to attempts that have not multicollinearity problem, as well as adj.R^2 is high (0.891) as appear in table 5. These results are concluding to be the best model to predict *SCP* among all the models.

Both models 16 and 17 (forward and backward) formulated with stepwise technique are categorized with *MLR*, and each one posing two independent variables (*C* and *LL*) while the rest independent variables (*PL* and *PLS*) are canceled according to forward and backward procedure (by another mean, presence of *PL* and *PLS* are not important in models 16 and 17(Table 4). Despite the low value of *RMSE* and high rate of adj.R^2 (table, 5) they are not recommended to be applied in prediction of *SCP*, as they both have multicollinearity problem by having *VIF* value more than 10 (Table 4). Additionally, the backward stepwise regression has a large value of intercept (-181.384,Table 4).

Out of all *MLR* models, only three have less *VIF* value, as they are models 6, 8, and 10 while others have multicollinearity problem (all with high *VIF* value more than 10, table 4) and they are not recommended to be used in predicting *SCP*. From the three chosen models, model no.8 has the minimum *RMSE*, with highest adj.R^2 value. Whereby, it considered as the best model among *MLR* models to predict *SCP* without treating for multicollinearity effect.

While, in case of focusing on *SLR* with no multicollinearity problem because they formulated from one independent variable, models no. 1, 2, 3, and 4 were formulated (Table 4). The second model has less *RMSE* value to predict *SCP* value from *LL* value; at the same time, it has the lowest value of adj.R^2 (Table 5). The mentioned performance criteria make this model be the best among all *SLR* models, while the only issue with this model is the large negative value of intercept (-169.614, table 4).

Table-4: Maximum variance inflation factor (VIF) values according to the regression typeand choosing (k) in ridge regression.

Type of Model		R^2	*Var.*	*SE*	*MSE*	VIF_{MAX}	*Models*	*Models No.*
Univariate		0.823	C	1.676	18324.321	-	SCP = 46.491+25.073*C	1
		0.836	LL	1.760	17006.415	-	SCP= -169.614+27.543*LL	2
		0.663	PL	4.915	34965.326	-	SCP= -119.616+47.775*PL	3
		0.785	PLS	5.220	22284.195	-	SCP = 113.431+69.164*PLS	4
Multivariate		0.851	LL	5.553	15830.865	10.697	SCP=-104.046+10.883*C+16.248*LL	5
		0.824	PL	7.597	18616.170	4.488	SCP=17.432+23.504*C+3.779*PL	6
		0.829	PLS	15.198	18112.498	10.430	SCP=45.976+18.681*C+18.990*PLS	7
		0.837	PL	8.018	17303.093	5.378	SCP= -156.781+29.106*LL-3.373*PL	8
		0.842	PLS	13.338	16746.201	8.689	SCP= -124.0431+21.145*LL+17.624*PLS	9
		0.798	PLS	9.743	21374.755	3.632	SCP= -1.062+12.774*PL+54.698*PLS	10
		0.853	LL	6.192	15945.294	13.204	SCP = -75.5014+11.605*C+18.4443*LL-6.359*PL	11
		0.851	C	6.077	16151.364	14.924	SCP=-98.718+10.043*C+15.661*LL+4.0156*PLS	12
		0.819	C	6.059	18444.983	12.990	SCP=23.173+17.629*C+2.968*PL+18.453*PLS	13
		0.843	LL	6.318	17036.296	12.868	SCP=-110.245+22.754*LL-3.555*PL+17.728*PLS	14
		0.853	C	6.217	16286.821	15.489	SCP=-72.371+10.973*C+17.965*LL-6.208*PL+2.937*PLS	15
Multivariate	*FW*	0.851	C	5.094	15830.865	10.697	SCP=-104.046+10.883*C+16.248*LL	16
	BW	0.853	C	6.084	15593.764	15.489	SCP = -181.384+10.976*C+17.970*LL	17
RR	*k=1*	0.657	PL	1.827	37969.360	0.127	SCP=137.157+5.577*C+6.249*LL+8.485*PL+14.673*PLS	18
	k=0.31395	0.771	PL	3.191	25337.350	0.582	SCP=0.009595+7.052*C+8.242*LL+7.109*PL+16.808*PLS	19
	k=0	0.853	C	6.217	16286.821	15.489	SCP=-72.371+10.973*C+17.965*LL-6.208*PL+2.937*PLS	20

Table-3: Performance criteria (adj.R^2 and RMSE) in the validation of formulated models.

Regression types		*No.*	*adj.R^2*	*RMSE $cm^2\ m^{-2}$*
SLR		1	0.774	176.245
		2	**0.884**	**126.194**
		3	0.826	154.835
		4	0.784	172.536
MLR		5	0.897	125.433
		6	0.802	165.252
		7	0.823	164.586
		8	**0.883**	**126.751**
		9	0.889	123.908
		10	0.867	142.470
		11	0.874	131.536
		12	0.880	135.656
		13	0.840	156.385
		14	0.897	125.391
		15	0.888	131.149
SMR	*FW*	16	0.897	125.433
SMR	*BW*	17	0.899	124.506
RR	*k=1*	18	0.894	120.956
	k=0.31395	**19**	**0.891**	**122.786**
	k=0	20	0.888	131.149

Conclusions

In this paper, ridge regression technique based on accuracy indicators (RMSE and adj.R^2) and presence of multicollinearity problem was compared to *SLR*, *MLR*, and *SMR* in the prediction of *SCP* from clay, liquid limit, plastic limit, and percent of shrinkage limit. All these variables were positively correlated with *SCP*; the most effective variable is the liquid limit for predicting *SCP*. Ridge regression technique at k=0.31395 gave the best model has the lowest value of intercept and *VIF*(fewer multicollinearity problems) as well as offered the best predictive performance (*RMSE*; adj.R^2). Therefore, *RR* considered as a good technique to overcome multicollinearity effects by rectifying signs and values of regression coefficient when the modelis made for prediction purposes.

References

[1] Khurshid, Sh.H., "*Characterization of cracked soils and strategies for their management in Sulaimani Governorate*", PhD dissertation, University of Sulaimani, Sulaimani. (2010).

[2] Stirling, R.A., Davie, C.T., and Glendinning, S., "*Numerical modeling of desiccation crack induced permeability*", 18th International Conference on Soil Mechanics and Geotechnical Engineering, pp. 813-816, Paris. (2013).

[3] Mitchell, J.K., "*Fundamentals of Soil Behavior*", John Wiley and Sons, Inc., New York, NY. (1993).

[4] Nahlawi, H. and Kodikara, J.K., "*Laboratory experiments on desiccation cracking of thin soil layers*", Geotechnical and Geological Engineering. Vol. 24, pp. 1641–1664. (2006).

[5] Omer, M.E.T., and Mohd, R.T., "*Cracks in Soils Related to Desiccation and Treatment*", Australian Journal of Basic and Applied Sciences, pp. 1080-1089, Malaysia. (2011).

[6] Reeve, M.J., Hall, D.G.M., and Bullock, P., "*The effect of soil composition and environmental factors on the shrinkage of some clayey British soils*", European Journal of soil scienc. Vol. 31, No. 3, pp. 429-442. (1980).

[7] Karim, T.H., and Khalid, A.S., "*Survey on different soil properties in northern Iraq*". (2000).

[8] Gray, C.W. and Allbrook, R., "*A relationship between shrinkage induces and soil properties in some New Zealand soils*", Geoderma. Vol. 108, No. 3-4, pp. 287-299. (2002).

[9] Tang, Ch., Bin, Sh., Liu, Ch., Zhao, L., and Wang, B., "*Influencing factors of geometrical structure of surface shrinkage cracks in clayey soils*", Engineering geology. Vol. 101, No. 3-4, pp. 17. (2008).

[10] Fattah, M.A., Khurshid Sh. H., and Ahmad R. A., "*Soil Cracking Depth as Influenced By Soil Physical Properties*", Special Issue, 2nd Int. Conference of Agricultural Sciences. Journal of Zankoi Sulaimani. pp. 105-114. (2018).

[11] Maulood. Y. I., "*Control of Cracks due to Drying Shrinkage of an Expansive Soil Using Different Drying Mechanisms and Filler Additive*", Zanco Journal of Pure and Applied Sciences. Vol. 27, No. 1, pp. 31-40. (2015).

[12] Alvis, A. and Marcelo, S., "*Analysis of Cracking Behavior of Drying Soil*", 2nd International Conference on Environmental Science and Technology. IPCBEE, 6, 66-69, IACSIT Press, Singapore. (2011).

[13] Yesiller, N., Miller, C. J.,Inci, G. and Yaldo, K., "*Desiccation and cracking behaviour of three compacted landfill liner soils*", Wayne State University, Department of Civil. and Environmental Engineering, Detroit, USA. (2000).

[14] Huawang, S., "*Application of Unascertained Measure to Classification of Expansive Soils*", International Colloquium on Computing, Communication, Control, and Management, ISECS, pp. 57-60, China. (2009).

[15] Masto, R.E., Sheik, S., Nehru, G., Selvi, V.A., George, J., Ram, L.C., "*Assessment of environmental soil quality around SonepurBazari mine of Raniganj coalfield, India*", Solid Earth Vol. 6, No. 3, pp. 811–821. (2015).

[16] Van Leeuwen, J.P., Lehtinen, T., Lair, G.J., Bloem, J., Hemerik, L., Ragnarsdóttir, K.V., Gísladóttir, G., Newton, J.S., de Ruiter, P.C., "*An ecosystem approach to assess soil quality in organically and conventionally managed farms in Iceland and Austria*", Soil Vol. 1, pp. 83–101. (2015).

[17] Costa, J.L., Aparicio, V., Cerdà, A., "*Soil physical quality changes under different management systems after 10 years in the Argentine humid pampa*", Solid Earth Vol. 6, No. 1, pp. 361–371. (2015).

[18] Pulido, M.M., Gabriels, D., Cornelis, W., Lobo, D., "*Comparing aggregate stability tests for soil physical quality indicators*", Land Degradation and deployment. Vol. 26, No. 8, pp. 843–852. (2015).

[19] Frisch, R., "*Statistical Confluence Analysis by Means of Complete Regression Systems*", Oslo: University Institute of Economics, Publication No. 5. (1934). Cited from [Spyroglou I., Chatzimichail, E.Spanou, E.N. Paraskakis, E. andRigas, A., "*Ridge regression and bootstrapping in asthma prediction*", Conference Paper, New Developments in Pure and Applied Mathematics. (2015).

[20] Pimentel, E. CG., Queiroz, S A. Carvalheiro, R. Fries, L. A., "*Use of ridge regression for the prediction of early growth performancein crossbred calves*", Genetics and Molecular Biology. Vol. 30, No. 3, pp. 536-544. (2007).

[21] O'brien R.M., "*A caution regarding rules of thumb for variance inflation factors*", Quality and Quantity. Vol. 41, pp. 673–690. (2007).

[22] Klute, A., "*Methods of Soil Analysis*", Part 1. 2nded. Agron.Monogr. 9. ASA and SSSA. Madison. WI. (1986).

[23] ASTM-D423. "*American Society for testing and materials*", 1916 Race street, Phiadelphia, P.A, 19103. (1986).

[24] ASTM-D424., "*American Society for testing and materials*",1916 Race street, Phiadelphia, P.A, 19103. (1986).

[25] British Standards Institution BS 1377., "*Methods of test for soils for civil engineering purposes*", 2 Park Street, London. (1975).

[26] Blake, G.R. and K.H. Hartge, "*Bulk density. In method of soil analysis*", Part 1: Physical and mineralogical methods. 2nd. edited by A.Klute. pp. 363-375. (1986).

[27] Rowell, D.L., "*Soil Science: Methods and applications*", Longman Group UK Limited. pp. 50-51. (1996).

[28] Nelson, D. W. and Sommers, L. E., "*Total carbon, organic carbon and organic matter*", p. 539-580, *In:* page, A.L., Miller, R.H. and Keeney, D.R. (eds) "*Methods of soil analysis*" part 2. Chemical and microbiological properties.Agron.Monogr.9 ASA Madison, WI. (1982).

[29] Page, A.L., Miller, R.H. and Keeney, D.R., "*Methods of soil analysis*", part 2. Chemical and microbiological properties.2nd edition Agronomy No.9.ASA, Madison, Wisconsin USA. (1982).

[30] Fattah, M., A., "*Some practical tests to reduce seepage from earthen ponds for different texture soils*", MSC. Thesis, University of Salahaddin, College of Agriculture, Soil and water science department. (2004).

[31] Sharma, P. K., T. S. Verma, and R. M. Bhagat, "*Soil structural improvements with addition of lantana Camara biomass in rice-wheat cropping*", Soil use and management, Vol. 11, pp. 199-203. (1995).

[32] Dasog, G. S., Acton, D. I., Mermute, A. R., de Jong, E., "*Shrink-swell potential and cracking in clay soils of Saskatchewan*"Canadian Journalof Soil Science. Vol. 68, pp. 251-260. (1988).

[33] Dasog, G. S. and G. B. Shashidhara."*Measurements of cracking intensity in a vertisol under different crop cover*", Soil Science. Vol. 156, pp. 424-428. (1993).

[34] Novak, V., "*Soil crack characteristics-estimation methods applied to heavy soils in the NOPEX area*", Inst. Of hydrology, Slovak academy of Sci. P. O. Box 94, Racianska 75, 830 08 Bratislava, Slovak Republic.Agr.And Forest Meteorology Vol. 98-99. pp. 501-507. (1999).

[35] Novak, V., J. Simunek, and M. Th., van Genuchten, "*Infiltration of water in to soil with cracks*", J. of Irr. And Drain., Eng., ASCE, Vol. 126, pp. 41-47. (2000).

[36] Sequeira, C.H., Wills, S.A., Seybold, C.A., West, L.T., "*Predicting soil bulk density for incomplete databases*",Geoderma. Vol. 213, pp. 64–73. (2014).

[37] Hoerl, A.E. and Kennard, R.W., "*Ridge Regression: Biased estimates for non-orthogonal problems*",Technometrics. Vol. 12, No. 55, pp. 55–67. (1970).

[38] Cessie, S. Le and Van Houwlingen, J. C., "*Ridge Estimators in Logistic Regression*", Journal of the Royal Statistical Society. Vol.41, No. 1, pp. 191–201. (1992).

[39]Hair Jr JF, Anderson RE, Tatham RL and Black WC, "*Multivariate Data Analysis*", 3rdedition. Macmillan Publishing Company, New York, pp.544. (1992).

[40] Draper NR.and Smith H., "*Applied Regression Analysis*", 3rdedition. John Wiley & Sons, New York, pp.706 (1998).

[41] Wilding, L.P., "*Spatial variability: Its documentation, accommodation and implication to soil surveys*", In: D.R. Nielson and J. Bouma,editors, Soil spatial variability. Pudoc, Wageningen, Netherlands. pp.166–196. (1985).