International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:5, No:2, 2011

Studding of Number of Dataset on Precision of Estimated Saturated Hydraulic Conductivity

M. Siosemarde and M. Byzedi

Abstract**—**Saturated hydraulic conductivity of Soil is an important property in processes involving water and solute flow in soils. Saturated hydraulic conductivity of soil is difficult to measure and can be highly variable, requiring a large number of replicate samples. In this study, 60 sets of soil samples were collected at *Saqhez* region of *Kurdistan* province-*IRAN*. The statistics such as Correlation Coefficient *(R)*, Root Mean Square Error *(RMSE)*, Mean Bias Error *(MBE)* and Mean Absolute Error *(MAE)* were used to evaluation the multiple linear regression models varied with number of dataset. In this study the multiple linear regression models were evaluated when only percentage of sand, silt, and clay content *(SSC)* were used as inputs, and when *SSC* and bulk density, *Bd***,** *(SSC+Bd)* were used as inputs. The *R*, *RMSE*, *MBE* and *MAE* values of the 50 dataset for method *(SSC)*, were calculated 0.925, 15.29, -1.03 and 12.51 and for method *(SSC+Bd)*, were calculated 0.927, 15.28,-1.11 and 12.92, respectively, for relationship obtained from multiple linear regressions on data. Also the *R*, *RMSE*, *MBE* and *MAE* values of the 10 dataset for method *(SSC)*, were calculated 0.725, 19.62, -9.87 and 18.91 and for method *(SSC+Bd)*, were calculated 0.618, 24.69, -17.37 and 22.16, respectively, which shows when number of dataset increase, precision of estimated saturated hydraulic conductivity, increases.

Keywords**—**dataset, precision, saturated hydraulic conductivity, soil and statistics.

I. INTRODUCTION

THE saturated hydraulic conductivity of soil is a measurement parameter of the soil's ability to transmit water when submitted to a hydraulic gradient that is defined by Darcy's law. The hydraulic conductivity depends on the soil grain size, the structure of the soil matrix, the type of soil fluid, and the relative amount of soil saturation present in the soil matrix. The important properties relevant to the solid matrix of the soil include pore size distribution, pore shape, tortuosity, specific surface, and porosity [11 & 14]. Numerous field and laboratory techniques have been developed to directly measure this quality [1 & 25].

Since, direct measurement of hydraulic conductivity is time consuming and costly, the cost-effectiveness of obtaining soil hydraulic conductivity can be improved by using indirect methods, which allow the estimation of hydraulic conductivity from more easily measured parameters [2, 3, 5, 13, 17, 20, 24, 26, 27, 28, 29, 32, 33, 34, 35 & 36].

Knowledge about the variability of soil properties is probably as old as the soil classification system [12 & 30]. Variability in the hydraulic properties of soil units has been studied by several researchers [7, 12 & 15]. Field observations show that the hydraulic properties of soils vary significantly with spatial location even within a given soil type [12 & 22].

Hazen (1982) proposed the relationship between saturated hydraulic conductivity and soil particle diameter such that 10% of all soil particles are finer (smaller) by weight [6 & 9]. Shepherd (1989) extended Hazen’s research by performing power regression analysis [21]. Puckett et al. sampled six soils at seven different locations in the Alabama lower coastal plain [18], and used regression analysis to determine that percentage of clay sized particles was the best predictor of Ks. Rawls and Brakensiek used field data across the U.S. to develop a regression equation that relates porosity, and the percentages of sand and clay-sized particles in the sample to K_s [19]. Jabro estimated K_s from grain-size and bulk density data [10]. Sperry and Peirce (1995) developed a linear model to estimate K_s based on grain size, shape, and porosity, also they performed an evaluation by comparing the measured hydraulic conductivity values of different porous materials with those determined from their own equations as well as the equations of Hazen, Kozeny-Carman and Alyamani and Sen [23]. Cronican and Gribb (2004) developed a multiple linear regression for southeastern U.S. sandy soils based on regional soil data [6]. Han et al. (2008) developed a new model to estimate saturated hydraulic conductivity from soil structural properties derived from water retention curve [8]. Jadczyszyn and NiedŸwiecki (2005) reported that the lower content of both silt and organic matter and lower values of bulk density had increased K_s [11]. The results showed that the hydraulic conductivities calculated by the USBR and Slitcher methods are in all cases lower than for the other methods [4, 16 & 31].

Although many researches carried out for prediction of saturated hydraulic conductivity from soil texture and particle diameter but in this research an attempt has been made to study the effect of dataset number on precision of estimated saturated hydraulic conductivity.

II. MATERIALS AND METHODS

In this study, 60 sets of soil samples were collected to measure hydraulic conductivity based on parameters of soil texture at location of *Saqhez* region of *Kurdistan* province-*IRAN*. Standard methods were applied to investigate parameters of soil texture (clay, silt and sand). Soil texture

M. Siosemarde is with the Sama technical and vocational training school, Islamic Azad University, Mahabad Branch, Mahabad, Iran. Corresponding author to provide phone: +98-918-8725212; fax: +98-442-2336187; (e-mail: maroof_33m@yahoo.com).

M. Byzedi is with the Sama technical and vocational training school, Islamic Azad University, Mahabad Branch, Mahabad, Iran. (e-mail: M.byzedi@gmail.com).

was classified according to the International Society of Soil Science *(ISSS)* classification system. The textural classification of soils was clay, silty clay, silty clay loam, silty loam and loam. Bulk density was determined from the mass of dry soil contained in 120 cm^3 steel sampling cylinders. The samples were oven-dried at 105°C for 48 hours. The results were calculated as g/cm^3. The content of clay, silt & sand and values of bulk density and saturated hydraulic conductivity are summarized in Table 1. The mean contents of clay, silt and sand were 36.00, 50.45 and 13.12 [g (100g)$^{-1}$], respectively, also the mean values of bulk density and saturated hydraulic conductivity were 1.31 (g/cm^3) and 50.40 (cm/day).

In this study saturated hydraulic conductivity was measured by the falling-head method. The falling-head method is very similar to the constant head methods in its initial setup; however, the advantage to the falling-head method is that can be used for both fine-grained and coarse-grained soils. The samples were first wetted by capillarity for 24 hours. This was done from the bottom so that air could escape from the upper surface. The water is then allowed to flow through the soil without maintaining a constant pressure head and saturated hydraulic conductivity was measured when the rate of head loss was constant.

TABLE 1
SUMMARIZE OF STATISTICS OF SILT, CLAY AND SAND CONTENT AND VALUES OF BULK DENSITY AND SATURATED HYDRAULIC CONDUCTIVITY

Statistics	PARAMETERS				
	Bd	*Silt*	*Clay*	*Sand*	K_s
Mean	1.31	50.45	36.00	13.12	50.40
Median	1.31	51.00	36.50	12.00	41.15
Std. Error of Mean	0.012	1.267	1.370	0.897	4.91
Minimum	1.06	29.0	10.0	3.0	0.10
Maximum	1.50	69.0	57.0	36.0	157.0
Std. Deviation	0.089	9.813	10.61	6.946	30.04
Skewness	-0.034	-0.212	-0.354	1.043	1.085

Silt is the percentage of silt-sized particles; *Clay* is the percentage of clay-sized particles; *Sand* is the percentage of sand-sized particles; K_s, saturated hydraulic conductivity is expressed in cm/day and *Bd* is the soil bulk density (g/cm^3).

In this study multiple linear regression models obtained by different number of dataset including 50, 40, 30, 20 and 10 series of dataset and the obtained models were evaluated using the other 10 series of dataset. The multiple linear regression models were developed when only percentage of sand, silt, and clay content *(SSC)* were used as inputs, and when SSC and bulk density, *Bd*, *(SSC+Bd)* were used as inputs. In this study the models were studied by changing the number of dataset.

The results were analyzed with *SPSS 16.0* and *EXCEL* software with statistics of Correlation Coefficient *(R)*, Root Mean Square Error *(RMSE)*, Mean Bias Error *(MBE)* and Mean Absolute Error *(MAE)* that calculated using equation (1), (2), (3), and (4) respectively, where O_i and P_i represents measured and predicted, and O_{ave} and P_{ave} represents mean values of measured and predicted respectively, and *n* represents the number of instances presented to the model that it is total number of testing pattern (10 series of dataset).

$$R=\left[\frac{\sum_{i=1}^{n}(O_i-O_{ave})(P_i-P_{ave})}{\sqrt{\sum_{i=1}^{n}(O_i-O_{ave})^2(P_i-P_{ave})^2}}\right]^{1/2} \quad (1)$$

$$RMSE=\left[\frac{\sum_{i=1}^{n}(P_i-O_i)^2}{n}\right]^{1/2} \quad (2)$$

$$MBE=\sum_{i=1}^{n}\left[(P_i-O_i)/n\right] \quad (3)$$

$$MAE=\sum_{i=1}^{n}\left[\left|P_i-O_i\right|/n\right] \quad (4)$$

III. RESULT AND DISCUSSION

The following equations for K_s, saturated hydraulic conductivity (cm/day), were obtained from multiple linear regressions on different dataset when only percentage of sand, silt, and clay content *(SSC)* were used as inputs, and when *SSC* and bulk density, *Bd*, *(SSC+Bd)* were used as inputs. The equation 5, 6, 7, 8 and 9 for K_s, were obtained from multiple linear regressions on 50, 40, 30, 20 and 10 series of dataset when *(SSC)* were used as inputs data respectively, and the equation 10, 11, 12, 13 and 14 for K_s, were obtained from multiple linear regressions on 50, 40, 30, 20 and 10 series of dataset when *(SSC+Bd)* were used as inputs data respectively.

$$K_s=-82.38-14.19(Bd)+0.19(C)+1.84(Si)+4.96(S) \quad (5)$$

$$K_s=-184.43-24.35(Bd)+0.89(C)+3.06(Si)+6.22(S) \quad (6)$$

$$K_s=-290.58-33.33(Bd)+2.08(C)+4.21(Si)+7.42(S) \quad (7)$$

$$K_s=-265.50-30.33(Bd)+1.94(C)+3.77(Si)+6.96(S) \quad (8)$$

$$K_s=-83.57-163.42(Bd)+1.37(C)+3.84(Si)+6.95(S) \quad (9)$$

$$K_s=-115.48+0.01(C)+1.95(Si)+5.11(S) \quad (10)$$

$$K_s=-237.07+1.19(C)+3.20(Si)+6.41(S) \quad (11)$$

$$K_s=-410.10+2.97(C)+4.90(Si)+8.09(S) \quad (12)$$

$$K_s=-386.56+2.88(C)+4.53(Si)+7.65(S) \quad (13)$$

$$K_s=-534.13+4.60(C)+5.90(Si)+8.69(S) \quad (14)$$

Where *C* is the percentage of clay-sized particles, *Si* is the percentage of silt-sized particles, *S* is the percentage of sand-sized particles and *Bd* is the soil bulk density (g/cm^3).

For this case-study, the statistics of different multiple linear regression models related to various number of dataset, when only three *(SSC)* inputs were used and when four *(SSC+Bd)* inputs were used in predicting K_s, are presented in Table 2 and Table 3 respectively. The best model of multiple linear regressions was selected based on the four statistics e.g. Correlation Coefficient *(R)*, Root Mean Square Error *(RMSE)*, Mean Bias Error *(MBE)* and Mean Absolute Error *(MAE)*.

Values of statistics presented in Table 2 and Table 3 show when *SSC* and bulk density, *(SSC+Bd)* were used as inputs

TABLE II
STATISTICS OF VARIOUS EQUATIONS BASED ON DIFFERENT NUMBER OF DATASET FROM (SSC) MODEL

Number of dataset	STATISTICS PARAMETERS			
	R	RMSE	MBE	MAE
50	0.925	15.29	-1.03	12.51
40	0.915	16.71	-1.51	13.73
30	0.882	16.72	-1.86	13.83
20	0.852	17.03	-4.56	13.96
10	0.725	19.62	-9.87	18.91
Mean	0.860	17.07	-3.77	14.59

R is the Correlation Coefficient; RMSE is the Root Mean Square Error; MBE, is the Mean Bias Error; and MAE, Mean Absolute Error.

TABLE III
STATISTICS OF VARIOUS EQUATIONS BASED ON DIFFERENT NUMBER OF DATASET FROM (SSC+Bd) MODEL

Number of dataset	STATISTICS PARAMETERS			
	R	RMSE	MBE	MAE
50	0.927	15.28	-1.11	12.92
40	0.918	16.55	-1.82	13.74
30	0.897	16.75	-1.96	14.32
20	0.849	17.91	-5.30	14.72
10	0.618	24.69	-17.37	22.16
Mean	0.842	18.24	-5.51	15.57

R is the Correlation Coefficient; RMSE is the Root Mean Square Error; MBE, is the Mean Bias Error; and MAE, Mean Absolute Error.

data estimated K_s, almost such when only percentage of sand, silt, and clay content *(SSC)* were used as inputs data. Values of statistics presented in Tables 2 and 3 resulted in under prediction (negative *MBE*) for both methods of *SSC* and *SSC+Bd*.

The results presented in Table 2 and Table 3 showed that when number of dataset increase, the statistics of *RMSE*, *MBE* and *MAE* decrease and statistic of *R* increases, therefore the results of this study showed the equations were obtained from multiple linear regressions on more dataset estimated K_s, better (less error) than equations from less dataset.

IV. CONCLUSION

The research investigated to study the effect of dataset number on precision of estimated saturated hydraulic conductivity. The results of multiple linear regressions showed when *SSC* and bulk density, *(SSC+Bd)* were used as inputs data estimated K_s, almost such when only percentage of sand, silt, and clay content *(SSC)* were used as inputs data. It is concluded that equations were obtained from multiple linear regressions on more dataset estimated K_s, better than equations from less dataset.

ACKNOWLEGMENT

The authors wish to thank the Sama technical and vocational training school, Islamic Azad University, Mahabad Branch, for supporting this research.

REFERENCES

[1] Arya, L. M., Feike J. L., Peter J. S., and M. Th. van Genuchten. 1999. Relationship between the hydraulic conductivity function and the particle-size distribution. Soil Sci. Soc. Am. J., 63, 1063–1070.

[2] Boadu, F. K. 2000. Hydraulic conductivity of soils from grain-size distribution: new models. Journal of Geotechnical and Geoenvironmental Engineering.

[3] Chakraborty, D., *A.* Chakraborty, P. Santra, R. K. Tomar1, R. N. Garg, R. N. Sahoo1, S. G. Choudhury, M. Bhavanarayana and N. Kalra. 2006. Prediction of hydraulic conductivity of soils from particle-size distribution. Current Science, VOL. 90, NO. 11, 1527-1531.

[4] Cheng, C., and Chen, X. 2007. Evaluation of methods for determination of hydraulic properties in an aquifer- aquitard system Hydrologically connected to river. Hydrogeology Journal. 15: 669-678

[5] Cirpka, O. A.2003. Environmental Fluid Mechanics I: Flow in Natural Hydrosystems.

[6] Cronican A. E. and M. M. Gribb. 2004. Literature review: Equations for predicting hydraulic conductivity based on grain-size data. Supplement to Technical Note entitled: Hydraulic conductivity prediction for sandy soils. Published in Ground Water, 42(3):459-464.

[7] Gupta, R. K., R. P. Rudra, W. T. Dickinson, and G. J. Wall. 1992. Stochastic analysis of groundwater levels in a temperate climate. Trans. Am. Soc. Agric. Engrs. 35(4), 1167-1172.

[8] Han, H., D. Gimenez, L. Lilly. 2008. Textural averages of saturated soil hydraulic conductivity predicted from water retention data. Geoderma, 146:121-128.

[9] Hazen, A. 1892. Some physical properties of sands and gravels. Massachusetts State Board of Health, Annual Report, 539-556.

[10] Jabro, J.D. 1992. Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data. Journal of the American Society of Agricultural Engineers 35, no. 2: 557-560.

[11] Jadczyszyn, J. and J. NiedŸwiecki. 2005. Relation of Saturated Hydraulic Conductivity to Soil Losses. Polish Journal of Environmental Studies Vol. 14, No 4, 431-435.

[12] Kumar A., R.S. Kanwar and G. R. Hallberg. 1994. Modelling spatial variability of saturated hydraulic conductivity using Fourier series analysis. Hydrological Sciences Journal. 39 (2): 143-156.

[13] Leij, F., M.G. Schaap, and L.M. Arya. 2002. Water retention and storage: Indirect methods. p. 1009–1045. In J.H. Dane and G.C. Topp (ed.) Methods of soil analysis. Part 4. SSSA Book Ser. No. 5. SSSA, Madison, WI.

[14] Mallants D., Jaques D., Tseng P. H., Van Genuchten M. Th., Feyen J. 1997. Comparison of three hydraulic property measurement methods. J. Hydrol. 199, 295.

[15] Mohanty, B. P., R. S. Kanwar, and R. Horton. 1991. A robust-resistant approach to interpret spatial behavior of saturated hydraulic conductivity of a glacial till soil under no-tillage system. Wtt. Resour. Res. 27(11), 2979-2992.

[16] Odong, J., Hubei, P. and R. China. 2007. Evaluation of empirical formulae for determination of hydraulic conductivity based on grain-size analysis. Journal of American Science, 3(3), 54-60.

[17] Parasuraman, K., A. Elshorbagy, and B. C. Si. 2006. Estimating saturated hydraulic conductivity in spatially variable fields using neural network ensembles. Soil Sci. Soc. Am. J. 70:1851–1859.

[18] Puckett, W.E., J.H. Dane, and B.F. Hajek. 1985. Physical and mineralogical data to determine soil hydraulic properties. Soil Science Society of America Journal 49, no. 4:831-836.

[19] Rawls, W.J. and D.L. Brakensiek. 1989. Estimation of soil water retention and hydraulic properties. Unsaturated flow in Hydrologic Modeling Theory and Practice, ed. H. J.

[20] Shao, M. and Robert, H., 1998. Integral method of soil hydraulic properties. Soil Sci. Soc. Am. J., 62, 585–592.

[21] Shepherd, R.G. 1989. Correlations of Permeability and Grain Size. Ground Water 27, no. 5: 633-638.

[22] Sisson, J. B. and P. J. Wierenga. 1981. Spatial variability of steady state infiltration rates as stochastic process. Soil Sci. Soc. Am. J. 45, 699-704.

[23] Sperry, J.M. and J.J. Peirce. 1995. A model for estimating the hydraulic conductivity of granular material based on grain shape, grain size, and porosity. Ground Water 33, no. 6: 892-898.

[24] Tietje, O., and V. Hennings. 1996. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. Geoderma 69:71–84.

[25] Todd, D. K., and Mays, L.W. 2005.Groundwater hydrology. John Wiley & Sons, New York.

[26] Van Dam J. C., Stricker, J. N. M. and Droogers, P., 1992. Inverse method for determining soil hydraulic function from one-step outflow experiments. Soil Sci. Soc. Am. J., 56, 1042–1050.

[27] Van Genuchten, M. Th. and Leji, F., 1989. On estimating the hydraulic properties of unsaturated soils. In Proceedings of the International Workshop on Indirect Method of Estimating Hydraulic Properties of Unsaturated Soils (eds van Genuchten, M. Th. et al.), 11–13 October, US Salinity Laboratory and Department of Soil and Environmental Science, Univ. of California, Riverside, 1992, pp. 1–14.

[28] Van Genuchten, M.Th., F.J. Leij, and L.J. Lund. 1992. On estimating the hydraulic properties of unsaturated soils. p. 1–14. In M.Th. van Genuchten et al. (ed.) Indirect methods for estimating the hydraulic properties of unsaturated soils. Proc. Int.Workshop. Riverside, CA. 11–13 Oct. 1989. Univ. of California, Riverside.

[29] Vereecken, H., J. Maes, and J. Feyen. 1990. Estimating unsaturated hydraulic conductivity from easily measured soil properties. Soil Sci. 149:1–12.

[30] Vieira, S. R., D. R. Nielsen, and J. W. Biggar. 1981. Spatial variability of field measured infiltration rate. Soil Sci. Soc. Am. J. 45, 1040-1048.

[31] Vukovic, M., and Soro, A. 1992. Determination of hydraulic conductivity of porous media from grain-size composition. Water Resources Publications, Littleton, Colorado.

[32] Wosten, J.H.M. 1990. Use of soil survey data to improve simulation of water movement in soils. Ph. D. thesis. Univ. of Wageningen, the Netherlands.

[33] Wosten, J.H.M., P.A. Finke, and M.J.W. Jansen. 1995. Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics. Geoderma 66:227–237.

[34] Wosten, J.H.M., A. Lilly, A. Nemes, and C. Le Bas. 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90:169–185.

[35] Wosten, J.H.M., Y.A. Pachepsky, and W.J. Rawls. 2001. Pedotransfer functions: Bridging gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 251:123–150.

[36] Zeleke, T.B., and B.C. Si. 2005. Scaling relationships between saturated hydraulic conductivity and soil physical properties. Soil Sci. Soc. Am. J. 69:1691–1702.