EVALUATING AN APPLICABILITY OF NEURAL NETWORKS FOR ESTIMATION OF FIELD CAPACITY AND WILTING POINT IN LOWER KOSI BASIN, INDIA

Ray Singh Meena[1]
[1]Assistant Professor, Department of Civil Engineering, Poornima Group of Institutions Jaipur, India-302022,
Email: ray.meena@poornima.org

ABSTRACT

Soil water content at field capacity (FC) and wilting point (WP) is an important input parameter. Direct measurements of FC and WP are difficult, time-consuming and expensive. This study is conducted to evaluate an applicability artificial neural networks program Rosetta and k-NN for its validity. FC and WP obtained by applying hierarchical rules in Rosetta and k-NN, Rosetta performed well for SSCBDFC (R^2=0.957) and SSCBDFCWP (R^2=0.934) input level and k-NN SSCOM (R^2=0.399) and SSCBDOM (R^2=0.391). Results are significant because FC and WP data are in the development stage for Kosi floodplain.

Keywords: Artificial neural network, field capacity, pedotransfer functions, and wilting point.

1. INTRODUCTION

Researchers studying various phenomena across the soil-water-plant continuum have always been interested in measuring the amount of water a soil can hold. This measure has importance because of its usefulness in a range of fields from hydrology to plant sciences. Many hydrological models require the estimation of soil water content (θ) at FC (-33 kPa) and at WP (-1500 kPa). Soil water is intrinsically linked with dynamic natural processes; removal of water occurs due to drainage, evaporation, and transpiration and addition of water occurs with dewdrops, rainfall, and irrigation (Taylor and Ashcroft, 1972). The movement of water downward does not cease, but continues at a reduced rate for a long time. There is no real value for FC and WP. Therefore, a range of values (soil water contents) are associated with FC and WP. Currently, data-mining techniques are gaining popularity in the PTF-research field with the application of nonconventional statistical methods, e.g., Artificial Neural Networks (ANNs), Classification and Regression Trees (CART), k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), Genetic Algorithms (GA), and Genetic Programming (GP). Many investigators have been used neural network to establish empirical PTFs (Pachepsky, et al., 1996), (Schaap and Leij, 1998), (Schaap, et al., 1998), (Minasny and McBratney, 2002), and (Nemes et al., 2008). An advantage of neural networks over traditional PTFs is that they do not require a priori model concept. The optimal and possibly nonlinear relations that link input data to output data are obtained and implemented in an iterative calibration procedure. As a result, neural network models typically extract the maximum amount of information from the data (Schaap, et al., 2001). In this study Rosetta and K-NN are used to predict FC and WP. To facilitate application of the PTFs, (Schaap, et al., 2001) developed “Rosetta”, a computer program that implements some of the models published by (Schaap and Bouten, 1996), (Schaap and Leij, 1998), (Schaap, et al., 1998) and (Schaap and Leij, 2000). This stand-alone software combines neural network analyses with the bootstrap approach to provide uncertainty estimates of the predicted hydraulic parameters (Efron & Tibshirani, 1993). (Nemes, et al. 2008) developed a computer program for estimating FC and WP from basic soil properties like sand, silt, clay fraction, BD and OC/OM in hierarchical order. The software combines k-NN algorithm with the bootstrap data-subset selection technique to allow the development of model ensembles; that can be used to estimate the uncertainty of the final model. The objective of this study is to evaluate the general applicability and the prediction accuracy of Rosetta and k-NN models to predict FC and WP.

2. STUDY AREA AND DATA COLLECTION

2.1 The Kosi basin

The Kosi basin is an important sub-basin of the Ganga basin. Upper catchment of the basin lies in Nepal and Tibet at great heights of the Himalayan range. The total drainage area of the Kosi River is 74,030 km^2 out of which 11,410 km^2 lies in India and the rest 62,620 km^2 lies in Tibet and Nepal (FMIS, 2013). The location of the lower Kosi basin in India lies between 86°22’24’’- 87°37’40’ East and 25°19’25’’- 26°35’16’’ North. The upper catchment of the Kosi basin lies totally in mountainous region. The soils encountered in these regions are usually classified as (GFCC, 1983): Mountain Meadow Soil, Sub-Mountain Meadow Soil, and Brown Hill Soil. The entire lower area of the Kosi Basin in the plains can be regarded as a large inland delta formed by the huge sandy deposit of the Kosi River. Figure 1 shows the location map of Kosi basin situated in India, main Kosi river and its tributaries extracted from Shuttle Radar Topography Mission-Digital Elevation Model (SRTM-DEM) 30meter resolution.

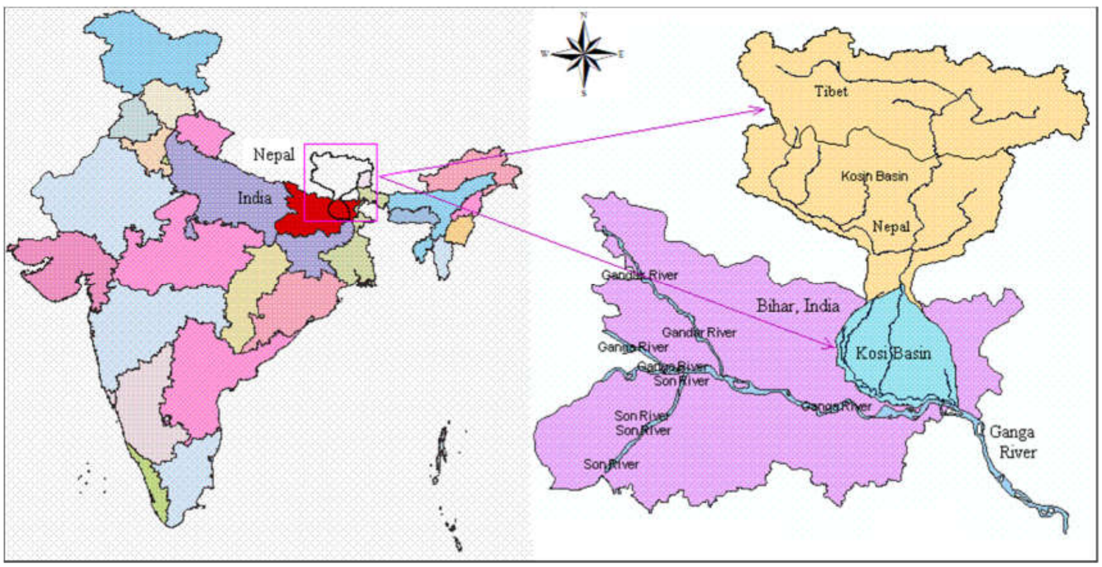


Figure 1: Index Map of Kosi Basin in Tibet, Nepal and India

Disturbed and undisturbed soil core samples have been collected from 14 different locations of the Kosi floodplain (at the depth of 50cm) (Figure 2). Soil physical properties i.e. bulk density, particle size distribution, specific gravity, porosity, organic carbon, and water content at different pressures are calculated in the laboratory.

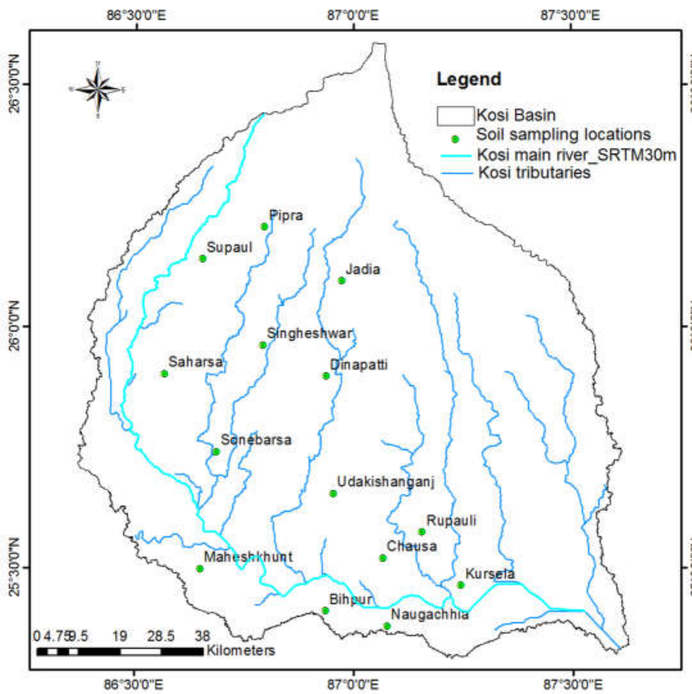


Figure 2. Study area and location map of soil sampling sites

3. METHODOLOGY

3.1 ROSETTA

ROSETTA is able to estimate the van Genuchten (1980) water retention parameters and saturated hydraulic conductivity (K_s), as well as unsaturated hydraulic conductivity parameters, based on Mualem's (1976) pore-size model (Schaap, et al., 2001). In this study four distinct hierarchical levels of input data are used for evaluation of Rosetta model:

1. Input level-3: SSC
2. Input level-4: SSCBD
3. Input level-5: SSCBDFC; and
4. Input level-6: SSCBDFCWP

where, SSC is sand, silt and clay fraction in %, BD is bulk density in g/cc, FC is the water content at field capacity in (m^3/m^3), and WP is the water content at wilting point (m^3/m^3).

VG parameters estimated by each of the hierarchical rule are used for FC and WP estimation. However, Rosetta can be used to obtain parameters of the van Genuchten (1980) soil water retention function alone. Although VG parameters can be converted to parameters of other soil water retention functions, accuracy could be compromised. Thus, it was essential to assess the VG function for its ability to describe SWRC of the Kosi floodplains soils in comparison with other analytical functions.

3.2 k-Nearest Neighbor (k-NN) technique

Nemes et al. (2008) developed k-Nearest Neighbor (k-NN) technique to estimate soil water retention at field capacity and wilting point from basic soil properties i.e. particle size distribution, bulk density, and organic matter in hierarchical order. The technique is based on pattern-recognition rather than on fitting equations of data. The technique is based on pattern-recognition and using similarities rather than on fitting equations to data. Application of the k-

NN means identifying and retrieving the nearest (most similar) instances, based on their input attributes, to the target object from a known set of stored instances (Nemes et al., 2008). Bootstrap technique combined to obtain a measure of uncertainty around the estimate.

FC and WP are estimated using four level of input information:
1. Input level-3: SSC
2. Input level-4: SSCBD
3. Input level-4: SSCOM; and
4. Input level-5: SSCBDOM

where, SSC is sand, silt and clay fraction in %, BD is the bulk density in g/cc, OM is the organic matter in %.

3.3 Performance evaluation

Measured and estimated water retention data are compared and the performance of the functions was evaluated using statistical indices to identify the best-suited function for describing FC and WP of Kosi basin soils. The following statics are used to performance evaluation:

Root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(E_i - M_i)^2}{n}} \tag{1}$$

Mean Absolute Error (MAE)

$$MAE = \sum_{i=1}^{N} \frac{|E_i - M_i|}{N} \tag{2}$$

Maximum absolute error

$$ME = max|E_i - M_i| \tag{3}$$

Index of agreement

$$d = 1 - \frac{\sum_{i=1}^{n}(E_i - M_i)^2}{\sum_{i=1}^{n}(|E_i - \bar{M}| + |M_i - \bar{M}|)^2} \tag{4}$$

Linear correlation coefficient (r)

$$r = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(E_i - \bar{M})(E_i - \bar{E})}{S_M S_E} \tag{5}$$

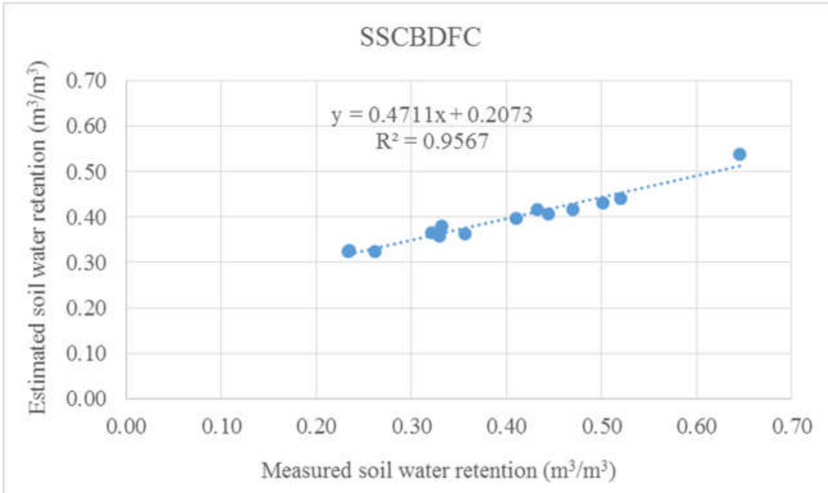
Coefficient of determination (R^2) is the square of the correlation coefficient (r).

where, E_i is estimated water content (m^3/m^3) at pressure h for i_{th} value, M_i is measured water content (m^3/m^3) at i_{th} value, N is total number of observations, $\bar{M}$ is mean of measured water content, $\bar{E}$ is mean of estimated water content, S_M is sum of measured water content and S_E is sum of estimated water content.

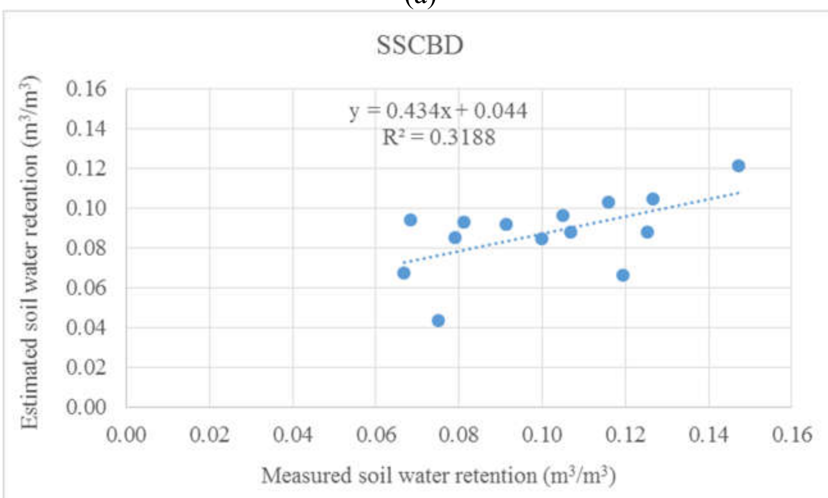
4. RESULTS AND DISCUSSION

4.1 ROSETTA

VG parameters (θ_s, θ_r, α, and n) estimated by each of the hierarchical rule are used for estimation of FC and WP. Water retention estimates by Rosetta, in general improved with the increase in input variables, as reported in earlier studies (Nemes, et al., 2003), (Rawls, et al., 2001), (Wosten, et al., 2001). In this study it is observed that the improvement trend with an increase in the number of predictors was not uniform. Table 1 summarize the different indices for both cases FC and WP. It has been observed that SSCBDFC input level perform well with highest R^2=0.957, lowest RMSE = 0.001 (m^3/m^3) and with d = 0.934. Rosetta model performed better in wet condition (FC) compare to dry condition (WP). Degree of agreement found from 0.799 to 0.974 for FC and WP. Figure 3 shows the comparison of measured and estimated soil water retention at FC and WP.



(a)



(b)

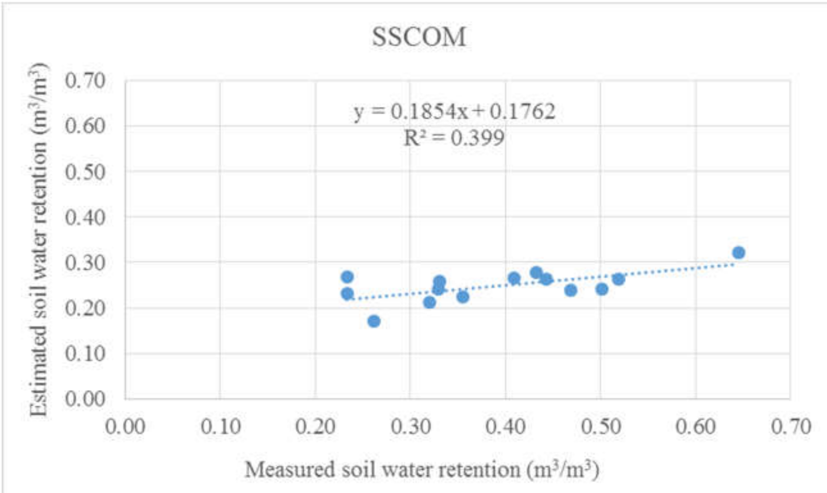
Figure 3. Measured and estimated soil water retention at (a) FC using SSCBDFC and (b) WP using SSCBD by Rosetta

Table 1. Evaluation indices indicating accuracy of Rosetta in predicting FC and WP

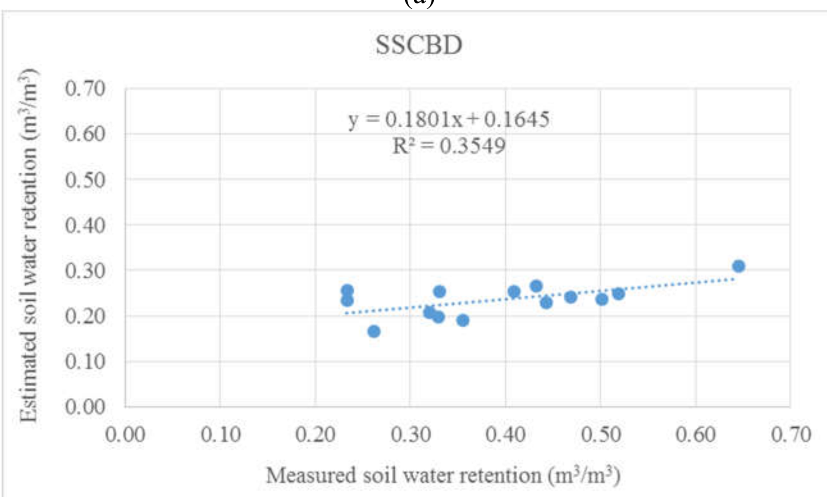
	Field capacity				Wilting point			
Input	SSC	SSCBD	SSCBDFC	SSCBDFCWP	SSC	SSCBD	SSCBDFC	SSCBDFCWP
R2	0.438	0.321	0.957	0.934	0.246	0.319	0.216	0.211
r	0.662	0.566	0.978	0.966	0.496	0.565	0.465	0.459
RMSE	0.182	0.457	0.001	0.013	0.009	0.048	0.498	0.638
MAE	0.084	0.133	0.054	0.049	0.020	0.019	0.133	0.170
ME	0.135	0.045	0.093	0.097	0.044	0.026	0.367	0.415
d	0.885	0.844	0.934	0.937	0.974	0.971	0.826	0.799

4.2 k-Nearest Neighbor (k-NN) technique

Field capacity and wilting point estimated using k-NN technique using soil physical properties. Four type of inputs; SSC, SSCBD, SSCOM and SSCBDOM are provided to estimate the water retention at FC and WP. SSCOM input level showed highest R^2 = 0.399 for field capacity, further addition of BD could not increase the performance for FC estimation. Figure 4 shows the comparison between measured and estimated soil water retention at FC and WP.



(a)



(b)

Figure 4. Measured and estimated water retention at (a) FC using SSCOM, and (b) WP using SSCBD by k-NN

Table 2. Evaluation indices indicating accuracy of Rosetta in predicting FC and WP

	Field capacity				Wilting point			
Input/Index	SSC	SSCBD	SSCOM	SSCBDOM	SSC	SSCBD	SSCOM	SSCBDOM
R^2	0.213	0.355	0.399	0.391	0.010	0.350	0.086	0.373
r	0.461	0.596	0.632	0.625	0.102	0.592	0.294	0.611
RMSE	0.399	0.585	0.534	0.597	0.064	0.137	0.135	0.144
MAE	0.123	0.160	0.148	0.162	0.026	0.037	0.038	0.039
ME	0.076	0.022	0.035	0.016	0.025	-0.002	0.013	-0.005
d	0.847	0.824	0.832	0.823	0.963	0.955	0.953	0.954

5. CONCLUSIONS

The neural network PTFs Rosetta and K-NN are evaluated to predict FC and WP for the soils of Kosi floodplains India. These PTFs are evaluated using basic soil information sand, silt, clay, BD and OC/OM. Water retention estimates by Rosetta, in general, improved with the increase in input variables (Schaap et al. 2001). In this study the performance was not uniform, with increasing in input levels, Rosetta model performed well for the combination of SSCBDFC (R^2 = 0.957) and SSCBDFCWP (R^2 = 0.934) for FC estimation. The k-NN technique also performed better in FC estimation than WP. This method could be useful to provide information of water content at FC and WP, for crop-water management and hydrological models.

6. REFERENCES

Efron B, Tibshirani R.J. (Eds.). (1993) An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability, Chapman and Hall, New York.

Flood Management Information System (FMIS). (2013). Water Resources Department Government of Bihar, Flood Report.

Ganga Flood Control Commission (GFCC), Government of India. Report on Comprehensive Plan of Flood Control for the Kosi Sub-Basin (1983)

Minasny B, McBratney AB (2002) The Neuro-m method for fitting neural network parametric pedotransfer functions. Soil Science Society of America Journal 66: 352-361.

Mualem Y (1976) A new model predicting the hydraulic conductivity of unsaturated porous media. Water Resources Research 12: 513–522.

Nemes A, Roberts RT, Rawls WJ, Pachepsky YA, van Genuchten MT (2008) Software to estimate -33 and -1500 kPa soil water retention using the non-parametric k-nearest neighbor technique. Environmental Modelling & Software 23(2): 254-255.

Pachepsky YA, Timlin D, Varallyay GY (1996). Artificial neural networks to estimate soil water retention from easily measurable data. Soil Science Society of America Journal 60: 727-773.

Rawls WJ, Brakensiek DL (1989) Estimation of soil water retention and hydraulic properties. In Unsaturated flow in hydrologic modeling. Springer Netherlands 275-300.

Schaap MG, Bouten W (1996) Modeling water retention curves of sandy soils using Neural Networks. Water Resources Research 32: 3033-3040.

Schaap MG, Leij FJ (2000) Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten model. Soil Science Society of American Journal 64: 843-851.

Schaap MG, Leij FJ, van Genuchten MTh (1998) Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Science Society of America Journal 62: 847-855.

Schaap MG, Leij FJ, van Genuchten MTh (2001) Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. Journal of Hydrology 251: 163-176.

Schaap MG, Leij FJ (1998) Database-related accuracy and uncertainty of pedotransfer functions. Soil Science 163: 765-779.

Taylor SA, Ashcroft GL (1972) Physical Edaphology: The Physics of Irrigated and Nonirrigated Soils. San Francisco. W.H. Freeman

Van Genuchten MT (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil science society of America journal 44(50): 892-898.