Soil depth modelling using terrain analysis and satellite imagery: the case study of Qeshlaq mountainous watershed (Kurdistan, Iran)

Salahudin Zahedi,[1] Kaka Shahedi,[2] Mahmod Habibnejad Rawshan,[2] Karim Solimani,[2] Kourosh Dadkhah[3]

[1]Kurdistan Agricultural and Natural Resources Research and Education Centre, AREEO, Sanandaj; [2]Watershed Management Department, Sari University of Agricultural Sciences and Natural Resources, Sari; [3]Department of Statistics, Faculty of Sciences, University of Kurdistan, Sanandaj, Iran

Abstract

Soil depth is a major soil characteristic, which is commonly used in distributed hydrological modelling in order to present watershed subsurface attributes. This study aims at developing a statistical model for predicting the spatial pattern of soil depth over the mountainous watershed from environmental variables derived from a digital elevation model (DEM) and remote sensing data. Among the explanatory variables used in the models, seven are derived from a 10 m resolution DEM, namely specific catchment area, wetness index, aspect, slope, plan curvature, elevation and sediment transport index. Three variables *landuse*, *NDVI* and *pca*1 are derived from Landsat8 imagery, and are used for predicting soil depth by the models. Soil attributes, soil moisture, topographic curvature, training samples for each landuse and major vegetation types are considered at 429 profiles within four sub-watersheds. *Random forests* (*RF*), *support vector machine* (*SVM*) and *artificial neural network* (*ANN*) are used to predict soil depth using the explanatory variables. The models are run using 336 data points in the calibration dataset with all 31 explanatory variables, and soil depth as the response of the models. Mean decrease permutation accuracy is performed on Variable selection. Testing dataset is done with the model soil depth values at testing locations (93 points) using different efficiency criteria. Prediction error is computed for both the calibration and testing datasets. Results show that the variables *landuse*, *specific surface area*, *slope*, *pca1, NDVI* and *aspect* are the most important explanatory variables in predicting soil depth. RF and SVM models are appro-priate for the mountainous watershed areas that have been limited in the depth of the soil and ANN model is more suitable for watershed with the fields of agricultural and deep soil depth.

Introduction

Soil depth or regolith depth is defined as the depth from the land surface to un-weathered bedrock (Kuriakose *et al*., 2009; Mehnatkesh *et al*., 2013; Sarkar *et al*., 2014; Yang *et al*., 2014). The spatial distribution of soil depth is controlled by complex interactions of many factors, including topography, parent material, climate, biological, chemical and physical processes (Catani *et al*., 2010). Soil is related to topography and land cover due to the role played by topography and vegetation in affecting soil forming processes (Moore *et al*., 1993). Vegetation, landuse pattern and surface material are other important parameters of the soil depth estimation (Sarkar *et al*., 2014). Estimation of soil depth is crucial for the assessment of watershed hydrological processes (Tromp van Meerveld and Mcdonnell, 2006). Statistical approach is one of the most important methods for predicting soil depth using different environmental variables. Tesfa *et al*. (2009) applied two statistical prediction models, random forest (RF) and general additive model to identify the relationship between soil depth and landscape topographic variables derived from a digital elevation model (DEM) and land cover attributes extracted from satellite imagery. Results showed that these models could explain about 50% of the observed soil depth variability in an out of sample data on study area. In a case study by Kuriakose *et al*. (2009), landuse and land cover maps served as a predictor for predicting soil depth using different statistical methods. The results showed that landuse parameter was the most effective explanatory variable for estimating the depth of the soil. In order to determine the relationship between the soil depth and terrain attributes using multiple linear stepwise regressions model, Yang *et al*. (2014) conducted a study over Guohua Karst Ecological Experimental Area in Southwest China. The results showed that slope, TWI, and elevation can explain about 61.4% of the total variability in soil depth. Mehnatkesh *et al*. (2013) in a study over hilly regions in western Iran showed that slope, wetness index, catchment area and sediment transport index can explain about 76% of total variability in soil depth at the selected site. The main objective of this study is to determine the relationship between soil depth and environmental variables derived from DEM, satellite images and field data measured to predict soil depth variation on a semi-arid hard accessible mountainous watershed in west of Iran. Our study also aims at determining the most effective statistical prediction model in the study area.

Materials and methods

Study area

The study area is located on between 46° 45' and 46° 57' E, and 35° 25' and 35° 38' N in the North part of Sanandaj city, Kurdistan province, Iran, and covers an area of about 27,000 ha (Figure 1). Elevation in the study area varies from 1550 to 2850 m above sea level. This area is constituted of Chile, Andesite and Calcareous rocks. Except for approximately 4% of its which is hilly, all domains are mountainous, and sediment of the rivers in form of plain are seldom. In these mountainous areas due to the high slope alluvium, farmlands have very little expansion. The main landuse of this area is therefore rangelands with more than 23,000 ha.

Field data

Four sub-basins were selected to show the topographic and land cover variability present within the study area. Surveys were done at 436 points within the four sub-watersheds. At each survey place the *GPS* location was recorded. To determine the depth of the soil profile, the digging method was used, in some cases along with a metal bar with 2 meters in length and 2 cm in diameter. Topographic curvature was recorded at sample points as concave (–1), convex (1) or intermediate (0). Landuse and training samples for each landuse class were recorded. Vegetation classification was done using fieldwork with Physiognomic-Floristic method along with vegetation sampling in a Stratified random sampling form.

The percentage of the area covered by vegetation was determined within 1×1 meter plots, and rangelands were divided to three classes named dense, semi- and low-dense range. Due to the large number of data points and high variety of measured parameters, and with regard to the easier digging profiles after precipitation, field data collection was done at three different dates. The first phase of the surveys was performed in May 2015, the second one in November 2015, and the last one in April 2016. The soil depth survey was carried out during early and mid-spring and mid-autumn when the soil was moist, and it was easier to dig the soil profile or penetrate by rod.

Geospatial data

The 1:25,000 scale topography maps from *National Survey Organisation* (*NSO*) was used for preparing a DEM with 10-meter grid resolution using *ARCGIS* 10.3 software. The Landsat8 image of May 6, 2015, path 167, row 35, was used to derive various land cover attributes.

Data derived from digital elevation model

Terrain attributes accessible from DEM, which have potential for soil depth estimation in a watershed, are divided into two primary and secondary categories (Moore *et al.*, 1993). Primary attributes are extractable directly from DEM. Some of the most important of them are elevation, slope, aspect, specific catchment area, plan curvature, and profile curvature (Wilson and Gallant, 2000). Secondary or compound attributes involve combinations of the primary attributes. These indices describe the spatial variability
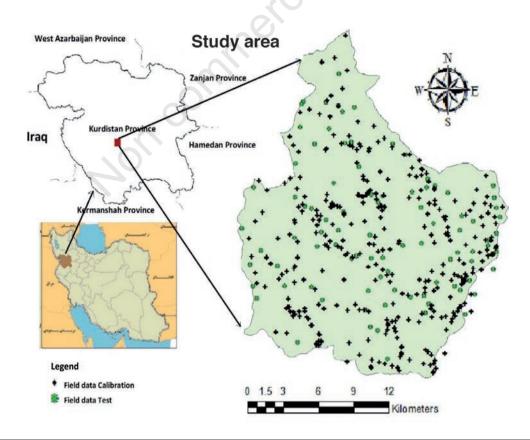

Figure 1. Study area and distribution of soil depth sampling site calibration and test data.

of specific processes, which are occurring in the watershed, such as soil water content or the potential for erosion (Moore *et al*., 1993). Some of the secondary categories used in this study are stream power index (Florinsky *et al*., 2002), wetness index and sediment transport index (Moore *et al*., 1993; Wilson and Gallant, 2000), the D8 (deterministic eight-node) algorithm and its derivatives (D8 contributing area, D8 slope, D8 distance to stream, D8 longest upslope length, D8 total upslope length, D8 slope averaged and D8 flow direction grid), and finally D∞ algorithm and its derivatives (D∞ flow direction and D∞ slope) (Tarboton, 1997; Tesfa *et al*., 2009; Hass, 2010) (Table 1).

Data derived from satellite image

The images were pre-processed for both radiometric calibration and geometric corrections. A thematic map of landuse was created through supervised classification of the Landsat8 image. *Principal component analysis* (*pca*) was used to identify orthogonal components from the Landsat8 input bands that explain significant variance (Pearson, 1901). Different vegetation indices such as *soil adjust vegetation index* (*SAVI*) (Huete, 1988), *normalised difference vegetation index* (*NDVI*) (Rouse *et al*., 1973), *soil moisture index* (*SMI*) (Lingli and Qu, 2009), *vegetation index* (*VI*) and *canopy cover* (*CC*) (Tesfa *et al*., 2009), were used and analysed to reveal the relationship between these indices and soil depth. The *Tasseled Cap transformation* (*tc*) was used to convert the Landsat8 bands into three components (*tc1, tc2, tc3*) designated as brightness, greenness and wetness (Tesfa *et al*., 2009) (Table 2).

Statistical analysis

Normalisation

Box Cox transformations were used to transform the measured soil depth and all explanatory variables. The results showed that their distribution was close to normal.

Models

We applied three types of prediction methods, *i.e*. RF, artificial neural network (ANN), and support vector machines (SVMs) to predict the soil depth using the explanatory variables.

Random forest

Random forests is a statistical classification and regression model that combines many classification and regression trees such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. Each tree is built from a bootstrap sample drawn from the training dataset with replacement. The RF package in the R software was used in this study to develop RF prediction models (Breiman, 2001).

Artificial neural network

In modelling and forecasting nonlinear and impermanent time series of processes where there is no exact solution and clear relationship to recognise and describe them, artificial neural networks have shown good performance. In this study, we apply a *back propagation neural network*, a multi-layer feed-forward neural network capable of predicting based on topography and land cover data sources (Mustafa *et al*., 2012).

Table 1. Data derived from the digital elevation model.

Variable	Definition
Elevation (elev.)	Elevation above sea level
Aspect (asp)	The direction of topographic slope faces in terms of degrees from the north
Slope (slp.)	Topographic slope of survey's points
Specific catchment area (sca)	From the D∞ method: computed by contributing area divided to the grid cell size
Profile curvature (prcurve)	Curvature of the surface in the direction of maximum slope
Plan curvature (plancurve)	The curvature of the surface perpendicular to the direction of the maximum slope
General curvature (gcurve)	The second derivative of the surface
D∞ flow direction	Direction of the steepest outwards slope from the triangular facets centered on each grid cell and is reported as the angle in radians counter-clockwise from east
D8 Contributing Area (d8a)	Number of grid cells draining through each grid cell using the single flow direction model
D8 slope (d8s)	Steepest outwards slope from a grid cell to one of its eight neighbors reported as drop/distance
D8 Distance to Stream (d8dis)	Horizontal distance from each grid cell to a stream grid cell traced along D8 flow directions
D∞ slope	The steepest outwards slope from the triangular facets centered on each grid cell
D8 Longest Upslope Length (D8LUL)	The length of the flow path from the furthest cell that drains to each cell along D8 flow directions
D8 Total Upslope Length (D8TUL)	The total length of flow paths draining to each grid cell along D8 flow directions
D8 Slope averaged (D8SA)	Slope averaged over a 100 m path traced downslope along D8 flow directions
D8 flow direction grid (D8FDG)	Representing the flow direction from each grid cell to one of its adjacent or diagonal neighbors, encoded as 1 to 8 counter-clockwise starting at east
Wetness inverse index (WII)	An index calculated as slope/specific catchment area
Wetness index (Compound topographic index)	Sets catchment area in relation to the slope gradient. It shows the extent of flow accumulation
Sediment transport index (STI)	Effects of topography on erosion and soil loss

Support vector machines

Support vector machine has recently been introduced as a relatively new statistical learning technique. Due to its strong theoretical statistical framework, SVM has proved much more robust in several fields. It has brought forth high expectations in the past few years, as it has been successful in classification problems, regression and forecasting. This is due to the fact that it includes aspects and techniques from machine learning, statistics, mathematical analysis and convex optimisation. The support vector regression solves the estimation problem by mapping between an input and output data set (Sujay and Paresh, 2014).

Variable selection and model complexity

To evaluate suitable model complexity, a completely random split method was used to split calibration data points into two parts, the training and validation datasets as illustrated in Figure 2. The average of permutation importance index was computed by 50 RF model runs for all variables and field data point (Genuer *et al.*, 2010). All computations were done by the random Forest package in the R system programming language. The default parameters in the R package random forest were used, *i.e.* n=336, p=32, mtry=p/3=11, ntree=1000, and nfor=50. The number of explanatory variables in the model is a measure of the model complexity. In order to compute the optimum model complexity, a variation of mean square prediction error for training and validation of datasets in terms of the number of input variables was computed for the three prediction models (RF, SVM and ANN).

Testing the models

All models were used to predict soil depth for the study area with topographic variables. Then the testing data points with the model soil depth were compared using the following model evaluation methods (Nash and Sutcliffe, 1970; Krause *et al.*, 2005).

Coefficient of determination r^2: is defined as the squared value of the coefficient of correlation according to Bravais-Pearson. It is calculated as:

$$R = \frac{\sum_{i=1}^{n}((y_{obs})_i - (\bar{y}_{obs})) \times ((y_{est})_i - (\bar{y}_{est}))}{\sqrt{\sum_{i=1}^{n}((y_{obs})_i - (\bar{y}_{obs}))^2 \sum_{i=1}^{n}((y_{est})_i - (\bar{y}_{est}))^2}} \qquad (1)$$

Root-mean-square error (RMSE): is a frequently used measure of the difference between the values (sample and population values) predicted by a model or an estimator and the values actually observed. It is calculated for the data set as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}((y_{est})_i - (y_{obs})_i)^2}{n}} \qquad (2)$$

Mean absolute error (MAE): is another useful measure widely used in model evaluations. It is calculated for the data set as:

$$MAE = \frac{\sum_{i=1}^{n}|(y_{est})_i - (y_{obs})_i|}{n} \qquad (3)$$

Nash-Sutcliffe efficiency coefficient (NSE): is calculated as:

$$E = 1 - \frac{\sum_{i=1}^{n}(y_{obs,i} - y_{est})^2}{\sum_{i=1}^{n}(y_{obs,i} - \overline{y_{obs}})^2} \qquad (4)$$

where in all four equations y_{obs} is the observed value, is the mean of the observed values, y_{est} is the estimated or modelled value, is the mean of the estimated or modelled values, and *n* is the number of data points.
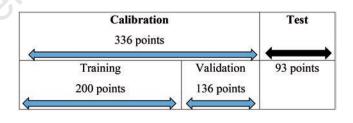

Figure 2. Data division: training, validation and testing sets.

Table 2. Data derived from satellite images.

Variable	Definition
Landuse	Landuse map derived from satellite image and field data
PCA1	First principal component analysis
PCA2	Second principal component analysis
PCA3	Third principal component analysis
TCA1	First tasseled cap component analysis (brightness index)
TCA2	Second tasseled cap component analysis (greenness index)
TCA3	Third tasseled cap component analysis (wetness index)
NDVI	Normalised difference vegetation index
VI	Vegetation index
CC	Canopy cover index
SAVI	Soil adjust vegetation index
SMI	Soil moisture index

Results and discussion

Variable selection and model complexity

The most important variable for soil depth prediction is landuse. Another five important variables from the most to the least important are specific catchment area, NDVI, slope, aspect, and pca1 respectively (Figure 3). The least validation error for the RF, SVM and ANN models occurred with 15, 18 and 18 input variables respectively. After variable 13, validation error in all 3 models began to fluctuate. Therefore 13 explanatory variables were selected as the optimum complexity of models. Training and validation mean square errors for RF and SVM models were rather similar. Both errors decreased up to the first 10 variables as the input variables increased gradually. After that there were minor swings when input variables increased. In the SVM model, in both training and validation data, increases accrued with a steady routine from variable 25. In the ANN model with first 17 input variables validation error decreased, then with more input variables for the validation model the error increased slightly. In the training model, the error decreased until up to variable 25. After that there were minor swings when input variables increased (Figure 4).

Model evaluation

RF, ANN and SVM models were developed using variables with all calibration data points. Figure 5 shows the scatter plots of RF (A), ANN (B) and SVM (C) predicted *versus* the measured soil depth respectively, for the calibration data. In this figure the central lines represent the 1:1 line (predicted = measured). The two diverging dash lines, above and below the 1:1 line, show the predicted soil depth with five standard errors representing 95% confidence intervals. These lines diverge as a result of the Box-Cox back transformation. Figure 6 shows similar scatter plots for the testing data that was not used in models development. The two models SVM and ANN were compared to RF model in this study. Applying Nash-Sutcliffe efficiency (E), RF, SVM and ANN models showed to be able to explain respectively 83, 79 and 62% of the modelled soil depth variability in a calibration of models. These values were 69, 62 and 58 in an out of sample test (Table 3). Applying the adjusted coefficient of determination (r^2), RF, SVM and ANN models showed to be able to explain respectively 83, 82 and 79% of the modelled soil depth variability in a calibration of models. These values were 75, 72 and 65 in an out of sample test (Table 3). R^2, MAE and E indicate that RF model is better than the two other models for predicting soil depth at point scale for both calibration and out of sample statistical data test. But results of the three models were approximately equal for the root mean squared errors (RMSE) criterion. Figures 4 and 5 indicate that the RF model underestimates the soil depth for 0-30 cm, overestimates for 30-50 cm, and underestimates for above 60 cm. SVM model
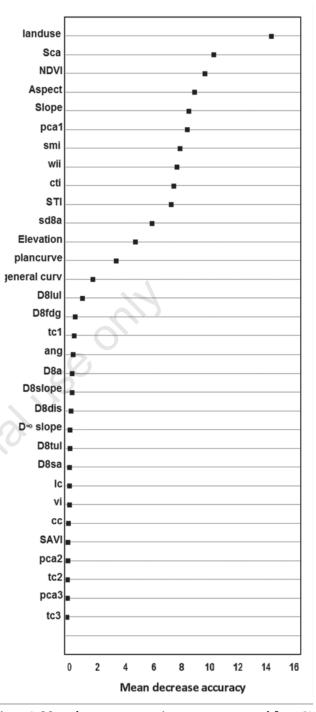

Figure 3. Mean decrease permutation accuracy averaged from 50 random forests model runs.

Table 3. Results of different efficiency criteria for statistical prediction models.

Efficiency criteria	RF model		ANN model		SVM model	
	Calibration	Test	Calibration	Test	Calibration	Test
R2	0.83	0.75	0.82	0.72	0.79	0.65
RMSE	10.8	10.4	12.1	10.4	13.5	10.4
MAE	9.9	12.7	11.5	14.7	12.6	14
E	0.83	0.69	0.79	0.58	0.74	0.62

RF, random forests; ANN, artificial neural network; SVM, support vector machine.

underestimates the soil depth less than 10 cm and overestimates for 10-40 cm, but soil depth estimate above 50 cm has a normal dispersal. Unlike the other two models, predictions of different soil depths were consistent using the ANN model. RF and SVM models testing data could not estimate soil depths of more than about 100 cm. It can be concluded that RF and SVM models are appropriate for the mountainous watershed areas that are limited in the depth of the soil and ANN model is more suitable for watershed with agricultural fields and high soil depth. Mean error in both model calibration and model test of RF model is less than with the other two models. This is in line with the result of the study by Tesfa *et al*. (2009). Statistical prediction models were applied to predict soil depth over a mountainous watershed using different environmental variables derived from DEM and satellite image. These variables included 10 topographic variables: specific catchment area (*sca*), Aspect (*asp*), slope (*slp*), wetness inverse index (wii), wetness index (wi), sediment transport index (sti), D8 Slope averaged (D8SA), elevation, plan curvature, D8 Longest Upslope Length (D8 LUL) and general curvature. Three variables extracted from satellite image are landuse map (lum), NDVI, and first principal component analysis (pca1). Six most important variables for predicting soil depth were landuse map (lum), *specific catchment area*, wetness index, NDVI, aspect and slope gradient. Results indicated that topographic variables were generally more important than satellite extracted variables in predicting and mapping of soil depth in our study of the mountainous watersheds in west of Iran. The importance of the top 3 variables landuse, slope, and D8 Slope averaged (D8SA) in estimating the depth of the soil reveals the impact of human intervention and management on this parameter, and have negative correlation with soil depth. Because of proximity to villages and residential areas, class3 rangelands are destructed due to the negative factors such as grazing livestock, foliage harvesting, fires and change to rainfed landuse. Therefore, the depth of the soil is low in comparison with class2 rangelands, which have greater distance from these areas. Rangelands class1 is much further away from residential areas and due to the topographical situation, it is hard to access them. For this reason, the soil depth is higher in this class. The soil depth in protected and

Figure 5. Predicted and measured soil depth data calibration with 5 SD: A) random forests (RF); B) artificial neural network (ANN); C) support vector machine (SVM).
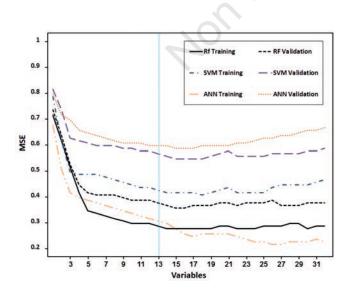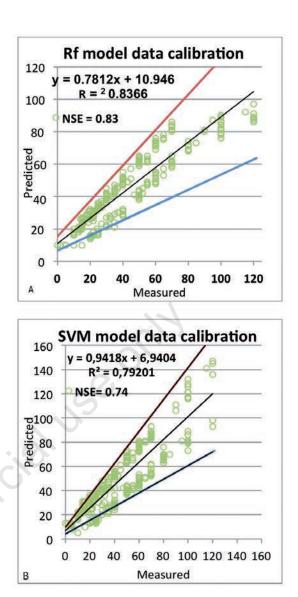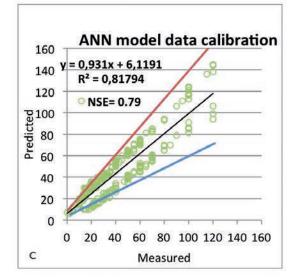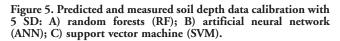

Figure 4. Model complexity (number of input variables) against mean square error. RF, random forests; SVM, support vector machine; ANN, artificial neural network.

enclosure areas is also significant. A large part of class2 rangelands has been converted to rainfed. But due to plough in the direction of the steep slope and thus an increased soil erosion rate, these areas have lost their soil depth. In the course of a few years, these areas have been abandoned due to their reduced production capability, and classified as bare lands. Therefore, the effects of human intervention variable (landuse) in relation to the depth of the soil indicate that abandoned rainfed located on gradient more than 50% have the second lowest value of the soil depth after the edges and rocks. This is consistent with the literature (Kuriakose *et al*., 2009; Sarkar *et al*., 2014; Yang *et al*., 2014). Slope and D8 Slope averaged (D8SA) quantify relationships between soil depths and slope. The importance of the variables *sca*, *wi*, plan curvature, and wetness inverse index reveals the role of flat and concave areas in preservation of surface runoff due to the higher soil depth in these areas. Similar result was also reported by other authors (Tesfa *et al*., 2009; Mehnatkesh *et al*., 2013; Sarkar *et al*., 2014; Yang *et al*., 2014). The depth of the soil in the northern slope is more than other directions due to having more moisture and less sunny hours. Minimum depth of soil is at the southern slope. This is consistent with the results reported by Penížek and Borůvka (2006) and Tesfa *et al*. (2009). The three major indices extracted from satellite images: landuse map, NDVI and pca1 are among the important variables to estimate the depth of the soil. This shows that the importance of vegetation in forecasting soil depth in the study area, which has also been reported by previous researchers (Tesfa *et al*., 2009; Gastaldi *et al*., 2012; Seid *et al*., 2013). Carriero *et al*. (2005) showed that there was a fairly good correlation between soil depth and slope gradient, wetness index and mean annual precipitation for the study basin. They showed that this correlation can be improved, if information about vegetation characteristics is added.

Conclusions

In this study, RF, ANN and SVM models were developed using environmental variables derived from DEM and satellite image in order to soil depth prediction. Because of the topography of the study area, RF model has the best performance than other two models. The results showed that landuse is the most important variable in soil depth prediction for the case study. Optimum quantity (point sample number) and quality (method of sampling and point samples distribution on study area) of field data are significant parameters for accuracy of soil prediction models. In comparison with prior researches, a large number of field data with adequate distribution on study area was used in order to improve the accuracy of prediction in this study. The study was performed on mountainous areas with an area of 270 square kilometres. 336 samples were used for calibration and 93 were used for testing the models. On average, one sample was picked per 0.63 km², which means that the number of samples per area unit is higher in comparison to previous studies with similar study areas (Mehnatkesh *et al*., 2013; Sarkar *et al*., 2014). For sampling, profile drilled method and metal bar as an ancillary tool were used. This increased the accuracy of the results in comparison with previous studies (Tesfa *et al*., 2009; Mehnatkesh *et al*., 2013). In our study, variable selection for all the models was performed based on the RF model. Therefore, a possibility for further work is to perform specialised variable selection for the other two models. In addition, the models developed in this study should be used and validated in other mountainous watersheds with similar environmental conditions to evaluate its overall accuracy for model transportability.
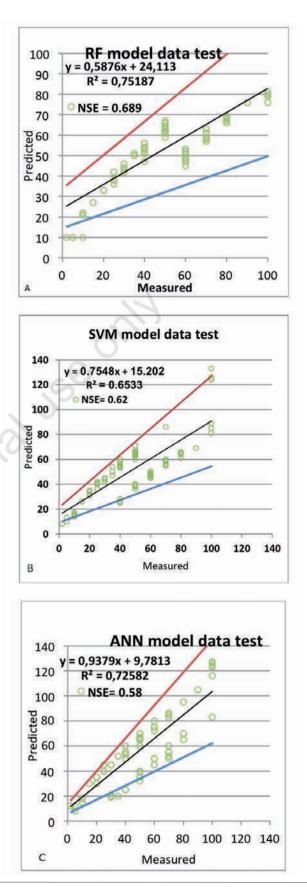

Figure 6. Predicted and measured soil depth data test with 5 SD: A) random forests (RF); B) artificial neural network (ANN); C) support vector machine (SVM).

References

Breiman L. 2001. Random forests. Machine Learning. 45:5-32.

Carriero D., Romano N., Fiorntion M. 2005. A simplified approach for determining hydrological behavior and depth of soils at basin scale. J. Agr. Eng. 2:1-10.

Catani F., Segoni S., Falorni G. 2010. An empirical geomorphology-based approach to the spatial prediction of soil thickness at catchment scale. Water Resour. Res. 46:W05508.

Florinsky I.V., Eilers R.G., Manning G.R., Fuller L.G. 2002. Prediction of soil properties by digital terrain modeling. Environ. Model. Software 17:295-311.

Gastaldi G., Minasny B., McBratney A.B. 2012. Mapping the occurrence and thickness of soil horizons within soil profiles. Digital soil assessments and beyond. Taylor & Francis Group, London, UK.

Genuer R., Poggi J.M., Tuleau-Malot C. 2010. Variable selection using random forests. Pattern Recogn. Lett. 31:2225-36.

Hass J. 2010. Soil moisture modeling using TWI and satellite imagery in the Stockholm region. Royal Institute of Technology (KTH), Stockholm, Sweden.

Huete A.R. 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25:295-309.

Krause P.D., Boyle P., Base F. 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5:89-97.

Kuriakose S.L., Devkota S., Rossiter D.G. 2009. Prediction of soil depth using environmental variables in an anthropogenic landscape, a case study in the Western Ghats of Kerala, India. Catena. 79:27-38.

Lingli W., Qu J.J. 2009. Satellite remote sensing applications for surface soil moisture monitoring: a review. Front. Earth Sci. 3:237-47.

Mehnatkesh A., Ayoubi S., Jalalian A., Kanwar S. 2013. Relationships between soil depth and terrain attributes in a semi-arid hilly region in western Iran. J. Mount. Sci. 10:163-72.

Moore I.D., Gessler P.E., Nielsen G., Peterson G.A. 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57:443-52.

Mustafa M.R., Isa M.H., Rezaur R.B. 2012. Artificial neural networks modeling in water resources engineering, infrastructure and applications. Int. J. Civil Environ. Struct. Constr. Architect. Eng. 6:128-36.

Nash J.E., Sutcliffe J.V. 1970. River flow forecasting through conceptual models. Part 1 - A discussion of principles. J. Hydrol. 10:282-90.

Pearson K. 1901. On lines and planes of closest fit to systems of points in space. Philos. Magaz. 2:559-72.

Penizek V., Boruka L. 2006. Soil depth prediction supported by primary terrain attributes: a comparison of methods. Plant Soil Environ. 52:424-30.

Rouse J.W., Haas R.H., Schell J.A., Deering D.W. 1973. Monitoring vegetation systems in the Great Plains with ERTS. Int. 3rd ERTS Symposium, NASA SP- 351 I, pp. 309-317.

Sarkar S., Archana K.R., Tapas R.M. 2014. Soil depth Estimation through soil and landscape modeling using regression kriging in a Himalayan terrain. Int. J. Geogr. Inf. Sci. 27:2436-54.

Seid N.M., Yitaferu B., Kibebew K., Ziadat F. 2013. Soil-landscape modeling and remote sensing to provide spatial representation of soil attributes for an Ethiopian watershed. Appl. Environ. Soil Sci. 2013:798094.

Sujay R.N., Paresh C.D. 2014. Support vector machine applications in the field of hydrology: a review. Appl. Soft Comput. 19:372-86.

Tarbaton D. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resour. Res. 33:309-19.

Tesfa T.K., Tarbaton D.G., Chandler D.G., Macnamara J.P. 2009. Modeling soil depth from topographic and land cover attributes. Water Resour. Res. 45:W10438.

Tromp-van Meerveld H.J., McDonnell J.J. 2006. Threshold relations in subsurface storm flow: 2. the fill and spill hypothesis. Water Resour. Res. 42:W02411.

Wilson J.P., Gallant J.C. 2000. Terrain analysis, principals and applications. John Willey & Sons, New York, NY, USA, pp 87-131.

Yang Q., Fawang Z., Zhongcheng J., Wenjun L., Jianbing Z., Faming Z., Hui L. 2014. Relationship between soil depth and terrain attributes in karst region in southwest China. J. Soil Sedim. 14:1568-76.