

Soil–Landscape Estimation and Evaluation Program (SLEEP) to predict spatial distribution of soil attributes for environmental modeling

Feras M Ziadat^{1*}, Yeganantham Dhanesh², David Shoemate³,
Raghavan Srinivasan³, Balaji Narasimhan⁴, Jaclyn Tech³

(1. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy; 2. Department of Civil Engineering, Texas A&M University, College Station, USA; 3. Spatial Sciences Laboratory, Texas A&M University, College Station, USA; 4. Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India)

Abstract: The spatial distribution of surface and subsurface soil attributes is an important input to environmental modeling. Soil attributes represent an important input to the Soil and Water Assessment Tool (SWAT), which influence the accuracy of the modeling outputs. An ArcGIS-based tool was developed to predict soil attributes and provide inputs to SWAT. The essential inputs are digital elevation model and field observations. Legacy soil data/maps can be used to derive observations when recent field surveys are not available. Additional layers, such as satellite images and auxiliary data, improve the prediction accuracy. The model contains a series of steps (menus) to facilitate iterative analysis. The steps are summarized in deriving many terrain attributes to characterize each pixel based on local attributes as well as the characteristics of the contributing area. The model then subdivides the entire watershed into smaller facets (subdivisions of subwatersheds) and classifies these into groups. A linear regression model to predict soil attributes from terrain attributes and auxiliary data are established for each class and implemented to predict soil attributes for each pixel within the class and then merged for the entire watershed or study area. SLEEP utilizes Pedo-transfer functions to provide the spatial distribution of the necessary unmapped soil data needed for SWAT prediction. An application of the tool demonstrated acceptable accuracy and better spatial distribution of soil attributes compared with two spatial interpolation techniques. The analysis indicated low sensitivity of SWAT prediction to the number of field observations when SLEEP is used to provide the soil layer. This demonstrates the potential of SLEEP to support SWAT modeling where soil data is scarce.

Keywords: GIS, remote sensing, terrain analyses, watershed, SWAT, inverse distance weighted, Kriging

DOI: 10.3965/j.ijabe.20150803.1270

Citation: Ziadat F M, Dhanesh Y, Shoemate D, Srinivasan R, Narasimhan B, Tech J. Soil–Landscape Estimation and Evaluation Program (SLEEP) to predict spatial distribution of soil attributes for environmental modeling. Int J Agric & Biol Eng, 2015; 8(3): —.

1 Introduction

Providing information about the vertical and lateral distribution of soil characteristics is a challenging task for most environmental modeling applications^[1]. This is partially due to the complexity of soils and their spatial

distribution and the cost and effort of collecting detailed information. However, the accuracy of soil information determines the accuracy of these applications and the decisions made based on these^[2,3]. The use of small-scale maps is a possibility but the heterogeneous representation of soil characteristics is a limitation^[4,5]. The gradual change in soil characteristics is not perfectly

Received date: 2014-05-15 **Accepted date:** 2015-03-15

Biographies: **Yeganantham Dhanesh**, MS, Research interests: statistical hydrology, scaling, watershed modeling, GIS. Email: dhaneshy@tamu.edu; **David Shoemate**, MS, Research interests: GIS, spatial modeling, natural resources management. Email: shoe2@neo.tamu.edu; **Raghavan Srinivasan**, PhD, Research interests: hydrology, watershed management, GIS. Email: r-srinivasan@tamu.edu; **Balaji Narasimhan**, PhD, title, Research interests: water resources management and planning, remote

sensing and GIS in hydrology. Email: nbalaji@iitm.ac.in; **Jaclyn Tech**, BS, Research interests: Web-services, advanced software development. Email: jbttech@ag.tamu.edu.

***Corresponding author: Feras M Ziadat**, PhD, Research interests: soil-landscape modeling, land evaluation, land degradation and soil conservation. FAO, Viale delle Terme di Caracalla 00153 Rome, Italy. Tel: +39 06 57054079. Email: feras.ziadat@fao.org

reflected in the polygon representation provided by most available traditional soil maps^[6,7]. Therefore, another approach that considers these limitations is needed^[4,8,9].

Topographic variables play an important role in soil differentiation^[5,10]. Soil scientists use qualitative relationships between topography and soil variation in soil mapping^[6,7,9]. Some researchers have used quantitative relationships to estimate the spatial distribution of different soils^[10,11]. The use of GIS and remote sensing provide promising tools to quantify these relationships and aid digital soil mapping efforts through the relationships between soils and topographic and remote sensing variables^[2,3,6,12–16]. Digital elevation models (DEMs) are used to derive many topographic variables, which are used to predict the distribution of soil characteristics^[7,10,17,18].

Many researchers have found satisfactory statistical relationships between different soil attributes and terrain attributes easily derived from DEM. Some of the promising indicators are pH, organic matter, carbonates, particle size distribution, color, bulk density, and depth to specific horizon boundaries^[5,10,19]. Soil depth was significantly correlated ($R^2 = 0.30$) with slope angle and absolute and relative height^[20]. Soil depth and A-horizon depth were correlated with plan curvature, compound topographic index (CTI) and upslope mean plan curvature^[17]. Models that utilized only CTI explained 84% and 71% of variation in soil depth and A-horizon depth, respectively^[19]. Slope and wetness index accounted for half of the variability in A-horizon depth, sand content and other soil properties^[4]. Research indicated that slope, tangential and profile curvatures were good predictors of soil texture^[5]. The increasing availability of high-resolution remote sensing data provides a new window for predicting soil characteristics with acceptable accuracy^[21]. Researchers have provided evidence regarding the contribution of remote sensing data in providing acceptable prediction of soil characteristics^[22–24].

Soil data represent a basic input of the Soil and Water Assessment Tool (SWAT). SWAT is a semi-distributed process based ecohydrological model used to simulate stream flow, crop yield, sediment transport and nutrient transport, which has been applied worldwide across a

broad range of watershed scales and environmental conditions^[25,26]. In-depth descriptions of the theoretical underpinnings of the model have been provided elsewhere^[27,28]. Major inputs needed to setup the SWAT model and simulate hydrologic processes include spatially distributed Digital Elevation Map (DEM), land use and soil data, along with weather data. Cropping system, fertilizer applications and other management data are also important inputs when simulating agricultural-generated diffuse pollution. With the advancement in remote sensing, satellite precipitation data and other data sources, most of these data are becoming increasingly available worldwide for at least coarse resolutions. However, larger gaps exist regarding availability of adequate soil data in many global subregions. Therefore, the need exists for tools to be developed that support easier preparation of soil input data for SWAT.

Currently, the majority of soil data are available as soil maps, in polygon format, which tend to aggregate the individual soil attributes and present soil information in a form of soil classification that generalizes soil variability within one polygon into one value or class. The extraction of layers of individual soil attributes, which also reflect the spatial variability within these polygons is, in most cases, not possible. Environmental modeling as well as other soil applications require proper representation of the spatial distribution of soil attributes and favor the representation of attributes as individual layers for each soil parameter to facilitate the integration with other layers of information. The approach described in this study is designed to help users generate higher resolution soil information to cover areas where soil data is not available or available at low resolution. Using digital elevation model and soil observations, the model generates spatially continuous representation of soil attributes that are available in format ready for use as an input to SWAT. This will also benefit users who are demanding spatially distributed soil information for various applications.

Generally, the prediction accuracy of environmental models, such as SWAT, depends on how well the inputs describe the spatial characteristics of the watershed^[29–33]. For example, the use of different resolutions of soil data, such as the U.S. State Soil Geographic (STATSGO)^[34]

versus the Soil Survey Geographic (SSURGO)^[35] databases, may give different simulation results for water, sediment, and agricultural chemical yields^[36,37]. The effect of the spatial resolution of soil data in the prediction accuracy of runoff and soil erosion was investigated in several previous studies^[29–33,36–39]. Results indicated differences in runoff prediction as a result of using different soil data inputs, with SSURGO-based predictions the most accurate. The results also indicate that the accuracy of modeled runoff is reduced when lower resolution soils data are used and when the model parameters are lumped to larger spatial units of analysis^[38]. The use of STATSGO resulted in assigning a single classification to areas that may have different soil types if SSURGO were used. This resulted in different number and size of HRUs, the number of HRUs generated when STATSGO and SSURGO soil data were used is 261 and 1301, respectively, which influences sediment yield parameters (slope and slope length)^[29]. Therefore, the predicted stream flow was higher when SSURGO was used compared to STATSGO. Furthermore, the predicted sediment and sediment-attached nutrients was less in the case of SSURGO. Therefore, modelers need to select the optimum inputs with a suitable resolution to ensure proper outputs.

Researchers have used different statistical relationships to predict individual soil characteristics or soil classes with promising results^[4,6,17,19]. However, these methods are not reproducible owing to the complexity of applying these relationships and the necessary iterations to reach an acceptable result. Therefore, automation of these analyses and the ease of running several iterations in a relatively short time, through a user-friendly program, will aid the application of SWAT (and other models) at larger scales and for various environmental condition^[40]. Thus, the objectives of this work is to describe the Soil–Landscape Estimation and Evaluation Program (SLEEP), a user-friendly GIS-based program, to investigate different options to use SLEEP to provide high resolution soil attribute layers, and to explore the quality of the outputs. This will foster the use of better soil information in many land and water resources management and modeling efforts.

2 Theoretical background

The key SLEEP model software requirements and functions are outlined in Figure 1 and Box 1. SLEEP uses measured soil properties (viz. soil depth and percentage content of clay, silt, sand, stone and organic matter) at different locations in a watershed along with the geographical co-ordinates of the measurement locations, to produce the spatially distributed soil properties for the whole watershed in the form of raster data (Figure 1). These watershed distributed soil properties are also input into an Microsoft Excel Macro Tool (MS-Excel Tool)^[41] to convert the above soil properties into the soil database required for SWAT by using Pedo-transfer functions^[37]. Both the SLEEP model and MS-Excel Tool are built as standalone programs; they can be used in combination to produce the SWAT soil database but they can also be executed independently. Existing spatially distributed soil properties are required if a user intends to use the MS-Excel Tool to generate tabulated soil properties for SWAT without executing SLEEP within ArcGIS 10.1^[42].

Box 1 SLEEP software requirements, documentation and anticipate internet access information

Key user aspects	Additional requirements/description
<u>Required software</u>	
ArcGIS 10.1	Arc-Hydro Tools and Spatial Analysis Extensions need to be enabled (Note: SLEEP will be updated in the future for new versions of ArcGIS)
Microsoft Excel	Macro capabilities need to be enabled (can be executed independently of ArcGIS 10.1 depending on user objectives)
<u>Documentation</u>	
SLEEP User Guide ^[43]	Will be posted on SWAT website (see URL below)
<u>Internet access</u>	
SWAT website	http://swat.tamu.edu/software/links/
Anticipated release date	June 1, 2015

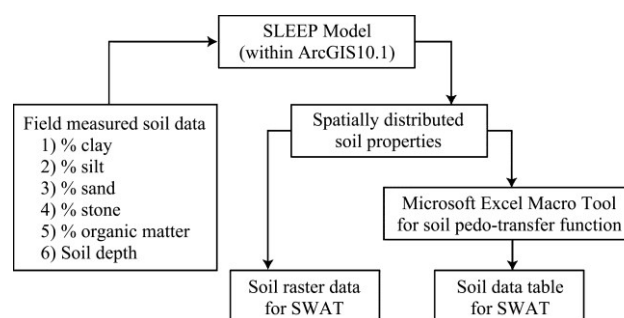


Figure 1 Flow chart showing the process of generating a SWAT user-soil database using SLEEP software in ArcGIS 10.1 in combination with the Microsoft Excel macro Tool

The SLEEP model can be used independently to convert any measured soil properties, in addition to those mentioned above, into spatially distributed forms for applications other than SWAT. An example of these applications is the use of SLEEP outputs in land suitability and land use planning analyses. Previous research has shown that the use of predicted soil attributes using soil-landscape modeling improved the accuracy of suitability maps compared with traditional sources of soil data^[41]. The SLEEP tool will facilitate and speed up the production of detailed soil attributes and therefore provide detailed suitability maps to aid better land use planning. The tool is sufficiently flexible to include steps that will automate production of suitability maps in future versions.

The theoretical foundation of SLEEP is based on previously published work^[44–46] and is summarized in the flow chart shown in Figure 2 and Table 1. SLEEP allows users to use a DEM to automatically generate

terrain attributes required to predict the soil parameters. The SLEEP utilizes the DEM and available soil observations to generate spatially continuous layers of soil attributes. The difference between these layers and those generated by spatial interpolation is that the latter uses the distance as a major factor in interpolating soil attributes between two points, without considering the factors that govern soil variation between the points. However, SLEEP allows for the consideration of the topography between two points and hence accounting for soil forming factors in the interpolation process. Assuming that two observations were sampled at the top of two adjacent hills, the spatial interpolation will assign values to the pixels between the two points without considering the low areas between the two hilltops. The SLEEP model considers these changes in the topography, using the DEM-derived topographic attributes, and therefore assigns values that simulate the variations in the soil forming factors.

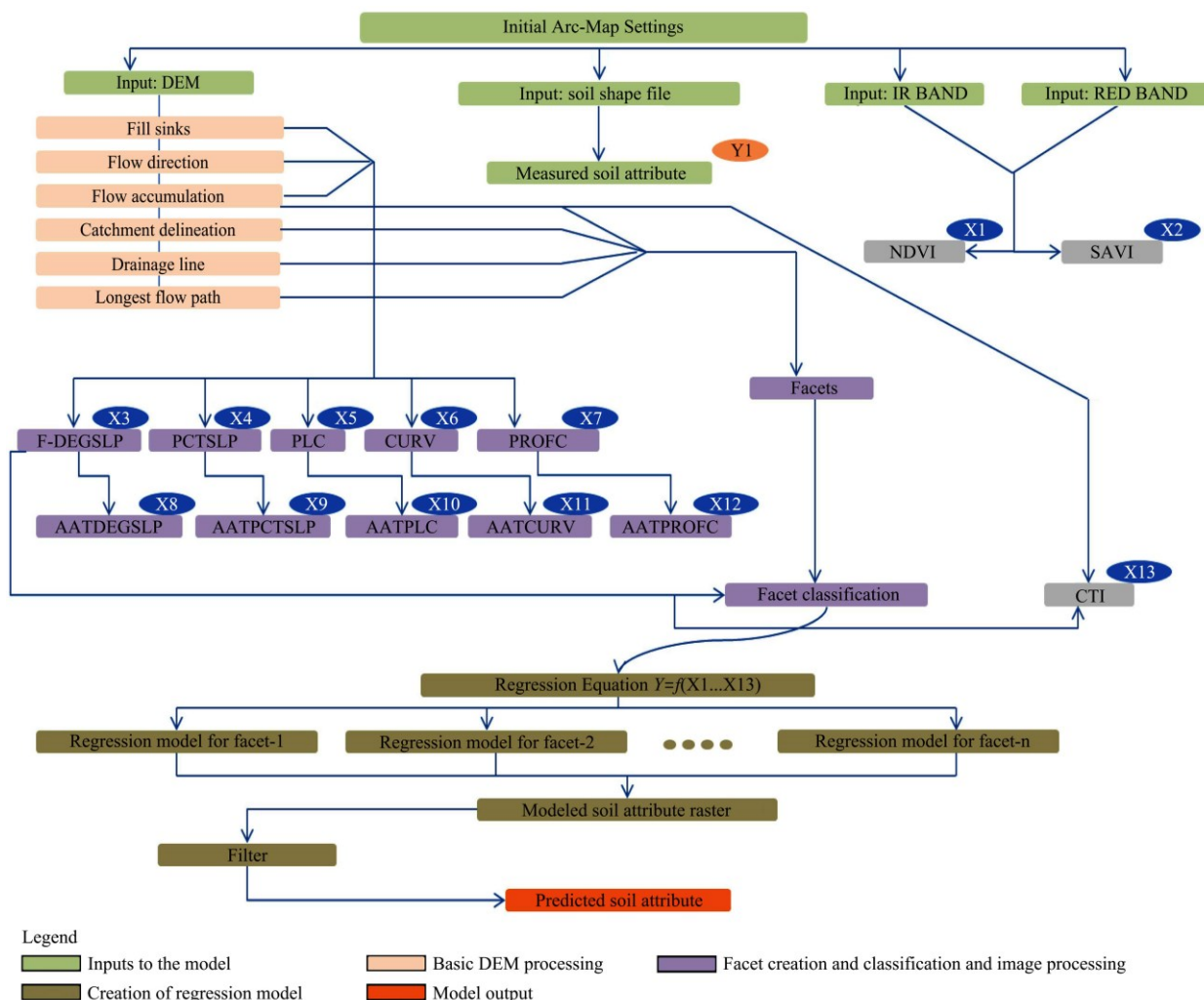


Figure 2 Flow chart showing the methodology of SLEEP model

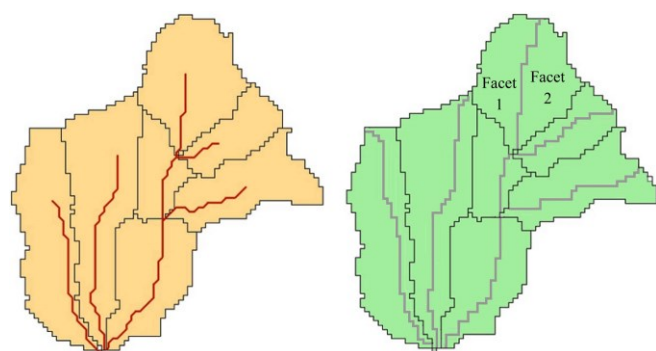
Table 1 List of SLEEP variables, attributes processing steps and corresponding descriptions for Figure 2

Process step	Variable, attribute or processing name	Description
1	Initial Arc-Map Settings	In this step the Geoprocessing Settings and the map document properties are set.
2a	Input: DEM	Digital Elevation map of the watershed or region is loaded
2b	Input Soil Shape File	Point shape file of the locations where the soil observations exist is loaded
2c	Measured Soil Attribute	Measured soil properties are tied to attribute table of the above shape file.
2d	Input IR Band	Input Infra-red Band Image
2e	Input Red Band	Input Red Band Image
3a	Fill Sinks	Step used to fill the sinks in the DEM and create a seamless DEM
3b	Flow Direction	Step used to create flow direction of each pixel from the DEM
3c	Flow Accumulation	Step used to calculate the flow accumulation using the Flow direction raster
3d	Catchment Delineation	Step used to delineate the catchment using the DEM
3e	Drainage Line	These are streams created while delineating the catchment
3f	Longest Flow Path	This is a polyline feature which represents the longest flow path
4	Facets	Subwatershed polygons divided by the drainage line feature or the longest flow path feature
5a	NDVI	Normalized Difference Vegetation Index
5b	SAVI	Soil Adjusted Vegetation Index
5c	F-DEGSLP	Slope grid measured in degrees and converted to an integer grid
5d	PCTSLP	Slope grid measured in percent and converted to an integer grid
5e	PLC	Perpendicular curvature to direction of maximum slope, influences flow convergence and divergence; interpreted as convex, concave or flat.
5f	CURV	Curvature of the surface at each cell with respect to the eight surrounding neighbors, the slope of the slope
5g	PROFC	Curvature in the direction of the maximum slope, affects acceleration and deceleration of flow, interpreted as convex, concave or flat.
5h	AAT ^a -DEGSLP	Accumulated slope degree
5i	AAT ^a -PCTSLP	Accumulated slope percent
5j	AAT ^a -PLC	Accumulated perpendicular curvature
5k	AAT ^a -CURV	Accumulated curvature
5l	AAT ^a -PROFC	Accumulated profile curvature
5m	CTI	Compound Topographic Index
6	Facet Classification	Unsupervised classification of the Facets
7	Regression Equation	Equation relating the soil properties and the above parameters
8	Modeled Soil Attribute Raster	Raster-based spatially distributed soil properties for entire watershed
9	Filter	Smoothing of the raster layer
10	Predicted Soil Attribute	Final predicted soil attribute required for SWAT (or other application)

Note: ^aAAT refers to the Outputs of the Basic Terrain Attributes model that represent the average upstream value accumulated in each cell. To calculate the average, the attribute values are accumulated, and then divided by the number of cells accumulated.

Within SLEEP, the entire watershed is subdivided into subwatersheds and each subwatershed is further divided into two subdivisions or “facets” (Figure 3). These facets are made by subdividing a subwatershed into two parts by a delineated main stream. In the case of a subwatershed located along the boundary of an entire watershed, the longest flow path is used to divide it into facets. Figure 3 shows an example of the facets created by the SLEEP tool. The facet is viewed as an infinite number of hillslope units where the flow of material is expected in the direction from the upper part of the

hillslope (boundaries between two subwatersheds (ridges)) to the lower part (stream line at the end of the hillslope). This enables the establishment of a relationship between terrain attributes within each facet and soil attributes because of the uniformity of soil forming processes within each particular facet. This is because the relationships are generally weak if established for the entire watershed or regional study area. Within these facets the relationships are stronger and enable the prediction of soil characteristics from terrain and satellite variables with acceptable accuracy.



a. Delineated subwatersheds with streams b. Subwatersheds divided into facets
Note: facet 1 and facet 2 derived from one subwatershed.

Figure 3 Deriving facets from subwatersheds

The slope and area of subwatersheds (facets) are used to subdivide the entire watershed or study area into homogeneous classes of subwatersheds. Statistical relationships between soil attributes, derived from field observations, and the derived topographic attributes and remote sensing parameters are established for each class. Regarding soil attributes, it can be any soil attribute that was measured in the field or analyzed in the laboratory at a particular location in the field (the exact location is determined by the geographic coordinates of the sampling site). This gives the user of SLEEP high flexibility in predicting soil attributes that was recorded at particular site(s), which also widen the applications of SLEEP for users other than SWAT users. For the terrain attributes, the possible attributes that could be derived from any DEM are listed in Figure 2. However, SLEEP also enables the user to add further attributes deemed useful for the prediction. This includes additional DEM-derived attributes or any other auxiliary layers that are important in determining soil variations under certain conditions. In Figure 2 there are 13 independent variables (attributes; X1 through X13), but the user can add to these further important independent attributes to enhance the prediction. These relationships are applied to predict soil attributes for each pixel within the class as shown in Figure 2. The predictions are then merged together to provide predictions of soil attributes for the entire watershed or study area. The predicted soil attributes are converted as inputs to SWAT. The results are presented in digital format and with high resolution (based on the resolution of DEM and satellite data).

This makes these outputs an attractive input for SWAT and many other environmental-related models^[44]. More details can be found on the help and tutorial documentation within the program.

3 SLEEP description

The only software required to install and run SLEEP is ArcGIS 10.1 along with the ArcHydro tool^[45] (although users will also need to install Microsoft Excel to perform selected functions as described earlier). The data requirement for using SLEEP are: (1) measured soil attributes at various observation points in the field, which should be converted into a shape file (in case field survey data for this are not available, legacy soil data/maps can be used to derive observations to run the tool), (2) DEM of the areas under considerations, and (3) an optional infrared and red band of the field to calculate the normalized difference vegetation index (NDVI) and the soil adjusted vegetation index (SAVI). The users can access the help documentation and tutorial when the tool is downloaded (Box 1).

The overall aim of SLEEP software is to construct a relationship between the measured soil attributes and the landscape and environmental attributes, and then predict the soil attributes for the entire watershed or study area using these relationships. SLEEP automatically assists the user to calculate the landscape and environmental attributes and to establish linear relationships between these and the soil attributes. However, the user is expected to have a basic understanding of the relationships between landscape and environmental properties as well as soil attributes, in order to make decisions as to how these relationships are used. However, the tool is designed to allow easy iterations of the analysis to achieve desirable results.

The whole package was created with ArcGIS tool box options and code written in Visual Basic. The tool is divided into five major steps and each step is divided into some sub-tasks. The major steps are (Figure 4): (1) Initial ArcMap Setup, (2) Basic DEM Processing, (3) Facet and Attribute Processing, (4) Image Processing and (5) Soil Attribute Prediction.

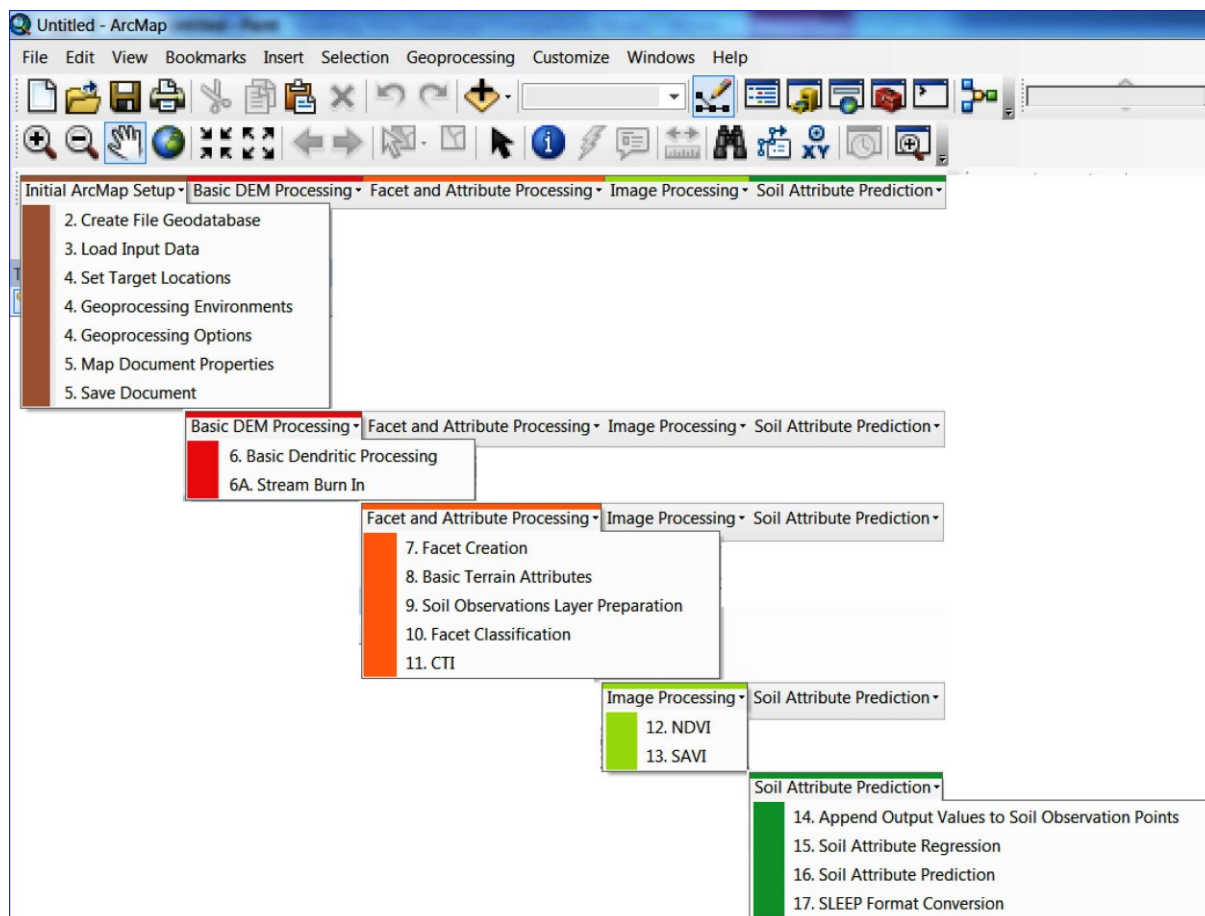


Figure 4 Main and dropdown menus of SLEEP

3.1 Initial ArcMap setup

The first menu, Initial ArcMap Setup (Figure 4), provides guidance on how to set up a new project including the following steps: enable proper access to input data, organize outputs to specific target databases and avoid duplication of files in case of running many iterations. Some of these steps are not directly accessible through SLEEP and need to be done through other menus available in ArcMap or other ArcGIS modules. However, the user will be directed to the necessary action when clicking in the dropdown menu.

3.2 Basic DEM processing

In this menu all the tasks required for the delineation of the catchment and the streams are done (Figure 4). A more reliable shape file of the streams can be directly entered as input if available. In this step various flow-related layers are derived automatically using basic dendritic processing (flow direction, flow accumulation, stream line and subwatersheds within the entire watershed). The necessary input for this step is a digital elevation model (DEM). The outputs of this step are a

series of files with known names that describe their contents. An important decision to be made by the user is the number of pixels needed to start a stream. This is determined by many factors such as the topography of the area, relief and general environmental conditions. As a guideline, 5% of the total number of pixels in the whole DEM is initially used. The user then checks the generated stream layer and if it is not suitable, the step can be repeated until a satisfactory result is achieved.

3.3 Facet and attribute processing

The first step in this menu is the creation of the facets (Figure 4). Each subwatershed is divided into two facets by the stream connecting the outlet and the farthest point in the subwatershed (Figure 3). In cases where the stream does not join these two points, then the longest flow path is used to divide the subwatershed into two facets.

The second step in this menu is basic terrain processing (Figure 2 and Table 1). Some of the terrain attributes used in the regression are calculated here. The terrain attributes calculated are degree slope, percentage

slope, aspect, curvature, profile curvature and plan curvature.

The next step under this menu is the facet classification (Figure 2 and Table 1). Here the facets are classified based on the overall slope value and the area of facets. Previous research indicated that the average slope and area are suitable criteria for this classification, which enable the generation of homogeneous groups of facets^[44]. There is no rule to choose the number of classes. Each time the user runs the classification with a selected number of classes, the tool will return the number of field observations within each class of facets. The statistical relationship depends on the number of observations within each class. This step ensures that a sufficient number of field observations are available within each class to establish rigorous statistical relationships between both terrain and remote sensing parameters versus soil attributes. The user is advised to repeat the classification until the minimum number of observations within each facet is sufficient to build a rigorous statistical model and/or the number of observations for all facets is, as much as possible, even. This will depend on the total number of observations available, the total area under consideration and the complexity of the terrain. The user will need to consider these issues in selecting a suitable number of classes to group the facets.

The last step in this menu is the compound topographic index (CTI) layer creation (Figure 2 and Table 1). CTI is calculated for each pixel using the equation $CTI = \ln (A_s / \tan D)$, where A_s is the average upslope contributing area and D is the average slope degree^[4]. Several researchers have indicated the potential of this variable in predicting various soil characteristics^[4,17,19,47].

3.4 Image processing

The fourth menu, Image Processing (Figure 4), allows the user to derive remote sensing indices, such as NDVI and SAVI (Figure 4 and Table 1).

$$NDVI = [(NIR - Red) / (NIR + Red)]$$

where, NIR is the near infrared band and Red is the red band derived from satellite images. SAVI is calculated using NDVI but multiplied by a factor between zero and

one, using the equation $SAVI = [(NIR - Red) / (NIR + Red + L)] \times [1 + L]$. The factor 'L' is adjusted for the effect of bare soil on deriving the NDVI. In areas with good vegetation cover $SAVI = NDVI$, and $L = 1$. In areas with low vegetation or more bare soil then L is approximately 0.5. Both NDVI and SAVI improve the accuracy of predicting soil characteristics^[46]. In areas with low vegetation cover, the SAVI index provides information about the variations in soil color, which is directly linked to soil properties. Conversely, NDVI data provide information about the vegetation conditions in areas with dense vegetation cover, which is indirectly reflecting the variations in soil properties. Therefore, proper use of these indices provides information about soil variability and improves the prediction accuracy. The user can also use any auxiliary information/layers that might improve the prediction, such as land use/cover and geology.

3.5 Soil attribute prediction

The fifth and last menu, Soil Attribute Prediction, is where all generated information in the previous steps are collated, analyzed statistically and used to generate the final product, a predicted soil characteristic (Figure 2). The first step in this menu is to append all output layers that were previously generated at the soil observation points. The result of this step is a point shape-file with the above calculated terrain attributes and processed satellite images (NDVI and SAVI). The facet class to which the points belong is also appended to all field observations. The point file, for each class, contains a column for each terrain and satellite image parameter and column for each soil attribute. One row is presented for each soil observation. This will enable the establishment of a statistical model (regression) between terrain and satellite image indices in one hand and soil characteristics in the other hand, for each class of facets.

The regression is performed externally by a Visual Basic code and is linked to the toolbox as a step in SLEEP. The attribute table generated in the previous step is used as an input into this tool. The regression tool gives the correlation coefficient between the chosen soil parameter and the various calculated terrain attributes and remote sensing attributes. From these correlation

coefficient values the user can choose the variable(s) to be used for generating the regression equation to produce the chosen soil attribute. Here different regression equations are generated for different facet classes. An example of these equations is:

$$\text{Clay content (class 3)} = 39.2 - 0.53(\text{slope}) + 1.34(\text{CTI}) + 3.93(\text{aspect}) + 0.04(\text{NDVI}) \quad (1)$$

The last step is the Soil Attribute Prediction. Here the regression equation for different facets formed in the previous steps is used to generate the soil attribute for each pixel in the corresponding facets. For example the above equation is applied within class three, by using the values of slope, CTI, aspect and NDVI, for each pixel, to derive the predicted clay content for that pixel and for all pixels within the class. The regression coefficient of predicting clay content (equation (1)) using slope, CTI, aspect and NDVI was ($R^2 = 0.80$), which is in agreement with previous research^[5,45,46]. Since these terrain parameters are already calculated from the DEM, clay content could be estimated for any other pixel within class 3, with an acceptable accuracy. The model will then apply each equation for the relevant class to derive the predicted soil attributes. The results will then be combined for all classes of facets to generate one layer of predicted soil characteristics for the entire watershed or study area. This represents the ultimate product of this program. The users are advised to run at least one smoothing filter of 3×3 or 5×5 windows to smooth any extreme values generated during the prediction process, especially at the edges between different classes.

An independent set of observation points which is not used in the above process can be used for validation of the soil attribute layer predicted by SLEEP. The root mean square error (*RMSE*) can be used to assess the agreement between predicted and observed data. The whole process can be iterated any number of times until a satisfactory result is obtained. During each iteration the number of facet classes and the variables chosen for the regression equation can be changed. In case sufficient field observations are not available, the users are encouraged to search for legacy soil data, either from previous surveys or to derive these from existing soil maps with attached soil observations.

The model also includes a step known as “SLEEP format conversion” within the “Soil Attribute Prediction” menu. The output is produced in the form of raster image. This is converted into ASCII format and then into a tabular form. The conversion to this tabular form will help in exporting the data to any other model or tool which requires data in tabular format. The MS-Excel Tool developed uses this output and applies the Pedo-transfer function to form the soil attributes required by the SWAT model.

4 SLEEP application: An example

4.1 Description of the methodology

SLEEP was tested in a 54-km² watershed located in the Tana River basin in northwest Ethiopia (Figure 5). The watershed has 203 field observations, where soil attributes were collected in the field and/or analyzed in the laboratory^[46]. The measured soil attributes, which were used in this application, included the depth of the soil layer and percentages of silt, sand, clay and organic matter in different layers of the soil profile. These were selected because these are the basic soil attributes needed by SWAT, as well as by many other environmental models. There were two trials done using SLEEP. In the first, nearly 80% of the observations were used for the calculation of the seamless soil attribute raster map, and the remaining points were used for validation of the model. Therefore, 167 points were used for calculation and 36 points were kept aside for validation. To ensure robustness of the model, the second trial used fewer points for calculation (only 30 points out of the total of 167 were used for calculation) and the same 36 points as used in the previous step were used for validation, thus ensuring cross comparison of model performance.

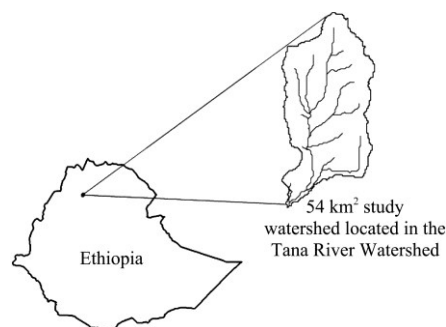


Figure 5 Location of the study area (Tana River basin) in northwest Ethiopia

The model results were also compared with the spatial interpolation techniques available in the ArcGIS package. Inverse distance weighted (IDW) and Kriging methods were used to interpolate the soil attributes using the same 167 and 30 sample points for input as used for SLEEP.

4.2 Results and discussion

The accuracy of the results derived from SLEEP and from the two spatial interpolation techniques were compared, based on their agreement with the same set of verification observations and on their merit in reflecting the spatial distribution of the predicted/interpolated soil attributes. The coefficients of determination (R^2) of the established regression models using either 167 or only 30 observations (Table 2) were generally in agreement with previous studies^[48–50]. Within SLEEP, the slope and area of subwatersheds (facets) are used to subdivide the entire watershed or study area into homogeneous classes of subwatersheds. The statistical relationships between soil attributes and topographic attributes and remote sensing parameters are established for each class. The grouping of facets into classes improves the prediction compared with establishing statistical relationships for the

entire watershed. However, the number of observations for each class should be enough to establish a robust regression model. Using 167 observations, it was possible to group all facets into six classes. However, only one class was used for the subset of 30 observations because there are not enough observations. The R^2 could be improved using stepwise regression. However, at this stage SLEEP allows only simple multiple linear regressions between soil attributes as dependent variables and terrain and remote sensing attributes as independent variables. In future versions of SLEEP, this will be improved to allow better fine-tuning of the regression models to improve the results. At present, SLEEP users can view the correlation coefficients and use these to select the terrain and satellite attributes that will be used to build the regression model and predict soil attributes. An example of the regression models to predict the soil organic matter content of the surface layer is presented in Table 3. The R^2 for the different classes were in the range of 0.23–0.57, and these models were applied within SLEEP to predict soil attributes.

Table 2 Coefficient of determination (R^2) of the predicted soil attributes for different classes using different observation densities

No. of observations used	Facet class	Organic content/%	Clay content/%	Silt content/%	Sand content/%	Stone content/%	Soil depth/cm
30	1	0.33	0.28	0.26	0.19	0.33	0.50
	1	0.57	0.45	0.19	0.34	0.24	0.49
	2	0.53	0.50	0.33	0.38	0.33	0.38
	3	0.33	0.23	0.27	0.08	0.41	0.25
	4	0.49	0.41	0.43	0.23	0.66	0.56
	5	0.29	0.27	0.17	0.05	0.09	0.13
167	6	0.23	0.30	0.11	0.33	0.18	0.38

Table 3 Regression model used to calculate organic matter content

Facet class	Intercept	NDVI	CTI	Accum ¹ slope D	Profile curvature	Accum ² aspect	Accum ³ slope P	Percent slope	R^2
1	2.51	1.02	−0.23	0.60	−0.35	0.42	0.08	0.02	0.57
2	1.20	−0.68	0.10	1.70	−3.72	0.17	−0.04	0.08	0.53
3	18.55	0.17	−1.13	−0.38	−0.84	−0.15	0.04	−0.02	0.33
4	10.57	5.94	−0.50	0.23	−1.71	−0.25	0.12	−0.02	0.49
5	0.68	−3.14	−0.15	0.81	0.81	0.33	0.09	−0.02	0.29
6	14.76	−0.13	−0.83	−0.57	0.13	−0.86	0.09	−0.02	0.23

Note: ¹ Accum slope D: average slope degree of the upslope contributing pixels; ² Accum aspect: average aspect of the upslope contributing pixels; ³ Accum slope P: average slope percent of the upslope contributing pixels

The root mean square errors (RMSEs) for the predicted soil attributes were generally comparable to those generated using the two interpolation techniques (Table 4). This differed slightly compared to results

from previous studies, which indicated better RMSE and accuracy of the soil–landscape prediction models^[46,51]. Previous research has showed that the spatial interpolation methods are usually data-specific or even

variable-specific and indicated that the predictive performance of the methods depends on many factors^[1]. The accuracy of the predicted soil attributes using SLEEP could be improved significantly using stepwise multiple linear regressions, which allows the selection of the most important factors to predict soil attributes. It also appears that for each soil attribute, within each facet class, there are certain terrain and satellite attributes suitable to generate an optimum accuracy of prediction; this will be considered in further development of SLEEP to improve prediction accuracy.

Nevertheless, comparison of the spatial distribution of

the predicted soil attributes using SLEEP with those derived from interpolation techniques indicate an obvious advantage of the former (Figure 6). While Kriging classifies the entire watershed into two classes, within which many verification observations are different from that class, the prediction using SLEEP classifies the area into many classes, within which the verification observations are in closer agreement with the spatial distribution of soil attributes. Hence, the application of SLEEP, using careful selection of independent terrain and satellite attributes, can lead to better mapping of soil attributes.

Table 4 Root mean square error between the observed soil parameters and predicted attributes using SLEEP and interpolated attributes using Kriging and inverse distance weighted (IDW) techniques using 167 or 30 observations.

Soil Property	SLEEP 167	Kriging 167	IDW 167	SLEEP 30	Kriging 30	IDW 30
Soil depth	24.8	23.2	24.5	29.3	29.9	29.8
Organic content	1.6	1.4	1.4	1.4	1.3	1.5
Clay	8.9	10.0	9.0	10.6	10.8	11.1
Silt	7.7	6.8	6.0	6.4	6.4	6.4
Sand	8.5	7.4	10.0	7.9	7.5	9.3
Stone	12.7	13.1	16.0	14.1	17.4	13.7

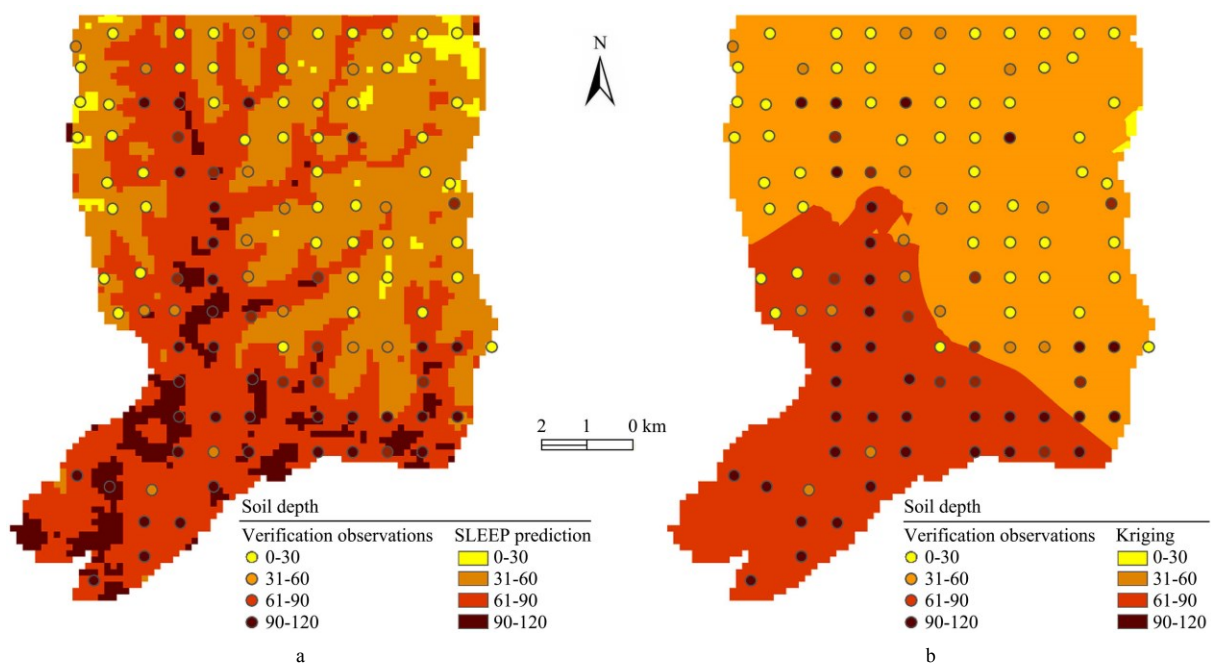


Figure 6 Comparison of the spatial distribution of predicted soil attributes using (a) SLEEP and (b) interpolated soil attributes using Kriging algorithm

The SLEEP model gives the output of the soil attributes in the form of raster image and also in table format as a MS-Excel file. If the organic carbon and silt, sand, clay, and stone percentages are estimated using the SLEEP tool for a specific area/watershed, then the soil

attributes required in the soil database for SWAT can be calculated using the Pedo-transfer functions^[39,52]. A special tool to convert the predicted soil data to SWAT input formats will be added in future releases of SLEEP. This will enhance the application of SLEEP for

environmental modeling such as SWAT. One important application would be using SLEEP to generate detailed soil layers and testing the improvement of SWAT results compared with traditional soil data.

5 SLEEP application for SWAT

For the same watershed located in the Tana basin (Figure 5) five SWAT models were setup using the same DEM, landuse data and weather data but different soil data. The DEM was obtained from the Shuttle Radar Topography Mission (SRTM)^[53], the landuse data was derived from Moderate Resolution Imaging Spectroradiometer (MODIS) land use data^[54] and the weather data used were from the Climate Forecast System Reanalysis (CFSR)^[55]. The five sets of soil data used were: (1) soil layer data derived from the 167 SLEEP field data points, (2) soil layer data derived from the 30 SLEEP field data points, (3) soil layer data derived from Kriging interpolation techniques of 167 field data points, (4) soil layer data derived from Kriging interpolation techniques of 30 field data points, and (5) soil layer data based on the Food and Agriculture Organization (FAO) Harmonized World Soil Database^[56]. This analysis allows additional investigation of the performance of SLEEP and it is usefulness to improve SWAT outputs compared with interpolation methods and the FAO data. Ideally, these SWAT outputs should be compared with observed data. However, the absence of sufficient runoff data for long time to enable good calibration of SWAT is a limitation in this area.

Daily stream flow time series generated by the above mentioned five SWAT models are plotted in Figure 7 for the wet periods during August of 2012. There is nearly no discernible difference in the predicted stream flows using the soil data generated from the 30 points (SLEEP30) versus the soil data based on the 167 points (SLEEP 167). This shows the consistency in the SWAT model output while using soil data developed using the different empirical relations in the SLEEP 167 and SLEEP 30 soil datasets. In addition, this indicates that when SLEEP is used extensive field observation are not required to generate adequate soil data for SWAT simulations of the Tana River basin experimental

subwatershed. However, further testing of the SLEEP software is needed to determine accurate thresholds for other watershed conditions. The SWAT stream flow output from the SLEEP soil data is in better agreement with the output using the FAO soil data compared with the SWAT output using the Kriging method. This small watershed in the Tana basin is an experimental watershed and has an advantage of having soil properties measured at very close spatial intervals which should have resulted in the SWAT's performance of SLEEP data being close to that of the FAO in predicting stream flow. However, using the same observations to derive soil data for SWAT using interpolation techniques doesn't seem to provide comparable SWAT outputs. These results point to the robustness of the SLEEP methodology. However, the performance of SLEEP needs to be tested in multiple watersheds with varying environmental conditions and different spatial distributions of the measured soil data, and with more in-depth calibration and validation of SWAT applications using SLEEP-derived soil data.

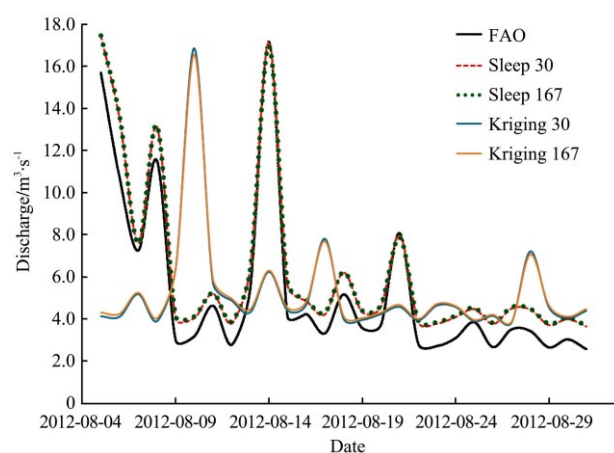


Figure 7 Comparison of the SWAT modeled stream flow (not calibrated) by using different soil inputs

6 Conclusions

The SLEEP tool presented here enables users to use terrain attributes, remote sensing data and auxiliary information to predict the lateral and vertical distribution of soil characteristics. Using GIS capabilities, the user can run many iterations in a reasonable time and produce satisfactory results. The five menus of the tool will guide the user in a systematic style in analyzing the relationship between factors that govern soil variations under specific environmental conditions – this will

facilitate the understanding of the cause–effect relationships that aid this prediction. This approach is an attempt to quantify and automate the qualitative approach that has long been used by surveyors to map the distribution and characteristics of soils. The tool provides outputs in a format suitable for SWAT, which enables the integration of the outputs to enhance the prediction of watershed processes. The application of SLEEP in an Ethiopian watershed produced promising results in generating accurate predictions coupled with a reasonable spatial distribution of soil attributes that better resembled the field situation than interpolation techniques. The analysis indicated low sensitivity of SWAT prediction when SLEEP was used to derive soil inputs from 167 observations compared with only 30 observations. This highlights the potential of SLEEP to provide satisfactory soil inputs for SWAT in areas with scarce soil information. Yet, the tool represents a starting point and is designed to be dynamic and to evolve in response to users’ needs and feedback. Additional features that will be developed are the ability to use stepwise multiple linear regressions or other algorithms to replace the simple regression so as to strengthen the predictions and incorporating means to verify the spatial and attribute accuracy of the predictions. The stepwise multiple linear regression allows the user to select the independent variables that are best predictors of certain soil attribute, which improve the prediction accuracy as compared to simple regression. It is anticipated that as more users from various disciplines use this public domain tool, increased development will strengthen and widen its use for environmental applications.

Acknowledgements

This model is a result of collaborative efforts between the International Center for Agricultural Research in the Dry Areas (ICARDA) and Texas A&M University. The authors would like to acknowledge the financial support by the CGIAR Research Program on Water, Land and Ecosystems (WLE), USAID – linkages program, Middle East Water and Livelihood Initiative - WLI, and the Coca-Cola Foundation. The authors would also like to

thank Dr. Philip Gassman for reviewing the manuscript and providing substantial comments and suggestions.

[References]

- [1] Li J, Heap A D. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, 2014; 53: 173–189. doi: 10.1016/j.envsoft.2013.12.008.
- [2] Mermut A., Eswaran H. Some major developments in soil science since the mid-1960s. *Geoderma*, 2001; 100: 403–426. doi: 10.1016/S0016-7061(01)00030-1.
- [3] Salehi M H, Eghbal M K, Khademi H. Comparison of soil variability in a detailed and a reconnaissance soil map in central Iran. *Geoderma*, 2003; 111: 45–56. doi: 10.1016/S0016-7061(02)00252-5.
- [4] Moore I D, Gessler P E, Nielsen G A, Peterson G A. Soil Attribute Prediction Using Terrain Analysis. *Soil Science Society of America Journal* Soil Science Society of America, 1993; 57: 443. doi: 10.2136/sssaj1993.03615995005700020026x.
- [5] Pachepsky Y A, Timlin D J, Rawls W J. Soil Water Retention as Related to Topographic Variables. *Soil Science Society of America Journal* Soil Science Society: 2001; 65: 1787 doi: 10.2136/sssaj2001.1787.
- [6] A. X. Zhu, B Hudson, Burt j, Lubich K, Simonson D. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science society of America Journal*. 2001; 65: 1463–1472.
- [7] Esfandiarpour B I, Salehi M H, Toomanian N, Mohammadi J, Poch R M. The effect of survey density on the results of geopedological approach in soil mapping: A case study in the Borujen region, Central Iran. *Catena*, 2009; 79: 18–26 doi: 10.1016/j.
- [8] Cook S E, Corner R J, Grealish G, Gessler P E, Chartres C J. A Rule-based System to Map Soil Properties. *Soil Science Society of America Journal* Soil Science Society of America, 1996; 60: 1893 doi: 10.2136/sssaj1996.03615995006000060039x.
- [9] McKenzie N J, Gessler P E, Ryan P J, O’Connell D A. The role of terrain analysis in soil mapping, In: Wilson J P, Gallant J C (Eds.) *Terrain Analysis: Principles and Applications* New York, NY: John Wiley and Sons, 2000; Chapter 10.
- [10] Girgin B N, Frazier B E. Landscape position and surface curvature effects on soils developed in the Palouse area, WA. Pullman, WA: Washington State University, Department of Crop and Soil Sciences. 1996.
- [11] Gessler P E. and Chadwick O A. Quantitative soil–landscape modeling: a key to linking ecosystem

- processes on hillslopes. *Pedometrics '97 International Workshop University of Wisconsin-Madison, Madison, Wisconsin, August 18–20. 1997.*
- [12] McBratney A, Mendonça S M, Minasny B. On digital soil mapping. *Geoderma*, 2003; 117: 3–52. doi: 10.1016/S0016-7061(03)00223-4.
- [13] Jiang H T, Xu F F, Cai Y, Yang D Y. Weathering Characteristics of Sloping Fields in the Three Gorges Reservoir Area, China. *Pedosphere*, 2006; 16: 50–55. doi: 10.1016/S1002-0160(06)60025-8.
- [14] Zhang X Y, Sui Y, Zhang X D, Meng K, Herbert S J. Spatial Variability of Nutrient Properties in Black Soil of Northeast China. *Pedosphere*, 2007; 17: 19–29 doi: 10.1016/S1002-0160(07)60003-4.
- [15] Florinsky I, Eilers R, Manning G, Fuller L. Prediction of soil properties by digital terrain modelling. *Environmental Modelling & Software*, 2002; 17: 295–311 doi: 10.1016/S1364-8152(01)00067-6.
- [16] Klingseisen Bernhard, Metternicht Graciela, Paulus Gernot. Geomorphometric landscape analysis using a semi-automated GIS-approach. *Environmental Modelling & Software*, 2008; 23: 109–121 doi: 10.1016/j.envsoft.2007.05.007.
- [17] Gessler P E, Moore I D, McKenzie N J, Ryan P J. Soil-landscape modelling and spatial prediction of soil attributes. *International Journal of Geographical Information Systems Taylor & Francis Group*, 2007;
- [18] Kuriakose S L, Devkota S, Rossiter D G, Jetten V G. Prediction of soil depth using environmental variables in an anthropogenic landscape, a case study in the Western Ghats of Kerala, India. *Catena*, 2009; 79: 27–38 doi: 10.1016/j.catena.2009.05.005.
- [19] Gessler P E, Chadwick O A, Chamran F, Althouse L, Holmes K. Modeling Soil–Landscape and Ecosystem Properties Using Terrain Attributes. *Soil Science Society of America Journal Soil Science Society*, 2000; 64: 2046 doi: 10.2136/sssaj2000.6462046x.
- [20] Goodman A. Trend surface analysis in the comparison of spatial distributions of hillslope parameters. Ph. D. Dissertation. Deakin University, 1999.
- [21] Browning D M, Duniway M C. Digital soil mapping in the absence of field training data: A case study using terrain attributes and semi automated soil signature derivation to distinguish ecological potential. *Applied and Environmental Soil Science*, 2011.
- [22] Mishra U. Predicting storage and dynamics of soil organic carbon at a regional scale. PhD Dissertation, Ohio State University, 2009.
- [23] Marchetti A, Piccini C, Santucci S, Chiuchiarrelli I, Francaviglia R. Simulation of soil types in Teramo province (Central Italy) with terrain parameters and remote sensing data. *Catena*, 2011; 85: 267–273 doi: 10.1016/j.catena.2011.01.012.
- [24] Mulder V L, de Bruin S, Schaepman M E, Mayr T R. The use of remote sensing in soil and terrain mapping — A review. *Geoderma*, 2011; 162: 1–19. doi: 10.1016/j.geoderma.2010.12.018.
- [25] Arnold J G, Fohrer N. SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes*, 2005; 19: 563–572 doi: 10.1002/hyp.5611.
- [26] Williams J R, Arnold J G, Kiniry J R, Gassman P W, Green C H. History of model development at Temple, Texas. *Hydrological Sciences Journal Taylor & Francis Ltd*, 4 Park Square, Milton Park, Abingdon Ox14 4rn, Oxon, England: 2008; 53: 948–960 doi: 10.1623/hysj.53.5.948.
- [27] Arnold J G, Srinivasan R, Muttiah R S, Williams J R. Large Area Hydrologic Modeling And Assessment Part I: Model Development. *Journal of the American Water Resources Association*, 1998; 34: 73–89 doi: 10.1111/j.1752-1688.1998.tb05961.x.
- [28] Neitsch S L, Arnold J G, Kiniry J R, Williams J R. Soil & Water Assessment Tool Theoretical Documentation Version 2009; 2011.
- [29] Geza M, McCray J E. Effects of soil data resolution on SWAT model stream flow and water quality predictions. *Journal of Environmental Management*, 2008; 88: 393–406 doi: 10.1016/j.jenvman.2007.03.016.
- [30] Boluwade A, Madramootoo C. Modeling the Impacts of Spatial Heterogeneity in the Castor Watershed on Runoff, Sediment, and Phosphorus Loss Using SWAT: I. Impacts of Spatial Variability of Soil Properties. *Water, air, and soil pollution*, 2013; 224, 1692 doi: 10.1007/s11270-013-1692-0.
- [31] Bossa A Y, Diekkrüger B, Igué A M, Gaiser T. Analyzing the effects of different soil databases on modeling of hydrological processes and sediment yield in Benin (West Africa). *Geoderma*, 2012; 173–174 doi:10.1016/j.geoderma.2012.01.012.
- [32] Moriasi D N, Starks P J. Effects of the resolution of soil dataset and precipitation dataset on SWAT2005 streamflow calibration parameters and simulation accuracy. *Journal of Soil and Water Conservation*, 2010; 65: 63–78 doi: 10.2489/jswc.65.2.63.
- [33] Li R, Zhu A X, Song X, Li B, Pei T, Qin C. Effects of spatial aggregation of soil spatial information on watershed hydrological modelling. *Hydrological Processes*, 2012; 26: 1390–1404 doi: 10.1002/hyp.8277.
- [34] USDA-NRCS. U.S. General Soil Map (STATSGO2). U.S. Department of Agriculture, Natural Resources Conservation Service: Washington, DC, USA. 2009.

- [35] USDA-NRCS. Soil Survey Geographic (SSURGO) Database. U.S. Department of Agriculture, Natural Resources Conservation Service: Washington, DC, USA. 2009.
- [36] Wang X, Melesse A M. Effects of statsgo and ssurgo as inputs on swat model's snowmelt simulation1. *Journal of the American Water Resources Association*, 2007; 42: 1217–1236 doi: 10.1111/j.1752-1688.2006.tb05296.x.
- [37] Mukundan R, Radcliffe D E, Risse L M. Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. *Journal of Soil and Water Conservation*, 2010; 65: 92–104 doi: 10.2489/jswc.65.2.92.
- [38] Mednick A C. Does soil data resolution matter? State Soil Geographic database versus Soil Survey Geographic database in rainfall-runoff modeling across Wisconsin. *Journal of Soil and Water Conservation Soil Water Conservation Soc*, 945 Sw Ankeny Rd, Ankeny, Ia 50023-9723 USA: 2010; 65: 190–199 doi: 10.2489/jswc.65.3.190.
- [39] Romanowicz A A, Vanclooster M, Rounsevell M, La Junesse I. Sensitivity of the SWAT model to the soil and land use data parametrisation: a case study in the Thyle catchment, Belgium. *Ecological Modelling*, 2005; 187: 27–39 doi: 10.1016/j.ecolmodel.2005.01.025.
- [40] Thorp K R, Bronson K F. A model-independent open-source geospatial tool for managing point-based environmental model simulations at multiple spatial locations. *Environmental Modelling & Software*, 2013; 50: 25–36 doi: 10.1016/j.envsoft.2013.09.002.
- [41] Microsoft Office Excel 2010". <https://products.office.com/en-us/excel>
- [42] ArcGIS 10.1 Simplifies Sharing of Geographic Information: New Tools and Infrastructure Extend the Reach of GIS throughout Organizations" (Press release). Esri. 2012-06-11. Archived from the original on 2012-06-15.
- [43] Ziadat, F., Shoemate, D., Dhanesh, Y., Srinivasan R, Tech, J. User's Guide for SLEEP (Soil-Landscape Estimation and Evaluation Program) for ArcGIS 10.1. Spatial Sciences Laboratory, Texas A&M University, College Station, USA.
- [44] Ziadat F M. Land suitability classification using different sources of information: Soil maps and predicted soil attributes in Jordan. *Geoderma*, 2007; 140: 73–80 doi: 10.1016/j.geoderma.2007.03.004.
- [45] Ziadat F M. Analyzing Digital Terrain Attributes to Predict Soil Attributes for a Relatively Large Area. *Soil Science Society of America Journal SOIL SCI SOC AMER*, 677 South Segoe Road, Madison, WI 53711 USA: 2005; 69: 1590 doi: 10.2136/sssaj2003.0264.
- [46] Nurhussen M S, Yitaferu B, Kibret K, Ziadat F M. Soil-Landscape Modeling and Remote Sensing to Provide Spatial Representation of Soil Attributes for an Ethiopian Watershed. *Applied and Environmental Soil Science*, 2013; Article ID 798094.
- [47] Ziadat F M. Prediction of Soil Depth from Digital Terrain Data by Integrating Statistical and Visual Approaches. *Pedosphere*, 2010; 20: 361–367 doi: 10.1016/S1002-0160(10)60025-2.
- [48] Kunkel M L, Flores A N, Smith T J, McNamara J P, Benner S G. A simplified approach for estimating soil carbon and nitrogen stocks in semi-arid complex terrain. *Geoderma*, 2011; 165: 1–11 doi: 10.1016/j.geoderma.2011.06.011.
- [49] Sumfleth K, Duttman R. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. *Ecological Indicators Elsevier Science Bv*, Po Box 211, 1000 Ae Amsterdam, Netherlands: 2008; 8: 485–501 doi: 10.1016/j.ecolind.2007.05.005.
- [50] Van de W J, Baert G, Moeyersons J, Nyssen J, De Geyndt K, Taha N. Soil-landscape relationships in the basalt-dominated highlands of Tigray, Ethiopia. *Catena*, 2008; 75: 117–127 doi: 10.1016/j.catena.2008.04.006.
- [51] Selige T, Böhner J, Schmidhalter Urs. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma*, 2006; 136: 235–244 doi: 10.1016/j.geoderma.2006.03.050.
- [52] Saxton K E, Rawls W J. Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal Soil Sci Soc Amer*, 677 South Segoe Road, Madison, WI 53711 USA: 2006; 70: 1569 doi: 10.2136/sssaj2005.0117.
- [53] Jarvis A, Reuter H I, Nelson A, Guevara E. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database. 2008.
- [54] NASA Land Processes Distributed Active Archive Center (LP DAAC) as producer), Land Cover Type Yearly L3 Global 0.05Deg CMG, 2011.
- [55] Environmental Modeling Center/National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, NCEP Climate Forecast System Reanalysis (CFSR) Selected Hourly Time-Series Products, January 1979 to December 2010.
- [56] FAO/IIASA/ISRIC/ISSCAS/JRC. Harmonized World Soil Database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria. 2012.