

Digital soil mapping

for precise land management

P C Moharana* and S Dharumarajan

ICAR-National Bureau of Soil Survey and Land Use Planning, Nagpur, Maharashtra 440 033

Digital soil mapping (DSM) has evolved from traditional soil survey to take advantage of advances in computing and geographic data handling. DSM methodology has been adopted by FAO so that digital soil maps can be produced reliably for sustainable land management. DSM has moved from research phase to operational phase across the nations but no fully fledged research has been carried out in India. ICAR-National Bureau of Soil Survey and Land Use Planning (ICAR-NBSS&LUP), Nagpur has recently worked on Digital soil mapping with the objective to develop soil properties map as per GlobalSoilMap specifications. DSM will be regarded as key for guiding decision-makers in natural resource assessment, environmental modelling, and land use studies.

Keywords: Digital soil mapping, Digital soil assessment, Environmental covariate, Legacy soil database

Soil plays an integral part in the global environmental sustainability challenges of food security, water security, climate stability, biodiversity, and ecosystem service delivery. Accurate mapping of soil properties is crucial for adequate management both at global and local levels. For many years, digital soil mapping (DSM) represented an alternative to conventional soil survey, to produce soil prediction maps and exploiting associated uncertainties in large sets of spatial data combined to recent or ancillary data. DSM has evolved from traditional soil survey to take advantage of advances in computing and geographic data handling, as well as increased availability of environmental covariate data from digital elevation models and remotely sensed imagery. DSM can be regarded as key for guiding decision-makers in natural resource assessment, environmental modelling, and land use studies. However, it requires the knowledge of experienced pedologists and experts in modelling. DSM methodology has been adopted by FAO so that digital soil maps can

be produced reliably for sustainable land management. Traditional soil mapping techniques mostly depend on ground-based surveys and rarely provide information about the spatial distribution of soil properties at the desired resolution over the landscape. Furthermore, mapping soil spatial variations by traditional field surveys is time-consuming and expensive and at the same time it is difficult to capture the real status of the dynamic soil properties, especially at national, regional or global scales. Therefore, it is necessary to have robust methods and models to predict soil properties at a given location or scale. In India, ICAR-National Bureau of Soil Survey and Land Use Planning (ICAR-NBSS&LUP), Nagpur has recently worked on digital soil mapping with the objective to develop soil properties map as per GlobalSoilMap specifications.

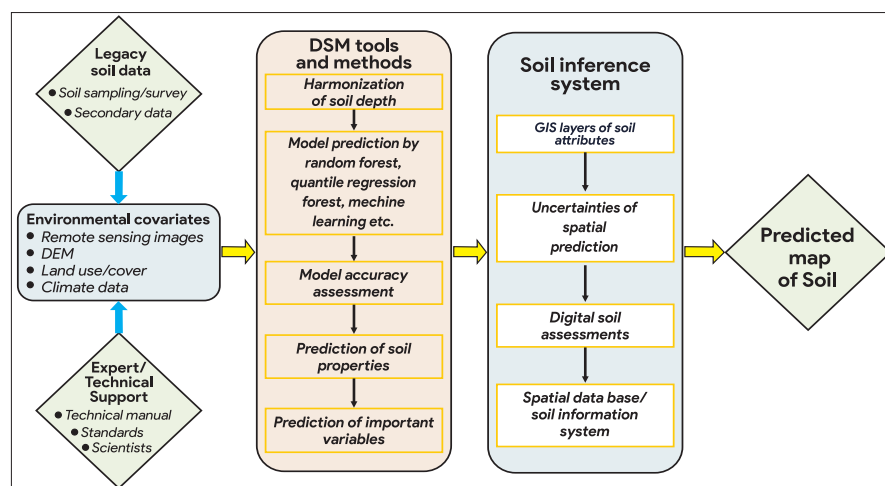
Digital Soil Mapping

DSM is the generation of geographically referenced soil databases based on quantitative relationships between spatially explicit environmental data and

measurements made in the field and laboratory. The digital soil map is a raster composed of two-dimensional cells (pixels) organized into a grid in which each pixel has a specific geographic location and contains soil data. Digital soil maps illustrate the spatial distribution of soil classes or properties and can document the uncertainty of the soil prediction. The scientific foundation of soil mapping is Hans Jenny's (1941) conceptual model according to which soils (S) on a landscape are a function of five environmental factors, namely climate (cl), organisms (o), relief (r), parent material (p), and time (t):

$$S = f(cl, o, r, p, t)$$

While this model, sometimes known as CLORPT, has been useful in conventional soil mapping, it is not quantitative neither spatially explicit. To represent soil and the related environmental factors in a spatial context and to express these relationships quantitatively, SCORPAN model was proposed, where soil (as either soil classes, Sc, or soil attributes, Sa) at a point in space and time is an empirical quantitative function of seven environmental covariates: soil (s),



Schematic representation of the DSM methodology used in prediction of soil properties

and geological maps. Recent advances in remote sensing and geomorphometry have made the quantification of such ancillary information more readily accessible in formats that can be manipulated with GIS. Perhaps the most easily quantified and directly correlated state factors are relief, as digital elevation model (DEM) derivatives have been the most popular predictor (Table 1).

Global Soil Map Project

Global Soil Map Project is an initiative of the Digital Soil Mapping Working Group of the International Union of Soil Sciences (IUSS) and is led by academic and research centres in all continents (GlobalSoilMap.net). This project aims to make a new digital soil map of the world using state-of-the-art and emerging technologies for soil mapping and predicting soil properties at fine resolution. It will be supplemented by interpretation and functionality options, and could assist in better decisions in a

climate (c), organisms (o), relief (r), parent material (p), age (a), and spatial location (n):

$$Sc,a = f(s, c, o, r, p, a, n)$$

The important advances of the SCORPAN model for use in DSM are: (1) the recognition that the environmental factors are not necessarily independent of each other and are thus defined as environmental covariates, (2) the inclusion of soil as an environmental covariate, (3) the spatially explicit nature of the model, and (4) the quantitative nature of the functional relationships. In the SCORPAN model, soil as point observational data, existing soil maps, or remotely sensed spectral properties, can be used as input data. The SCORPAN model facilitates the quantification of the relationships between spatially explicit digital environmental covariates and the soil classes or attributes to be predicted in a spatial context. It also facilitates the estimation of error or uncertainty of the spatial prediction of soil classes or properties.

Environmental covariate

Environmental covariates are the variables characterizing soil forming factors of climate, parent material, topography, vegetation, human activities and time. In the last decade, many efforts have been made in developing effective environmental covariates according to the targeted soil properties and landscapes. Climate and terrain factors have been widely used in DSM. Common climate

variables such as air temperature and precipitation are often generated through interpolating the observations from meteorological stations. With the advance of remote sensing techniques, remotely sensed surface moisture, temperature and evapotranspiration products can also be used to characterize climate conditions. The SCORPAN factors have been interpreted from aerial photography, topographic

Table 1. Summary of important DEM-derived terrain attributes

Terrain attribute	Definition	Pedologic significance
Slope gradient (degrees) (sg)	Maximum rate of change (degrees)	Flow velocity
Slope aspect (sa)	Direction of slope gradient (degrees)	Flow velocity
Profile curvature	Curvature of slope gradient (radians)	Flow acceleration
Tangential curvature (radians)	Curvature perpendicular to slope gradient (radians)	Flow convergence
Multiresolution valley bottom flatness index	Measure of flatness and lowness	Depositional environments
Multiresolution ridge top flatness index	Measure of flatness and upness	Stable uplands settings
Contributing area (ca)	Upslope area (m ²)	Effective precipitation
Specific contributing area (sca)	ca/grid size (m)	Effective precipitation
Topographic wetness index (twi)	ln (sca/sg)	Water accumulation or soil saturation
Catchment height	Average upslope height	Potential energy
Catchment slope	Average upslope gradient	Flow velocity
Overland flow distance to stream	Relative topographic position (m)	Potential energy
Vertical distance to channels	Relative topographic position (m)	Cold air drainage
Solar radiation or insolation	Amount of incoming solar energy (kWh/mA ²)	Soil temperature and evapotranspiration

range of global issues, such as food production and hunger eradication; climate change and environmental degradation. Coordinated by FAO, several universities and institutes integrated the regional soil maps to a new comprehensive Harmonized World Soil Database (HWSD) which has been widely used. As per recent GlobalSoilMap specifications, there are 12 parameters to be predicted along with uncertainty over six depth intervals e.g. 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm and 100–200 cm.

Soil property values and uncertainty to be predicted are: Total profile depth (cm), Plant exploitable (effective) soil depth (cm), Organic carbon (g/kg), pH ($\times 10$), Sand (g/kg), Silt (g/kg), Clay (g/kg), Gravel (m^3/m^3), ECEC ($\text{cmol (c+)}/\text{kg}$), Bulk density of the fine earth fraction (excludes gravel) (mg/m^3), Bulk density of the whole soil in situ (includes gravel) (mg/m^3), available water capacity (mm).

Soil sampling strategy for DSM

Soil sampling is an important part of DSM. It largely determines the accuracy of DSM, and directly influences the cost of the mapping. In the last decade, significant progress has been made in developing soil sampling methods, with the purpose of obtaining the highest soil mapping performance with the least soil samples. Design-based sampling (simple random sampling, stratified random sampling, systematic random sampling, etc.), and model-based sampling (geostatistical sampling and centered grid sampling, etc.) are used in soil mapping strategy. Besides, the conditioned Latin hypercube sampling method has been popular for soil mapping in recent years. This method is a type of stratified random sampling that accurately represents the variability of environmental covariates in feature space. Lots of potential access constraints may prevent sampling at our desired locations. The constraints can be some kinds of land use, land cover, high-relief terrain, and unexpected road blocking, etc. Several scientists have

proposed a flexible Latin hypercube sampling strategy to deal with operational sampling problems in a large and remote region. It can be expected that more flexible and efficient soil sampling methods would be developed for DSM.

Predictive models

Numerous techniques were applied in the field of soil mapping to deal with the regression/classification of soil properties or soil classes, including statistical, geostatistical, hybrid and machine learning methods. A lot of case studies have been conducted to compare the performance of different predictive techniques, such as geostatistics and machine learning. The geostatistics approaches often generate similar or better performance than regression models. Some geostatistical techniques, such as kriging with an external drift, regression kriging and can incorporate environmental variables, and achieve better performance than ordinary/simple kriging. Classification and Regression Tree (CART) recursively produces a tree-based structure storing predictors and soil data. This technique has been widely adopted in the soil mapping case study due to its robust predictions. Random Forest (RF) is an up-to-date ensemble approach using recursive partitioning for exploring the relationships between predictors and the variable of interest. RF has the same model building process as CART, sharing similar model advantages (i.e. flexible with different types of datasets, and insensitive to missing data). Many trees can be built in RF, in which variables are randomly selected and only a subset of the variables can be adopted to identify splits in each tree. Therefore, compared with CART, RF can provide a better error measurement and is less susceptible to over-fitting, which has been validated in the mapping of regional soil type. Some hybrid approaches, such as Random Forest plus residuals kriging, were proposed in an effort to generate a reliable prediction.

Legacy soil database in India

Legacy soil data mainly includes polygon-based soil maps and soil profile databases. They both can serve as the input of digital soil mapping procedures. Different types of soil surveys at different scales and intensities have been carried out in the country by the ICAR-NBSS&LUP in collaboration with the various state soil survey and land use organizations. NBSS&LUP has completed a soil survey of the entire country at a scale of 1:250,000 as a part of the mission project on Soil Resource Mapping. Reconnaissance soil survey was conducted at 1:50,000 scale using toposheets and Landsat/IRS (Indian Remote Sensing) imagery by the ICAR-NBSS&LUP and State Land Use Survey of India (SLUSI) under which many districts were surveyed in different states of the country. Generally, the reconnaissance maps show an association of two or three soils in one map unit or delineation. Detailed soil survey was carried out by SLUSI and other state soil survey organizations at different scales (1:4000 to 1:10,000) to survey many catchments and command areas. ICAR-NBSS&LUP have also surveyed many research farms, watersheds across the country and recently >50 blocks were covered in different agro-ecological regions of the country at 1:10000 scales under a Land Resource Inventory programme.

Major progresses in India

DSM has moved from research phase to operational phase across the countries but no full-fledged research has been carried out in India. In India, DSM approach was used in predicting soil organic carbon and inorganic carbon stocks at 250 m resolution using limited observations. Quantitative soil information is important to identify the vulnerable areas of land degradation/desertification. This type of information helps in modelling soil losses and highlights the areas with the greatest degradation susceptibility. Recently, land degradation/desertification was mapped using global gridded

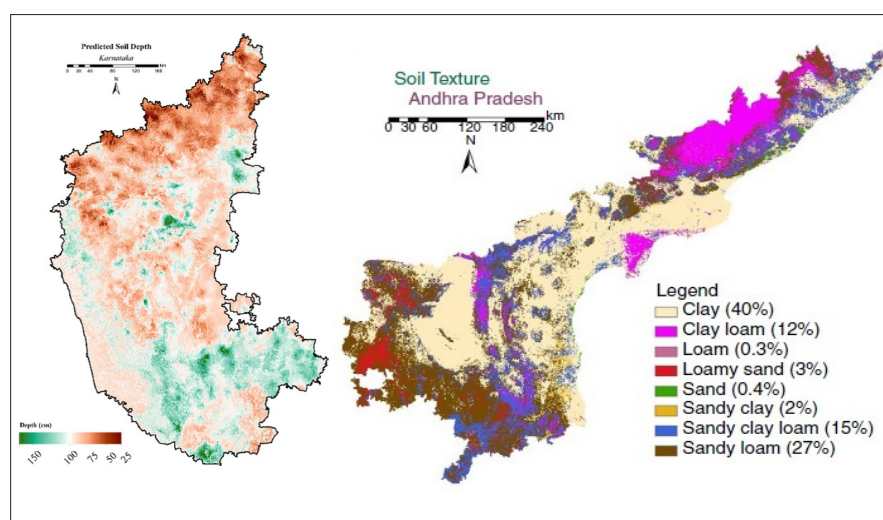
soil information such as coarse fragments, organic carbon, *pH* and clay content. DSM has moved towards an operational activity from research phase which has been concretized by the GlobalSoilMap project. It aims to map several key soil properties onto a three dimensional grid at fine spatial resolution with local uncertainty estimates throughout the world. First versions of GlobalSoilMap products have already been produced in various countries. There are few examples of DSM activities in India (Table 2). The DSM of major soil properties (OC, *pH* and EC) was predicted using Random Forest techniques.

Application of DSM

DSM originally focused on the soil-landscape modelling, which quantified the relationships between soil properties and environmental variables. The emphasis of soil mapping is gradually shifting from the soil variation and predictive techniques design to the application in various fields like agricultural management, ecosystem services assessment, land evaluation etc. The up-to-date multi-scale soil maps have been the new global imperative and have attracted much attribution. Like conventional soil mapping, the major stakeholders of digital soil maps are farmers, land use planners, researchers and policy makers. Carbon sequestration studies, land degradation assessment, water resource management and climate change analysis are some of the major fields in India where quantitative soil information is needed. For instance, soil carbon studies have become increasingly important for environmental reasons to know not just the carbon storage at particular time but also the changes or trend with time. The widespread availability of digital elevation data and other DEM attributes has led to soil organic carbon mapping possible at higher resolutions. DSM can be useful in regular monitoring of soil health card data launched by Government of India to help the farmers access the status of soil, to identify the crop in suitability, nutrient availability and

Table 2. Major DSM work progresses in India

State/Region	Soil properties	Model used
Western Ghats	Soil organic carbon stocks	Quantile Regression Forest
Karnataka	Soil Depth, Field capacity and permanent wilting point	Quantile Regression Forest and Regression Kriging
Bukkarayasamudrum Mandal, Andhra Pradesh	OC, <i>pH</i> and EC	Random Forest
Andhra Pradesh	Soil texture classes and soil depth	Random Forest, Quantile Regression Forest
Arid western India	Sand content	Ordinary Kriging, Universal Kriging/Kriging with external drift, Random Forest and Regression Kriging
Bikaner district of Rajasthan	Carbon content	Quantile Regression Forest



Predicted soil depth and texture map of Karnataka and Andhra Pradesh

providing further recommendation for improving the soil quality. Availability of digital soil maps helps in preparing soil health cards by providing site specific quantitative information and also in identifying the suitability of land resources for different crops. DSM can be used to create initial soil survey maps, refine or update existing soil surveys, generate specific soil interpretations, and assess risk. It can facilitate the rapid inventory, re-inventory, and project-based management of lands in a changing environment.

SUMMARY

Demands on soil related information have been significant worldwide and are still increasing. Due to the relatively high costs of new data collection, DSM approaches have taken over the role of traditional soil maps in the

field of data service. In the future, the regional and national soil maps should be harmonized together for the global issues. Several challenges need to be settled in future DSM studies, such as simulation of the soil heterogeneity at the regional scale, soil mapping at the flat terrain and linking soil mapping and soil spectroscopy. Since urbanization results in rapid land use change, urban areas exhibit high soil variation in space, which is a challenge to conventional models. There is a need to design more sophisticated technologies for modelling soil properties and processes at fine resolution and with high accuracy.

*Corresponding author's e-mail: pravashiari@gmail.com