A comparative Digital Soil Mapping (DSM) study using a non-supervised clustering analysis and an expert knowledge based model - A case study from Ahuachapán, El Salvador

Martín-López, Javier M. [1]**; da Silva Mayesse**[1]**; Jefferson Valencia**[1]**, Marcela Quintero**[1]**, Keough, Adam**[2] **Casares, Francisco**[2]

[1]International Center for Tropical Agriculture (CIAT), Agroecosystems and Sustainable Landscapes (ASL), Cali, Colombia
[2] Catholic Relief Services (CRS), San Salvador, El Salvador.
E-mail for contact: j.m.martin@cgiar.org.

DSM is the inference of spatial and temporal soil property variations using mathematical models based on quantitative relationships between environmental information and soil measurements. The quality of DSM information depends on the method and environmental covariates used for its estimations. We compared two DSM methods to predict soil properties such as Organic Matter "MO" (%), Sand (%), Clay (%), pH (H_2O), Phosphorus (mg/kg), Effective Cationic Exchange Capacity "CICE" (cmol/L), Potassium (cmol/L) and Water Holding Capacity (mm/m) for the department of Ahuachapán in El Salvador to support the activities of the Agriculture Landscape Restoration Initiative (ALRI) in the country.

Legacy soil data composed by 258 points at 30 cm depth, auxiliary information like geology, relief, landform classification, soil maps, and terrain attributes were used in both DSM methodologies. The terrain attributes (Cross-sectional and Longitudinal Curvature, Relative Slope Position, SAGA Wetness Index, Slope and Valley Depth) were derived from a 10 m-resolution Digital Elevation Model (DEM) using SAGA-GIS.

In the first method, we used a fuzzy k-means clustering algorithm performed in R-Studio. This methodology groups information into clusters based on its similarities, preserving that data from different clusters are as dissimilar as possible. Geology, relief, landforms (developed in Germorphons add on in GRASS-GIS) and terrain attributes were the inputs for generating 15 fuzzy soil units. Following DSM approach based on fuzzy membership logic, an R-Studio script was developed for estimating soil property maps employing a random sample of 70% points from legacy soil data.

For the second method, we used SIE, which is an expert knowledge model based on fuzzy membership functions for characterizing soil relationships that runs in ArcGIS (ArcSIE). SIE can create soil classes based on a classification given by a soil expert. This method was implemented using ArcSIE, which provides tools to prepare data, define soil-environment classification rules, and perform final maps. A first inference-ruled map was defined following the taxonomic description established by the University of El Salvador in: "Las nuevas clasificaciones y los suelos de El Salvador, 1974". In this part, we grouped five soil units (Entisol, Vertisol, Inceptisol, Alfisol and Andisol) in terms of its geology, relief, landform, and elevation description. A second inference-ruled classification was implemented using the zonal statistics (mean and standard deviation) from the terrain attributes as a rule-parameter for each soil unit. Property maps were obtained following the procedure described for the first method.

For the maps evaluation and comparison in both cases, the Root Mean Square Error (RSME) was performed employing 30% of legacy data points not used for mapping. Soil property maps presented errors of MO (%): 1.78 and 1.42, Sand (%): 6.7 and 5.5; Clay (%): 6.8-5.7; pH (H_2O): 0.58 and 0.55, Phosphorus (mg/kg): 21.12 and 16.14, CICE (cmol/L): 3.61 and 3.58, Potassium (cmol/L): 0.62 and 0.58 and Water Holding Capacity (mm/m): 12.59 and 10.17 for Fuzzy k-means and Expert knowledge method respectively. Results showed that the expert knowledge method provided a higher soil property prediction for all soil properties than the given by the fuzzy k-means clustering.

Keywords: k-means, Soil Inference Engine (SIE), fuzzy logic, terrain attributes, soil property maps.

Acknowledgements: This study was funded in part through the Agriculture Landscape Restoration Initiative (ALRI), financed by the Howard G. Buffett Foundation.