EGU2020-20873
https://doi.org/10.5194/egusphere-egu2020-20873
EGU General Assembly 2020

Digital soil mapping: the challenge to obtain the best soil dataset and create a precise environmental model to support land use management at a national level (Ecuador).

Daphne Armas[1], Mário Guevara[2], Fernando Bezares[3,4], Rodrigo Vargas[2], Pilar Durante[4,5], and Cecilio Oyonarte[3,4]

[1]Universidad de Almeria, Departamento de agronomia, Ecuador (daphne.armas@gmail.com)
[2]Department of Plant and Soil Science, University of Delaware, 531 South College Ave., 152, Townsend Hall, Newark DE, EEUU, 19716. (rvargas@udel.edu)
[3]Centro Andaluz para la Evaluación y Seguimiento del Cambio Global (CAESCG). Universidad de Almería. Crta. Sacramento s/n. 04120 La Cañada, Almería. España. (coyonarte@ual.es)
[4]Departamento de Agronomía, Edif. CITEIIB. Universidad de Almería. 04120 La Cañada, Almería. España. (fernandobezares95@hotmail.es)
[5]Agresta Sociedad Cooperativa. C/ Duque de Fernán Núñez, 2, 1º. 28012 Madrid. España. (pdurante@egrata.org)

One of the biggest challenges for digital soil mapping is the limited of field soil information (e.g., soil profile descriptions, soil sample analysis) for representing soil variability across scales. Global initiatives such as the Global Soil Partnership (GSP) and the development of a **Global Soil Information System** (GloSIS), World Soil Information Service (WoSis) or SoilGrids250m for global pedometric mapping highlight new opportunities but the crescent need of new and better soil datasets across the world. Soil datasets are increasingly required for the development of soil monitoring baselines, soil protection and sustainable land use strategies, and to better understand the response of soils to global environmental change. However, soil surveys are a very challenging task due to their high acquisition costs such data and operational complexity. The use of legacy soil data can reduce these sampling efforts.

The main objective of this research was the rescue, synthesis and harmonization of legacy soil profile information collected between 2009 and 2015 for different purposes (e.g., soil or natural resources inventory) across Ecuador. This project will support the creation of a soil information system at the national scale following international standards for archiving and sharing soil information (e.g., GPS or the GlobalSoilMap.net project). This new information could be useful to increase the accuracy of current digital soil information across the country and the future development of digital soil properties maps.

We provided an integrated framework combining multiple data analytic tools (e.g., python libraries, pandas, openpyxl or pdftools) for the automatic conversion of text in paper format (e.g., pdf, jpg) legacy soil information, as much the qualitative soil description as analytical data, to usable digital soil mapping inputs (e.g., spatial datasets) across Ecuador. For the conversion, we used text data mining techniques to automatically extract the information. We based on regular

expressions using consecutive sequences algorithms of common patterns not only to search for terms, but also relationships between terms. Following this approach, we rescued information of 13.696 profiles in .pdf, .jpg format and compiled a database consisting of 10 soil-related variables.

The new database includes historical soil information that automatically converted a generic tabular database form (e.g., .csv) information.

As a result, we substantially improved the representation of soil information in Ecuador that can be used to support current soil information initiatives such as the WoSis, Batjes et al. 2019, with only 94 pedons available for Ecuador, the Latin American Soil Information System (SISLAC, http://54.229.242.119/sislac/es), and the United Nations goals towards increasing soil carbon sequestration areas or decreasing land desertification trends. In our database there are almost 13.696 soil profiles at the national scale, with soil-related (e.g., depth, organic carbon, salinity, texture) with positive implications for digital soil properties mapping.

With this work we increased opportunities for digital soil mapping across Ecuador. This contribution could be used to generate spatial indicators of land degradation at a national scale (e.g., salinity, erosion).

This dataset could support new knowledge for more accurate environmental modelling and to support land use management decisions at the national scale.