2019 Joint workshop for Digital Soil Mapping and GlobalSoilMap
12-16 March 2019 – Universidad de Chile, Santiago, Chile

Advances in soil data availability for large scale digital soil mapping (GlobalSoilMap) in Brazil

Alessandro Samuel-Rosa[1]

[1] *Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brazil. E-mail: alessandrorosa@utfpr.edu.br*

Digital soil mapping (DSM) has benefited from the rapid increase in the power of computers to deal with large volumes of base data to produce new soil information. These base data include the environmental covariate data, which have become widely available at various spatial resolutions thanks to the quick development of remote sensing technology. Unfortunately, in several countries, the increase in the availability of the point soil data needed to calibrate DSM models has not been as fast. In Brazil, soil surveys have been carried out since the 1930's. But most of the data collected along the last 90 years still is in paper format. Some of them may already have been lost forever. Very few institutions worked to compile and openly share the existing legacy soil data. Until 2016, the largest soil database in Brazil – Sistema de Informação de Solos Brasileiros, SISB – contained only about 9000 soil observations. Less than half of those contained spatial coordinates. In comparison, Australia, that is about the size of Brazil, contains data from almost 300 000 soil observations in its national soil database – National Soil Site Collation, NSSC. In the beginning of 2017, several Brazilian soil scientists from various universities and research institutions decided to change this scenario. The Free Brazilian Repository for Open Soil Data, febr, was born – building on the work done in SISB. By the end of 2018, febr already contained data from about 15 000 soil observations. About half of these data was collected before the 1990's, when the official national soil survey program was stopped. Different from previous efforts, febr also aims at improving the quality and usability of legacy soil data for DSM. For example, existing spatial coordinates are checked for positional accuracy and missing ones are estimated – with a precision of a few hundred meters – using auxiliary environmental data. Only 20% of the observations still remain without spatial coordinates. Moreover, a series of (combined) automated and manual data validation routines have been implemented to guarantee the correctness of the soil property data. Most of the manual tasks are performed by undergraduate students under the supervision of soil scientists. Some undergraduate soil science courses are already including the development of activities in febr in their syllabus – such as legacy soil data compilation and validation. The formal involvement of undergraduate students is seen as the key to promote a cultural change towards a richer soil data sharing environment in Brazil.

Keywords: Legacy soil data, Open soil data, Soil data validation, Soil data cleansing, Soil data wrangling