Legacy data: exploratory analysis to digital soil class mapping.

Carvalho Junior, Waldir[1]; Bhering, Silvio Barge[1]; Pereira, Nilson Rendeiro[1]; Chagas, Cesar da Silva[1]; Lopes, Carlos Henrique Lemos[2].
[1]: Embrapa Solos, {waldir.carvalho; silvio.bhering; nilson.pereira; cesar.chagas}@embrapa.br; [2]: Secretaria de Estado de Meio Ambiente, Desenvolvimento Econômico, Produção e Agricultura Familiar – SEMAGRO, chlopes@semagro.ms.gov.br

Thematic Session: Legacy Data – how to turn it usefull

Abstract

The use of legacy data improve the value of these data and promote the input of new data on the database. Also, reduces the primary data collection and analytical procedures. In MS State, the legacy data was 1325 soil profiles. We are looking for correlations with previus soil map, and litology and vegetation maps to reveal the profiles distribution. The DEM and derivatives was used also to compare with the profiles distribution. The spatial distribution of soil profiles follow the categorical maps units occurrence. The statistical values of DEM and slope are not significantly different. The spatial distribution of soil profiles can represent the covariates of the all area and are ready to be used to produce digital soil maps.
Keywords: soil database; Mato Grosso do Sul; Tropical.

Introduction

The use of legacy data for soil surveys is important to reducing primary data collection, and enhance available data that could otherwise be neglected.

Looking for a methodology to verify the use of these legacy data against some predictor covariates, an exploratory spatial analysis of legacy data over the Paraguay river basin, in the state of Mato Grosso do Sul, with the exception of the wetland, was carried out (Figure 01).

The objective of this work is to verify the adequacy of the spatial distribution of the legacy data as a function of covariates such as altimetry, slope, lithology, biomes, vegetation and soils maps in the study area.

Methodology

The dataset has 1325 soil profiles collected in previous works, without the use of statistical sampling techniques, important step in digital soil mapping. These data belong to the ZAE MS project of Embrapa Solos and will soon be available in the institution's databases. Numerical altimetry and slope covariates were obtained from NASA JPL (2020). Thematic covariates on Biomes, Lithology, Soils and vegetation were obtained from BDiA – IBGE (2021).

The study area is included in the cerrado biome, and represents approximately 27% of the state's area, with 96,960 km2. Figure 01 shows the study area and thematic covariates used in this study, namely, lithology, soils and vegetation.

Results and discussion

The DEM and slope covariates has the follow characteristics (Table 01). The DEM values are close between study area and soil samples. Apparently the slope has different distribution, but areas with slope above 45 % are less than 01% of the study area, and we can consider that the relation are maintained.

Table 02 shows 1325 soil samples grouped by class according to FAO (2021). Note an unbalanced distribution of classes, with predominancy of Ferralsols, Arenosols and Acrisols. The covariates are showed in Figure 01.

Table 01. Statistics values of DEM (meters) and slope (%) of study area and soil samples.

	min	max	mean	SD
DEM study area	73	942	334	146
Slope study area	0	370	7.5	7.9
DEM soil profiles	76	890	344	146
Slope soil profiles	0	52	5.5	4.8

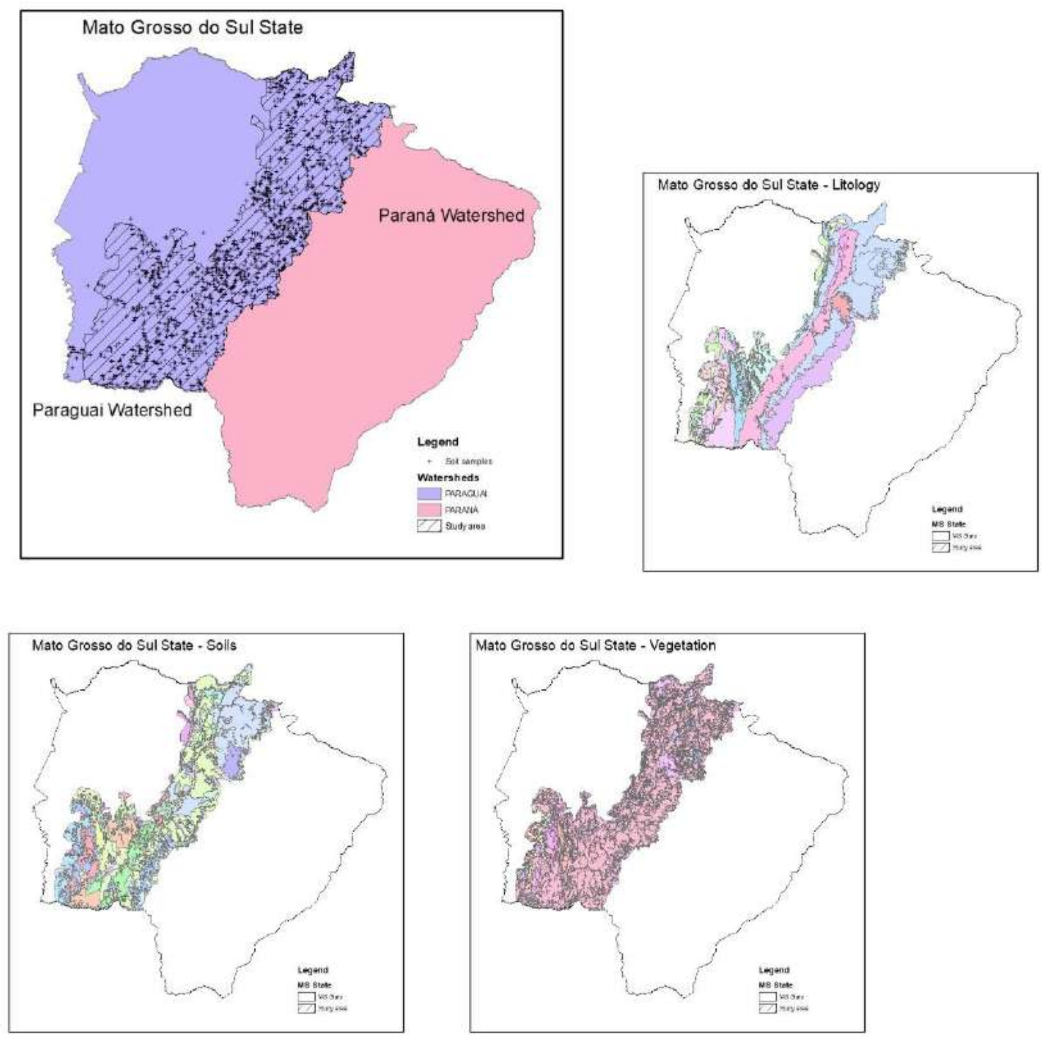

Figure 1. Study area with soil samples locations and covariates

Table 02. Samples Soil Class distribution according FAO soil taxonomy (FAO, 2021)

FAO Class of soil samples	count	%
Acrisols	145	10.9
Arenosols	250	18.9
Cambisols	45	3.4
Chernozems	42	3.2
Ferralsols	540	40.8
Fluvisols	1	0.1
Gleysols	46	3.5
Leptosols	69	5.2
Luvisols	4	0.3
Nitisols	34	2.6
Planosols	50	3.8
Plinthosols	46	3.5
Regosols	35	2.6
Vertisols	18	1.4

The result of join the soil samples and soil map, shows that the soil samples distribution has a quite relation with soils polygons distributions. The Table 3 shows this relation. We can note that only one soil unit map (Histosols) doesn´t have soil samples and that the greater soil units have the greater account of soil samples. The idea is to analysis which soil samples are within each soil unit, to best understand the soil sample distribution in relation to soil map, but the quantitative aspects reveals that the soil samples distribution follow the same distribution of soil unit map.

Table 3. Soil units map, percent and km2 distribution and soil samples within each soil unit.

Soil units	%	km2	soil samples
Rock outputs	0.6	595	3
urban	0.0	20	0
water	0.1	89	0
Plinthosols	1.8	1,785	21
Gleysols	2.0	1,976	20
Ferralsols	25.9	25,120	482
Chernozems	3.7	3,619	27
Nitisols	3.5	3,352	66
Histosols	0.0	8	0
Acrisols	12.4	12,056	205
Leptosols	10.9	10,571	73
Arenosols	20.2	19,625	272
Regosols	8.9	8,659	72
Planosols	7.6	7,343	64
Vertisols	2.2	2,143	20

The vegetation covariate also has a relation between map units and soil samples spatial distribution. The same correlation occurs between soil samples and lithology map units. These correlations can be noted in Table 4.

Table 4. Correlation between soil samples spatial distribution and map units of vegetation and lithology. Note that only the lithology map units with soil samples are showed.

Lithology map units	%	km^2	soil samples	Vegetation map units	%	Km^2	soil samples
Sand deposits	1.1	1,049	15	Contact	32.2	31,260	370
amphibolite	0.1	57	1	Water	0.1	86	
Arches, Conglomerate	0.3	323	6	Deciduous Seasonal Forest	2.9	2,777	18
Sandstone	37.0	37,300	524	Seasonal Semideciduous Forest	1.0	928	7
Biotite	9.0	8,750	81	Savannah	62.6	60,704	920
Limestone	4.1	3,999	57	Savannah-estepe	1.2	1,195	10
dacitus	12.2	11,806	205	Savannah-estepe	1.2	1,195	10
clay deposit	6.5	6,260	62				
Diamictite, Shale	15.2	14,705	217				
Philito	4.8	4,607	59				
shale	2.0	1,943	44				
Marble	1.1	1,043	20				
quartzite	1.4	1,342	22				
Schist	3.4	3,312	12				

Conclusions

The spatial location of soil profiles follows the spatial distribution of the covariates, which denotes that the spatial distribution of soil profiles can represent the covariates of the all area and can be used to produce digital soil maps.

References.

BDia – IBGE (2021). Banco de Informações Ambientais. from https://bdiaweb.ibge.gov.br/#/home. Accessed August 18, 2021

FAO (2021). World Reference Base For soil Resources 2014. http://www.fao.org/soils-portal/data-hub/soil-classification/world-reference-base. Accessed August 19, 2021.

NASA JPL (2020). *NASADEM Merged DEM Global 1 arc second nc V001*. NASA EOSDIS Land Processes DAAC. Accessed 2020-08-28 from https://doi.org/10.5067/MEaSUREs/NASADEM/NASADEM_NC.001. Accessed August 28, 2020