USE OF PROXIMAL SENSOR FOR SOIL CLASSES SEPARATION USING PRINCIPAL COMPONENT ANALYSIS (PCA)

Valéria Ramos Lourenço[1], David Bruno Sousa Teixeira[2], Carlos Alexandre Gomes Costa[3], Márcio Rocha Francelino[4]

[1]Departamento de Solos e Nutrição de Plantas - Universidade Federal de Viçosa, Av. Peter Henry Rolfs s/n, Viçosa – Minas Gerais, ramos.valeria74@yahoo.com; [2]Departamento de Engenharia Agrícola - Universidade Federal de Viçosa, Av. Peter Henry Rolfs s/n, Viçosa – Minas Gerais, teixeiradbs@gmail.com; [3]Departamento de Engenharia Agrícola – Universidade Federal do Ceará, Av Mister Hull s/n, Fortaleza – Ceará , costacag@gmail.com[4]Departamento de Solos e Nutrição de Plantas - Universidade Federal de Viçosa, Av. Peter Henry Rolfs s/n, Viçosa – Minas Gerais, marcio.francelino@gmail.com.

ABSTRACT

The spectrally active components of the soil allow the realization of integrative analyzes of soil aspects such as their classification. The objective of this study was to evaluate the separation of soil classes from spectral reflectance data using main components analysis (PCA). The study was carried out in the Aiuaba Experimental Basin located in the municipality of Aiuaba-CE. Soil samples were collected in Ustalfs, Ustults and Ustorthents profiles. The samples were submitted to spectral analysis and subsequent PCA analysis, CPs were used to separate the soil classes using the two-dimensional graphical analysis. From the analysis of spectral behavior data from the different soil classes, it was possible to separate the class of Ustorthents from the Ustalfs and Ustults.

Key words — Ustalfs, Ustults, Ustorthents, Reflectance.

1. INTRODUCTION

Due to the exponential growth in data availability in several areas, a new era of information processing, the big data, has started, with which new needs raised, such as the processing capacity of these data. Facing this, there were proposed techniques and algorithms of processing that aim to meet this need, and that can return information from this data density. [1], [2].

In the field of soil science, the great variability of soils and its attributes have always involved large databases and, with the advent of new technologies that make it easier to acquire, this characteristic is more and more prominent. Studies that depart from surveys using orbital [3] - [6] or proximal sensors, such as those of [7] that estimated the concentration of heavy metals in the soil correlating the spectra of these elements with their reflectance, or [8] that predicted textural classes from soil reflectance data, both produce dense databases that require large processing capacities.

Another soil aspect dependent on numerous variables is its classification from a pedogenic point of view. This has been the subject of several studies using proximal sensors, as done by [9] that aimed to predict soil classes through their spectral characteristics. The fact of the spectrally active components of the soil in the visible and near infrared (Vis-NIR) regions are mainly iron oxides, organic matter, clay minerals, carbonates and water [10], [11] integrative analyzes of soil aspects such as degradation [12], soil quality [13] and their classification [14], [15] are possible.

However, the problem with the high density of the databases generated by analyzes performed with proximal sensors still persists. As one of the alternatives, we have the main components analysis (PCA), which transforms high density datasets into components representing the initial data, having a smaller size and preserving the data variability as much as possible [1], [2].

In this way, the objective of this study was to evaluate the separation of soil classes from spectral reflectance data using main components analysis.

2. MATERIAL AND METHODS

The study was conducted in the Aiuaba Experimental Basin (BEA), located in the municipality of Aiuaba-CE in the Inhamuns microregion, currently considered the largest Federal reserve of the Caatinga biome. The current area of the basin is 12 km² (Figure 1) and it is fully inserted in the ecological station (ESEC), which means that it is a preserved area.

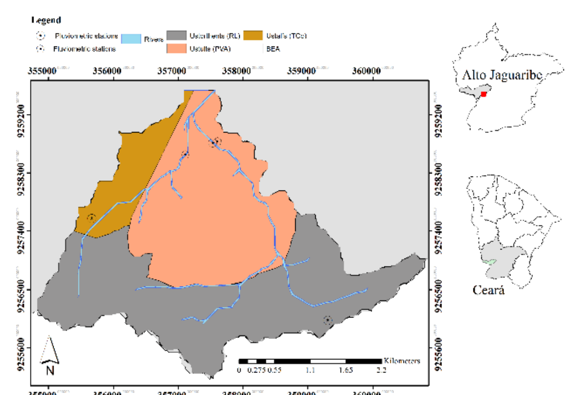

Figure 1. Aiuaba Experimental Basin (BEA), its location in the State of Ceará, in the Upper Jaguaribe Basin, and in the Bengué Representative Basin, with the distribution of soil classes and basin hydrography.

The climate of the region is defined as 'Bs' according to the Köppen's classification, presenting average precipitation of 560 mm year^{-1}; evaporation of the 2500 mm year^{-1} by class A tank [16].

Based on the assessment of the environmental dynamics in the basin, the area was divided into three associations between soil and vegetation (ASV's) that were defined as homogeneous units for studies of environmental variables. [17], [18] classified the ASVs based on their predominance of soil and vegetation, which can be observed in Table 1.

Table 1: Characterization of the components of the Soil Vegetation Associations (ASV's) in the Aiuaba Experimental Basin (BEA).

ASV	Predominant vegetation	Soil group	Area in BEA (%)
ASV 1	Catingueira (*Caesalpinia pyramidalis* Tul)	Ustalfs (TCo)	20
ASV 2	Angelim (*Piptadenia obliqua*)	Ustults (PVA)	34
ASV 3	Jurema-preta (*Mimosa tenuiflora* (Willd.) Poir)	Ustorthents (RL)	46

Source: [18].

Four soil samples were collected in characteristic profiles at each ASV. The samples were collected at a soil depth of 0-0.2 m. Subsequently, they were stored and identified for further analysis.

Analyzes were carried out to determine the texture and organic matter (OM) content according to the methodology [19].

In order to obtain the soil reflectance data using a proximal sensor, the FielSpec 3 sensor was used according to the methodology of [20]. This sensor has resolution of 1 nm (350-1100 nm) and 2 nm (1000-2200 nm). The system geometry was based on the perpendicular positioning of the sensor in relation to the sample, maintaining the distance of 6 cm. The light source was positioned at 50 cm from the sample, forming a 45° angle with the zenith. The spectral reference standard used was a white spectral plate. The reflectances were obtained from the average of three readings for each sample.

After obtaining the data, they were submitted to main component analysis (PCA), using the software R and the FactorMaineR package scripts provided by [21], in order to identify the most important components and the separation of the soil classes.

3. RESULTS

The results of the textural class and OM can be observed in Table 2.

Table 2: Texture and organic matter, and their deviations, for the three soil classes analyzed.

Soil Group	Organic matter	Texture			
		Sand	Silt	Clay	Coarse sand
	(g.kg^{-1})	(g.kg^{-1})			
Ustalfs (TCo)	15,1 (13,5)	0,36 (±2,2)	0,45 (±4,2)	0,18 (±3,0)	0,18 (±1,5)
Ustults (PVA)	31,7 (4,5)	0,35 (±5,5)	0,25 (±1,7)	0,32 (±12,9)	0,14 (±4,0)
Ustorthents (RL)	12,8 (3,0)	0,68 (±15,8)	0,22 (±5,7)	0,09 (±7,6)	0,35 (±7,9)

The highest OM contents were observed in the Ustults, while the lowest in the Ustorthents. A highlight is given to the silt content of Ustalfs, which may be associated with its poor pedogenetic development. The highest sand contents, as expected, were observed in the Ustorthents.

In relation to the spectral response curves of the soils obtained with the proximal sensor, the soil classes of the three ASVs presented a quite different behavior as can be observed in Figure 2.

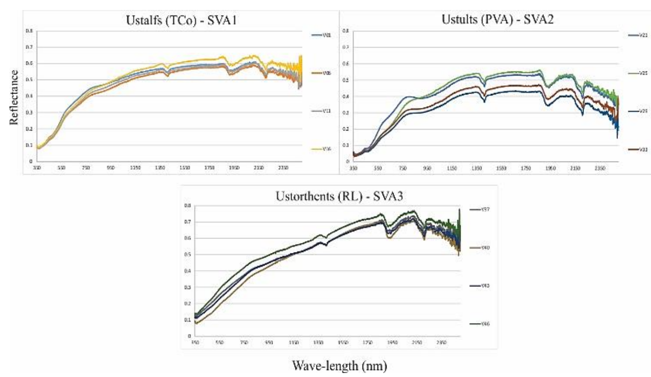

Figure 2. Spectral curve of the soil classes obtained with proximal sensor.

It was observed a greater homogeneity of spectral response for the samples belonging to the Ustalfs, a greater differentiation was observed only after the wavelength of 1350 nm. While in the other two classes, greater heterogeneity of responses was observed at most wavelengths.

Regarding the separation of the wavelengths that compose the soil reflectance, the following data separation in main components (CP) was observed in the PCA analysis (Figure 3).

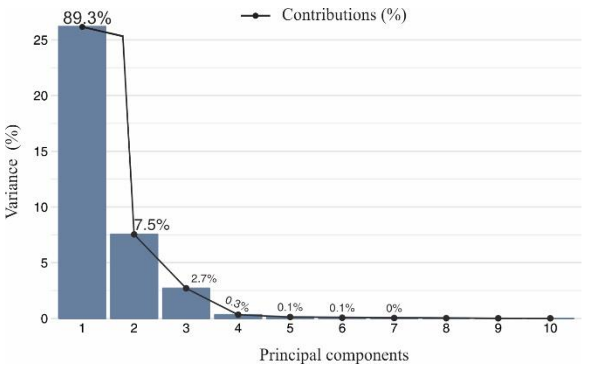

Figure 3. Variation of the first seven major components of the reflectance spectrum of the soils and the contribution of the spectral data in the construction of these.

It was observed that the greatest variability of the data is in the first main component, and significant variability is observed until the third CP, which explains 99.4% of the variation of the spectral data. The contribution of the number of reflectance bands in the construction of the CPs is strongly concentrated in the first two components, which indicates a low contribution in the variability when considering the spectral data that constitute the other main components.

After the treatment of the spectral data through the PCA, the first three CPs were used to separate the soil classes using the two-dimensional graphical analysis (Figure 4).

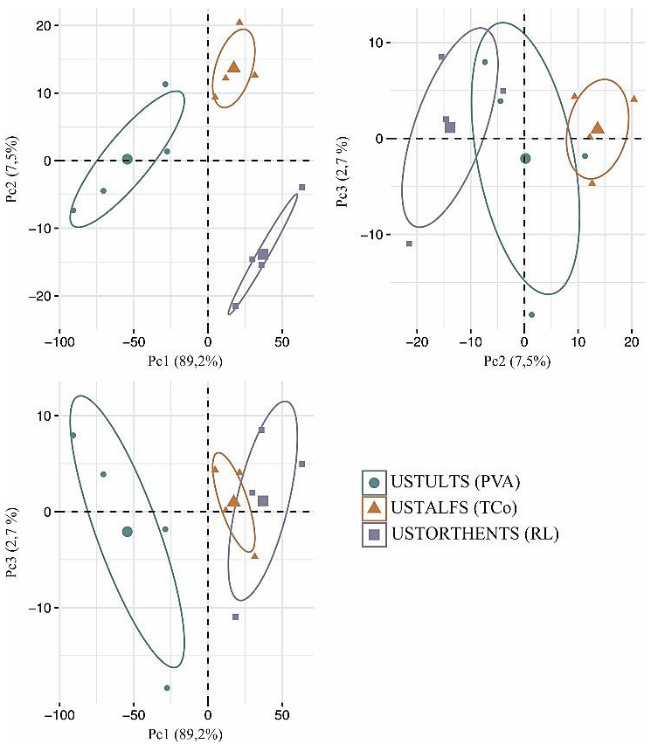

Figure 4. Two-dimensional analysis of the spectral data of Ustalfs, Ustults and Ustorthents.

It can be observed in the comparison between the first two CPs that the spectral data are sufficiently discriminatory to perform the separation between the Ustalfs, Ustults and Ustorthents. The distinction between Ustorthents and Ustalfs was not well performed in the other two-dimensional analyzes. It was observed that, in the analyzes between CPs with lower variability, the separation between soil classes could not be performed.

4. DISCUSSION

The selective absorption by the soil components, mainly iron oxides, OM and the constituents of the clay fraction, make possible the use of Vis-NIR in the evaluation of soil properties [9], [10]. The wavelengths that presented the greatest contribution in the formation of the three main components (PCs) were the intervals of 1412-1420; 750-760 and 350-380 nm, respectively. These wavelengths are mainly associated with the reflectance of clay minerals, 1:1 (kaolinite), and 2:1 (smectite, mica and illite) which have a strong signal between 1400-2200 nm. However, organic matter is characterized in the range of 750-870 nm, its chromophore characteristic is mainly given by the combination C-H, N-H and by the vibration of these elements [9]. The range of 380-430 nm and 480-550 nm are absorption ranges characteristic of iron oxides such as hematite and goethite [22], [23].

When the greatest variability in the first CP is observed, this variation should be associated with greater heterogeneity of the reflectances in all soils evaluated in the range of 1350-1450 nm (Figure 2), which corroborates with the bands that contributed the most in the construction of the first component, which certainly cooperated for the distinction of soil classes in the two-dimensional analysis.

The Ustorthents class, which presented the greatest differentiation both in the reflectance curve and in the two-dimensional analysis, is characterized by a higher albedo response and thus an upward growth of its curve (Figure 2) as well as an increased reflectance near the infrared (SWIR, 1200-2500 nm), which could explain the lower amount of iron oxides, and the higher amounts of quartz present in the soil [20].

The separation of the classes using PC2 and PC3 were not possible due to the low accumulated variability of their constituents, whereas the non-separation of Ustrorthents and Ustalfs in the third moment may be associated to the spectral behavior of the soils in the range 350-450 nm that contributed more strongly to the construction of CP3. This range makes the characterization of iron oxides, and due to the incipient degree of pedogenetic development of the classes [24], [25] the reflectance variation may not have been representative enough for the separation.

5. CONCLUSIONS

From the use of spectral behavior data from different soil classes, submitted to main components analysis, it was possible to separate the class of Ustorthents from the Ustalfs and Ustults soil classes.

6. REFERENCES

[1] I. T. Jolliffe, J. Cadima, and J. Cadima, "Principal component analysis : a review and recent developments Subject Areas : Author for correspondence :," *Philos. Trans. a*, 2016.

[2] A. Wu, X; Wai, H; Li, L; Scaglione, "A Review of Distributed Algorithms for Principal Component Analysis," *Proceeding IEEE*, vol. 106, no. 8, pp. 1321–1340, 2018.

[3] J. George Kalambukattu, S. Kumar, and R. Arya Raj, "Digital soil mapping in a Himalayan watershed using remote sensing and terrain parameters employing artificial neural network model," *Environ. Earth Sci.*, vol. 77, p. 203, 2018.

[4] A. Laborczi, G. Szatmári, K. Takács, and L. Pásztor, "Mapping of topsoil texture in Hungary using classification trees," *J. Maps*, vol. 12, no. 5, pp. 999–1009, 2016.

[5] C. Liddicoat, D. Maschmedt, D. Clifford, R. Searle, T. Herrmann, L. M. Macdonald, and J. Baldock, "Predictive mapping of soil organic carbon stocks in South Australia's agricultural zone," *Soil Res.*, vol. 53, no. 8, pp. 956–973, 2015.

[6] M. L. Mendonça-Santos, R. O. Dart, H. G. Santos, M. R. Coelho, R. L. L. Berbara, and J. F. Lumbreras, "Digital Soil Mapping of Topsoil Organic Carbon Content of Rio de Janeiro State, Brazil," in *Digital Soil Mapping*, Dordrecht: Springer Netherlands, 2010, pp. 255–266.

[7] H. Cheng, R. Shen, Y. Chen, Q. Wan, T. Shi, J. Wang, Y. Wan, Y. Hong, and X. Li, "Geoderma Estimating heavy metal concentrations in suburban soils with re fl ectance spectroscopy," *Geoderma*, vol. 336, no. August 2018, pp. 59–67, 2019.

[8] D. Curcio, G. Ciraolo, F. D. Asaro, and M. Minacapilli, "Prediction of soil texture distributions using VNIR-SWIR reflectance spectroscopy," *Procedia Environ. Sci.*, vol. 19, pp. 494–503, 2013.

[9] X. Xie and A. Li, "Geoderma Identi fi cation of soil pro fi le classes using depth-weighted visible – near- infrared spectral re fl ectance," *Geoderma*, vol. 325, no. April, pp. 90–101, 2018.

[10] E. Ben-Dor, "Quantitative remote sensing of soil properties," pp. 173–243, 2002.

[11] B. Stenberg, R. A. Viscarra Rossel, A. M. Mouazen, and J. Wetterlind, "Visible and Near Infrared Spectroscopy in Soil Science," pp. 163–215, 2010.

[12] A. O. Awiti, M. G. Walsh, K. D. Shepherd, and J. Kinyamario, "Soil condition classification using infrared spectroscopy: A proposition for assessment of soil condition along a tropical forest-cropland chronosequence," 2007.

[13] M. S. Askari, S. M. O'rourke, and N. M. Holden, "Evaluation of soil quality for agricultural production using visible-near-infrared spectroscopy," *Geoderma*, vol. 243–244, pp. 80–91, 2015.

[14] Y. Ogen, N. Goldshleger, and E. Ben-Dor, "3D spectral analysis in the VNIR–SWIR spectral region as a tool for soil classification," *Geoderma*, vol. 302, pp. 100–110, Sep. 2017.

[15] R. A. Viscarra Rossel and R. Webster, "Discrimination of Australian soil horizons and classes from their visible-near infrared spectra," *Eur. J. Soil Sci.*, vol. 62, no. 4, pp. 637–647, Aug. 2011.

[16] J. C. de Araújo and J. I. González Piedra, "Comparative hydrology: analysis of a semiarid and a humid tropical watershed," *Hydrol. Process.*, vol. 23, no. 8, pp. 1169–1178, Apr. 2009.

[17] C. A. Gomes Costa, J. Wellington, B. Lopes, E. Alves, R. Pinheiro, J. Carlos De Araújo, R. Rodrigues, and G. Filho, "Spatial behaviour of soil moisture in the root zone of the Caatinga biome 1 Comportamento espacial da umidade do solo na zona das raízes do Bioma Caatinga," pp. 685–694, 2013.

[18] E. Alves Rodrigues Pinheiro, K. Metselaar, Q. de Jong van Lier, and J. C. de Araújo, "Importance of soil-water to the Caatinga biome, Brazil," *Ecohydrology*, vol. 9, no. 7, pp. 1313–1327, Oct. 2016.

[19] A. TEIXEIRA, P C.; DONAGEMMA, G K.; FONTANA and W. G. TEIXEIRA, *Manual de métodos de análise de solo*, EMBRAPA. 2017.

[20] D. J. Romero, E. Ben-dor, J. A. M. Demattê, A. Barros, L. Eduardo, T. Rodrigues, M. Martello, T. Fernanda, P. Paulo, P. Ricardo, B. Cristina, M. Vinicius, and M. Tonini, "Geoderma Internal soil standard method for the Brazilian soil spectral library : Performance and proximate analysis," *Geoderma*, vol. 312, no. September 2017, pp. 95–103, 2018.

[21] S. Lê, J. Josse, and F. Husson, "FactoMineR : An R Package for Multivariate Analysis," *J. Stat. Softw.*, vol. 25, no. 1, 2008.

[22] S. J. Parikh, K. W. Goyne, A. J. Margenot, F. N. D. Mukome, and F. J. Calderón, "Soil Chemical Insights Provided through Vibrational Spectroscopy," *Adv. Agron.*, vol. 126, pp. 1–148, Jan. 2014.

[23] D. M. Sherman and D. T. Waite, "Electronic spectra of Fe3 + oxides and oxide hydroxides in the near IR to near UV," *Am. Mineral.*, vol. 70, pp. 1262–1269, 1985.

[24] Oliveira, L. B.; Fontes, M. P. F.; Ribeiro, M. R.; Ker, J. C. "Morfologia e classificação de Luvissolos e Planossolos desenvolvidos de rochas metamórficas no Semiárido do nordeste brasileiro," vol. 33, pp. 1333–1345, 2009.

[25] Oliveira, L. B.; Fontes, M. P. F.; Ribeiro, M. R.; Ker, J. C. "Micromorfologia e gênese de Luvissolos e Planossolos desenvolvidos de rochas metamórficas no semi-árido brasileiro," *Rev. Bras. Ciência do Solo*, vol. 32, pp. 2407–2423, 2008.