

Principal component analysis (PCA) of buried archaeological remains by VIS-NIR spectroscopy

Yoon Jung Choi^{a, b}, Johannes Lampel^b, David Jordan^c, Sabine Fiedler^a and Thomas Wagner^b

KEY-WORDS: soil spectroscopy, visible-to-near infrared soil spectroscopy, *in situ* measurements

INTRODUCTION

Buried archaeological remains are sometimes visible on the ground surface after ploughing. Such features are detectable to the naked eye, aerial photography or airborne remote-sensing, due to a difference in soil colour between the archaeological remains and the surrounding natural background soil. This work tries to analyse the spectral difference between buried archaeological remains and natural soil in the visible-to-near infrared range using the principal component analysis (PCA) method.

METHODOLOGY

The study was carried out on *in-situ* soil spectra of the excavated sections of six pits from Calabria, Italy. One of the pits contained a red archaeological stratum in the profile, which was clearly distinguishable from the natural soil. For each pit, at least three spectral measurements were taken for every layer using an Analytical Spectral Devices (ASD) spectrometer with artificial halogen light. The instrument has a spectral range of 350–2500 nm, but only wavelengths between 400–2400 nm were applied to minimise the influence of noise. These spectra were normalised by a continuum removal (Clark and Roush 1984) method to emphasise the absorption features and then analysed with the principal component analysis method.

Principal component analysis (PCA) is a multivariate chemometric method. It is a common tech-

^aJohannes Gutenberg University, Mainz, Germany

^bMax Planck Institute for Chemistry, Mainz, Germany

^cLiverpool John Moores University, Liverpool, United Kingdom

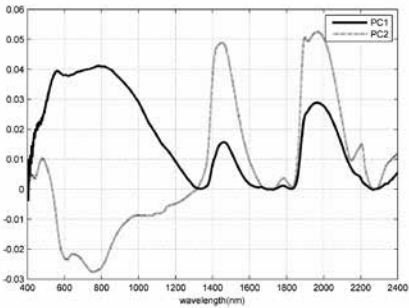


Fig. 1. Intensity of the first (PC1) and second (PC2) principal component

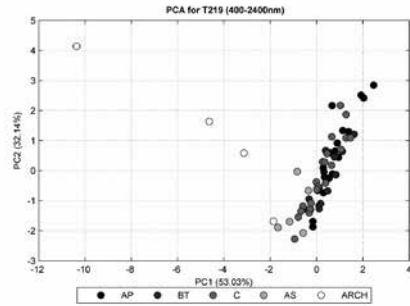


Fig. 2. PC1–PC2 plotting with different soil horizons: AP – distinct topsoil disturbed by ploughing; BT – B horizon (subsoil) enriched with clay; C – underlying unconsolidated material (parent material); AS – archaeological horizon; ARCH – archaeological material (burned soil and ceramics)

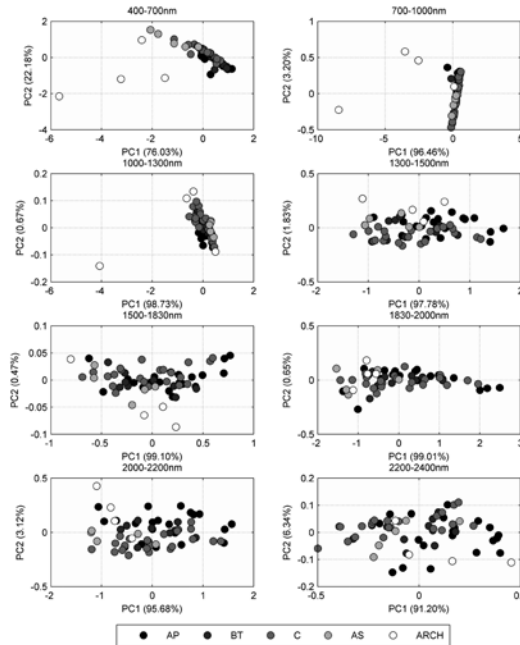


Fig. 3. PC1–PC2 plot at different wavelengths: AP – distinct topsoil disturbed by ploughing; BT – B horizon (subsoil) enriched in clay; C – underlying unconsolidated material (parent material); AS – archaeological layer; ARCH – archaeological material (burned soil and potsherds)

nique used to reduce the dimensionality of a dataset (Smith *et al.* 1985) and to discover new variables called principal components (PCs), which account for most of the variability in the data. The first principal component (PC1) represents the most dominant features among the spectra and the second principal component (PC2) represents the second most common features and so on. In soil spectroscopy, PCA is widely used for spectra comparison (Viscarra Rossel *et al.* 2009), mineral determination (Smith *et al.* 1985) and to select the end member spectra for the subsequent application of the linear spectral unmixing model (Galvão *et al.* 2001). In this study, soil reflectance spectra will be analysed using the PCA method to distinguish spectral features related to archaeological remains and natural soil.

RESULTS AND DISCUSSION

The first three principal components accounted for about 53%, 32%, and 10% of the total variation of the measured spectra, respectively, with a cumulative variance of about 95%. Figure 1 shows the first and second principal PC features. Here, the most dominant feature (PC1) is shown as a broadband feature in the 400–1200 nm spectral range. The near infrared region shows common features at the water absorption bands (1400 and 1900 nm). PC2 shows some absorption features in the visible spectral range. The near infrared region of PC2 is dominated by the water absorption bands (1400, 1900, 2200 nm) and small features (e.g., at 1800 nm and 2100 nm) are also included. By looking at the intensity of PC1 and PC2, we can expect that most spectral differences of soil spectra are concentrated below 1200 nm.

Figure 2 shows the scores of PC1 plotted versus those of PC2. We can observe that the archaeological materials (ARCH), which are burned materials and ceramics, are well separated from the “soil cluster” on the right. Although soil spectra of different horizons were not well separated, we can see that archaeological materials (ARCH) and archaeological soil (AS) are gathered together.

To investigate under which conditions we can most clearly distinguish archaeological remains and natural soil, PCA was performed over different wavelengths (Fig. 3). The wavelength ranges were chosen according to various factors which affected the measurement. The separation of wavelength windows was performed at the spectrometer boundaries (ASD spectrometer is comprised of three different spectrometers) and close to strong water absorption bands. Since the archaeological horizon was visually distinguishable, we expected to see a clear separation between archaeological features and natural soil in the visible spectral range (400–700 nm). However, despite the small difference between archaeological materials and natural soil, PC1 did not make a clear separation. Also, the different horizons (which were seen visually) are not very well separated in the visible range. This is probably related to the normalisation procedure applied to the spectra. By looking at the PCA at different wavelengths, we can see that the archaeological materials (ARCH) are separated for wavelength ranges below 1000 nm. Beyond this wavelength, it becomes difficult to find the difference between archaeological material and natural soil.

CONCLUSION

The paper shows preliminary result of PCA application to the reflection spectra of archaeological remains. The result indicates that archaeological materials are well separated from the natural soil through PCA. The PCA result can probably be improved by using

a larger dataset (spectra) over a wide range of archaeological sites. This will improve the statistical results and perhaps be used to separate different horizons as well. Currently, more research is going on with the PCA application to archaeological sites to distinguish archaeological remains through spectroscopy.

ACKNOWLEDGEMENTS

We extend our thanks to the University of Groningen and the National Museum of Hungary for their help and support during the fieldwork.

REFERENCES

- Clark, R.N. and Roush, T.L. 1984. Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications. *Journal of Geophysical Research* 89: 6329–6340.
- Galvão, L.S., Pizarro, M. A. and Epiphanyo, J.C.N. 2001. Variations in reflectance of tropical soils: Spectral-chemical composition relationships from AVIRIS data. *Remote Sensing of Environment* 75: 245–255.
- Smith, M.O., Johnson, P.E. and Adams, J.B. 1985. Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. *Journal of Geophysical Research* 90: 797–804.
- Viscarra Rossel, R.A., Cattle, S.R., Ortega, A. and Fouad Y. 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* 150: 253–266.