

INTERNATIONAL SOCIETY FOR SOIL MECHANICS AND GEOTECHNICAL ENGINEERING



This paper was downloaded from the Online Library of the International Society for Soil Mechanics and Geotechnical Engineering (ISSMGE). The library is available here:

<https://www.issmge.org/publications/online-library>

This is an open-access database that archives thousands of papers published under the Auspices of the ISSMGE and maintained by the Innovation and Development Committee of ISSMGE.

The paper was published in the proceedings of the 12th Australia New Zealand Conference on Geomechanics and was edited by Graham Ramsey. The conference was held in Wellington, New Zealand, 22-25 February 2015.

Development and use of low cost spectroscopy for soil identification

Truong Le¹, Nikki Akbari², Jason Surjadinata³ and David Airey⁴.

¹School of Civil Engineering, University of Sydney, NSW, 2006; email: truong.le@sydney.edu.au

²Partridge, 1 Chandos St, St Leonards, NSW 2065, PH (612) 94609000; email: nikki.akbari@partridge.com.au

³Douglas Partners Pty Ltd, 96 Hermitage Rd, West Ryde, NSW 2114; PH (612) 88780659; email: jason.surjadinata@douglaspartners.com.au

⁴School of Civil Engineering, University of Sydney, NSW, 2006 PH (612) 93513342; email: david.airey@sydney.edu.au

ABSTRACT

Over the past two decades there has been an increase in research into the use of spectroscopy for rapid analysis of basic soil composition. The interest has been driven principally by the fast and simple sample preparation, rapid testing procedures, overall efficiency and the ability to use the process both in the laboratory and in situ. Previous soil analysis using reflectance spectroscopy has used expensive commercial grade spectrometers limiting its appeal as a soil analysis technique. This paper describes the design, construction and evaluation of an inexpensive, visible to near infrared diffused reflectance spectrometer. Experimentation focused on developing a technique for capturing reliable diffused reflectance spectra of artificially constructed reconstituted soils. The set of results were used to create a calibration spectral library that enabled a predictive program to be developed. It was found that the predictive model was able to estimate selected soil properties to 4% of the value obtained by standard laboratory testing. It is demonstrated that low cost spectroscopy is capable of determining some soil properties, however more research is required to demonstrate the reliability of the low cost device.

Keywords: rapid soil analysis, spectroscopy, soil classification

1 INTRODUCTION

1.1 Overview

As soil is a complex matrix composed of minerals, organic matter, water and air with high degrees of variation, it is difficult and laborious to classify and characterise the large volumes impacted by proposed construction activities. The process of classification can involve a wide range of procedures to obtain information on the physical, chemical and mineralogical properties of the soil, and the task of understanding the soil properties can grow to involve intensive sampling strategies, which can include complex, time-consuming and expensive laboratory analysis.

With the demand for increased efficiency throughout project lifetimes, it is desirable to be able to rapidly analyse soil properties. The combination of improved technology and a growing body of research has led to soil spectroscopy being investigated as an alternative to traditional laboratory classification tests. Research has been mainly focused on the use of visible-near infrared (vis-NIR) diffused reflectance spectroscopy for the simultaneous classification of various soil properties (Viscarra Rossel et al., 2006). Spectroscopy has the advantage of being non-destructive, simple, quick and requiring little to no sample preparation.

Uses of spectroscopy have focused on basic soil composition, particularly soil organic matter, soil texture, clay mineralogy and moisture content (Stenberg et al., 2010). There has also been interest in the use of spectroscopy for the identification of specific geotechnical properties such as swelling index, Atterberg limits and shrinkage limit (Yitagesu, 2006). The results from the collective research have indicated that diffused reflectance spectroscopy, using the visible (VIS), near infrared (NIR), and mid infrared (MIR) spectral ranges, may be used to efficiently acquire soil information (Viscarra Rossel et al., 2006).

Technological advances and the availability of both hardware and software components, has led to the development of open-source software that is intended to enable spectroscopy to be used to support community investigations into environmental issues. These developments have been adopted to

develop and investigate the suitability of low cost spectroscopy for geotechnical classification purposes. The construction and operation of the low cost spectrometer is described, and results of the diffused reflectance spectra are provided. The calibration procedure required to obtain useful soil parameters is explained, and the performance of a simple predictive model is evaluated.

1.2 Fundamentals of diffused reflectance spectroscopy in soil analysis

Fundamentally, spectroscopy is a study of light, as a function of wavelength, that is emitted, absorbed, reflected or scattered from a material either solid, liquid or gas (Clark, 1999). As the energy of each photon is directly related to the frequency and inversely related to the wavelength, the absorbance, and in effect, reflectance, spectra produce a characteristic shape that can be used for an analytical purpose (Stenberg et al., 2010). This is a result of the mineral constituents and chemical matrix causing different individual molecular bonds to vibrate, absorbing the source radiation to various degrees (Stenberg et al., 2010). For example, the presence of smectites produces absorption features at a wavelength of 1400 nm due to the OH stretching and at 1900 nm due to a combination of H-O-H bending and OH stretching (Yitagesu, 2006). Additionally, in the 2200 nm to 2300 nm range the presence of metal ions such as aluminium, magnesium and iron can be detected (Yitagesu, 2006). As absorption features of the soil spectrum are influenced by both physical and chemical properties, these features can lead to quantitative compositional information of mineral constituents as well as estimates of soil properties (e.g. Viscarra Rossel et al. 2006)

The spectral range of the source radiation and sensitivity of the sensor critically affect the predictive ability of the results (Viscarra Rossel et al., 2006). Comparisons of VIS, NIR, MIR and VIS-NIR-MIR spectrometers for different soil properties have indicated that VIS spectra are generally less accurate and are unable to discriminate specific minerals. The ability to capture spectra in the infra-red range allows for better characterisation of various soil chemical and physical properties. Nevertheless, Viscarra Rossel et al. (2006) concluded that spectrometers in the VIS, NIR and MIR range can all be used for the simultaneous assessment of various soil properties. The choice of which spectral range to use depends on: the necessary degree of accuracy; the availability of the equipment, especially considering that NIR and MIR spectrometers are considerably more expensive; and the amount of sample preparation required.

1.3 Spectroscopic multivariate calibration

The complex nature of soil composition and variability in soil properties leads to the diffused reflectance spectra of soils being largely nonspecific (Chang et al., 2001). Conventional means of calibration with simple univariate methods such as simple linear regression have proven to be inadequate as a result of the overlap between weak overtones and the combination of fundamental vibrational bands in the spectra (Stenberg et al., 2010). Predictions using standard regression models are made more complex as a result of scattering effects caused by soil structure and nonspecific constituents (i.e. no distinguishable peaks) such as quartz (Stenberg et al., 2010).

To discriminate between different properties and to provide quantitative results, spectral data is often analysed using multivariate analysis. The most common method of calibration for soil identification has been based on forms of linear regression, namely multiple linear regression (MLR), principal components regression (PCR) and partial least squares regression (PLSR) (Chang et al., 2001) with more recent studies adopting data mining techniques such as neural networking (NN), multivariate adaptive regression splines and boosted regression trees (Stenberg et al., 2010). A summary of soil identification studies listing the soil attributes tested has been provided by Viscarra Rossel et al. (2006). Stenberg et al. (2010) noted that there are inconsistencies in the predictability of various soil attributes between studies, which they attributed to the lack of awareness among analysts of full spectrum data compression techniques such as PCR and PLSR. Chang et al. (2001) have also noted that no one proposed calibration technique has achieved universal acceptance due primarily to the variability in application and the degrees of success in their ability to predict soil properties.

1.4 Spectroscopy for geotechnical applications

Geotechnical classification tests such as Atterberg limits and grain size distribution are cheap and relatively easy to perform in addition to providing an indication of the mechanical behaviour and soil constituents, both of which are important in the development and understanding of the soil profile and

behaviour. However, they provide limited information about the fundamental mineralogy and soil structure. Additional classification tests, such as X-ray diffraction and differential thermal analysis, are available to provide this information but are rarely conducted due to their expense, labour and specificity to sample location.

The features of the spectral response of soils, discussed above, suggest that spectroscopy should be applicable for geotechnical application (e.g. Yitagesu 2006). For example, successful qualitative and quantitative predictions of clay mineralogy and clay content have been reported (Chang et al. 2001, Stenberg et al. 2010), although these have required infra-red detectors which are beyond the capability of the device investigated in this paper. Moisture content is another attribute that has been successfully estimated with vis-NIR spectroscopy (e.g. Sørensen and Dalsgaard 2005, Chang et al. 2001, Stenberg et al. 2010), but the accuracy is affected by variations in soil colour, texture and soil origin. Yitagesu (2006) has also demonstrated good predictive ability for soil cation, liquid limit, plastic limit and plastic index using a data driven partial least squares regression model.

2 MATERIALS AND METHODS

2.1 Experimental setup

The spectrometer is shown in Figure 1. The detecting component was made by using the foldable mini-spectrometer kit supplied by Public Laboratory (2013) which has a cost of approximately \$10. The kit includes a laser-cut thick, light-blocking black card stock, a DVD-R (to be used as the diffraction grating), and instructions for assembly. The spectrometer was assembled using a Logitech C300 webcam with the infrared filter removed as the detector and a segment of a DVD-R as the diffraction grating for splitting the dispersed light. The infrared filter was removed from the webcam to extend the spectral range of the camera into the near-infrared region.

The complete spectrometer setup consisted of timber framing constructed from material costing approximately \$10, four 15 watt 'Cool White' 4000K LED lights costing approximately \$200 AUD, and the detecting component, as described previously.

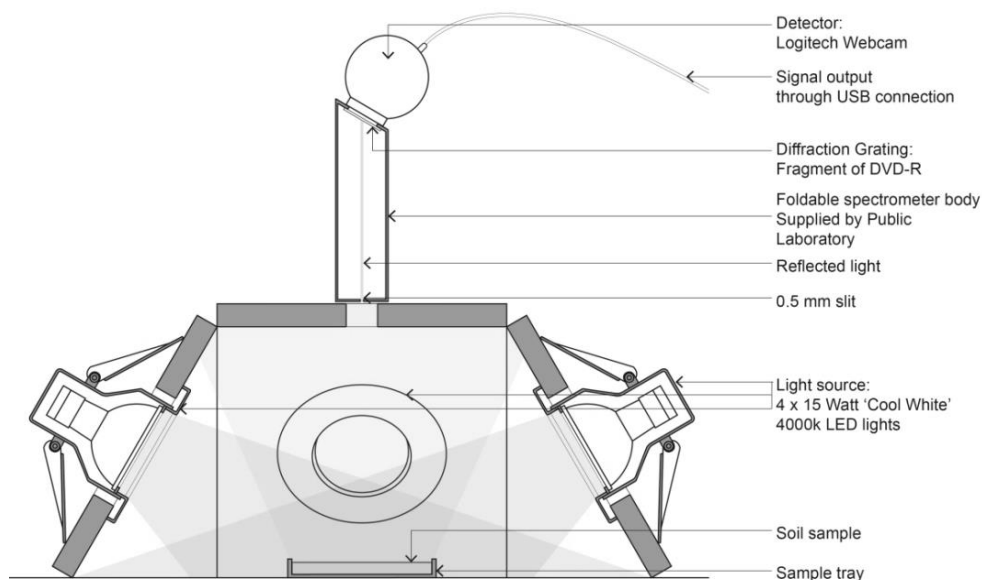


Figure 1. Experimental reflectance spectrometer setup

The camera and light source specifications control the spectral range that can be investigated. The camera CCD with the infrared filter removed is capable of detecting wavelengths between 400 nm to 1000 nm, with peak sensitivity between 550 nm and 800 nm (Spring et al., 2000). The radiation source for the spectrometer was chosen to be an LED light source due to its wider spectral emission range, between 425 nm and 675 nm. The spectrometer is therefore expected to have a reliable spectral range where the detectable range overlaps with the emission range, that is, between 425 nm and 675 nm.

2.2 Soil sampling and laboratory analyses

To determine the feasibility of the technique, unknown soil variation was kept to a minimum. All soils used in testing were therefore artificially constructed with controlled quantities and reconstituted under laboratory conditions. Principal soil constituents were selected to have high values of purity.

To achieve a uniform and consistent soil composition for testing, all soils were initially ground in a mortar and pestle and then sieved to obtain the <150 µm fraction. The materials were then dried in an oven at 99°C for two days and stored in sealed containers in a cool dry room until required. Dry specimens were thoroughly mixed and placed loosely into small containers and the surface was scraped level. The wet specimens were mixed with distilled water and allowed to cure in sealed plastic bags for several days, before placing into the test containers. The soil compositions of the samples are summarised in Table 1.

Table 1: Soil properties tested

Soil property	Principal soil constituent	Percentage
Moisture content	Goethite	0, 10, 20, 25, 30, 35, 40, 45
	Bentonite	0, 10, 20, 30, 40, 50, 60, 70, 80
	Kaolinite	0, 10, 20, 30, 40
	Ball clay	0, 10, 20, 30, 35, 40
Sand content	Kaolinite	0, 25, 50, 75, 85, 90, 95, 100
Carbon content	Kaolinite	0, 20, 40, 60, 80, 100
Iron oxide content	Kaolinite	0, 25, 50, 75, 100
	Bentonite	0, 25, 50, 75, 90, 100
	Ball clay	0, 25, 50, 75, 100

2.3 Analytical technique

Soil samples were placed in the apparatus as shown in Figure 1 and the soil spectra, plots of normalised light intensity against wavelength, recorded seven times. Data stacking was used to remove random variations. It was observed that the response was reasonably repeatable as after averaging 7 spectra, minimal change was observed. This process took less than 1 minute and this ensured no appreciable drying of the moist specimens.

The spectra were captured using the free Spectral Workbench web-based software (Public Laboratory, 2013). This allowed the user to observe the spectrum on the screen and to download the data for further processing. The output was provided at a wavelength resolution of 1.1 nm. The software also requires a simple calibration process to establish the wavelength scale using two of the peaks in the intensity, wavelength response from a fluorescent light. It was observed that this process was not very precise and an additional calibration was performed using two lasers with known wavelength.

To obtain the reflectance, the normalised light intensity, of the soil the averaged soil spectra were normalised against the absolute reflectance of the radiation source. To minimise the experiment cost, the absolute reflectance of the spectrometer has been taken to be the reflectance recorded from an 18% grey card. High quality commercial spectrometers use a spectralon, a material capable of producing high reflective indexes across a wide range of wavelengths, to determine the absolute reflectance of the spectrometer. However, the cost of this is several times the cost of the spectrometer and is not justified if only the relative magnitudes of the intensities are being compared as described below.

3 RESULTS AND DISCUSSION

3.1 General Observations

Interpretation of the data requires the setting up of a spectral library, sets of reflectance spectra that are grouped together to form series related to a particular parameter, an example is shown in Figure 2a. Each series of spectra, for all the soil properties examined, had unique features and shapes (Figures 2 b – d). When comparing all the spectral series there was no sign of any consistent peak suggesting there were no experimental artefacts or common environmental interferences in the

wavelength specific reflectance values. It is therefore considered that the complete range of any reflectance spectra is relevant to the soil property tested. To better demonstrate, graphically at least, the change in reflectance for each property, the spectra have been plotted as baseline removed reflectance spectra.

$$R_{bv} = R_v - R_0 \quad (1)$$

Where R_{bv} is the value of the baseline removed spectra, R_v is the value of the reflectance spectra for a given value of a soil property and R_0 is the value of the reflectance spectra for the base case with zero content of the property being varied. For example, Figure 2b shows a baseline removed spectral series for increasing carbon content using the data from the spectra in Figure 2a, with the baseline representing 100% kaolinite.

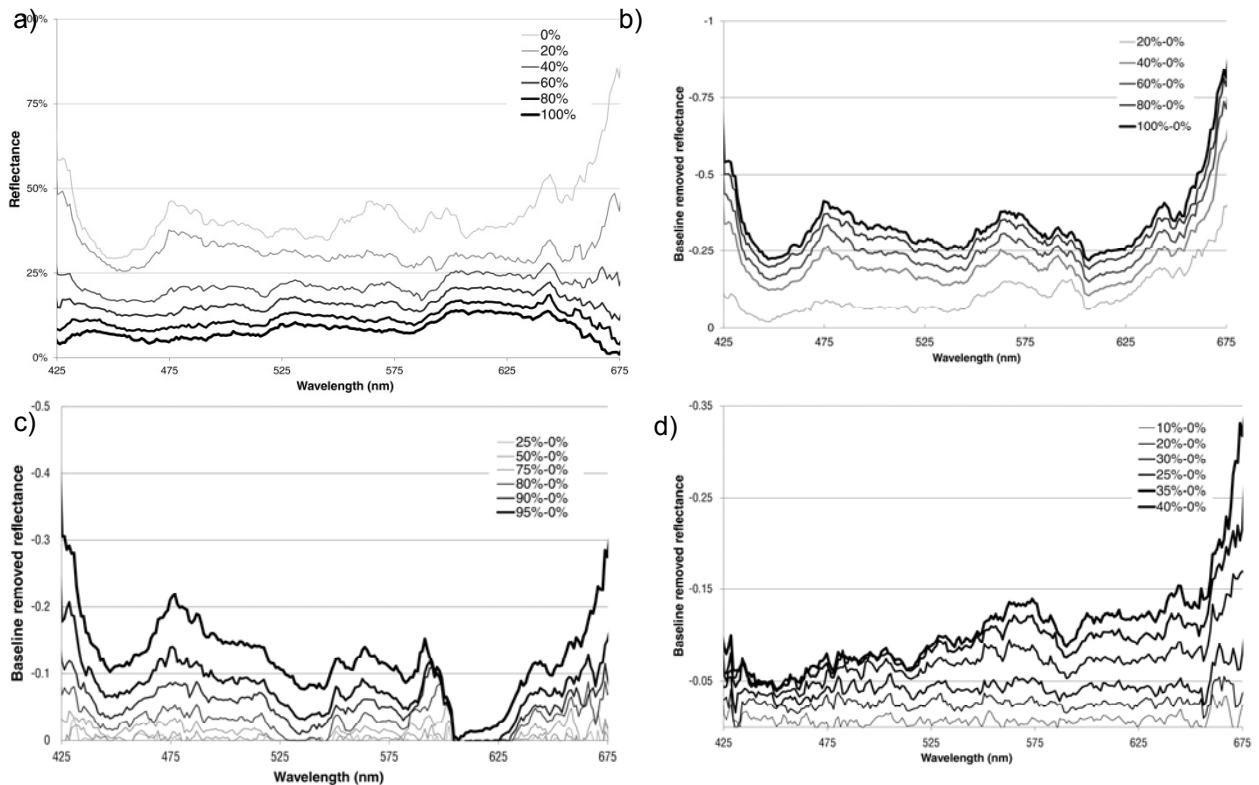


Figure 2. a) Typical series of reflectance curves (increasing carbon content in kaolinite), Series of baseline removed curves for b) increasing carbon content in kaolinite c) increasing sand content in kaolinite d) increasing moisture content in goethite

To develop a calibration model capable of determining, from an unknown soil specimen, a set of soil properties, a large number of wavelengths needs to be captured and used. This is principally due to differences in the correlations between spectral variations at specific wavelengths and the measured soil properties. For some properties discrimination can best occur at specific wavelengths and these can lie within the NIR or MIR range as discussed previously. Further research will be necessary to uniquely determine soil attributes using the method proposed in this paper. However, there are some unique peaks evident in the spectra for the investigated soil attributes suggesting that discrimination should be possible.

An automated analysis process was developed as manually selecting appropriate wavelengths for analysis would be laborious and ultimately defeat the objective of a low cost approach. The final procedure adopted was to record the spectral data for the soil, perform a baseline removal, and to superimpose the spectra onto the series selected. This process was automated by a simple Visual Basic code in Microsoft Excel. In analysing an unknown soil, the baseline removed reflectance values at a specified number of equi-spaced wavelengths, within the range from 425 nm to 675 nm, were taken from the unknown soil spectrum, compared with the calibration data and the closest saved values larger and smaller than the unknown at each wavelength collected. An example of this result is shown in Figure 3a. The process of selecting wavelengths at equal intervals did not bias the results to

particular peaks but nevertheless it was found that with only 5 equi-spaced wavelengths in the range the model was capable of reasonably predicting single soil attributes. This estimate will increase in reliability as the number of wavelengths increases, but this did not affect the actual range of the estimated values significantly. For each wavelength an estimate of the unknown property was obtained by linearly interpolating between the greater, lower and unknown value. The model gives estimates of the soil property from each of the wavelengths, provided the unknown values are inside the values in the calibration data, as shown in Figure 3b. Values that lie outside the calibration data range are often associated with ranges in the spectral response that have weak or inconsistent variation with the property in question, as seen in Figure 2c between 580 nm and 625 nm.

At some wavelengths the predictions were significantly different from the set of estimated values. These selected outliers were thought to be a result of spectral variation, as they commonly occurred near the ends of the spectral range or near regions of rapid change in reflectance, and were thus determined to be unreliable results. To selectively remove outlier results from the set of estimated values; a statistical process was established to provide a best estimate. To avoid these values skewing the averages the outlier analysis method, a modified z-score model as proposed by Iglewicz and Hoaglin (1993) was adopted. A detailed discussion into both selection and justification of the rejection parameters and the process of the outlier analysis can be found in Le and Akbari (2013). In summary, the process was to analyse the set of estimated values and remove values that lie outside the general average of the set, for example the points shown as empty squares in Figure 3b. The final output of the spectral analysis returned a single numerical estimation of the soil property which was the average of the final set of estimated values, shown as the dotted line in Figure 3b.

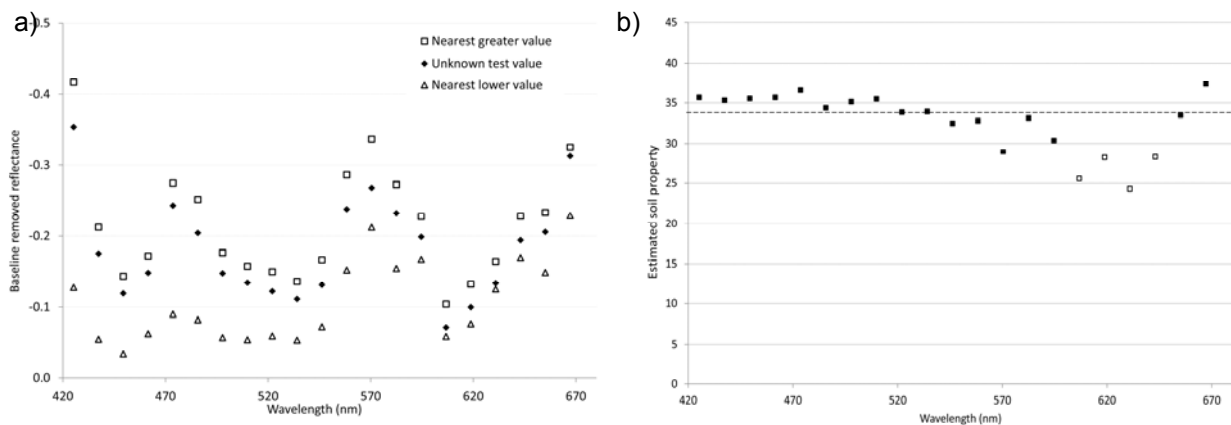


Figure 3. a) Selected wavelengths/spectral values for interpolation b) Estimated soil properties for each selected wavelength and final estimation

3.2 Qualitative description of spectral data

The reflectance spectra of all soil mixtures analysed displayed unique variations across the measured range of the spectrometer, as shown in Figure 2. As discussed by Viscarra Rossel et al. (2006), the spectra can be characterised by the shape and values of the peaks in the data. Generally, the spectral reflectance decreased as soil colour darkened; a trend that was observed in all varying soil properties. The experimental data indicated that the most reliable and significant results were from soil properties that caused the greatest change in colour as the soil property varied, a consequence of this spectrometer operating only in the visible spectral range. Thus, for this experiment, spectroscopic analysis was primarily inferring the co-variance between soil colour and composition.

Moisture content variations demonstrated a consistent trend for all soils tested, with the reflectance generally decreasing as the moisture content increased. This trend was seen to be consistent up to the plastic limit after which the soil became more reflective. For the dry mixtures of soil, as the soil mixture varied from one pure constituent to the other, the reflectance spectra shifted in proportion to the relative percentages, for example Figure 2c.

It was observed that not all spectra responded uniformly with variation. Sand content in clay was seen to produce the weakest change in spectral response as soil composition varied. The sand content, as seen in Figure 2c, was seen to only vary sufficiently to be detectable, in the 75-100% range. Between

sand contents of 0-75%, due to the similarities in colour, there was no significant change in the reflectance spectra. This is a consequence of the fine clay particles filling in the void space and obscuring the sand particles as well as the scattering effect associated with quartz particles.

3.3 Calibration model

To demonstrate the process and possibilities of spectroscopy, a calibration model was developed to automate the process of estimating soil attributes. As previously noted, the premise of the model is to perform very simple, linear interpolations between the nearest greater and lesser value to arrive at an estimated value of the soil attribute. This set of values was then passed through an outlier filtering algorithm to narrow down the estimated value to a more consistent value.

The calibration model developed was then used to predict the constituents of samples with soil properties other than those described in Table 1. For all the validation samples tested, the absolute difference in the estimated and actual values differed by at most 4% and the standard deviation of the set of estimated values was less than 5%. These values demonstrate the potential for the technique to be used as an analytical method in future geotechnical characterisation studies. Nevertheless, much more experience and investigation is required to establish the reliability of spectroscopy as an analytical tool. Previous studies have shown that the lack of universally accepted standards for the capture and analysis of spectral data often results in inconsistencies in the prediction capabilities of spectrometers, estimates of soil attributes and the overall success of the technique. It was also observed that to obtain repeatable spectra considerable care and strict control measures were required.

3.4 Comparison with high quality spectrometer

Comparisons between the reflectance spectra for identical soil specimens obtained from the low cost spectrometer and a high quality commercially available ASD Fieldspec spectrometer are shown in Figure 4. The results show significant differences in the reflectance values across the range of wavelengths where comparison is possible.

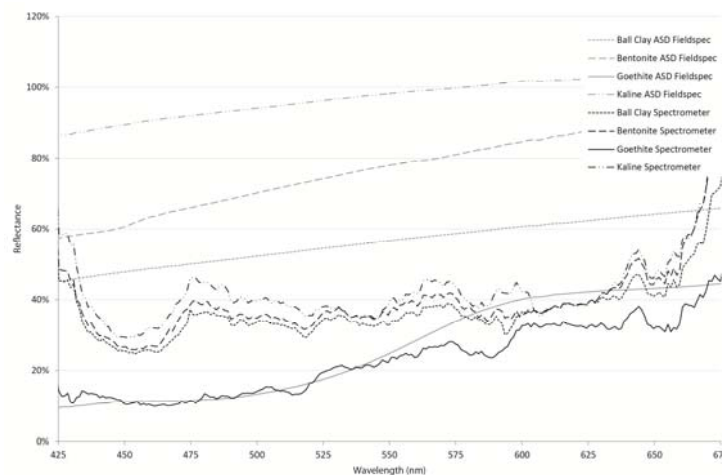


Figure 4. Comparison of low cost spectrometer and ASD Fieldspec reflectance spectra

These differences are thought to be the result of the different techniques used in capturing reflectance spectra. The ASD Fieldspec device allows the receiver and radiation source to be directly in contact with the sample, whereas with the low cost spectrometer, the reflected light has to travel some distance before passing into the detector. Stenberg et al. (2010) have noted how the apparent absorbance of a sample increases as path length increases, suggesting that the apparently lower reflectances measured in the low cost device are a consequence of the apparatus design.

The differences in the reflectance indicate that there is unlikely to be a universal calibration, but provided suitable calibration procedures are adopted both spectrometers can provide qualitative and quantitative data on the soil constituents. Given the 2 orders of magnitude difference in cost it is considered that further experimentation with the low cost device to investigate reliability and repeatability are warranted.

3.5 The future of inexpensive spectroscopy

One of the motivations for the current study was to explore the possibility of using spectroscopy during site investigation to provide an immediate indication of mineralogy, water content and other soil attributes. The technique has the advantages of speed, simplicity and relatively low cost when compared with traditional analytical methods. However, further study is required to develop procedures that are reliable and repeatable, and to determine suitable calibration techniques that can lead to reliable qualitative and quantitative estimates of soil properties. At present variability in results may be attributed to the selection of pre-treatment techniques, sample library and regression model.

The ability of spectroscopy to pick changes in water content, iron content and carbon content suggest direct application in certain specific applications, and with suitable calibration data it may be possible to provide rapid soil classification over sites or small regions. Alternatively spectral data can be combined from varied soils but as noted by Stenberg et al. (2010) the resulting calibration models may be limited in their use as they lack the precision and accuracy required for small projects. If spectroscopy is to develop, data from a wide range of soils is required and initiatives such as Global Soil Spectral Library (Rossel, 2009) and Spectral Workbench which allow users to share and combine spectra are of vital importance to encourage researchers, reduce costs and expand the community.

4 CONCLUSION

Our research has demonstrated the potential of a low cost spectrometer as a geotechnical analysis tool by estimating soil attributes with generally high predictive coefficients. With existing research generally agreeing on the relevance and robustness of diffused vis-NIR spectroscopy in estimating soil attributes, whether through direct correlation with spectral signatures or co-variation with water, organics and/or minerals, the focus of the experiment was to demonstrate the feasibility of a low cost spectrometer by producing similar predictive capabilities.

Results have indicated that under optimal conditions, the low cost spectrometer is able to accurately estimate various soil properties with a simple calibration model. This leads to the conclusion that, as there is a great potential for diffused reflectance spectroscopy in geotechnical applications, further research should be conducted to further encourage the use of spectroscopy as an analytical technique.

REFERENCES

- Chang, C.-W., Laird, D. A., Mausbach, M. J. & Hurburgh, C. R. 2001. Near-infrared reflectance spectroscopy—principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65, 480-490.
- Clark, R. N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. *Manual of remote sensing*, 3, 3-58.
- Iglewicz, B. & Hoaglin, D. C. 1993. *How to detect and handle outliers*, ASQC Quality Press Milwaukee (Wisconsin).
- Le, T. & Akbari, N. 2013. *An investigation into estimating soil properties using an inexpensive visible spectrometer*. Civil Engineering, University of Sydney.
- Public Laboratory. 2013. *Spectral Workbench* [Online]. Available: <http://spectralworkbench.org/> [Accessed 11 April 2013].
- Rossel, R. The Soil Spectroscopy Group and the development of a global soil spectral library. EGU General Assembly Conference Abstracts, 2009. 14021.
- Sørensen, L. & Dalsgaard, S. 2005. Determination of clay and other soil properties by near infrared spectroscopy. *Soil Science Society of America Journal*, 69, 159-167.
- Spring, K. R., Fellers, T. J. & Davidson, M. W. 2000. *Introduction to Charge-Coupled Devices (CCDs)* [Online]. Available: <http://www.microscopyu.com/articles/digitalimaging/ccdintro.html> [Accessed 08 May 2013].
- Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M. & Wetterlind, J. 2010. Chapter five-visible and near infrared spectroscopy in soil science. *Advances in agronomy*, 107, 163-215.
- Viscarra Rossel, R., Walvoort, D., McBratney, A., Janik, L. J. & Skjemstad, J. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131, 59-75.
- Yitagesu, F. A. 2006. Spectroscopy to derive engineering parameters of expansive soils. Msc, ITC.