DIFFUSE REFLECTANCE SPECTROMETRY AS A PROXIMAL SENSING TOOL FOR PRECISION AGRICULTURE

D. J. J. WALVOORT

Wageningen University & Research Centre, Wageningen, The Netherlands
E-mail: Dennis.Walvoort@oio.beng.wag-ur.nl

A. B. McBRATNEY

Australian Centre for Precision Agriculture, The University of Sydney, Sydney, Australia
E-mail: Alex.McBratney@acss.usyd.edu.au

ABSTRACT

The aim of this paper is to evaluate diffuse reflectance spectrometry as a data acquisition method for precision agriculture in general and variable rate nitrogen fertiliser application in particular. Partial least squares regression models were fitted to predict clay, carbon, and nitrogen contents of soil. Diffuse reflectance spectra (Vis-NIR region) were used as inputs for these models. Results indicate that clay and carbon contents can be predicted reasonably well. However, more work has still to be done to make the models more robust.

INTRODUCTION

A condition for the successful implementation of precision agriculture is the availability of high quality soil and crop data sampled at high resolution and minimal cost. Conventional sampling and laboratory methods are generally too expensive and time consuming for this purpose. Therefore, much research is aimed at developing alternative data acquisition methods that are fast, cheap and accurate. Proximal sensing, *i.e.*, the application of sensors in close proximity to the property of interest, appears to be promising in this respect (Viscarra Rossel and McBratney, 1998; Walvoort *et al.*, 2001; Viscarra Rossel *et al.*, 2001).

Potentially, diffuse reflectance spectrometry may provide a relatively fast and cheap proximal sensing method without the need for elaborate sample preparation. Dalal and Henry (1986) applied near infrared (NIR) diffuse reflectance spectrometry to predict soil moisture content, total nitrogen content and organic carbon content of soil. They concluded that NIR spectrometry could be successfully applied in soils with a narrow colour range and with moderate organic matter contents. Viscarra Rossel and McBratney (1998) applied NIR spectrometry to predict the moisture content, the clay content, and the organic matter content of soil. They found significant response-surface models, except for organic matter content. Masserschmidt *et al.* (1999) predicted soil organic matter content by analysing mid infrared diffuse reflectance soil spectra. Ehsani *et al.* (1999) successfully applied NIR diffuse reflectance spectrometry to predict nitrate concentrations in soil.

For soil nutrient management, the farmer is concerned about where and when to apply a specific fertiliser and in what amounts. The optimal application rate should maximise crop yield and crop quality subject to environmental and economic constraints. In addition to

information collected by means of proximal sensing, we also have to rely on mathematical models to compute this optimal rate. Processes like denitrification and nitrogen leaching are hard to quantify without models. Van Alphen (2001), for example, used a mechanistic model to optimise nitrogen fertiliser application rates for distinct management units. The approach he pursued not only predicts optimal fertiliser application rates but also forecasts the optimal application dates. For the farm he studied, four application dates were needed.

As an extension of Van Alphen's (2001) approach, we may also consider the variation within management units. Proximal sensing may provide a means for this. Optimal fertiliser application rates not only depend on the amount of mineral N in the soil at the start of the growing season, but also on processes like mineralisation, denitrification, and leaching. To quantify these processes by means of mathematical models, properties like the amount of N (both mineral and organic), the soil organic matter content, and the clay content have to be known. In this paper, we will apply diffuse reflectance spectrometry as a proximal sensing method to predict these soil properties.

MATERIALS AND METHODS

Soil sampling

Soil sampling was performed at West Creek, an 80 ha field near Moree, New South Wales, Australia. An aerial colour photograph of the bare soil surface showed distinct patterns of spatial variation. The soils encountered in this field are classified as red chromosols and grey and brown vertosols (Isbell, 1996). Soil samples were taken along a NW-SE transect at 59 equidistant sampling points.

Laboratory analysis

After drying (air dry), grinding, and sieving (<2 mm), all soil samples were analysed for clay, carbon (C) and nitrogen (N) content. Clay content was measured by means of a hydrometer. C and N were measured using a LECO CHN-1000 analyser. A Varian Cary 500 scan spectrophotometer equipped with a Labsphere DRA-CA-50D diffuse reflectance accessory was used to collect diffuse reflectance spectra of the soil. The spectral range of interest was from 250 nm to 2450 nm. The spectral resolution was 1.1 nm for wavelengths up to 800 nm and 3 nm for wavelengths greater than 800 nm.

Statistical methods

Given a soil sample, we want to use its diffuse reflectance spectrum to predict soil property y. For that purpose we need a calibration model. Suppose we have taken n soil samples i=1..n. Each soil sample i is analysed for property y by means of a conventional (expensive, and/or time consuming) laboratory method. In addition, diffuse reflectance spectra $\mathbf{x}_i = [x_1, x_2, ..., x_k]'$ consisting of diffuse reflectances x are collected at k wavelengths. Let $\mathbf{y} = [y_1, y_2, ..., y_n]'$ be a vector of n measurements of soil property y, and let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]'$ be a matrix of corresponding diffuse reflectance spectra. A calibration model can now be constructed by means of partial least squares regression (PLS). PLS decomposes $\mathbf{X}$ and $\mathbf{y}$ into factor scores ($\mathbf{T}$) and factor loadings ($\mathbf{P}$ and $\mathbf{q}$) according to:

$$\mathbf{X} = \mathbf{TP'}+\mathbf{E}$$
$$\mathbf{y} = \mathbf{Tq}+\mathbf{f}$$

X and **y** should be zero-centred prior to decomposition. Decomposition is performed in such a way that the first p ($< \min(n,k)$) bilinear factors explain most of the variation in **X** and **y**. The remaining bilinear factors resemble noise and can thus be ignored, hence the addition of residuals **E** and **f**. The resulting matrices and vectors generally have a much lower dimension than **X** and **y**. The estimator $\hat{y}$ given a new spectrum **x** is then a linear combination of the factors scores and factor loadings of **x**. For more details on PLS see *e.g.*, Martens and Næs (1989) and Denham (1995). For an application of PLS with respect to soil nitrate, see Ehsani *et al.* (1999). A paper on predicting soil organic matter content by PLS is written by Masserschmidt *et al.* (1999).

Leave-one-out cross-validation (Martens and Næs, 1989) was applied to estimate the number of bilinear factors to retain in the model. This method consists of calibrating a model n times, each time leaving out one of the n observations. Each calibrated model is used to predict the removed observation and the error is computed as the difference between its observed and predicted value. A statistic like the Root Mean Squared Error of prediction (RMSE) is then computed to summarise the results. This procedure is repeated for models with different numbers of bilinear factors. The model with the lowest RMSE is usually selected.

RESULTS AND DISCUSSION

Fig. 1 shows the cumulative distributions of clay, C, and N content. It reveals only a slight variation in C, N, and clay content, despite a distinct variation in pedology. It should be noted that calibration models conditioned on reference data with low variability like the West Creek samples only have local validity.

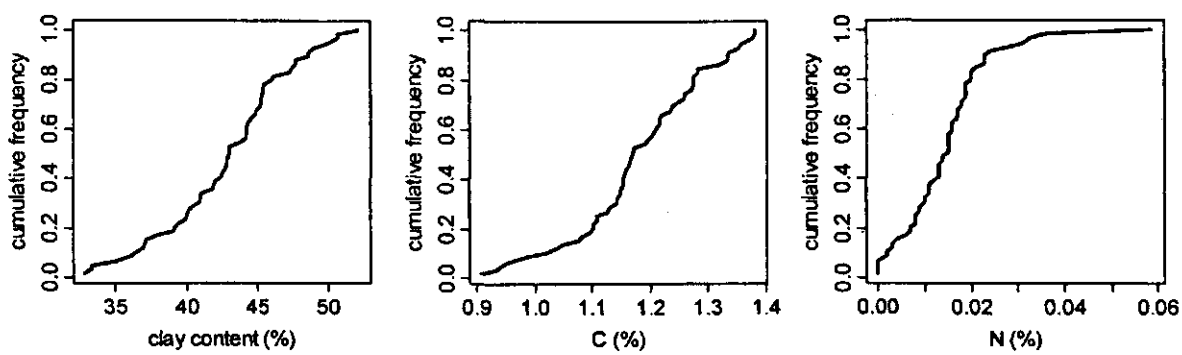

FIGURE 1. Graphs of cumulative distributions of clay content (%), carbon content (%), and nitrogen content (%).

Leave-one-out cross-validation was applied to find the number of bilinear factors to retain in the model. The results are given in Fig. 2. We selected 5 bilinear factors for clay content, 6 for C, and 11 for N. However, it can be argued that a smaller number of bilinear factors would yield more robust and parsimonious models at the expense of only a small increase in RMSE. More work has to be done on this issue.

Potential outlier wavelengths were identified by means of procedures described in Martens & Næs (1989, chapter 5). Outlier wavelengths were only removed when they were caused by artefacts of the spectrophotometer, *i.e.*, changing sources, filters or gratings.

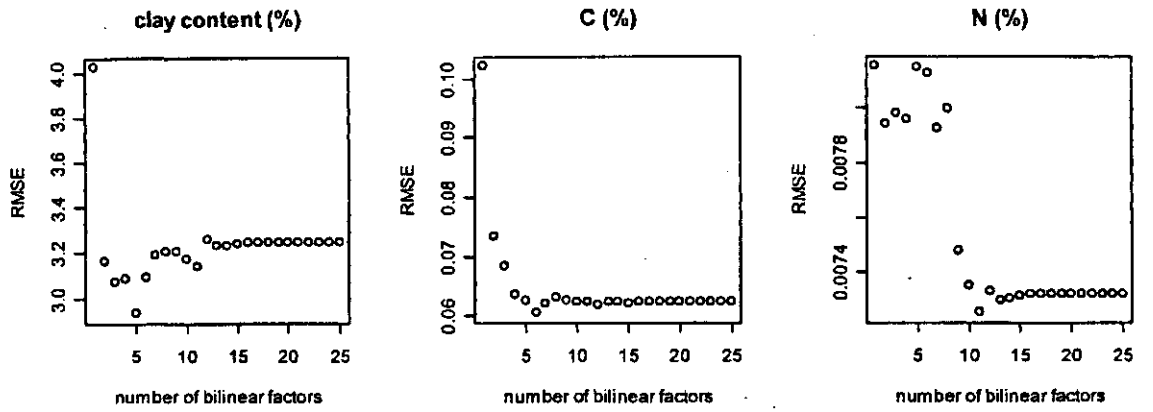

FIGURE 2. RMSE as a function of the number of bilinear factors retained in the model. Predictions are obtained by means of leave-one-out cross-validation.

The resulting models were validated by predicting clay, N, and C contents of soil samples by using the diffuse reflectance spectra as inputs. The results are given in Fig. 3. The RMSE for clay content is 2.3%, the RMSE for C is 0.04% and the RMSE for N is 0.0008%. However, these predictions are too optimistic, because the observations are used for both calibration and validation.

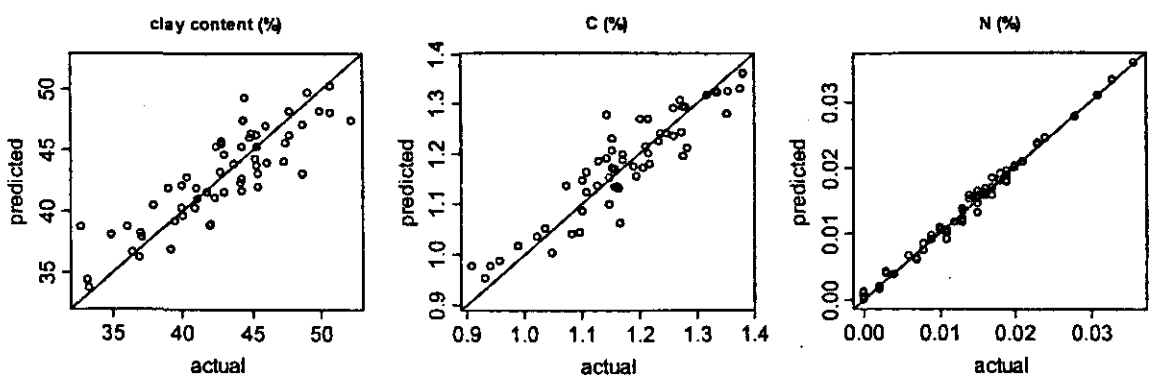

FIGURE 3. Predicted versus observed clay, C, and N contents. Samples are used for calibration and validation.

Independent validation can be achieved by means of leave-one-out cross-validation. In this way, validation is performed with samples not used for calibration. The results are given in Fig. 4. As can be expected, the RMSEs are greater, *i.e.*, 2.9% for clay content, 0.06% for C and even 0.007%. for N. Although clay and C content only show a small increase in RMSE, the increase for N is quite significant. These results probably suggest that a larger calibration set is needed to obtain more robust calibration models.

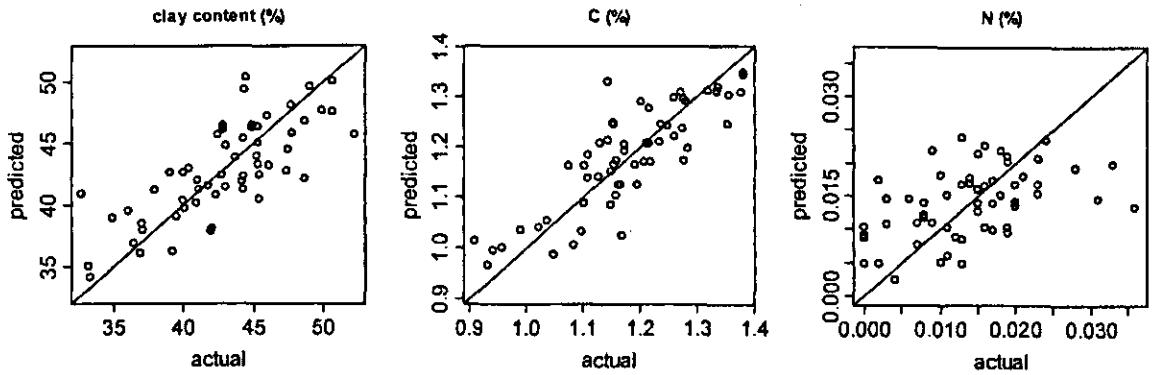

FIGURE 4. Predicted versus observed clay, C, and N contents. Predictions are obtained by means of leave-one-out cross-validation.

CONCLUSIONS

Calibration models were derived for predicting clay, C, and N contents in soil. The models were calibrated by means of partial least squares regression with diffuse reflectance soil spectra as inputs. Although the data set was limited in size, reasonable predictions could be obtained. However, in order to obtain more robust models, a larger data set is probably needed. Furthermore, the issue of finding the optimal number of bilinear factors to attain in the model has still to be addressed.

ACKNOWLEDGEMENTS

We would like to thank Matthew Clutterham for soil sampling and particle size analysis and Raphael Viscarra Rossel and two anonymous reviewers for their useful comments on an earlier draft of this paper. This research was supported by the Technology Foundation STW, *i.e.*, the applied science division of NWO and the technology program of the Ministry of Economic Affairs of the Netherlands.

REFERENCES

Dalal, R.C., Henry, R.J. (1986). Simultaneous Determination of Moisture, Organic Carbon, and Total Nitrogen by Near Infrared Reflectance Spectrometry. *Soil Science Society of America Journal* 50: 120-123.

Denham. M.C. (1995). Implementing Partial Least Squares. *Statistics and Computing* 5: 191-202.

Ehsani, M.R., Upadhyaya, S.K., Slaughter, D. Shafii, S., Pelletier, M. (1999). A NIR Technique for Rapid Determination of Soil Mineral Nitrogen. *Precision Agriculture* 1: 217-234.

Isbell, R.F. (1996). The Australian Soil Classification. CSIRO Publishing, Melbourne, 160 pp.

Masserschmidt, I. Cuelbas, C.J., Poppi, R.J., De Andrade, J.C., De Abreu, C.A., Davanzo, C.U. (1999). Determination of Organic Matter in Soils by FTIR/Diffuse Reflectance and Multivariate Calibration. *Journal of Chemometrics* 13: 265-273.

Martens H., Næs, T. (1989). Multivariate Calibration. John Wiley & Sons, Chichester, 419 pp.

Van Alphen, B.J. (2001). Precision Nitrogen Fertilization: A Case Study for Dutch Arable Farming. *In*: Proceedings of the 5th International Conference on Precision Agriculture and Other Research Management (CD-ROM). Precision Agriculture Center, University of Minnesota, ASA/CSSA/SSSA.

Viscarra Rossel R.A., McBratney, A.B. (1998). Laboratory Evaluation of a Proximal Sensing Technique for Simultaneous Measurement of Clay and Water Content. *Geoderma* 85: 19-39.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O. (2001). Proximal Sensing of Soil pH and Lime Requirement by Mid Infrared Diffuse Reflectance Spectrometry. These proceedings.

Walvoort, D.J.J., Bouma, J., de Gruijter, J.J., Peters, P.D. (2001). Using ISFETs for Proximal Sensing in Precision Agriculture. *In*: Proceedings of the 5th International Conference on Precision Agriculture and Other Research Management (CD-ROM). Precision Agriculture Center, University of Minnesota, ASA/CSSA/SSSA.