

# Extending predictions of soil parameters for a national dataset

A.H.J. Robertson,<sup>a\*</sup> E. Perez-Fernandez,<sup>b</sup> N.J. Baggaley,<sup>a</sup> B.M. McKenzie,<sup>a</sup> I.J. Owen<sup>a</sup> and A. Lilly<sup>a</sup>

<sup>a</sup>The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK. E-mail: [jean.robertson@hutton.ac.uk](mailto:jean.robertson@hutton.ac.uk); [nikki.baggaley@hutton.ac.uk](mailto:nikki.baggaley@hutton.ac.uk); [blair.mckenzie@hutton.ac.uk](mailto:blair.mckenzie@hutton.ac.uk); [jason.owen@hutton.ac.uk](mailto:jason.owen@hutton.ac.uk); [allan.lilly@hutton.ac.uk](mailto:allan.lilly@hutton.ac.uk)

<sup>b</sup>BÜCHI Labortechnik AG, Flawil, Switzerland. E-mail: [PerezFernandez.E@buchi.com](mailto:PerezFernandez.E@buchi.com)

The National Soil Inventory of Scotland (NSIS) is an objective dataset and represents the most common soil types across Scotland. The soils in this unique dataset have been extensively characterised, with a wide range of chemical and physical parameters measured, including the near infrared reflectance (NIR) spectra. Significantly, the parameters measured have often been carried using more than one analytical method. In this work, we are looking to build on previously developed NIR calibrations for prediction of soil parameters and extend the useful information that can be gained from the NIR soil spectral data. We have examined in more detail the differences in the NIR correlations between elemental concentrations for some of the soil nutrients (Al, Fe, K, Mg, Mn, Na and P) measured in different ways. In addition, we are increasing the range of parameters used for calibration development to ones more directly linked to soil function. Here results are reported for correlations to soil aggregate stability data, which we are looking to apply to improving prediction of erosion risk.

## Introduction

The National Soils Inventory of Scotland (NSIS) is an objective dataset of Scottish soils originally sampled on a 10km grid across the entire country between 1978 and 1988 and repeated for a subset of these locations on a 20km grid between 2007 and 2009. At each of 183 locations on the 20km grid throughout Scotland all the major horizons of a soil profile were sampled on both visits and the samples taken in the 1978–88 campaign were dried and archived. These samples represent the soil types common in Scotland and are also representative of many parts of northern Europe. The soils in this unique dataset have been extensively characterised, with a wide range of chemical and physical parameters measured, including the near infrared reflectance (NIR) spectra. Significantly, the parameters measured have often been carried out using multiple analytical methods e.g. total P, oxalate extractable and aqua regia extracted P have all been measured on all these soils. Previously local and global

calibrations for prediction of many of the soil parameters from the NIR spectra have been developed and, as for other studies, have shown very promising results for a range of parameters (total elemental carbon and nitrogen, loss on ignition (450 °C and 900 °C), exchangeable H and Mg, moisture content, pH (in H<sub>2</sub>O and CaCl<sub>2</sub>) and dry bulk density).<sup>1</sup> In this work we look to extend the useful information that can be gained for land managers by using the NIR soil spectral data. There are two different aspects to this work: firstly, examining in more detail the differences in the NIR correlations between element concentrations measured in different ways and secondly extending the range of parameters used for calibration development to ones more directly linked to soil function. In particular, developing the work on predicting soil physical properties, we will describe correlating NIR spectral data to soil aggregate stability data with a view to better prediction a soil's susceptibility to erosion.

### Correspondence

A.H.J. Robertson ([jean.robertson@hutton.ac.uk](mailto:jean.robertson@hutton.ac.uk))

doi: 10.1255/nir2017.123

**Citation:** A.H.J. Robertson, E. Perez-Fernandez, N.J. Baggaley, B.M. McKenzie, I.J. Owen and A. Lilly, "Extending predictions of soil parameters for a national dataset", in Proc. 18<sup>th</sup> Int. Conf. Near Infrared Spectrosc., Ed by S.B. Engelsens, K.M. Sørensen and F. van den Berg. IMPublications Open, Chichester, pp. 123–127 (2019). <https://doi.org/10.1255/nir2017.123>

© 2019 The Authors

This licence permits you to use, share, copy and redistribute the paper in any medium or any format provided that a full citation to the original paper is given, the use is not for commercial purposes and the paper is not changed in any way.



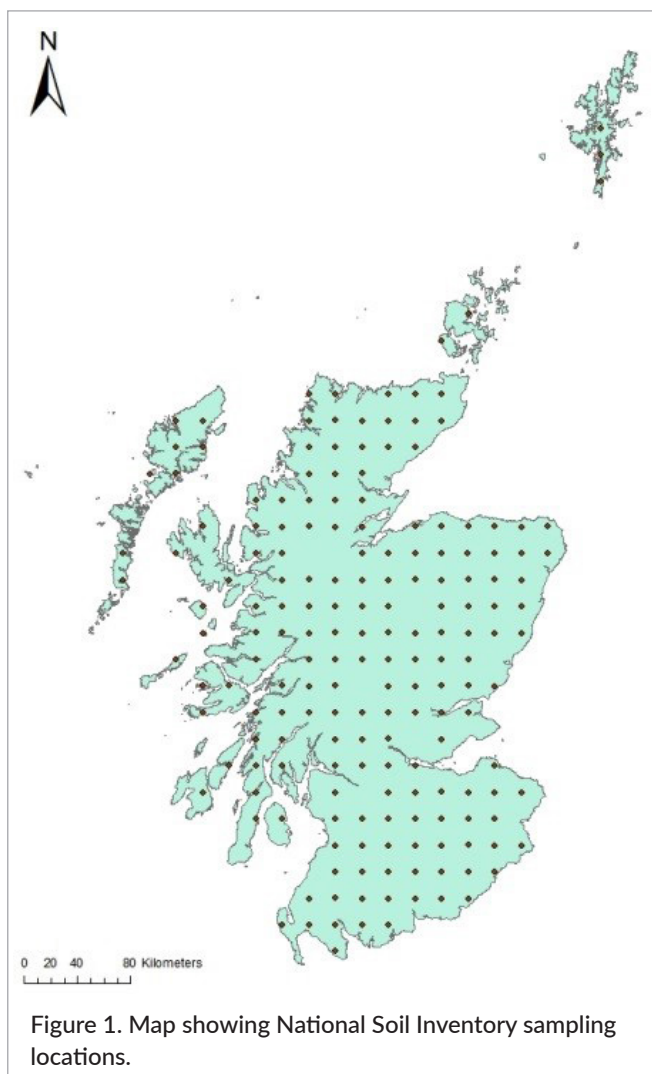
ISBN: 978-1-906715-27-4

## Methods

### Dataset

Spectral analysis was carried out on a spatial dataset of soils sampled on a 20 km grid throughout Scotland, as part of the National Soil Inventory of Scotland (NSIS),<sup>2-4</sup> see Figure 1.

Between 1978 and 1988 an inventory of soils in Scotland was undertaken by locating and describing the soil profile at each 5 km intersection between the easting and northing grid lines. The major soil horizons in each soil profile were also sampled at each 10 km grid intersection giving a dataset of 721 sampled profiles (NSIS1 samples). Between 2007 and 2009, 183 of these locations which had soil were revisited on a 20 km sample frame, that is, 25% of the original sites. At each location the site characteristics such as slope, rockiness and aspect were matched with the previous description and a soil profile similar to the original was identified and each



of the major soil horizons were sampled giving a total of 700 samples (NSIS2 samples). For spectral analysis, soils from the original NSIS1 profiles that matched or were the equivalent of the 700 NSIS2 samples were retrieved from the National Soils Archive. However, some of the archived samples had been depleted, and only 546 of the original NSIS1 samples were available for analyses giving a total of 1246 samples from both sampling campaigns. The soils were sampled from a range of different land uses such as: arable, both improved and semi-natural grassland, moorland, bog and woodland.

Aggregate stability analysis was only undertaken on the mineral soils from the NSIS2 campaign giving a total of 69 mineral topsoils and 69 mineral subsoils.

### Spectroscopic analysis

The 1246 Samples for NIR analysis were air dried (30°C) and sieved (<2 mm). NIR spectra in reflectance mode were recorded at 2 nm intervals in the range from 1100 nm to 2500 nm using a FOSS NIRSystems 5000 spectrophotometer (FOSS NIRSystems, Silver Springs, MD, USA), with a transport module sampling accessory and a quarter cup sample holder. Infracore International ISI Scan software, version 2.85.3 (Infracore International LLC., Port Matilda, PA, USA) was used to operate the spectrophotometer. The NSIS1 and NSIS2 samples were scanned at the same time, but in a randomised order, to avoid bias within the spectra due to the effect of varying laboratory conditions (e.g. temperature, humidity).

Partial least squares regression (PLSR) was used to develop calibrations to predict 32 soil parameters<sup>1</sup> from the total NSIS dataset of 1246 samples. A number of spectral pre-treatments were used, including first and second derivatives, as well as scatter correction (e.g. standard normal variate and detrending SNVD, multiplicative scatter correction MSC).

Some elements had more than one calibration developed as there were multiple analytical approaches or definitions used for their presence in the soil (e.g. available vs total). For these calibrations, a validation set of 250 samples representative of the population as a whole (approximately 20% of the dataset, including 125 samples from each NSIS1 and NSIS2) was selected using Shenk and Westerhaus's sample selection algorithms.<sup>5</sup> The remaining library of 996 samples was used to develop Global and Local PLS.

The calibrations developed for the aggregate stability had a total number of 132 samples, all of which were

used as calibration samples. Cross validation was the method of choice to evaluate calibration performance. A three-dimensional plot of Principal Component Analysis (PCA) scores of this set (Figure 2) shows that the majority of samples for this NSIS calibration subset are well clustered, with a relatively small number lying outside the main cluster. Some of the outliers were excluded from the calibration during cross validation.

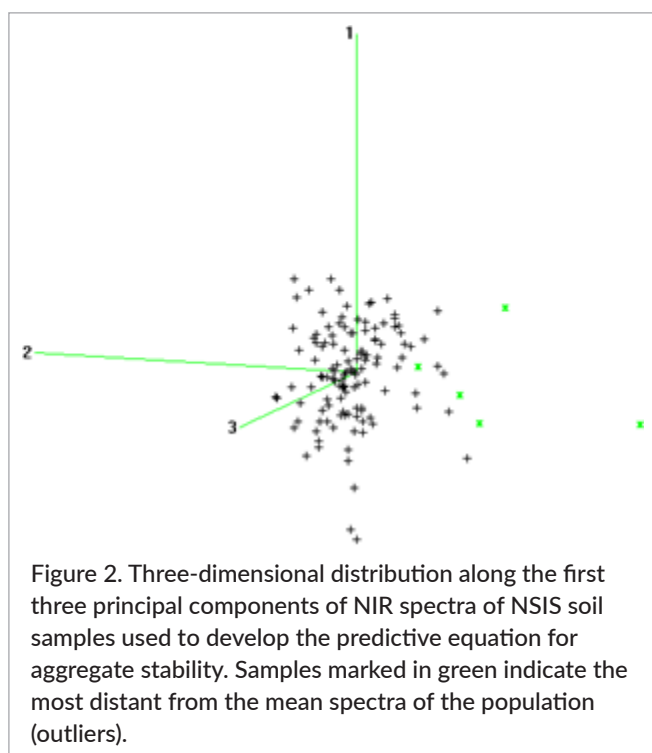
### Chemical analysis of the soil samples

Sample processing and laboratory analyses both for NSIS1 archived soil samples and NSIS2 samples were carried out between 2007 and 2010 at the James Hutton Institute, using identical protocols and instrumentation for both sets.

Soils were analysed for over 30 different parameters.<sup>1</sup> Analysis of Al, Fe, K, Mg, Mn, Na and P was carried out using at least two of the following methods: exchangeable (ammonium acetate extraction), ammonium oxalate extraction and aqua regia digest. The data used in this work are contained within the Scottish Soils Database.<sup>6</sup>

### Exchangeable cations

The exchangeable cations (such as Al, Fe, K, Mg, Mn, Na) in a soil are removed in solution by exchanging them with  $\text{NH}_4^+$  from a 1 M ammonium acetate solution (pH 7).



**Figure 2.** Three-dimensional distribution along the first three principal components of NIR spectra of NSIS soil samples used to develop the predictive equation for aggregate stability. Samples marked in green indicate the most distant from the mean spectra of the population (outliers).

Ammonium acetate was added to approximately 10g of air-dried soil (30°C), sieved to pass a 2 mm aperture, to form a soil slurry and left overnight to equilibrate. The resulting soil slurry was filtered and then leached with approximately 100 mL of 1 M ammonium acetate and left overnight prior to leaching with ammonium acetate to 250 mL. The concentrations of the extracted cations were determined using inductively coupled plasma-optical emission spectroscopy (ICP-OES).

### Ammonium oxalate extraction

Acid ammonium oxalate extractions were used to determine Al, Fe, Mn and P from poorly crystalline aluminosilicates, ferrihydrite, and Al- and Fe-humus complexes. Milled soil (60 mg) was placed into a centrifuge test tube along with acid ammonium oxalate (30 mL; 0.2 M) extractant (prepared from ammonium oxalate and oxalic acid, pH 3). The samples were first shaken and then centrifuged for 15 min at room temperature. The samples were filtered prior to analysis by inductively coupled plasma-mass spectrometry (ICP-MS).

### Aqua regia digest

Approximately 2 mg of air-dried (30°C), milled, soil samples were digested in a mixture of 21 mL of hydrochloric and 7 mL of nitric acid (aqua regia) under reflux, for two hours (130°C). The digest was allowed to cool and was filtered and made up to a known volume with 0.5 M  $\text{HNO}_3$ . The digest was then analysed by ICP-MS to determine element concentration (K, Mg, Na and P expressed as  $\text{mg kg}^{-1}$ ).

### Aggregate stability

Aggregate stability for the 132 samples described earlier was determined by wet sieving using the Kemper & Rosenau (1986) method.<sup>7</sup>

The sample for analysis was allowed to air-dry until they were in a friable state and then passed through an 8 mm sieve. The <8 mm fraction was then left to “totally” air-dry. The proportion of water stable aggregates was determined using an Eijkelkamp wet sieving machine as described by Kemper & Rosenau.<sup>7</sup> Four subsamples of approximately 4 g were taken and tipped onto sieves (aperture sizes 2 mm, 1 mm, 0.5 mm and 0.25 mm). No pre-moistening was done prior to wet sieving. The samples were immersed in water repeated for three minutes (giving a total of around 100 immersions) using an Eijkelkamp wet sieving machine. The material remaining on the sieve was

collected into a weighing tin and dried at 105 °C for 24 hours. The wet-sieving procedure was repeated three times. The proportion of dried material remaining on the sieve in relation to the starting weight gives a measure of the stability of the aggregates, the greater the soil remaining, the stronger the aggregates.

## Results

### Correlations to soil nutrients

The results in relation to the correlation of elements analysed in different ways is highly complex, but patterns do emerge, with implications for soil nutrient analysis. Al and Fe when measured by ammonium oxalate extraction are better correlated to the NIR spectral data ( $R^2$  global = 0.57, 0.71;  $R^2$  local = 0.66, 0.75) than when measured by ammonium acetate extraction (or exchangeable— $R^2$  global = 0.32, 0.55;  $R^2$  local = 0.55, 0.50). The ammonium oxalate extraction is thought to remove the amorphous, inorganic and complexed metals whereas the exchangeable method gives the plant available metals, i.e. the cations that plant roots can absorb by swapping them for  $H^+$ . It therefore appears that the NIR spectra are better reflecting the Al and Fe in the amorphous, inorganic and complexed states than that which is plant available. Although Mn concentrations in the soil can also be measured using these two methods they are generally much more poorly correlated to the NIR spectra than the Al and Fe were.

When we consider the exchangeable K, Mg and Na, all appear quite well correlated to the NIR spectra ( $R^2$  global = 0.69, 0.89, 0.84;  $R^2$  local = 0.75, 0.83, 0.72). The local calibrations, however, appear to give much better standard error of prediction (SEP) for these exchangeable values than the global calibrations do. The global validation results for Na, however, are unexpectedly poor and, although this is not reflected in the standard error of cross validation (SECV) for the global calibration and the SEP for the local calibrations, will be further investigated in future. The relationship between the exchangeable values and plant available nutrients make these local calibrations potentially amongst the most useful. Generally, the correlations between the NIR spectra and the aqua regia digest values (or “pseudo totals”) for these elements are similar to (K, Mg) or worse (Na) than those for the exchangeable values. In future work we intend to compare results for Fourier transform infrared spectra,

which better reflect the overall geology of the soil, and it will be interesting to see whether those spectra correlate better to the aqua regia values.

For P, comparing the results of the aqua regia and ammonium oxalate extractable correlations to the NIR spectra suggests similar correlations to both for the global calibrations, but with relatively poor validation results ( $R^2$  global = 0.63, 0.60, respectively). However, for the local calibrations, correlations were generally lower ( $R^2$  local = 0.39, 0.49) but with better SEP values than for the global calibrations. The ammonium oxalate extractable local calibration performed relatively well with a SEP close to the SECV for the global calibration. Much remains to be investigated but there is potential for development of some useful methods.

### Correlations to soil aggregate stability

We found promising correlations between NIR spectra and aggregate stability values. Correlation between NIR spectra and aggregate stability values varied with aperture size of the sieve, i.e. with the amount of material passing through different sieve sizes. We found better accuracy and precision for samples passing through smaller apertures ( $R^2$  = 0.48 and SECV = 0.11, for the 0.25 mm sieve size and  $R^2$  = 0.62, SECV = 0.15 for 0.5 mm sieve size) than for those that passed through the larger sieve sizes ( $R^2$  = 0.29 and SECV = 0.17, for the 2 mm sieve size and  $R^2$  = 0.43, SECV = 0.18 for 1 mm sieve size). In all cases, the best results were achieved using a 1<sup>st</sup> derivative transformation together with standard normal variate and de-trending (SNVD) scatter correction.

The scatter plots of predicted vs reference aggregate stability values for the 0.5 mm sieve and the 0.25 mm sieve (Figure 3A and 3B, respectively) show similar patterns with many samples well predicted across the range from high to low aggregate stability but also a number of groups of samples which are more poorly predicted. In particular, there is a group of low measured values which are over-predicted and a group of higher aggregate stability values which are under-predicted. The nature of these samples will be investigated in further work to improve the calibrations, as will the differences in results between the different sieve sizes. However, the initial results appear promising and an ability to predict the aggregate stability of mineral soil samples from NIR spectra would replace a time-consuming process and make aggregate stability measures much more widely

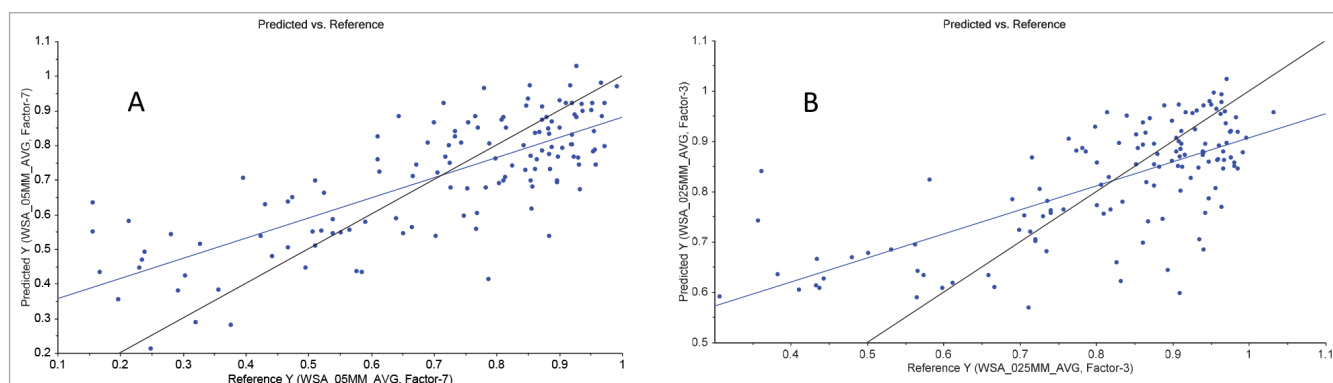


Figure 3. Scatter plots of Predicted vs Reference aggregate stability values, displaying calibration (blue) and target (black) regression lines for A) 0.5 mm sieve and B) 0.25 mm sieve.

available. Further work will include using the NIR spectral data, and aggregate predictions in assessment of erosion risk.

## Conclusions

Correlations between the NIR spectral data of soils in the NSIS dataset and soil nutrients, measured in different ways, reveal some interesting trends. It is clear that understanding of the methods used in the analysis is important and the various relationships to the NIR spectra require more study. Further work on this complex dataset could provide valuable information relevant to soil nutrient analysis.

Initial results for correlations between the NIR spectral data and soil aggregate stability are promising, and will be pursued in relation to prediction of erosion risk.

## Acknowledgements

Special thanks go to all the NSIS team who carried out the soil sampling, processing and analyses; and Carrie Donald, Rachael Hill and Caroline Thomson for spectra collection. This work was funded by the Rural & Environment Science & Analytical Services Division of the Scottish Government.

## References

1. E. Perez-Fernandez and A.H.J. Robertson, "Global and local calibrations to predict chemical and

physical properties of a national spatial dataset of Scottish soils from their near infrared spectra", *J. Near Infrared Spectrosc.* **24**, 293–316 (2016). <https://doi.org/10.1255/jnirs.1229>

2. A. Lilly, J.S. Bell, G. Hudson, A.J. Nolan and W. Towers, *National Soil Inventory of Scotland 1 (NSIS\_1): site location, sampling and profile description protocols (1978-1988)*. Technical Bulletin, Macaulay Institute (2010).
3. S.J. Chapman, J.S. Bell, C.D. Campbell, G. Hudson, A. Lilly, A.J. Nolan, A.H.J. Robertson, J.M. Potts and W. Towers, "Comparison of soil carbon stocks in Scottish soils between 1978 and 2009", *Eur. J. Soil Sci.* **64**, 455–465 (2013). <https://doi.org/10.1111/ejss.12041>
4. A. Lilly and S.J. Chapman, "Assessing changes in carbon stocks of Scottish soils: lessons learnt", *Earth Environ. Sci.* **25**, Article No. 012016 (2015).
5. J.S. Shenk and M.O. Westerhaus, "Population definition, sample selection, and calibration procedures for near-infrared reflectance spectroscopy", *Crop Sci.* **31**, 469 (1991). <https://doi.org/10.2135/cropsci1991.0011183X003100020049x>
6. Scottish Soils Database, <http://www.soils-scotland.gov.uk>
7. W.D. Kemper and R.C. Rosenau, "Stability and size distribution", in *Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods*, 2<sup>nd</sup> Edn, Ed by A. Klute. Agronomy Monographs, 9, ASA and SSSA, Madison, WI, pp. 425–442 (1986).