Use of handheld mid-infrared spectroscopy and partial least-squares regression for the prediction of the phosphorus buffering index in Australian soils

Sean T. Forrester[A], *Les J. Janik*[A], *José M. Soriano-Disla*[A,B,C,G], *Sean Mason*[C], *Lucy Burkitt*[D], *Phil Moody*[E], *Cameron J. P. Gourley*[F], *and Michael J. McLaughlin*[A,C]

[A]Contaminant Biogeochemistry and Environmental Toxicology Group, CSIRO Land and Water,
  Waite Campus, Waite Road, Urrbrae, SA 5064, Australia.
[B]Department of Agrochemistry and Environment, University Miguel Hernández of Elche.
  Avenuenida de la Universidad S/N, 03202. Elche, Spain.
[C]University of Adelaide, PMB 2, Glen Osmond, SA 5064, Australia.
[D]Fertiliser and Lime Research Centre, Massey University, Private Bag 11222, Palmerston North,
  4442, New Zealand.
[E]Science Division, DSITIA, 41 Boggo Road, Dutton Park, Qld 4102, Australia.
[F]Victorian Department of Environment and Primary Industries, RMB 2460, Ellinbank, Vic. 3821, Australia.
[G]Corresponding author. Email: jose.sorianodisla@csiro.au

Abstract. The development of techniques for the rapid, inexpensive and accurate determination of the phosphorus (P) buffer index (PBI) in soils is important in terms of increasing the efficiency of P application for optimum crop requirements and preventing environmental pollution due to excessive use of P fertilisers. This paper describes the successful implementation of partial least-squares regression (PLSR) from spectra obtained with bench-top and handheld mid-infrared (MIR) spectrometers for the prediction of PBI on 601 representative Australian agricultural soils. By contrast, poor predictions were obtained for available (Colwell) P. Regression models were successfully derived for PBI ranges of 0–800 and 0–150, the latter range resulting in the optimum model considering the dominance of low PBI soils in the sample set. Concentrations of some major soil minerals (mainly kaolinite and gibbsite content for high PBI, and smectites or illites for low PBI), quartz (representative of low surface area of soils) and, to a lesser extent, carbonate and soil organic matter were identified as the main drivers of the PBI models. Models developed with soils sieved to <2 mm presented an accuracy similar to those developed using fine-ground material. The accuracy of the PLSR for the prediction of PBI by using bench-top and handheld instruments was also similar. Our results confirm the possibility of using MIR spectroscopy for the onsite prediction of PBI.

Additional keywords: Al and Fe oxyhydroxides, available phosphorus, particle size, phosphorus buffering capacity, phosphorus sorption, portable MIR.

Received 9 May 2014, accepted 14 September 2014, published online 12 January 2015

Introduction

Australian soils contain low inherent quantities of plant-available phosphorus (P), leading to the application of P fertilisers (Colwell 1963). Excessive use of P fertilisers leads to P contamination of surface or groundwater via runoff and leaching, respectively, and this has been cited as an important source of eutrophication, algal blooms, and displacement of native aquatic systems (Bomans *et al*. 2005; Burkitt *et al*. 2010). Thus, it is imperative to optimise P fertiliser input in order to increase the available P pool above the critical level required to meet plant demand (McLaughlin *et al*. 2011).

The availability of P in soil and the effectiveness of P fertilisers are determined by the ability of a soil to sorb the applied P (Barrow 1978). Soil P sorption has traditionally been evaluated by the P buffer capacity (PBC; Ozanne 1980), which provides a useful tool for improved management of P in agricultural systems (Bolland *et al*. 1996; Burkitt *et al*. 2002; Moody 2007). A simple, single-addition P-buffering index (PBI) was successfully developed as a surrogate of PBC in Australian soils (Burkitt *et al*. 2002) and it is widely used by research and commercial soil testing laboratories throughout Australia. This index is routinely used to modify 'critical' Colwell P levels in determining optimal yields for pastures and various crop types across a wide range of soil types (Gourley *et al*. 2006; Moody 2007; Speirs *et al*. 2013). Additionally, the PBI is increasingly being used to assess the risk of P loss in surface runoff (Burkitt *et al*. 2010; Hart and Cornish 2012). The determination of PBI is relatively time-consuming (Burkitt *et al*. 2002) and values

measured on Australian agricultural soils vary widely (Burkitt *et al.* 2002; Moody 2007). Thus, the development of a precise but more rapid technique for the prediction of PBI would be a considerable advantage.

Diffuse-reflectance infrared Fourier-transform (DRIFT) spectroscopy, using the mid-infrared (MIR) spectral range coupled with partial least-squares regression (PLSR), has been shown to have potential for the prediction of PBI (Forrester *et al.* 2003, 2010*a*). The DRIFT-PLSR technique (previously described, for example, in Janik *et al.* 1998) has been used for the prediction of many soil properties; it is rapid and inexpensive, does not require chemical reagents, allows for the simultaneous prediction of multiple analytes, and has the potential to be easily adaptable to portable instrumentation for in-field measurement.

The MIR frequency region is sensitive to the fundamental vibrations of major soil components such as quartz, 2 : 1 and 1 : 1 clays (e.g. smectites/illites and kaolinite), aluminium (Al) and iron (Fe) oxides/oxyhydroxides (e.g. gibbsite and goethite), carbonates and soil organic matter (Van der Marel and Beutelspacher 1976; Nguyen *et al.* 1991; Janik *et al.* 1998; Reeves *et al.* 2001). Apart from quantitative prediction, the interpretation of the spectral peaks responsible for multivariate models assists in an understanding of the most important soil properties controlling P sorption.

The P-sorption capacity of soils has been related to the presence of Al and Fe oxides and hydroxides, which adsorb phosphate by an exchange reaction (McKeague and Day 1966; Scheinost and Schwertmann 1995; Bertrand *et al.* 2003; Burkitt *et al.* 2006; Moody *et al.* 2013). In high-pH soils, calcium carbonate, with high surface area, is a primary sorbent for P-forming calcium phosphate (Samadi and Gilkes 1998; Bertrand *et al.* 2003; Bomans *et al.* 2005), whereas larger carbonate and kaolinite micro-aggregates can have reactive Al and Fe oxide surface coatings responsible for P sorption (Ryan *et al.* 1984; Bomans *et al.* 2005). Soil organic matter (SOM) has also been shown to be important in P sorption (Kudeyarova *et al.* 1991; Bainbridge *et al.* 1995).

A recent review by Soriano-Disla *et al.* (2014) noted that only a few studies predicted P sorption via infrared spectroscopy. Bainbridge *et al.* (1995) predicted the P sorption isotherm slope ($R^2 = 0.82$) for soils from South Africa by using near-infrared (NIR) spectroscopy. However, Cohen *et al.* (2007) were less successful ($R^2 = 0.69$, $n = 300$) in the prediction of P sorption in wetlands of Florida, USA, using the visible–NIR region. Successful models for Australian soils were previously reported for PBI by using MIR spectroscopy, with prediction R^2 values ranging from 0.87 to 0.82 (Janik *et al.* 1998; Forrester *et al.* 2003, 2010*a*), and for residual P ($R^2 = 0.84$–0.79) after equilibration with 40 mg P/L (Minasny *et al.* 2009). However, apart from the study by Forrester *et al.* (2003), which used soils throughout Australia, the other studies reported on samples from a specific region, e.g. New South Wales, and did not disclose the main spectral features corresponding to the P sorption model.

With the recent availability of handheld MIR technology, DRIFT spectroscopy could be adapted to on-site analysis, thus reducing the time associated with the transport of the sample and, in some cases, the need for sample pre-treatment (Reeves 2010). However, there are questions about the relative performance of the handheld devices compared with bench-top instruments. Furthermore, inter- and intra-particle variability, as found in field conditions, can be a major source of loss of accuracy (Stumpe *et al.* 2011) due to the small sampling spot (~4 mm diameter) of the incident infrared beam of most MIR spectrometers.

To our knowledge, the single previous attempt to derive a model for PBI using MIR, handheld instruments has been made by Forrester *et al.* (2010*a*), using a small set of 30 soils modelled by PLSR cross-validation, but without discussing the specific soil mechanisms responsible for deriving the models. Thus, there is a need to develop a robust PBI calibration using a handheld device for a much larger range of soil types, and to test the influence of soil sample heterogeneity on the PBI predictions. With regard to the development of PBI models across a range of soils, this may present a challenge, because the soil mechanisms responsible for such calibrations may vary depending on the characteristics of the soils being examined (Burkitt *et al.* 2006; McLaughlin *et al.* 2011; Moody *et al.* 2013). Such indirect calibrations may lead to instability problems for global calibrations over a large variety of soil types, or over large geographical areas (Stenberg and Viscarra-Rossel 2010).

The main aim of this study was therefore to test the performance of a handheld MIR spectrometer for the prediction of PBI across a wide range of Australian soil types and PBI values. We also tested the performance of the MIR technique to predict available (Colwell) P because of the requirement to evaluate both Colwell P and PBI for adjusting P fertiliser applications rates. We also set out to test the comparative performance of bench-top and handheld MIR spectrometers, to study the effect of sample heterogeneity, and to identify the main soil components responsible for the PBI calibration models.

Materials and methods

Soils

Two sets of Australian soils were collected: a relatively small set (set A, $n = 90$) and a large set (set B, $n = 516$). Set A also contained soil types representative of set B. The soils for both sets were collected from throughout Australia from sites with varying rainfall and climatic conditions, and had been under permanent pasture, cropping, and horticultural production. Soils ranged from those that had never received P fertiliser to those that had been heavily fertilised.

As previously described by Burkitt *et al.* (2002), the distribution of samples for set A was: 41 from Victoria, 15 from New South Wales, 13 from Queensland, 11 from Western Australia, 5 from South Australia, and 5 from Tasmania. The majority of soils were sampled to a depth of 0.10 m, with the exception of one soil that was sampled to a depth of 0.15 m. The soils were largely classified as Chromosols and Podosols, with moderate contributions from Dermosols, Ferrosols, Kandosols, Kurosols and Vertosols (Fig. 1; Isbell 1996).

Set B comprised soils from throughout Australia as above (excluding the Northern Territory), and including some from the archives of the National Soil Fertility Program
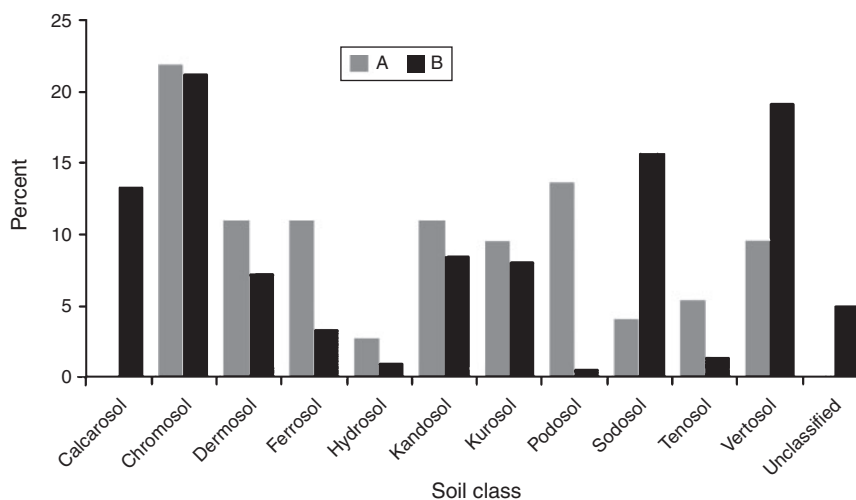
Fig. 1. Distribution of soil classes (%) according to the Australian Soil Classification system (Isbell 1996) for samples in set A (grey bars) and set B (black bars).

(in Moody *et al.* 2013). The majority of the soils ($n = 331$) were collected from a depth of 0.10 m, and a few were collected from other depths ($n = 66$ samples from 0.05 m, $n = 72$ from 0.075 m, and $n = 47$ from 0.15 m). The soils varied widely, with the largest proportion classified as Chromosols, Vertosols, Sodosols and Calcarosols, with moderate contributions from Kandosols and Kurosols and Dermosols (Fig. 1).

The primary purpose of using set A was to test the relative performance of regression models for PBI prediction derived using the soils sieved to <2 mm (unground) and the same soils finely ground to <0.1 mm. The two sets were combined (termed set A + B) to derive a more extensive calibration. This combined set was also used for testing two alternative spectrometer options: a bench-top and a handheld instrument.

Laboratory determination of PBI

The reference laboratory values of PBI were calculated according to Burkitt *et al*. (2002). A single addition of 1000 mg P kg^{-1} (as KH$_2$PO$_4$) was added at a 1:10 soil:solution ratio in 0.01 M CaCl$_2$. The suspensions were equilibrated by shaking in an end-over-end shaker at 25°C for 17 h. After equilibrium, P remaining in solution was determined, and PBI, referred to in Burkitt *et al*. (2002) as PBI$_{+Col}$, was calculated as:

$$PBI = \left(\frac{Ps + \text{initial Colwell P}}{c^{0.41}} \right)$$

where Ps is P sorbed (mg P kg^{-1} soil), c is final solution P concentration (mg P L^{-1}), and Colwell P (used as an estimation of the available P) has units of mg P kg^{-1} (Colwell 1963). Values of PBI are dimensionless.

Infrared spectra

The soils were oven-dried at 40°C for 12 h and sieved to <2 mm, then cooled to room temperature in a desiccator before infrared analysis. Additionally, soils from set A were further ground to <100 µm using a steel vibrating puck mill. Samples were

scanned in duplicate and the average spectra of replicates was used for model development.

Two Fourier transform infrared (FTIR) spectrometers were used: a bench-top Spectrum One spectrometer (Perkin Elmer Inc., Waltham, MA, USA) in the frequency range 7800–450 cm^{-1}, and a handheld Agilent 4100 spectrometer (Agilent, Santa Clara, CA, USA) in the frequency range 6000–650 cm^{-1}, both with a resolution of 8 cm^{-1}. The handheld instrument showed excessive noise below 800 cm^{-1} in the MIR and above 4000 cm^{-1} in the NIR (standard deviation of noise = 0.0049 and 0.80 absorbance units for the two ranges, respectively) compared with that of the bench-top spectrometer (standard deviation of noise = 9.4×10^{-5} and 4.7×10^{-5} absorbance units, respectively). To permit a comparison of modelling performance between the handheld and bench-top instruments, the final frequency range used for both instruments was restricted to the MIR spectral region only (4000–800 cm^{-1}).

Silicon carbide references (Perkin Elmer Inc.) were used as a background for both spectrometers (assumed to have a reflectance $R_0 = 1$), and absorbance spectra, in pseudo-absorbance (A) units (where A = $\log_{10} R_0 R_s^{-1}$), were calculated from the reflectance units of the sample (R_s). For multivariate analysis, the spectra were imported from the Perkin Elmer and the Exoscan format (initially in .SP and .ASP units, respectively) into Grams file format (.SPC) using the Grams-Al converter software (Thermo Fisher Scientific, Waltham, MA, USA) before importing the spectra into Unscrambler™ V9.8 (Camo, Oslo, Norway) spreadsheets. The spectra were pre-processed with the Standard Normal Variate (SNV) function to remove the multiplicative interference of scatter and particle size (Barnes *et al*. 1993), and baseline-corrected to remove slope and baseline offset. The baseline correction used was the Unscrambler 'De-trending' pre-processing application, by fitting a first-order polynomial to the full spectrum and then removing this as a baseline correction, thus resulting in a more robust correction than the use of a single point on the spectrum for estimating the degree of offset.

Multivariate modelling

A principal component analysis (PCA) was carried out using the spectral data for all samples. The PCA loadings and score distributions were used to discern the spectral features contributing to the spectral variance. PLSRs were derived from the DRIFT spectra (*X*-predictor variables) and experimental PBI (*Y*-dependent variables) values.

The following PLSR models were developed:

(i) For set A, cross-validation (with 10 sample rotation) using fine-ground (<100 μm) and unground (<2 mm sieved) soil material scanned on the bench-top instrument. Additionally, models for unground samples were developed by using an average of two repeat (replicate) samples and compared with those using only a single scan (repetition 2 *v.* 1).

(ii) For a composite set A + B, cross-validation using spectral data from only the unground soils, scanned with the bench-top and handheld instruments. In addition, this set was used for the development of models by cross-validation for the prediction of Colwell P with unground soils scanned in the bench-top machine.

(iii) For the composite set, using the 0–800 and 0–150 PBI ranges (the latter range being characteristic of Australian soils; Burkitt *et al.* 2002), again with spectral data from only the unground soils, scanned with the handheld device.

For the selected models corresponding with (*iii*), the samples were randomly split, two-thirds for developing calibration models and the remaining one-third for validation. The calibration samples were used for the development of the PLSR models for PBI prediction by cross-validation. During cross-validation, an optimum number of PLSR factors was established that minimises the root-mean-square error (RMSE) of the cross-validation without leading to overfitting. The resultant models were used to predict the concentrations of the samples in the test set. Models based on samples from 0–800 PBI values were used for the prediction of samples in the 0–800 and 0–150 PBI ranges. Similarly, models developed with samples from 0–150 PBI were used for the prediction of samples with 0–800 and 0–150 PBI range values. PLSR statistics for infrared predictions are generally reported as an average over the range of values being predicted. They are described in terms of the coefficient of determination (R^2), RMSE, and the ratio of the standard deviation of the reference PBI values to the RMSE of the prediction (RPD; Williams 1987). The prediction of PBI error statistics, namely RMSE, standard error (s.e.), bias, and the values of bias ± s.e. were calculated for PBI ranges used in the (*iii*) models.

The quality of PLSR for prediction purposes can be roughly ascertained from the RPD, with values <1.5 considered poor, 1.5–1.9 suggesting indicator quality, 2.0–2.9 suggesting good quality, and ≥3.0 analytical quality (Sudduth and Hummel 1996).

Results and discussion

Soil variability

The soils used in this study varied greatly in soil type and composition. This variability was manifested by changes in the infrared peaks due to the range of compositional soil minerals and soil organic matter in each of the soils. Plots of PCA scores 2 *v.* scores 1 and the corresponding PCA loadings can be found in Fig. 2, with the scores of the set B samples superimposed onto those of the set A samples. The score plot suggests a high degree of variability within the set B samples, with most of the set A samples lying within the bounds of PC1 and PC2 for set B, but with somewhat fewer samples from set A lying in the negative PC1/PC2 quadrant. The first two PCA loading weights, depicted in Fig. 2, identified the dominant soil components within the soil set that are associated with the main spectral features. The first PCA loading was dominated by peaks due to quartz (negative peaks near 2000–1750, 800 and 700 cm^{-1}, and positive peaks near 1250–1100 cm^{-1}) and kaolinite plus illite–smectite clays (positive peaks 3695–3620, 3620 and 3450–3200 cm^{-1}; Janik and Skjemstad 1995). Loading 2 was dominated by a strong, broad negative peak due to Al–OH in clays and absorbed water –OH attributed to 2 : 1 layer silicates such as smectite (3620–3000 cm^{-1}). A weak positive peak due to carbonate was observed near 2515 cm^{-1}. Thus, as shown in Fig. 2, quartz and clay were the main contributors to the variability in these spectra.

Partial least-squares regressions for Colwell P

The range of values found for the Colwell-P determination varied widely across the soils (range <1 to 563 mg kg^{-1}) with a standard deviation of 47 mg kg^{-1}. According to previous reports, PLSR models developed using only the MIR spectral range generally outperform those developed using the NIR (Janik *et al.* 1998; Reeves *et al.* 2001; Soriano-Disla *et al.* 2014). For this reason, and also because of the excessive spectral noise in the NIR region with the handheld instrument below 800 cm^{-1}, the spectral range used for Colwell P and PBI models in the present study was restricted to the MIR (4000–800 cm^{-1}) spectral range. The PLSR models developed for Colwell P showed a poor performance with the full range of data ($R^2 = 0.28$ and RMSE of cross-validation = 40 mg kg^{-1}). Restriction of the range of Colwell P values used in the calibrations, or data transformation (e.g. square-root transform), did not improve the regressions (data not shown). Similar poor results were found by Viscarra Rossel *et al.* (2006) for the prediction of Colwell P using MIR for a smaller sample set ($R^2 = 0.20$, $n = 49$). This is in accordance with previous studies using MIR for the prediction of extractable P (Soriano-Disla *et al.* 2014).

Partial least-squares regressions for PBI

Both set A ($n = 88$, PBI range 1–732, median = 91, standard deviation = 160) and set B ($n = 513$, PBI range 1–779, median = 59, standard deviation = 97) comprised a predominance of low PBI values (five samples with PBI values higher than 800 were removed). The combined dataset (A + B) resulted in distributions similar to set B, but with a higher number of samples available ($n = 601$) and a PBI range 1–779 (median = 62, standard deviation = 115). Because of the very large range of PBI values found in these soils, and the heavily skewed distribution towards low values (skewness coefficient = 2.2, 4.4 and 3.8 for sets A, B and A + B, respectively),

normalisation or linearisation (e.g. by square root or log) of the PBI data could be an advantage (Varmuza and Filmoser 2009). A uniform or normal distribution of reference data is preferable for deriving PLSR models and generally results in optimum predictions (Smith 1993; Maindonald and Braun 2010). For these data, the square-root transformation considerably reduced the distribution asymmetry, as illustrated in Fig. 3, and was useful in reducing the influence of the few high PBI samples on the PLSR models. Consequently, all PLSR modelling for the

0–800 PBI range was performed on the square-root transformed data.

Despite the demonstrated advantages of a square-root transform in modelling highly asymmetrical or non-linear data distributions, there are issues with reporting the errors associated with predicted data that have been previously transformed (Smith 1993). Generally, the requirement is for prediction results to be expressed in terms of their raw values to be useful. This involves back-transforming the predicted
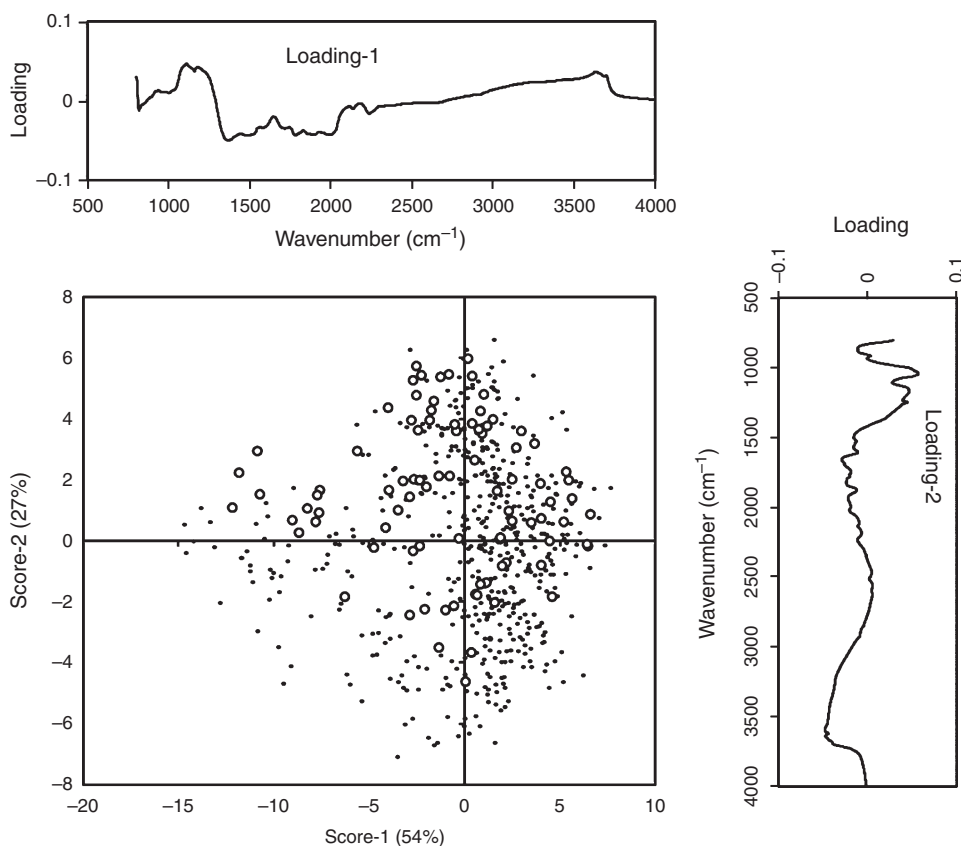
Fig. 2. Principal components analysis (PCA) plot of score 2 *v.* score 1 for the samples of set A (○) and set B (●). Corresponding PCA loading 2 and loading 1 are also shown.
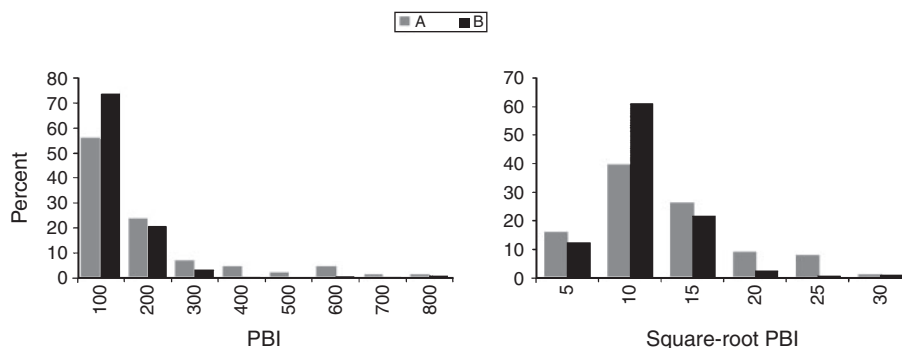
Fig. 3. Raw percent distribution of phosphorus buffer index (PBI) (left) and square-root transform (right) for soil set A (grey symbols) and set B (black symbols).

data by, in this case, squaring the predicted data (in this study taking into account the sign of the predicted values). Consequently, the PLSR predicted data that were based on square-root values were back-transformed for reporting the regression statistics.

Although data transformation was expected to result in improved PLSR models for the 0–800 PBI range, 66 of the 88 (75%) set A samples and 457 of the 513 (89%) set B samples had PBI values <150. The restriction of the PBI values to 0–150 normalised the distribution, with skewness values dropping sharply to 0.22, 0.55 and 0.49 for sets A, B and A+B, respectively. In addition, this restricted set included almost all of the original samples, covering almost the entire range of PBI categories as proposed by Moody (2007), ranging from extremely low (PBI <15) to high (PBI >281). Non-linear transformation, therefore, may not be the best approach or required for the reduced 0–150 PBI range.

Testing the effect of soil grinding

Given the heterogeneous nature of soils, some sample preparation is often performed to limit sample heterogeneity and other sources of non-systematic variability, thus improving reproducibility and accuracy of predictions from the spectra. The objective of sample pre-treatment is mainly to reduce inter- and intra-particulate heterogeneity, but may also involve drying the samples to reduce the effects of variable and uncontrolled moisture contents and grinding or pulverising to reduce the effects of varying particle size.

Fine grinding of a soil sample reduces inter- and intra-particular heterogeneity, resulting in more accurate PLSR (Baldock *et al.* 2013). Fine grinding (usually to <0.10 mm) helps to ensure a more homogenous mixture of soil micro-aggregates, improving access to the infrared incident beam by

exposing the inner composition of soil aggregates (Janik *et al.* 1998; Stumpe *et al.* 2011). For example, Brunet *et al.* (2007) found better results for the prediction of total carbon (C) and nitrogen (N) by using NIR spectroscopy in finely ground soils. However, soil grinding is a relatively time-consuming activity, and for this reason, we set out to test the relative performance of PLSR using finely ground and unground material scanned in a bench-top instrument (Table 1).

Figure 4 shows representative spectra, recorded with the bench-top instrument, for unground and finely ground soil selected from near the centre of the PCA score plot (Fig. 4 also includes the spectrum recorded with the handheld instrument, discussed later). The most notable feature was a significant reduction of the overall spectral intensity for the finely ground soil compared with the unground soil, probably because of increased reflection (and therefore reduced absorption) of the incident infrared radiation. However, similar cross-validation models were found for both sample pretreatments in set A ($R^2 = 0.87$, RMSE = 59, RPD = 2.7 for finely ground; and $R^2 = 0.89$, RMSE = 55, RPD = 2.9 for unground soil).

Reports in the literature on the effect of sample grinding with regard to the prediction performance of PLSR show mixed results and depend on the specific soil property of interest (Stenberg *et al.* 2010). Our results are in agreement with Reeves *et al.* (2010), who found similar results for total C and N with MIR by comparing calibrations based on air-dried samples with air-dried + ground samples. Using NIR spectroscopy, Nduwamungu *et al.* (2009) found that sieving soil samples to <2 mm appeared sufficient to reduce inter-sample heterogeneity, because sieving to <0.2, 0.5, or 1.0 mm did not improve calibration accuracy for a range of properties.

As discussed by Brunet *et al.* (2007), the fine grinding of sandy soils can result in the breaking of coarse sand particles

Table 1. Partial least-squares regression (PLSR) statistics for the prediction of phosphorus buffer index (PBI) using the bench-top and handheld spectrometers for the mid-infrared spectral range (800–4000 cm⁻¹)

G, Ground (<100 µm); UG, unground (<2 mm). Statistics are for median, standard deviation (s.d.), PLSR factors (PCs), coefficient of determination (R^2), root-mean-square error (RMSE), and ratio of the standard deviation of the reference PBI values to the RMSE of the prediction (RPD). PLSR cross-validation used a 10-sample rotation in the 0–800 PBI range. Validation was tested in the 0–800 PBI range using the 0–800 and 0–150 PBI calibrations, and on the 0–150 PBI range using the 0–800 and 0–150 PBI calibrations as indicated. Data for PBI used in PLSR modelling were square-root transformed, unless indicated otherwise (raw); square-root transformations applied for PLSR modelling were then back-transformed to show the final statistics

Soil set	G or UG	N	PBI range	Median	s.d.	Instrument type	PCs	R^2	RMSE	RPD
Cross-validation										
A	G	88	1–732	91	160	Bench-top	6	0.87	59	2.7
A	UG	88	1–732	91	160	Bench-top	6	0.89	55	2.9
A (replicate-1 only)	UG	88	1–732	91	160	Bench-top	6	0.85	63	2.5
A (replicate-2 only)	UG	88	1–732	91	160	Bench-top	6	0.88	59	2.7
A+B	UG	601	1–779	64	110	Bench-top	18	0.87	43	2.6
A+B	UG	601	1–779	64	110	Portable	18	0.82	50	2.2
Validation		Cal, Val[A]								
A+B (calibration 0–800 PBI)	UG	397, 204	3–769	61	115	Portable	17	0.83	53	2.2
A+B (calibration 0–800 PBI)	UG	397, 177	3–150	54	35	Portable	17	0.69	24	1.4
A+B (calibration 0–150 PBI)	UG	346, 177	3–150	57	35	Portable	18	0.77	17	2.1
A+B (calibration 0–150 PBI) (raw)	UG	346, 177	3–150	56	35	Portable	18	0.79	16	2.1
A+B (calibration 0–150 PBI) (raw)	UG	346, 204	3–769	61	115	Portable	18	0.38	98	1.2

[A]Number of calibration (Cal) and validation (Val) samples.
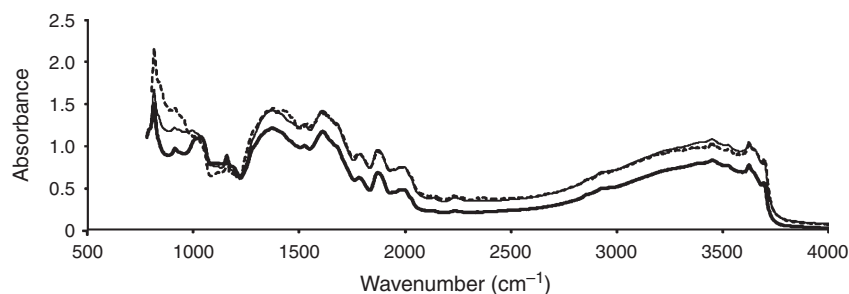
Fig. 4. Infrared spectra of fine ground (——) and unground (——) samples scanned with the laboratory bench-top instrument, and the unground (...) samples scanned with the handheld instrument.

and the creation of new reflection planes. Those authors found that grinding was less beneficial for the prediction using NIR of total C and N in sandy soils than clayey soils (Brunet *et al.* 2007). A similar discussion on the effect of grinding in sandy and clayey soils can be found in Stumpe *et al.* (2011), who recommended developing calibrations by using soil samples with similar or uniform texture.

Testing the effect of replicate sampling

In order to reduce the impact of sample heterogeneity and thus achieve optimum results, PLSR models were derived using the average of replicate scans for each sample, which were compared with those derived from a single scan. The optimum model was achieved when the averaged spectrum was used ($R^2 = 0.89$, RMSE = 55, RPD = 2.9). By contrast, when taken separately, there was some small variation in accuracy between the two replicates, with a range of R^2 0.85–0.88, RMSE 59–63 and RPD 2.5–2.7, suggesting that averaging for these soils for PBI was unnecessary. These results were assumed to be due to the reduced variability from averaging the data. The results presented here suggest that both inter- and intra-particulate heterogeneity in these samples is relatively low, an important implication for direct measurement of field samples where sample grinding is often not possible or practical.

Handheld MIR v. bench-top spectrometer

The data predicted using infrared spectroscopy can vary according to the type of instrumentation used (Reeves 2010). Different instrument specifications, such as, for example, resolution, spectral range, sample accessory and instrument performance, might influence the accuracy of the resulting spectra, or the sensitivity of the spectra to the soil property being studied (Soriano-Disla *et al.* 2014). In particular, most FTIR bench-top infrared spectrometers have very high signal-to-noise ratios compared with handheld instruments. Handheld spectrometers must be easy to handle, resistant to field conditions including high humidity, temperature and rough handling, and be usable for extended periods without recharging the battery. To achieve these requirements, the instrument energy needs are kept to a minimum, and for MIR capability, ZnSe optics are often used which, although being resistant to moisture, reduce beam throughput and limit the spectral range to ~5500–600 cm^{-1} (compared with 7800–

400 cm^{-1} for bench-top instruments, which use KBr optics). Thus, and in further developing infrared PLSR for the prediction of PBI for *in-situ* or field applications, the question arises as to whether handheld FTIR devices can provide performance similar to bench-top spectrometers.

A comparison of MIR spectra taken with both the bench-top and handheld instruments is shown in Fig. 4. Very few differences were observed between spectra of <2 mm soil in the 4000–1200 cm^{-1} spectral range. There was, however, an enhancement of the quartz peaks below 1000 cm^{-1} in the sample scanned with the handheld instrument, possibly due to the different optical configuration of the DRIFT accessory or differences in the reflectivity of the background material. Prediction statistics, for the comparison of PLSR cross-validation of the 0–800 range PBI data for the sample set A + B and using the bench-top *v.* handheld spectrometers, are summarised in Table 1. Samples were scanned unground (<2 mm) because this is more representative of the way in which samples are prepared in the laboratory and the state of samples potentially scanned in the field.

Results, using a PLSR model with 18 factors, showed that the cross-validation accuracy for PBI in set A + B was slightly reduced for the handheld instrument from $R^2 = 0.87$ and RMSE = 43 for the bench-top spectrometer, to $R^2 = 0.82$ and RMSE = 50 for the handheld instrument, thus still producing predictions of similar accuracy.

Few studies have reported the use of portable MIR instruments (Forrester *et al.* 2010*a*; Kuang *et al.* 2012; Soriano-Disla *et al.* 2014). In one study (Reeves *et al.* 2010), predictions of total C and N by using MIR spectroscopy with a handheld FTIR spectrometer (SOC-400; Surface Optics Corp., San Diego, CA, USA) were equal to those achieved with a bench-top instrument (Digilab FTS-7000; Varian, Palo Alto, CA, USA). The lack of use of MIR handheld spectrometers has been partly a result of the lack of commercial availability of MIR handheld instruments and the easier sampling and remote sensing capabilities of NIR instruments. The situation has changed recently, in that truly handheld, high-performance FTIR instruments, using diffuse reflectance sample scanning, have become commercially available.

Although both the bench-top and handheld instruments covered the spectral range 6000–600 cm^{-1}, tests to determine the optimum PLSR spectral range for PBI prediction showed that the range 4000–800 cm^{-1} could be used effectively on either

instrument, and that a spectral range beyond this frequency span was not required. This was fortuitous because the spectral noise level of the handheld instrument becomes very high above $5000\,cm^{-1}$ and below $800\,cm^{-1}$. These results expand the applications of the technique to the on-site scanning of samples. Further work is required to validate the models with samples under field conditions.

Development of the optimum range for the PBI prediction model

The ultimate objective of the present study was to derive a viable PBI prediction model typical of soils in Australia. For this purpose, we describe the validation results for two PLSR calibration ranges: 0–800 PBI units covering the full range of PBI encountered, and 0–150 PBI units accounting for >75% of the available soils. To this end, calibrations for these two sets were tested using the randomly selected 'test' set. A statistical summary is provided in Table 1 and the error distribution statistics for these models is shown in Table 2. For this study, Colwell P was included in the calculation of PBI and the values were compared with uncorrected PBI (as suggested by Burkitt *et al.* 2008). This resulted in a Pearson correlation coefficient of 0.99 and an RMSE of 20, the high correlation suggesting that models for PBI uncorrected for Colwell P would be a good approximation for Colwell P corrected PBI.

Initially, a PLSR calibration model of 397 samples in the 0–800 PBI range, using the handheld spectrometer spectra, was used to predict the values (also in the 0–800 PBI range) of the validation set. The regression plots of predicted *v.* reference PBI values, in both the 0–800 and 0–150 PBI ranges but using the 0–800 PBI calibration, are illustrated in Fig. 5. The prediction accuracy for the 0–800 PBI range validation samples was similar ($R^2 = 0.83$, RPD = 2.2, RMSE = 53) to that of the cross-validation of the full 0–800 PBI samples ($R^2 = 0.82$, RPD = 2.2, RMSE = 50), thus confirming the robustness of our models and the low risk of overfitting. The same model was then used for the prediction of samples within the 0–150 range, but in this case, there was a general decrease in validation accuracy ($R^2 = 0.69$, RPD = 1.4, RMSE = 24).

Our results confirmed previous successful results of Janik *et al.* (1998, 2009) and Minasny *et al.* (2009) for samples scanned on a bench-top spectrometer. However, in this study we used unground material in a handheld device and for a wider range of Australian soils. Specifically, our full-range model covered a range of values similar to those used by Janik *et al.* (1998) of 50–900 PBI; Janik *et al.* (2009) of 3–901 PBI, and a mean ± standard deviation P sorption 377 ± 235, RMSE = 93; and Minasny *et al.* (2009) of 0–900 PBI, and mean P sorption of 432, RMSE = 79. As shown for the 0–800 PBI range, our RMSE of prediction (53) was considerably lower than those observed in the two latter studies, possibly as a consequence of the considerably higher proportion of low PBI values in our sample set (median PBI = 64, mean PBI = 95). Our RMSE of prediction was also lower than the standard error (s.e. = 90) reported by Forrester *et al.* (2003) but their results included a higher proportion of PBI values >800.

The reported error values in the prediction of low PBI values were increased by inclusion of high PBI samples in the calibration (Fig. 5*a*). This observation thus prompted the question of whether different soil compositions were responsible for the low and high PBI models (specifically discussed earlier). It is evident that the 0–800 PBI calibration model contained some spectral information correlating with the successful prediction of samples with PBI from 0–150 units. However, the 0–800 PBI model also contained information relevant to the prediction of high-PBI samples, but such high PBI correlations may be irrelevant or confounding for predicting low PBI data.

The question of error distribution with respect to the values of predicted PBI is now considered. Included in Fig. 5 are plots of the prediction error distribution in the PBI ranges 0–50, 50–100, 100–150, 150–300, 300–500 and 500–800 units. The statistical indicators accounting for prediction accuracy are commonly calculated over the full range of soil property values. Whereas this may be valid for normal or uniformly distributed linear data, it does not hold for many soil applications. For example, in most soil datasets, organic C, carbonate, and in this case PBI demonstrate heavily skewed

Table 2. Validation error statistics for phosphorus buffer index (PBI) using the handheld spectrometer in the mid-infrared spectral range (800–$4000\,cm^{-1}$)

Three calibrations were tested: (1) the back-transformed 0–800 PBI range using the square-root-transformed 0–800 PBI calibration; (2) the 0–800 PBI range using the square-root transformed 0–150 PBI calibration; and (3) the 0–800 PBI range using the raw 0–150 PBI calibration. Error statistics expressed as root-mean-square error (RMSE), standard error (s.e.) and bias calculated for the 0–800, 0–150, and the separate 0–50, 50–100, 100–150, 150–300, 300–500 and 500–800 PBI ranges

Validation set	Validation	PBI range							
		0–800	0–150	0–50	50–100	100–150	150–300	300–500	500–800
(1) A + B (calibration	RMSE	53	24	24	17	36	38	142	232
0–800 PBI)	s.e.	33	24	22	18	36	33	68	46
	Bias	−6	6	7	2	8	12	−132	−210
(2) A + B (calibration	RMSE		17	12	13	32	90	258	507
0–150 PBI)	s.e.		16	10	12	30	18	19	23
	Bias		0	5	0	−13	−75	−252	−496
(3) A + B (calibration	RMSE		16	15	12	27	89	267	500
0–150 PBI) (raw)	s.e.		15	13	11	26	13	11	17
	Bias		2	3	1	−9	−76	−262	−489
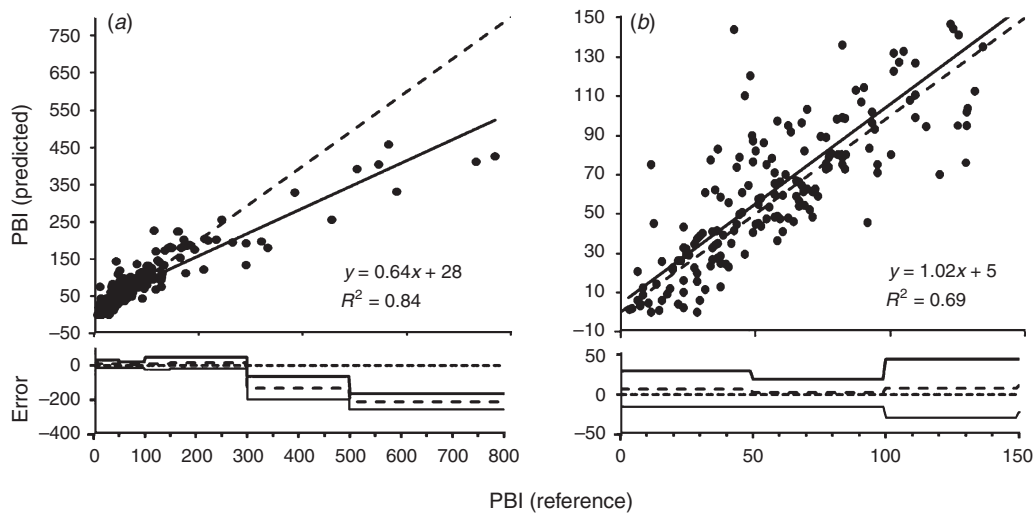
Fig. 5. Comparison of the prediction of back-transformed phosphorus buffer index (PBI) values for set A + B for the (*a*) 0–800 and (*b*) 0–150 PBI ranges, using the 0–800 PBI range calibration from the handheld spectrometer spectra. The dotted line is the 1 : 1 line.

distributions towards low values. Models developed for soil variables following such distributions can result in prediction residuals with the variance distribution depending in some way on the fitted value (heteroscedasticity; for example, when residuals tend to fan out as fitted values increase, giving a 'funnel' effect). We then face a major statistical problem when deriving confidence intervals and when testing the regression parameters, as well as in reporting the prediction errors for new observations (Varmuza and Filmoser 2009; Maindonald and Braun 2010).

In an attempt to avoid such problems, data transformation (e.g. square-root and log) is applied to the raw data before PLSR modelling (Crawley 2007; Janik *et al.* 2009) to normalise or linearise the data values. Alternatively, non-linear algorithms such as artificial neural networks can be used to model the raw data. Although transforms may achieve the desired linearisation and normal distribution, back-transformation of the predicted results to raw data format is subsequently required, leading to skewness and/or increasing bias in the back-transformed error distribution. We therefore propose that the usual regression statistics are inappropriate for describing prediction performance, and that a more detailed assessment of error for a particular analyte value may be required.

Figure 5*a* shows the increasing negative error for validation values from 0 to 800 PBI units, partly due to bias and partly due to s.e. That is, we have a highly non-uniform error distribution, increasing with the size of the predicted PBI values. The resulting error distributions, with respect to the various specific PBI ranges, are confirmed by the results in Table 2. The lowest prediction errors for the 0–800 calibration were found at the two ends of the 0–100 PBI range (s.e. 18–22, bias 2–7 and RMSE 17–24), but then doubled at the two ends of the of the 100–300 range (s.e. 33–36, bias 8–12 and RMSE 36–38) and increased even further in the 300–500 range (s.e. = 68, bias = −132 and RMSE 142). The s.e. for the 500–800 PBI range was lower (46) but this value was due to the very few

samples in this range. Bias appeared to have the strongest contribution to the increased error.

Because of the high proportion of samples in the low PBI range, it was thought that a more accurate prediction might result from a calibration limited to the 0–150 PBI range rather than one using the full 0–800 range. This option was tested by a separate validation set, using both the square-root transformed data (as for the regressions described above) and raw, non-transformed PBI data. In the case of the prediction of PBI in the 0–150 range from the square-root transformed data, we found overall regression statistics of $R^2 = 0.77$, RMSE = 17, RPD = 2.1 (see Table 1). The low RMSE suggested that this model could be useful for prediction purposes despite the low RPD.

The error in the 0–100 PBI range of predicted values using the 0–150 PBI square-root transformed model resulted in a range of s.e. 10–12, bias 0–5 and RMSE 12–13. The error increased to s.e. = 30, bias = −13, and RMSE = 32 for the 100–150 PBI values (see Table 2). A very large increase in error occurred for the 150–800 PBI range of values (s.e. 18–23, bias −75 to −496 and RMSE 90–507), mostly as a result of increasing bias at higher PBI values.

The error distribution for the 150–800 PBI range values (Fig. 6*a*), using the 0–150 PBI raw data model, showed a large increase in negative bias (−76, −262 and −489 for the 150–300, 300–500 and 500–800 PBI ranges, respectively; Table 2). The increase in RMSE observed for the overall model in the 150–800 PBI range could thus be attributed mostly to bias. The use of this model is, therefore, not advisable for the prediction of high-PBI samples (Table 1; Fig. 6*a*), clearly demonstrating that the spectral information required to predict low PBI samples is not sufficient for the prediction of high PBI samples. In a practical situation, if an unknown is predicted as high PBI (i.e. >150), we would use the 0–800 PBI model. The validation of the 0–150 PBI range model derived from raw (non-transformed) PBI data (see Table 1) resulted in $R^2 = 0.79$, RMSE = 16 and RPD = 2.1.

This was of similar accuracy to that of the square-root transformed data. The prediction errors from the 0–150 PBI raw (Table 2), non-transformed calibration resulted in a range of s.e. 11–13, bias 1–3 and RMSE 12–15 in the 0–100 range, and s.e. = 26, bias = –9, and RMSE = 27 for the 100–150 PBI value range. This represented a slight improvement over the square-root transformed data in this PBI range (RMSE = 32).

Soil mechanisms driving PBI predictions

Apart from quantitative predictions, MIR spectroscopy allows a qualitative description of the soil properties and compositions responsible for controlling PBI, by examining the PLSR loading weights. The first few loading weights show spectral peaks that correspond to the 'pure' soil components most strongly related to the soil property of interest (Janik *et al.* 1998).

In Fig. 7, we depict the loading weights for the first four PLSR loading weights for both the 0–800 and 0–150 PBI ranges. There was a strong negative correlation between PBI and quartz in loading weight for PLSR factor 1. This contribution is the most important in both the 0–800 and 0–150 ranges, along with a positive correlation with smectite in the 0–150 range and with kaolinite and gibbsite in the 0–800 range. Loading weights for PLSR factors 2 and 3 are more difficult to assign but appear to show further correlations with kaolinite and quartz and an unidentified negative peak at $1300\,cm^{-1}$. Also interesting is the weak contribution of carbonates ($2520\,cm^{-1}$), which is stronger for the 0–150 PBI model. This weak contribution may be related to the low proportion of highly calcareous
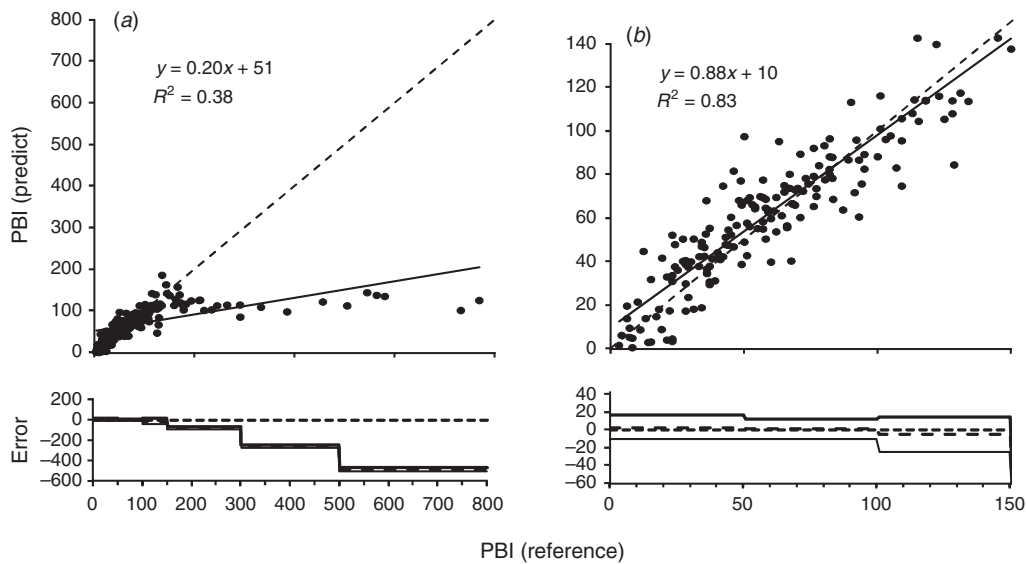
Fig. 6. Comparison of the prediction of raw phosphorus buffer index (PBI) values for set A + B for the (*a*) 0–800 and (*b*) 0–150 PBI ranges, using the 0–150 PBI range calibration from the handheld spectrometer spectra. The dotted line is the 1 : 1 line.
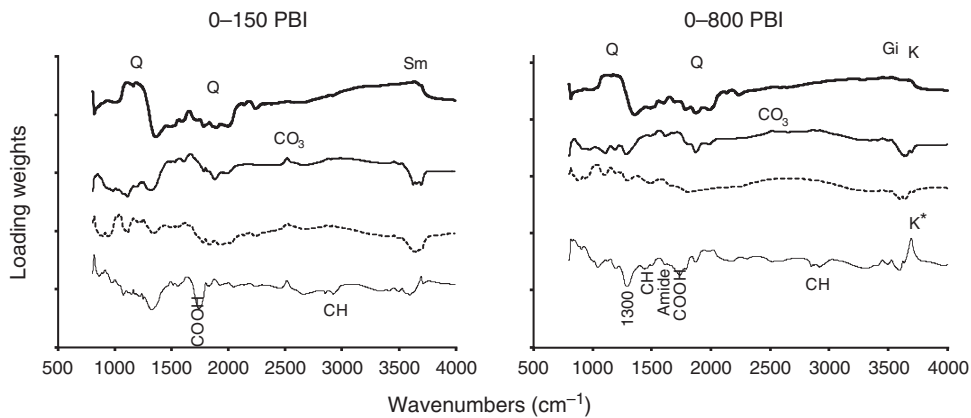
Fig. 7. Comparison of partial least-squares regression cross-validation loading weight plots for the 0–150 phosphorus buffer index (PBI) (left) and 0–800 PBI (right) with the handheld instrument. Loading weights shown for (top to bottom) PC1, PC2, PC3 and PC4. Major soil components shown are for carbonate (CO_3), gibbsite (Gi), kaolinite (K), organic matter (amide, CH, COOH) and quartz (Q).
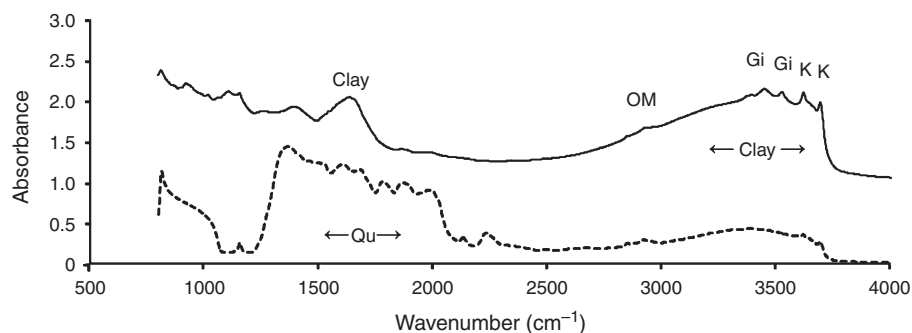
Fig. 8. Averaged mid-infrared spectra for the >400 phosphorus buffer index (PBI) (——) and the 0–5 PBI (- - -) ranges. Major soil components shown are for quartz (Qu), gibbsite (Gi), kaolinite (K) and organic matter (OM).

soils in our set of samples. The loading weight for PLSR factor 4 is characterised by kaolinite (the kaolinite peak is less intense in the 0–150 range) and negative contributions due to SOM. Peaks were observed near 2930–2850 cm^{-1} (aliphatic alkyl stretching, –CH_2), 1720 cm^{-1} (alkyl acid, –COOH) and 1670 cm^{-1} (protein amide, CO–NH). Despite the known sorption of P by goethite (FeOOH, peaks at 3141, 1791, 1656, 936, 836, 685 and 552 cm^{-1}), these peaks were not observed in the loading weights for either the low or high PBI ranges.

Changes observed in the loading weights of the models for the full and low range are indicative of different mechanisms responsible for P sorption. The spectral characteristics observed in the loading weights for the low and high PBI range calibrations support the results found using PLSR. This was possibly why we were unsuccessful in the prediction of 0–800 PBI samples using the 0–150 PBI calibration. Different sorption mechanisms appeared to be operating in high and low PBI samples, as illustrated in Fig. 8 by the average MIR spectra taken from samples with low PBI (<4) and high PBI (>400). In agreement with the interpretation of loading weight 1 for the PLSR models, the high PBI samples were characterised by clay (and thus low quartz as sand) and gibbsite. The contribution of organic matter in both average spectra was similar.

As discussed by Lins and Cox (1989) and Cox (1994), P sorption in soils is strongly dependent on clay content and/or surface area. Soils with high sand contents consequently have low clay contents, and thus typically have low P-sorption capacities. The quartz content is related to relative low surface area, and high surface area is related to the type and amount of clay minerals contained in the soil.

Apart from the surface area, the variable charge characteristics of Al and Fe oxy-hydroxides (e.g. goethite and gibbsite) and 1 : 1 lattice clays such as kaolinite are particularly important in P sorption by soils (Uehara and Gillman 1981). Both the surface area and the variable charge sorption mechanism apply to gibbsite, which has a high reactive surface area (Lins and Cox 1989), resulting in a high capacity to sorb P. When assessing the soil properties correlated with PBI in set B soils, Moody et al. (2013) found Mehlich-extractable Al correlated with PBI for the acidic soils in the set. Likewise, Cohen et al. (2007) found oxalate-extractable Al the primary driver of P sorption in soils of USA

wetlands. The important contribution of kaolinite as shown in Fig. 7 is supported by Bainbridge et al. (1995), who found that highly weathered, acidic red and yellow-brown soils with predominantly 1 : 1-type clays and high OM, occurring mainly in high-rainfall areas of South Africa, sorbed large amounts of P. As discussed by McLaughlin et al. (1981), in the case of silicate layer minerals, the layer silicates *per se* are of minor significance in the sorption of P by soils. More important is the presence of reactive (variable charge) Al/Fe oxide coatings on kaolinite minerals, which appear to be largely responsible for the sorption of P by increasing the number of reactive sites (McLaughlin et al. 1981).

A negative contribution of OM was found in the fourth loading. This contribution of OM functional groups for explaining PBI is weak compared with other soil components, partly because of the low concentration of OM in the soils. This is in contrast to the finding of Moody et al. (2013) who reported a positive correlation between PBI and OM for 142 soils in set B that were neutral–alkaline, non-calcareous and predominantly of 2 : 1-type clay mineralogy. A possible explanation for the existence of an overall weak negative relationship between OM and PBI when all set A and set B soils are considered may be the complexation of Al/Fe oxy-hydroxides by functional OM groups, thereby decreasing soil P-sorption capacity (Sibanda and Young 1986; Whitehead 2000). These apparently contrasting effects of OM on PBI, depending on which soil suite is considered, highlight the danger of generalising about P sorption processes across soils that vary in clay mineralogy.

Conclusions

This study demonstrates the feasibility of applying MIR-DRIFT spectroscopy, coupled with PLSR, to the cost-effective and rapid prediction of PBI in soils across Australia. However, it failed to predict Colwell P. Mechanisms responsible for the PLSR models suggested that for high PBI, the soil surface area, represented by the inverse of quartz content, gibbsite and alumino-silicate clay minerals such as kaolinite, was the main soil characteristic driving the PBI models. However, the models developed for the prediction of low PBI samples, besides being dependent on surface area, were also influenced by 2 : 1 clays such as smectites–illites, thus

confirming the existence of different mechanisms responsible for weaker P sorption.

The accuracy of the calibrations was not markedly affected by fine grinding the soils. A handheld FTIR produced similar calibration accuracy to that of the bench-top spectrometer.

According to the PBI ranges analysed and the model error distributions, an optimum model was developed for the prediction of PBI from 0 to 150 units. We also developed a successful model with PBI values ranging from 0 to 800, proving that the MIR-DRIFT technique can predict samples with much higher PBI values. The PBI distribution was heavily skewed towards low values and corrected by applying a square-root transformation of the PBI values. This modification, however, did not prevent the non-uniform distribution of errors across the PBI range when the results were back-transformed to raw PBI units.

Thus, in this manuscript we have successfully developed an Australia-wide model for the prediction of PBI in the 0–150 range, using a handheld spectrometer for soils that have been dried and sieved at 2 mm. It is acknowledged that not all soil types may have been included in this study. The handheld instrument capability will potentially enable paddock-scale mapping, which, in conjunction with yield maps, could be used for more efficient P-fertiliser application. The calibration model developed in this study can be used for the prediction of PBI in Australian soils from spectra scanned on similar handheld spectrometers, although calibration transfer procedures may first need to be applied. For accurate, routine applications of soils with PBI > 150, we may need to include more high-range PBI values in the calibration model. Further research on the effects of variable moisture contents and heterogeneity is needed to enable in-field sampling with a handheld MIR spectrometer.

Acknowledgements

References

Bainbridge SH, Miles N, Praan R, Johnston MA (1995) Phosphorus sorption in Natal soils. *South African Journal of Plant and Soil* **12**, 59–64. doi:10.1080/02571862.1995.10634338

Baldock JA, Hawke B, Sanderman J, MacDonald LM (2013) Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Research* **51**, 577–595. doi:10.1071/SR13077

Barnes RJ, Dhanoa MS, Lister SJ (1993) Correction of the description of Standard Normal Variate (SNV) and De-Trend Transformations in Practical spectroscopy with applications in food and beverage analysis, 2nd edn. *Journal of Near Infrared Spectroscopy* **1**, 185–186.

Barrow NJ (1978) The description of phosphate adsorption curves. *Journal of Soil Science* **29**, 447–462. doi:10.1111/j.1365-2389.1978.tb00794.x

Bertrand I, Holloway RE, Armstrong RD, McLaughlin MJ (2003) Chemical characteristics of phosphorus in alkaline soils from southern Australia. *Australian Journal of Soil Research* **41**, 61–76. doi:10.1071/SR02021

Bolland MD, Gilkes RJ, Brennan RF, Allen DG (1996) Comparison of seven phosphorus sorption indices. *Australian Journal of Soil Research* **34**, 81–89. doi:10.1071/SR9960081

Bomans E, Fransen K, Gobin A, Mertens J, Michiels P, Vandendriessche H, Vogels N (2005) Addressing phosphorus related problems in farm practice. Final Report to the European Commission. Available at: http://ec.europa.eu/environment/natres/pdf/phosphorus/AgriPhosphorus Report%20final.pdf

Brunet D, Barthès BG, Chotte JL, Feller C (2007) Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma* **139**, 106–117. doi:10.1016/j.geoderma.2007.01.007

Burkitt LL, Moody PW, Gourley CJP, Hannah MC (2002) A simple phosphorus buffering index for Australian soils. *Australian Journal of Soil Research* **40**, 497–513. doi:10.1071/SR01050

Burkitt LL, Gourley CJP, Hannah MC, Sale PWG (2006) Assessing alternative approaches to predicting soil phosphorus sorption. *Soil Use and Management* **22**, 325–333. doi:10.1111/j.1475-2743.2006.00042.x

Burkitt LL, Sale PWG, Gourley CJP (2008) Soil phosphorus buffering measures should not be adjusted for current phosphorus fertility. *Australian Journal of Soil Research* **46**, 676–685. doi:10.1071/SR06126

Burkitt LL, Dougherty WJ, Carlson SM, Donaghy DJ (2010) Effect of variable soil phosphorus on phosphorus concentrations in simulated surface runoff under intensive dairy pastures. *Australian Journal of Soil Research* **48**, 231–237. doi:10.1071/SR09025

Cohen MJ, Paris J, Clark MW (2007) P-sorption capacity estimation in southeastern USA wetland soils using visible/near infrared (VNIR) reflectance spectroscopy. *Wetlands* **27**, 1098–1111. doi:10.1672/0277-5212(2007)27[1098:PCEISU]2.0.CO;2

Colwell JD (1963) The estimation of the phosphorus fertiliser requirements of wheat in southern New South Wales by soil analysis. *Australian Journal of Experimental Agriculture and Animal Husbandry* **6**, 105–120.

Cox FR (1994) Predicting increases in extractable phosphorus from fertilizing soils of varying clay content. *Soil Science Society of America Journal* **58**, 1249–1253. doi:10.2136/sssaj1994.03615995005800040036x

Crawley MJ (2007) 'The R Book.' (John Wiley & Sons: Chichester, UK)

Forrester S, Janik L, Beech A, Merry R, McLaughlin M, Burkitt L, Gourley C (2003) Prediction of a range of phosphorus buffering techniques using mid-infrared spectroscopy. In 'Tools for nutrient and pollutant management: Applications to agriculture and environmental quality'. Occasional Report No. 17. (Eds LD Currie, JA Hanly) (Fertiliser and Lime Research Centre, Massey University: Palmerston North, New Zealand)

Forrester S, Janik L, McLaughlin M (2010a) Mid-infrared spectroscopic analysis of soils—from the laboratory to the field. In 'The environment—the future'. Australasian Soil & Plant Analysis Council Inc. (ASPAC) Conference, Canberra. (ASPAC: Carapook, Vic.)

Gourley CJP, Melland AM, Waller RA, Awty IM, Smith AP, Peverill KI, Hannah MC (2006) Making better fertiliser decisions for grazed pastures in Australia. Australian Soil Resource Information System. Available at: www.asris.csiro.au/themes/nutrient.html (accessed 12 August 2014)

Hart M, Cornish P (2012) Available soil phosphorus, phosphorus buffering and soil cover determine most variation in phosphorus concentration in

runoff from pastoral sites. *Nutrient Cycling in Agroecosystems* **93**, 227–244. doi:10.1007/s10705-012-9512-2

Isbell RF (1996) 'The Australian Soil Classification.' (CSIRO Publishing: Melbourne)

Janik LJ, Skjemstad JO (1995) Characterization and analysis of soils using mid-infrared partial least-squares. II. Correlations with some laboratory data. *Australian Journal of Soil Research* **33**, 637–650. doi:10.1071/SR9950637

Janik LJ, Merry RH, Skjemstad JO (1998) Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture* **38**, 681–696. doi:10.1071/EA97144

Janik LJ, Forrester ST, Rawson A (2009) The prediction of soil chemical and physical properties from mid-infrared spectroscopy and combined partial least-squares regression and neural networks (PLS-NN) analysis. *Chemometrics and Intelligent Laboratory Systems* **97**, 179–188. doi:10.1016/j.chemolab.2009.04.005

Kuang B, Mahmood HS, Quraishi MZ, Hoogmoed WB, Mouazen AM, van Henten EJ (2012) Sensing soil properties in the laboratory, *in situ*, and on-line. A review. *Advances in Agronomy* **114**, 155–223. doi:10.1016/B978-0-12-394275-3.00003-1

Kudeyarova AY, Davydkina LV, Kvartskheliya MZ, Korbacheva II (1991) Ability of phosphate to replace organic compounds in soil organomineral complexes. *USSR Academy of Sciences, Translated from Pochvovedeniye* **2**, 31–44.

Lins IDG, Cox FR (1989) Effect of extractant and selected soil properties on predicting the optimum phosphorus fertiliser rate for growing soybeans under field conditions. *Communications in Soil Science and Plant Analysis* **20**, 319–333. doi:10.1080/00103628909368085

Maindonald J, Braun J (2010) 'Data analysis and graphics using R: An example-based approach.' (Cambridge University Press: Cambridge, UK)

McKeague JA, Day JH (1966) Dithionite and oxalate extractable Fe and Al as aids in differentiating various classes of soils. *Canadian Journal of Soil Science* **46**, 13–22. doi:10.4141/cjss66-003

McLaughlin JR, Ryden JC, Syers JK (1981) Sorption of inorganic phosphate by iron- and aluminium- containing components. *Journal of Soil Science* **32**, 365–378. doi:10.1111/j.1365-2389.1981.tb01712.x

McLaughlin MJ, McBeath TM, Smernik R, Stacey SP, Ajiboye B, Guppy C (2011) The chemical nature of P accumulation in agricultural soils-implications for fertiliser management and design: An Australian perspective. *Plant and Soil* **349**, 69–87. doi:10.1007/s11104-011-0907-7

Minasny B, Tranter G, McBratney AB, Brough DM, Murphy BW (2009) Regional transferability of mid-infrared diffuse reflectance spectroscopic prediction for soil chemical properties. *Geoderma* **153**, 155–162. doi:10.1016/j.geoderma.2009.07.021

Moody PW (2007) Interpretation of a single-point P buffering index for adjusting critical levels of the Colwell soil P test. *Australian Journal of Soil Research* **45**, 55–62. doi:10.1071/SR06056

Moody PW, Speirs S, Scott BJ, Mason SD (2013) Soil phosphorus tests I: What soil phosphorus pools and processes do they measure? *Crop & Pasture Science* **64**, 461–468. doi:10.1071/CP13112

Nduwamungu C, Ziadi N, Tremblay GF, Parent LÉ, Tremblay GF, Thuriès B (2009) Near-infrared reflectance spectroscopy prediction of soil properties: Effects of sample cups and preparation. *Soil Science Society of America Journal* **73**, 1896–1903. doi:10.2136/sssaj2008.0213

Nguyen TT, Janik LJ, Raupach M (1991) Diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy in soil studies. *Australian Journal of Soil Research* **29**, 49–67. doi:10.1071/SR9910049

Ozanne PG (1980) Phosphate nutrition of plants—A general treatise. In 'The role of phosphorus in agriculture'. pp. 559–589. (American Society of Agronomy/Crop Science Society of America/Soil Science Society of America: Madison, WI, USA)

Reeves JB III (2010) Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? *Geoderma* **158**, 3–14. doi:10.1016/j.geoderma.2009.04.005

Reeves JB III, McCarty GW, Reeves VB (2001) Mid-infrared diffuse reflectance spectroscopy for the quantitative analysis of agricultural soils. *Journal of Agricultural and Food Chemistry* **49**, 766–772. doi:10.1021/jf0011283

Reeves JB, McCarty GW, Hively WD (2010) Mid- versus near-infrared spectroscopy for on-site analysis of soil. In 'Proximal soil sensing'. (Eds RA Viscarra-Rossel, AB McBratney, B Minasny) pp. 133–142. (Springer Science+Business Media: New York)

Ryan J, Curtin D, Cheema MA (1984) Significance of iron oxides and calcium carbonate particle size in phosphate sorption by calcareous soils. *Soil Science Society of America Journal* **48**, 74–76.

Samadi A, Gilkes RJ (1998) Forms of phosphorus in virgin and fertilized calcareous soils of Western Australia. *Australian Journal of Soil Research* **36**, 585–601. doi:10.1071/S97060

Scheinost AC, Schwertmann U (1995) Predicting phosphate adsorption–desorption in a soilscape. *Soil Science Society of America Journal* **59**, 1575–1580. doi:10.2136/sssaj1995.03615995005900060010x

Sibanda HM, Young SD (1986) Competitive adsorption of humus acids and phosphate on goethite, gibbsite and two tropical soils. *Journal of Soil Science* **37**, 197–204. doi:10.1111/j.1365-2389.1986.tb00020.x

Smith RJ (1993) Logarithmic transformation bias in allometry. *American Journal of Physical Anthropology* **90**, 215–228. doi:10.1002/ajpa.1330900208

Soriano-Disla JM, Janik L, Viscarra Rossel RA, McDonald LM, McLaughlin MJ (2014) The performance of visible, near and mid-infrared spectroscopy for prediction of soil physical, chemical and biological properties. *Applied Spectroscopy Reviews* **49**, 139–186. doi:10.1080/05704928.2013.811081

Speirs SD, Scott BJ, Moody PW, Mason SD (2013) Soil phosphorus tests II: A comparison of soil test-crop response relationships for different soil tests and wheat. *Crop & Pasture Science* **64**, 469–479. doi:10.1071/CP13111

Stenberg B, Viscarra-Rossel RA (2010) Diffuse reflectance spectroscopy for high-resolution soil sensing. In 'Proximal soil sensing. Progress in soil science 1'. (Ed. RA Viscarra-Rossel) pp. 29–47. (Springer Science +Business Media: New York)

Stenberg B, Viscarra-Rossel RA, Mouazen AM, Wetterling J (2010) Visible and near-infrared spectroscopy in soil science. *Advances in Agronomy* **107**, 163–215. doi:10.1016/S0065-2113(10)07005-7

Stumpe B, Weihermüller L, Marschner B (2011) Sample preparation and selection for qualitative and quantitative analyses of soil organic carbon with mid-infrared reflectance spectroscopy. *European Journal of Soil Science* **62**, 849–862. doi:10.1111/j.1365-2389.2011.01401.x

Sudduth KA, Hummel JW (1996) Geographic operating range evaluation of a NIR soil sensor. *Transactions of the American Society of Agricultural Engineers* **39**, 1599–1604. doi:10.13031/2013.27674

Uehara G, Gillman G (1981) 'The mineralogy, chemistry, and physics of tropical soils with variable charge clays.' pp. 31–95. (Westview Press: Boulder, CO, USA)

Van der Marel HW, Beutelspacher H (1976) Clay and related minerals. In 'Atlas of infrared spectroscopy of clay minerals and their admixtures'.

(Eds HW Van der Marel, H Beutelspacher) (Elsevier Scientific: Amsterdam, The Netherlands)

Varmuza K, Filmoser P (2009) 'Introduction to multivariate statistical analysis in chemometrics.' (Taylor & Francis Group: Boca Raton, FL, USA)

Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO (2006) Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **131**, 59–75. doi:10.1016/j.geoderma.2005.03.007

Whitehead DC (2000) Phosphorus. In 'Nutrient elements in grassland: Soil-plant-animal relationships'. (Ed. DC Whitehead) pp. 126–153. (CAB International: Wallingford, UK)

Williams PC (1987) Variables affecting near-infrared reflectance spectroscopy. In 'Near-infrared technology in the agricultural and food industries'. (Eds PC Williams, KH Norris) pp. 143–167. (American Association of Cereal Chemist Inc.: St Paul, MN, USA)