Author Query Form

Journal: EJSS

Article: 13202

Dear Author,

During the copyediting of your manuscript the following queries arose.

Please refer to the query reference callout numbers in the page proofs and respond.

Please remember illegible or unclear comments and corrections may delay publication.

Many thanks for your assistance.

AUTHOR: Please note that missing content in references have been updated where we have been able to match the missing elements without ambiguity against a standard citation database, to meet the reference style requirements of the journal. It is your responsibility to check and ensure that all listed references are complete and accurate.

Query reference	**Query**	**Remarks**
Q1	Please resupply the first names in expanded form to comply with the journal style.	
Q2	Please verify that the linked ORCID identifiers are correct for each author. The ORCID ID for 'Raphael A. Viscarra Rossel' seems to be invalid. Please check and supply the correct ORCID ID.	
Q3	Please confirm that given names (blue) and surnames/family names (vermilion) have been identified correctly.	
Q4	Please check that authors and their affiliations are correct.	
Q5	Please check that the correct author has been identified as the contact for correspondence.	
Q6	Journal style requires that the first forenames of all authors. Please provide missing details.	
Q7	Please provide the "volume number, page range" for reference Greenberg, Seidel, Vohland & Ludwig, 2021.	
Q8	Please provide the "volume number" for reference van Leeuwen, Mulder, Batjes & Heuvelink, 2021.	

Journal Code: EJSS | Article ID: 13202 | Dispatch: 18-DEC-21 | CE:
No. of Pages: 17 | ME:

Received: 24 June 2021 | Revised: 17 November 2021 | Accepted: 1 December 2021

DOI: 10.1111/ejss.13202

ORIGINAL ARTICLE

European Journal of Soil Science WILEY

The cost-effectiveness of reflectance spectroscopy for estimating soil organic carbon

S. Li[1] | **R. A. Viscarra Rossel**[2] | **R. Webster**[3]

[1]Key Laboratory for Geographical Process Analysis & Simulation of Hubei Province, College of Urban & Environmental Sciences, Central China Normal University, Wuhan, China

[2]School of Molecular & Life Sciences, Faculty of Science & Engineering, Curtin University, Perth, Australia

[3]Rothamsted Research, Harpenden, UK

Correspondence
R. A. Viscarra Rossel, School of Molecular & Life Sciences, Faculty of Science & Engineering, Curtin University, GPO Box U1987, Perth WA 6845, Australia.
Email: r.viscarra-rossel@curtin.edu.au

Abstract

Reflectance spectra of soil can be used to estimate the concentrations of organic carbon in soil (SOC). The estimates are more or less imprecise, but spectroscopy is quicker, less laborious and cheaper than conventional dry combustion analysis. Are the greater economy and efficiency sufficient to justify the loss of information arising from errors in estimation? We measured soil spectra with three instruments: a bench-top mid-infrared (mid-IR) (mid-$\mathrm{IR_b}$) spectrometer, a portable mid-IR (mid-$\mathrm{IR_p}$) spectrometer and a portable visible–near infrared (vis–$\mathrm{NIR_p}$) spectrometer. We calculated a quantity E to express the cost-effectiveness of spectroscopic estimates relative to the conventional analysis, by accounting for their inaccuracy, their cost and their capacity, namely the maximum number of samples that can be prepared and measured daily. In all, 562 samples of soil were collected from 150 locations at four depths on a farm. The samples were dried and ground to particle sizes of $\leq$2 and $\leq$0.5 mm before measurements were made by dry-combustion analysis. The machine learning algorithm Cubist was used to derive spectroscopic models of SOC concentrations and their uncertainties. We found that the mid-$\mathrm{IR_b}$ on the $\leq$0.5 mm samples was the most accurate and expensive but nevertheless sufficiently cost-effective (large value of E) for determining the organic C. The mid-$\mathrm{IR_p}$ was somewhat more accurate, but its E was smaller than vis–$\mathrm{NIR_p}$ on corresponding samples because it required more time to record the spectra. We also found that, with the portable spectrometers, the SOC predictions made on the $\leq$0.5 mm samples were somewhat more accurate than those made on the $\leq$2 mm samples, but their E was smaller because of the additional cost of sample preparation. The vis–$\mathrm{NIR_p}$ on the $\leq$2 mm samples was the most cost-effective for estimating SOC because it is cheap, accurate and has a large capacity for measurements.

Highlights

- Concentrations of soil organic carbon (SOC) were determined by standard dry combustion and estimated from reflectance spectra recorded by three instruments.
- The labour required for each of the techniques and the cost, including that of the equipment, were recorded.

- A quantity *E*, expressing the cost-effectiveness relative to dry combustion was calculated for each spectral technique, taking into account both accuracy and cost.
- Dry combustion was always more accurate than estimates from spectra for individual samples, and the technique was also more cost-effective for small numbers of samples.
- The cost-effectiveness of the spectral techniques varied among themselves, but all were more cost-effective than dry combustion for large numbers of samples.

KEYWORDS

accuracy, capacity, cost, economy, mid-infrared spectroscopy, soil organic carbon, visible–near-infrared spectroscopy

1 | INTRODUCTION

Scientists have for many years measured the concentration of organic carbon in the soil (SOC). They have wanted to understand the role of carbon (C) in the soil's behaviour, especially as the result of changes under cultivation. In recent years emphasis has switched towards the carbon cycle in ecosystem processes (Davidson & Janssens, 2006; Fontaine et al., 2007; Schmidt et al., 2011; Stell et al., 2021) and for carbon accounting (Atwood et al., 2017; Mishra et al., 2021; Viscarra Rossel et al., 2014). Scientists have taken soil material from the field into the laboratory, dried it, crushed and sieved it, and then analysed sub-samples of it chemically. The methods are time-consuming, laborious, and expensive, and they can be unaffordable when many samples are needed, for example, to monitor SOC on farms and larger regions such as catchments and nations. Modern spectroscopy in the near- and mid-infrared regions of the electromagnetic spectrum offers an alternative means of analysing soil quickly and cheaply, and there is growing interest in it for estimating SOC (Paustian et al., 2016).

The spectra contain a great deal of information, not only on the organic C but also on soil minerals, particle size and water (if measured under field conditions). The relevant information for estimating the SOC has, therefore, to be teased out. This requires some fairly complex statistical processing and modelling, but once the spectra have been recorded digitally they too can be rapid. Questions then arise concerning the accuracy of estimates from the technique.

There is relevant spectral information on C in the mid-infrared (mid-IR, 2500–25,000 nm) and visible–near-infrared (vis–NIR, 350–2500 nm) regions (Reeves et al., 2006; Viscarra Rossel & Hicks, 2015), and there are now modestly priced spectrometers for sensing and recording reflectance in these regions. Quantification of SOC by spectroscopy in the vis–NIR and mid-IR requires reference soil samples with both analytical data and spectra from which to derive a spectroscopic model. Despite the successful estimation of SOC by laboratory-based spectroscopy (for both bench-top and portable instruments) (Soriano-Disla et al., 2014), the preparation of samples (drying and grinding) can still be time-consuming.

Measurements with portable vis–NIR (vis–NIR$_p$) and mid-IR (mid-IR$_p$) spectrometers can be made in the laboratory on air-dried and ground soil samples and on wet soil under field conditions (Dhawale et al., 2015; Greenberg et al., 2021; Hutengs et al., 2019; Ji et al., 2016; Li et al., 2015). Hand-held vis–NIR$_p$ and mid-IR$_p$ spectrometers are available from several manufacturers, and they are becoming increasingly affordable. Their advantage over bench-top instruments is that they are small, portable, can be deployed in the field, and can be incorporated into multi-sensor systems (e.g., Viscarra Rossel et al., 2017): there is thus growing motivation for their further development.

It has been shown that spectroscopy in the mid-IR provides 10%–40% more accurate estimates of organic C than that in the vis–NIR (Baumann et al., 2021; Bellon-Maurel & McBratney, 2011; Viscarra Rossel et al., 2006). With few exceptions, however, the use of portable mid-IR spectrometers has not been sufficiently explored (Hutengs et al., 2018) because until recently, few mid-IR$_p$ spectrometers have been commercially available. Also, the preparation of samples strongly affects the estimates of soil properties from mid-IR spectra. Soil for mid-IR spectroscopy is commonly ground more finely than that for vis–NIR, to between 0.1 and 0.5 mm (Janik et al., 1998; Le Guillou et al., 2015). For the latter, the soil is usually ground to ≤2 mm, though measurements can be made on unprepared samples. We need to be sure that estimates of SOC from spectra recorded by mid-IR$_p$ spectrometers on soil ground only to ≤2 mm are sufficiently accurate before such instruments are used routinely.

Estimates of SOC must be accurate, of course. But what do we mean by accuracy, or, more importantly in the current context, inaccuracy? van Leeuwen et al. (2021) address the question in some detail in this journal, and they illustrate their answer with a statistical analysis of some data on organic C in soil. They, and we, treat 'inaccuracy' in general as the departure of an estimate or measurement from a true value. This departure comprises systematic deviation from a true value, that is, bias, plus apparently random fluctuation about the true value or its estimate. Many statisticians distinguish by restricting the word 'accuracy' to refer to bias and its avoidance and use the word 'precision' for the random component. Others use the word to embrace both, and we do that following the usage of Kempthorne and Allmaras (1965).

We take the dry combustion method as a standard. We have assumed that the errors in the measurements by dry combustion provided by the accredited laboratories (NATA and ASPAC, listed below), were small enough to negligible, though recognising that such errors can be substantial (van Leeuwen et al., 2021). Inaccuracies arise almost entirely from variation among replicate samples from the field or analyses on replicate sub-samples in the laboratory or both. We can determine their contributions from the replicate measurements, and we can diminish the error, that is, increase precision, by increased replication—at a price in terms of labour and cost.

Variation from sample to sample contributes to an error in the measurement of the spectra. Again, one can increase precision by increasing the replication, but now at substantially less cost. Nonetheless, such variations contribute to errors in spectral estimates of SOC. Additionally, its estimates contain systematic errors, that is biases, arising from calibration and modelling, and these must be taken into account when one judges the merits of spectroscopy. Thus, estimates of SOC from spectroscopy are less accurate than those from dry combustion, but by how much? Given that spectroscopy is much cheaper than dry combustion, can we put a price on its inaccuracy?

We can certainly document the costs. We can record the times taken to obtain measurements by the various techniques; we can convert them into monetary units knowing the wages of the labour by trained technicians. We know the prices of instruments and other equipment and their running costs, and we can take into account their depreciation in time. In other words, we can know how much it costs to obtain data. If those data are inaccurate, however, there is an additional hidden cost that we should take into account.

Though previous investigators have discussed the cost (O'Rourke & Holden, 2011), or cost–benefit (Nocita et al., 2015), of soil spectroscopy, they seem not to have done so quantitatively, or set the cost of loss of information against the operational savings. Can we do that, or alternatively, can we increase the sampling for spectroscopy to counteract the inaccuracy of individual measurements? In the simplest situations the cost of obtaining a mean value of SOC is directly related to the size of the sample, say n; that is, the cost is a linear function of n, and the error variance is inversely related to n (Cochran, 1977). One can then define the effectiveness of a procedure in terms of its cost relative to some standard. The situation we investigate is far more complex, yet we still want a measure of relative effectiveness. We designate a quantity E, the cost-effectiveness of spectroscopy compared to conventional laboratory analysis. This quantity combines time, cost and inaccuracy, and it is intended to reveal the most economical means of obtaining sufficient accurate data for a particular application over a region of interest; the larger it is the better. It would enable an investigator to choose a technique that gives the best value within a limited budget or to estimate the resources required to achieve some given goal of accuracy. This is the background to the study we describe below.

Our aims here were as follows:

1. To compare spectra recorded with three spectrometers: a bench-top mid-IR spectrometer (mid-IR$_b$), a portable mid-IR$_p$ spectrometer and a portable vis–NIR$_p$ spectrometer, all capturing spectra from soil ground to ≤2 and ≤0.5 mm;
2. To compare spectroscopic predictions of SOC made with the spectra from the three spectrometers on the two-particle sizes;
3. To compare the cost-effectiveness, E, of the spectroscopic predictions of SOC relative to measurements made by dry combustion as standard, after quantifying their inaccuracy and costs.

2 | MATERIALS AND METHODS

2.1 | Soil sampling

We took soil for this study from a 600-ha cattle farm in Northern New South Wales, Australia (30.69°S, 151.48°E). It is predominantly a Kurosol in the Australian Soil Classification (Isbell, 2002) approximately equivalent to Acrisols and Planosols in the World Reference Base (WRB) for Soil Resources (IUSS Working Group WRB, 2006).

We sampled to a stratified random design to obtain 150 soil cores to a depth of 1 m, or to the limit of penetration if that was less. The cores were 50 mm in diameter and were retained in PVC tubes for storage. Viscarra Rossel et al. (2017) have recorded the details of the soil and its variation over the farm.

The word 'sample' is used in two ways by soil scientists. One is the strictly statistical meaning of a set of units, say cores as above, drawn from a population; the other, and more usual, is a mass of soil material on which measurements are made. In this paper, we use the word in this second sense unless stated otherwise.

2.2 | Sample preparation and laboratory analyses

Once in the laboratory, all the cores were air-dried, sub-sampled and stored before measurements were made. The soil cores were divided into sections at depths of 0–5.5 cm, 13–17 cm, 28–32 cm and 58–62 cm to produce a total of 562 sub-samples for the experiments. The samples were dried in a drying room at 40°C for approximately 3 days. Portions were then crushed and ground to particle sizes ≤2 and ≤0.5 mm. Some of each portion was analysed for SOC by the DC elemental method (Rayment & Lyons, 2011) with a LECO analyser (Laboratory Equipment Corporation, St Joseph, MI, USA). We sent the prepared samples to a National Association of Testing Authorities (NATA) and Australasian Soil and Plant Analysis Council (ASPAC) accredited environmental analysis laboratory for the measurement of SOC. The analytical accuracy of these laboratories has been independently assessed by "blind" duplicate analysis.

2.3 | Spectroscopic measurements

Figure 1 shows three spectrometers used in this study, the corresponding spectra in absorbance were also given underneath. The mid-IR$_b$ spectra of the soil samples were recorded with a Vertex 70 bench-top spectrometer (Bruker Optik GmbH, Ettlingen, Germany) with liquid nitrogen-cooled mercury cadmium telluride (MCT) detector. The spectral range was 7500–600 cm^{-1} (1333–16,667 nm). Samples with particle size ≤0.5 mm were placed into a 48-well microplate with a well diameter 10 mm. A gold standard was placed on the first well. The robotic arm (Twister Microplate Handler, also from Bruker Optik GmbH) was used for the automated loading of sample plates into the spectrometer. Before recording the soil spectra, once every plate, a background measurement was made with the gold standard. We used a spectral resolution of 4 cm^{-1} and took 64 measurements per sample. Soil samples were filled in the next 44 wells (with the last three wells empty) for each plate. To minimise errors in spectroscopic measurement to obtain a robust estimate, four sub-samples of each soil sample were measured, and the four replicate spectra were then averaged into one spectrum for each sample.

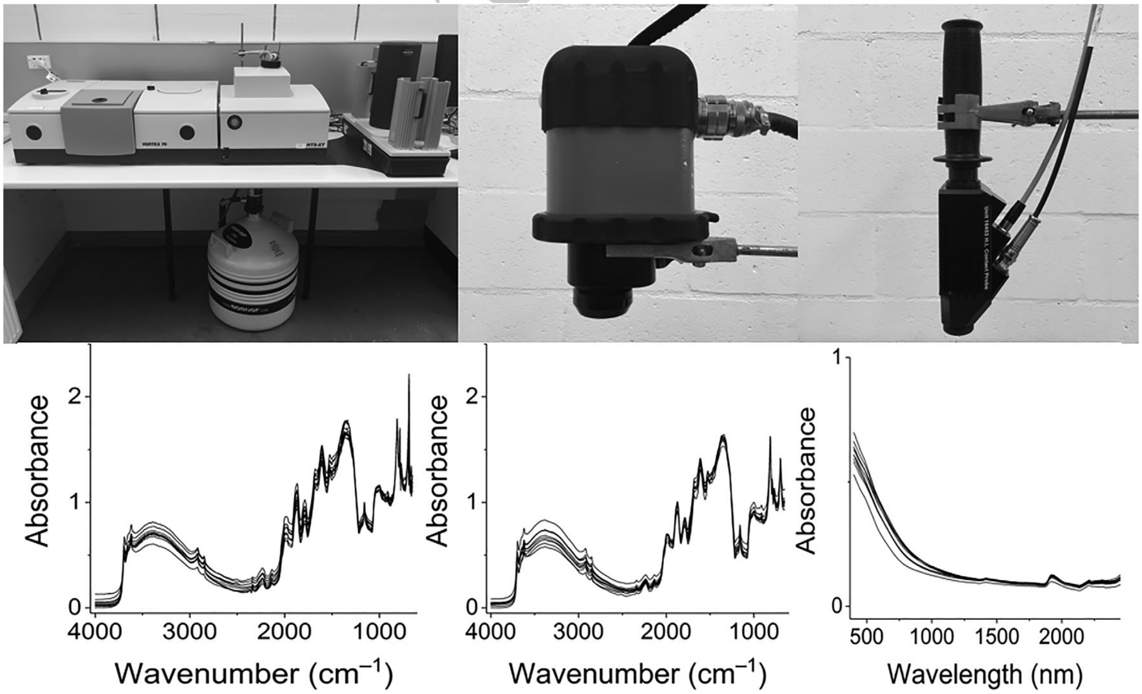


FIGURE 1 Three spectrometers with their corresponded spectra. From the left to the right: Mid-IR$_b$, mid-IR$_p$, and Vis–NIR$_p$

The mid-IR$_p$ spectra were recorded manually with a hand-held Agilent 4200 spectrometer with a deuterated triglycine sulfate (DTGS) detector cooled by a thermo-electric cooling module (Agilent Technologies, Santa Clara, CA). The instrument measures diffuse reflectance over the range 4000–650 cm^{-1} (2500–15,385 nm) with a spectral resolution of 4 cm^{-1}. Samples ground to ≤2 and ≤0.5 mm were placed in a petri dish (5 cm diameter, 1 cm depth). We recorded a gold reference background spectrum once every 10 measurements. Each recorded spectrum was the product of 64 internal measurements.

The vis–NIR$_p$ spectra were measured manually with a portable Labspec® spectrometer (PANalytical Inc., formerly Analytical Spectral Devices–ASD, Boulder, CO) with a high-intensity contact probe attached to an external fibre-optic cable. The contact probe measures a spot of diameter 10 mm, and is designed to minimise errors associated with stray light. The spectra were recorded at a resolution of 1 nm from 350 to 2500 nm. Samples ground to ≤2 and ≤0.5 mm were placed in a petri dish (5 cm diameter, 1 cm depth). Standardised measurements were made with the protocols described by Viscarra Rossel, Behrens, et al. (2016). Briefly, we calibrated the sensor with a Spectralon® white reference once every 10 measurements. Thirty spectra were averaged into one spectrum for each soil sample.

2.4 | Pre-processing of spectra

The spectral ranges from both mid-IR$_b$ and mid-IR$_p$ were reduced to span the 4000–650 cm^{-1} range. The vis–NIR$_p$ spectra were reduced to 400–2450 nm to eliminate noise at the ends of the spectra. Because the spectra are highly collinear, the vis–NIR spectra were sampled to a resolution of 10 nm. Both mid-IR and vis–NIR reflectances (R) were later transformed to the logarithmic apparent absorbance, $A = log_{10}(1/R)$.

To reduce noise further and to enhance the signals, we used Savitzky–Golay smoothing combined with a first-derivative treatment (Savitzky & Golay, 1964) and then centred the resulting spectra before modelling them. We then offset the spectra to a baseline for qualitative description by subtracting the absorbance of the first wavenumber or wavelength.

To summarise the spectra and examine their structure, we decomposed the spectra using principal components analysis (PCA) with the iterative non-linear partial least squares (NIPALS) algorithm on the covariance matrix of the spectral data (Martens & Næs, 1989). This enabled us to condense the information contained in the samples to a few scores and to identify those variables that load heavily on the leading principal components.

We projected the spectra from the mid-IR$_p$ (≤2, ≤0.5 mm particle sizes) onto the space of the mid-IR$_b$, and the spectra from vis–NIR$_p$ (≤2 mm particle size) onto that of ≤0.5 mm particle size.

2.5 | Modelling and assessment

As above, we had 562 samples of soil from 150 cores available for our analysis. To validate our models so that they could be used with confidence, we randomly selected 30 cores (96 samples) to represent an independent test set of data. From the remaining 120 cores, we selected a training set of 82 cores (320 samples) using the Kennard–Stone (KS) algorithm (Kennard & Stone, 1969). Because of the correlations of the measurements from the same core, we selected cores rather than samples by defining one core as one subgroup in the KS algorithm. The remaining 38 cores (146 samples) were used for validation. Table 1 lists the basic statistics of the calibration, validation and test sets of data.

We used the machine-learning algorithm Cubist (Quinlan, 1992) to build the spectroscopic models for SOC. The algorithm divides the variable space into smaller regions and creates a multivariate linear least-squares model for each of the partitions, and it divides the response data into subsets in which their characteristics are similar to their spectra and other predictors that might be used. It is the same approach as Viscarra Rossel and Webster (2012) described and used to model soil spectra.

We took 50 bootstrap samples of the calibration data when we implemented Cubist to ensure that the model was stable and to assess the uncertainty of our predictions. The method relies on repeated random sampling with replacement from the original units of calibration datasets to obtain 50 bootstrap samples. Each bootstrap sample is the same size as the initial set of calibration data, though it contains replicates of some units, while others are absent and represent the out-of-bag (OOB) samples (Hastie et al., 2005). The bootstrap estimates provide robust predictions because the final predictions are calculated from the mean of the 50 bootstraps. Briefly, these improvements result from the aggregation of the 50 different bootstrapped models, each of which provides unique information (Viscarra Rossel, 2007).

To express the uncertainty of the estimation of the independent test set of data we derived the upper and lower 95% confidence intervals by calculating,

$$y = \bar{y} \pm 1.96 \times \sqrt{\frac{1}{N}\sum_{i=1}^{N} s_i^2}\,, \qquad (1)$$

TABLE 1 Statistics of SOC concentrations (in %) measured by dry combustion in the laboratory

	1st				3rd			
Data set	***N***	**Min.**	**Quartile**	**Median**	**Mean**	**Quartile**	**Max.**	**Skew**
Calibration data	320	0.08	0.22	0.40	0.79	0.91	3.30	1.40
Validation data	146	0.08	0.19	0.35	0.70	0.83	2.66	1.26
Independent test data	96	0.08	0.22	0.43	0.80	0.85	3.20	1.44

Abbreviations: Max, maximum; Min, minimum; *N*, sample size; Skew, skewness coefficient.

where $\bar{y}$ is the average value of predicted SOC concentrations at a particular depth, s_i^2 is the variance of 50 bootstraps for sample i, and N is the number of soil samples.

We used the coefficient of determination (R^2) to evaluate and compare the performances of the spectroscopic models, which is simply the square of the Pearson correlation coefficient between the predictions and the measurements, the root mean square error (RMSE) to assess the inaccuracy of the estimations, the mean error (ME) to evaluate the bias, and the ratio of performance to interquartile distance (RPIQ) to assess the quality of the spectroscopic model (Bellon-Maurel et al., 2010), defined as the ratio of the interquartile ($\text{IQ} = \text{Q3} - \text{Q1}$) to the RMSE. The mathematical pre-processing and chemometric analyses were done in R (R Core Team, 2020). For the DC method, the inaccuracy was calculated by RMSE from the results of duplicated soil samples.

2.6 | The cost effectiveness, *E*, of spectroscopy

We calculated the quantity E to account for the inaccuracy of the spectroscopic SOC predictions from the three spectrometers and their cost, relative to the measurements made with dry combustion (DC) analyser. We assumed that all measurements began with the work done for the sample preparation (drying and grinding). We defined the capacity of each method as the maximum number of samples that could be ground and measured by one analyst using a particular method during a working day, which is 7 h. We assumed that as the number of samples increases, E approaches a maximum determined by the capacity of the procedure to acquire data. Thus, we calculated,

$$E = \frac{U_\mathrm{r} \times C_\mathrm{r}}{U_\mathrm{s} \times C_\mathrm{s}} \times n^{\log_{\alpha_\mathrm{s}}\left(\frac{\alpha_\mathrm{s} \times \beta_\mathrm{s}}{\alpha_\mathrm{r} \times \beta_\mathrm{r}}\right)} \quad \text{for} \quad 1 \leq n \leq \alpha_\mathrm{s}\,, \tag{2}$$

where n is the number of samples. The quantity α is the capacity of the laboratory to prepare samples for analysis or spectroscopy; it is the maximum number of samples that can be ground and prepared daily. The quantity β is the capacity of data acquisition; it is the maximum number of soil samples that can be measured by the method. These two quantities are the most important in determining the cost-effectiveness of the techniques. The subscript r refers to the DC method, and subscript s refers to the particular spectrometer. The quantities U_r and U_s are the mean square errors (MSEs) of the measurements (in %); they express both the assumed random fluctuation and systematic error about the mean estimates. The quantities C_r and C_s are the total costs, in AU \$, of the measurements, see Equations (3)–(8) below. These equations also tell how the time taken was transferred into cost and capacity.

We calculated the total time taken, T, and total cost, C, as,

$$\begin{aligned} T &= T_\mathrm{sp} + T_\mathrm{da} \\ C &= C_\mathrm{sp} + C_\mathrm{da}\,, \end{aligned} \tag{3}$$

where T_sp and T_da are the times for preparation of the samples and for acquiring the data C_sp and C_da are the costs.

The C_sp was calculated as,

$$C_\mathrm{sp} = \frac{7 \times \nu}{\alpha}\,, \tag{4}$$

where ν is the hourly salary (AU \$45) and there are 7 h in the working day. The quantity α is the capacity of sample preparation; it is the maximum number of samples that can be prepared in 1 day and is given by,

$$\alpha = \frac{7 \times 3600}{T_\mathrm{sp}}\,, \tag{5}$$

The value 3600 converts hours to seconds, and 7 is again the number of working hours in a day. The quantity T_sp is the time taken for crushing and grinding samples to ≤2 or to ≤0.5 mm. In practice, operators need to take substantial breaks from such work, and the effective working day for crushing and grinding the samples is only about 60% of the 7 h (Collewet & Sauermann, 2017).

The time and cost for data acquisition were calculated as,

$$\begin{aligned} T_{\mathrm{da}} &= T_{\mathrm{ld}} + T_{\mathrm{rf}} + T_{\mathrm{mm}} \\ \text{and} \quad C_{\mathrm{da}} &= C_{\mathrm{mo}} + C_{\mathrm{mi}}\ , \end{aligned} \tag{6}$$

where T_{ld} and C_{ld} are the cost and time to load sample from containers into plates or Petri dishes after samples have been mixed thoroughly in the container. The portable sensors required extra cleaning between the measurements to avoid contamination from one sample to another. The quantity T_{rf} is the time spent for measuring of the reference material; T_{mm} is the time for measurement, C_{mo} is the cost of measurement by the operator; and C_{mi} is the cost of measurement by the selected instrument.

Specifically, we calculated each part of C_{da} by,

$$\begin{aligned} C_{\mathrm{mo}} &= \frac{7 \times \nu}{\beta} \\ \text{and} \quad C_{\mathrm{mi}} &= \frac{1.1 \times \Omega}{240 \times \beta}\ , \end{aligned} \tag{7}$$

where β is the capacity of data acquisition, that is, the maximum number of samples that can be measured within one working day. The quantity Ω is the estimated one-off capital cost of the equipment, which we assume to have a life of 10 years and to depreciate at the rate of 10% per annum. The 240 is the number of working days in 1 year.

With a DC analyser, one operator can measure 50 samples per day on average. So $\beta = 50$. As mentioned in Section 2.3, we have 132 or 220 for mid-IR$_{\mathrm{b}}$ with Bruker Vertex 70 spectrometer as its 60% and 100% β, respectively. The β of each portable spectrometer was calculated as,

$$\beta = \frac{T_{\mathrm{w}} - 1800}{T_{\mathrm{da}}}\ , \tag{8}$$

where T_{w} is the length of a working day in seconds (25,200 s). The 1800 s is the half-hour warm-up time for before the spectrometer can begin scanning.

To measure the difference in the capacity of the two methods with increasing n, Equation (2), we introduce the partition coefficient (Kwon, 2001) as the power. A value of $E > 1$ indicates that the spectroscopic method is more effective than DC. Conversely, $E < 1$ suggests that the spectroscopic method is less so.

The quantities α and β are set for a working day of 7 h with a single break (for lunch, for example). In practice, many technicians find such continuous, repetitive work tedious and require more frequent breaks. Therefore, we used 60% of the working day as a conservative estimate of the effective time spent to work on one or both of those tasks (Collewet & Sauermann, 2017). We, therefore, define four levels of work as follows:

Normal (60% α with 60% β),
Medium (100% α with 60% β),
High (60% α with 100% β), and
Extreme (100% α with 100% β).

3 | RESULTS

3.1 | Soil spectroscopy

Figure 2 shows, for each spectrometer, the average logarithmic apparent absorbance of all the 562 ground samples for each of the four layers of soil and twice their standard deviations, representing the variation of the soil. The vis–NIR spectra contain fewer absorption bands than the mid-IR spectra; the latter effectively contain more information on the composition of the soil.

For all three spectrometers (mid-IR$_{\mathrm{b}}$, mid-IR$_{\mathrm{p}}$, and vis–NIR$_{\mathrm{p}}$), the magnitude of the spectral variation in the two deeper layers was generally larger than that of in the top two layers. The average absorbance spectra in the mid-IR$_{\mathrm{p}}$ were less for the ≤2 mm particles than for the ≤0.5 mm particles. This might be because the larger particles reflect more radiation from their surfaces than do the smaller particles.

The spectra from the same ≤0.5 mm particle size, both mid-IR$_{\mathrm{b}}$ and mid-IR$_{\mathrm{p}}$ were similar; both had sharp absorbances in the 4000–2000 cm^{-1} region. In the range between 2000 and 650 cm^{-1}, however, the spectra from the mid-IR$_{\mathrm{b}}$ have more pronounced peaks than those of the mid-IR$_{\mathrm{p}}$. The spectra of the mid-IR$_{\mathrm{p}}$ were more varied in each layer (shown by the shaded regions in Figure 2). This variation was smaller for measurements made on the larger ≤2 mm fraction. The opposite is true for the vis–NIR$_{\mathrm{p}}$ spectra; that is, there was more variation in the spectra of the larger ≤2 mm fraction.

As the concentrations of SOC gradually decrease with increasing depth, the mid-IR absorptions near 2920 and 2850 cm^{-1} associated with stretching —CH$_2$ vibrations in organic matter gradually disappear (Figure 2). Similarly, there were other absorptions around 1400 cm^{-1} that are attributed to organic compounds that were less pronounced in the deeper layers (Capriel et al., 1995). As the clay content of the soil samples increased with increasing depth, the absorptions of the vis–NIR spectra that are

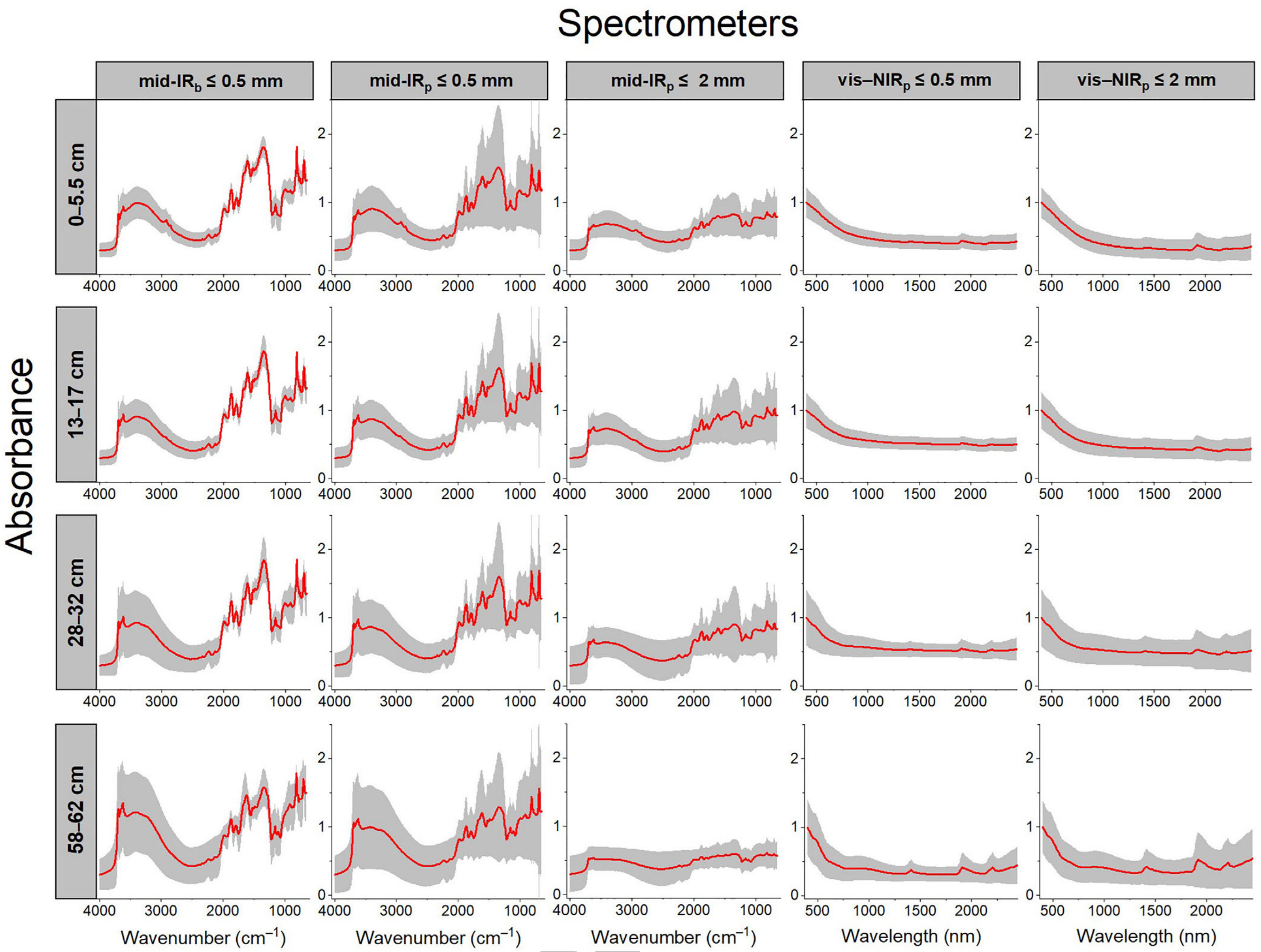


FIGURE 2 The average mid-IR$_b$, mid-IR$_p$ and Vis–NIR$_p$ spectra of four depths (0–5.5 cm, 13–17 cm, 28–32 cm, and 58–62 cm) for the two particle sizes, ≤0.5 and ≤2 mm. The absorbance spectra are shown as red curves with their twice standard deviations show as shaded grey bands about the means. The standard deviation here shows the variation of the samples

associated with iron oxides in the range 400–1000 nm and water near 1400 and 1900 nm and clay minerals such as kaolinite near 2200 nm.

Figure 3 shows the scatter of the scores of the first two principle components plotted against one another for the spectra from the mid-IR$_b$, mid-IR$_p$ (Figure 3a) and vis–NIR$_p$ (Figure 3b) for the two-particle fractions. The two leading components accounted for almost 90% of the total variance of the mid-IR spectra. All three types of spectra are clustered, though with some overlap. As described above (Figure 2), the spectra of the mid-IR$_p$ ≤0.5 mm are more varied and thus have a wider spread in the vector space of the first two components.

For the vis–NIR$_p$ spectra, the first two principal components accounted for 94.3% of the variance in the absorbance spectra. Although the scores of the ≤2 and ≤0.5 mm overlap, those of the ≤2 mm are more variable (see Figure 2) and thus have a wider spread (Figure 3b).

3.2 | Spectroscopic model evaluation and estimation

Figure 4 shows the estimated values in the independent test set of data and their uncertainties. The 95% lower and upper confidence limits (CLs) of each statistic are also given in brackets. The thick red lines down the profiles represent the means of the observations, and black dashed lines depict the predicted concentrations of organic C.

The mid-IR$_b$ (Figure 4a) produced the best estimates of SOC ($R^2 = 0.96$, RMSE $= 0.17$, RPIQ $= 3.70$) with the narrowest 95% confidence limits. For both mid-IR$_p$ and vis–NIR$_p$ instruments, the >2% SOC concentrations tended to be under-estimated. Estimates of SOC from these portable spectrometers were more accurate for particles ≤0.5 mm than for the larger fraction, ≤2 mm.

The 95% confidence interval for mid-IR$_p$ on ≤2 mm, Figure 4c, was narrower than that of the ≤0.5 mm,

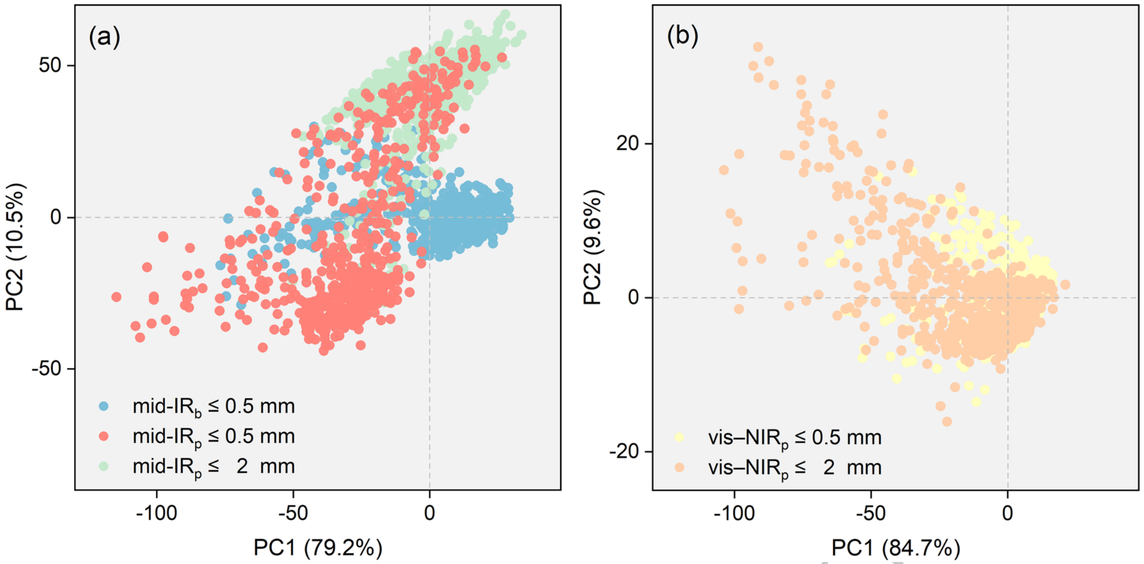


FIGURE 3 Scatter plots of the scores of second principal component (PC2) against those of the first principal component (PC1) for soil absorbance spectra: (a) PCA transformed mid-IR spectra values and (b) PCA transformed Vis–NIR spectra values

Figure 4b, which accords with our results shown in Figures 2 and 3a. We found no clear difference between the estimates of SOC from the mid-IR$_p$ ($R^2 = 0.84$, RMSE = 0.35) and vis–NIR$_p$ ($R^2 = 0.81$, RMSE = 0.38) for the ≤2 mm fraction. However, the estimates of SOC with the vis–NIR$_p$ on the ≤0.5 mm fraction ($R^2 = 0.88$, RMSE = 0.30) were more accurate than those with the mid-IR$_p$ on ≤2 mm fraction.

Overall, estimates from the mid-IR$_p$ were somewhat more accurate than those from the vis–NIR$_p$ on soil ground to the same particle size. There is one poorly predicted sample for the ≤2 mm particle fraction shown in Figure 4c,e (shown as red discs) but which was well predicted on the ≤0.5 mm particle fraction. We checked the sample quality and further examined this by remeasuring the sample five times with all of the spectrometers to avoid any measurement error (see Figure A1).

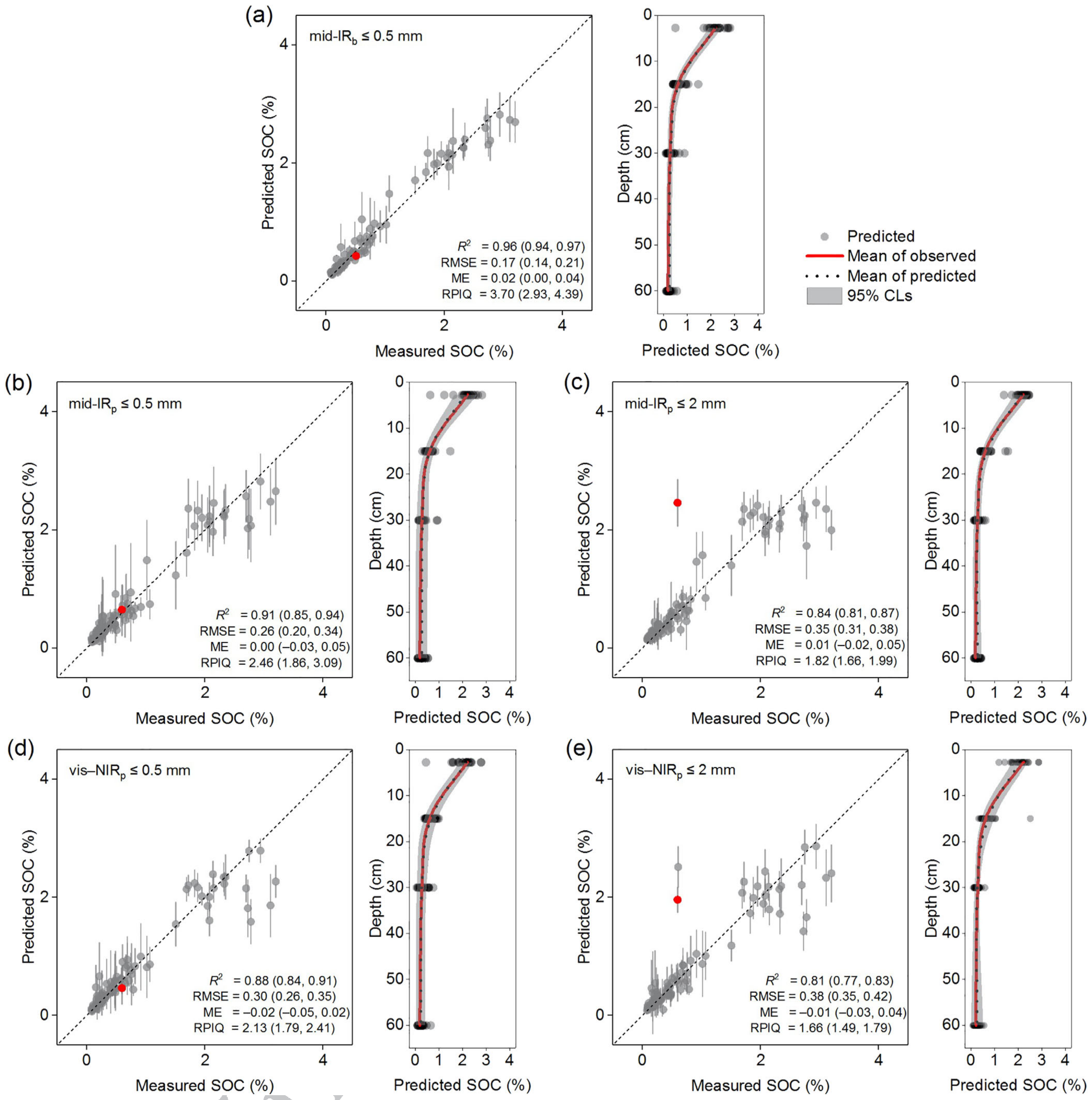
3.3 | Comparison of cost-effectiveness

Table 2 lists the approximate time taken for one individual sample by DC and each of the spectroscopic methods. Overall, the DC was the most time-consuming, requiring 500 s for sample preparation (T_{sp}) and 780 s for data acquisition (T_{da}). The total time for portable instruments accounted for around 27%–45% relative to that of DC. The T_{sp}, accounted from more than half to 90% of the total time for spectroscopic methods, but only around 39% for the DC. This was mainly because spectroscopic measurement itself is quick; the time taken to prepare the samples, T_{sp}, is not changed. As one of three components of the T_{da}, loading samples (T_{ld}) for the mid-IR$_b$ takes the longest time (240 s) among all methods; this is because four replicates are needed (60 s for each) and loading a spot of sample into each well of the sample plates requires great care and patience.

Table 3 lists the values of α and β for both a realistic 60% working day and a whole working day, 100%, as ultimate limits for the methods. The capital cost (Ω, in AU $) and inaccuracy (in %) for each method are also given. The times taken, listed in Table 2, and so it is evident that more samples can be measured in a full day, 100%, than at 60%. There is one exception, namely DC for which β is constant. As the sample racks with boats are fixed, β is the same for 60% and 100% of the working day. Determination of SOC with the spectrometers is less accurate than by DC, but because spectroscopy is quicker its values of β are greater than that of the DC (50 samples per day).

Table 4 lists the approximate costs of determining the SOC by DC and spectroscopy for a comfortable working day (60%) and a full day (100%). See also Table 2. For all methods working at 60% of the day costs more than at 100%. For example, the DC was the most expensive method with a total cost AU $ 35 (60% α) or AU $ 31 (100% α) per sample. Although mid-IR$_b$ on the ≤0.5 mm ranked the second, it accounted for less than half to one-third that for DC. The total costs for determining the SOC with portable spectrometers are much less than those of the more refined techniques because they are quicker; their values of β are much larger.

Figure 5 shows that for each spectrometer its E relative to DC increases up to its full capacity, β (i.e., for DC



Color online, B&W in print

FIGURE 4 Validation of independent test data of the SOC estimates from cubist showing the observed values against the estimated ones and their uncertainties (bars), and estimations of SOC on cores with their 95% confidence intervals (shaded grey bands). The thick (red) line and (black) short dash line down the profiles represent the mean of observed and predicted values of SOC, respectively. Values in brackets are 95% lower and upper confidence limits (CLs)

$E = 1$). Figure 5a,b are the results of E under 60% and 100% α of the working day respectively. Lower and upper boundaries of the coloured area distinguish the 60% and 100% β. Each interval of area is the condition between 60% and 100% β. Figure 5c shows the intersection among methods under four levels depicted as a confusion matrix.

The spectroscopic methods are less accurate than dry combustion (DC). Nevertheless, they are quicker and cheaper, and many more samples can be measured per working day. That reduces the cost per datum and can make them more cost-effective. All the spectroscopic methods were more cost-effective than DC as long as

TABLE 2 Approximate time taken (in seconds) for individual measurements of SOC concentrations made with the DC (dry combustion) analyser and three spectrometers (mid-IR$_b$, mid-IR$_p$, and Vis–NIR$_p$)

Method	Sample preparation		Data acquisition			Total time/s
	≤2 mm	≤0.5 mm	Loading	Reference	Measurement	
DC	–	500	120	300	360	1280
Mid-IR$_b$	–	500	240[a]	4[a]	120[a]	864
Mid-IR$_p$	–	500	35	5	30	570
	300	–	35	5	30	370
Vis–NIR$_p$	–	500	35	5	10	550
	300	–	35	5	10	350

[a]Comprised four replicates.

TABLE 3 Approximate numbers of samples under 60% and 100% capacities of the sample preparation (α) and data acquisition (β) within one working day made by a DC (dry combustion) analyser and three spectrometers (mid-IR$_b$, mid-IR$_p$, and Vis–NIR$_p$)

Method	Particle diameter/mm	Capital cost/AU $	α 60%	α 100%	β 60%	β 100%	RMSE/%
DC	≤0.5	200,000	30	50	50	50	0.09
Mid-IR$_b$	≤0.5	120,000	30	50	132[a]	220[a]	0.17
Mid-IR$_p$	≤0.5	80,000	30	50	200	334	0.26
	≤2	80,000	50	84	200	334	0.35
Vis–NIR$_p$	≤0.5	80,000	30	50	280	468	0.30
	≤2	80,000	50	84	280	468	0.38

[a]Comprised four replicates.

more than some threshold number of samples were measured.

For each spectroscopic method, the E for 100% β always exceeds the corresponding E for 60% β for the same number of samples, n, whatever the values of α are. This is simply because more samples can be measured per day (larger β) thereby reducing the cost of data acquisition (C_{da}) and therefore the total cost. The entries is Table 3 confirm this. It is the same for α; the E at 100% α is always greater than that of the 60% with same β. This is because when β is fixed, the greater capacity for preparing samples means that more samples are ready for measurement (see Table 3), and that helps to reduce the cost for preparation of sample (C_{sp}) significantly and further reduce the total cost of the whole operation. Not only that, larger α or β diminished the number of intersections (see Figure 5c). More details of the maximum E from Figure 5a, b are listed in Table 5. The value of E for DC remains at 1. With greater throughput, faster data acquisition and good accuracy, the vis–NIR$_p$ on the ≤2 mm samples was the most cost-effective for estimating organic C; its E ranged from 2.11 to 5.24 times that of the DC measurement at its full capacity among four levels of work.

4 | DISCUSSION

4.1 | Performance of the spectroscopic models

The laboratory measurements with the mid-IR$_b$ spectrometer produced sharper spectra than did mid-IR$_p$ instrument. This is because the spectrometers have different detectors. The detector in the mid-IR$_p$ Agilent 4200 spectrometer is a thermoelectrically regulated DTGS, the sensitivity of which is roughly 1000th of that of the liquid-cooled MCT detector in the mid-IR$_b$ Bruker Vertex 70 spectrometer. This might contribute to the accuracy of the estimates of SOC for the ≤0.5 mm samples from the mid-IR$_b$ spectra. Figures 2 and 3a show a larger variation of the mid-IR$_p$ spectra on the ≤0.5 mm soil samples, particularly in the region between 2000 and 650 cm^{-1}. It might be the reason that the estimates of SOC with the spectra of the ≤0.5 mm samples are less certain (Figure 4b).

The mid-IR$_p$ spectra produced somewhat more accurate estimates of SOC than those from the vis–NIR$_p$ on the same particle fraction (Figure 4). One reason is that

TABLE 4 Approximate cost (in AU \$) for individual measurements of SOC concentrations made with the DC (dry combustion) analyser and three spectrometers (mid-IR$_{\mathrm{b}}$, mid-IR$_{\mathrm{p}}$, and Vis–NIR$_{\mathrm{p}}$)

	Sample preparation		Data acquisition[a]		
Method	≤2 mm	≤0.5 mm	Measurement[b]	Measurement[c]	Total cost/AU \$
DC	–	10.5	6.3	18.3	35
Mid-IR$_{\mathrm{b}}$	–	10.5	(2.4, 1.4)	(4.2, 2.5)	(17, 14)
Mid-IR$_{\mathrm{p}}$	–	10.5	(1.6, 0.9)	(1.8, 1.1)	(14, 13)
	6.3	–	(1.6, 0.9)	(1.8, 1.1)	(10, 8)
Vis–NIR$_{\mathrm{p}}$	–	10.5	(1.1, 0.7)	(1.3, 0.8)	(13, 12)
	6.3	–	(1.1, 0.7)	(1.3, 0.8)	(9, 8)
DC	–	6.3	6.3	18.3	31
Mid-IR$_{\mathrm{b}}$	–	6.3	(2.4, 1.4)	(4.2, 2.5)	(13, 10)
Mid-IR$_{\mathrm{p}}$	–	6.3	(1.6, 0.9)	(1.8, 1.1)	(10, 8)
	3.8	–	(1.6, 0.9)	(1.8, 1.1)	(7, 6)
Vis–NIR$_{\mathrm{p}}$	–	6.3	(1.1, 0.7)	(1.3, 0.8)	(9, 8)
	3.8	–	(1.1, 0.7)	(1.3, 0.8)	(6, 5)

Note: The upper section of the table is for $\alpha = 60\%$ and lower section for $\alpha = 100\%$. The entries in the square brackets are for $\beta = 60\%$ (the first entry) and for $\beta = 100\%$ (second entry).

[a] From 60% (left in the bracket) to 100% β (right in the bracket).

[b] By operator.

[c] By instrument.

the absorptions in mid-IR range correspond to intense fundamental vibrations of bonds in molecules (Williams & Norris, 1987), whereas the absorption in the NIR range corresponds to weak overtones and combinations of the components at characteristic wavelengths or in particular wavelengths in that range and in the visible range they are electronic transitions of atoms (Vohland et al., 2014). The estimates from the vis–NIR$_{\mathrm{p}}$ spectra of the ≤0.5 mm samples were more accurate than those from the mid-IR$_{\mathrm{p}}$ spectra on ≤2 mm fractions, however. This indicates that sample preparation and surface conditions of the soil have a significant effect on spectroscopic predictions. Grinding to a fine particle size homogenises the soil and increases the overall accuracy of SOC estimates (Le Guillou et al., 2015; Stumpe et al., 2011).

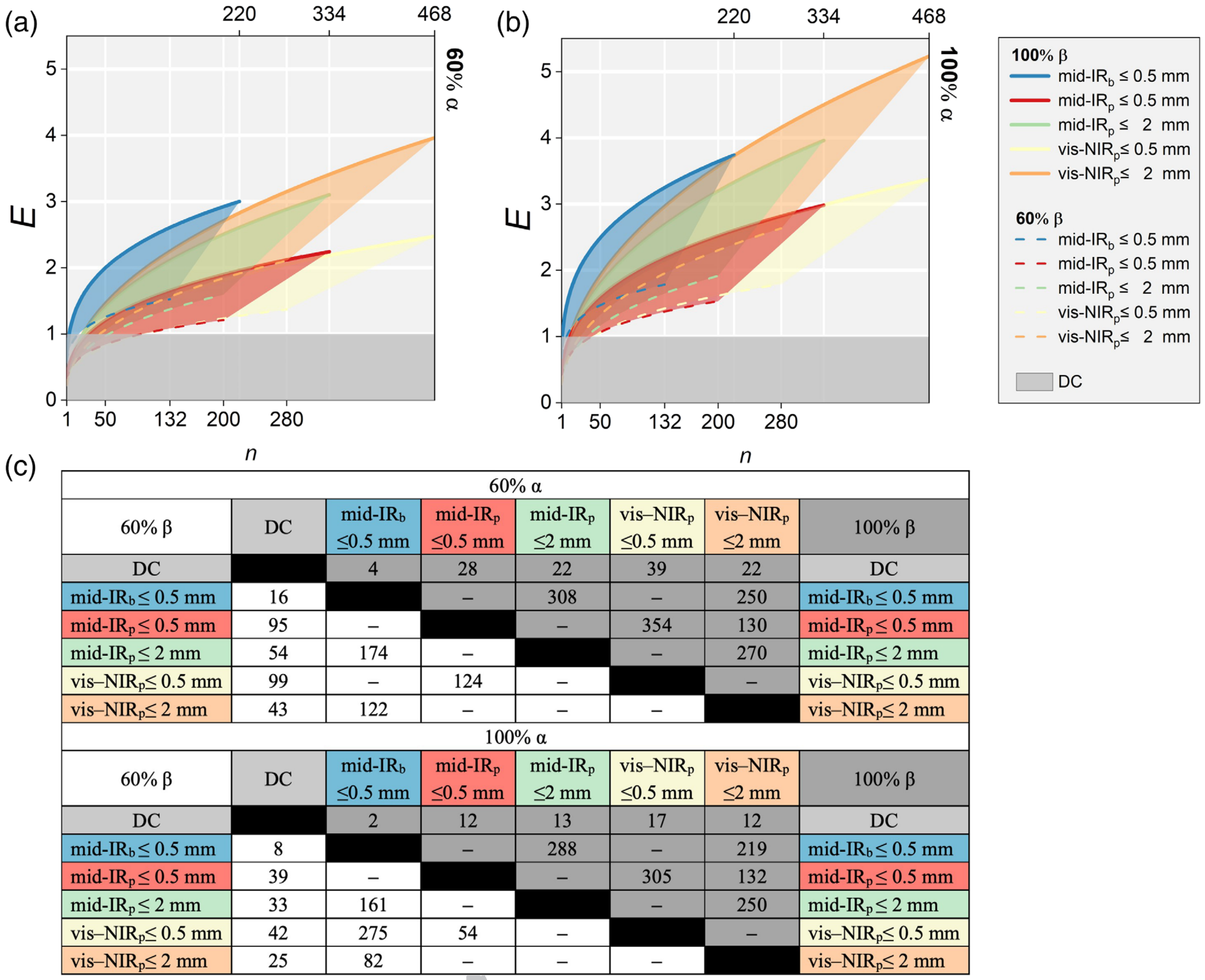
We believe that the overestimates of the SOC for the ≤2 mm samples (red discs), shown in Figures 4c,e and A1, are likely to be caused by poor penetration of light into the soil so that diffuse reflection of the light from the soil is also poor. Stenberg et al. (2010) observed larger errors in estimates of SOC by vis–NIR spectroscopy when the samples contained larger amounts of sand (which are coarser than clay and silt) and that the errors in the sandiest soil were dominated by overestimates of SOC. We, therefore, suspect the mid-IR suffers similarly; it is a matter that needs further investigation.

4.2 | Factors affecting the cost effectiveness

In our cost-effectiveness formula, Equation (2), four main factors affecting E with the numbers of samples (n) increasing are the inaccuracy (U), the total cost (C) and the capacities of sample preparation (α) and data acquisition (β). Having found that all spectroscopic methods at full capacity are more cost-effective (i.e., $E>1$) than dry combustion, we discuss below only those situations for which E exceeds 1 (see Figure 5).

As one of two components of the C, the cost on sample preparation (C_{sp}) for dry combustion accounted for 20%–30% of its C, but more than 50% of that for spectroscopy (one exception is the mid-IR$_{\mathrm{b}}$ under a medium working level, which was 48%). This suggests that the effort it takes to grind the samples can diminish E and thus the overall performance of the spectroscopic methods.

The cost of measurement by particular instrument (C_{mi}) was determined by the depreciation, capital cost (Ω) and the β (see Equation 7). These C_{mi} were less than around 30% of the total cost per sample for the spectroscopic methods compared with around 50% (AU \$ 18.3) for DC (AU \$ 35 at 60% α, or AU \$ 31 at 100% α). The purchase price of an instrument as the Ω can vary from one supplier to another, and this variation will positively



60% α

60% β	DC	mid-IR_b ≤0.5 mm	mid-IR_p ≤0.5 mm	mid-IR_p ≤2 mm	vis–NIR_p ≤0.5 mm	vis–NIR_p ≤2 mm	100% β
DC	■	4	28	22	39	22	DC
mid-IR_b ≤ 0.5 mm	16	■	–	308	–	250	mid-IR_b ≤ 0.5 mm
mid-IR_p ≤ 0.5 mm	95	–	■	–	354	130	mid-IR_p ≤ 0.5 mm
mid-IR_p ≤ 2 mm	54	174	–	■	–	270	mid-IR_p ≤ 2 mm
vis–NIR_p ≤ 0.5 mm	99	–	124	–	■	–	vis–NIR_p ≤ 0.5 mm
vis–NIR_p ≤ 2 mm	43	122	–	–	–	■	vis–NIR_p ≤ 2 mm

100% α

60% β	DC	mid-IR_b ≤0.5 mm	mid-IR_p ≤0.5 mm	mid-IR_p ≤2 mm	vis–NIR_p ≤0.5 mm	vis–NIR_p ≤2 mm	100% β
DC	■	2	12	13	17	12	DC
mid-IR_b ≤ 0.5 mm	8	■	–	288	–	219	mid-IR_b ≤ 0.5 mm
mid-IR_p ≤ 0.5 mm	39	–	■	–	305	132	mid-IR_p ≤ 0.5 mm
mid-IR_p ≤ 2 mm	33	161	–	■	–	250	mid-IR_p ≤ 2 mm
vis–NIR_p ≤ 0.5 mm	42	275	54	–	■	–	vis–NIR_p ≤ 0.5 mm
vis–NIR_p ≤ 2 mm	25	82	–	–	–	■	vis–NIR_p ≤ 2 mm

Color online, B&W in print

FIGURE 5 The cost-effectiveness (E) of measuring soil organic carbon concentrations with a DC analyser and three spectrometers: Mid-IR_b, mid-IR_p, and Vis–NIR_p under (a) 60% α, (b) 100% α, and (c) the intersections over or equal to $E = 1$, combined with 60% β and 100% β, respectively. The dashed and solid boundaries of the coloured area represent the 60% β and 100% β, respectively. Columns represent the 60% α (upper matrix) and 100% α (lower matrix); rows represent the 60% β (lower triangular matrix in white) and 100% β (upper triangular matrix in grey), for example, no 16 at DC column and mid-IR_b ≤0.5 mm row means the E of former is equal to that of latter when 16 samples were measured

TABLE 5 Maximum values of E made with the DC dry combustion) analyser and three spectrometers (mid-IR_b, mid-IR_p, and Vis–NIR_p) at full capacity under four levels of work

Level	α (%)	β (%)	DC ≤0.5 mm	Mid-IR_b ≤0.5 mm	Mid-IR_p ≤0.5 mm	Mid-IR_p ≤2 mm	Vis–NIR_p ≤0.5 mm	Vis–NIR_p ≤2 mm
Normal	60	60	1	1.52	1.21	1.60	1.37	2.11
Medium	100	60	1	1.78	1.53	1.92	1.78	2.64
High	60	100	1	3.00	2.24	3.10	2.48	3.96
Extreme	100	100	1	3.73	2.97	3.96	3.36	5.24

affect the value of E. However, Ω is insignificant and is masked by the greater capacities of the spectroscopic methods. Moreover, the price of spectrometers is decreasing with greater technological development, for example, development of spectrometers based on microelectromechanical systems (MEMS). The cost of labour (e.g., the hourly wage) has a large effect on the *E*. The salary could be less, perhaps much less, in countries other than Australia (e.g., 10% of AU $ or even less). For the spectroscopic methods, however, the effect on *E* is likely to be small because of the greater capacities of the instruments than those of dry combustion.

The rank order of maximum values of *E*, see Table 5, increases as the working day intensifies. The greater capacity leads to smaller costs on sample preparation (C_{sp}) or data acquisition (C_{da}), and finally to the substantially smaller total cost. For spectroscopy, however, increasing the capacity of data acquisition (β) has a larger effect on increasing *E* than does increase the capacity of sample preparation (α). For example, in the mid-IR$_b$ increasing the working day for sample preparation from 60% to 100% increased *E* from 1.52 to 1.78, a difference of 0.26; increasing the working day for measurement increased *E* from 1.52 to 3.00, a difference of 1.48.

The upper half of Figure 5c is additionally instructive. We found the estimates of SOC with the mid-IR$_b$ spectra on ≤0.5 mm samples to be consistently more cost-effective ($E > 1$) than the DC measurements provided more than four soil samples are being measured at the high working level or more than 16 samples are being measured at the normal working level. We attribute this partly to the accuracy of the technique (RMSE = 0.17%) and partly to the automation and consequent small operational costs (C_{mo}). Automation with the robotic arm means that it can be operated comfortably for 100% of the working day to measure 220 samples per day rather than only 132 samples at 60%; in other words its capacity for data acquisition, β, is much larger.

For portable spectrometers, the final ranking of their *E* values (under each level) was the reverse of their accuracy. For example, the mid-IR$_p$ on ≤0.5 mm was the most accurate, but least cost-effective among the mid-IR$_p$ and vis–NIR$_p$ on ground samples. The reason is that more time is needed to prepare the ≤0.5 mm samples (T_{sp} 500 s) than the ≤2 mm samples (300 s), and to measure them (mid-IR$_p$ $T_{mm} = 30$ s; vis–NIR$_p$ $T_{mm} = 10$ s) (Table 2).

We do not claim that the advantage of the larger throughput with a portable spectrometer is necessarily more cost-effective regardless of its accuracy. First, a tolerable error is arbitrary (Lark & Knights, 2015) and depends on the data and the application. For instance, an RMSE of 0.5% SOC might be acceptable, when the mean SOC is large or when many measurements are made for mapping by kriging, for example, kriging would diminish the prediction error variances further. It would be too large, however, if the mean SOC was small or for carbon accounting in a region or farm (Viscarra Rossel, Brus, et al., 2016).

Although spectroscopic estimates on any one sample might be less precise than measurement by conventional chemical analysis, many more measurements can be made for the same cost and at the same time. The final estimates might well then be more precise and hence more informative, (e.g., the many spectral estimates better characterising spatial variation compared with only a few measurements made by conventional laboratory methods). Another important advantage of spectroscopy is that a single spectrum allows for simultaneous estimation of other soil properties Viscarra Rossel and Webster (2012), although our study here does not account for this.

5 | CONCLUSIONS

We have determined the inaccuracy of estimating the concentration of SOC from mid-IR$_b$, mid-IR$_p$ and vis–NIR$_p$ spectra measured with commercial spectrometers on ground samples. We have also measured the time required to prepare the samples beforehand and to capture the spectra themselves from the instruments. From these measurements and the costs of the instruments, we have calculated the cost per estimate and its effectiveness relative to the standard laboratory technique with the DC analyser.

We conclude the following:

1. Dry combustion, which is well established as a standard, should be the first choice when fewer than 50 samples are to be measured per day.
2. The mid-IR$_b$ was the most accurate of the spectroscopic techniques ($R^2 = 0.96$, RMSE = 0.17, RPIQ = 3.70); its cost-effectiveness also was excellent, so it would be a good choice for fairly accurate determinations of SOC on 100–200 soil samples per day with four or more replicates.
3. The cost-effectiveness of mid-IR$_p$ on ≤2 mm is comparable with that of the mid-IR$_b$.
4. Measurements on soil ground to ≤2 mm with the vis–NIR$_p$ spectrometer were the most cost-effective when more than 250 samples were measured per day.
5. One can increase the *E* of spectroscopy, that is, improve its performance, either by reducing the time of sample preparation and data acquisition to reduce the cost and increase the capacities or by improving the accuracy of the spectroscopic modelling.
6. We do not recommend finer grinding of soil samples for either portable vis–NIR or mid-IR spectrometers because the small gain in accuracy does not justify the additional cost of the labour.

Looking ahead, we think that automation with portable instruments recording spectra directly in the field should increase efficiency markedly. We recognise that such measurements would be strongly affected by soil water content. It would be necessary to use either direct standardisation (DS, Ji et al., 2015) or external parameter orthogonalization (EPO, Minasny et al., 2011) to remove the effects. Promising results have been reported previously by (Viscarra Rossel et al., 2017), but are undeveloped in mid-IR$_p$, which should also be tested in future.

ACKNOWLEDGEMENTS

We thank the International Postdoctoral Exchange Fellowship Program from the office of the China Postdoctoral Council for supporting S. Li, and Agilent Technologies for lending us the 4200 portable mid-infrared spectrometers for the tests we report.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

S. Li https://orcid.org/0000-0001-8675-4070
R. A. Viscarra Rossel https://orcid.org/0000-0003-1540-4748

REFERENCES

Atwood, T. B., Connolly, R. M., Almahasheer, H., Carnell, P. E., Duarte, C. M., Lewis, C. J. E., Irigoien, X., Kelleway, J. J., Lavery, P. S., Macreadie, P. I., Serrano, O., Sanders, C. J., Santos, I., Steven, A. D. L., & Lovelock, C. E. (2017). Global patterns in mangrove soil carbon stocks and losses. *Nature Climate Change*, *7*, 523–528.

Baumann, P., Helfenstein, A., Gubler, A., Keller, A., Meuli, R. G., Wächter, D., Lee, J., Viscarra Rossel, R., & Six, J. (2021). Developing the Swiss soil spectral library for local estimation and monitoring. *The Soil*, *7*, 525–546.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends in Analytical Chemistry*, *29*, 1073–1081.

Bellon-Maurel, V., & McBratney, A. (2011). Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils–critical review and research perspectives. *Soil Biology and Biochemistry*, *43*, 1398–1410.

Capriel, P., Beck, T., Borchert, H., Gronholz, J., & Zachmann, G. (1995). Hydrophobicity of the organic matter in arable soils. *Soil Biology and Biochemistry*, *27*, 1453–1458.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.

Collewet, M., & Sauermann, J. (2017). Working hours and productivity. *Labour Economics*, *47*, 96–106.

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, *440*, 165–173.

Dhawale, N. M., Adamchuk, V. I., Prasher, S. O., Viscarra Rossel, R. A., Ismail, A. A., & Kaur, J. (2015). Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer. *European Journal of Soil Science*, *66*, 661–669.

Fontaine, S., Barot, S., Barré, P., Bdioui, N., Mary, B., & Rumpel, C. (2007). Stability of organic carbon in deep soil layers controlled by fresh carbon supply. *Nature*, *450*, 277–280.

Greenberg, I., Seidel, M., Vohland, M., & Ludwig, B. (2021). Performance of field-scale lab vs in situ visible/near- and mid-infrared spectroscopy for estimation of soil properties. *European Journal of Soil Science*. https://doi.org/10.1111/ejss.13180

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: Data mining, inference and prediction. *The Mathematical Intelligencer*, *27*, 83–85.

Hutengs, C., Ludwig, B., Jung, A., Eisele, A., & Vohland, M. (2018). Comparison of portable and bench-top spectrometers for mid-infrared diffuse reflectance measurements of soils. *Sensors*, *18*, 993.

Hutengs, C., Seidel, M., Oertel, F., Ludwig, B., & Vohland, M. (2019). In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma*, *355*, 113900.

Isbell, R. F. (2002). *The Australian soil classification: Revised Edition*. CSIRO Publishing.

IUSS Working Group WRB. (2006). *World reference base for soil resources* (World soil resources Reports No. 103). FAO.

Janik, L. J., Merry, R. H., & Skjemstad, J. O. (1998). Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture*, *38*, 681–696.

Ji, W., Adamchuk, V. I., Biswas, A., Dhawale, N. M., Sudarsan, B., Zhang, Y., Viscarra Rossel, R. A., & Shi, Z. (2016). Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering*, *152*, 14–27.

Ji, W., Viscarra Rossel, R. A., & Shi, Z. (2015). Accounting for the effects of water and the environment on proximally sensed Vis-NIR soil spectra and their calibrations. *European Journal of Soil Science*, *66*, 555–565.

Kempthorne, O., & Allmaras, R. R. (1965). *Errors of observation*. ASA/SSSA.

Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, *11*, 137–148.

Kwon, Y. (2001). *Handbook of essential pharmacokinetics, pharmacodynamics and drug metabolism for industrial scientists* (2nd ed.). Kluwer Academic/Plenum Publishers.

Lark, R. M., & Knights, K. V. (2015). The implicit loss function for errors in soil information. *Geoderma*, *251*, 24–32.

Le Guillou, F., Wetterlind, W., Viscarra Rossel, R. A., Hicks, W., Grundy, M., & Tuomi, S. (2015). How does grinding affect the mid-infrared spectra of soil and their multivariate calibrations to texture and organic carbon? *Soil Research*, *53*, 913–921.

Li, S., Shi, Z., Chen, S., Ji, W., Zhou, L., Yu, W., & Webster, R. (2015). In situ measurements of organic carbon in soil profiles using Vis-NIR spectroscopy on the Qinghai–Tibet plateau. *Environmental Science & Technology*, *49*, 4980–4987.

Martens, H., & Næs, T. (1989). *Multivariate calibration*. John Wiley & Sons.

Minasny, B., McBratney, A. B., Bellon-Maurel, V., Roger, J. M., Gobrecht, A., Ferrand, L., & Joalland, S. (2011). Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma*, *167-168*, 118–124.

Mishra, U., Hugelius, G., Shelef, E., Yang, Y., Strauss, J., Lupachev, A., Harden, J. W., Jastrow, J. D., Ping, C.-L., Riley, W. J., Schuur, E. A. G., Matamala, R., Siewert, M., Nave, L. E., Koven, C. D., Fuchs, M., Palmtag, J., Kuhry, P., Treat, C. C., …

Orr, A. (2021). Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks. *Science Advances*, *7*, eaaz5236. https://doi.org/10.1126/sciadv.aaz5236

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Dor, E. B., Brown, D. J., Clairotte, M., Csorba, Á., Dardenne, P., Demattê, J., Génot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., … Wetterlind, J. (2015). Soil spectroscopy: An alternative to wet chemistry for soil monitoring. *Advances in Agronomy*, *132*, 139–159.

O'Rourke, S. M., & Holden, N. M. (2011). Optical sensing and chemometric analysis of soil organic carbon–a cost effective alternative to conventional laboratory methods? *Soil Use and Management*, *27*, 143–155.

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G. P., & Smith, P. (2016). Climate-smart soils. *Nature*, *532*, 49–57.

Quinlan, J. R. (1992). Learning with continuous classes. In A. Adams & L. Sterling (Eds.), *Proceedings AI'92, 5th Australian Conference on Artificial Intelligence* (pp. 343–348). World Scientific.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rayment, G. E., & Lyons, D. J. (2011). *Soil chemical methods: Australasia*. CSIRO Publishing.

Reeves, J. B., III, Follett, R. F., McCarty, G. W., & Kimble, J. M. (2006). Can near or mid-infrared diffuse reflectance spectroscopy be used to determine soil carbon pools? *Communications in Soil Science and Plant Analysis*, *37*, 2307–2325.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*, 1627–1639.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., & Trumbore, S. E. (2011). Persistence of soil organic matter as an ecosystem property. *Nature*, *478*, 49–56.

Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., Macdonald, L. M., & McLaughlin, M. J. (2014). The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews*, *49*, 139–186.

Stell, E., Warner, D., Jian, J., Bond-Lamberty, B., & Vargas, R. (2021). Spatial biases of information influence global estimates of soil respiration: How can we improve global predictions? *Global Change Biology*, *27*, 3923–3938. https://doi.org/10.1111/gcb.15666

Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, *107*, 163–215.

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.

van Leeuwen, C., Mulder, V., Batjes, N. M. G. B., & Heuvelink, G. (2021). Statistical modelling of measurement error in wet chemistry soil data. *European Journal of Soil Science*, 1–17. https://doi.org/10.1111/ejss.13137

Viscarra Rossel, R. A. (2007). Robust modelling of soil diffuse reflectance spectra by "bagging-partial least squares regression". *Journal of Near Infrared Spectroscopy*, *15*, 39–47.

Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., Zhou, S., Bo, S., Antoine, S., Viacheslav, A., Hamouda, A., Barthès, B., Harm, B., Anita, B., Martial, B., Kristin, B., Lukas, B., Du Changwen, C. A., & Wenjun, J. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, *155*, 198–230.

Viscarra Rossel, R. A., Brus, D., Lobsey, C., Shi, Z., & McLachlan, G. (2016). Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference. *Geoderma*, *265*, 152–163.

Viscarra Rossel, R. A., & Hicks, W. S. (2015). Soil organic carbon and its fractions estimated by visible–near infrared transfer functions. *European Journal of Soil Science*, *66*, 438–450.

Viscarra Rossel, R. A., Lobsey, C. R., Sharman, C., Flick, P., & McLachlan, G. (2017). Novel soil profile sensing to monitor organic C stocks and condition. *Environmental Science & Technology*, *51*, 5630–5641.

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflactance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, *131*, 59–75.

Viscarra Rossel, R. A., & Webster, R. (2012). Predicting soil properties from the Australian soil visible-near infrared spectroscopic database. *European Journal of Soil Science*, *63*, 848–860.

Viscarra Rossel, R. A., Webster, R., Bui, E. N., & Baldock, J. A. (2014). Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change. *Global Change Biology*, *20*, 2953–2970.

Vohland, M., Ludwig, M., Thiele-Bruhn, S., & Ludwig, B. (2014). Determination of soil properties with visible to near-and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma*, *223*, 88–96.

Williams, P., & Norris, K. (1987). *Near-infrared technology in the agricultural and food industries*. American Association of Cereal Chemists.

How to cite this article: Li, S., Viscarra Rossel, R. A., & Webster, R. (2021). The cost-effectiveness of reflectance spectroscopy for estimating soil organic carbon. *European Journal of Soil Science*, e13202. https://doi.org/10.1111/ejss.13202

APPENDIX A

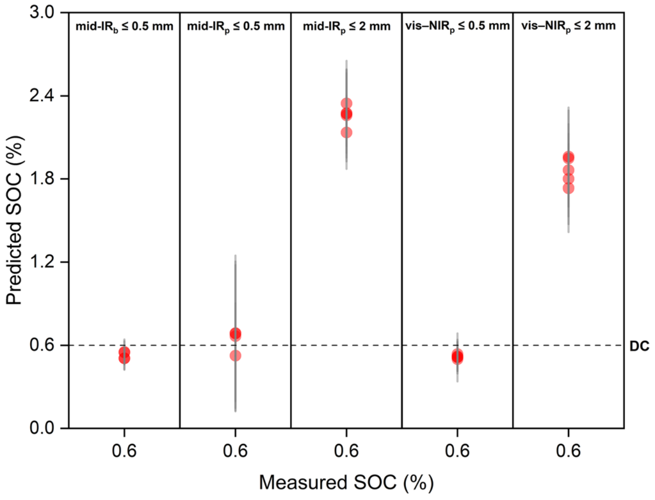


FIGURE A1 The predictions of soil organic carbon concentrations measured with a DC analyser and three spectrometers: Mid-IR$_b$, mid-IR$_p$, and Vis–NIR$_p$