Near and mid infrared soil spectroscopy

Wartini Ng[a], Brendan Malone[b], Budiman Minasny[a], and Sangho Jeon[c], [a]Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Eveleigh, NSW, Australia; [b]CSIRO Agriculture and Food, Black Mountain, ACT, Australia; [c]National Institute of Agricultural Sciences, Rural Development Administration, Wanju, Republic of Korea

Abstract

Soil Infrared spectroscopy uses electromagnetic energy to characterize the soil. In the past decades, visible, near to mid infrared spectroscopy has gained considerable interest as a viable alternative for in-field and laboratory measurements, especially where a high spatial density of measurements is needed, because of its speed and cost-effectiveness. Once chemometric models have been calibrated, a single spectrum can be used to predict a range of soil properties simultaneously. This chapter provides an overview of spectroscopy in soil research, including predicting physical, chemical, and biological properties, as well as determination of soil contaminants.

Key points

- Infrared spectroscopy has been used to predict a range of physical, chemical and biological soil properties successfully.
- Soil contaminants have also been successfully quantified with infrared spectroscopy.
- Developments in artificial intelligence, computational power and big data have made more complex machine learning algorithms possible.
- Smaller, portable and lower-cost spectrometers are now available.

Introduction

Obtaining quantitative soil information rapidly and objectively is essential for monitoring current soil conditions and maintaining if not improving them, to ensure the current management practice does not lead to soil degradation. Quantitative soil data, which vary in time and space, are conventionally obtained from laboratory analysis. However, this is expensive, time-consuming, and requires the use of chemical reagents.

Research has been conducted worldwide to investigate proximal sensors that can measure soil properties from a close distance to the soil surface using a various range of wavelengths of the electromagnetic spectrum as shown in Fig. 1. The electromagnetic energy ranges from gamma rays which have higher frequencies to radio waves with the lowest frequencies. The interaction of energy with matter depends on its molecular composition. A material can reflect, scatter, or emit electromagnetic radiation, which results in a unique spectral signature. These sensors measure the response of the material at certain wavelengths, providing information about its constituents. Soil responds uniquely to the infrared spectrum, and the technology has made infrared spectrometers suitable for the analysis of soil samples as they can measure rapidly, cost-effectively, and non-destructively.

Spectroscopy in soil science

Optical sensors utilize various wavelengths from the electromagnetic spectrum (Fig. 1) to characterize properties (visible (vis): 400–700 nm; near-infrared (NIR): 700–2500 nm; mid-infrared (MIR): 2500–25,000 nm). The absorption in the MIR region can be readily identified as it is related to the fundamental vibrations of molecules. The absorbance in the vis–NIR is more difficult to
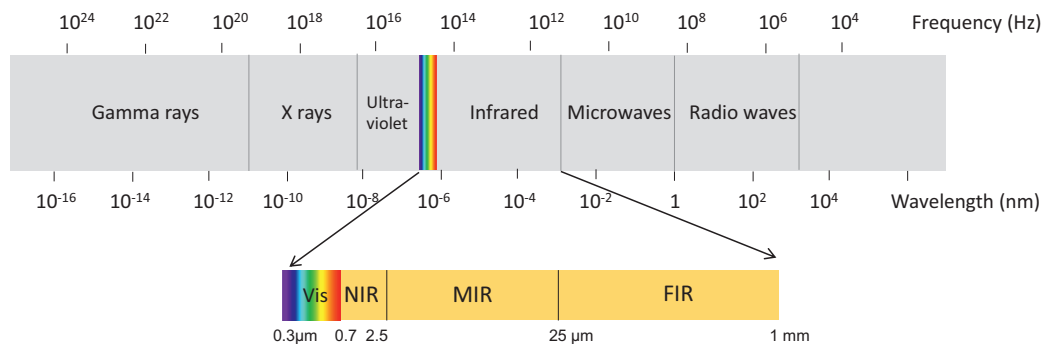
Fig. 1 The electromagnetic spectrum showing the boundaries between different regions, including the visible region (380–700 nm), the near infrared (NIR; 700–2500 nm), the mid-infrared (MIR: 2500–25,000 nm), and the far infrared (FIR 25,000–1,000,000 nm).
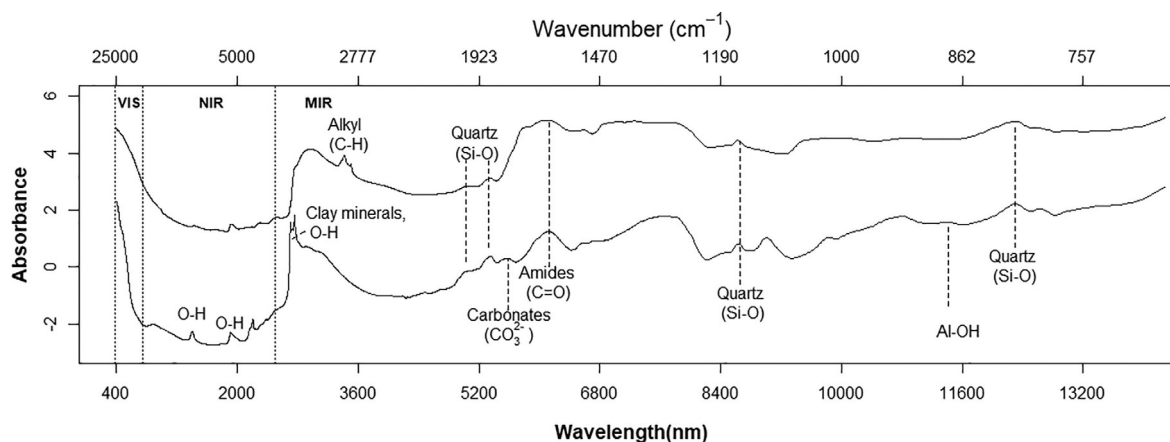
Fig. 2 The vis–NIR–MIR spectra of two soil samples showing fundamental peaks on MIR and overtones on NIR.

interpret as there are fewer and broader absorptions related to overtones and combinations of the fundamental vibrations in the MIR region. Absorptions in the visible region are associated with minerals that contain iron and organic matter, while absorptions in the NIR region (780–2500 nm) result from the overtones of hydroxyl (OH^-), sulfate (SO_4^{2-}), and carbonate (CO_3^{2-}) groups, and combinations of fundamental features of water (H_2O) and carbon dioxide (CO_2) (Stenberg et al., 2010). An illustration of spectra from the vis-NIR-MIR region is included in Fig. 2.

Discovery of the infrared region dates back to 1800 by Herschel (1800). However, the practical applications did not occur until 150 years later. In 1950s, researchers explored the use of infrared to characterize soil minerals (Hunt et al., 1950). In the 1980s, although NIR was mainly used in agriculture and the food industries, it became popular the decade after when practical applications started to expand to other industries, including pharmaceutical, earth science and environmental analyses. In soil science, early use of near infrared spectroscopy was by Bowers and Hanks (1965) where the influence of moisture, organic matter and particle size on the soil absorbance spectrum were observed.

Both vis–NIR and MIR have been used to determine successfully various chemical (pH, cation exchange capacity, carbon contents (organic, inorganic, and total) and nutrient contents), physical (texture (sand, silt and clay) and minerals (kaolinite, gibbsite, montmorillonite and iron oxides)), and biological (microbial biomass and enzyme activities) soil properties (Soriano-Disla et al., 2014).

The use of infrared spectroscopy is gaining interest for applications in various settings due to its possible application on different platforms, ranging from proximal sensing in the field, within the laboratory setting to remote sensing platforms with hyperspectral capabilities. Improvement in instrumentation also leads to the development of portable and handheld spectrometers. Recently, lower-cost portable and miniaturized NIR spectrometers with limited wavelength ranges have also become available, making it a cost-effective method of analysis for farmers, agricultural consultants, and researchers. It has been demonstrated that the miniaturized spectrometer provided comparable accuracy to the standard vis–NIR spectrometer (Tang et al., 2020).

Over the past two decades, the application of spectroscopy to complement soil analyses has increased rapidly. This increase is correlated with progress in the multivariate statistical data analysis and the development of both infrared and computer technologies.

Fundamental absorption of clay minerals

The vis–NIR and MIR spectra can be used for detecting clay mineral species and iron oxides. The most common soil forming clay minerals include kaolinite, montmorillonite (also commonly referred to as smectite) and illite. The most common iron oxide species include hematite and goethite. These minerals are distinct in composition, structural arrangements, and physicochemical characteristics. Consequently, each mineral has specific spectral responses at specific wavelengths in the electromagnetic spectrum (Clark et al., 1990). Earlier contributions from Gadsden (1975) comprehensively summarized diagnostic spectral features of minerals. Some diagnostic wavelengths in both the vis–NIR and MIR spectral regions for the key clay minerals and iron oxides are provided in Table 1.

The relatively simple way to assess the mineral composition of a soil sample directly from infrared spectra is to compare the reflectance of the diagnostic wavelengths from the reference spectra with those at the same wavelengths of the soil samples. As established, the wavelength ranges specifically diagnostic to each clay mineral and iron oxide specimen are known and they can be isolated and standardized by fitting a convex hull to them. The convex hull fits a straight line that connects the local maxima of the spectrum of a specific wavelength range and represents the background information. Fig. 3 shows an example of the fitted convex hulls to the diagnostic reflectance ranges from the reference spectra sourced from the United States Geological Survey (USGS) spectral library for end member specimens of smectite and kaolinite, and the associated spectra once the continuum is removed.

Soil infrared spectra are similarly standardized followed by the application of comparative metrics between reference and soil spectra to derive relative abundances. For example, shape match algorithms can be used and are implemented in the USGS Tetracorder decision making framework for mineral detection (Clark et al., 2003). Other metrics include relative area and depth of the absorbance feature and slope of the convex hull. Relative abundance of the minerals can then be inferred. For example, Fig. 4 provides statistical summaries of the relative abundances of kaolinite and smectite minerals derived from soil survey samples from the Hunter valley region, NSW Australia used in Malone et al. (2014). Here, there is a clear relationship between estimated abundance and its corresponding spectral response feature relative to the reference specimen.

The alternative to direct inference on mineral composition from infrared spectra is with soil spectral calibration models, which are described in more detail further on. Robinson and Kitching (2016) demonstrated this where partial least squares regression (PLSR) and support vector machine models were trialled and constructed to estimate soil clay mineral compositions (kaolinite, smectite and illite) based on associated soil MIR spectra. These models were extended to a larger spectral library to enable digital soil mapping of these variables in their study area in Victoria, Australia. The authors reported that their model can have an accuracy of 70% or better for each of the mineral species.

Soil chemometrics

As soil is a complex mixture of materials, it is difficult to assign particular features of the spectra to specific physical, chemical, or biological components. Multivariate calibration techniques are used to predict soil properties, relating the spectra to observed soil properties (properties that have been measured using conventional laboratory analysis).

Digital soil spectra usually contain hundreds or thousands of reflectance values as a function of wavelength. Since there are typically more predictor variables than the observed variables, methods that reduce the dimensions of the predictors are required. Principal component regression and PLSR are commonly utilized methods. Partial least squares regression extracts successive linear combinations of the predictors, optimizing the combined goals of explaining response variation and predictor variation.

Another way to handle large dimensional data is to use variable selection techniques. When the high dimensional data have been reduced to several components or important variables have been selected, they are used for prediction using either linear regression or data-mining tools. Regression trees, neural networks and support vector machines have been used for such predictions.

There are also machine learning algorithms that can handle a large dimension of inputs and extract information from many samples. Models such as Cubist models, random forest, neural networks, support vector machines have been trialled successfully for predicting soil properties from vis–NIR and MIR data. While such models have been used successfully for prediction, each model

Table 1 Diagnostic wavelength ranges of clay minerals and iron oxides as reported by Gadsden (1975).

Mineral	Diagnostic wavelength range(s) in wavenumber (cm^{-1})	
	Visible and near infrared	Mid infrared
Kaolinite	4812–4411	3636, 2710, 1923–1818, 1587, 1064, 855
Smectite	4721–4372	3636, 1961, 1852, 1639, 1111, 1064
Illite	4640–4413, 4336–4192	3226, 2874, 2141, 2604–2513, 1799, 1639
Goethite	21,881–17,761; 12,887–7899	912–882, 812–793
Hematite	21,978–16,340; 13,072–9524	560–550
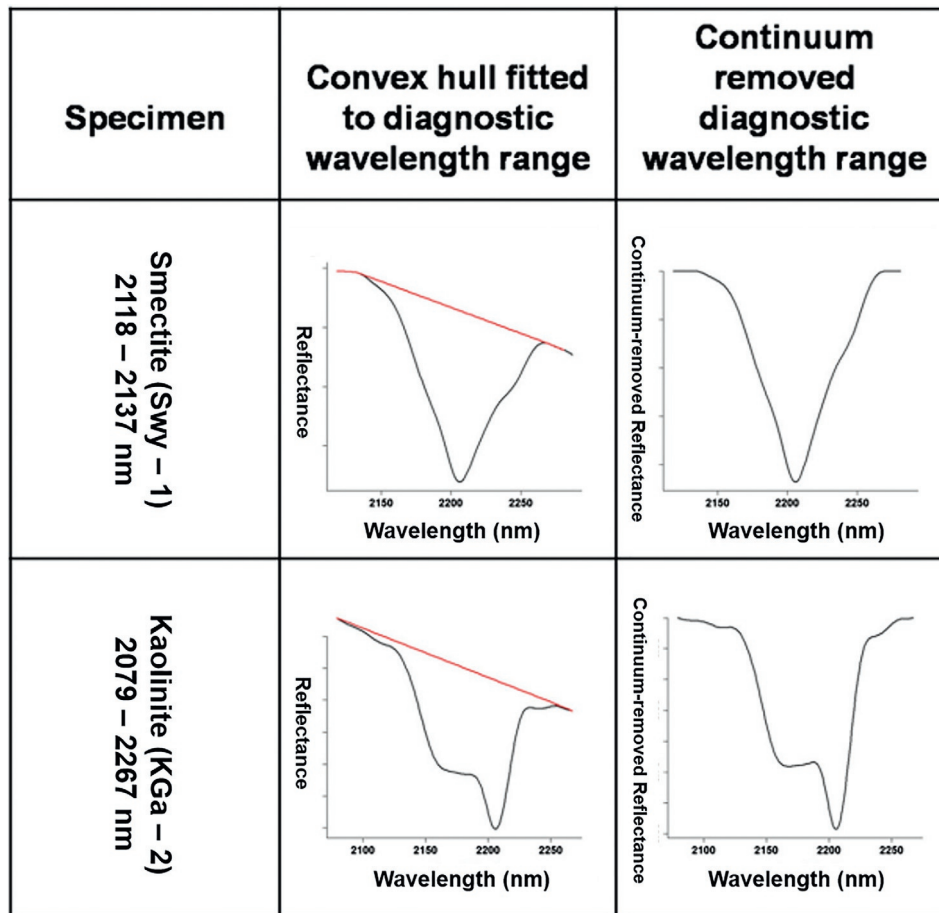
Fig. 3   Diagnostic wavelength ranges for smectite and kaolinite. To prepare for analyses, spectra need to be standardized by fitting a convex hull (red line). Soil spectra and reference spectra are then compared based on the continuum removed spectra. Sourced from Malone BP, Hughes P, McBratney AB, Minasny B (2014) A model for the identification of terrons in the Lower Hunter Valley, Australia. *Geoderma Regional* 1: 31–47.
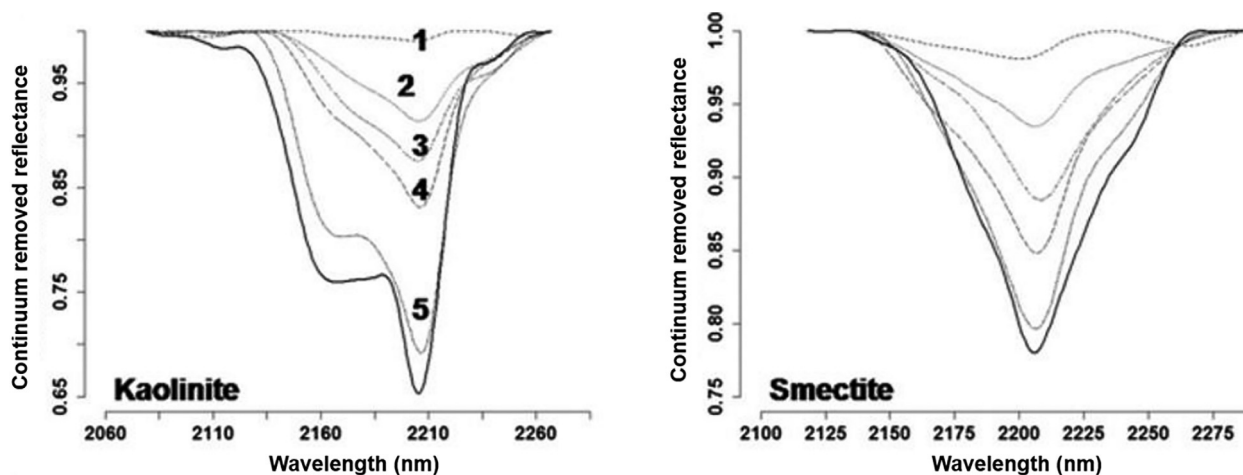
Fig. 4   Continuum removed soil spectra (dashed lines) with comparison to reference spectra at their given diagnostic wavelength ranges. The dashed lines represent (as numbered for kaolinite) the spectra for soils where relative mineral abundance is (1) the minimum, (2) first quartile, (3) second quartile, (4) third quartile, and (5) maximum. Partial figure sourced from Malone BP, Hughes P, McBratney AB, Minasny B (2014) A model for the identification of terrons in the Lower Hunter Valley, Australia. *Geoderma Regional* 1: 31–47.

form varies in complexity and ability of interpretation, and these features should be considered in detail in the exploratory stages of soil spectral model development.

Large soil spectral libraries are expanding as more spectral data are collected at a regional or national scale. An emerging method used for large data and the processing complex images, and which is capable of handling high-dimensional data is the deep learning model. One form is the convolutional neural network (CNN) which is widely used in handling high dimensional data, such as images. Ng et al. (2019) explored the use of the CNN model for predicting soil properties on a large vis-NIR and MIR spectral database. The architecture of the CNN model is quite flexible and can accept multi-dimensional input and provide multi-task predictions. The study reported that deep learning on an extensive MIR spectra library could predict fundamental soil properties (texture, organic and inorganic C, CEC, and pH) very accurately with $R^2 > 0.95$. Furthermore, only one model needed to be trained to provide predictions of multiple properties. Data fusion can also improve the prediction of soil properties.

Soil properties

Soil physical properties

Soil physical properties regulate the movement, retention and availability of water and nutrients to plants, and determine the flow of heat and air into, within and out of the soil. Despite the importance of these properties, a common feature in their measurement is that they often require considerable time to collect the data, are expensive to acquire, and require very specialized equipment. Consequently, alternative and less costly approaches with pedotransfer functions, which exploit easier to acquire attributes coupled with statistical modelling, have been proposed and have demonstrated their utility in numerous applications. Similarly, soil spectral modelling of these attributes has been demonstrated. Soriano-Disla et al. (2014) provide an all-encompassing review of the applicability of vis–NIR and MIR spectroscopy for the prediction of various soil properties such as texture, aggregation and soil water content. Stenberg et al. (2010) provide a similar account but focus specifically on vis–NIR soil spectroscopy. It is recommended that readers review those contributions and the references cited therein to appreciate not only the breadth and variation of studies that have been undertaken, but also the comparative assessment of their model performances.

In general, MIR spectroscopy outperforms vis–NIR for most soil physical properties. Soil spectral models based on properties like clay, sand and water content using either vis–NIR or MIR spectra are generally comparable ($R^2 = 0.7$–0.8). Soil aggregation as measured by wet sieving also been reported as being well-predicted, for example, mean aggregate diameter and mean weight diameter with R^2 values between 0.66 and 0.79. Despite seemingly good modelling outcomes reported for soil physical properties mentioned above, there are also corresponding examples where outcomes have not been so good such as bulk density and water retention at field capacity. There are many potential causes for this variability, with some stemming from the data used and the extent of soil variability captured, to differences in models and how the data are processed and handled. However, the more probable reason is that soil spectral calibrations of soil physical properties are really about the utilization of variables related to these properties that are known to be well-predicted, such as clay, soil carbon, and mineralogy. This co-relationship is acknowledged in the reporting of very good results in the prediction of COLE and Atterberg limits with vis–NIR spectroscopy and reported R^2 of 0.84 and 0.48–0.71, respectively (Rehman et al., 2020). Utilizing both soil spectroscopy and pedotransfer functions together, i.e. soil spectral inference (McBratney et al., 2006), is seen as a more appropriate way to estimate some soil physical properties, particularly those based on soil pore–space relationships, e.g. volumetric water retention, hydraulic conductivity and penetration resistance, and bulk density. This somewhat echoes the conclusions from Minasny et al. (2008) that soil physical properties based on soil pore space relations are not in general well predicted directly with infrared spectroscopy, therefore more focus should be given to soil spectral inference systems to improve what can be predicted well and use those estimates in existing or, newly calibrated pedotransfer functions.

Soil chemical properties

Near- and mid-infrared spectroscopy is mostly used to predict soil chemical attributes. These properties affect soil biological activity and nutrient dynamics in the soil and have been used as soil quality assessment indicators. Soil chemical properties include pH, organic carbon, nitrogen, nutrients, as well as cation exchange capacity.

Numerous studies had been carried out to measure soil carbon. This interest is due to the recognition that storing additional carbon in the soil serves as one of the options to mitigate anthropogenic climate change. Soil organic carbon is a component of soil organic matter that affects the physical, chemical, and biological properties of soil, influencing soil structure, water and nutrient retention. In addition to total carbon, organic and inorganic soil carbon predictions were also found to be successful, with R^2 ranging from 0.94–0.98 (McCarty et al., 2002). The use of infrared spectroscopy for the prediction of various carbon fractions, including labile and recalcitrant carbon fractions ($R^2 = 0.89$–0.97), had been proven to be successful (Zimmermann et al., 2007). In the MIR range, organic carbon may also be identified by absorption peaks at approximately 2920 and 1230 cm^{-1} (Zimmermann et al., 2007), while carbonates may be identified by strong and numerous absorption bands at 1450, 1060, 880 and 700 cm^{-1} (Clark et al., 1990) (as shown in Fig. 2). In the NIR range, organic carbon can be related to absorption peaks at approximately 600, 1910 and 2100 nm, while carbonates may be identified by peaks at 2300–2350 nm (Clark et al., 1990).

Although soil pH is non spectrally active, it is one of the soil properties that is predicted quite well ($R^2 = 0.57$ 0.79) because of its relationship with exchange sites related to clay minerals and organic matter (Soriano-Disla et al., 2014).

Infrared spectroscopy has been used successfully to predict exchangeable cation related properties. While cation exchange capacity (CEC) is not a spectrally active component, it has a strong correlation with clay mineralogy and soil organic matter content (Stenberg et al., 2010), particularly absorbance of peaks near 2926–2850 cm^{-1}, and carboxylate species 1633–1413 cm^{-1}. Greater prediction accuracy for certain cations (Ca^{2+}, Mg^{2+}, and K$^+$) with R^2 of 0.70–0.84 compared to Na$^+$ ($R^2 = 0.32$) was most likely because those cations are dominant on the exchange sites.

Infrared spectroscopy has also been used to predict nutrients in the soil, particularly the nitrogen (N), phosphorus (P), and potassium (K) concentrations in the soil. However, most studies have shown that only the total amount of nutrients can be well predicted as this relates to soil minerals and organic matter. Available nutrients cannot be well-predicted because they are determined by extraction and solution from soil. Integration of the prediction of plant nutrients and fertilizer recommendations maximizes the nutrient management in the agroecosystems.

Soil biological properties

The correlation of soil biological properties and organic carbon enables the biological properties to be successfully predicted using spectral features. Infrared spectroscopy has been used for the rapid assessment of a range of biological properties, including microbial biomass, soil respiration, microbial groups, and soil enzymatic activities. No spectral contribution is expected from the microbial biomass because it represents roughly less than 5% of the total organic carbon content (Janik et al., 1998). Successful predictions of biological properties were most likely due to the correlation of biological properties with organic matter or other soil properties that have spectral signatures.

Chang et al. (2001) obtained R^2 of 0.6 for biomass C using NIR. Meanwhile, Janik et al. (1998) and Ludwig et al. (2015) reported greater accuracy for the microbial biomass C predictions with MIR ($R^2 = 0.69$ and 0.79 respectively). In addition to biomass C, biomass N ($R^2 = 0.79$) and P ($R^2 = 0.62$) were also determined with moderate accuracy by Ludwig et al. (2015).

A better prediction was reported by Chang et al. (2001) using NIR for microbial respiration than the soil microbial biomass with R^2 of 0.72, 0.66, and 0.82 for mineralizable N, total respiration rate and basal respiration rate (BSR), respectively. van Groenigen et al. (2003) reported greater accuracy for mineralizable N using NIR ($R^2 = 0.46$) compared to MIR ($R^2 = 0.21$) in flooded soils. Similarly, Ludwig et al. (2015) reported an acceptable accuracy of $R^2 = 0.63$ for BSR using MIR.

Microbial biomarkers have also been utilized as a biological indicator of change in the microbial community because the size of microbial biomass itself is not sufficient to explain the effect of certain management regimes. Zornoza et al. (2008) successfully predicted several variables related to soil microbial groups based on PLFA biomarkers: bacteria ($R^2 = 0.93$), G+ bacteria ($R^2 = 0.91$), actinomycetes ($R^2 = 0.92$), VAM fungi ($R^2 = 0.91$) and PLFAs ($R^2 = 0.91$).

The presence of soil enzymes can be linked to soil physico-chemical properties. Although 500 enzymes might be present, only a handful have been identified and measured to date. Reeves et al. (2000) achieved an R^2 within the range of 0.8 with NIR to predict four enzymes (dehydrogenase, phosphatase, arylsulfatase and urease).

Soil contaminants

Contaminants include a wide range of natural and synthetic materials that can be found in soil. Given that exposure to contaminants leads to detrimental effects on human health and ecosystems, there is a need to develop rapid, low-cost methods to assess the spatial and temporal variation of soil contaminants. The application of spectroscopy in detecting contaminants is beneficial for rapid screening to determine whether contamination exceeds a certain threshold and if remediation is required. The prediction of organic contaminants can be directly related to the absorption peaks in the spectra because of sensitivity of the IR method to organic materials. Nonetheless, these absorbances of the organic contaminants overlapped with other organic and inorganic components that exist in the soil throughout the electromagnetic spectrum. Several examples of organic pollutants include petroleum hydrocarbons, microplastics and pesticides.

Absorbance peaks at 2222–2439 nm and 3000–2600 cm^{-1} in the NIR and MIR region have been used to predict the total petroleum hydrocarbon concentration (Forrester et al., 2013) with high accuracy. The peaks in the MIR region are related to two —CH$_3$ peaks (2950 and 2730 cm^{-1}), which were absent in the natural soil organic matter absorption —CH$_2$ region at 2930–2850 cm^{-1}. A similar range in the NIR region at 2250–2350 nm was reported by Chakraborty et al. (2010) for the prediction of total petroleum hydrocarbon, which achieved good accuracy in moist intact soil. The absorption at 2200 nm is due to the O—H bonds in the clay minerals, such as kaolinite, illite, and smectite, while absorptions between 2250 and 2450 nm were associated with the C—H bonds in the organic matter and carbonates (Clark et al., 1990).

Absorptions between 1130 and 1720 nm in the NIR range have been reported in various studies (e.g. Sato et al., 2003) in the identification of plastic polymers, such as polyethylene terephthalate (PET) and low density polyethylene (LDPE). Absorbance at 1210, 1420 and 1730 nm is due to C—H (CH$_2$) first and second overtones and C—H deformation (Sato et al., 2003), while absorbance at 1660 nm is related to overtone stretching of aromatic C—H groups which is present in PET but not LDPE.

Hermansen et al. (2020) successfully utilized NIR for predicting glyphosate (an active ingredient in herbicides) sorption in soil ($R^2 = 0.93$). Bengtsson et al. (2007) also successfully utilized NIR as a tool to predict other pesticide sorption in soil, particularly lindane ($R^2 = 0.85$) and linuron ($R^2 = 0.84$), despite no specific wavelengths being identified.

Although inorganic contaminants, such as heavy metals are not spectrally active, diagnostic screening is made possible due to the relationship with iron oxides, clay, and organic matter. Gholizadeh et al. (2015) using the vis–NIR range detected heavy metal

concentrations (Cu, Pb, Zn and Cd), which had low to moderate to correlations with clay and/or organic matter content, whereas Mn did not have a dependence on clay and organic matter. Similarly, Kemper and Sommer (2002) achieved good R^2 accuracy for the prediction of As (0.84), Fe (0.72), Hg (0.96), Pb (0.95) and Sb (0.93).

Lower cost - miniaturized spectrometer

Lower cost and miniaturized spectrometers with a limited NIR spectral range are now widely available and can potentially enable NIR measurements of soil properties to be used more broadly. For example, a miniaturized Michelson interferometer can be built using monolithic micro-electro-mechanical systems (MEMS). A case study in Australia demonstrates that a miniaturized spectrometer operating at 1250–2500 nm provides comparable accuracy with the complete vis–NIR spectrometer (400–2500 nm) in predicting soil pH, CEC and exchangeable Ca and Mg ($R^2 > 0.63$–0.78), and slightly less accurate predictions of total carbon, sand, and clay ($R^2 > 0.70$–0.80) (Tang et al., 2020).

Conclusions

Soil spectroscopy is a viable measurement technology with untapped possibilities. Soil reflectance in the visible-, near- and mid-infrared offers rapid, cheap, and relatively accurate predictions for several soil properties. A single infrared spectrum can offer the prediction of numerous soil properties. The MIR spectra have been shown to provide signatures that can be used to derive many physical, chemical and biological properties. The mechanisms for the prediction of soil properties are due to the direct and indirect interaction between the soil properties and the organic matter and mineral components of soil. Infrared can predict fundamental soil properties related to the soil matrix or solid constituents and soil solution concentrations in equilibrium with the solid phase. Soil properties that do not have active spectral components can be predicted if they are correlated to the mineral or organic components of the soil. However, soil spectroscopy has been mainly used in research studies and has not been widely used as a routine analysis method in agriculture.

References

Bengtsson S, Berglöf T, and Kylin H (2007) Near infrared reflectance spectroscopy as a tool to predict pesticide sorption in soil. *Bulletin of Environmental Contamination and Toxicology* 78(5): 295–298.

Bowers SA and Hanks RJ (1965) Reflection of radiant energy from soils. *Soil Science* 100(2).

Chakraborty S, Weindorf DC, Morgan CLS, Ge YF, Galbraith JM, Li B, and Kahlon CS (2010) Rapid identification of oil-contaminated soils using visible near-infrared diffuse reflectance spectroscopy. *Journal of Environmental Quality* 39(4): 1378–1387.

Chang CW, Laird DA, Mausbach MJ, and Hurburgh CR (2001) Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal* 65(2): 480–490.

Clark RN, King TVV, Klejwa M, Swayze GA, and Vergo N (1990) High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research - Solid Earth and Planets* 95(B8): 12653–12680.

Clark RN, Swayze GA, Livo KE, Kokaly RF, Sutley SJ, Dalton JB, McDougal RR, and Gent CA (2003) Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *Journal of Geophysical Research, Planets* 108(E12).

Forrester ST, Janik LJ, McLaughlin MJ, Soriano-Disla JM, Stewart R, and Dearman B (2013) Total petroleum hydrocarbon concentration prediction in soils using diffuse reflectance infrared spectroscopy. *Soil Science Society of America Journal* 77(2): 450–460.

Gadsden JA (1975) *Infrared Spectra of Minerals and Related Inorganic Compounds.* London: Butterworths.

Gholizadeh A, Borůvka L, Saberioon MM, Kozák J, Vašát R, and Němeček K (2015) Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features. *Soil & Water Research* 10(4): 10.

Hermansen C, Norgaard T, Wollesen de Jonge L, Moldrup P, Müller K, and Knadel M (2020) Predicting glyphosate sorption across New Zealand pastoral soils using basic soil properties or vis–NIR spectroscopy. *Geoderma* 360: 114009.

Herschel W (1800) Observations of the powers of the prismatic colours to heat and illuminate objects. *Philosophical Transactions of the Royal Society of London* 90: 225.

Hunt JM, Wisherd MP, and Bonham LC (1950) Infrared absorption spectra of minerals and other inorganic compounds. *Analytical Chemistry* 22(12): 1478–1497.

Janik LJ, Merry RH, and Skjemstad JO (1998) Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture* 38(7): 681–696.

Kemper T and Sommer S (2002) Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. *Environmental Science & Technology* 36(12): 2742–2747.

Ludwig B, Sawallisch A, Heinze S, Joergensen RG, and Vohland M (2015) Usefulness of middle infrared spectroscopy for an estimation of chemical and biological soil properties – Underlying principles and comparison of different software packages. *Soil Biology and Biochemistry* 86: 116–125.

Malone BP, Hughes P, McBratney AB, and Minasny B (2014) A model for the identification of terrons in the Lower Hunter Valley, Australia. *Geoderma Regional* 1: 31–47.

McBratney AB, Minasny B, and Viscarra Rossel R (2006) Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. *Geoderma* 136(1): 272–278.

McCarty GW, Reeves JB, Reeves VB, Follett RF, and Kimble JM (2002) Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society of America Journal* 66(2): 640–646.

Minasny B, McBratney AB, Tranter G, and Murphy BW (2008) Using soil knowledge for the evaluation of mid-infrared diffuse reflectance spectroscopy for predicting soil physical and mechanical properties. *European Journal of Soil Science* 59(5): 960–971.

Ng W, Minasny B, Montazerolghaem M, Padarian J, Ferguson R, Bailey S, and McBratney AB (2019) Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma* 352: 251–267.

Reeves JB, McCarty GW, and Meisinger JJ (2000) Near infrared reflectance spectroscopy for the determination of biological activity in agricultural soils. *Journal of Near Infrared Spectroscopy* 8(3): 161–170.

Rehman HU, Arthur E, Tall A, and Knadel M (2020) Estimating coefficient of linear extensibility using vis–NIR reflectance spectral data: Comparison of model validation approaches. *Vadose Zone Journal* 19(1): e20057.

Robinson NJ and Kitching M (2016) The 3D distribution of phyllosilicate clay minerals in western Victoria. *Geoderma* 284: 152–177.

Sato H, Shimoyama M, Kamiya T, Amari T, Šašiç S, Ninomiya T, Siesler HW, and Ozaki Y (2003) Near infrared spectra of pellets and thin films of high-density, low-density and linear low-density polyethylenes and prediction of their physical properties by multivariate data analysis. *Journal of Near Infrared Spectroscopy* 11(4): 309–321.

Soriano-Disla JM, Janik LJ, Rossel RAV, Macdonald LM, and McLaughlin MJ (2014) The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews* 49(2): 139–186.

Stenberg B, Viscarra Rossel RA, Mouazen AM, and Wetterlind J (2010) Visible and near infrared spectroscopy in soil science. In: Sparks DL (ed.) *Advances in Agronomy*. vol. 107, pp. 163–215. Burlington, MA, USA: Academic Press.

Tang Y, Jones E, and Minasny B (2020) Evaluating low-cost portable near infrared sensors for rapid analysis of soils from South Eastern Australia. *Geoderma Regional* 20.

van Groenigen JW, Mutters CS, Horwath WR, and van Kessel C (2003) NIR and DRIFT-MIR spectrometry of soils for predicting soil and crop parameters in a flooded field. *Plant and Soil* 250(1): 155–165.

Zimmermann M, Leifeld J, and Fuhrer J (2007) Quantifying soil organic carbon fractions by infrared-spectroscopy. *Soil Biology and Biochemistry* 39(1): 224–231.

Zornoza R, Guerrero C, Mataix-Solera J, Scow KM, Arcenegui V, and Mataix-Beneyto J (2008) Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biology and Biochemistry* 40(7): 1923–1930.