3-27-2022

A global soil spectral calibration library and estimation service

Keith D. Sheperd

Rich Ferguson

David Hoover

Fenny van Egmond

Jonathan Sanderman

See next page for additional authors

Authors
Keith D. Sheperd, Rich Ferguson, David Hoover, Fenny van Egmond, Jonathan Sanderman, and Yufeng Ge

Soil Security 7 (2022) 100061

Contents lists available at ScienceDirect

Soil Security

journal homepage: www.sciencedirect.com/journal/soil-security

A global soil spectral calibration library and estimation service

Keith D. Shepherd[a,*], Rich Ferguson[b], David Hoover[b], Fenny van Egmond[c], Jonathan Sanderman[d], Yufeng Ge[e]

[a] Innovative Solutions for Decision Agriculture (iSDA), Rothamsted Campus, West Common, Harpenden. Hertfordshire AL5 2JQ, United Kingdom
[b] United States Department of Agriculture (USDA), Natural Resources Conservation Service (NRCS), Soil and Plant Science Division, National Soil Survey Center, which contains the Kellogg Soil Survey Laboratory (KSSL), Lincoln, NE, U.S
[c] ISRIC World Soil Information, Droevendaalsesteeg 3, 6708 PB, Wageningen, the Netherlands
[d] Woodwell Climate Research Center, 149 Woods Hole Road, Falmouth MA 02540, USA
[e] Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska 68583, USA

ARTICLE INFO

Keywords:
Mid-infrared spectroscopy (MIR)
Global soil spectral library
Soil property estimation
Standard Operating Procedures (SOPs)
Measurement error
Soil Information Systems (SIS)

ABSTRACT

There is growing global interest in the potential for soil reflectance spectroscopy to fill an urgent need for more data on soil properties for improved decision-making on soil security at local to global scales. This is driven by the capability of soil spectroscopy to estimate a wide range of soil properties from a rapid, inexpensive, and highly reproducible measurement using only light. However, several obstacles are preventing wider adoption of soil spectroscopy. The biggest obstacles are the large variation in the soil analytical methods and operating procedures used in different laboratories, poor reproducibility of analyses within and amongst laboratories and a lack of soil physical archives. In addition, adoption is hindered by the expense and complexity of building soil spectral libraries and calibration models. The Global Soil Spectral Calibration Library and Estimation Service is proposed to overcome these obstacles by providing a freely available estimation service based on an open, high quality and diverse spectral calibration library and the extensive soil archives of the Kellogg Soil Survey Laboratory (KSSL) of the Natural Resources Conservation Service of the United States Department of Agriculture (USDA). The initiative is supported by the Global Soil Laboratory Network (GLOSOLAN) of the Global Soil Partnership and the Soil Spectroscopy for Global Good network, which provide additional support through dissemination of standards, capacity development and research. This service is a global public good which stands to benefit soil assessments globally, but especially developing countries where soil data and resources for conventional soil analyses are most limited.

1. Introduction

Up-to-date information on soil properties and the ability to track changes in soil properties over time are critical for improving multiple decisions on soil security at various scales, ranging from global climate change modelling and policy to national level environmental and development planning, to farm and field level resource management. This need is important everywhere, but greatest in resource poor countries where soil information and resources are most limited and where policies for protecting soils to achieve soil sustainability and security are not well developed. Steady advances in digital soil mapping (Hengl et al., 2021; Minasny and McBratney, 2016; Searle et al., 2021; Wadoux and McBratney, 2021) are providing solutions for planning and precision agriculture but accuracy is generally limited by the spatial density and quality of ground observations used for training and the limits on the power of remote sensing covariates to predict spatial variation in soil properties. Conducting conventional soil analysis on large numbers of samples required for digital approaches is generally cost prohibitive and there is need for low-cost measurement of soil properties, both in the laboratory and in the field using proximal sensing (Wadoux and McBratney, 2021). This need extends beyond digital mapping to many other applications, such as scaling soil testing services to smallholder farmers, particularly in sub-Saharan African and South Asia, benchmarking and tracking soil health for regenerative agriculture, and soil carbon trading.

There is increasing evidence that the need for low-cost soil measurements can be partly met through the use of soil diffuse reflectance spectroscopy in the visible (Vis), near-infrared (NIR) and mid-infrared (MIR) ranges, providing a rapid and reproducible method for

* Corresponding author.
E-mail address: keith.shepherd@isda-africa.com (K.D. Shepherd).

https://doi.org/10.1016/j.soisec.2022.100061
Received 30 September 2021; Received in revised form 7 March 2022; Accepted 27 March 2022
Available online 1 April 2022

Abbreviations

KSSL	Kellogg Soil Survey Laboratory (KSSL) of the Natural Resources Conservation Service of the USDA
USDA	United States Department of Agriculture
GLOSOLAN	Global Soil Laboratory Network of the Global Soil Partnership
Vis	Visible
NIR	Near infrared
MIR	Mid infrared
LUCAS	The Land Use/Cover Area frame Statistical Survey of Europe
GSCLES	Global Soil Spectral Calibration Library & Estimation Service
FAO	Food & Agriculture Organization
GSP	Global Soil Partnership
SOC	Soil organic carbon
SIS	Soil Information System
OGC	Open Geospatial Consortium
GloSIS	Global Soil Information System

estimating soil properties (Bellon-Maurel and McBratney, 2011; Nocita et al., 2015; Shepherd and Walsh, 2007: Stenberg et al., 2010; Viscarra-Rossel et al., 2006; 2016). Benchtop MIR, NIR and VisNIR instruments are now in common use in soil laboratories (Benedetti and van Egmond, 2021). Soil spectroscopy has shown the ability to estimate a wide range of soil physical, chemical and biological properties (Barra et al., 2021; Dangal et al., 2019; Nocita et al., 2015; Stenberg et al., 2010; Terhoeven et al., 2010). Soil properties that relate directly to mineral and organic composition tend to be very well estimated (R^2 >0.8), for example soil organic and inorganic carbon, total soil nitrogen, exchangeable calcium, while other properties may be estimated to varying degrees of accuracy through indirect associations (e.g., extractable nutrients) (Towett et al., 2015). However, care should be taken to ensure that calibrations hold up adequately when soil properties are estimated solely due to correlation with properties that have primary associations with spectral absorption features.

The main requirement of soil spectroscopy is for calibration to reference soil property measurements for a given population of soils. This requires building databases of spectra (spectral libraries) that represent the soil diversity in a target geographical area in combination with a representative subset of soil samples measured with standard methods (i.e., reference samples) with known accuracies (McBratney et al., 2006; Shepherd and Walsh, 2007). The spectral and soil property reference data comprise a spectral calibration library. Calibration modelling of the relationships between the soil properties measured with standard methods and spectra is done using multivariate regression methods, such as partial least squares regression, or increasingly using machine learning methods, such as Random Forests, neural networks, deep learning, and ensemble models (Hengl et al., 2021; Ng et al., 2020b; Sila et al., 2016; Yang et al., 2021). With large spectral calibration libraries, there has been even greater success with local models that subset spectral nearest neighbours in spectral data space and develop an individual model for each sample for which an estimation is to be made.

Despite major advances in building spectral libraries at continental and national levels (see Section 2), a major limitation for wide deployment of soil spectroscopy is the lack of availability of consistent calibration libraries and soil property estimation models (Dangal et al., 2019; Gomez et al., 2020). The high cost of building calibrations, particularly of analysing large numbers of soil samples for a wide suite of properties (the reference data) that are representative of the region of interest and a lack of consistency and reliability of reference analyses present major obstacles. Comprehensive physical archives of recent samples are also rare. Where these do exist, often either the soil analytical data are old and inconsistent, or there is insufficient sample quantity or budget to allow new analysis of a suite of soil properties. These limitations present large obstacles especially for developing countries, where resources for collecting, storing and analysing soil samples to a high quality are scarce. The lack of calibration libraries is halting the utilization of the technology (Benedetti and van Egmond, 2021) by existing national soil spectroscopy laboratories in Africa (Soil Plant Spectral Diagnostics Lab, 2021) even though these countries stand to gain the most from deployment of soil spectral technology.

This paper describes a proposal for establishing a Global Soil Spectral Calibration Library and Estimation Service (GSCLES) to help overcome these obstacles. The concept has evolved through a partnership amongst several institutions[1] and is being fostered under the umbrella of the Global Soil Laboratory Network (GLOSOLAN) of the Global Soil Partnership of the Food & Agriculture Organisation of the United Nation (FAO) and Soil Spectroscopy for Global Good.[2] The initiative builds on the extensive existing MIR spectral calibration library of the Kellogg Soil Survey Laboratory (KSSL) of the Natural Resources Conservation Service of the United States Department of Agriculture (USDA) (See Section 4.2). KSSL prioritised MIR over other spectral ranges due to its high prediction accuracy for a wide range of soil properties (e.g., Janik et al., 2007; Ng et al., 2019, 2022a; Reeves, 2010). However, the service could be extended to include other spectral ranges.

The objectives of the GSCLES initiative are to:

1 Build a globally representative soil spectral calibration library (database) based on soil mid-infrared diffuse reflectance (MIR) spectra initially, with accompanying soil property reference data recorded in one gold-standard reference laboratory.
2 Provide a freely available and easy-to-use soil property estimation service based on the global spectral library using open-source models.
3 Support countries to contribute to the global spectral calibration library and use the soil property estimation service with local measured soil spectra.

2. Soil spectroscopy uptake

Soil spectroscopy is being increasingly taken up, including in Africa and Asia. For example, World Agroforestry, at the request of national programs, has helped establish 30 benchtop spectrometers in 16 Africa countries and provided training support (Soil Plant Spectral Diagnostics Lab, 2021). Four Sub-Saharan African countries have deployed MIR spectroscopy in the establishment of their national soil information systems under the Africa Soil Information Service project (National Soil Services, 2021), joined recently by the Rwanda Soil Information Service (RwaSIS, 2021). A new initiative is underway to establish a Soil Information System for Africa based on soil spectral technology including the Forum for Agricultural Research in Africa and several African national programs (Soils4Africa, 2021). In South Asia, the Government of India has approved the deployment of soil spectroscopy (ICAR, 2020). The National Soil Conservation Service (NRCS) has established benchtop MIR at the Kellogg Soil Survey Laboratory and at 15 regional centres. MIR spectroscopy is also used extensively in Australia for digital soil mapping (Searle et al., 2021). Additionally, there has been rapid development of portable spectral devices that show increasing potential

[1] FAO under the Global Soil Laboratory Network of the Global Soil Partnership, United States Department of Agriculture's Natural Resources Conservation Service, World Agroforestry (CIFOR-ICRAF), Innovative Solutions for Decision Agriculture (iSDA), ISRIC World Soil Information, Woodwell Climate Research Center, University of Nebraska, and the University of Sydney.

[2] https://soilspectroscopy.org

for field use (Minasny et al., 2009; Ng et al., 2020a; Tang et al., 2020). Commercial soil testing laboratories are now adopting infrared spectroscopy for some soils analyses (e.g., CNLS, 2021) and NIR portable scanning services have been in commercial use for several years (e.g., AgroCares, 2021).

Soil spectral libraries are being developed, at national level (e.g., Grinand et al., 2012; Viscarra Rossel and Webster 2012), regional level (e.g., Stevens et al., 2013; Vågen et al., 2020) and global level (Dangal et al., 2019; Brown et al., 2006; Terhoeven-Urselmans et al., 2010; Viscarra Rossel et al., 2016). Based on 4184 soil samples from 37 countries, although only 416 samples were from outside of the USA, Brown et al. (2006) estimated that 5.2×10^9 carefully selected calibration samples would be required to span the known global soil compositional space. Ten years later, the largest and most diverse current available global Vis-NIR soil spectral library was reported by Viscarra Rossel et al. (2016). This global dataset consists of 23,631 soil spectra from 92 countries, and all spectra were voluntarily contributed by around 45 soil scientists and researchers from 35 countries and institutions. The study showed that the global spectral library could estimate soil organic carbon (R^2 0.89), extractable Fe (R^2 0.86), calcium carbonate content (R^2 0.77), CEC (R^2 0.73), clay (R^2 0.71), and silt (R^2 0.68) contents, and pH (R^2 0.62). Terhoeven et al. (2010), using a globally distributed MIR spectral library, demonstrated predictions for spatially independent validation samples for pH (R^2 0.89), cation exchange capacity (R^2 0.82), and organic C content (R^2 0.77).

The most complete and consistent continental Vis-NIR soil spectral library has been developed in the framework of the European Land Use/Cover Area frame Statistical Survey (LUCAS) during which ~20,000 geo-referenced top-soil samples were collected to assess the state of the soils across Europe (Stevens et al., 2013). Twelve chemical and physical properties were estimated. The Africa Soil Information Service conducted the first stratified random sampling of sub-Saharan Africa soils, providing a MIR calibration library of over 1900 samples (Towett et al., 2015; Vågen et al., 2020>). Soil properties that were well estimated (R^2>0.8) included organic carbon, total nitrogen, pH, and Mehlich-3 Al and Ca. Johnson et al. (2019) assessed soil fertility properties in rice fields based on a large regional MIR soil spectral library from sub-Saharan Africa. A total of 2845 topsoil samples from 42 sites in 20 sub-Saharan African countries were collected from three different target rice production systems (irrigated lowland, rainfed lowland and rainfed upland). Thirty soil properties were measured by conventional wet chemistry analysis. Their results suggested that NIR-MIR spectroscopy can offer an alternative to conventional wet chemistry methods for assessing those soil fertility properties in rice fields and could also be used to develop soil fertility indicators.

One of the most comprehensive national (Vis-NIR-MIR) soil spectral libraries was developed by the French national soil quality monitoring network (Arrouays et al., 2002). The French soil spectral library samples were collected from regular, nationwide 16 km grids which consist of over 2200 sites and 3800 samples. Recently, Gomez et al. (2020) successfully used French MIR soil dataset (2178 topsoil samples) to calibrate soil inorganic carbon and soil organic carbon, then tested the models using 96 topsoil samples from a Tunisian MIR dataset. This work highlighted the very high applicability of MIR for soil inorganic carbon determination and the robustness of soil inorganic carbon prediction models, even when the training and testing set come from different pedologic and climatic contexts. Australia has extensive VNIR (Viscarra Rossel and Webster, 2012) and MIR spectral libraries (CSIRO, 2021) which have been widely used for digital soil mapping and routine soil analytical purposes. Brazil has developed a VNIR spectral library (Dematte et al., 2019) consisting of 39,284 soil samples from all 26 states and a MIR spectral library consisting of 4309 soil samples from different depths and from across the country (Mendes et al., 2022). The GSCLES aims to build on these successes and help enable wider uptake of soil spectroscopy.

3. Main components of the GSCLES

The purpose of soil spectroscopy is to model the relationship between soil properties measured by standard/reference methods (reference measurements) and spectral variables so that soil properties can be rapidly and inexpensively estimated from spectra for new samples. The main purpose of the GSCLES (Fig. 1) is to enable a user in any locality to upload soil spectra, recorded using an approved standard operating procedure (SOP)), and obtain estimates for a suite of soil properties together with uncertainty estimates. A fundamental principle of the GSCLES is that the spectral and reference measurements that provide the foundation for the calibration library and estimation service are performed in one primary laboratory with well-established and rigorous quality control protocols. To achieve this, the primary laboratory needs access to a physical sample archive of globally diverse soil samples, which can be further developed over time.

The estimation service is an Application Programming Interface (API) that provides spectral quality checks, spectral distance measures and calibration models. The versioned calibration models are available online with an option provided for download of local offline versions. The degree to which a submitted spectrum is an outlier to the calibration library provides an indication of the value of that sample for improving the overall calibration model. Users may then submit high value samples to the central laboratory for spectral and reference analysis, which are then added into the global calibration library. In this way the value of the estimation service iteratively improves as new samples are submitted and characterised and the calibration models are updated.

The main value of the GSCLES is in the efficiency it creates by: (i) reducing the amount of investment a country or local laboratory needs to make in reference measurements for constructing spectral calibrations, (ii) allowing countries to leverage a much larger and more diverse spectral calibration library than they would have access to by themselves, and (iii) simplifying the process of building, maintaining, and deploying soil spectral libraries and calibrations. The following sections provide more details on each component of GSCLES.

4. Calibration library

4.1. Quality of soil reference & spectral data

For a soil spectroscopy estimation service, the obtained prediction error such as the Root Mean Square Error of Prediction (RMSEP) for unknown samples is estimated from the calculated differences between the prediction values and "known" reference values of the validation set. A major source of model error is the precision and accuracy of the reference data (i.e., conventionally measured with standard methods) to which spectra are modelled during calibration. Laboratory reference analysis of soil chemical, biological or physical properties contain measurement error derived from cumulative errors through the chain of sample preparation, sub-sampling, and instrument readings (BIPM et al., 2008; Taylor, 1997). Many laboratories routinely monitor repeatability (precision) over time, using internal standards, and absolute error (bias) using externally supplied standards, such as those supplied by WEPAL (2021). However, variability in reference data remains a big obstacle to the development of spectral calibration libraries due to combined effect of differences in the analytical methods and standard operating procedures (SOPs) used, and the large inter-laboratory variation even when the same methods and SOPs are used. The inter-lab variability problem is illustrated by the results of a global ring test conducted by GLOSOLAN (Hartmann and Suvannang, 2020). A set of control soil samples were sent out to 120 laboratories around the world together with a standard operating procedure (SOP) for conducting the soil test. The results returned from 82 laboratories show extreme variation (Fig. 2). The spread of the values obtained across different laboratories spans the typical range of values obtained in agricultural soils. Such large errors will obviously significantly degrade the performance of a spectral

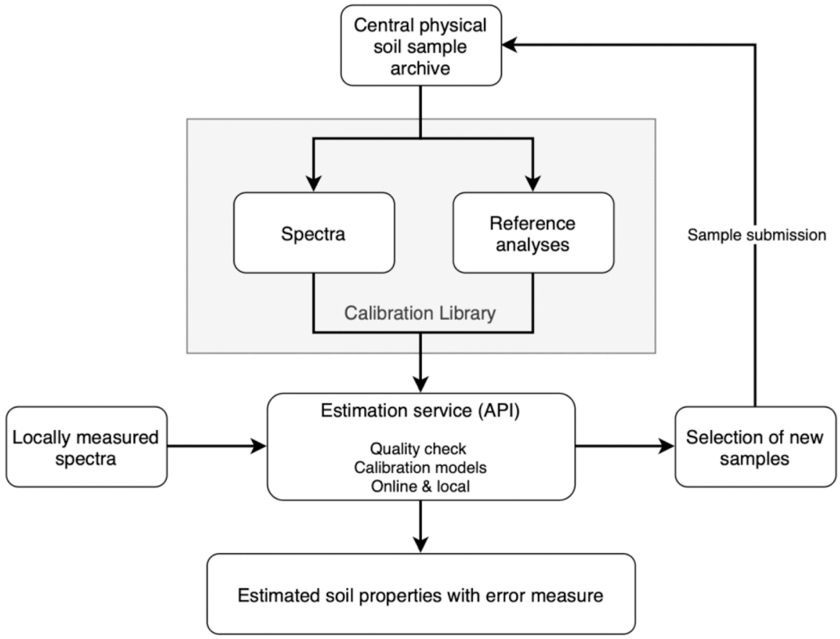

Fig. 1. Conceptual flow chart of Soil Spectral Calibration Library and Estimation Service.

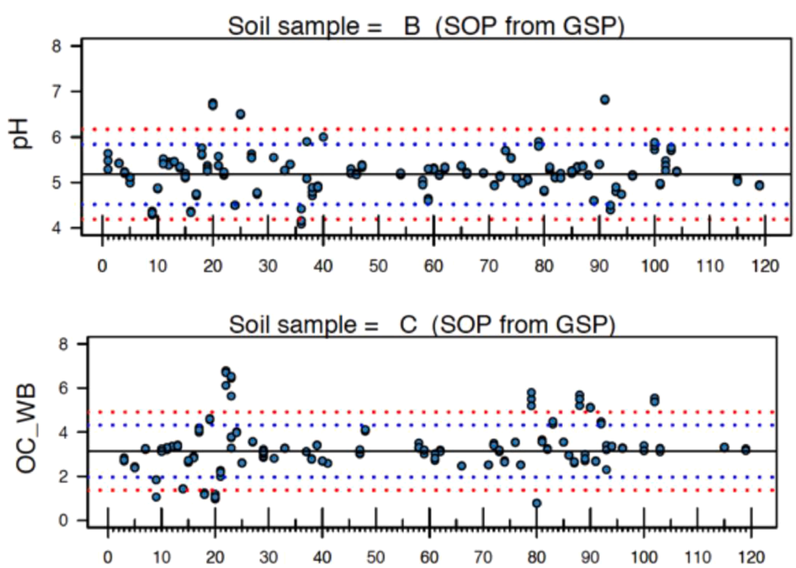

Fig. 2. Results of triplicate analysis of control samples returned by 82 participating laboratories using the same standard operating procedure for soil pH (mean 5.2, SD 0.25) and organic carbon by the Walkley-Black method
Source: Hartmann and Suvannang (2020)

estimation model (Aastveit and Marum, 1991; Faber and Kowalski, 1997; Mark et al., 1989; Sørensen, 2002). The importance of high quality and consistent reference measurements coupled with the difficulty of achieving them has been a major motivation for the development of the GSCLES centred on one 'gold standard' laboratory. With high quality reference data from a large, representative, and high-quality global spectral library, random errors associated with the reference measurements can be reduced due to averaging during the spectral calibration regression (Abrams et al., 1987; Difoggio, 1995). This means that for some soil properties, the prediction error can approach the error associated with the standard laboratory method. However, this does not apply to systematic errors, such as those observed amongst different laboratories.

Like standard methods, soil spectroscopy is also prone to errors due to differences in: (i) spectroscopy equipment and SOPs used (Knadel et al., 2013; Ge et al., 2011; Pimstein et al., 2011); (ii) soil preparation and sub-sampling (Ben-Dor et al., 2015); (iii) the temperature and humidity of the environment (Challibrat et al., 2019); and (iv) the date of measurement of the spectral data relative to the reference data due to change in soil properties during storage. Slight variations in instrument set up and scanning conditions can lead to quite significant differences in absorption features (Ge et al., 2011) ultimately resulting in spurious predictions. For example, Gholizadeh et al. (2021) compared estimations of soil organic carbon (SOC) for spectra recorded on different Vis-NIR spectrometers within and across laboratories. Merging raw reflectance spectra from multiple spectrometers and laboratories resulted in poor model performance ($R^2 = 0.48$, RMSE = 0.33%). However, use of an internal standard and spectral pre-processing minimized variations between scanning environments enabling the merging of the spectral libraries and significantly improved model performance (R^2 = 0.70, RMSE = 0.25%). A major thrust of the GSCLES initiative is to develop and disseminate SOPs, standards and tools and enhance capacity to minimise these errors (see Section 7).

A final source of error is the modelling approach chosen, which includes the choice of modelling technique (e.g., partial least squares regression, memory-based learning, neural networks, random forests, Cubist, etc.), and details such as wavelength range selections and data pre-processing methods. There are many modelling approaches and options to choose from and these topics are an ongoing area of research. Which method is best depends on factors such as available computational capacity and the level of error that is acceptable for the particular application for which the model is intended to serve.

4.2. The KSSL spectral calibration library

To address some of the problems described above that are related to variable reference and spectral data, the GCSLES uses one global reference laboratory for building the soil spectral calibration library. The USDA-NSSC Kellogg Soil Survey Laboratory (KSSL) has successfully demonstrated the use of mid-infrared (MIR) spectral calibrations for key soil properties across a very wide range of soil types across the continental USA (Table 1; Dangal et al., 2019; Sanderman et al., 2020). Ng et al. (2022a) based on 45,000 samples from the KSSL library demonstrated that MIR could infer 50 soil properties with high accuracy (R^2 centroid 0.76- 0.88) and 44 properties with moderate accuracy (R^2 centroid 0.59). They concluded the properties estimated can be used to evaluate a range of soil functions, including food production, carbon storage, water storage, nutrient cycling, and habitat function. The foundation for the high performance of the calibrations based on the KSSL library is the consistency in methods and standard operating procedures and consistency in quality control of the reference methods sustained over many years. This is coupled with consistency from the use of primary MIR instrument and associated SOP. While other laboratories with gold standard reference analyses and large spectral libraries could be considered for this role, this initiative selected the KSSL due to the broad compositional diversity of its soil archive, its quality as shown in the spectroscopy literature (e.g., Dangal et al., 2019; Ng et al., 2022a., Sanderman et al., 2020; Seybold et al., 2019), and its open and unencumbered data sharing policy according to U.S. law (https://www.congress.gov/bill/115th-congress/house-bill/1770).

The current calibration library of over 80,000 soil samples, which is still growing, represents a significant coverage of global soil variation, including 292 globally distributed samples from ISRIC World Soil Information. KSSL has more than 120,000 additional analysed samples by standard methods, including international samples, still to be scanned using MIR (Table 2, Appendix 1). Therefore, the spectral calibration library proposed under GSCLES is based on the KSSL laboratory.

Soils are compositionally highly variable across the globe, and it is the extent to which the variability of soil properties for a given region is captured by the predictive model that determines its suitability for reliably estimating soil properties of samples collected in that region. If only part of regional variability is captured in the calibration, the model is not expected to perform as well for under-represented samples compared with a model that captures all regional soil variability. This is one of the impetuses for developing a global calibration library, as it will likely capture a wide range of spectral variability of global soil resources for improved model performances at local scales, possibly in addition to the use of high quality local libraries, if available.

The use of a wide representative spectral library, in addition to improving model performance, also increases model efficiency due to the fact that spectral properties of soils, for example with similar mineralogy, from one part of the world can be similar to those from other parts of the world. For example, Fig. 3 shows the Africa Soil Information Service MIR spectral library overlaid in principal components space on the KSSL spectral library. There is considerable overlap suggesting that even the existing KSSL library could provide reasonable soil property estimates for much of sub-Saharan Africa. Furthermore, samples that fall outside the KSSL space can be flagged as of high value for

Table 1
R-square values for spectrally estimated versus reference soil properties for a 20% hold-out validation set from the KSSL laboratory using mid-infrared spectroscopy and memory-based learning for a diverse set of soils from the USA (Dangal et al., 2019; Sanderman et al., 2020).

Property	n	Units	Min	25th percentile	Mean	75th percentile	Max	R^2	RMSEP
Physical indicators									
Water retention (1/3 bar)	10,996	wt%	1.6	18.4	27.9	30.4	2124.9	0.83	5.98
Water retention (15 bar)	27,116	wt%	−6.4	6.6	15	17.3	354.9	0.94	3.10
Bulk density (clod)	10,553	g cm-3	0.47	1.2	1.35	1.5	2.1	0.81	0.10
Bulk density (core)	7003	g cm-3	0.08	0.5	0.93	1.3	2.2	0.80	0.21
Sand	34,912	wt%	0	11.5	38.6	61.8	100	0.96	5.72
Silt	34,913	wt%	0	22.8	38.1	53.7	94.5	0.92	6.23
Clay	34,913	wt%	0	9.2	22.5	32.4	96.1	0.96	2.83
Aggregate stability	1912	wt%	0	10	38.2	65	100	0.71	15.2
Al (DCB extract)	22,892	wt%	0	0.04	0.19	0.21	4.2	0.97	0.04
Fe (NH4OAC extract)	21,318	wt%	0	0.1	0.44	0.57	6.7	0.81	0.22
Chemical indicators									
Cation exchange capacity	39,600	cmol(+) kg-1	0	8.2	22.6	26.9	584.6	0.98	3.12
Exchangeable Ca	38,068	cmol(+) kg-1	0	4.5	25.1	31.7	507.3	0.94	6.59
Exchangeable Mg	38,122	cmol(+) kg-1	0	1.1	5.9	7	172.6	0.88	1.88
Exchangeable K	37,702	cmol(+) kg-1	0	0.2	0.7	0.8	32.3	0.83	0.34
Exchangeable Na	16,259	cmol(+) kg-1	0	0.1	11.5	3.7	868.4	0.94	6.05
Base saturation	14,658	%	1.5	32.9	59.3	87.5	99	0.86	10.1
EC (paste)	6400	dS m-1	0	0.7	5.4	4.5	247	0.82	4.26
EC (water)	614	dS m-1	0	0.1	1.6	2.4	25	0.84	0.63
pH (water)	37,123	–	2.7	5.5	6.3	7.6	10.7	0.88	0.36
CaCO3	19,171	wt%	0	0.2	7.9	10.7	105.8	0.98	1.41
Biological indicators									
Organic carbon	53,673	wt%	0	0.4	7.7	4.9	65.6	0.99	0.64
Total nitrogen	51,641	wt%	0	0.1	0.5	0.5	41.9	0.97	0.13
Plant available nutrients									
P (Bray-1)	3527	mg kg-1	0	2.3	26.5	30.2	1436.7	0.74	19.9
P (Olsen)	10,000	mg kg-1	0	1.9	13.7	16.5	223.6	0.72	13.7
P (Mehlich3)	19,139	mg kg-1	0	2.6	30.8	36.6	825.2	0.70	34.6
K (Mehlich3)	952	mg kg-1	0	71.7	156.1	212.5	1150.9	0.72	50.5

Table 2
Number of samples in major soil groups and horizons in the KSSL spectral library and additional archive samples.

Soil order	No of samples			Soil horizon	No of samples		
	Spectral library	**Archive**	**Total**		**Spectral library**	**Archive**	**Total**
Alfisols	5816	20,885	26,701	A	12,681	25,652	38,333
Andisols	1543	3037	4580	B	20,403	57,193	77,596
Aridisols	1675	10,522	12,197	C	6219	19,018	25,237
Entisols	2788	8720	11,508	E	1144	3912	5056
Gelisols	617	303	920	O	2603	3076	5679
Histosols	999	655	1654	R	25	364	389
Inceptisols	4818	11,770	16,588	Undefined	40,801	10,795	51,596
Mollisols	12,041	24,241	36,282				
Oxisols	23	1128	1151				
Spodosols	1949	4432	6381				
Ultisols	3184	7923	11,107				
Vertisols	886	4642	5528				
Undefined	47,537	21,752	69,289				
Total	83,876	120,010	203,886	Total	83,876	120,010	203,866

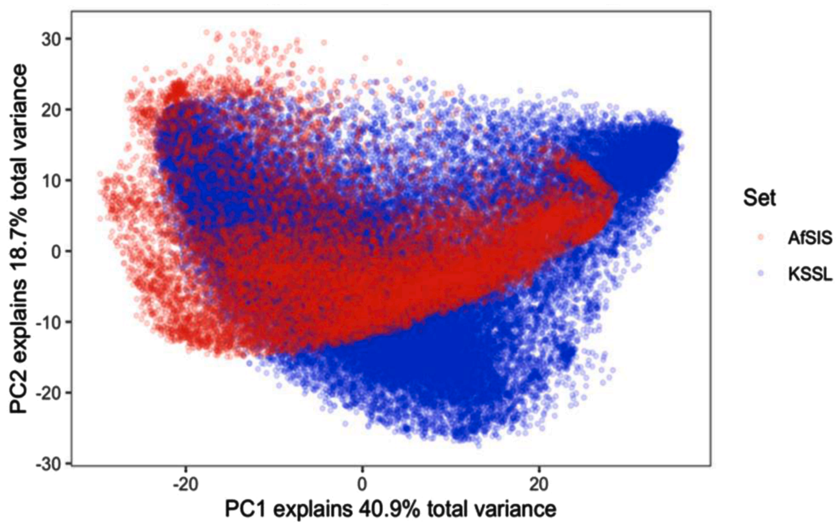

Fig. 3. Principal components score plot for first derivative MIR spectra for the Africa Soil Information Service library overlaid on the KSSL library. Data source: Vagen et al. (2020).

inclusion in the global library. To this end, countries are invited to submit a subset of nationally representative samples for analysis by KSSL to the GSCLES through the GLOSOLAN network. This will increase the GSCLES predictive power and provide a means to assess and improve local laboratory quality by comparing the results for both spectral and reference measurements. Padarian et al. (2019) have illustrated how transfer learning, a machine learning technique that transfers some of the rules learnt by the more general global models to a local domain, can enhance the use of global spectral libraries for local application, in future possibly also including high quality local libraries.

The global service is initially focused on MIR diffuse reflectance, which has shown the best spectral range for soil property estimation, however there is potential to extend the service to other spectral ranges at a later stage. The reference properties to be included in the spectral calibration library are listed in Table 3. There are plans to extend the calibrations to Vis-NIR benchtop and portable instruments over time.

5. Estimation service

Building a soil spectral library for a region of interest, for example at national level, and developing and applying appropriate statistical models to make spectral predictions of various soil properties requires a high level of specialised knowledge and skills and this poses a significant barrier to the adoption of soil spectroscopy (Benedetti and van Egmond, 2021). The proposed GSCLES is designed to remove much of this complexity, so that a user anywhere in the world could upload a set of soil spectra recorded locally using a prescribed SOP and obtain soil property estimates with uncertainty estimates.

While the core features of an estimation service can be quite simple, ensuring high quality predictions requires thoughtful development. Spectra must be compatible with the existing spectral library. First, spectra either need to be collected on an instrument and in a manner compatible with the library. Differences in sample preparation and instrument parameters can often be minimized by following standard operating procedures including routine use of reference materials (Ben Dor et al., 2015). Incompatibilities can also be minimized through judicious use of spectral pre-processing (Naes et al., 2017) or calibration transfer. Calibration transfer applies a spectral model developed from a primary instrument to a spectral dataset measured by a secondary instrument with statistically retained accuracy and precision (Pittaki--Chrysodonta et al., 2021). This requires scanning a set of standard samples on primary and secondary instruments, which is a limitation requiring exchange of physical samples. Seybold et al. (2019) illustrated direct use of spectra across different MIR instruments without the need for standardization when models were built using a subset of a large national library appropriate for the target region.

In addition, new spectra must fall within the feature space of the existing spectral library. Otherwise, predictions can be highly biased, particularly from machine learning algorithms (Dangal and Sanderman, 2020). There are several relatively simple statistical methods for

Table 3
Soil reference properties to be analysed for the global spectral calibration library.

	Property	Method code
1.	Total carbon, nitrogen, and sulfur	4H2a1–3a1
2.	Inorganic carbon (if appropriate)	4E1a1a1a1–2
3.	Organic carbon (calculated from total carbon and inorganic carbon)	4H2a + 4E1a1a1a1–2
4.	Gypsum	4E2b1a1a1–2
5.	pH: 1:1 water	4C1a2a1a-b1
6.	pH: 1:2 0.01-M calcium chloride	4C1a2a2a-b1
7.	pH: 1:1 1-N potassium chloride	C1a2a3a-b1
8.	pH: 1:50 1-N sodium fluoride (if appropriate)	4C1a1a1a-b1
9.	Cation exchange capacity, pH 7	4B1a1a1a1a-b1
10.	Ammonium acetate (pH7) exchangeable calcium, magnesium, potassium, sodium	4B1a1c1–4a-b1*
11.	1500 kPa water holding capacity	3C2a1a-b
12.	Dithionite-citrate extractable iron, aluminium	4G1b1–4a-b1*
13.	Ammonium oxalate extractable iron, aluminium	G2a1a1–5a-b1
14.	Clay, silt, sand	3A1a1a
15.	Exchangeable aluminium	4B3b1a1-b1*
16.	Mehlich III phosphorus	4D6a1a-b1
17.	Olsen phosphorous (if appropriate)	4D5a1a-b1
18.	Electrical conductivity; method	4F1a1a1a1
19.	Sodium adsorption ratio and exchangeable sodium percentage (derived quantities)	4F3b + 4F3a2

Note: Except as noted with an asterisk (*), method codes are from "Kellogg Soil Survey Laboratory Methods Manual, SSIR-42, v. 5, USDA-NRCS (2021)." *to-be-published in v. 6 of SSIR-42; pending publication of v. 6 in 2022, SOPs are available on request to: christopher.lee@usda.gov.

calculating whether a new spectrum should be considered an outlier (Hicks et al., 2015). Samples flagged as outliers are of high value to send to an analytical laboratory to further build out a wide and representative spectral library. Such samples should be included in the library before similar spectra can be used in the estimation service.

The most straightforward way to build an estimation service is to host pretrained models. Several multivariate and machine learning approaches applied to large diverse soil spectral libraries have been shown to produce reasonably accurate and precise predictions for a number of soil properties (Dangal et al., 2019; Ng et al., 2020b; Wijewardane et al., 2018). While precalculated "global" models can perform very well, "local" modelling approaches, where only the most relevant samples in the spectral library are used to build models on-the-fly, typically outperform the most advanced and sophisticated global models (Ramirez-Lopez et al., 2013) even after spiking the global models with local samples (Ng et al., 2022b). In an estimation service, the downside of local modelling approaches is that they can be computationally demanding, requiring a thoughtful approach to hosting the service. Regardless of model choice, any estimation service should provide predictions with estimates of uncertainty, especially when there are no local validation datasets. Numerous methods exist for estimating prediction level uncertainty depending on the modelling approach used (Dangal et al., 2019; De Vries and Braak, 1995).

GSCLES aims to provide options to users in terms of the trade-off between accuracy and computational resource requirements. Local modelling can be made available to those who have access to the computational resources required. We foresee global models being made available where resources are insufficient to permit local modelling. In addition to a cloud-based service, options will be explored that would enable users to download and run models on their own computer resources. In each case, uncertainty estimates will be provided and strict versioning protocols observed.

6. Growing the calibration library

Building spectral calibrations requires matching reference soil property data (i.e., measured with standard methods) with spectral data to represent the diversity of soils in participating regions. For the GSCLES to have a global predictive power it should therefore be extended with samples that cover the global diversity in soils. For this reason, countries are invited to submit a selected set of samples to the USDA-NSSC Kellogg Soil Survey Laboratory (KSSL). Guidelines for sample submission to KSSL will help to ensure consistency, practical utility and comparability of data, in order to improve the predictive power of the library, are proposed in Appendix 2.

Prior to sample submission, the sample submitter must accept that submitted and measured data and metadata become open and free public information according to USA law (Congress, 2017). USDA cannot omit the data or metadata from public view. Without exception, the results of all analysis conducted by the KSSL will be distributed without copyright restriction. Prior to inclusion of the data in the library and during the quality control process at KSSL, the sample submitter will be asked to review the measurement results as an additional quality check. It is encouraged that local laboratories analyse the samples via reference methods and spectra measurements for quality comparison and improvement and to allow for the development of transfer functions (see Section 8). The standard operating procedures for soil spectral measurements have been developed by the initiative and will be published soon. Continued efforts are needed to develop and disseminate standards for sample preparation, spectral measurements, spectral data storage and exchange.

In terms of data management, the GSCLES initiative aligns with the goals of the Global Soil Information System (GloSIS, 2021), established by the Global Soil Partnership and partners, and the two initiatives will likely be linked in the future. The GloSIS initiative aims to help countries structure and provide their soil data online using customised open source software for a multitude of applications. The initiative aims to stimulate data sharing and the availability of soil information for local, national, continental, and global decision making. Data is stored and shared through national, regional, and institutional nodes or soil information systems (SIS) connected to a central portal through Open Geospatial Consortium (OGC) webservices. This allows national customisation, control of data access and maintenance of the datasets, while increasing its findability and accessibility through the national and global portals. GloSIS envisages a distributed system design and open-source approach, where all linked SISs retain full control of their data. Soil data in GloSIS consists of point data such as soil descriptions, samples analysed by conventional/standard laboratory methods, and maps on various scales and topics, including soil spectra and soil spectral libraries. In the first phase (meta)data will be provided with a required minimum set of metadata. In the second stage, standardised data will be provided according to accepted vocabularies and ontology. This will facilitate seamless exchanges and combination of datasets, improving the efficiency for data analysis and soil mapping. The standardisation of soil data exchange, for both reference samples and associated spectral soil data, and the facilitation of a network of nodes that provide this data by the GloSIS initiative is relevant to this GCLES initiative because it will allow easier linkage of other soil spectral libraries for localised studies. This in turn will facilitate growing local predictive power, inclusiveness and the use of existing valuable data by allowing easier submission of spectra to the estimation service and translation of the result to local contexts.

7. Capacity development and sustaining the service

A growing realisation of the importance of improved soil information coupled with recent technical developments has fuelled a strong sense of urgency within the soil spectroscopy community to combine forces in global initiatives to foster organised science and help operationalise soil spectroscopy. Global coordination in soil spectroscopy is being fostered through the large established networks of GLOSOLAN of the Global Soil Partnership and the Soil Spectroscopy for Global Good network. Additional initiatives include the International Network of Soil Information Institutes (INSII) of the Global Soil Partnership and the IEEE P4005 initiative Standards and Protocols for Soil Spectroscopy. There is a need for securing funding for building and maintaining the services and the required infrastructure, shipping and analysing samples, developing standards and protocols, assisting laboratories to voluntarily submit soil samples from different countries, and through capacity development. The open science approach of Soil Spectroscopy for Global Good and the fostering of an inclusive approach to participation of countries by GLOSOLAN are all essential for rapid progress.

8. Areas for further research

The Global Soil Spectral Calibration Library & Estimation Service initiative is fostering collaborative research in several key areas to improve the efficiency and usability of the service through research networks of the Soil Spectroscopy for Global Good and GLOSOLAN initiatives. Several priority areas have been identified and are under active research by the global soil spectroscopy community.

Transfer of spectra or calibrations from a primary instrument to other instruments of the same type or to different instruments remains a bottleneck to building global spectral libraries and calibrations. The challenge increases further down the chain of different sample preparation, sample presentation or fore-optics, instrument technology (e.g., Fourier-Transform vs dispersive spectrometers), and different spectral ranges (e.g. NIR to MIR). Machine learning approaches based on a limited set of standard soil samples holds promise (e.g., Pittaki-Chrysodonta et al., 2021), but ideally transfer would be based on a set of synthetic standards that are not prone to change in storage and effects of particle settling or changes in particle size through repeated use. For example, the Soil Spectroscopy for Global Good network (SoilSpec4GG, 2022) is organizing the exchange of a set of standard samples amongst spectral laboratories as part of the research on calibration transfer. Standards developed by other groups should also be considered (e.g., Baldock et al., 2013). Being able to efficiently transfer calibrations across instruments would be transformative, allowing new instruments to be rapidly deployed.

Although there has been rapid development in modelling techniques based on machine learning (see Section 5), further work is required on how spectral calibration libraries can be most effectively built and used. Challenges include how to best: (i) select and use local samples to spike the global model for optimal local application; (ii) select local samples for approaches that build calibrations on-the-fly using subsets of spectra; (iii) include spatial correlation effects; (iv) include environmental covariates obtained through earth observation and remote sensing; and (v) use various existing soil spectral libraries of different quality in a unified modelling effort to allow joint use of the global and of local libraries.

9. Limitations of the approach

There are several potential limitations of the approach, and the success of the global service will depend on the degree to which these can be overcome. First, is whether there is sufficient demand for soil spectroscopy, especially in countries with limited resources, to justify the global service. The demand for soil spectroscopy technology in countries in Africa and Asia is well established by the high level of uptake and training requests, articulated in Section 2. This demand has been further confirmed by the GLOSOLAN needs and capacities assessment (Benedetti and van Egmond, 2021) based on a survey of 97 laboratories and experts from 56 different countries, which confirmed a strong interest in improving or starting use of soil spectroscopy, in training and tools, and support for sharing and using shared soil spectroscopy data, including through the decentralized approach offered by GLOSIS. The establishment of GLOSOLAN Regional Champion spectral laboratories and the strong participation in GLOSOLAN webinars (over 2600 participants from 142 countries) provide further evidence of the demand. Soil spectroscopy has also been the prime tool behind estimation of soil properties in the new 30 m resolution soil properties map of Africa (Hengl et al., 2021; iSDASoil, 2021), which is being routinely used for land use planning and nutrient management planning. The fact that commercial soil testing services are also now deploying spectral technology at scale in Europe (Reijneveld et al., 2019) and Africa (e.g., Agrocares, 2021; CNLS, 2021) further strengthens this evidence.

Second, is whether a global initiative is justified in addition to efforts to build local capacity for developing calibrations. We have presented several arguments for why a centralised global calibration initiative offers a good alternative, or supplement, where countries find it difficult to commit resources to establishing a gold standard wet chemistry laboratory capable of analysing a wide suite of soil properties to a high degree of consistency, or to invest in establishing and maintaining their own spectral calibrations. The results of the GLOSOLAN ring test (Fig. 2) demonstrate the difficulty of obtaining consistent reference data even for the commonly measured soil properties, and the authors have observed this constraint to be a major impediment to the application of spectral technology from over 20 years' experience in assisting many resource-constrained tropical countries to develop spectral laboratories. While local calibrations based on high quality reference soil property data will often outperform a global calibration, approaches such as memory-based learning and transfer learning have potential to provide the best of both worlds, benefiting from a wider range of samples and properties than would otherwise be available to a local laboratory. This is illustrated by Fig. 3, which shows the potential for transfer of calibrations, even across continents.

Third, the current global spectral library under-represents soil orders such as oxisols and key geographic areas such as Africa and South Asia. While users of the service should be made aware of the current limits of the geographical distributions of the calibration samples, the best test will be for potential users to upload their spectra and evaluate how well they project onto the calibration spectra data space, and to examine whether prediction errors are acceptable for the intended purpose.

Fourth, countries may find it difficult to find resources for collecting and submitting samples or be unwilling to do so for various reasons, such as phytosanitary regulations or data protection concerns, which may limit the calibration coverage of the estimation service. Support from international efforts such as Soils4Africa and GLOSOLAN to help with submission of samples can benefit and contribute to this global public good. RwaSIS (2021), for instance, has demonstrated how poorly resourced countries can implement a national level sampling campaign, collecting 5750 samples from 2875 sites in less than one year. Even if a country does not submit samples, there might be overlaps in that samples submitted by neighbouring countries, or even continents, with similar soil conditions leading to local calibration value, such as illustrated in Fig. 3. For example, soils from Australia or South America may have calibration value for some Africa soils. To illustrate this, Janik et al. (2007) found MIR calibrations for soil organic carbon fractions using samples from soil types and parent materials from all States in Australia produced only a slight bias when estimating soil carbon fractions for Kenyan soils and provided reasonable estimates of charcoal carbon, which is a difficult and expensive property to measure. In addition, with the inclusion of a subset of Kenya samples in the calibration the bias was completely removed. This implies that not all countries need to submit samples to obtain reasonable predictions, provided that sufficient geographic and soil feature space coverage is present in the library.

Fifth, some countries may not favour a centralised approach, prefer sharing spectral libraries only, prefer to not share data and samples, or prefer for their soil samples not to be archived. Participation in the service is of course completely voluntary. The main purposes of centralising the reference analysis are to provide consistent high quality calibration data and a wider calibration database than available to local laboratories leading to an efficiency and accuracy gain, with secondary benefits of having access to an easy-to-use spectral exploration and estimation tool. The enhanced calibrations are intended to aid the development of local and regional solutions. We hope that individual countries will see the value of participating, including the value of leveraging calibration samples from outside their own country, and that contributors and the international community will see the efficiencies generated from having a centralised soil spectral calibration service. In addition, having soil samples archived at KSSL at no cost to contributing countries, has potential future value for countries, for example calibrations can be extended to included new reference methods that may become available in the future. We see examples of international and regional cooperation working in other sectors, such as agricultural research and trade, that provide mutual benefits and generate efficiencies. An example, is the sharing of soils data for the production of digital soil maps for the benefit of all (iSDASoil, 2021; SoilGrids, 2021).

Sixth, countries that already have well established and high quality spectral libraries, such as France and Australia, may see limited added value in contributing to and using GSCLES since the quality of their local predictions is likely to be better than current global estimates. Benefits of participating are having a wider, more extended calibration set at limited extra costs after applying spectral and reference data transfer functions for in country and international use, and the possibility of a quality comparison with another high-quality laboratory. Participation will help neighbouring countries to achieve better geographic and feature space coverage, and therefore increase the usability of GCSLE for countries that are not able or willing to send in samples and/or to develop an extensive local library themselves. GCSLE will also provide a common reference to aid the development of transfer functions of reference measurements or to use as standard analysis method to alleviate transboundary challenges in regional mapping as is experienced in country-driven approaches as for example applied by the Global Soil Partnership (GLOSIS, 2021) and the EJP SOIL (EJP SOIL, 2021) project. Systematic offsets in maps at country borders limit their use for international and regional policy and decision making and reporting to e.g., UN bodies. In the longer run, an aim is to develop algorithms that allow transfer across spectral libraries as mentioned in Section 8. This strengthens the use of spectroscopy as technique, the GSCLES and developed and developing local libraries, allowing more localisation than the GCLES alone can afford while using the added value of similar samples from other countries.

Seventh, laboratories with limited resources may not be able to invest in the same soil processing facilities or spectral equipment used by KSSL, which will limit the utility of the estimation service. The accompanying research needs outlined in Section 8, to develop algorithms to transfer across soil preparation methods and instruments, are critical for the success not only for the GSCLES but for calibration transfer and regular calibration updating in any initiative.

Lastly, the objective of GSCLES is not to supplant ongoing efforts to standardise and improve the quality of conventional soil analytical laboratories or to improve the capacity of laboratories to develop their own spectral calibrations. On the contrary, GSCLES can provide a valuable benchmark against which laboratories can compare both their spectral and chemical reference analyses, and a valuable tool for cases where laboratories have insufficient resources to develop gold standard calibrations. There may also be a viable model whereby over time gold standard regional laboratories are developed, and again cross-referencing with GSCLES can assist such a model.

10. Conclusion

The proposed Soil Spectral Calibration Library and Estimation Service, facilitated by GLOSOLAN and Soil Spectroscopy for Global Good, could generate enormous efficiencies and would constitute an important global public good. Most importantly, participating laboratories will have access to high quality soil reference data and will benefit from the global soil spectral library, which may contain similar soils from other regions. The coverage and value of the global calibration library will increase over time with smart selection and addition of new samples. The capacity and competitiveness of national soil laboratories on spectroscopy will be enhanced through participation in the initiative. Developing countries with limited laboratory resources stand to gain the most from the service as they can take advantage of soil samples that have been collected and characterized by other countries, with a minimal investment required to submit samples and thereby further localise the calibrations. Since inherent limitations in the reproducibility of laboratory reference measurements currently affects reliability everywhere, this centralized effort would result in more reliable and lower cost spectral estimations and prevent enormous wastage of resources spent on sub-optimal calibrations in laboratories across the world, while increasing the potential productivity of laboratories by increasing the number of samples that can be analysed from the same budget. Access to a high-quality spectral calibration library and estimation service would support unprecedented high quality and quantity of soil data collection with lower costs, which in turn would improve evidence-based decision-making in many fields, including sustainable soil management, food security and nutrition, and climate adaptation and mitigation. Urgent research is required on ways to easily transfer calibrations across different instrument types and sample preparation methods to reap the potential benefit of the global service.

Funding

Staff time was contributed from Soil Spectroscopy for the Global Good, funded through USDA NIFA Award No. 2020-67021-32467, Innovative Solutions for Decision Agriculture (iSDA), funded by the Bill & Melinda Gates Foundation, and by USDA-KSSL, ISRIC World Soil Information and the University of Nebraska.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge the substantial support for this initiative of Lucrezia Caon and Yi Peng of GSP, FAO through GLOSOLAN. We thank Andrew Sila (CIFOR-ICRAF) for assistance in preparing Fig. 3.

Appendix 1

Number of samples from different countries in the KSSL spectral library and additional archive samples with analytical data.

Country	Spectral Library	Archive	Total
Albania		235	235
Antarctica	5	52	57
Argentina		78	78
Australia	264		264
Belgium	17		17
Belize		119	119
Bhutan	32		32
Bolivia		11	11
Botswana		128	128
Brazil		185	185
Bulgaria		168	168
Burundi		55	55
Cameroon	8	30	38
Canada	121	140	261
Chile	38	108	146
China		203	203
Colombia		89	89
Congo (Democratic Republic of the)		89	89
Costa Rica	72	42	114
Denmark		31	31
Ecuador		7	7
Egypt		15	15
El Salvador		27	27
Estonia		21	21
Federated States of Micronesia		83	83
Finland		96	96
France		17	17
Gambia		13	13
Georgia	28		28
Ghana	25	68	93
Guatemala		51	51
Haiti	256	53	309
Honduras		13	13
Hungary		28	28
Iceland		26	26
India		153	153
Indonesia	152	196	348
Iraq	23		23
Japan		159	159
Jordan	13	82	95
Kenya	10	128	138
Korea, Democratic People's Republic of		19	19
Korea, Republic of		127	127
Latvia		33	33
Lebanon		16	16
Lesotho		79	79
Lithuania		33	33
Malawi		140	140
Malaysia	1	20	21
Mali		70	70
Mauritania		43	43
Mexico		274	274
Mongolia		43	43
Morocco	20	60	80
Nepal		51	51
Netherlands	87		87
New Zealand		13	13
Nicaragua		19	19
Niger		77	77
Nigeria		41	41
Norway		23	23
Oman		40	40
Pakistan		312	312
Palau		133	133
Panama	2	17	19
Papua New Guinea		25	25
Philippines		472	472
Poland		298	298
Russia	40	251	291
Rwanda		118	118
Samoa		106	106
Senegal		1	1
Somalia		29	29
South Africa		6	6
Spain	49		49
Sri Lanka		7	7

(*continued on next page*)

(*continued*)

Country	Spectral Library	Archive	Total
Sudan		185	185
Suriname	15		15
Syria		151	151
Taiwan		263	263
Thailand		169	169
Togo		98	98
Tunisia		86	86
Turkey		15	15
Uganda		93	93
United Arab Emirates	59	38	97
United Kingdom	4		4
United States	82,535	112,314	194,849
Uruguay		134	134
Venezuela		58	58
Yemen		38	38
Zambia		118	118
Zimbabwe		255	255
Total	**83,876**	**120,010**	**203,886**

Appendix 2

Sample submission guidelines for inclusion in the Global Soil Spectral Calibration Library and Estimation Service.

1 Submitted samples should represent benchmark soil series or be selected according to your choice of statistical or otherwise representative sampling frames across targeted landscapes, for which predictive models are intended. The overriding consideration in sample selection should be to capture as much of the compositional variability of the soil resources that local calibrations would serve. The program will initially accept up to 300 samples per country, if the samples represent the soil diversity of the entire country or a major region.
2 The sample submitter must complete a pre-formatted spreadsheet and include the data discussed in the guidelines below; to request the spreadsheet, e-mail spectralsamples@isric.org. Type "Global MIR Spectroscopy Initiative" in the subject line. After appropriate screening, the KSSL will contact the sample submitter to facilitate compliant shipments for analysis.
3 Although not required, the sample submitter is encouraged to submit samples representing whole pedons or soil profiles to a depth of 200 cm (or less if bedrock or undisturbed parent material is shallower). Because the number of samples per country must be controlled, you may also choose to select samples that emphasize the upper 100 cm.
4 No oils (e.g., WD-40) or other chemicals should come into contact with the samples. Extra care may be needed during sample collection and transport.
5 Prior to shipment, all samples must:
  a Be thoroughly air-dried at 30 to 35°C for 3 to 7 days.
  b Be hand-sieved (not machine processed) through a 2-mm stainless steel (not brass) sieve. Machine processing is acceptable for samples that have a high clay content and are not thought to contain coarse fragments.
  c Be devoid of coarse fragments and organic material that is greater than 2 mm.
  d Have a minimum mass of 200 gs; 500 gs is preferred so that sample may be archived for future additional analyses that might later be requested by the customer.
    - If less than 200 gs is submitted, the customer may be asked to prioritize analysis requests to ensure that soil properties of highest importance are measured first.
6 Include the following information on the spreadsheet. Samples without sufficient and quality field data will not be authorized for shipment.
  a Sample collection date (mandatory)
  b Sample submitter contact information (for handling the shipment purpose only), including:
    i Name (mandatory)
    ii Professional affiliation (mandatory)
    iii Professional email address (mandatory)
    iv Professional phone number (mandatory) Note: Do NOT furnish personal contact information, only professional information.
  c Sample provenance:
    i Country of origin (mandatory)
    ii State or province of origin (mandatory)
    iii Site coordinates as decimal latitude and longitude in WGS84 (mandatory for newly collected samples, optional but highly desirable for previously collected samples)
    iv Pedon or location IDs (mandatory)
    v Soil classification, highest order, including classification system used, e.g., FAO-World Reference Base for Soil Resources, USDA Soil Taxonomy, etc.
    vi Land cover type/land use: 2 levels according to USDA-NRCS classification (mandatory if available for newly collected samples, optional for previously collected samples)
    vii Land use history (mandatory if available for newly collected samples, optional for previously collected samples)
  d Sampling depth range, in centimeters below the mineral soil surface; e.g., 0–5 cm, 5–25 cm, 25–75 cm, 75–100 cm (mandatory)
  e Horizon designations, including classification system used (mandatory if available)
  f Layer IDs (mandatory), such as the ID or unique identifier of a specific horizon or sampling depths of layer

g Photograph of site and surrounding area (mandatory for newly collected samples unless a camera is unavailable, optional for previously collected samples)

7 The sample submitter is responsible for all costs of sampling, packaging, and shipping. Note: It is NOT necessary to sterilize samples prior to shipment to the USA.

8 The sample submitter must accept that submitted and measured data and metadata will become free public information according to U.S. law. USDA cannot omit the data or metadata from public view. Without exception, the results of all analysis conducted by the KSSL will be distributed without copyright restriction.

9 Once samples are received, they become the property of USDA-NRCS. When submitting samples, the sample submitter should keep subsamples in case they need some of the soil for future use.

References

Aastveit, A.H., Marum, P., 1991. On the effect of calibration and the accuracy of nir spectroscopy with high-levels of noise in the reference values. Appl. Spectrosc. 45 (1), 109–115. https://doi.org/10.1366/0003702914337740.

Abrams, S.M., Shenk, J.S., Westerhaus, M.O., Barton F.E., II, 1987. Determination of forage quality by near infrared reflectance spectroscopy: efficacy of broad-based calibration equations. J. Dairy Sci. 70 (4), 806–813. https://doi.org/10.3168/jds.s0022-0302(87)80077-2.

AgroCares, 2021. AgroCares Nutrient Intelligence. https://www.agrocares.com.

Arrouays, D., Jolivet, C., Boulonne, L., Bodineau, G., Saby, N., Grolleau, E., 2002. A new initiative in France: a multi-institutional soil quality monitoring network. Comptes Rendus de l'Académie d'Agriculture de France 88, 93–105.

Baldock, J., Hawke, B., Sanderman, J., Macdonald, L., 2013. Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. Soil Res. 51, 577–595, 10.1071/SR13077.

Barra, I., Haefele, S.M., Sakrabani, R., Kebede, F., 2021. Soil spectroscopy with the use of chemometrics, machine learning and pre-processing techniques in soil diagnosis: recent advances - a review. Trends Anal. Chem. 135, 116166 https://doi.org/10.1016/j.trac.2020.116166.

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 Biol. Biochem. 43 (7), 1398–1410. https://doi.org/10.1016/j.soilbio.2011.02.019.

Ben Dor, E., Ong, C., Lau, I.C., 2015. Reflectance measurements of soils in the laboratory: standards and protocols. Geoderma 245–246, 112–124. https://doi.org/10.1016/j.geoderma.2015.01.002.

Benedetti, F., van Egmond, F., 2021. Global Soil Spectroscopy Assessment. Spectral Soil Data – Needs and Capacities. FAO, Rome. https://doi.org/10.4060/cb6265en.

BIPM, I., IFCC, I., ISO, I., IUPAP, O., 2008. International vocabulary of metrology–basic and general concepts and associated terms (VIM. J.C.G.M. 200, 2012. https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf/f0e1ad45-d337-bbeb-53a6-15fe649d0ff1.

Brown, D., Shepherd, K.D., Walsh, M.G., 2006. Global soil characterization using a VNIR diffuse reflectance library and boosted regression trees. Geoderma 132, 273–290. https://doi.org/10.1016/j.geoderma.2005.04.025.

Chabrillat, S., Gholizadeh, A., Neumann, C., Berger, D., Milewski, R., Ogen, Y., Ben-Dor, E., 2019. Preparing a soil spectral library using the Internal Soil Standard (ISS) method: influence of extreme different humidity laboratory conditions. Geoderma 355, 113855. https://doi.org/10.1016/j.geoderma.2019.07.013.

CNLS, 2021. Crop Nutrition Laboratory Services Ltd, Nairobi. https://cropnuts.com.

Congress, 2017. H.R.1770 - OPEN Government Data Act, 115th Congress (2017-2018) of the USA. https://www.congress.gov/bill/115th-congress/house-bill/1770.

CSIRO, 2021. CSIRO National Soil Site Database. Commonwealth Scientific and Industrial Research Organisation, Australia. https://researchdata.edu.au/csiro-national-soil-site-database/444301.

Dangal, S.R.S., Sanderman, J., 2020. Is standardization necessary for sharing of a large mid-infrared soil spectral library? Sensors 20 (23), 6729. https://doi.org/10.3390/s20236729.

Dangal, S.R.S., Sanderman, J., Wills, S., Ramirez-Lopez, L., 2019. Accurate and precise prediction of soil properties from a large mid-infrared spectral library. Soil Syst 3 (1), 11. https://doi.org/10.3390/soilsystems3010011.

De Vries, S., Ter Braak, C.J., 1995. Prediction error in partial least squares regression: a critique on the deviation used in The Unscrambler. Chemom. Intell. Lab. Syst. 30 (2), 239–245. https://doi.org/10.1016/0169-7439(95)00030-5.

Demattê, J.A.M., Dotto, A.C., Paiva, A.F.S., Sato, M.V., Dalmolin, R.S.D., de Araújo, M.do S.B., Couto, H.T.Z., 2019. The Brazilian Soil Spectral Library (BSSL): a general view, application and challenges. Geoderma 354, 113793. https://doi.org/10.1016/j.geoderma.2019.05.043.

Difoggio, R., 1995. Examination of some misconceptions about near-infrared analysis. Appl. Spectrosc. 49 (1), 67–75. https://doi.org/10.1366/0003702953963247.

EJP SOIL, 2021. European Joint Programme Towards Climate-Smart Sustainable Management of Agricultural Soils. https://ejpsoil.eu/about-ejp-soil/structure-work-packages/.

Faber, K., Kowalski, B.R., 1997. Improved prediction error estimates for multivariate calibration by correcting for the measurement error in the reference values. Appl. Spectrosc. 51 (5), 660–665. https://doi.org/10.1366/0003702971941061.

Ge, Y., Morgan, C.L.S., Grunwald, S., Brown, D.J., Sarkhot, D.V., 2011. Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. Geoderma 161 (3–4), 202–211. https://doi.org/10.1016/j.geoderma.2010.12.020.

Gholizadeh, A., Neumann, C., Chabrillat, S., van Wesemael, B., Castaldi, F., Borůvka, L., Hohmann, C., 2021. Soil organic carbon estimation using VNIR–SWIR spectroscopy: the effect of multiple sensors and scanning conditions. Soil Tillage Res. 211 https://doi.org/10.1016/j.still.2021.105017.

GloSIS, 2021. The Global Soil Information System. http://www.fao.org/global-soil-partnership/pillars-action/4-information-data/insii/en/.

Gomez, C., Chevallier, T., Moulin, P., Bouferra, I., Hmaidi, K., Arrouays, D., Jolivet, C., Barthès, B.G., 2020. Prediction of soil organic and inorganic carbon concentrations in Tunisian samples by mid-infrared reflectance spectroscopy using a French national library. Geoderma 375, 114469. https://doi.org/10.1016/j.geoderma.2020.114469.

Grinand, C., Barthes, B.G., Brunet, D., Kouakoua, E., Arrouays, D., Jolivet, C., Caria, G., Bernoux, M., 2012. Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS). Eur. J. Soil Sci. 63 (2), 141–151. https://doi.org/10.1111/j.1365-2389.2012.01429.x.

Hartmann, C., Suvannang, N., 2020. Global Soil Laboratory Assessment, 2018 Online Survey. FAO, Rome. http://www.fao.org/publications/card/en/c/CA7091EN/.

Hengl, T., Miller, M., Krizan, J., Shepherd, K., Sila, A., Kilibarda, M., Crouch, J., 2021. African soil properties and nutrients mapped at 30–m spatial resolution using two-scale ensemble machine learning. Sci Rep. 11 (1), 6130. https://doi.org/10.1038/s41598-021-85639-y.

Hicks, W., Viscarra Rossel, R.A., Tuomi, S., 2015. Developing the Australian mid-infrared spectroscopic database using data from the Australian Soil Resource Information System. Soil Res 53 (8), 922–931. https://doi.org/10.1071/SR15171.

ICAR, 2020. In: Proceedings of an International Webinar on Soil Spectroscopy: An Emerging Technique for Rapid Soil Helath Assessment. ICAR-India Institute of Soil Science and World Agroforestry, 1 October 2020. https://www.worldagroforestry.org/file-download/download/public/22578.

iSDASoil, 2021. ISDASoil 30-m Resolution Soil Properties Map of Africa. https://www.isda-africa.com/isdasoil/.

Janik, L.J., Skjemstad, J.O., Shepherd, K.D., Spouncer, L.R., 2007. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. Aust. J. Soil Res. 45, 73–81 https://doi.org/10.1071/SR06083.

Johnson, J.M., Vandamme, E., Senthilkumard, K., Sila, A., Shepherd, K.D., Saito, K., 2019. Near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for assessing soil fertility in rice fields in sub-Saharan Africa. Geoderma 354, 113840. https://doi.org/10.1016/j.geoderma.2019.06.043.

Knadel, M., Stenberg, B., Deng, F., Thomsen, A., Greve, M.H., 2013. Comparing predictive abilities of three visible-near infrared spectrophotometers for soil organic carbon and clay determination. J. Near Infrared Spec. 21 (1), 67–80. https://doi.org/10.1255/jnirs.1035.

McBratney, A.B., Minasny, B., Viscarra Rossel, R.A., 2006. Spectral soil analysis and inference systems: a powerful combination for solving the soil data crisis. Geoderma 136 (1–2), 272–278. https://doi.org/10.1016/j.geoderma.2006.03.051.

Mark, H., Norris, K., Williams, P.C., 1989. Methods of determining the true accuracy of analytical methods. Anal. Chem. 61 (5), 398–403. https://doi.org/10.1021/ac00180a004.

Mendes, W., de, S., Demattê, J.A.M., Rosin, N.A., Terra, F., da, S., Poppiel, R.R., Urbina-Salazar, D.F., Souza Valladares, G., 2022. The Brazilian soil mid-infrared spectral library: the power of the fundamental range. Geoderma 415, 115776. https://doi.org/10.1016/j.geoderma.2022.115776.

Minasny, B., McBratney, A.B., 2016. Digital soil mapping: a brief history and some lessons. Geoderma 264, 301–311. https://doi.org/10.1016/j.geoderma.2015.07.017.

Minasny, B., McBratney, A.B., Pichon, L., Sun, W., Short, M.G., 2009. Evaluating near infrared spectroscopy for field prediction of soil properties. Aust. J. Soil Res. 47 (7), 664–673. https://doi.org/10.1071/sr09005.

Naes, T., TIsaksson, T., Fearn, T., Davies, T., 2017. A User-friendly Guide to Multivariate Calibration and Classification, 2nd edition. NIR Publications, Chichester, UK.

National Soil Services, 2021. National Soil Services Established Under the Africa Soil Information Service (AfSIS). https://www.isda-africa.com/national-soil-services/.

Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., McBratney, A.B., 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. https://doi.org/10.1016/j.geoderma.2015.04.017.

Ng, W., Husnain, H., Anggria, L., Siregar, A., Hartatik, W., Sulaeman, Y., Jones, E., Minasny, B., 2020a. Developing a soil spectral library using a low-cost NIR

spectrometer for precision fertilization in Indonesia. Geoderma Reg 22, e00319. https://doi.org/10.1016/j.geodrs.2020.e00319.
Ng, W., Minasny, B., de Sousa Mendes, W., Demattê, J.A.M., 2020b. The influence of training sample size on the accuracy of deep learning models for the prediction of soil properties with near-infrared spectroscopy data. Soil 6 (2), 565–578. https://doi.org/10.5194/soil-6-565-2020.
Ng, W., Minasny, B., Jeon, S.H., McBratney, A., 2022a. Mid-infrared spectroscopy for accurate measurement of an extensive set of soil properties for assessing soil functions. Soil Security 6, 100043. https://doi.org/10.1016/j.soisec.2022.100043.
Ng, W., Minasny, B., Jones, E., McBratney, A., 2022b. To spike or to localize? Strategies to improve the prediction of local soil properties using regional spectral library. Geoderma 406, 115501. https://doi.org/10.1016/j.geoderma.2021.115501.
Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthes, B., Ben Dor, E., Brown, D.J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J.A.M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai, H., Soriano-Disla, J.M., Shepherd, K.D., Stenberg, B., Towett, E. K., Vargas, R., Wetterlind, J., 2015. Soil spectroscopy: an alternative to wet chemistry for soil monitoring. Adv. Agron. 132, 139–159 https://doi.org/10.1016/bs.agron.2015.02.002.
Padarian, J., Minasny, B., McBratney, A.B., 2019. Transfer learning to localise a continental soil vis-NIR calibration model. Geoderma 340, 279–288. https://doi.org/10.1016/j.geoderma.2019.01.009.
Pimstein, A., Notesko, G., Ben-Dor, E., 2011. Performance of three identical spectrometers in retrieving soil reflectance under laboratory conditions. Soil Sci. Soc. Am. J. 75 (2), 110–174. https://doi.org/10.2136/sssaj2010.0174.
Pittaki-Chrysodonta, Z., Hartemink, A.E., Sanderman, J., Ge, Y., Huang, J., 2021. Evaluating three calibration transfer methods for predictions of soil properties using mid-infrared spectroscopy. Soil Sci. Soc. Am. J. 85 (3), 501–519. https://doi.org/10.1002/saj2.20225.
Ramirez-Lopez, L., Behrens, T., Schmidt, K., Stevens, A., Demattê, J.A.M., Scholten, T., 2013. The spectrum-based learner: a new local approach for modeling soil vis-NIR spectra of complex datasets. Geoderma 195–196, 268–279. https://doi.org/10.1016/j.geoderma.2012.12.014.
Reeves, J.B., 2010. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: where are we and what needs to be done? Geoderma 158 (1–2), 3–14. https://doi.org/10.1016/j.geoderma.2009.04.005.
Reijneveld, J.A., Brolsma, K.M., van Vliet, P.C.J., van Oostrum, M.J., Oenema, O., 2019. Using NIRS, multi-nutrient 0.01M CaCl2 extractions and the intensity, buffering capacity and quantity concept as tools for fertilization recommendations. In: Proceedings International Symposium of Soil and Plant Analysis (ISSPA), Wageningen. https://cdnmedia.eurofins.com/eurofins-germany/media/2857008/eurofins-agro-the-science-behind-the-tools.pdf.
RwaSIS, 2021. Rwanda Soil Information Service. https://www.worldagroforestry.org/project/rwanda-soil-information-service.
Sanderman, J., Savage, K., Dangal, S.R.S., 2020. Mid-infrared spectroscopy for prediction of soil health indicators in the United States. Soil Sci. Soc. Am. J. 84 (1), 251e261. https://doi.org/10.1002/saj2.20009.
Searle, R., McBratney, A., Grundy, M., Kidd, D., Malone, B., Arrouays, D., Stockman, U., Zund, P., Wilson, P., Wilford, J., Padarian Campusano, J., Minasny, B., Jones, E., et al., 2021. Digital soil mapping and assessment for Australia and beyond: a propitious future. Geoderma Reg. 24, e00359. https://doi.org/10.1016/j.geodrs.2021.e00359.
Seybold, C.A., Ferguson, R., Wysocki, D., Bailey, S., Anderson, J., Nester, B., Thomas, P., 2019. Application of mid-infrared spectroscopy in soil survey. Soil Sci. Soci. Am. J. 83 (6), 1746–1759. https://doi.org/10.2136/sssaj2019.06.0205.
Shepherd, K.D., Walsh, M.G., 2007. Infrared spectroscopy—Enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. J. Near Infrared Spectrosc. 15 (1), 1–19. https://doi.org/10.1255/jnirs.716.
Sila, A.M., Pokhariyal, G.P., Shepherd, K.D., 2016. Evaluating the utility of mid-infrared spectral subspaces for predicting soil properties. Chemom. Intell. Lab. Syst. 153, 92–105. https://doi.org/10.1016/j.chemolab.2016.02.013.
Soil-Plant Spectral Diagnostics Lab, 2021. Network of Dry Spectroscopy Labs. https://worldagroforestry.org/sites/default/files/u897/ICRAF%20Network%20of%20Dry%20Spectroscopy%20Labs_2018.04.03.pdf.
Soils4Africa, 2021. Soils-4-Africa Project. https://www.soils4africa-h2020.eu.
SoilGrids, 2021. SoilGridsTM – global Gridded Soil Information. https://www.isric.org/explore/soilgrids.
SoilSpec4GG, 2022. Soil sSpectroscopy for Global Good. https://soilspectroscopy.org.
Sørensen, L., 2002. True accuracy of near infrared spectroscopy and its dependence on precision of reference data. J. Near Infrared Spectrosc. 10 (1), 15–26. https://doi.org/10.1255/jnirs.317.
Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., & Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. In Donald L. Sparks, Adv. Agron. 163–215. https://doi.org/10.1016/S0065-2113(10)07005-7.
Stevens, A., Nocita, M., Tóth, G., Montanarella, L., van Wesemael, B., 2013. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. PLoS ONE 8 (6), e66409. https://doi.org/10.1371/journal.pone.0066409.
Tang, Y., Jones, E., Minasny, B., 2020. Evaluating low-cost portable near infrared sensors for rapid analysis of soils from South Eastern Australia. Geoderma Reg 20, e00240. https://doi.org/10.1016/j.geodrs.2019.e00240.
Taylor, J.R., 1997. An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books, Sausalito, California, USA.
Terhoeven-Urselmans, T., Vagen, T.-.G., Spaargaren, O., Shepherd, K.D., 2010. Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. Soil Sci. Soc. Am. J. 74 (5), 1792–1799. https://doi.org/10.2136/sssaj2009.0218.
Towett, E.K., Shepherd, K.D., Sila, A., Aynekulu, E., Cadisch, G., 2015. Mid-infrared and total x-ray fluorescence spectroscopy complementarity for assessment of soil properties. Soil Sci. Soc. Am. J. 79 (5), 1375–1385. https://doi.org/10.2136/sssaj2014.11.0458.
USDA-NRCS, 2021. Kellogg Soil Survey Laboratory Methods Manual, SSIR-42, v. 5, USDA-NRCS. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2_054247.
Vågen, T.-.G., Winowiecki, L.A., Desta, L., Tondoh, E.J., Weullow, E., Shepherd, K., Sila, A., 2020. Mid-Infrared Spectra (MIRS) from ICRAF soil and plant spectroscopy laboratory: Africa Soil Information Service (AfSIS) Phase I 2009-2013. World Agroforestry - Research Data Repository, V1. https://doi.org/10.34725/DVN/QXCWP1.
Viscarra Rossel, R.A., Webster, R., 2012. Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. Eur. J. Soil Sci. 63 (6), 848–860. https://doi.org/10.1111/j.1365-2389.2012.01495.x.
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 reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131 (1–2), 59–75. https://doi.org/10.1016/j.geoderma.2005.03.007.
Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aïchi, H., Barthès, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Böttcher, K., Brodský, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, M., Knadel, H.J.M., Morrás, M., Nocita, L., Ramirez-Lopez, P., Roudier, E.M., Rufasto Campos, C.B., Sanborn, P., Sellitto, V.M., Sudduth, K.A., Rawlins, B.G., Walter, C., Winowiecki, L.A., Hong, S.Y., Ji, W., 2016. A global spectral library to characterize the world's soil. Earth Sci. Rev. 155, 198–230. https://doi.org/10.1016/j.earscirev.2016.01.012.
Wadoux, A.M.J.C., McBratney, A.B., 2021. Digital soil science and beyond. Soil Sci. Soc. Am. J. https://doi.org/10.1002/saj2.20296 (in press).
WEPAL, 2021. WEPAL Proficiency Testing. https://www.wepal.nl/en/wepal.htm.
Wijewardane, N.K., Ge, Y., Wills, S., Libohova, Z., 2018. Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library. Soil Sci. Soc. Am. J. 82 (3), 722–731. https://doi.org/10.2136/sssaj2017.10.0361.
Yang, Y., Shen, Z., Bisset, A., Viscarra Rossel, R.A., 2021. Estimating soil fungal abundance and diversity at a macroecological scale with deep learning spectrotransfer functions. Soil. https://doi.org/10.5194/soil-2021-79 (in press).