Analyst

PAPER

View Article Online
View Journal | View Issue

Cite this: *Analyst*, 2025, **150**, 2934

Qualitative and quantitative soil characterization on an agricultural field using a portable shifted excitation Raman difference spectroscopy instrument

Kay Sowoidnich, *[a] Stefan Pätzold,[b] Markus Ostermann,[c] Bernd Sumpf[a] and Martin Maiwald [a]

Site-specific farmland management requires comprehensive information about the soil status to derive informed treatment decisions, *e.g.* for liming or fertilizer recommendations. Standard laboratory methods relying on sample collection have only limited ability to adequately capture the spatial variability of typical agricultural fields. Here, on-site analytical techniques with the potential to measure the soil properties on a substance-specific level and at the required spatial resolution could be very beneficial. Raman spectroscopy is a very promising technique for this purpose as it provides a molecular fingerprint of soil constituents. However, intrinsic soil fluorescence and daylight interference can be major issues masking characteristic Raman signals. Here, we apply an in-house developed portable shifted excitation Raman difference spectroscopy (SERDS) instrument based on a dual-wavelength diode laser emitting around 785 nm to effectively separate the Raman signals of soil from such interferences. SERDS investigations on a selected agricultural field in Germany demonstrate that the Raman spectroscopic signature of 9 soil minerals and organic carbon could successfully be separated from intense backgrounds. Using partial least squares regression against reference analyses, a successful prediction of the soil carbonate (R^2 = 0.86, root mean squared error of cross validation RMSECV = 2.49%) and soil organic carbon content (R^2 = 0.89, RMSECV = 0.32%) as important soil parameters is realized. The results obtained on-site with the portable instrument were confirmed by SERDS laboratory experiments of collected soil samples thus highlighting the capability and reliability of portable SERDS as promising and complementary tool for precision agriculture.

Received 17th February 2025,
Accepted 18th May 2025

DOI: 10.1039/d5an00178a

rsc.li/analyst

Introduction

Productive agricultural soil is a crucial precondition to meet the increasing food demand of a steadily growing world population. However, soil conditions can be adversely affected, *e.g.* by inadequate soil management, excessive or insufficient fertilizer application as well as the effects of the climate change. To address these issues, the concept of precision agriculture[1] to establish site-specific farmland treatment aiming for efficient and sustainable soil nutrient management is gaining importance on a global level. The approach can contribute to improve crop productivity, increase overall soil productivity, and optimize usage of fertilizers. To exploit the full potential of site-specific agricultural soil management, detailed information about the present soil properties is essential to derive informed treatment decisions, *e.g.* for liming, fertilizer or irrigation recommendations.

Current procedures for soil analysis are based on the collection of composite samples from rather large areas at the ha-scale[2] followed by labor-intensive and time-consuming analysis using established laboratory methods. This approach does unfortunately have only limited ability to adequately capture the spatial soil heterogeneity of typical agricultural fields[3] where in many cases a spatial resolution on the order of 10 m would be beneficial. This corresponds to a spatial scale at which modern agricultural machinery can be adapted to changing soil properties. Here, on-site analytical techniques with the potential to measure the soil composition on a substance-specific level and at the desired spatial resolution could be very beneficial, *e.g.* to contribute essential information for evidence-based decision support tools.[4]

[a]*Ferdinand-Braun-Institut (FBH), Gustav-Kirchhoff-Straße 4, 12489 Berlin, Germany. E-mail: kay.sowoidnich@fbh-berlin.de*
[b]*Institute of Crop Science and Resource Conservation (INRES), Soil Science and Soil Ecology, University of Bonn, Nussallee 13, 53115 Bonn, Germany*
[c]*Federal Institute for Materials Research and Testing (BAM), Process Analytical Technology, Richard-Willstätter-Straße 11, 12489 Berlin, Germany*

Raman spectroscopy is a well-established analytical method providing a molecular fingerprint of the sample under investigation thus permitting for qualitative and quantitative measurements. The technique is, however, currently still underexplored in the area of soil compositional analysis. This is mainly due to the fact that Raman scattering is a weak process and that the characteristic signals can easily be masked by interfering contributions, *e.g.* fluorescence originating from soil organic matter[5] or clay minerals.[6] In case of on-site experiments, an additional challenge beside intrinsic soil fluorescence is interference by daylight that can also easily mask the characteristic Raman signals of soil constituents.

A physical approach to effectively address these issues is Shifted Excitation Raman Difference Spectroscopy (SERDS).[7,8] The basic concept behind SERDS is that the sample under investigation is consecutively excited at two slightly different laser wavelengths. While the Raman signals will follow the small shift in excitation wavelength, *e.g.* 10 cm^{-1} which corresponds to a spectral shift of 0.6 nm at 785 nm, background contributions remain essentially unchanged. A following subtraction of both recorded Raman spectra from each other can thus efficiently separate the characteristic Raman spectroscopic fingerprint of the probed specimen from various backgrounds, including interferences from fluorescence and ambient lights.[9,10]

Our group has previously demonstrated and evaluated the capability of SERDS for laboratory soil analysis. Beside qualitative analysis,[11] also quantitative measurements, *e.g.* for the prediction of the important parameters soil organic matter content[12] and soil carbonate content[13] are possible. Moreover, a portable SERDS system was realized and its feasibility for *in situ* soil analysis was shown during an initial proof-of-concept study on a test plot. For demonstration, selected soil minerals could be identified on this artificially created plot with a footprint of 12 m × 12 m containing a well-defined soil composition (silt loam).[9]

In contrast to our previous work, this paper now goes a significant step ahead and presents, for the first time, systematic on-site SERDS investigations for qualitative and quantitative soil analysis directly on an agricultural field in Northeast Germany. Contrary to the above-mentioned test plot, this field is used for agricultural production and undergoes processing by agricultural machinery on a regular basis. These measurements are performed to generate a set of reference Raman spectra of soil constituents using 25 selected measurement points along a distance of 480 m across the field, chosen to capture its spatial heterogeneity. Moreover, SERDS combined with partial least squares regression is evaluated as a tool for quantitative predictions in case of the soil calcium carbonate content, a relevant parameter for the assessment of liming requirements, and the soil organic carbon content that is an important indicator for the overall soil condition and in particular the nutrient status. To validate the results obtained with the portable SERDS instrument during the field measurement, a comparison with SERDS laboratory experiments of samples collected from the measurement positions that were probed on-site is carried out.

Materials and methods

Portable SERDS instrument

The rugged and portable SERDS system specifically designed for soil analysis is described in detail in our previous publication[9] and only a brief overview of key components is given here. As excitation light source, a dual-wavelength Y-branch distributed Bragg reflector ridge waveguide (DBR-RW) diode laser[14] with two distinct emission lines at 783.7 nm and 784.3 nm is chosen. With respect to its application for SERDS, this one-chip device is designed to provide a spectral distance between the two excitation wavelengths of 10 cm^{-1} (0.6 nm). The dual-wavelength diode laser is implemented into an in-house developed compact turnkey diode laser system that provides an operating environment with integrated current sources and a heat sink with included Peltier element for temperature management and is controlled *via* USB 2.0 interface.[15] The excitation laser radiation is then launched into an optical fiber (core diameter: 105 μm, NA 0.22) contained in a stainless-steel sleeve and connected to a customized probe (Raman probe II™, InPhotonics).

The probe head is equipped with a sapphire optical window to protect the internal optical components from the sample and environmental conditions, and to maintain a constant distance between the focusing lens (working distance: 5 mm) and the sample. The laser light is focused onto the sample generating an excitation spot diameter of about 115 μm and the back-scattered light is collected by the same lens in 180° geometry. Inside the probe, a Raman long-pass filter assembly only transmits the Raman Stokes light (longer wavelengths compared to the excitation radiation) that is then launched into an optical fiber (core diameter: 300 μm, NA 0.22). This fiber is encapsulated inside a stainless-steel sleeve and connected to a compact OEM spectrometer with an improved light throughput by means of a high-throughput virtual slit (Hyperflux P2, Tornado Spectral Systems). The spectrometer provides a spectral resolution of 6 cm^{-1} within the accessible spectral region between 215 and 3280 cm^{-1}. A research-grade spectroscopy CCD camera (iVac 316, Andor) cooled down to an operating temperature of −20 °C is connected to the spectrometer and serves as multichannel detector.

System control and data handling is realized by means of in-house developed software running on a robust laptop equipped with a Intel(R) Core(TM) i5 3320M processor operating at 2.60 GHz and 4 GB of RAM (Toughbook CF-53, Panasonic).

Study site and experimental parameters

On-site SERDS experiments were conducted on an agricultural field located in the Younger Moraine Landscape in northeast Germany (latitude: 52.394316N; longitude: 14.461156E) in 2021. Further information regarding selected soil parameters of the field are provided in our previous publication.[12] It is well-known from preceding investigations that there exists a pronounced spatial variability across the field, particularly with respect to the distribution of carbonates and organic

matter. To capture this heterogeneity, the measurement points for the on-site SERDS investigations (n = 25) were selected along a straight line along the tractor tracks across the field covering a total distance of 480 m. Within the region of the largest variation, adjacent points are separated by 15 m while the remainder of the line was probed with a distance of 30 m between neighboring points. As some plant residuals were still present on the surface of the field, small holes (*ca.* 10 cm depth and diameter) were created with a stainless-steel scoop to expose the bare soil without interference from plant litter. During the SERDS measurements the Raman probe was slightly pressed to the exposed soil surface and hand-held by the operator.

Short integration times of 50 ms had to be used to avoid saturation of the CCD detector and 50 spectra were accumulated for each of the two excitation wavelengths. The heatsink temperature was set to 25 °C and the injection currents at the diode laser were selected to achieve an optical output power of 36 mW at the sample position. These parameters are identical those selected in our previously published pilot investigations and are discussed there in more detail.[9] To account for possible spatial variations at the millimeter and centimeter scale, 10 randomly selected spot were probed at each of the 25 measurement points.

Processing and analysis of SERDS spectra

A detailed outline of the data processing for SERDS by means of an in-house developed algorithm implemented in MATLAB (R2017a, MathWorks, Natick, MA, USA) is presented in our previous publications[12,13] and only an overview of the major steps is given here. Initially, the mean Raman spectra recorded at each of the two excitation wavelengths are subtracted from each other to calculate a SERDS difference spectrum. Following cubic spline fitting to obtain a derivative-shaped difference spectrum centered around zero-baseline, a numerical integration is performed to generate a reconstructed SERDS spectrum in a conventional form. To achieve a straight horizontal baseline, finally a baseline correction of the reconstructed SERDS spectrum is computed.

Prior to multivariate regression, the data set containing average spectra for each of the 25 measurement points on the agricultural field is truncated to the 250–1650 cm^{-1} spectral region (corresponding to Raman-Stokes scattering in a wavelength range of 800–900 nm at the chosen excitation wavelength around 785 nm) that contains prominent Raman signals of relevant major soil constituents and is subsequently normalized to the intensity of the strongest Raman signal originating from the sapphire measurement window at 418 cm^{-1}. Partial least squares regression (PLSR) of the SERDS data is performed using the MATLAB function "plsregress" that is based on the SIMPLS algorithm[16] and included within the Statistics and Machine Learning Toolbox. Mean-centering of the data as pre-processing step is performed as part of the applied PLS algorithm. Correlations are established with calcium carbonate contents as derived from soil calcium contents determined by XRF analysis as well as soil organic carbon contents measured by elemental analysis. Leave-one-out cross validation was selected as suitable cross validation approach considering the rather small sample size of 25 spectra. The number of PLS factors included into the regression models is identified by determining the number of factors giving the minimum root mean squared error of cross validation (RMSECV).[17]

Sample collection and laboratory analyses

Along with the on-site SERDS measurements, soil samples from the probed 25 locations were collected for additional laboratory investigations. According to established sample preparation standards in soil science, these specimens were air-dried at room temperature, pestled in a mortar and sieved to grain sizes smaller than 2 mm. As outlined in our previous publications, soil calcium contents were determined using X-ray fluorescence (XRF) analysis,[13] while the soil organic carbon (SOC) content is measured by means of elemental analysis.[12] To confirm the results obtained on-site with the portable SERDS system, soil samples were additionally investigated with a customized SERDS laboratory setup using an established measurement protocol previously identified as being suitable for qualitative and quantitative soil analysis by probing a regular 10 × 10 grid pattern within an area of 1 cm^2.[12,13]

Results and discussion

Raman and SERDS spectra of soil acquired on an agricultural field

As displayed in Fig. 1, due to intrinsic fluorescence originating from soil organic matter or clay minerals only a Raman signal of the sapphire window at 418 cm^{-1} can be identified in the Raman spectra recorded at the two slightly shifted excitation wavelengths (top curves). Here, each curve represents the average of 50 spectra recorded at 50 ms integration time for a single Raman spectrum. It can be seen that the background shape in both recorded Raman spectra remains unchanged while the sapphire Raman signal follows the spectral shift between both excitation wavelengths.

The reconstructed SERDS spectrum that is displayed in the bottom part of Fig. 1 shows that individual soil constituents can be identified at the selected measurement point. Beside quartz (SiO_2) as frequently occurring component in a wide range of soils (Si–O–Si symmetric stretching vibration at 465 cm^{-1})[18] this also includes the calcium carbonate ($CaCO_3$) polymorph aragonite (symmetric C–O stretching vibration of the CO_3 group at 1084 cm^{-1})[19] and organic carbon (D-band at 1350 cm^{-1} and G-band at 1590 cm^{-1}).[20] Moreover, two other Raman signals originating from the sapphire window can be identified at 577 cm^{-1} and 749 cm^{-1} (indicated by asterisks in Fig. 1).

In the next section, an overview of representative spectra for all soil components identified on-site is presented.

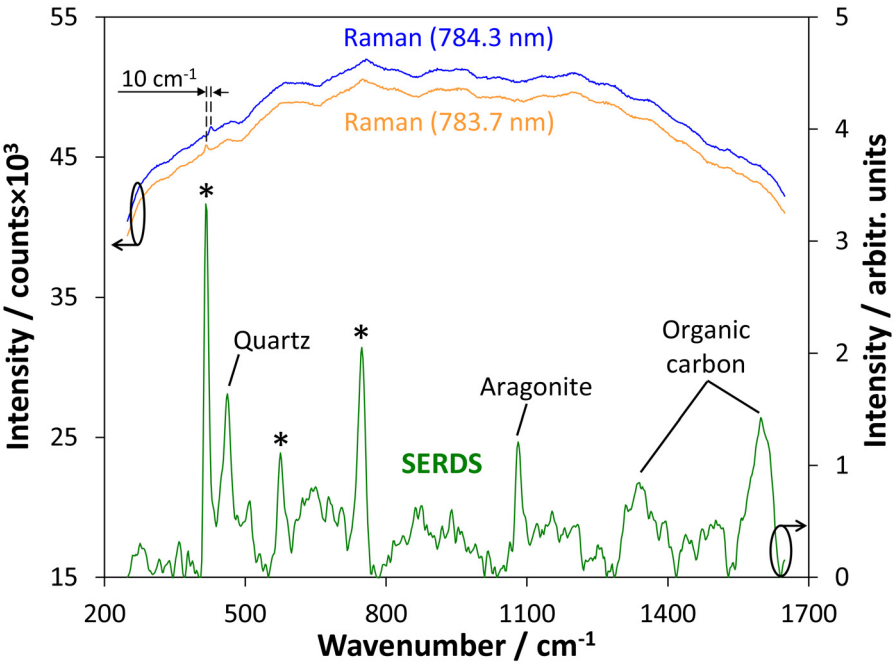

Fig. 1 Average of 50 Raman spectra (top curves) excited at 783.7 nm and 784.3 nm using integration times of 50 ms, and corresponding reconstructed SERDS spectrum (bottom curve) obtained on-site from one single measurement position at the field under study. Asterisks indicate Raman signals originating from the sapphire window of the Raman probe.

Assessment of molecular soil composition

In Fig. 1 it can be seen that multiple soil constituents can be detected at a single measurement spot. This outcome is not unexpected due to the applied measurement spot size on the order of 100 μm diameter and has already been observed in our previous SERDS laboratory investigations.[13] Nevertheless, it was possible to generate a set of reference Raman spectra of soil constituents using the total number of 250 spectra acquired from the investigated measurement points along the field. For this purpose, an average of two selected sampling positions (exception: almandine that was only detected at one single spot) with the strongest spectral contribution of the respective target substance but small or negligible spectral signatures of other soil constituents was calculated. For clarity, the spectral signature of quartz was subtracted from the selected spectra of the other detected components to remove the strong contribution of its prominent Raman signal at 465 cm^{-1}. The weaker contributions originating from the sapphire window of the Raman probe were not removed from the spectra and the corresponding signal positions[21,22] at 418 cm^{-1}, 577 cm^{-1}, and 749 cm^{-1} are indicated by asterisks.

Silicate minerals contained within various soils are of particular interest due to their relative abundance and importance.[23] Representative average SERDS spectra obtained from quartz and feldspar are presented in the top part of Fig. 2A. Quartz (SiO_2) displays the strongest Raman signal due to the Si–O–Si symmetric stretching vibration at 465 cm^{-1}, a weak signal originating from the SiO_4 asymmetric stretching vibration at 1159 cm^{-1}, and two other weak Raman signals attributed to lattice modes at 266 cm^{-1} and 356 cm^{-1}.[18,24] In case of feldspar, Na-rich and K-rich modifications can be identified and distinguished based on the recorded spectra. The Na-feldspar ($NaAlSi_3O_8$) in the form of albite with a triclinic crystal structure is characterized by a prominent Raman signal at 509 cm^{-1} and a minor signal at 292 cm^{-1}, while the strongest Raman signal in case of the K-feldspar ($KAlSi_3O_8$) present in the form of microcline with a triclinic crystal structure is located at 513 cm^{-1} and a minor signal can be identified at 285 cm^{-1}. For both types of feldspar, the strongest bands around 510 cm^{-1} originate from breathing modes of the four-membered tetrahedral rings whereas the minor signals around 290 cm^{-1} can be attributed to rotation-translation modes of the rings.[25,26] The last identified silicate mineral is almandine ($Fe_3Al_2Si_3O_{12}$) which is a member of the garnet group. The strongest Raman signal in this case is originating from a Si–O symmetric stretching vibration at 913 cm^{-1}. Moreover, weaker bands at 861 cm^{-1} and 1037 cm^{-1} due to O–Si–O asymmetric stretching can be identified. A Raman signal at 554 cm^{-1} is originating from a O–Si–O symmetric bending vibration while the two bands at 500 cm^{-1} and 631 cm^{-1} are caused by O–Si–O asymmetric bending vibrations. Finally, two Raman signals assigned to rotations of the SiO_4 tetrahedron are observed at 346 cm^{-1} and 373 cm^{-1}.[27,28]

Hydroxyapatite ($Ca_5(PO_4)_3OH$) as intrinsic phosphate species within the soil can be identified as well based on its characteristic Raman signal originating from the symmetric P–O stretching vibration of the PO_4 group at 960 cm^{-1}.[29,30]

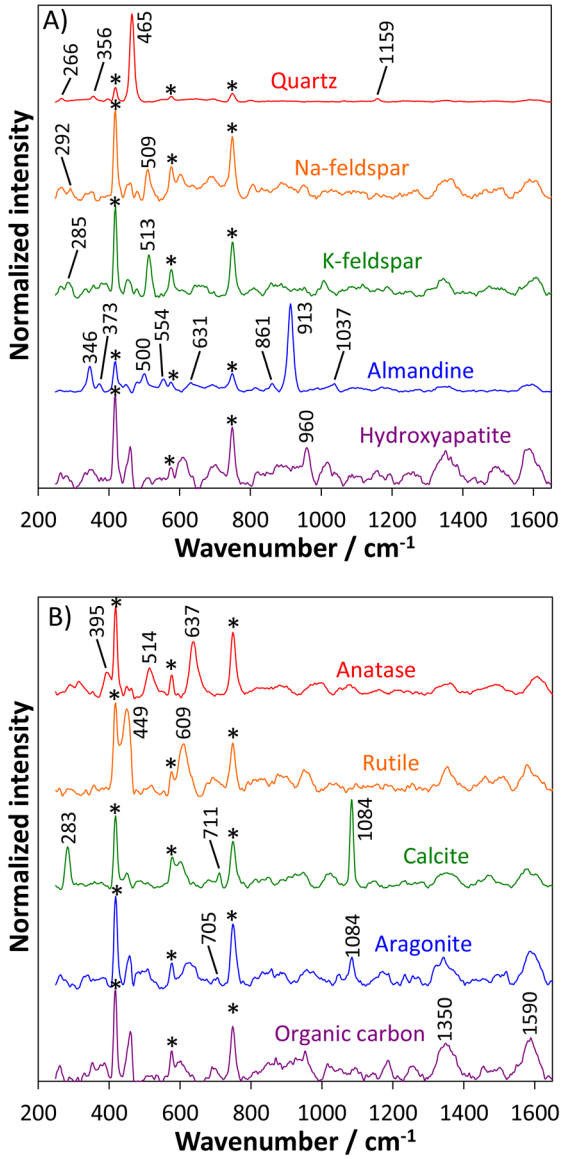

Fig. 2 Averaged SERDS spectra of soil constituents measured with the portable SERDS instrument including silicates and phosphate (A) as well as titanium dioxides, carbonates, and organic carbon (B). Each spectrum represents the average of 2 measurement positions at the field, except for almandine that was only detected at one single position. For clarity, the spectrum of quartz was subtracted from the other acquired spectra. Asterisks indicate Raman signals of the sapphire measurement window contained in the Raman probe. Spectra are normalized to their respective maximum and are vertically offset for clarity.

Beside the identified silicates and phosphate, a few other substances were detected during the on-site investigations with the portable SERDS system and their spectra are displayed in Fig. 2B. Titanium dioxide can appear in several polymorphic forms that share the same chemical composition (TiO_2) but possess different crystal structures.[31] Thus, a clear discrimination between the polymorphs anatase and rutile can be realized by means of the molecular fingerprint obtained by Raman spectroscopy and SERDS as exemplarily depicted in the top part of Fig. 2B. Anatase as first detected TiO_2 polymorph is characterized by three prominent Raman signals. The symmetric Ti–O stretching vibration is located at 637 cm^{-1} while O–Ti–O bending vibrational bands can be observed at 395 cm^{-1} and 514 cm^{-1}.[32,33] Rutile is the second TiO_2 polymorph that has been identified on-site and its SERDS spectrum is characterized by two prominent Raman signals. The Ti–O symmetric stretching vibration can be found at 609 cm^{-1} while the antisymmetric O–Ti–O bending vibration is located at 449 cm^{-1}.[33–35]

Another group of substances showing polymorphic behavior that has been detected and identified during the field measurement are two calcium carbonate ($CaCO_3$) species with distinct Raman spectroscopic signatures as displayed in the center of Fig. 2B. Calcite (trigonal crystal system) and aragonite (orthorhombic crystal system) both exhibit their strongest Raman signal attributed to the symmetric C–O stretching vibration of the CO_3 group at 1084 cm^{-1}. A distinction can, however, be achieved considering additional weaker Raman signals. The asymmetric bending vibration of the CO_3 group can be observed at 711 cm^{-1} (calcite) and 705 cm^{-1} (aragonite). In case of calcite, a lattice vibration involving libration of the CO_3 groups is visible at 283 cm^{-1}.[19,36]

Organic carbon can be detected on-site as well. Disordered carbonaceous matter is characterized by two broad Raman signals. The G-band observable at 1590 cm^{-1} is attributed to a C–C stretching vibration whereas the D-band located at 1350 cm^{-1} arises from a breathing vibration of benzene rings.[20]

Overall, nine mineral soil components as well as organic carbon could be identified from the SERDS spectra recorded on-site. These finding are in accordance with our previous laboratory-based SERDS studies[12,13] on soil samples collected from the same agricultural field as probed in this study. As expected, the previous SERDS laboratory investigations probing 15.000 different measurement positions in total were able to identify some additional materials that were present at very low abundance, *e.g.* the silicate minerals spessartine, zircon and diopside, as well as the carbonate mineral dolomite. It should also be noted that for the previous investigations soil samples collected down to depths of 1 m were analyzed while the present on-site study was only accessing the upper 10 cm. Consequently, soil constituents primarily found in deeper soil layers were more unlikely to be detected in the uppermost soil layer in this study. Nevertheless, it is remarkable that based on SERDS spectra obtained on-site from only 250 measurement positions probed across the agricultural field a clear identification of 10 most abundant soil constituents is feasible.

The demonstrated on-site detection of the major soil components present on the investigated agricultural field is an important step to derive informed conclusions about the soil composition and its overall state. This molecule-specific characterization also provides important insights for the quantitative soil analysis presented below, *e.g.* about the presence of specific components as well as their relative abundance.

Taking Ca-bearing species as an example, calcium carbonates (calcite and aragonite) were detected much more frequently (30 out of 250 spectra) than calcium phosphates (hydroxyapatite) (2 out of 250 spectra) while calcium sulfates were not detected at all on the investigated field. Moreover, characteristic indicator Raman signals present for individual substances can be used as input for a database providing the basis for an automatable evaluation of recorded soil spectra in terms of detected molecular species. Qualitative analysis has shown the capability of SERDS to discriminate between the inorganic (*i.e.* carbonates) and organic soil carbon fractions and these two parameters were exemplarily selected for quantitative analysis.

Determination of soil carbonate content

Within the soil, calcium (*e.g.* present in the form of calcium carbonates) has several functions and plays an important role as a pH regulator and to promote plant growth.[37] Soil calcium content has also implications for liming requirements, *e.g.* in precision agriculture.[38] Moreover, calcite has been shown to be a nucleation and growth site influencing the transformation and mobility of dissolved calcium orthophosphate species in soils.[39] Consequently, the average SERDS spectra obtained from each measurement point on the field will be correlated with calcium carbonate contents calculated from the elemental calcium contents determined by XRF analyses to evaluate if the spectral variations present in the data set can be used to assess this important soil parameter quantitatively.

Total calcium contents of soil samples can be assessed using XRF as an established technique for the analysis of elements. Outcomes are integral contents of Ca but without the ability to obtain molecule-specific information about individual species, *e.g.* calcium phosphates, calcium sulfates, or calcium carbonates. The key benefit of Raman spectroscopy as complementary analytical technique is the capability to distinguish between these Ca species based on their distinct molecular fingerprints. As mentioned above, our qualitative on-site SERDS investigations have shown that at the investigated measurement positions calcium carbonate is the only calcium-containing molecular species that is present in significant amounts while calcium phosphate in the form of hydroxyapatite is only present at low abundance. Consequently, in our case it is reasonable to estimate soil calcium carbonate contents based on the elemental calcium contents determined by XRF.

The calcium carbonate contents predicted from the SERDS data using a PLS model with four factors are depicted in dependence of the calcium carbonate contents derived from the XRF analysis in Fig. 3A. Overall, a good prediction of the soil calcium carbonate contents with a coefficient of determination of $R^2 = 0.86$ can be realized based on the SERDS spectra acquired on-site with the portable SERDS instrument. The root mean squared error of cross-validation (RMSECV) of 2.49% highlights that quantitative predictions of the soil calcium carbonate content can be realized based on a rather small number of 10 probed spots for each measurement position on the field. The largest uncertainty can be observed in the range

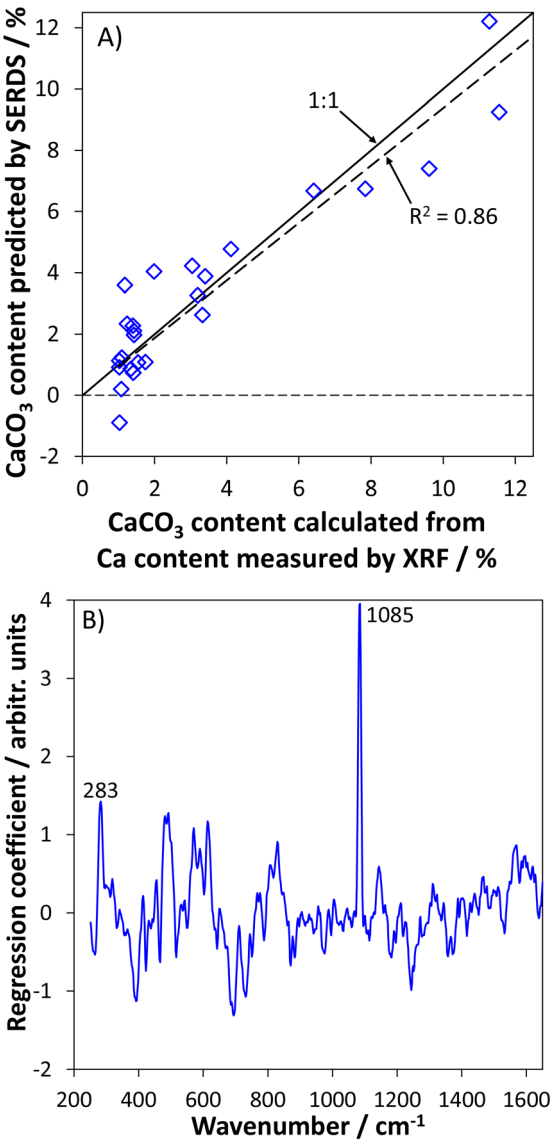

Fig. 3 Soil $CaCO_3$ content predicted from SERDS spectra acquired on-site with the portable SERDS system using PLS regression model with four factors (spectral range 250–1650 cm^{-1}) plotted in dependence of soil $CaCO_3$ content calculated from Ca content measured by XRF analysis (dashed line: linear fit, solid line: 1 : 1 dependence) (A) and corresponding regression coefficient (B).

of the lowest investigated concentrations below 2% $CaCO_3$. The slope of the regression line is 0.94 indicating a trend towards a minimal under-estimation of the actual calcium carbonate contents. A potential explanation for this behavior is that XRF measures the overall concentration of all Ca species present in the sample. In this way, it is possible that Ca species other than carbonates, *e.g.* phosphates or sulfates, contribute to the values measured by XRF as well. Such other species (among others also elemental Ca in dissolved or adsorbed but exchangeable form) are not necessarily detected in the recorded Raman spectra due to low abundance and/or as they do not give rise to intense Raman signals. However, in

sum this contribution of Ca species being detected by XRF but being not observable in the SERDS spectra acquired on site could be a potential reason for the obtained slight under-estimation of the calcium carbonate values when using XRF integral Ca contents as reference values. Moreover, it should also be noted that SERDS spectra were recorded on untreated soil directly on the field whereas XRF analysis was performed on air-dried and sieved specimens under laboratory conditions, *i.e.* not exactly the same soil was probed in both cases.

To assess the underlying spectral contributions that are responsible for the obtained prediction of the calcium carbonate content, a plot of the regression coefficient is given in Fig. 3B. Here, the two strongest signal contributions are due to the calcium carbonate polymorph calcite. The Raman signal at 1085 cm^{-1} originates from the symmetric C–O stretching vibration while the signal at 285 cm^{-1} can be attributed to a lattice vibration.

Assessment of soil organic carbon content

Soil organic carbon (SOC) is the dominating element in soil organic matter (SOM),[40] an important soil component influencing a wide range of soil physical, chemical, and biological properties. It is therefore considered an important indicator of soil quality, functionality, health, as well as agronomic productivity and sustainability. Selected parameters related to SOC include soil structure and aggregation, bulk density, water and plant nutrient retention and use efficiency, microbial biomass and diversity, rhizospheric processes, invertebrate bioindicators (earthworms), and yield.[41–43] Moreover, SOM has the potential to become a sink for atmospheric carbon dioxide thus playing an important role in mitigating greenhouse gas emissions to reduce the adverse effects of climate change.[44]

Fig. 4A presents the SOC contents predicted from the SERDS data against the SOC contents directly determined by reference analysis using a PLS model with four factors. The SERDS spectra acquired on-site with the portable SERDS instrument permit for a good prediction of the SOC contents with a coefficient of determination of R^2 = 0.89 and a RMSECV of 0.32%. The slope of the regression line is 0.99 and thus very close to the ideal value of one. In contrast to the slight under-estimation observed in the case of the carbonate prediction shown above using Ca contents determined by XRF, here no under-estimation of SOC contents is observed. This is likely due to the fact that no unknown carbon sources are detected by the applied reference analysis method. Consequently, an almost perfect 1 : 1 relation between SOC contents predicted from the SERDS spectra (calibrated on elemental analyses) and those values obtained by the direct reference analysis is achievable.

These outcomes underline that quantitative predictions of the SOC content can be achieved despite the rather small number of 10 probed spots for each measurement point on the field. Similar as in the case of the soil carbonate content, intrinsic soil heterogeneity will always present an obstacle to any quantitative analysis but it is remarkable that even under field conditions SERDS measurements with the achieved accuracy can be conducted.

A plot of the regression coefficient is given in Fig. 4B to assess the spectral features responsible for the obtained correlation. A pronounced contribution from the G-band of organic carbon can be identified around 1585 cm^{-1}. Additional strong positive contributions are visible at 281 and 1085 cm^{-1} and these can be attributed to the calcium carbonate polymorph calcite as outlined in the previous section. Negative contributions in the regression coefficient can be attributed to quartz (464 cm^{-1}) and the titanium dioxide polymorph anatase (394 cm^{-1}, 517 cm^{-1}, and 635 cm^{-1}) indicating that these soil constituents do not play a significant role for the SOC content prediction.

It is not surprising to retrieve Raman signals of carbonates in the prediction of the SOC content as a coincidence between soil organic matter (SOM; content directly proportional to SOC

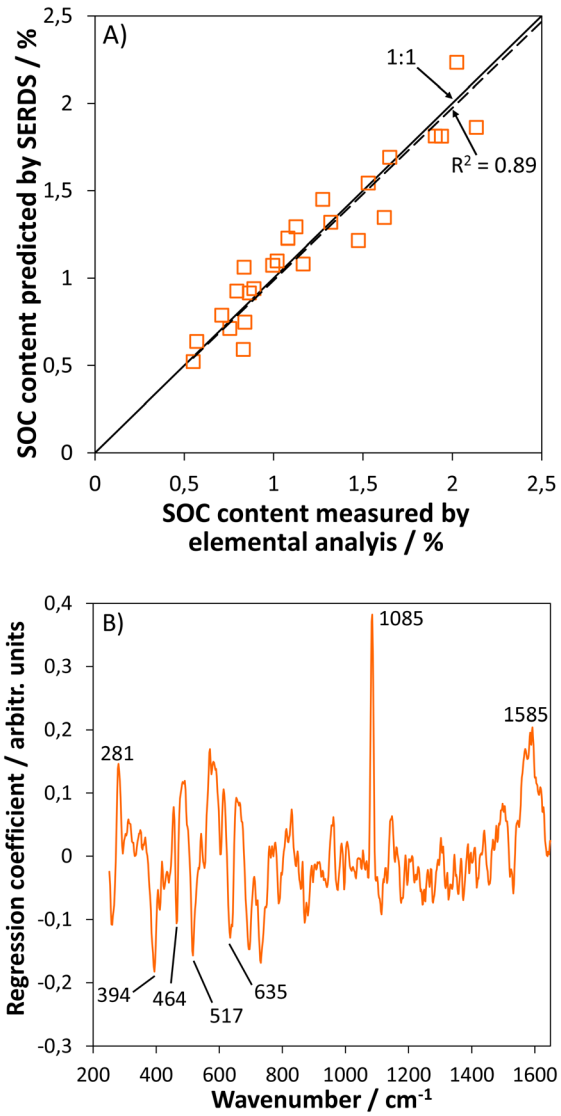

Fig. 4 Soil organic carbon (SOC) content predicted from SERDS spectra acquired on-site with the portable SERDS system using PLS regression model with four factors (spectral range 250–1650 cm^{-1}) plotted in dependence of soil organic carbon content determined by elemental analysis (dashed line: linear fit, solid line: 1 : 1 dependence) (A) and corresponding regression coefficient (B).

content[40]) and calcite has been reported in the literature for certain soil types.[29] Moreover, our previous SERDS laboratory investigations on soil samples collected from the same agricultural field have shown a spatial correlation between the occurrence of calcite and the presence of SOM.[12] Below, it is illustrated that also in the present investigations the measurement positions with the highest carbonate contents coincide with regions where elevated SOC concentrations are observed.

Comparison with SERDS laboratory measurements

For comparison, soil samples were collected from the 25 measurement points probed on the field. SERDS investigations by means of a laboratory setup were conducted on these samples probing 100 individual spots each[12,13] to validate the results obtained on-site. For qualitative soil analysis, the 2.500 SERDS spectra recorded in total were then assessed for the presence of individual soil constituents.

The results obtained with the SERDS laboratory investigations demonstrate that the same major soil components could be detected as for the field measurements thus highlighting the capability of the portable SERDS system for rapid qualitative on-site soil analysis. In case of the minor constituents with an abundance well below 1% there is some difference between the on-site and the laboratory SERDS measurements. While almandine was only detected on-site (1 out of 250 spectra), the investigations conducted with the SERDS laboratory setup probing a 10-times larger number of spots revealed the additional presence of spessartine ($Mn_3Al_2Si_3O_{12}$) (1 out of 2.500 spectra) and zircon ($ZrSiO_4$) (3 out of 2.500 spectra).

In case of quantitative soil analysis, average SERDS spectra obtained from the 25 collected soil samples were subjected to PLSR to assess their capability to predict soil carbonate and soil organic carbon content. Key parameters to assess the model performance are compiled in Table 1.

For the prediction of the carbonate content, the model based on the SERDS laboratory investigations used the same number of 4 factors as the model obtained using the on-site acquired SERDS data. The data acquired with the laboratory setup permit to obtain a model with 10% more explained variance, an increased coefficient of determination R^2 and a slope of the linear fit that is slightly closer to the ideal value of one. Moreover, the RMSECV in case of the model generated using the SERDS laboratory data is improved by 0.75% leading to more accurate predictions of the soil carbonate content.

In case of the soil organic carbon content the PLSR model based on the SERDS data obtained on-site uses 4 factors while the model based on the spectra acquired using the SERDS laboratory setup incorporates only 3 factors. While in both cases an identical slope of the linear fit that is very close to the ideal value of one is achieved, the data acquired with the laboratory setup lead to a model with *ca.* 3% more explained variance and an increased coefficient of determination R^2. Furthermore, the RMSECV in case of the model obtained using the SERDS laboratory data is improved by 0.1% leading to more accurate predictions of the soil organic carbon content.

It is not unexpected that the PLSR models calculated from the data acquired with the SERDS laboratory setup show improved performance compared to the models obtained from the data set acquired on-site with the portable SERDS instrument. The most important reason for the observed difference is very likely due to the intrinsic soil heterogeneity at the sub-mm scale. While both measurement systems apply similar measurement spot sizes (115 μm for the portable system *vs.* 100 μm for the laboratory setup), the number of probed spots is one order of magnitude larger in case of the laboratory investigations (100 spots per sample) compared to the on-site study (10 spots per measurement point). Consequently, the total area probed per sample is 10-times larger for the laboratory investigations and can thus more accurately capture variations due to the intrinsic soil heterogeneity. This, in turn, leads to more accurate quantitative predictions of important soil parameters, as *e.g.* demonstrated in case of the soil carbonate and soil organic carbon content.

Regarding the application of Raman spectroscopy for soil carbonate quantification, only little research has been done previously. One example is our recent study applying SERDS for the prediction of the soil carbonate content.[13] Using a set of 117 samples collected from the same agricultural field in Germany that is also probed in the current study, a successful prediction of the carbonate content with R^2 = 0.94 and RMSECV = 2.1% was achieved. Based on a much smaller number of samples, our present on-site study shows slightly worse but comparable prediction performance for the soil carbonate content.

Table 1 Details of PLSR models applied for the prediction of soil $CaCO_3$ content and soil organic carbon (SOC) content based on SERDS data obtained on-site using a portable instrument or SERDS data acquired by means of a laboratory setup

	$CaCO_3$ content		SOC content	
Model parameter	Portable SERDS system (on-site)	SERDS setup (laboratory)	Portable SERDS system (on-site)	SERDS setup (laboratory)
Number of factors	4	4	4	3
Variance explained/%	87.2	97.3	89.8	93.1
R^2	0.86	0.97	0.89	0.93
Slope of linear fit	0.94	0.99	0.99	0.99
RMSECV/%	2.49	1.74	0.32	0.22

In case of the soil organic carbon content, also a brief comparison with selected literature using Raman spectroscopy for organic matter quantification should be given. As an example, conventional Raman spectroscopy applying mathematical background correction was used for the analysis of Chinese farmland soils.[45] Based on a data set of 146 spectra, the soil organic matter (SOM) content was predicted with R^2 = 0.71 and RMSECV = 0.81%. Using a much smaller number of 33 soil samples, our previous SERDS laboratory investigation could demonstrate a successful prediction of the SOM content with R^2 = 0.82 and RMSECV = 0.41%. It should be noted that contents of SOM and SOC are not equal but can be calculated from each other. Assuming that SOM contains approximately 58% of organic carbon, SOM content can be converted to SOC content by dividing it by a factor of 1.72.[46] In this way, our on-site SERDS study shows comparable prediction performance as the two above-mentioned laboratory-based Raman investigations.

Assessment of spatial variability of selected soil constituents

The obtained results show that soil characterization can be performed directly on site without any sample preparation using the portable SERDS instrument. This capability can be used to capture spatial heterogeneity of the soil composition across an agricultural field. This is exemplarily depicted in Fig. 5A where the calcium carbonate contents predicted by PLSR using SERDS on-site and laboratory measurements are plotted in dependence of the measurement position along the field. Both curves show a very similar behavior with elevated calcium carbonate concentrations above 4% between *ca.* 250–350 m located from the first measurement point. On average, the difference between the carbonate concentrations predicted from on-site measurements *versus* laboratory investigations amounts to only 0.88% while the maximum observed deviation is 2.95%.

A similar behavior can be observed for the SOC content whose spatial distribution as predicted by PLSR based on the on-site as well as SERDS laboratory investigations is displayed in Fig. 5B. Both curves present a similar shape and indicate that the highest present SOC concentrations above 1.5% are located in a range *ca.* 200–350 m away from the first measurement point. The average difference between the SOC concentrations predicted from on-site measurements *versus* laboratory investigations is only 0.14% while the maximum observed deviation amounts to 0.44%.

These findings highlight the capability of the portable SERDS system to detect present differences in carbonate and SOC contents even between adjacent measurement positions that are separated by only 15 m. Such mapping capabilities at high spatial resolution on the order of 10 m are an important prerequisite for site-specific liming and fertilization recommendations in the context of precision agriculture. It is expected that portable SERDS can provide a valuable contribution to spatially-resolved on-site soil characterization, *e.g.* to deliver valuable information for decision support systems for precision agriculture.

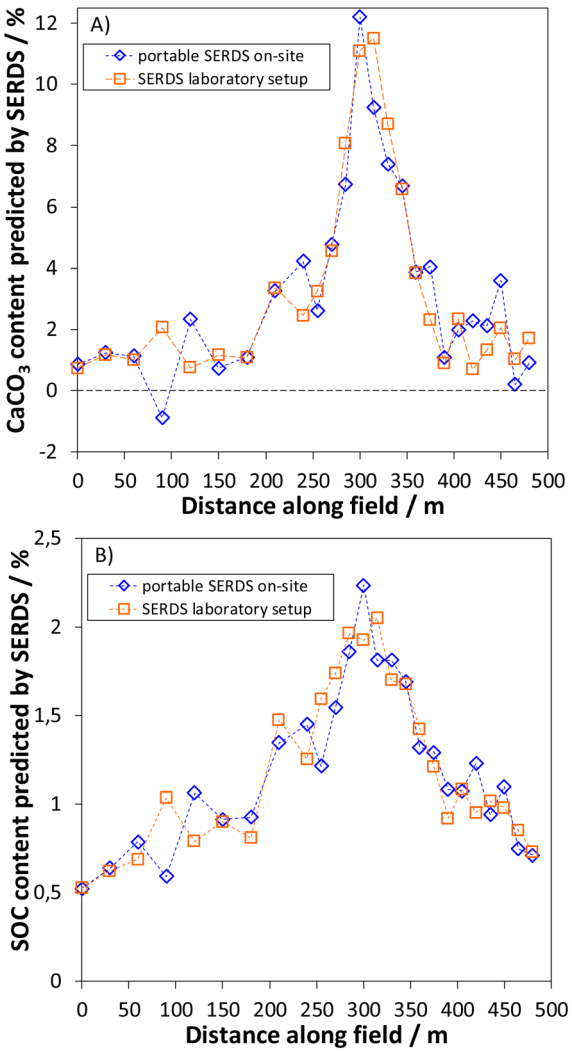

Fig. 5 Spatial distribution of $CaCO_3$ content (A) and soil organic carbon (SOC) content (B) predicted from SERDS spectra using PLSR plotted in dependence of the distance along the field (comparison between on-site measurements with portable SERDS system and investigations using a SERDS laboratory setup).

Conclusions

This work has shown the first application of an in-house developed portable SERDS instrument for on-site soil analysis directly on an agricultural field. SERDS permitted for an efficient separation of the Raman spectroscopic information from interfering contributions such as fluorescence and daylight. In this way, 9 mineral soil constituents (silicates, carbonates, titanium dioxides, and phosphate) as well as organic carbon could be detected and identified. In terms of quantitative analysis, successful predictions of the soil carbonate content (R^2 = 0.86, RMSECV = 2.49%) and of the SOC content (R^2 = 0.89, RMSECV = 0.32%) could be realized. The SERDS spectra acquired on-site from 25 measurement points also permitted to track spatial variations of these two parameters along a distance of 480 m across the field thus demonstrating

SERDS as substance-specific technique for spatially-resolved soil analysis. Overall, the presented results clearly demonstrate that the portable SERDS instrument has a large potential as rapid screening tool for qualitative and quantitative on-site soil characterization directly on agricultural fields, *e.g.* to provide input data for decision support systems in the context of precision agriculture. Combining SERDS with selected other sensing approaches tested for precision agriculture, such as near- and mid-infrared diffuse reflectance spectroscopy and gamma spectrometry,[47] may yield great benefit in the context of sensor fusion approaches, as the sensing principles provide complementary information.

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to confidentiality requirements but are available from the corresponding author upon reasonable request.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This study was funded by the Federal Ministry of Education and Research (BMBF) within the projects RaMBo (grant numbers 031A564C, 031B0513C and 031B1069C) within the consortium I4S (Intelligence for Soil) in the frame of the funding measure BonaRes (Soil as a Sustainable Resource for the Bioeconomy) and Research Fab Microelectronics Germany – FMD (grant number 16FMD02). We would like to thank all involved colleagues from I4S for sample collection and preparation, and the landowner Golo Philipp for access to the agricultural field. We are grateful to Anke Lindecke for conducting the SOC analyses and to Jeannine Gleim for technical assistance during the on-site field measurement.

References

1 I. Bhakta, S. Phadikar and K. Majumder, *J. Sci. Food Agric.*, 2019, **99**, 4878–4888.
2 M. R. Nanni, F. P. Povh, J. A. M. Demattê, R. B. de Oliveira, M. L. Chicati and E. Cezar, *Sci. Agric.*, 2011, **68**, 386–392.
3 S. Pätzold, F. M. Mertens, L. Bornemann, B. Koleczek, J. Franke, H. Feilhauer and G. Welp, *Precis. Agric.*, 2008, **9**, 367–390.
4 M. Debeljak, A. Trajanov, V. Kuzmanovski, J. Schröder, T. Sandén, H. Spiegel, D. P. Wall, M. Van de Broek, M. Rutgers, F. Bampa, R. E. Creamer and C. B. Henriksen, *Front. Environ. Sci.*, 2019, **7**, 115.
5 P. M. Nkebiwe, K. Sowoidnich, M. Maiwald, B. Sumpf, T. E. Hartmann, D. Wanke and T. Müller, *J. Plant Nutr. Soil Sci.*, 2022, **185**, 221–231.
6 K. Ikehata, Y. Arakawa and J. Ishibashi, *Vib. Spectrosc.*, 2021, **114**, 103247.
7 A. P. Shreve, N. J. Cherepy and R. A. Mathies, *Appl. Spectrosc.*, 1992, **46**, 707–711.
8 J. Zhao, M. M. Carrabba and F. S. Allen, *Appl. Spectrosc.*, 2002, **56**, 834–845.
9 M. Maiwald, K. Sowoidnich and B. Sumpf, *J. Raman Spectrosc.*, 2022, **53**, 1560–1570.
10 M. Maiwald, A. Müller, B. Sumpf, G. Erbert and G. Tränkle, *Appl. Opt.*, 2015, **54**, 5520–5524.
11 L. S. Theurer, M. Maiwald and B. Sumpf, *Eur. J. Soil Sci.*, 2021, **72**, 120–124.
12 K. Sowoidnich, S. Vogel, M. Maiwald and B. Sumpf, *Appl. Spectrosc.*, 2022, **76**, 712–722.
13 K. Sowoidnich, M. Maiwald, M. Ostermann and B. Sumpf, *J. Raman Spectrosc.*, 2023, **54**, 1327–1340.
14 A. Müller, M. Maiwald and B. Sumpf, *Proc. SPIE*, 2022, **12021**, 1202103.
15 B. Sumpf, L. Wittenbecher, T. Filler, D. Bandke, M. Krichler, A. Müller, K. Sowoidnich, A. Ginolas, U. Winterwerber and M. Maiwald, *Instrum. Sci. Technol.*, 2024, **52**, 138–150.
16 S. De Jong, *Chemom. Intell. Lab. Syst.*, 1993, **18**, 251–263.
17 C. M. Andersen and R. Bro, *J. Chemom.*, 2010, **24**, 728–737.
18 J. Jehlička, P. Vítek, H. G. M. Edwards, M. Heagraves and T. Čapoun, *Spectrochim. Acta, Part A*, 2009, **73**, 410–419.
19 M. Veneranda, A. Sanz-Arranz, J. A. Manrique, J. Saiz, C. Garcia-Prieto, E. Pascual-Sánchez, J. Medina, M. Konstantinidis, E. Lalla, A. Moral, L. M. Nieto, F. Rull and G. Lopez-Reyes, *J. Raman Spectrosc.*, 2022, **53**, 364–381.
20 A. Ferrugiari, M. Tommasini and G. Zerbi, *J. Raman Spectrosc.*, 2015, **46**, 1215–1224.
21 D. T. M. Phan, T. Häger and W. Hofmeister, *J. Raman Spectrosc.*, 2017, **48**, 453–457.
22 W. Zhu and G. Pezzotti, *J. Appl. Phys.*, 2011, **109**, 073502.
23 J. L. White, *Soil Sci.*, 1971, **112**, 22–31.
24 S. Kramar, M. Urosevic, H. Pristacz and B. Mirtič, *J. Raman Spectrosc.*, 2010, **41**, 1441–1448.
25 I. Aliatis, E. Lambruschi, L. Mantovani, D. Bersani, S. Andò, G. Diego Gatta, P. Gentile, E. Salvioli-Mariani, M. Prencipe, M. Tribaudino and P. P. Lottici, *J. Raman Spectrosc.*, 2015, **46**, 501–508.
26 M. Veneranda, J. A. Manrique-Martinez, C. Garcia-Prieto, A. Sanz-Arranz, J. Saiz, E. Lalla, M. Konstantinidis, A. Moral, J. Medina, F. Rull, L. M. Nieto and G. Lopez-Reyes, *Icarus*, 2021, **367**, 114542.
27 P. Makreski, G. Jovanovski and S. Stojančeska, *J. Mol. Struct.*, 2005, **744–747**, 79–92.
28 S. Kos, M. Dolenec, J. Lux and S. Dolenec, *Minerals*, 2020, **10**, 325.
29 C. Vogel, M. Ramsteiner, R. Sekine, A. Doolette and C. Adam, *J. Raman Spectrosc.*, 2017, **48**, 867–871.

30 A. M. Lanfranco, P. F. Schofield, P. J. Murphy, M. E. Hodson, J. F. W. Mosselmans and E. Valsami-Jones, *Mineral. Mag.*, 2003, **67**, 1299–1316.
31 U. Diebold, *Surf. Sci. Rep.*, 2003, **48**, 53–229.
32 M. A. Boda and M. A. Shah, *Mater. Res. Express*, 2017, **4**, 075908.
33 A. J. Pinto, N. Sanchez-Pastor, I. Callegari, B. Pracejus and A. Scharf, *Sci. Rep.*, 2020, **10**, 7445.
34 I. Lukačević, S. K. Gupta, P. K. Jha and D. Kirin, *Mater. Chem. Phys.*, 2012, **137**, 282–289.
35 Y. Zhang, C. X. Harris, P. Wallenmeyer, J. Murowchick and X. Chen, *J. Phys. Chem. C*, 2013, **117**, 24015–24022.
36 L. Borromeo, U. Zimmermann, S. Andò, G. Coletti, D. Bersani, D. Basso, P. Gentile, B. Schulz and E. Garzanti, *J. Raman Spectrosc.*, 2017, **48**, 983–992.
37 M. Rühlmann, D. Büchele, M. Ostermann, I. Bald and T. Schmid, *Spectrochim. Acta, Part B*, 2018, **146**, 115–121.
38 M. Leenen, G. Welp, R. Gebbers and S. Pätzold, *J. Plant Nutr. Soil Sci.*, 2019, **182**, 953–963.
39 L. Wang, E. Ruiz-Agudo, C. V. Putnis, M. Menneken and A. Putnis, *Environ. Sci. Technol.*, 2012, **46**, 834–842.
40 D. W. Pribyl, *Geoderma*, 2010, **156**, 75–83.
41 K. L. Page, Y. P. Dang and R. C. Dalal, *Front. Sustainable Food Syst.*, 2020, **4**, 1–17.
42 R. Lal, *Food Energy Secur.*, 2016, **5**, 212–222.
43 D. W. Reeves, *Soil Tillage Res.*, 1997, **43**, 131–167.
44 U. Stockmann, M. A. Adams, J. W. Crawford, D. J. Field, N. Henakaarchchi, M. Jenkins, B. Minasny, A. B. McBratney, V. D. R. de Courcelles, K. Singh, I. Wheeler, L. Abbott, D. A. Angers, J. Baldock, M. Bird, P. C. Brookes, C. Chenu, J. D. Jastrow, R. Lal, J. Lehmann, A. G. O'Donnell, W. J. Parton, D. Whitehead and M. Zimmermann, *Agric., Ecosyst. Environ.*, 2013, **164**, 80–99.
45 Z. Xing, C. Du, K. Tian, F. Ma, Y. Shen and J. Zhou, *Talanta*, 2016, **158**, 262–269.
46 K. I. Peverill, L. A. Sparrow and D. J. Reuter, *Soil analysis: An interpretation manual*, CSIRO Publishing, Collingwood, Australia, 1999.
47 H. Tavakoli, J. Correa, S. Vogel, M. Oertel, M. Zimne, M. Heisig, A. Harder, R. Wruck, S. Pätzold, M. Leenen and R. Gebbers, *Comput. Electron. Agric.*, 2024, **226**, 109443.