



openaccess

Using visible and near infrared spectroscopy to estimate carbonates and gypsum in soils in arid and subhumid regions of Isfahan, Iran

Fatemah Khayamim,^a Johanna Wetterlind,^b Hossein Khademi,^a A.H. Jean Robertson,^c Angel Faz Cano^d and Bo Stenberg^{b,*}

^aIsfahan University of Technology, Department of Soil Science, Isfahan 84156-83111, Islamic Republic of Iran

^bSwedish University of Agricultural Sciences, Department of Soil and Environment, PO Box 234, 532 23 Skara, Sweden.

E-mail: bo.stenberg@slu.se

^cThe James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, Scotland, UK

^dTechnical University of Cartagena, Department of Agricultural Production, Paseo Alfonso XIII, 48. 30203 Cartagena, Spain

Soils in arid and semi-arid regions are strongly affected by the accumulation of carbonates, gypsum and other, more soluble, salts. Carbonates and gypsum both have a considerable influence on soil properties, especially the chemical properties of the soil solution. The development of reliable, fast and inexpensive methods to quantify the amounts of carbonates and gypsum in soil is therefore important. Visible and near infrared (vis-NIR) spectroscopy is a non-destructive, rapid and cheap method for measuring several soil properties simultaneously. However, research on vis-NIR spectroscopy in quantifying carbonates and gypsum is limited. Therefore, this study evaluated the efficiency of vis-NIR spectroscopy in quantifying carbonates and gypsum in surface soils using partial least-squares regression (PLSR) compared with standard laboratory methods and compared PLSR with a feature-specific method using continuum removal (CR). Carbonates and gypsum in a total of 251 sieved and air-dried topsoil samples from Isfahan Province in central Iran were measured by standard laboratory methods and vis-NIR spectroscopy (350–2500 nm wavelength range). In parallel, PLSR and the feature-specific method based on CR spectra were used to predict carbonates and gypsum. The PLSR model efficiency (E) for carbonates and gypsum in the validation set was 0.52 and 0.80, respectively. The PLSR model resulted in better predictions than the feature-specific method for both soil properties. Because of the unique absorption features of gypsum, which did not overlap with other soil properties, predictions of gypsum resulted in higher E values and lower errors than predictions of carbonates.

Keywords: gypsum, carbonates, vis-NIR spectroscopy, continuum removal, partial least-squares regression (PLSR)

Introduction

Around 40% of the Earth's terrestrial surface is covered by arid and semi-arid ecosystems.^{1–3} Arid and semi-arid regions are characterised by a climate with no or insufficient rainfall

to sustain agricultural production. The soils in such regions are unique and are strongly affected by the accumulation of carbonates, gypsum and other, more soluble, salts because

relatively little water percolates deep enough to reach the groundwater.⁴ Calcium carbonate and gypsum are the most common carbonate and sulfate polymorphs in soils in general and are particularly abundant under arid, semi-arid and dry subhumid conditions.⁵

Carbonates are less soluble than gypsum and are therefore more resistant to leaching. Consequently, carbonates are more widely distributed in soils.⁶ The gypsum content ranges from about 1% to 100%, depending on climate, topography and physiography.^{7,8} Gypsum is found over a wide range of temperatures, but most gypsic soils occur in xeric, ustic and aridic soil moisture regimes.⁹ Gypsum usually occurs together with calcite and other soluble salts in soils. Carbonates and gypsum both have a considerable influence on soil properties.⁶ For example, studies on gypsiferous soils have shown that a concentration of >10% gypsum significantly interferes with soil characteristics such as structure, consistency and water-holding capacity. In soils containing 10–25% gypsum, the gypsum crystals tend to break the continuity of soil mass, while soils with more than 25% gypsum do not provide a good medium for plant growth, lack plasticity, cohesion and aggregation, and become completely unstable in water.¹⁰

Therefore, it is necessary to quantify accurately the amounts of carbonates and gypsum in soil. Development of more timely and cost-effective methods for detecting and quantifying carbonates and gypsum in soil with reliable precision is consequently important.

There are several methods available for measuring gypsum in soil, including the thermogravimetric method, wet chemical method, electroconductometric determination, determination of sulfate by ion chromatography and semi-quantitative X-ray diffraction.¹¹ All these methods have their limitations and advantages, but most are associated with large errors.^{11,12} The thermogravimetric and electroconductometric determination (standard acetone) methods are mostly commonly used by researchers because of their simplicity. The thermogravimetric approach is based on the loss of mass when gypsum is dehydrated and is recommended when the sample contains more than 8% gypsum.¹³ It can be a good semi-quantitative method but overestimates the gypsum content in many soils because other salts present are also dehydrated on heating, although to a lesser extent. In the electroconductometric method, the amount of gypsum dissolved in water can be determined by an electrical conductimeter.^{14,15} Sulfate ions from calcium sulfate dissolved in water (soil extract) are precipitated with calcium from calcium chloride in acetone. The precipitate is completely redissolved in water, and the electrical conductivity is measured and transformed to CaSO₄ content using an established relationship.

Soil carbonate is usually quantified by an acid-dissolution method involving determination of H⁺ consumption or Ca (and Mg) or CO₂ production.¹⁶ Alternatively, a dry combustion procedure, based on precombustion of organic matter at 575°C in an O₂ stream and subsequent combustion of carbonates at 1000°C and collection of CO₂, has been reported.¹⁷ Methods involving determination of CO₂ are generally preferred. The

CO₂ released in acid-dissolution methods can be measured gravimetrically,¹⁶ titrimetrically,¹⁸ manometrically,¹⁹ volumetrically,²⁰ spectrophotometrically by infrared spectroscopy or by gas chromatography.²¹ For other methods based on H⁺ consumption that involve reaction with a strong acid, such as HCl, back-titration of the unreacted acid is usually used for determination. The choice of procedure depends to a large extent on the equipment available to the researcher.²¹

Over the last 20 years, application of visible and near infrared (vis-NIR) diffuse reflectance spectroscopy in soil science has attracted great attention.^{22,23} Diffuse reflectance spectroscopy (DRS) is non-destructive and inexpensive, and information can be easily collected from small quantities of sample. Furthermore, DRS can be performed *in situ*, providing information under field conditions and within a short time. Many investigations have shown that vis-NIR spectroscopy is a useful and reliable method for evaluation of a number of soil properties.^{24–31}

Carbonates have spectral features in the vis-NIR spectrum at 2300–2350 nm and 2500–2550 nm, owing to combination and overtone bands of the CO₃ fundamentals.³² The absorption features of gypsum occur near 1000, 1200, 1400, 1600, 1740, 1900 and 2200 nm owing to combinations of O–H stretches and H–O–H bending fundamentals, as explained by Hunt *et al.*³³ While there are some reports estimating calcium carbonate in soils by vis-NIR spectroscopy,^{26,30,34–36} to our knowledge no such studies have been performed previously for gypsum.

Quantitative spectral analyses of soil properties using vis-NIR spectroscopy require techniques to differentiate the response of soil attributes from spectral characteristics. Partial least-squares regression (PLSR) is currently the most common calibration technique used for the prediction of soil properties, mainly owing to its robustness.³⁷ PLSR uses the full spectrum to establish a linear regression model for predicting soil properties. For soil properties with well-defined spectral features at specific bands, such as clay minerals, methods focusing on that specific feature instead of using the full spectrum could be an alternative.^{25,38} The continuum removal (CR) technique as a feature-specific method isolates the specific absorption features of materials at a specific wavelength, assuming that no other material has strong absorption features around this specific wavelength.³⁹ Continuum removal, which is known as a spectral normalisation method, has been used in soil science, mostly on remotely sensed spectra.⁴⁰ Using reflectance values at specific wavelengths from continuum-removed spectra, Viscarra Rossel³⁸ quantified the amounts of kaolinite, illite and smectite in surface soils of Australia, while Madeira Netto *et al.*⁴¹ applied continuum-removed spectra to map clay and calcite content from HYMAP imaging spectro-radiometer observations.

Owing to the importance of carbonates and gypsum in soils of arid and semi-arid regions, reliable, timely and cost-efficient methods for their detection and quantification need to be developed. Despite all the advantages of vis-NIR spectroscopy, its use in quantification of carbonates and especially gypsum is limited. Therefore, the objectives of this study were: (1) to

evaluate the efficiency of vis-NIR spectroscopy in quantifying carbonates and gypsum in surface soils using PLSR compared with standard laboratory methods; and (2) to compare PLSR with the feature-specific method using continuum removal for carbonate and gypsum predictions.

Materials and methods

Study area

The study area is located in the Isfahan Province, Central Iran ($30^{\circ}42'–34^{\circ}27'N$; $49^{\circ}38'–55^{\circ}32'E$) and covers an area of about $110,000\text{ km}^2$. The elevation ranges from 700 m to 2600 m above sea level. The mean annual temperature ranges from about 10°C to 21°C from west to east, and the mean annual precipitation ranges from 560 mm to 90 mm from west to east. The study area is divided into two distinct regions based on climate conditions. The eastern, northern and central parts are characterised by an arid climate, and the southern and western parts by a subhumid climate (Figure 1).

The soil parent materials in the study area are very variable. Based on the geology map,⁴² quaternary sediments, such as clay flats, salt lakes, sand dunes, and silty and clayey deposits, cover large parts of Isfahan Province. Limestone, sandstone, conglomerate and shale mostly occur in southern and western parts. Volcanic rocks occur as a narrow band from north to southeast, and metamorphic rocks are prominent soil parent materials in a small area located in the northwest of the province. Based on clay mineralogical studies, illite, smectite, chlorite and palygorskite are major clay minerals in the soils, and the mineralogy varies from east to west. Palygorskite is

the dominant clay mineral in the east, while smectite is abundant in western parts.

Soil sampling

A total of 251 topsoil samples (0–20 cm) were collected using random sampling within $20\text{ km} \times 20\text{ km}$ blocks (Figure 1). Samples were selected based on parent material by taking one sample from the dominant parent material within each grid cell. This ensured that sedimentary, volcanic and metamorphic rocks were represented. Each sample consisted of five subsamples, which were randomly collected from 2500 m^2 in each grid. All soil samples were air-dried and passed through a 2 mm sieve.

Chemical analyses

A back-titration method was performed to measure carbonates in soils by adding a measured excess of HCl solution, some of which was neutralised by carbonates in the soil. The remaining acid was then titrated with a standardised base, so that the amount neutralised by the carbonates could be calculated. In brief, 1 g of soil was placed in an Erlenmeyer flask, and 10 mL of 1 N HCl was added. The mixture was then heated and stirred until all the carbonates were dissolved (no more bubbles of CO_2 evolved). The unreacted acid in the flask was titrated to a phenolphthalein end-point with 0.5 N NaOH.⁴³ The oven-drying method was used to measure the gypsum content in soils from the arid regions of the study area, while the acetone method was used for soils from the subhumid regions. With the acetone method, soil gypsum was dissolved in a 1:5 soil:water ratio by shaking. Gypsum was precipitated from the soil extract by adding acetone. The precipitated gypsum was then completely dissolved by adding distilled water, and the electrical conductivity of the solution was measured. The gypsum content was determined by comparison with a standard curve relating gypsum content to electrical conductivity.¹⁵

Spectral measurements

The diffuse reflectance spectra of the samples were recorded with a FieldSpec Pro FR scanning instrument (Analytical Spectral Devices, Boulder, Colorado). Spectra were collected at 1.4–2 nm intervals with a spectral resolution of 3–10 nm. A wavelength interval of 1 nm was interpolated to the instrument output file. The spectral range covered both the visible and near infrared regions, 350–2500 nm. Approximately 30 mL of soil was poured into a Petri dish 9 cm in diameter and gently flattened. Measurements were made using a bare optic fibre, which was assembled together with a 20 W Al-coated halogen tungsten light source 7 cm above the soil sample, resulting in a field of view of $\sim 7.5\text{ cm}^2$. Each spectrum comprised 100 averaged subspectra from a rotating sample, covering a total sample area of about 50 cm^2 . Reflectance spectra were recorded in relation to an external white panel (Spectralon®, <http://www.labsphere.com>) scanned once every five soil samples. Three measurements were saved and, after quality control, averaged for each soil sample.

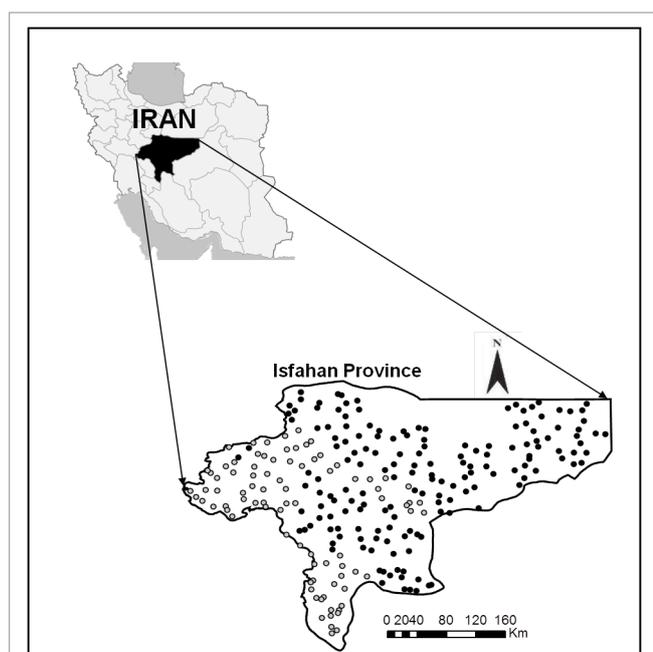


Figure 1. Spatial distribution of the sampling sites within the study area in Isfahan Province, Iran. Black and grey circles show the arid and subhumid climatic conditions, respectively.

Prediction of carbonates and gypsum by PLSR

Reflectance (R) measurements were transformed to apparent absorbance by $\log 1/R$. The first derivative with Savitzki–Golay smoothing over 11 data points was used to reduce the effect of random noise and improve the calibration model.⁴⁴ Owing to excessive noise, the 350–400 nm range was removed after derivation as well as transition domains between sensors (965–977 nm and 1783–1796 nm). A principal-component analysis (PCA) was conducted on the transformed spectra to visualise the data and identify possible spectral outliers.

Vis-NIR spectra were calibrated to carbonate and gypsum using PLSR.⁴⁵ Before making the PLSR calibrations, the dataset was randomly split into calibration ($n = 170$) and validation sets ($n = 81$). Leave-one-out cross-validation was used to select the optimal number of PLSR factors to be used in the model.

The performance of the prediction models with the cross-validation and independent validation sets was evaluated using the coefficient of determination (r^2) for linear regression between predicted and measured values, the model efficiency (E), the root mean squared error ($RMSE$) and the bias, calculated using Equations (1)–(4). The E value indicates the proportion of total variance explained by the model (the 1:1 line) and includes both the relationship between measured and predicted values and systematic errors. In Equations (1) and (2), y denotes the measured value, \hat{y} is the predicted value, n is the number of samples, and SD is the standard deviation of laboratory-measured values for the property in question:

$$E = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

$$Bias = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{n} \quad (4)$$

All spectral analyses, including the calibrations, were carried out using the Unscrambler 10.3 software (CAMO, PROCESS, AS, Oslo, Norway).

Feature-specific prediction of carbonates and gypsum

The feature-specific approach aims to quantify the absorption of materials at a specific band, assuming that no other material has strong absorption features interfering in this specific region.³⁹ The continuum is the background absorption onto which absorption features are superimposed. It can be approximated by a straight line joining two local reflectance maxima placed on both shoulders (λ_{\min} and λ_{\max}) of the feature of interest. CR was thus introduced as a function of reflectance value $R(\lambda)$ at wavelength λ , and its maximum value cannot be above 1.³⁹ This method removes the overall variance caused by albedo, thus increasing and centring the absorption bands and making it possible to compare specific absorption features.⁴⁶

The band depth (BD) of each absorption feature was calculated for each soil sample by subtracting the continuum-removed reflectance (CRR) at a particular wavelength (λ), using the equation $BD(\lambda) = 1 - CRR$. The abundance (A) of carbonates and gypsum in a soil sample was calculated as $A = BD_s / BD_R$, where BD_s is the band depth of the soil sample, and BD_R is the band depth of the sample with the highest amount of carbonate (80%) or gypsum (60%) according to laboratory analyses. Ideally, the reference should have been pure carbonate or gypsum, but that was not possible in our study. The abundance was therefore adjusted by multiplying A by 0.8 and 0.6 for carbonate and gypsum, respectively. CRR spectra were calculated for all samples using the statistical software environment R (R Development Core Team, 2012; <http://www.R-project.org>). The accuracy of prediction of carbonates and gypsum by the feature-specific method was assessed by r^2 , E , $RMSE$ and bias, similarly to the PLSR models.

Results

Summary of laboratory analyses

Descriptive statistics on carbonate and gypsum analyses in calibration and validation sets are presented in Table 1. Both

Table 1. Descriptive statistics of laboratory analyses for carbonates and gypsum in the calibration and validation sets.

	Calibration ($n = 170$)		Validation ($n = 81$)	
	Gypsum	Carbonates	Gypsum	Carbonates
Mean	5.1	29.2	6.2	25.1
Median	3.8	25.6	4.4	20.0
Maximum	61.7	80.0	38.8	67.0
Minimum	0.0	0.2	0.0	0.2
SD	7.7	18.0	7.8	17.0
Skewness	4.1	0.5	2.2	0.7

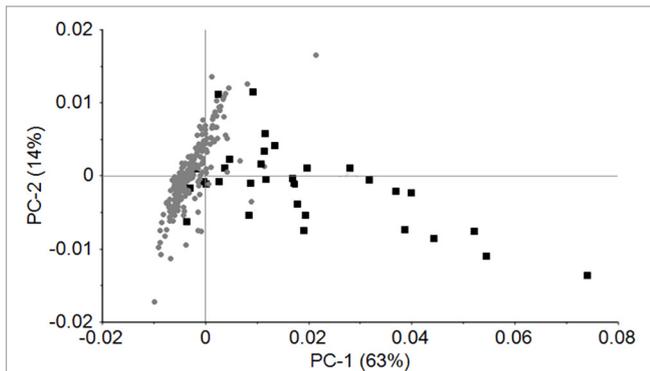


Figure 2. Score plot of the first and second principal components of the first-derivative spectra. Bold marking shows soil samples with more than 10% gypsum based on laboratory analyses.

soil properties vary widely within the study area. The means and standard deviations were similar for the calibration and validation sets, and the gypsum and carbonate contents in the validation set were within the range in the calibration set. Consequently, the validation set appeared to be represented accurately by the calibration set. The carbonate content in the study area was found to follow a normal distribution, whereas gypsum was positively skewed, as only some of the samples contained substantial amounts of gypsum.

Spectral characteristics of the soils

A score plot of the first and second principal components (PC1 and PC2) of the first-derivative spectra revealed that together they explained 77% of the variance in spectral data (Figure 2).

The most obvious spectral features of gypsum and carbonates are illustrated in Figure 3 using sample spectra high and low in gypsum and carbonate content. There were three dominant absorption features, near 1400 nm, 1900 nm and 2200 nm, in the soil sample without any carbonates or gypsum. These were caused by combination and overtone bands related to water and hydroxyls at 1400 nm and 1900 nm and Al–OH bonds in clay minerals at 2200 nm.³²

The vis-NIR spectrum of the soil sample with the highest amount of gypsum showed characteristic finger-shaped absorption features at 1400–1500 nm, as well as simple features at 1100 nm and 1750 nm (Figure 3).³³ The features in the NIR spectrum typical for gypsic soils are related to the di-hydrated gypsum mineral ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$). Absorption peaks were expected near 1750 nm and also at 1100 nm, 1400 nm and 1900 nm owing to the combination of O–H stretches, H–O–H bending and various overtones.³³ A highly pronounced absorption band near 2338 corresponding to the second overtone of the CO_3 antisymmetric stretching vibration characterised the soil sample with the highest amount of carbonates (Figure 3).³² For calculating carbonate abundance using the feature-specific method, 2338 nm was therefore selected. Viscarra Rossel *et al.*²⁵ also used 2338 nm as the spectral feature of calcite predictions. Despite the characteristic finger-shaped absorption features at 1400–1500 nm, the 1748 nm band was selected as the specific absorption feature to identify soils with high amounts of gypsum, because no other known features interfere in this region. The values of λ_{\min} and λ_{\max} were empirically determined and found to be $\lambda_{\min} = 2270$ nm and $\lambda_{\max} = 2375$ nm for carbonates and $\lambda_{\min} = 1680$ nm and $\lambda_{\max} = 1836$ nm for gypsum.

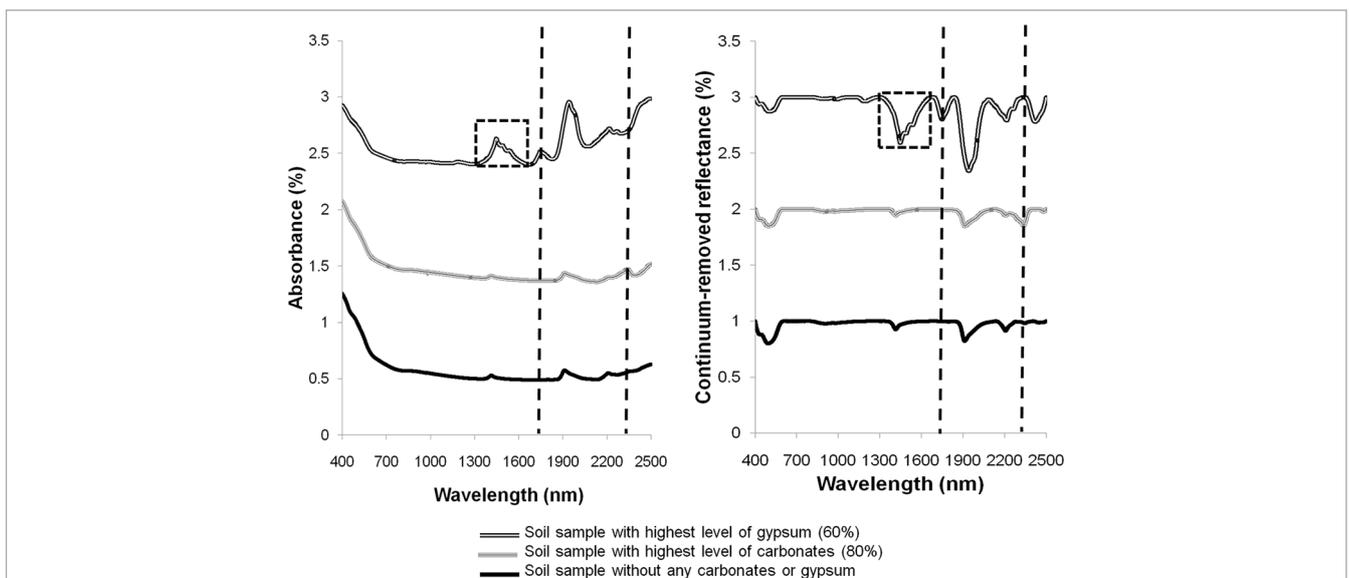


Figure 3. Finger-shaped absorption feature at 1400–1500 nm and absorption features at 1100 nm and 1750 nm, and a highly pronounced absorption peak near 2338 nm, which characterised the soil samples with the highest amount of gypsum and carbonates, respectively, in apparent absorbance and CRR vis-NIR spectra. The baseline is shifted between spectra for clarity. Double black lines, grey lines and simple black lines indicate the soil sample with the highest level of gypsum (60%), the soil sample with the highest level of carbonates (80%) and a soil sample without any carbonates or gypsum, respectively.

Table 2. PLSR model parameters for carbonates and gypsum estimation.

	Carbonates					Gypsum				
	n.f. ^a	E ^b	r ²	Bias	RMSE (%)	n.f. ^a	E ^b	r ²	Bias	RMSE (%)
PLSR										
Cross-validated calibration (n = 170)	7	0.58	0.58	0.29	11.7	2	0.73	0.73	-0.01	4.0
Validation (n = 81)	7	0.52	0.54	1.2	11.5	2	0.80	0.86	-0.2	3.4
Feature specific										
All samples (n = 251)		0.29	0.35	-3.4	14.9		0.56	0.71	-2.9	5.2
Validation samples (n = 81)		0.53	0.44	-2.3	13.6		0.60	0.85	-3.0	4.9

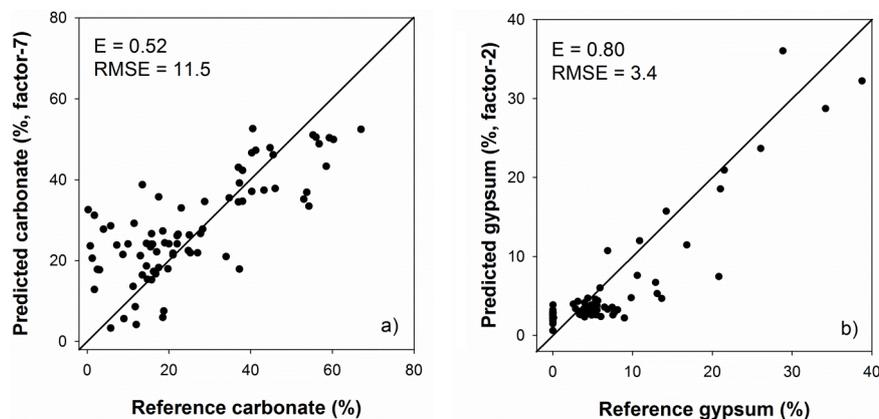
^aNumber of PLS factors^bModel efficiency

Figure 4. PLSR validation results showing predicted versus measured (a) carbonate and (b) gypsum content.

Prediction of carbonate and gypsum concentrations using PLSR and feature-specific estimates

Cross-validation statistics and validation results of the PLSR models for carbonates and gypsum in the soils from the study area are presented in Table 2 and Figure 4. The PLSR models were able to explain 52% and 80% of the variance in carbonate and gypsum content, respectively. As can be seen in Figure 4(b), there was a small gap between samples with a gypsum content close to zero and those with more gypsum. These low-gypsum samples originated from the subhumid region of the study area, where the standard acetone method was used for gypsum analysis.

In this study, 2338nm was used as the specific absorption feature of calcium carbonate based on the very sharp absorption features of soils with the highest amounts of calcium carbonate in the region (Figure 5). Despite these strong absorbance features in soils with carbonates, the feature-specific method did not prove satisfactory for predicting calcium carbonate. In addition to a lack of precision, the bias was quite substantial (Table 2). Figure 6 shows CRR spectra of soil samples with similar amounts of carbonates (34–41%) and high amounts of illite. A peak shift in the soil sample with the highest amounts of illite can be clearly distinguished. The

soil with the highest amounts of chlorite (black line) displayed the deepest absorption feature.

Figure 7 illustrates the vis-NIR CRR spectra for five soil samples with high amounts of gypsum. Results obtained

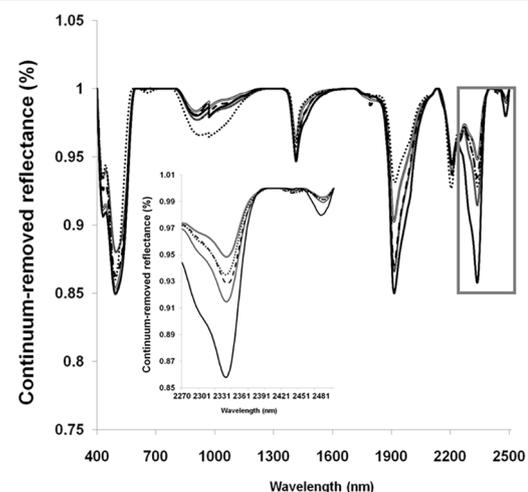


Figure 5. Diagnostic absorption feature of carbonates (2338 nm) in the CRR spectra for the five soils with the highest amounts of carbonates.

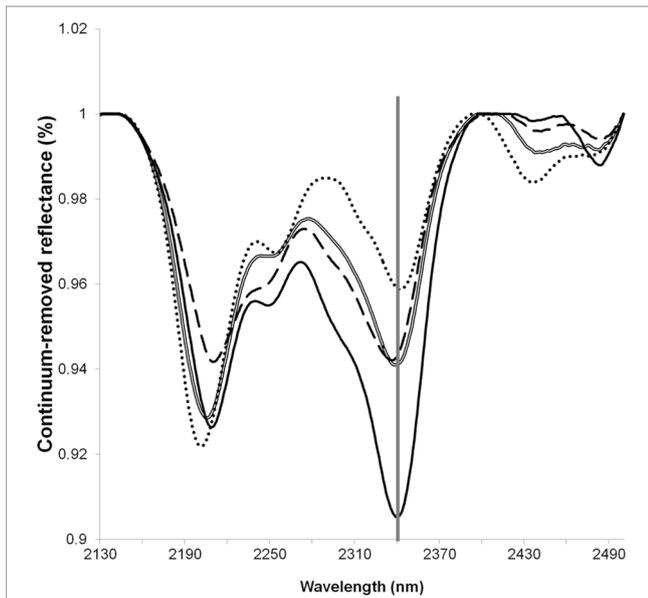


Figure 6. CRR spectra of soil samples with a similar amount of carbonates (34–41%) and high amounts of illite. The dotted and black lines represent the soils with the highest amount of illite and chlorite, respectively.

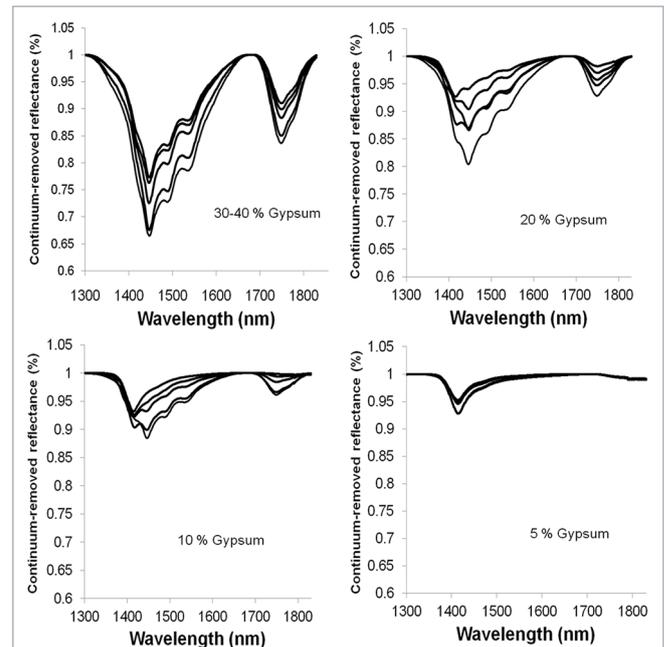


Figure 8. CRR spectra of soils with different percentages of gypsum.

using the feature-specific method to predict gypsum in all samples are presented in Table 2. The results show that PLSR predicted gypsum more accurately than the feature-specific method, but the difference was smaller than for carbonate (Table 2).

As can be seen in Figure 8, the typical finger-shaped absorption features and spectral feature at 1748 nm become apparent at about 10% gypsum according to the reference method. Owing to uncertainties in the laboratory analysis for samples with low-gypsum content, we tested the assumption that soils without the finger-shaped features did not contain any gypsum by assuming zero gypsum content in these soils.

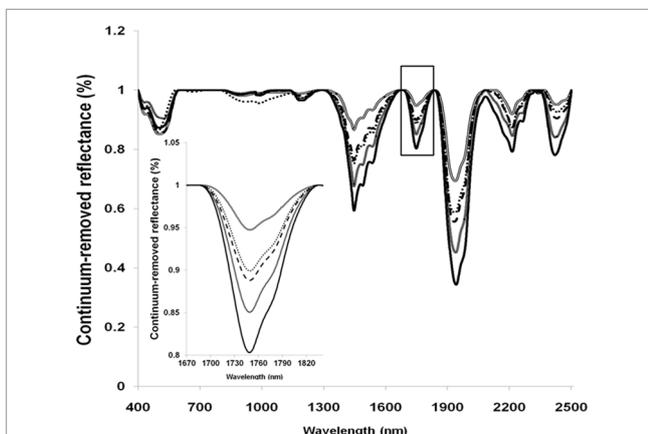


Figure 7. Diagnostic absorption feature of gypsum (1748 nm) in the CRR spectra of the five soils with the highest amounts of gypsum.

A new PLSR with full cross-validation on these data showed that E improved from 0.76 to 0.86, and $RMSE$ decreased from 3.8 to 3.0 (Figure 9).

Discussion

In the study area in arid and subhumid regions of Isfahan, Iran, there are different sources of variation, such as climate conditions, geology and elevation. Our PCA analyses showed that differences in gypsum content had a major influence on the vis-NIR spectra and constituted most of the variation in PC1 in the PCA analyses.

Other results showed that PLSR was able to predict gypsum in the study area accurately, whereas the predictions were not very accurate for carbonates. Volkan Bilgili *et al.*³⁴ and Summers *et al.*²⁶ reported R^2 values of 0.64 and 0.69 for calcium carbonate prediction in soils in Turkey and Australia, respectively, compared with 0.54 in the present study. The better results for carbonate predictions in their studies can probably be partly explained by differences in study area size and variability, variation in carbonate content and the experimental method used for carbonate measurement. Volkan Bilgili *et al.*³⁴ used a PLSR model to predict calcium carbonate in a 32-ha area in a semi-arid climate, whereas our study site was 110,000 km² with large climate and geological variations. Summers *et al.*²⁶ predicted calcium carbonate in 69 soil samples with 0–25.67% calcium carbonate by leave-one-out cross-validation, whereas the range of carbonates in our study was 0–80%. Moreover, the error of the calcimeter method used by Volkan Bilgili *et al.*³⁴ and Summers *et al.*²⁶ is less than

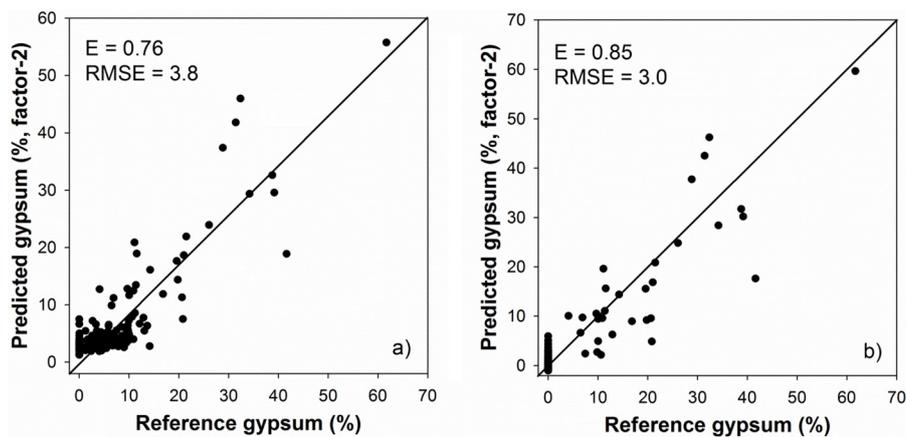


Figure 9. Full cross-validation in all soil samples with (a) laboratory gypsum data and (b) using the laboratory gypsum data only for soils with finger-shaped features and zero for other soils.

5%, whereas the error of the acid neutralisation or titration method used in the present study is more than 5%. It should be mentioned that the most suitable method is selected on the basis of precision (accuracy), equipment cost, operating and maintenance costs, staff requirements and sensitivity of the equipment.⁴⁷ The acid neutralisation or titration method was used in the present study because it needs minimal equipment and is very cheap.

In the feature-specific method, selecting unique absorption features for the soil parameters is crucial. This was sufficiently achieved for gypsum in the present study, where the strong absorption at 1748 nm was only present in soils containing gypsum. The very pronounced gypsum features apparent in the NIR spectra of gypsic soils from 1000 nm to 2500 nm are due to combinations of O–H stretches, H–O–H bending and overtones. The fundamental OH stretching and OH deformation absorption bands are very sharp distinct bands in the mid infrared region at 2809 nm, 2924 nm, 5935 nm and 6173 nm.⁴⁸ Liu *et al.*⁴⁹ have attributed the absorption band at 1748 nm to combination modes of $[\text{SO}_4]^{2-}$ and H_2O vibrations, which makes this band unique compared with most other minerals commonly occurring in soil.

For carbonates, a highly pronounced absorption peak near 2338 nm characterised the soil sample with the highest amount of carbonates. White⁵⁰ reported that the carbonate ion displays strong fundamental vibration bands: a stretching vibration near 7000 nm and two bending modes near 11,400 nm and 14,100 nm. According to Clark *et al.*,³² combination and overtone bands of these CO_3 fundamentals occur in the near infrared region, with the two strongest appearing near 2300–2350 nm and 2500–2550 nm. The bands have slightly different positions depending on the type of carbonate-containing mineral dominating in the soil.⁵¹ However, the absorbance feature of carbonate at 2338 nm is also similar to the absorption features of illite and chlorite. Therefore, the absorbance at 2338 nm is not solely attributable to carbonates, since the presence of different amounts of illite and chlorite affects

the peak height and slightly changes the peak position. This effect was observed in the present study (Figure 6) and probably contributed to the difficulties in using the feature-specific method for predictions of carbonate in the study area with its large differences in clay mineralogy.

PLSR obtains information from all parts of the spectrum to predict soil properties, so it may not be as sensitive to interfering absorbance features by other minerals as the feature-specific method. Comparing results obtained using PLSR and the feature-specific method demonstrated that PLSR was more capable of prediction for the soils of the study area. In a previous study, Gomez *et al.*⁴⁰ compared the ability of PLSR and CR to predict calcium carbonate in the soils of Montpellier in France, using the 2341 nm band as the specific feature for calcium carbonate. The results showed that the PLSR technique was only slightly more accurate than the feature-specific method for CaCO_3 content prediction compared with laboratory measurements. They also showed that the most important spectral features used by the PLSR, other than the feature at 2341 nm, were centred near 500 nm, 600 nm, 1430 nm, 1900 nm, 2230 nm and 2373 nm. Since these additional spectral features only slightly increase the prediction performance, selecting the PLSR approach in this situation can be questionable. This is particularly true because most of the added spectral features cannot be well explained, which hampers the use of the PLSR relations on a wider area.

Another difference of potential relevance between PLSR and the feature-specific method is that the latter relies on the representativeness of the chosen reference sample. The obvious bias of the feature-specific predictions obtained in the present study indicates issues related to this (Table 2). In the case of carbonate, there are several different types of carbonate within the study area, which undermines the possibility of finding samples representative for all. For gypsum, the bias is largely explained by the large number of samples with gypsum present according to the reference method, but without the specific gypsum features in the spectra. Other

forms of calcium sulfate can also be found in these types of soil with different degrees of hydration affecting the OH vibration spectral features.⁵²

It could be argued that pure minerals should be used as reference samples. This would most likely have given different results, but the problem with representativeness would persist. It can also be expected that spectral features will appear different when associated with the soil matrix. Lagacherie *et al.*⁵³ used the relationship between CaCO₃ and the CR₂₃₄₁, without a pure mineral as reference, when estimating calcium carbonate in an 80 km² catchment in southern France using laboratory, field and airborne hyperspectral measurements.

Although the calibration models obtained for gypsum in the present study were good, there is still room for improvement. The results are based on laboratory measurements of gypsum with associated uncertainties. Still, as was demonstrated with the improved predictions using the finger-shaped features as an indication of presence or not of gypsum in the soil, the results indicated that vis-NIR might be a more precise method for analysing soil gypsum content.

Conclusions

This study showed that both PLSR and a feature-specific method based on the continuum-removed spectra were able to predict gypsum content in soils with reasonable accuracy. Owing to the uniquely finger-shaped and highly pronounced absorption features of gypsum, with no overlap with other soil properties, prediction of gypsum content was more accurate than prediction of carbonate content. In contrast, the absorption feature used for carbonates was found to be too close to illite and chlorite absorption features, and therefore with the range of different clay minerals present in the soils studied, the prediction of carbonates by vis-NIR spectroscopy lacked precision. PLSR performed better than the feature-specific method based on the CR spectra for carbonates and gypsum. Despite uncertainties in the laboratory analyses of gypsum in this study, vis-NIR spectroscopy predicted gypsum in soils with good precision. However, more studies are needed on different methods of gypsum measurement in the laboratory in order to fully understand the ability of vis-NIR spectroscopy in gypsum content prediction for soils in arid and semi-arid regions.

Acknowledgements

All the vis-NIR spectroscopy analyses were performed during six months of sabbatical leave by the first author at the Swedish University of Agricultural Sciences. All the help and assistance of the faculty and staff at Swedish University of Agricultural Sciences is greatly appreciated. The Isfahan University of Technology and the Swedish Research Council Formas [Dnr 229-2010-951] is acknowledged for financial support.

References

1. H.N. Le Houérou, "Climate change, drought and desertification", *J. Arid Environ.* **34**, 133 (1996). doi: <http://dx.doi.org/10.1006/jare.1996.0099>
2. S. Kefi, M. Rietkerk and G.G. Katul, "Vegetation pattern shift as a result of rising atmospheric CO₂ in arid ecosystems", *Theor. Popul. Biol.* **74**, 332 (2008). doi: <http://dx.doi.org/10.1016/j.tpb.2008.09.004>
3. K.J. Naithani, B.E. Ewers and E. Pendall, "Sap flux-scaled transpiration and stomatal conductance response to soil and atmospheric drought in a semi-arid sagebrush ecosystem", *J. Hydrol.* **464–465**, 176 (2012). doi: <http://dx.doi.org/10.1007/s13592-013-0221-x>
4. H.C. Monger, "Arid soils", in *Encyclopedia of Soil Science*, vol. 1, 2nd edition, Ed by R. Lal. Taylor & Francis/CRC Press, Boca Raton, FL (2006).
5. H.E. Doner and W.C. Lynn, "Carbonate, halide, sulfate and sulfide minerals," in *Minerals in Soil Environments*, 2nd edition, Ed by J.B. Dixon and S.B. Weed. Soil Science Society of America, Madison, WI, p. 75 (1989).
6. F. Visconti, J.M. De Paz and J.L. Rubio, "Calcite and gypsum solubility products in water-saturated salt-affected soil samples at 25°C and at least up to 14 Ds M⁻¹", *Eur. J. Soil Sci.* **61**, 255 (2010). doi: <http://dx.doi.org/10.1111/j.1365-2389.2009.01214.x>
7. J. Herrero, J. Porta and N. Fedoroff, "Hypergypsic soil micromorphology and landscape relationships in northern Spain", *Soil Sci. Soc. Am. J.* **56**, 1188 (1992).
8. A. Watson, "Desert gypsum crusts as palaeoenvironmental indicators: a micropetrographic study of crusts from southern Tunisia and the Central Namib Desert", *J. Arid Environ.* **15**, 19 (1988).
9. FAO, "Management of gypsiferous soils," in *Soils Bulletin*, vol. 62. Food and Agriculture Organization of the United Nations, Rome (1990).
10. R. Smith and V.C. Robertson, "Soil and irrigation classification of shallow soils overlying gypsum beds, northern Iraq", *J. Soil Sci.* **13**, 106 (1962). doi: <http://dx.doi.org/10.1111/j.1365-2389.1962.tb00687.x>
11. J. Porta, "Methodologies for the analysis and characterization of gypsum in soils: a review", *Geoderma* **87**, 31 (1998). doi: [http://dx.doi.org/10.1016/S0016-7061\(98\)00067-6](http://dx.doi.org/10.1016/S0016-7061(98)00067-6)
12. J. Loveday, Ed. "Methods for analysis of irrigated soils", Technical Communication No. 54. Commonwealth Agricultural Bureaux, Farnham Royal, Buckinghamshire, UK (1974).
13. O. Artieda, *Factores Geológicos Que Inciden En El Desarrollo De Los Suelos En Un Medio Semiárido*. MS thesis, Universidad de Zaragoza (1993).
14. C.A. Bower and R.B. Huss, "Rapid conductometric method for estimating gypsum in soils", *Soil Sci.* **66**, 199 (1948). doi: <http://dx.doi.org/10.1097/00010694-194809000-00006>

15. L.A. Richards, Ed. *Diagnosis and Improvement of Saline and Alkali Soils*. United States Salinity Laboratory Staff, Washington, DC (1954).
16. L.E. Allison and C.D. Moodie, "Carbonate," in *Methods of Soil Analysis*, vol. 9, 2nd edition, Ed by C.A. Black. Agronomy Monograph, ASA, CSSA, and SSSA, Madison, WI (1965).
17. M.C. Rabenhorst, "Determination of organic and carbonate carbon in calcareous soils using dry combustion", *Soil Sci. Soc. Am. J.* **52**, 965 (1988). doi: <http://dx.doi.org/10.2136/sssaj1988.03615995005200040012x>
18. L.G. Bundy and J.M. Bremner, "Simple titrimetric method for determination of inorganic carbon in soils", *Soil Sci. Soc. Am. Proc.* **36**, 273 (1972). doi: <http://dx.doi.org/10.2136/sssaj1972.03615995003600020021x>
19. A.E. Martin and R. Reeve, "A rapid manometric method for determining soil carbonate", *Soil Sci.* **79**, 187 (1955). doi: <http://dx.doi.org/10.1097/00010694-195503000-00003>
20. A. Dreimanis, "Quantitative gasometric determination of calcite and dolomite by using Chittick apparatus", *J. Sediment. Petrol.* **32**, 520 (1962). doi: <http://dx.doi.org/10.1306/74d70d08-2b21-11d7-8648000102c1865d>
21. R.H. Loeppert and D.L. Suarez, "Carbonate and gypsum," in *Methods of Soil Analysis* vol. 5, D.L. Sparks, A.L. Page, P.A. Helmke, R.H. Loeppert, P.N. Soltanpour, M.A. Tabatabai, C.T. Johnson and M.E. Sumner, Eds. SSSA Madison, WI (1996).
22. C. Guerrero, R.A.V. Rossel and A.M. Mouazen, "Special issue 'Diffuse Reflectance Spectroscopy in Soil Science and Land Resource Assessment'", *Geoderma* **158**, 1 (2010). doi: <http://dx.doi.org/10.1016/j.geoderma.2010.05.008>
23. B. Stenberg, R.A. Viscarra Rossel, A.M. Mouazen and J. Wetterlind, "Visible and near infrared spectroscopy in soil science", in *Advances in Agronomy* **107**, 163, Ed by L.S. Donald. Academic Press, London (2010). doi: [http://dx.doi.org/10.1016/S0065-2113\(10\)07005-7](http://dx.doi.org/10.1016/S0065-2113(10)07005-7)
24. B. Stenberg, E. Nordkvist and L. Salomonsson, "Use of near infrared reflectance spectra of soils for objective selection of samples", *Soil Sci.* **159**, 109 (1995). doi: <http://dx.doi.org/10.1097/00010694-199502000-00005>
25. R.A. Viscarra Rossel, S.R. Cattle, A. Ortega and Y. Fouad, "In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy", *Geoderma* **150**, 253 (2009). doi: <http://dx.doi.org/10.1016/j.geoderma.2009.01.025>
26. D. Summers, M. Lewis, B. Ostendorf and D. Chittleborough, "Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties", *Ecol. Indic.* **11**, 123 (2011). doi: <http://dx.doi.org/10.1016/j.ecolind.2009.05.001>
27. R. Zornoza, C. Guerrero, J. Mataix-Solera, K.M. Scow, V. Arcenegui and J. Mataix-Beneyto, "Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils", *Soil Biol. Biochem.* **40**, 1923 (2008). doi: <http://dx.doi.org/10.1016/j.soilbio.2008.04.003>
28. Y.-W. Dong, S.-Q. Yang, C.-Y. Xu, Y.-Z. Li, W. Bai, Z.-N. Fan, Y.-N. Wang and Q.-Z. Li, "Determination of soil parameters in apple-growing regions by near- and mid-infrared spectroscopy", *Pedosphere* **21**, 591 (2011). doi: [http://dx.doi.org/10.1016/S1002-0160\(11\)60161-6](http://dx.doi.org/10.1016/S1002-0160(11)60161-6)
29. L. Xuemei and L. Jianshe, "Measurement of soil properties using visible and short wave-near infrared spectroscopy and multivariate calibration", *Measurement* **46**, 3808 (2013). doi: <http://dx.doi.org/10.1016/j.measurement.2013.07.007>
30. E. Ben-Dor and A. Banin, "Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties", *Soil Sci. Soc. Am. J.* **59**, 364 (1995). doi: <http://dx.doi.org/10.2136/sssaj1995.03615995005900020014x>
31. V. Bellon-Maurel and A. McBratney, "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**, 1398 (2011). doi: <http://dx.doi.org/10.1016/j.soilbio.2011.02.019>
32. R.N. Clark, T.V.V. King, M. Klejwa, G.A. Swayze and N. Vergo, "High spectral resolution reflectance spectroscopy of minerals", *J. Geophys. Res.* **95**, 12 (1990). doi: <http://dx.doi.org/10.1029/JB095iB08p12653>
33. G.R. Hunt, J.W. Salisbury and C.J. Lenhoff, "Visible and near infrared spectra of minerals and rocks: IV. Sulphides and sulphates", *Mod. Geol.* **3**, 1 (1971).
34. A. Volkan Bilgili, H.M. van Es, F. Akbas, A. Durak and W.D. Hively, "Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey", *J. Arid Environ.* **74**, 229 (2010). doi: <http://dx.doi.org/10.1016/j.jaridenv.2009.08.011>
35. E. Zelikman and E. Carmina, "The spectral response characteristics of the soils and their possible estimation by using partial least square regression (PLSR) analysis", *Int. J. Geomat. Geosci.* **3**, 438 (2013).
36. J.P. Gras, B.G. Barthès, B. Mahaut and S. Trupin, "Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils", *Geoderma* **214-215**, 126 (2014). doi: <http://dx.doi.org/10.1016/j.geoderma.2013.09.021>
37. G.M. Vasques, S. Grunwald and J.O. Sickman, "Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra", *Geoderma* **146**, 14 (2008). doi: <http://dx.doi.org/10.1016/j.geoderma.2008.04.007>
38. R.A. Viscarra Rossel, "Fine-resolution multi-scale mapping of clay minerals in Australian soils measured with near infrared spectra", *J. Geophys. Res.: Earth Surf.* **116**, F04023 (2011). doi: <http://dx.doi.org/10.1029/2011JF001977>
39. R.N. Clark and T.L. Roush, "Reflectance spectroscopy: quantitative analysis techniques for remote sensing

- applications”, *J. Geophys. Res.* **89**, 6329 (1984). doi: <http://dx.doi.org/10.1029/JB089iB07p06329>
40. C. Gomez, P. Lagacherie and G. Coulouma, “Continuum removal versus pls method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements”, *Geoderma* **148**, 141 (2008). doi: <http://dx.doi.org/10.1016/j.geoderma.2008.09.016>
41. J.S. Madeira Netto, J.M. Robbez-Masson and E. Martins, *Visible-NIR Hyperspectral Imagery for Discriminating Soil Types in the La Peyne Watershed (France)*, vol. 31 (2006).
42. A. Sorbi, *Geological Map of Isfahan Province, Scale (1/1000000)*. Geological Survey of Iran, Iran (2002).
43. R.E. Nelson, “Carbonate and Gypsum,” in *Methods of Soil Analysis*, 2nd edition, vol. 9, Ed by A.L.e.a. Page. SSSA, Madison, WI (1982).
44. A. Savitzky and M. Golay, “Smoothing and differentiation of data by simplified least squares procedures”, *Anal. Chem.* **36**, 1627 (1964). doi: <http://dx.doi.org/10.1021/ac60214a047>
45. H. Martens and T. Naes, *Multivariate Calibration*. John Wiley & Sons, Chichester, UK (1989).
46. K.S. Schmidt and A.K. Skidmore, “Spectral discrimination of vegetation types in a coastal wetland”, *Remote Sens. Environ.* **85**, 92 (2003). doi: [http://dx.doi.org/10.1016/S0034-4257\(02\)00196-7](http://dx.doi.org/10.1016/S0034-4257(02)00196-7)
47. F. Lamas, C. Irigaray, C. Oteo and J. Chacón, “Selection of the most appropriate method to determine the carbonate content for engineering purposes with particular regard to marls”, *Eng. Geol.* **81**, 32 (2005). doi: <http://dx.doi.org/10.1016/j.enggeo.2005.07.005>
48. V.C. Farmer (Ed), *The Infrared Spectra of Minerals*, Mineralogical Society, London (1974)
49. Y. Liu, A. Wang, and J.J. Freeman, “Raman, MIR and NIR spectroscopic study of calcium sulfates: gypsum, bassanite and anhydrite”, in *Proceedings of the 40th Lunar and Planetary Science Conference*, 23–27 March 2009, The Woodlands, TX. <http://www.lpi.usra.edu/meetings/lpsc2009/pdf/2128.pdf> (2009)
50. W.B. White, “The carbonate minerals”, in *The Infrared Spectra of Minerals*, Ed by V.C. Farmer, Mineralogical Society, London (1974)
51. G.R. Hunt and J.W. Salisbury, “Visible and near-infrared spectra of minerals and rocks: II. Carbonates”, *Mod. Geol.* **2**, 23 (1971).
52. E.G. Akpokdje, “The occurrence of bassanite in some Australian arid-zone soils” *Chem. Geol.* **47**, 361 (1984)
53. P. Lagacherie, F. Baret, J.-B. Feret, J. Madeira Netto and J.M. Robbez-Masson, “Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements”, *Remote Sens. Environ.* **112**, 825 (2008). doi: <http://dx.doi.org/10.1016/j.rse.2007.06.014>

