Near-infrared spectroscopy: alternative method for assessment of stable carbon isotopes in various soil profiles in Chile

María de los Ángeles Sepúlveda, Marcela Hidalgo, Juan Araya, Manuel Casanova, Cristina Muñoz, Sebastian Doetterl, Daniel Wasner, Ben Colpaert, Samuel Bodé, Pascal Boeckx, Erick Zagal

Near-infrared spectroscopy: Alternative method for assessment of stable carbon isotopes in various soil profiles in Chile

María de los Ángeles Sepúlveda [a], Marcela Hidalgo [a], Juan Araya [b], Manuel Casanova [c], Cristina Muñoz [a], Sebastian Doetterl [d], Daniel Wasner [d], Ben Colpaert [e], Samuel Bodé [e], Pascal Boeckx [e], Erick Zagal [a,*]

[a] Department of Soil and Natural Resources, Faculty of Agronomy, University of Concepción, Chile
[b] Department of Instrumental Analysis, Faculty of Pharmacy, University of Concepción, Chile
[c] Department of Engineering and Soil, Faculty of Agronomics Science, University of Chile, Santiago, Chile
[d] ETH-Zürich, Department of Environmental Systems Science, Zürich, Switzerland
[e] Department of Green Chemistry and Technology, Isotope Bioscience Laboratory (ISOFYS), Ghent University, Gent, Belgium

ARTICLE INFO

ABSTRACT

The role of soil in the global carbon cycle and carbon–climate feedback mechanisms has attracted considerable interest in recent decades. Consequently, development of simple, rapid, and inexpensive methods to support the studies on carbon dynamics in soil is of interest. Near-infrared spectroscopy (NIRS) has emerged as a rapid and cost-effective method for measurements of soil properties. The aim of this study was to develop and validate a predictive model for δ^{13}C values using NIRS in various soil profiles across Chile. Eleven sites were selected in the range of 30° to 50° S. These sites represent different soil moisture and soil temperature regimes, clay mineralogies, parent materials, and climates; in addition, they have prairie vegetation and contain C3-type vegetation. Air-dried soil samples were scanned in the NIR range at a resolution of 4 cm^{-1}. The carbon isotopic composition, expressed as δ^{13}C relative to the Vienna Pee Dee Belemnite standard, was analysed using an elemental analyser–isotope ratio mass spectrometer system. A prediction model for δ^{13}C values based on NIRS data was developed through a partial least-squares regression (PLS) model using ten latent variables. A second model was generated using a random forest (RF) approach. The model performances were acceptable. The RF model provided the best results. The values of the root mean square error of prediction for the validation runs for δ^{13}C obtained using the PLS and RF models were 1.38‰, and 1.15‰, respectively. These model performances indicate that NIRS can be used to predict δ^{13}C for the selected dataset. The results of this study support the use of NIRS as a predictive method in soil analyses and as a nondestructive waste-free method for studies on carbon dynamics in soil.

1. Introduction

Soil organic carbon (SOC) is one of the largest reserves of carbon in terrestrial ecosystems (Lal, 2006). However, studies indicate that the exchange of SOC with the atmosphere can vary depending on climatic conditions, which leads to questions whether the soil is a source or sink for atmospheric carbon dioxide (CO_2) (Carvalhais et al., 2014). Therefore, the understanding of SOC dynamics, particularly the carbon stabilisation, is crucial to predict the role of SOC in the carbon cycle under a changing climate (Crowther et al., 2016). The potential of soil to sequester carbon depends primarily on the soil development and interactions between weathering and biological processes that affect the nutrient availability (Doetterl et al., 2018).

Climatic factors control the SOC degradation (Carvalhais et al., 2014), while geochemical factors stabilise soil organic matter (SOM). The interaction between climatic and geochemical factors in soil carbon storage has attracted attention (Doetterl et al., 2015). Furthermore, the climate, vegetation, and geochemical soil composition affect the SOC dynamics (Finke et al., 2019). The influence of the parent material on the soil geochemistry reveals the importance of the paedological characteristics and soil types as factors of stabilising influences on the SOC (Finke et al., 2019).

An intensive and reliable mapping is required to monitor changes in soil organic pools (Bricklemyer et al., 2005; Mooney et al., 2004), as well as small scale studies. The SOM composition and carbon dynamics have been studied in topsoil horizons for a long time. Over the past decade, subsoil horizons have been actively investigated because the subsoil carbon contributes to the total carbon stocks. The understanding of the factors that stabilise carbon in deeper soil layers is still limited, which is essential to understand the feedback mechanisms between SOC stocks and atmospheric CO_2 during climate changes (Chabbi et al., 2009; Rumpel et al., 2002; Fontaine et al., 2007). Accurate and low-cost methods of soil analysis are required because the number of soil

* Corresponding author at: Vicente Méndez 595, Chillán, Chile.
  E-mail address: ezagal@udec.cl (E. Zagal).

samples typically involved in such studies is large. During the past two decades, visible–near-infrared (vis–NIR) diffuse reflectance spectroscopy has been developed as an easy-to-use method, suitable for prediction of several soil properties (e.g., % C, % N, pH, and texture) (Viscarra Rossel et al., 2006; Petisco et al., 2006; Zornoza et al., 2008).

Numerous methods can be used to study the dynamics of SOC. Techniques that can measure the isotopic abundance of ^{13}C in SOC are useful to elucidate C-process dynamics (Balesdent et al., 1993; Balesdent and Mariotti, 1996; Glaser, 2005; Trumbore, 2009; Accoe et al., 2003). Accoe et al. (2003) proposed that the change in ^{13}C content in soil profiles can be used as an indicator of the stability of SOM. The isotopic ratios of several elements (e.g., carbon, nitrogen, and oxygen) in the soil are typically determined using isotope ratio mass spectrometry (IRMS) (Balesdent and Mariotti, 1996; Glaser, 2005; Trumbore, 2009). However, IRMS is complex and requires both sophisticated equipment and trained personnel, which limits the number of samples that can be analysed.

Near-infrared reflectance spectroscopy (NIRS) has numerous characteristics of interest for agronomic and environmental studies. The sample preparation involves only drying and grinding. The analysis is nondestructive, without hazardous chemicals. In addition, the measurement period is only a few seconds. Furthermore, NIRS is suitable for analyses of large samples. Multiple soil properties can be estimated by a single scan. In recent years, NIRS has also been used to determine the abundance of stable carbon isotopes in soil. NIRS has been used to predict δ^{13}C in soil (Fuentes et al., 2012; Winowiecki et al., 2017). These studies indicate that infrared spectroscopy is a promising method to estimate δ^{13}C in soil, which can enable an increased number of analysed samples, often required for studies on carbon dynamics in soil (Accoe et al., 2003).

NIRS generates complex absorption patterns, which need to be mathematically processed to correlate latent variables with soil properties (Stenberg et al., 2010). Such analyses of soil spectra require multi-variate calibrations (Martens and Naes, 1989) to capture the information relevant to the calibration and validation of predictive models. Recently, the number of studies on the multi-variate analysis of NIRS data has largely increased. Good results have been obtained for multiple soil properties (Theo, 2005; Viscarra Rossel et al., 2006; Viscarra Rossel and McBratney, 1998; Zornoza et al., 2008). NIRS is a more accessible method for the analysis and a more suitable alternative to conventional chemical methods of soil analysis (Fuentes et al., 2012).

We report the development and validation of a predictive model to estimate δ^{13}C in topsoil and subsoils, which can improve our knowledge of the mechanisms of SOC stocks during climate changes. In contrast to those in other studies, this model attributed the changes in the signals to various soil types and profiles under different climatic conditions. The implementation of this alternative methodology using NIRS to assess stable isotopes of carbon in soil is presented as a viable and low-cost technique for studies on soil carbon dynamics in wide transects and at different depths, which are factors that increase the complexity and cost of the analysis.

2. Materials and methods

2.1. Soil sampling

From the Chilean Coquimbo Region to the Magallanes Region (30°–50° S), a total of eleven sites were carefully chosen to have a representative soil transect (see Doetterl et al. (2015) for the used selection criteria). The criteria included soils with a broad pH range (4.6–7.5) and null HCl reaction, soils having various soil moisture regimes (SMRs) (aridic, ustic, xeric, udic, and perudic), soil temperature regimes (STRs) (thermic, isothermic, mesic, isomesic, and cryic), and clay mineralogies (short range order mineral phases and crystalline), soils with different parent materials (volcanic ash, alluvial, fluvio-glacial material, marine sediment, etc.), soils located within different climatic zones (arid–semiarid, Mediterranean arid, Mediterranean humid, humid, and Magallanian), and soils under natural prairie vegetation conditions (C3-type vegetation). Field sampling campaigns were carried out primarily during summer (2017–2018). Sampling units were defined in plots of 50 × 50 m, from which six random soil cores were extracted. These campaigns were an extension of the transect reported by Doetterl et al. (2015) and included additional depths of soil layers. We chose these sites to maximise the climatic and physicochemical diversities of the soils to train models with the widest range of application. Site locations and basic physicochemical variables are listed in Tables 1, 2 and 3. (See Fig. 1.)

2.2. Sample preparation

Soil samples were collected in triplicate to a depth of 60 cm or 30 cm (until gravel material was encountered) using polyvinylchloride (PVC) tubes (height: 35 cm; diameter: 90 mm) to extract undisturbed soil samples. The samples were then transported to the University of Concepción, where they were stored at −20°C until further processing. Soil sample profiles were obtained at intervals of 2 cm to a depth of 10cm, at intervals of 5 cm in the depth range of 10–30 cm, and at intervals of 10cm in the depth range of 30–60 cm. To obtain detailed soil layer data, the tubes were cut using a custom-designed device, which actioned a steel saw unit at a high speed. The procedures were carried out very carefully to avoid contamination of the soil with PVC powder or other materials. The samples were air-dried and sieved at 2 mm. The fine roots were removed using electrostatic energy, as described by Kuzyakov et al. (2001).

Table 1
Description of the 11 sites used for the calibration and validation of the predictive model.

Soil series	WGS1984		STR	SMR	Soil suborder	Geomorphology
	X	Y				
	(ddd,ddd)	(ddd,ddd)			(Soil taxonomy)	
Calle Larga	−70.52162	−32.87609	Thermic	Xeric	Typic Argixeroll	Piedmont
Pimpinela	−70.72972	−34.32387	Thermic	Xeric	Mollic Haploxeralf	Piedmont
Bramaderos	−71.31464	−35.61330	Thermic	Xeric	Humic Haploxerand	High alluvial terraces
Santa Bárbara	−71.69721	−36.45816	Thermic	Xeric	Typic Haploxerand	Old fluvio-glacial terraces
Choshuenco	−72.11120	−39.85941	Isomesic	Udic	Andic Dystrudept	Hillocks and hills
Mayamó	−73.79915	−42.05300	Isomesic	Perudic	Acrudoxic Durudand	Gentle rolling hills
Aituí	−73.61712	−43.05791	Isomesic	Perudic	Hydric Fulvudand	High planes
Puerto Cisnes	−72.61337	−45.38105	Isomesic	Perudic	Acrudoxic Fulvudand	Fluvio-glacial terrace
Bahía Exploradores	−73.06868	−46.50487	Mesic	Udic	Oxyaquic Hapludand	Hills
Aguas Frescas	−70.98.860	−53.43267	Cryic	Udic	Inceptisol	Marine terraces
Santa Olga	−70.36106	−53.31478	Isomesic	Perudic	Inceptisol	Marine terraces

X and Y: coordinates; STR: soil thermic regime; SMR: soil moisture regime.

Table 2
Chemical and physical properties of the 11 sites used for the model construction.

Soil series	Nitrogen (g N · kg⁻¹ soil)			Carbon (g C · kg⁻¹ soil)			Bulk density (gr · cm⁻³)			pH (KCl)			TRB (cmol$_c$ · kg⁻¹)			Clay %		
Depth (cm)	0–10	10–30	30–60	0–10	10–30	30–60	0–10	10–30	30–60	0–10	10–30	30–60	0–10	10–30	30–60	0–10	10–30	30–60
Calle Larga	5,8	2,3	ns	52,2	20,9	ns	1,4	1,8	ns	5,3	5,3	ns	23,8	19,9	ns	32,3	32,7	ns
Pimpinela	2,0	1,6	ns	20,1	14,6	ns	0,9	1,1	ns	5,5	5,5	ns	13,4	17,0	ns	36,0	37,1	ns
Bramaderos	4,7	5,1	4,4	54,5	59,8	55,4	1,4	1,5	0,9	5,1	5,1	5,1	10,5	10,1	8,8	20,7	26,1	19,9
Santa Bárbara	5,1	3,3	2,1	62,4	40,6	23,8	0,8	0,7	0,6	4,9	5,3	5,8	2,1	5,4	5,8	29,5	37,8	35,5
Choshuenco	9,3	4,3	3,9	108,0	43,3	47,1	0,7	0,5	0,8	4,5	4,8	5,1	3,5	1,9	3,7	15,6	7,2	10,7
Mayamó	11,0	6,4	4,6	138,6	91,9	59,4	0,7	0,6	0,8	4,4	4,3	5,1	4,6	1,9	0,7	8,4	15,6	23,9
Aituí	13,7	7,5	5,4	171,9	105,2	71,1	0,6	0,4	0,5	4,6	4,7	5,2	5,3	1,9	1,0	16,8	19,2	31,6
Puerto Cisnes	16,1	10,4	6,6	163,9	126,5	82,8	0,5	0,6	0,3	4,2	4,7	4,6	2,7	1,1	2,1	17,5	10,4	21,8
Bahía Exploradores	4,9	1,0	1,0	60,7	15,8	14,5	1,7	1,7	1,0	4,0	4,7	4,4	2,1	0,3	0,5	24,3	25,5	22,0
Aguas Frescas	4,5	0,5	ns	67,7	13,3	ns	0,7	0,7	ns	4,8	4,3	ns	27,3	1,8	ns	7,6	9,9	ns
Santa Olga	7,7	4,8	ns	150,2	91,9	ns	0,6	0,8	ns	4,0	3,9	ns	12,9	9,5	ns	13,0	23,1	ns

ns: not sampled owing to excess of gravel.

Table 3
Descriptive statistics for chemical and physical properties of the 11 sites used for the model construction.

| Soil property	n[a]	Min	Max	Median
Nitrogen (g N · kg⁻¹ soil)	29	0,5	16,1	5
Carbon (g C · kg⁻¹ soil)	332	6,2	195,7	59,8
Bulk density (gr · cm⁻³)	29	0,3	1,8	0,7
pH (KCl)	29	3,9	5,8	4,8
TRB (cmol$_c$ · kg⁻¹)	29	0,3	27,3	3,7
Clay %	29	7,2	37,8	21,8

[a] For C and soil depth intervals 0–10 cm, 10–30 cm, and 30–60 cm (the last one, when available) all samples were used (n = 332). For the other physicochemical variables and soil depths, three composite samples per site was used (n = 29).

2.3. Laboratory analysis

The samples were air-dried and scanned at NIR wavelengths (800–2857 nm) by diffuse reflectance spectroscopy. The resolution was 4 cm⁻¹. The Fourier-transform NIR system was a Bruker Matrix-I (Bruker Optics, Rheinstetten, Germany) located at the Soil and Environmental Laboratory of the Department of Soil and Natural Resources, Faculty of Agronomy, Universidad de Concepción. The total carbon contents (% C), total nitrogen contents (% N), and isotopic ratios (^{13}C/^{12}C) of 332 samples were determined using an elemental analysis (EA)–IRMS (ANCA-SL, Sercon, Crewe, UK) system, coupled to a 2022 IRMS system (Sercon, Crewe, UK), at the Isotope Bioscience Laboratory (ISOFYS, www.isofys.be) of Ghent University, Belgium. The carbon isotopic ratio (^{13}C/^{12}C) of the soil sample is expressed relative to an international reference, using the delta notation (δ^{13}C). The delta value expresses the fractional difference in the isotopic ratio between the sample and international reference. For ^{13}C, the used international reference standard was Vienna Pee Dee Belemnite (VPDB), typically expressed in parts per thousand (‰) (Chen et al., 2005).

For soil chemical and physical analyses (three composite samples per site, at 0–10cm, 10–30 cm, and, where applicable, 30–60 cm), the bulk density, texture, pH, and total reserve in base cations (TRB) were determined (n = 29). The bulk density was determined using the cylinder method. The inner cylinder containing an undisturbed soil core was removed and trimmed to the end with a knife to yield a core whose volume could be easily calculated using its length and diameter. The weight of this soil core was then determined after drying in an oven at 105 °C for approximately 20h (Sandoval et al., 2012). The soil texture was measured using the hydrometer method proposed by Bouyoucos (1962), samples having an organic C contents >5% were pretreated with 10% H_2O_2. The soil pH was determined potentiometrically in 25 mL of KCl 1 M (soil: solution ratio: 1:2.5) with a glass electrode using an HI2550
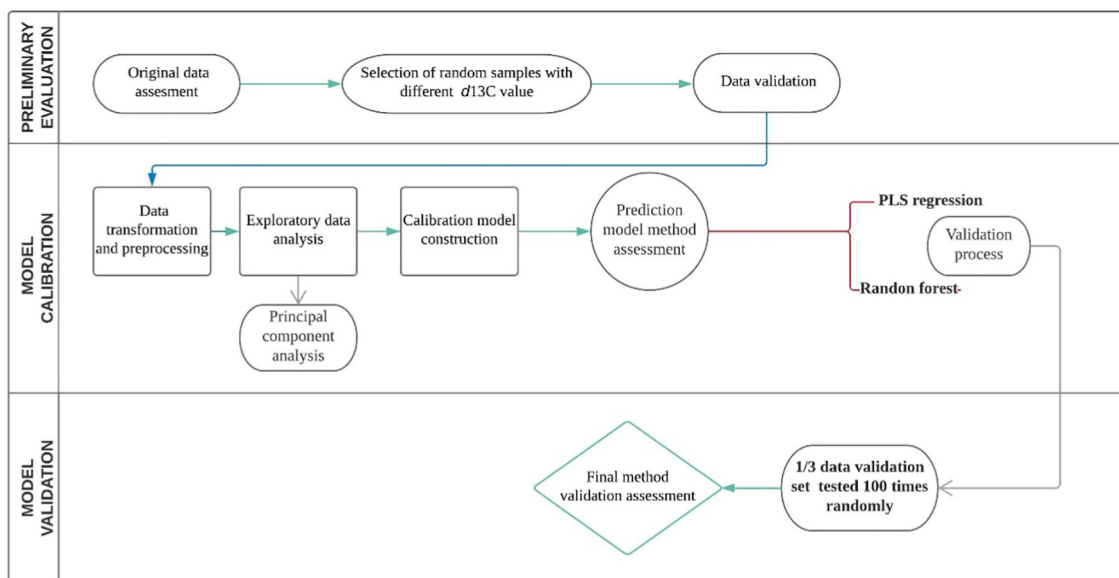
Fig. 1. Flowchart of the model generation.

meter (Hanna Instruments, US). The total reserve in base cations was measured (TRB, the sum of total Ca, Na, K, and Mg, in cmol$_c$ kg^{-1}) by following the protocols published by Herbillon (1986).

2.4. Data treatment, principal component analysis (PCA), model generation, and predictive model validation of $\delta^{13}C$

From the total dataset ($n = 332$), six samples were identified as outliers using a leverage × studentised residual plot, obtained using the Pirouette software (Infometrix, Bothell, WA, USA) and eliminated from the set. An initial investigation of the structure of the data was performed with the Pirouette software by a PCA. NIR spectra of all samples were analysed together to visualise spontaneous relationships and clustering among all samples, natural clustering in the data, and outlier samples.

For model generation, the spectroscopy data were pretreated to eliminate nonlinearities produced by light scattering. For NIRS of soil, these include variability in the light scattering due to soil roughness, aggregates, structure, and particle size. The raw NIRS data were treated by smoothing (Savitzky–Golay filter, 11 points), multiplicative scatter correction (MSC), normalisation, and mean centering. MSCs are the most widely used preprocessing techniques for NIRS data. Artefacts or imperfections (e.g., undesirable scatter effects) are removed from the data matrix prior to data modelling (Rinnan et al., 2009). In the total dataset (332 samples), six samples from the Choshuenco site were identified as outliers and eliminated from the set. Prediction models were created using the pretreated data. The first models were generated by partial least-squares (PLS) regression, a standard method in the multi-variate analysis (Martens and Naes, 1989). Leave-one-out cross validation was performed on the PLS model as an internal validation, which approximates the results that are likely to be obtained by an external validation. This method removes one sample from the training set, performs PLS regression on the remaining samples, predicts the value for the left-out sample, and then analyses the error. This process is repeated until every sample has been left out once. In this manner, the root mean square error of cross validation (RMSECV) is computed.

Using the pretreated data, a second set of models was generated using a random forest (RF) approach. RF implements the Breiman's RF algorithm for classification and regression based on a forest of trees using random inputs (Cutler et al., 2012). The RF was generated through regression using 500 trees and 1534 variables (i.e., 1/3 of the total number of used waves lengths) at each split.

The capability of the model to predict $\delta^{13}C$ of SOM for samples outside the training set was evaluated as follows: 1/3 of the data (111 samples) were selected randomly as the validation set, while the remaining 2/3 of the data were used to train the models. This was tested 100 times on randomly selected training and validation sets to evaluate the prediction abilities of the generated models within a single sample set.

The model performance was assessed based on the root mean square error of the prediction (RMSEP) of the validation set and based on the correlation between the predicted and measured $\delta^{13}C$ values. As mentioned above, the sites selected to validate our model were carefully chosen to maximise climatic and physicochemical diversity. Additionally, a test was added in order to observe if enough diversity had been covered with the selection of sites. Prediction models were constructed using PLS (1 to 10 latent variables) and RF. In this test; data from 10 sites was used to train the model and the remaining site used for validation, this was repeated for every site.

To investigate the origin of the predictability of $\delta^{13}C$ by NIR spectra, we carried out Pearson correlation analyses between $\delta^{13}C$ (measured by IRMS and predicted by the best PLS model) and several physicochemical variables (% C, % N, % clay, TRB). We used the NIR spectra (1100–2800 nm) to generate PLS models to predict these physicochemical variables. We visually compared the regression coefficients of these models to those of the best $\delta^{13}C$-predicting PLS model to identify potential shared wavelength bands between the models, which might indicate the causal correlation underlying the predictability of $\delta^{13}C$. As the physicochemical variables % N, % clay, and TRB were only available for the 0–10 cm, 10–30 cm, and in some cores, 30–60 cm intervals of one composite sample per site ($n = 29$), whereas NIR spectra and IRMS data were only available at a finer depth resolution ($n = 332$); the spectra and $\delta^{13}C$ values were averaged accordingly before the pretreatment and model validations were carried out as described above.

Unless indicated otherwise, all data pretreatments and model generations were carried out using the statistical software R 3.6.2 and packages "prospectr", "pls", "Chemospec", "BBmisc", "clusterSim", "plsRglm", and "radiant.data".

3. Results

3.1. Sites and soil-selected characteristics

The climatic and soil taxonomic variety of the soils included in this study is illustrated in Table 1. Table 2 shows the physicochemical variation that the soils in this study covered. Generally, the carbon and nitrogen contents decreased with the depth. Large differences between the locations were observed. The clay content also varied along the transect and it was in the range of 7 to 37%. Low bulk densities (< 1 g cm^{-3}) were associated to Andisols and Inceptisols (Table 3).

3.1.1. PCA as a clustering method

Fig. 2 shows a score plot of the first two principal components of a single PCA. The panels in Fig. 2a, b are coloured based on different criteria (for example, moisture and temperature regimes) using some of the data in Table 1. However, this information did not force sample distribution
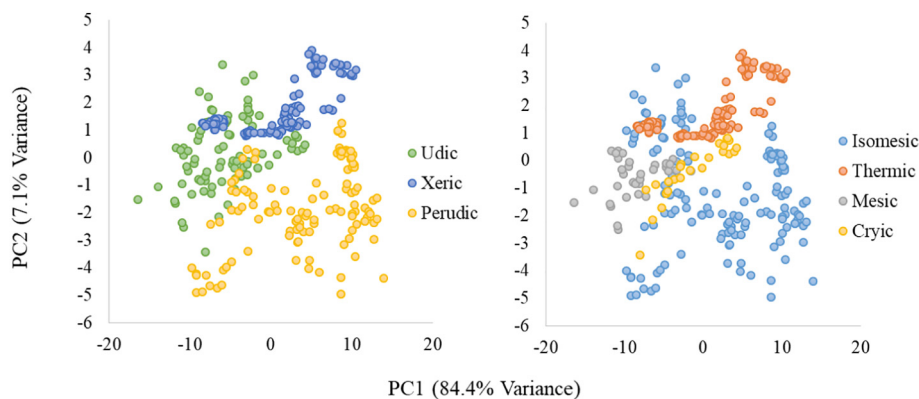

Fig. 2. PCA plots based on the NIR spectra of all samples used in this study ($n = 332$). a) SMR and b) STR plots.
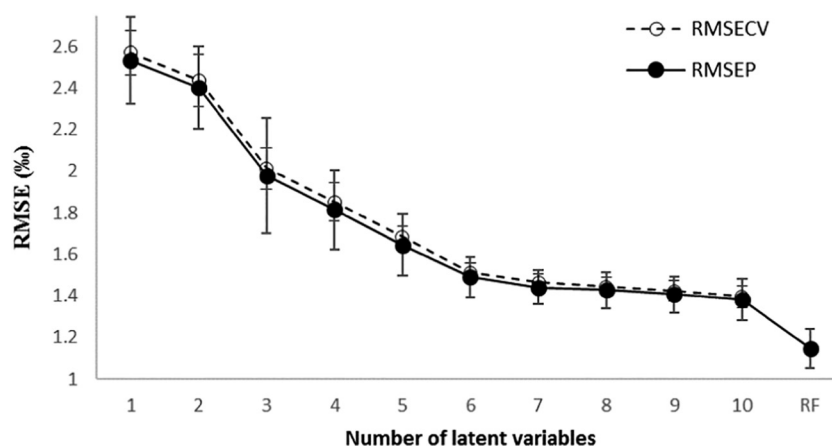
Fig. 3. Averages and SDs of RMSECV and RMSEP for the PLS models as functions of the number of LVs and for the RF model determined using 100 randomly selected validation (111 samples) sets in the entire dataset (332 samples).

Table 4
Calibration and external validation performances of selected models.

Model type	Validation	Random (1/3 of the dataset)		
	# LV	Var	RMSEP	R^2
PLS	1	$-87 \pm 16\%$	$2.53 \pm 0.21‰$	0.1%
	2	$-66 \pm 18\%$	$2.40 \pm 0.20 ‰$	2.9%
	3	$-6 \pm 10\%$	$1.98 \pm 0.28 ‰$	23.5%
	4	$11 \pm 7\%$	$1.81 \pm 0.19‰$	28.6%
	5	$26 \pm 5\%$	$1.65 \pm 0.15‰$	40.4%
	6	$40 \pm 4\%$	$1.49 \pm 0.10 ‰$	48.3%
	7	$44 \pm 4\%$	$1.44 \pm 0.08‰$	49.8%
	8	$48 \pm 3\%$	$1.43 \pm 0.09‰$	49.2%
	9	$55 \pm 4\%$	$1.41 \pm 0.09‰$	51.6%
	10	$59 \pm 3\%$	$1.38 \pm 0.10 ‰$	54.1%
RF	n.a.	73%	$1.15 \pm 0.09‰$	62.5%
Average	n.a.	n.a.	$1.70 \pm 1.29‰$	n.a.

Average: a model using the average of δ^{13}C of the training set as a common value for all validation samples, #LV: number of LVs used to construct the PLS model, RMSEP (mean $\pm$ SD of 100 training iterations), R^2 (Pearson correlation for the predicted δ^{13}C as a function of the measured δ^{13}C of the validation sets), n.a.: not applicable.

because PCA is an unsupervised pattern recognition method. In SMRs, it is possible to observe clustering between the same moisture regimes. Xeric soils (blue) seemed to be separated from perudic soils by PC2 (variance: 7.1%) and partially separated from udic soils by both PC1 and PC2. Udic soils were also separated from perudic soils by the contribution of the two PCs. In STRs, it is also possible to observe separate clusters in thermic, mesic, and cryic soils, while isomesic samples appear evenly scattered in the space described by the first two PCs. The similarity between soils with isomesic and mesic regimes is attributed mainly to their comparable mean annual soil temperature ranges. Another clustering pattern showed that mesic soils (grey) seemed to be separated from thermic soils by PC2. Similarly, the thermic regime (orange) in Fig. 2b and xeric regime (blue) in Fig. 2a are separated by PC2.

3.1.2. NIRS as a predictive method

The δ^{13}C values (-31 to $-21.5‰$), obtained by EA–IRMS, were used to train and validate the models. The average RMSECV and RMSEP of the PLS models for 100 randomly selected training (2/3 of the dataset) and validation (1/3 of the dataset) sets as functions of the number of latent variables (LVs) are shown in Fig. 3. Both RMSECV and RMSEP decreased with the increase in the number of LVs. The RMSEP of the model with five LVs ($1.65 \pm 0.15‰$) was only marginally better than that of a model where the average value of the calibration set was used as a single predictor for the validation set (RMSEP $= 1.70 \pm 1.29‰$). Therefore, only PLS models with six to ten LVs are analysed below. The average

absolute difference between RMSECV and RMSEP for more than six LVs was $0.11 \pm 0.08‰$. Thus, RMSECV is a good predictor of the RMSEP (Fig. 3).

In general, the RF model was better than the PLS models (Fig. 3 and Table 4). The average RMSEP for 100 randomly selected validation sets was $1.14 \pm 0.15‰$, considerably better than that obtained by the PLS models.

The removal of one site from the training set to predict the δ^{13}C of the omitted site had little to no effect on the model performance for 6 of the 11 sites, while for the remaining five sites omitting the site during the model training had a sever effect on its performance, see Fig. S1. The RF models, were much less affected by omitting one site during training than the models generated using PLS.

Notably, neither the δ^{13}C values measured by IRMS nor those predicted by the best PLS model correlated significantly with the soil C and N contents, as shown in Fig. 4. However, the δ^{13}C values correlated with the clay content ($r_{(PLS)} = 0.74$, $r_{(IRMS)} = 0.76$, $p < 0.05$, $n = 24$) and total reserve base cations ($r_{(PLS)} = 0.70$, $r_{(IRMS)} = 0.62$, $p < 0.05$, $n = 24$). The correlations between the δ^{13}C values derived from the IRMS and prediction slightly differed, likely owing to the loss of variation upon their averaging to match the composite samples for which soil physicochemical data were available.

The performances of the PLS models for the prediction of % C, % N, % clay, and TRB ranged from very good to mediocre, with the lower performances of the clay and N models possibly owing to the limited sample sizes (% C: $n = 332$, 10 LVs, $R^2 = 0.87$, RMSEP $= 1.92 \pm 0.09$ standard deviation (SD) %; % N: $n = 29$, 4 LVs, $R^2 = 0.50$, RMSEP $= 3.61 \pm 0.67$ SD %; % clay: n $= 29$, 3 LVs, $R^2 = 0.31$, RMSEP $= 8.86 \pm 0.82$ SD %; TRB: n $= 29$, 7 LVs, $R^2 = 0.81$, RMSEP $= 4.88 \pm 1.41$ SD $cmol_c$ kg^{-1}). However, this is not a major issue as these models were generated only to visually compare their regression coefficients to those of the δ^{13}C PLS model.

4. Discussion

Climate can affect soil carbon storage by changing, through photosynthesis, plant biomass inputs; by affecting rates of enzymatic microbial decomposition; and by altering geochemical properties that can protect soil organic matter from decomposition (Davidson, 2015). The interactions of climatic and geochemical factors control soil organic carbon storage and turnover in grasslands (Doetterl et al., 2015); but soil organic C density (SOCD) and its driving factors are also depending of the ecosystems and soil depth examined (Wang et al., 2013; Guan et al., 2019). Wang et al. (2013) found that soil organic carbon and δ^{13}C were correlated with soil characteristics across different
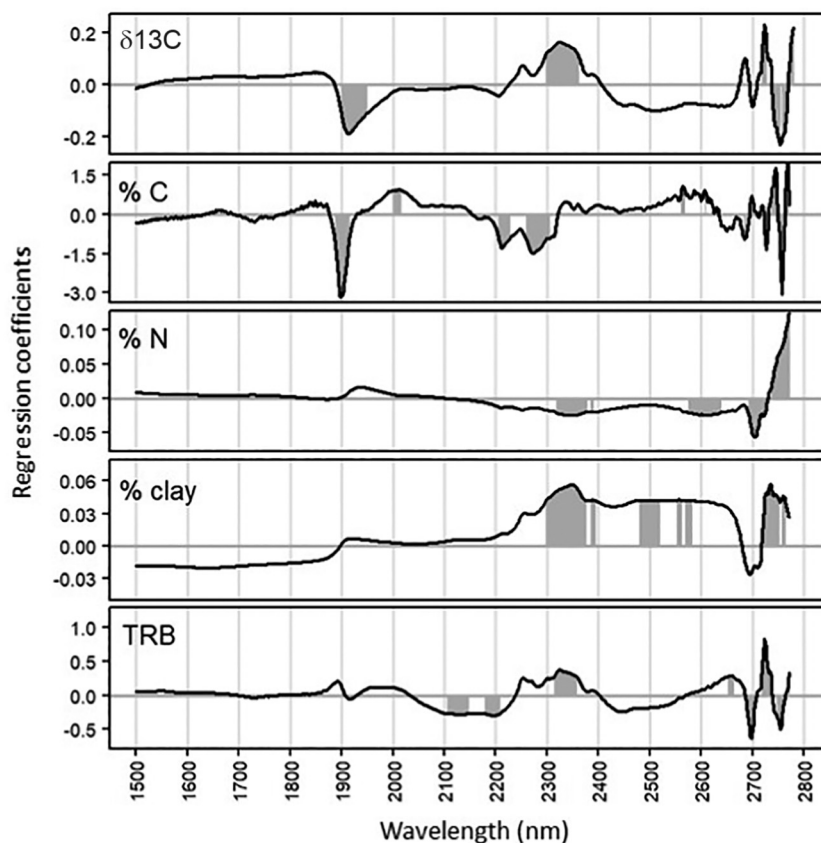
Fig. 4. Regression coefficients of the NIR spectrum wavelengths used in the best PLS model to predict $\delta^{13}C$ as well as for the PLS models, which predict the soil C content (% C), soil N content (% N), clay content (% clay), and TRB. Only the wavelength range of 1500 to 2750 nm is shown, as the wavelengths below 1500 nm did not contribute significantly to the models. The 10% of the wavelengths with the largest absolute regression coefficients in each model are highlighted in grey.

ecosystems (e.g. forest; meadow, steppe; croplands) and concluded that SOCD is a key contributor to the variation of soil $\delta^{13}C$. The spatial representation of the sampling sites in their study and in ours to cover spatial variability of soil organic carbon and $\delta^{13}C$ across wide transects is challenging.

The stable isotope ratios of the life science elements carbon, hydrogen, oxygen, and nitrogen vary slightly, but significantly in major compartments of the earth (Ghosh and Brand, 2003). They provide a method to quantify the contributions of different components to the ecosystem exchange. ^{13}C natural abundance measurements have helped understand soil carbon dynamics and develop kinetically defined SOC pool sizes and turnover rates (Paul, 2016). Better understanding of soil carbon dynamics is essential to understand the roles of soil carbon in the carbon cycle and feedback mechanisms in climate changes. Thus, the measurement of soil carbon dynamics requires an accurate assessment of isotopic variations, which can be distinguished by mass spectrometric measurements of soil samples (Ghosh and Brand, 2003). However, low-cost soil analysis alternatives are needed because the number of samples involved in such studies is large. Our study included different soil types and profiles under different environmental conditions to calibrate a predictive model and estimate $\delta^{13}C$ using NIRS, which covers a larger environmental gradient.

Exploratory analysis results (PCA) showed that the NIR spectra of the soils cluster is based on climatic regimes (Fig. 2) indicating that the climate shapes the soil physicochemistry, as reflected in the NIR spectra. Thus, the chemical information contained in the spectra is correlated with environmental variables. The PLS approach is a standard method in chemometrics (Wold et al., 2001), and it is a common regression method used to predict $\delta^{13}C$ (Martens and Naes, 1989). RFs have been successfully used for various applications in several disciplines.

They provide a multi-purpose method that is applicable to both regression and classification problems, including multi-class classification (Cutler et al., 2012). NIRS calibration data for ^{13}C EA–IRMS (Table 4) show that it is possible to generate a suitable predictive model. Our primary concern was to select the correct method for the generation of a model to predict $\delta^{13}C$ in Chilean soils. We used RMSEP to assess the number of chosen components, so that the model has adequate information to provide reliable predictions. Conversely, if too many components are chosen, the model would have noise as well as information, which would lead to less reliable predictions to be included to maximise the prediction capacity of the model and avoid over-fitting. The edaphoclimatic conditions in Chile are extremely variable. Thus, reliability is required for the development of a model for this type of transect. A model trained with soils having a large physicochemical variability is required. We evaluated the performances of the two methods for NIRS-based $\delta^{13}C$ prediction by their RMSEP values. RMSEP is a helpful measure of accuracy because it reflects the average differences between the measured and predicted values. Therefore, the model with more components was selected for PLS regression (Table 4, Fig. 3). As for both models (PLS and RF), the whole dataset was used and split into 2/3 (fit) and 1/3 (validation) sets randomly (repeated 100 times), and our approach provides a true assessment of the capability of the model to predict ^{13}C of samples within a sampling set. We evaluated the two methods for generation of models based on the validation set performances and qualities of the $\delta^{13}C$ values predicted using the NIRS data. The RMSEP values of the RF and PLS (with ten LVs) models were 1.15‰ and 1.38‰, respectively. Winowiecki et al. (2017) used similar models for soils in Africa under C3 and C4 plant species and obtained RMSEP = 1.95 and $R^2 = 0.80$ using PLS regression. For the RF approach they presented an average RMSEP of 1.77‰ and R^2 of

0.84. The R^2 values of the RF and PLS (with 10 LVs) models were 62.5 and 54.1, respectively (Table 4). Our results are similar to those reported by Winowiecki et al. (2017) and Fuentes et al. (2012). In the three studies, the samples represent various conditions, such as vegetation classes (Winowiecki et al., 2017), crop residues and rotations (Fuentes et al., 2012), and edaphoclimatic conditions (ours). The results obtained in these studies favour the use of NIRS as a predictive method, providing stable and rapid readings of ^{13}C in SOC. In this study, we further showed that the use of NIRS for δ^{13}C prediction is feasible at different soil depths along large transects with diverse soil types across Chile, which supports the studies on carbon dynamics in soils during climate changes.

We selected contrasting soils to test the spectral method capacity to adapt to the variability of conditions in the Chilean transect in order to found a general model. The selection of sites to train a predictive model has a large effect on the capability of the model to predict δ^{13}C of other sites. While some sites appeared to belong to a similar 'class' (i.e. they could still be well predicted when not included in the training set), other sites had unique features that interferes with the model's predictive capacity to estimate their δ^{13}C if they were not included in the training set (Fig. S1). Despite this, including samples of all sites for training a general model capable of predicting δ^{13}C from all the selected sites with acceptable accuracy could still be constructed. In general, the models constructed using RF, performed better to predict samples from sites outside the training set than PLS models. Generating separate models for per site 'class' could largely improve the performance of the predictive model within that class. Unfortunately, a criterion to separate the sites into classes 'a priori' (e.g. based on physicochemical, climatic, geographic parameters) could not be encountered using the current dataset. Therefore, we recommend a dedicated training data set per soil type (once is defined which types could be grouped in a site 'class') for a more robust and precise prediction of ^{13}C in soil profiles.

To investigate the origin of the predictability of δ^{13}C by NIR spectra, we analysed the regression vectors of the δ^{13}C-predicting PLS model and compared it to models predicting the soil C and N contents, clay content (Fig. 4) as a proxy for physical and chemical weathering, and total reserve base ions as a proxy for chemical weathering. The regression vectors of the models consist of the regression coefficients of all wavelengths in the NIR spectra, and thus reflect the importance of a given wavelength for the prediction of the various response variables. In Fig. 4 is also possible to observe differences in peaks of NIR spectrum, in the range of 2300 to 2400 nm, when comparing C% and δ^{13}C. Vibrational differences, probably due to mass and chemical environment interaction of the isotopes, could be the reason why this characteristic signal is produced that allows us to establish a clear difference between C content and ^{13}C abundance. In literature this range of wavelengths bands is known to correspond with combinations of vibrations of C–H and C–C (Mielke and Sobczyk, 2006; Stenberg et al., 2010). However, although it is clear that the presence of a ^{13}C molecule will clearly shift C–H and C–C vibration to lower shorter wavelengths, the effect that this has on PLS regression explanatory wavelength is much less clear. As it will be a result on the effect of multiple C–C and C–H vibration types simultaneously.

The spectra start at 2778 nm in the MIR range, which is known to be related to signals from O–H vibration of acid groups, which can be related to degradation. The δ^{13}C model draws influential information around 2753 nm and 2723 nm. These peaks are also important for prediction of soil C content, TRB, and clay. The influential pattern between 2150 and 2400 nm is shared with the models predicting clay and TRB, while it does not fit the pattern of C and N prediction. This band of wavelengths is known to correspond with combinations of vibrations of C–H and C–C (Mielke and Sobczyk, 2006; Stenberg et al., 2010). A strong negative peak between 2000 and 1800 nm is partly shared with the model predicting soil C. This region is typically considered to correspond with the overtone of C=O and combinations with O–H stretching (O–H stretching is not informative, but the resulting adsorption can be

combination with a C–C or C–H or C–O stretches that are relevant) (Stenberg et al., 2010). We thus found indications that the prediction of δ^{13}C values draws its most influential information from the same NIR ranges as the prediction of soil C, clay content and total reserve base ions. This, together with the strong correlation of δ^{13}C values with clay content and TRB, suggests that ^{13}C/^{12}C ratio of SOC may be related to soil mineralogical properties, and that thus also the spectral prediction of δ^{13}C works in parts via this correlative relationship. Possible mechanisms could be that soil mineralogical properties affect carbon decomposition via physical or chemical protection (Davidson, 2015; Doetterl et al., 2015) which in turn is known to affect δ^{13}C values of SOM (Accoe et al., 2003).

According to Zornoza et al. (2008), high demands exist for rapid and predictive soil data acquisitions in environmental monitoring, soil quality assessment, and emerging methods of soil analysis. They correlated the soil fertility and physical and biological properties using NIRS. In this study, we obtained acceptable predictions for the two selected models, and then, based on RMSEP and R^2, the RF model was chosen as the best model.

NIRS is considerably more affected by the physical structure variability caused by unrelated parameters, such as soil roughness, aggregates, structure, and particle size (size of aggregates, porosity, water) (Bellon-Maurel and Mc Bratney, 2011; Reyna et al., 2017), in comparison to Mid-Infrared (MIR) spectroscopy, which is another low-cost and easy-to-use alternative technique, NIR requires a simpler sample preparation than that of MIR. A strong further advantage of NIR is that is already widely established for quantification of a wide range of soil properties (Reeves and McCarty, 2001; Blanco and Villarroya, 2002; Gomez et al., 2008).

NIR spectra were used to obtain a regression model. Abundant information is contained in the spectral data. However, only a few latent variables are necessary to obtain a good correlation with the predicted property (in this case, δ^{13}C). For reliability of this technique, it is necessary to include a large number of samples from zones with wide ranges of values of soil properties (Zornoza et al., 2008). In this study, we covered various soils, which largely differed in different properties, and we used them to correlate the spectral data to ^{13}C. Thus, the calibrations are valid for important and different environmental systems. According to Eiler (2013), "historical precedence suggests that no emerging analytical capability grows to its full potential unless it meets a serious need in the applied sciences".

5. Conclusion

A prediction model for δ^{13}C based on NIRS data was developed using the PLS regression and RF approaches. Better results were obtained by the latter approach. The RMSEP parameters of both models indicated that NIRS can be used to predict δ^{13}C for various soil profiles. The stable and rapid readings of ^{13}C of SOC obtained in this study support the use of NIRS as a predictive method in soil analysis and as a nondestructive waste-free method for the studies on carbon dynamics in soil. As expected the selection of samples for model training has a large impact on the capability of the generated predictive model. We are aware that the quality of the model to predict samples decreases with sites that differ from the calibration samples in terms of their spectral properties, but they are very well suited to predict similar samples, within the sites with similar spectral characteristics to the sites described in the supplementary information, thus still useful to speed up analyses for large areas and or different depths. A country like Chile has a large environmental variability that hinders the accuracy of predictions. At this point, if this approach for training is followed, it has the advantage we can use the spectrometric method for different soil types, but the likely at the expense of lower prediction accuracy; the disadvantage is if we add another site (not included in the training) we fail to predict the ^{13}C. Generating separate models for per site 'class' could largely improve the performance of the predictive model within that class.

It was concluded that the predictability of δ^{13}C in soils may be linked to its correlation with soil mineralogical properties, for which variables such as the clay content and total reserve base ions are proxies.

Declaration of Competing Interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geodrs.2021.e00397.

References

Accoe, F., Boeckx, P., Cleemput, O.V., Hofman, G., 2003. Relationship between soil organic C degradability and the evolution of the δ^{13}C signature in profiles under permanent grassland. Rapid Commun. Mass Spectrom. 17 (23), 2591–2596.

Balesdent, J., Mariotti, A., 1996. Measurement of soil organic matter turnover using ^{13}C natural abundance. In: Boutton, T.W., Yamasaki, S. (Eds.), Mass Spectrometry of Soils. Marcel Dekker, New York, pp. 83–111.

Balesdent, J., Girardin, C., Mariotti, A., 1993. Site-related δ^{13}C of tree leaves and soil organic matter in a temperate forest. Ecology 74, 1713–1721.

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

Blanco, M., Villarroya, Ignacio, 2002. NIR spectroscopy: a rapid-response analytical tool. Trends Anal. Chem. 21, 240–250.

Bouyoucos, G.J., 1962. Hydrometer method improved for making particle size analysis of soils. Agron. J. 54, 464–465.

Bricklemyer, R.S., Miller, P.R., Paustian, K., Keck, T., Nielsen, G.A., Antle, J.M., 2005. Soil organic carbon variability and sampling optimization in Montana dryland wheat fields. J. Soil Water Conserv. 60, 42–51.

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Reichstein, M., 2014. Global covariation of carbon turnover times with climate in terrestrial ecosystems. Nature 514, 213–217.

Chabbi, A., Kögel-Knabner, I., Rumpel, C., 2009. Stabilised carbon in subsoil horizons is located in spatially distinct parts of the soil profile. Soil Biol. Biochem. 41, 256–261.

Chen, Q., Shen, C., Sun, Y., Peng, S., Yi, W., Li, Z.A., Jiang, M., 2005. Spatial and temporal distribution of carbon isotopes in soil organic matter at the Dinghushan biosphere reserve, South China. Plant Soil 273 (1–2), 115–128.

Crowther, K.E., Todd-Brown, O., Rowe, C.W., Wieder, W.R., Carey, J.C., Machmuller, M.B., Snoek, B.L., Fang, S., Zhou, G., Allison, S.D., Blair, J.M., Bridgham, S.D., Burton, A.J., Carrillo, Y., Reich, P.B., Clark, J.S., Classen, A.T., Dijkstra, F.A., Elberling, B., Emmett, B.A., Estiarte, M., Frey, S.D., Guo, J., Harte, J., Jiang, L., Johnson, B.R., Kröel-Dulay, G., Larsen, K.S., Laudon, H., Lavallee, J.M., Luo, Y., Lupascu, M., Ma, L.N., Marhan, S., Michelsen, A., Mohan, J., Niu, S., Pendall, E., Peñuelas, J., Pfeifer-Meister, L., Poll, C., Reinsch, S., Reynolds, L.L., Schmidt, I.K., Sistla, S., Sokol, N.W., Templer, P.H., Treseder, K.K., Welker, J.M., Bradford, M.A., 2016. Quantifying global soil carbon losses in response to warming. Nat. Lett. Res. 540, 104–108.

Cutler, A., Cutler, D.R., Stevens, J.R., Zhang, C., 2012. Random forests. In: Ma, Y.Q. (Ed.), Ensemble Machine Learning. Springer, Boston, MA.

Davidson, E.A., 2015. Soil carbon in a beer can. Nat. Geosci. 18, 748–749.

Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Casanova Pinto, M., Casanova-Katny, A., Muñoz, C., Boudin, M., Zagal Venegas, E., P., 2015. Soil carbon storage is controlled by interactions between geochemistry and climate. Nat. Geosci. 8, 780–783.

Doetterl, S., Berhe, A., Arnold, C., Bodé, S., Fiener, P., Finke, P., Fuchslueger, L., Griepentrog, M., Harden, W., Nadeu, E., Schnecker, J., Trumbore, S., Van Oost, K., Vogel, C., Boeckx, P., 2018. Links among warming, carbon, and microbial dynamics mediated by soil mineral weathering. Nat. Geosci. 11, 589–593.

Eiler, J.M., 2013. Isotopic Anatomies of Molecules and Minerals Annual Review of Earth and Planetary Sciences. 41 pp. 411–441.

Finke, P., Opolot, E., Balesdent, J., Berhe, A.A., Boeckx, P., Cornu, S., Harden, J., Hatté, C., Williams, E., Doetterl, S., 2019. Can SOC modelling be improved by accounting for paedogenesis? Geoderma 338, 513–524.

Fontaine, S., Barot, S., Barre, P., Bdioui, N., Mary, B., Rumpel, C., 2007. Stability of organic carbon in deep soil layers controlled by a fresh carbon supply. Nature 450, 277–281.

Fuentes, M., Hidalgo, C., González-Martín, I., Hernández-Hierro, J.M., Govaerts, B., Sayre, K.D., Etchevers, J., 2012. NIR spectroscopy: an alternative for soil analysis. Commun. Soil Sci. Plant Anal. 43 (1–2), 346–356.

Ghosh, P., Brand, W.A., 2003. Stable isotope ratio mass spectrometry in global climate change research. Int. J. Mass Spectrom. 228, 1–33.

Glaser, B., 2005. Compound-specific stable isotope (δ^{13}C) analysis in soil science. J. Plant Nutr. Soil Sci. 168 (5), 633–648.

Gomez, C., Rossel, R.A.V., McBratney, A.B., 2008. Soil organic carbon prediction by hyperspectral remote sensing and field Vis-NIR spectroscopy: an Australian case study. Geoderma 146 (3–4), 403–411.

Guan, Jin-Hong, Deng, Lei, Zhang, Jian-Guo, He, Qiu-Yue, Shi, Wei-Yu, Li, Guoqing, Du, Sheng, 2019. Soil organic carbon density and its driving factors in forest ecosystems across a northwestern province in China. Geoderma 352, 1–12.

Herbillon, A.J., 1986. In: Beinroth, F.H., Camargo, M.N., Eswaran, H. (Eds.), Proc. 8th Int. Soil Classification Workshop, pp. 39–48 (EMBRAPA-SNLCS, 1986).

Kuzyakov, Y., Biriukova, O., Turyabahika, F., Stahr, K., 2001. Electrostatic method to separate roots from soil. J. Plant Nutr. Soil Sci. 164, 541–545.

Lal, R., 2006. Soil carbon sequestration in Latin America pp. 49–64. In: Lal, R., et al. (Eds.), Carbon Sequestration in Soils of Latin America the Haworth Press. Binghamton, New York, USA.

Martens, H., Naes, T., 1989. Multivariate Calibration Johnson Wiley & Sons, Chichester.

Mielke, Z., Sobczyk, L., 2006. In: Kohen, A., Limbach, H.-H. (Eds.), In Isotope Effects in Chemistry and Biology, p. 281.

Mooney, S., Antle, J., Capalbo, S., Paustian, K., 2004. Influence of project scale and carbon variability on the costs of measuring soil carbon credits. Environ. Manag. 33, S252–S263.

Paul, E.A., 2016. The nature and dynamics of soil organic matter: plant inputs, microbial transformations, and organic matter stabilisation. Soil Biol. Biochem. 98, 109–126.

Petisco, C., García-Criado, B., Mediavilla, S., Vázquez De Aldana, B.R., Zabalgogeazcoa, I., García-Ciudad, A., 2006. Near-infrared reflectance spectroscopy is a fast and non-destructive tool to predict foliar organic constituents of several woody species. Anal. Bioanal. Chem. 386 (6), 1823–1833.

Reeves, J.B., McCarty, G.W., 2001. Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and a fibre-optic probe. J. Near Infrared Spectrosc. 9, 25–34.

Reyna, L., Dubé, F., Barrera, J.A., Zagal, E., 2017. Potential model overfitting in predicting soil carbon content by visible and near-infrared spectroscopy. Appl. Sci. 7, 708.

Rinnan, A., Van den Berg, F., Engelsen, B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trends Anal. Chem. 28 (10), 1201–1222.

Rumpel, C., Kögel-Knabner, I., Bruhn, F., 2002. Vertical distribution, age, and chemical composition of organic carbon in the two forest soils of different paedogenesis. Org. Geochem. 33, 1131–1142.

Sandoval, M., Dörner, J., Seguel, O., Cuevas, J., Rivera, D., 2012. Physical Soil Analysis Methods, Department of Soils and Natural Resources Universidad de Concepción Publication No 5 (80p).

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and near-infrared spectroscopy in soil science. Adv. Agron. 107 (163), 215.

Theo, J., 2005. Short introduction to infrared and Raman spectroscopy. In: Theo, J. (Ed.), The application of vibrational spectroscopy to clay minerals and layered double hydroxides. Clay Minerals Society, Chantilly, France, pp. 1–8.

Trumbore, S., 2009. Radiocarbon and soil carbon dynamics. Annu. Rev. Earth Planet. Sci. 37 (1), 47–66.

Viscarra Rossel, R.A., McBratney, A.B., 1998. Laboratory evaluation of a proximal sensing technique for simultaneous measurement of clay and water content. Geoderma 85 (1), 19–39.

Viscarra Rossel, R.A., Walvoort, D., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near-infrared, mid-infrared, or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131 (1–2), 59–75.

Wang, Shaoqiang, Fan, Jiangwen, Song, Minghua, Yu, Guirui, Zhou, Lei, Liu, Jiyuan, Zhong, Huaping, Gao, Lupeng, Hu, Zhongmin, Wu, Weixing, Song, Ting, 2013. Patterns of SOC and soil 13C and their relations to climatic factors and soil characteristics on the Qinghai–Tibetan Plateau. Plant Soil 363, 243–255.

Winowiecki, L.A., Vågen, T., Boeckx, P., Dungait, J., 2017. Landscape-scale assessments of stable carbon isotopes in soil under diverse vegetation classes in. East Africa 421 (1–2), 259–272.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool for chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130.

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., Mataix-Beneyto, J., 2008. Near infra-red spectroscopy for the determination of various physical, chemical, and biochemical properties in Mediterranean soils. Soil Biol. Biochem. 40, 1923–1930.