



DOI: <http://dx.doi.org/10.1590/1807-1929/agriambi.v23n6p460-466>

## Estimation of elemental composition of agricultural soils from West Azerbaijan, Iran, using mid-infrared spectral models

Maryam Ghebleh Goydaragh<sup>1</sup>, Ali Asghar Jafarzadeh<sup>1</sup>, Farzin Shahbazi<sup>1</sup>, Shahin Oustan<sup>1</sup>, Ruhollah Taghizadeh-Mehrjardi<sup>2,3</sup> & Marcos Lado<sup>4</sup>

<sup>1</sup> University of Tabriz/Faculty of Agriculture/Department of Soil Science. Tabriz, Iran. E-mail: maryam.ghebleh@tabrizu.ac.ir - ORCID: 0000-0001-8684-9942; jafarzadeh1396@yahoo.com - ORCID: 0000-0001-5681-7085; shahbazi@tabrizu.ac.ir - ORCID: 0000-0002-0024-7527; oustan@hotmail.com - ORCID: 0000-0002-4291-341X

<sup>2</sup> University of Tübingen/Department of Geosciences, Soil Science and Geomorphology. Tübingen, Germany. E-mail: ruhollah.taghizadeh-mehrjardi@mnf.uni-tuebingen.de - ORCID: 0000-0002-4620-6624

<sup>3</sup> Ardakan University/Faculty of Agriculture and Natural Resources. Yazd, Iran.

<sup>4</sup> University of A Coruña/Faculty of Sciences/Centro de Investigaciones Científicas Avanzadas. Coruna, Spain. E-mail: marcos.lado@udc.es (Corresponding author) - ORCID: 0000-0001-6996-0964

**ABSTRACT:** Characterizing the elemental composition provides useful information about the weathering degree of soils. In Miandoab County, Northern Iran, this characterization was missing, and thus the objectives of this work were to evaluate the weathering degrees for the most typical soils in the area from their elemental compositions, and to estimate this elemental composition using Fourier transform infrared spectroscopy and Random Forest models. Five soil profiles, including Aridisols and Inceptisols, were selected as the most representative of the area. Major elemental oxides were determined in each genetic horizon by X-ray fluorescence, showing that these soils were at early developmental stages. Only  $Al_2O_3$  and CaO were accurately estimated, with  $R^2$  values of 0.8, and out-of-bag mean square errors of 0.2 and 1.1, respectively. The other oxides were not predicted satisfactorily, probably due to small differences in their elemental compositions. Random Forest provided the important spectral bands related to the content of each element. For  $Al_2O_3$ , these bands were between 500 and 650  $cm^{-1}$ , which represent out-of-plane OH bending vibrations and Al–O gibbsite and alumino-silicate vibrations. For CaO, the most important bands are related to carbonate content. A combination of Fourier transform infrared spectra and Random Forest models can be used as a rapid and low-cost technique to estimate the elemental composition of arid and semi-arid soils of Northern Iran.

**Key words:** FTIR spectra, random forest, weathering

## Estimativa da composição elementar de solos do Azerbaijão oeste, Irã, utilizando-se modelos espectrais de infravermelho

**RESUMO:** A caracterização da composição elementar fornece informações úteis para caracterizar o grau de alteração dos solos. Em Miandoab, norte do Irã, esta caracterização não existe. Os objetivos deste trabalho foram avaliar o grau de intemperismo dos solos típicos da região usando a sua composição elementar e estimar esta composição usando espectroscopia infravermelha com transformada de Fourier (FTIR) e modelos Random Forest (RF). Foram selecionados cinco perfis de solo, incluindo Aridisolos e Inceptisolos, como os mais representativos da área. Os principais óxidos elementares foram determinados por fluorescência de raios-X em cada horizonte genético, mostrando que estes solos estavam em um estágio de baixo grau de desenvolvimento. Apenas o  $Al_2O_3$  e o CaO foram estimados com precisão, com valores de  $R^2$  de 0,8 e erro quadrático médio nos dados utilizados para validação de 0,2 e 1,1, respectivamente, enquanto os outros óxidos não foram preditos satisfatoriamente, provavelmente devido às pequenas diferenças na sua composição. O modelo Random Forest forneceu importantes bandas espectrais relacionadas com o conteúdo de cada elemento. Para o  $Al_2O_3$ , estes atingiram a região 500 a 650  $cm^{-1}$ , o que foi atribuído a vibrações de flexão de OH e vibrações de Al-O de gibbsite e alumino-silicatos. Para o CaO, as bandas mais importantes estavam relacionadas ao teor de carbonatos. Os resultados indicam que uma combinação de espectros infravermelho de transformada de Fourier e modelos Random Forest pode ser usada como uma técnica rápida e de baixo custo para estimar a composição elementar de solos do norte do Irã.

**Palavras-chave:** espectros de FTIR, Random Forest, intemperismo



## INTRODUCTION

Weathering is defined as the alteration of parent materials by physical, chemical and biotic processes (Osat et al., 2016). The most important processes are those related to the chemical transformation of existing minerals, which leads to their depletion (e.g., by hydrolysis, oxidation, hydration or dissolution) and the formation of secondary minerals and hydrous oxides (Jeleńska et al., 2008).

The determination of elemental composition is commonly used to assess the weathering intensity of soils, and it is usually performed using X-ray fluorescence (XRF) spectroscopy. However, this technique is expensive and the sample preparation is time-consuming.

An alternative is to predict the elemental composition from information obtained by other spectroscopic methods like Fourier transform infrared (FTIR), near infrared (NIR) or Raman spectroscopy, which allow spectral measurements of large numbers of samples at a lower cost and with minimal sample processing (Robertson et al., 2016).

The combination of spectral data and multivariate analysis techniques has gained popularity for quantifying soil properties (Rial et al., 2016). These include statistical methods such as multiple linear regression, principal component regression, partial least squares regression (PLSR), and Random Forest (RF). There are some studies relating FTIR spectra and elemental soil compositions using PLSR (e.g., Mohanty et al., 2016), but the potential use of RF for this purpose has not been analysed.

In the last years, several researchers have described the mineralogy and properties of the soils in Northern Iran (Moradi et al., 2012; Saraskanroud et al., 2017). However, the study of their elemental composition has not yet been done. Thus, the main objective of this study was to investigate the composition of the most representative soils in Miandoab County (West Azerbaijan, Northern Iran) and to predict their elemental composition using a combination of FTIR spectroscopy and RF.

## MATERIAL AND METHODS

The studied area is located in Miandoab County, West Azerbaijan, northern Iran (36° 55' 37" 0' N and 46° 0' 46" 10' E), at an altitude that ranges from 1283 to 1308 m. Soil moisture and temperature regimes are aridic and mesic, respectively. The mean annual rainfall and evaporation are 272.3 and 753.3 mm, respectively, with a mean annual air temperature of 12.8 °C. The main land use in this area is agriculture, and the dominant crops are fruits and nuts such as apples, grapes and walnuts.

A number of 17 soil profiles were characterized in order to study the soils in the area. As a result, the following taxons were identified according to the USDA soil classification system (Soil Survey Staff, 2014): two soil orders (Aridisol and Inceptisol), four suborders (Calcids, Cambids, Argids and Xerepts), four great groups (Haplocalcids, Haplocambids, Haplargids and Haploxerepts) and four subgroups (Typic Haplocalcids, Typic Haplocambids, Typic Haplargids and Typic Haploxerepts).

In each profile, genetic soil horizons were identified and described, and composite samples of each horizon were

obtained. In total, the number of samples collected was 44. The samples were transported to the laboratory, air-dried and ground to pass through a 2-mm mesh sieve and analysed for general soil properties.

Physical and chemical analyses of the fine-earth fraction were performed using standard methods. These included particle size distribution by the hydrometer method (Gee & Bauder, 1986), organic carbon content using a Flash EA 1112 elemental analyser (ThermoQuest, USA) after carbonate removal with HCl, calcium carbonate equivalent (CCE) using a titration method (Nelson & Sommers, 1996) and cation exchange capacity (CEC) by the Bower method (Bower et al., 1952).

Saturated paste extracts were prepared to measure soil pH using a pH meter (AZ-86502), electrical conductivity (EC) using an EC meter (AZ-8301), and soluble cations and anions by atomic absorption spectrophotometry using a Shimadzu AA-670 spectrophotometer (Shimadzu, Japan). The sodium adsorption ratio (SAR) was then calculated using Eq. 1 (Soil Survey Staff, 2014).

$$\text{SAR} = \frac{\text{Na}^+}{\sqrt{\text{Ca}^{2+} + \text{Mg}^{2+}}} \quad (1)$$

The 44 air-dried samples were ground in an Agatha mortar, and 1 mg of soil was combined with 200 mg of potassium bromide, mixed and ground together before being pressed to produce a pellet. Absorbance of the pellets was measured in the MIR region (wavenumber range from 400 to 4000 cm<sup>-1</sup>, with a resolution of 2 cm<sup>-1</sup>) using a FTIR Bruker Vector 22 spectrophotometer (Bruker, USA) in transmission mode. The obtained spectra were baseline corrected and normalized using the extended multiplicative signal correction method (EMSC) (Kohler et al., 2005).

Based on the results of the physical and chemical analyses of all samples, five profiles were selected as the most representative of the study area, which included a total number of 19 samples. Total elemental compositions were determined by X-ray fluorescence (XRF) for these samples using a S4 Pioneer XRF spectrometer (Bruker, USA).

Random Forest (RF) was used to find relationships between elemental composition as dependent variables and FTIR spectra as predictor variables, as well as to identify the spectral bands most relevant to elemental composition. Random Forest is a machine learning algorithm based on the calculation of multiple randomized regression trees (e.g., 500 to 2000 trees). It is becoming very popular because of its good modelling characteristics: 1) it is a non-parametric method; 2) it can be applied to datasets with a low number of observations and many predictors; 3) it can be applied to two and multi-class problems; 4) it does not overfit; 5) it ranks the most important variables for prediction and 6) it has good performance, low bias, low variance, and has low correlations between the classification trees (Wiesmeier et al., 2011).

For the construction of the RF model, the original data (n = 19) was split through bootstrapping into two sets: a training set that included 75% of the data, which was used to fit the model and a test set that comprised the remaining 25%, which was

used for validation of the fitted model. The non-sampled data in the training set was replaced until its size was the same as that of the original dataset. Then, a regression tree was constructed with the training set, and the non-sampled data, or “out-of-bag” (OOB) data, was used to estimate the prediction error. This process was repeated 500 times, and the final model parameters were selected averaging the results of all fitted regression trees.

To evaluate the performance of each RF model, the root mean square errors (RMSE), coefficients of determination ( $R^2$ ), and OOB mean square errors ( $MSE_{OOB}$ ) were calculated according to the following equations (Eqs. 2, 3, and 4): where:

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

$$R^2 = 1 - \frac{SS_{error}}{SS_{total}} \quad (3)$$

$$MSE_{OOB} = \frac{1}{n} \sum_{i=1}^n (y_i - y_p^{OOB})^2 \quad (4)$$

$n$  - total number of samples;

$y_i$  - observed value;

$y_p$  - estimated value;  
 $SS_{error}$  - sum of squared errors between estimated and observed values; and,

$SS_{total}$  - sum of squared deviations of each response variable from its mean.

All data manipulations and analyses were performed in the R-Studio environment (R Core Team, 2017). Baseline removal and spectral corrections were performed using the EMSC R package (Liland, 2017), and Random Forest models were constructed using the randomForest R package (Liaw & Wiener, 2002).

## RESULTS AND DISCUSSION

Some of the general properties of the five representative profiles are presented in Table 1. Calcium carbonate contents were high, which was common in both arid and semi-arid areas, and generally increased with depth, most probably because of mobilization from the upper horizons during wet periods and re-deposition in deeper parts of the profiles.

There were higher EC values in P1 and P4 and SAR values in P1, P3, and especially in P4, when compared to the other profiles because of the presence of a shallow water table near the surface. Soil reaction in all horizons ranged from moderately basic to

**Table 1.** General properties of the five reference profiles

Horizon	Depth (m)	Colour (moist)	Texture <sup>a</sup>	Structure <sup>b</sup>	Special feature	EC <sup>c</sup> (dS m <sup>-1</sup> )	pH	CEC (cmol <sub>c</sub> kg <sup>-1</sup> )	SAR (mmol <sub>c</sub> L <sup>-1</sup> ) <sup>0.5</sup>	CCE (g kg <sup>-1</sup> )	OC (g kg <sup>-1</sup> )
P1: Typic Haplocalcid											
A	0-0.30	10YR3/3	SiL	2mgr		4.9	7.9	12.6	4.5	115.0	18.0
Bk1	0.30-0.70	10YR2/2	L	2mabk	CaCO <sub>3</sub> nodules	9.5	7.8	13.3	10.3	133.0	12.0
Bk2	0.70-1.60	10YR4/4	L	3mabk	CaCO <sub>3</sub> nodules	12.2	8.5	8.3	11.9	140.0	1.0
C	1.60-2.00	10YR3/3	S	sg		5.0	7.9	3.1	8.6	141.0	8.0
P2: Typic Haplocambid											
A	0-0.30	10YR3/4	L	1fgr		0.8	7.5	14.6	0.4	125.0	10.0
Bw1	0.30-0.60	10YR3/4	SiL	1fabk	Fe mottles	0.7	7.8	9.3	0.4	145.0	7.0
Bw2	0.60-1.30	10YR3/4	L	1fabk	Fe mottles	0.6	8.0	8.4	0.5	155.0	2.0
C	1.30-2.00	10YR3/3	L	sg		0.5	8.0	5.1	0.5	157.0	2.0
P3: Typic Haplocambid											
A	0-0.30	10YR3/4	C	3cabk		1.4	8.0	29.4	4.8	70.0	10.0
Bw1	0.30-0.60	10YR4/3	CL	2cabk		1.4	8.0	16.4	8.0	115.0	8.0
Bw2	0.60-1.10	10YR4/6	L	2mabk		3.0	8.5	13.1	12.1	170.0	4.0
C	1.10-2.00	10YR4/4	L	ma		1.7	9.0	11.9	9.2	136.0	4.0
P4: Typic Haplargid											
A	0-0.40	10YR3/4	CL	1vfgr		3.8	7.9	19.7	10.6	128.0	9.0
Bt1	0.40-0.70	10YR4/3	SiC	2fabk	Clay coatings	3.8	8.3	17.1	12.0	134.0	5.0
Bt2	0.70-1.30	10YR3/4	C	2fabk	Clay coatings	8.7	8.4	18.7	32.3	131.0	5.0
C	1.30-2.00	10YR4/4	L	ma		7.4	8.5	16.7	37.5	132.0	3.0
P5: Typic Haploxerept											
A	0-0.40	10YR3/4	SL	1fgr		0.7	7.9	10.6	1.0	113.0	4.0
Bw	0.40-1.10	10YR3/4	L	1cabk		0.6	7.9	9.1	0.5	134.0	3.0
C	1.10-2.00	10YR3/6	SL	sg		0.6	8.0	8.2	0.4	143.0	2.0

<sup>a</sup>SiL - Silty loam, L - Loam, S - Sandy, C - Clay, CL - Clay loam, SiC - Silty clay, SL - Sandy loam;

<sup>b</sup>1 - Weak; 2 - Moderate; 3 - Strong; vf - Very fine; f - Fine; m - Medium; c - Coarse; gr - Granular; abk - Angular blocky; ma - Massive;

<sup>c</sup>EC - Electrical conductivity; CEC - Cation exchange capacity; SAR - Sodium adsorption ratio; CCE - Calcium carbonate equivalent; OC - Organic carbon

slightly alkaline, as expected from the presence of  $\text{CaCO}_3$ , with pH values between 7.5 and 9.0. The pH increased with depth, except for the transition between the Bk2 and C horizons of P1, because of the accumulation of carbonates in subsurface horizons. Considering their pH, EC, SAR and  $\text{CaCO}_3$  contents, P2, P3 and P5 were classified as non-saline-non-alkali, P1 was classified as saline and P4 was classified as saline-alkali.

The amount of organic carbon (OC) was low in all soils ( $< 20 \text{ g kg}^{-1}$ ), with the highest values being observed in the A horizon of P1. Clay contents were also low, except in P3 and P4. In P4, the translocation of clay from the surface formed argillic horizons that were the diagnostic feature of the soil.

Cation exchange capacity was low ( $< 25 \text{ cmol}_c \text{ kg}^{-1}$ ), except in P3, A and in general decreased with depth, even in the Haploargid profile, which indicated that the CEC was controlled mainly by the presence of organic matter.

The XRF elemental analyses included 10 elements (Si, Al, Ca, Fe, Mg, K, Na, Ti, P and Mn).  $\text{SiO}_2$  was the most abundant oxide, with values varying between 483.0-528.0  $\text{g kg}^{-1}$ , followed by  $\text{Al}_2\text{O}_3$ , with values between 128.0 and 159.0  $\text{g kg}^{-1}$ , as well as Ca, Fe and Mg in decreasing order (Table 2).

The amounts of Al were higher than those that were reported for other soils in the country (Broomandi et al., 2017), which were found to be depleted in this element. The relatively high amounts of Ca and Mg were most likely a result of the presence of carbonates in all horizons of these soils.

The low coefficient of variation of the contents of major elements (Si and Al), which are the building blocks of silicate minerals, suggests that all these soils, despite belonging to different soil orders, had a very homogeneous silicate composition, and that variations with depth should not be

expected to be pronounced.

Higher variability was observed for Ca and Mg composition, which was most likely a consequence of carbonate variations among soil profiles and horizons. A relatively high Na variability (Table 2) was also observed, which was attributed to the differences in soil sodicity shown in Table 1.

The FTIR spectra of the 19 reference samples after baseline and EMSC corrections are presented in Figure 1. In this figure, the spectra are shifted along the y-axis to avoid overlapping for illustrative purposes. Compared to other spectral regions, such as Vis-NIR, mid-infrared is considered better for the identification of inorganic and organic functional groups in soils, and therefore of soil constituents, due to their distinctive peaks (Hunt & Salisbury, 1970).

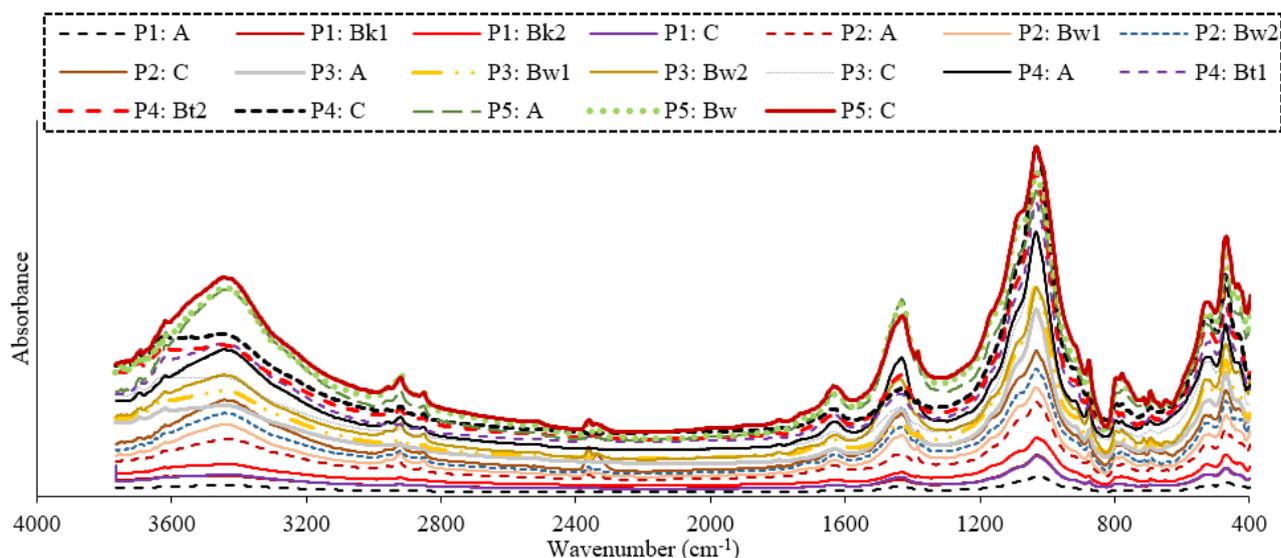
The spectra showed the characteristic bands of carbonates, as expected from the values for the general soil properties: a broad band around 1400-1500  $\text{cm}^{-1}$  with a peak at approximately 1440  $\text{cm}^{-1}$ , and two smaller bands at 875 and 712  $\text{cm}^{-1}$  (Müller et al., 2014). This last band indicated the presence of calcite instead of dolomite.

The bands between 3600-4000 and 1000-1200  $\text{cm}^{-1}$ , were mainly caused by silicates. The peak at 1030  $\text{cm}^{-1}$  corresponded to the Si-O-Si stretch vibration typical of these minerals (Madejová, 2003). However, the identification of silicates in a complex mixture, such as soil, is not straightforward, since characteristic peaks of different minerals often overlap.

Nevertheless, the presence of peaks at 3692 and 3627  $\text{cm}^{-1}$  are characteristic of 1:1 clays, which together with the absence of the typical combination of kaolinite peaks in the 900-1100  $\text{cm}^{-1}$  region, indicated that the dominant clay in these samples was likely halloysite (Joussein et al., 2005).

**Table 2.** Maximums, minimums, means, standard deviations and coefficients of variation (CV) of the elemental compositions measured by XRF

	$\text{SiO}_2$	$\text{Al}_2\text{O}_3$	CaO	$\text{Fe}_2\text{O}_3$	MgO	$\text{K}_2\text{O}$	$\text{Na}_2\text{O}$	$\text{TiO}_2$	$\text{P}_2\text{O}_5$	MnO
Mean	506.8	140.7	86.7	64.3	27.7	26.5	15.1	8.2	2.9	1.2
Standard deviation	11.8	9.1	21.8	5.7	4.0	1.5	3.6	0.7	1.5	0.2
Minimum	483.0	128.0	54.0	57.0	22.0	25.0	11.0	7.1	1.8	0.6
Maximum	528.0	159.0	119.0	78.0	38.0	30.0	25.0	9.3	7.8	1.5
C.V. %	2.3	6.5	25.2	8.9	14.5	5.5	24.2	8.7	52.0	16.6



**Figure 1.** Mid-infrared spectra of the analysed soil samples

The band at  $3682\text{ cm}^{-1}$  was also characteristic of smectites, which are typical in arid and semi-arid soils, and thus very likely to be present in these samples. In addition, bands at  $2850$  and  $2920\text{ cm}^{-1}$  were characteristic of C-H aliphatic chain vibrations, which are well known to be related to organic matter in the soil (Calderón et al., 2011).

The interpretation of bands below  $900\text{ cm}^{-1}$  is complicated, especially in mixtures, such as soils, since the absorbances of many organic and inorganic compounds overlap in this region (Calderón et al., 2011). However, some of the peaks described in the literature, such as those at  $800$  and  $780\text{ cm}^{-1}$ , have been identified as symmetric stretching vibrations of Si-O bonds, typical of silica and quartz (Müller et al., 2014).

Using the spectra from Figure 2 as independent variables, RF models were adjusted to the major oxides described in Table 2. The  $R^2$ , MSE and RMSE for each oxide are presented in Table 3. Only  $\text{Al}_2\text{O}_3$  and CaO were accurately estimated, with  $R^2$  values of 0.8 and MSE of 0.2 and 1.1, respectively.  $\text{TiO}_2$  was predicted with moderate accuracy, but all of the remaining oxides, including  $\text{MgO}$ ,  $\text{SiO}_2$ ,  $\text{P}_2\text{O}_5$ ,  $\text{MnO}$ ,  $\text{Na}_2\text{O}$ ,  $\text{K}_2\text{O}$  and  $\text{Fe}_2\text{O}_3$ , were poorly predicted.

These results could be considered less satisfactory than those that were found by other authors, like Mohanty et al.

**Table 3.** Results of the validation of the RF model calculated for each oxide

Element	$R^2$	$\text{MSE}_{\text{OoB}}$	RMSE
$\text{Al}_2\text{O}_3$	0.8	0.2	0.4
CaO	0.8	1.1	1.0
$\text{TiO}_2$	0.6	0	0
MgO	0.4	0.1	0.3
$\text{Fe}_2\text{O}_3$	0.3	0.2	0.5
$\text{Na}_2\text{O}$	0.2	0.1	0.3
$\text{P}_2\text{O}_5$	0	0.1	0.3
$\text{SiO}_2$	-0.1	1.4	1.2
MnO	-0.1	0	0
$\text{K}_2\text{O}$	-0.3	0.44	0.2

$R^2$  - Coefficient of determination;  $\text{MSE}_{\text{OoB}}$  - Mean square of error out-of-bag; RMSE - Root mean square error

(2016), who had an  $R^2$  of 0.9 when modelling oxide contents in soils from different regions of India using mid-infrared data and PLSR.

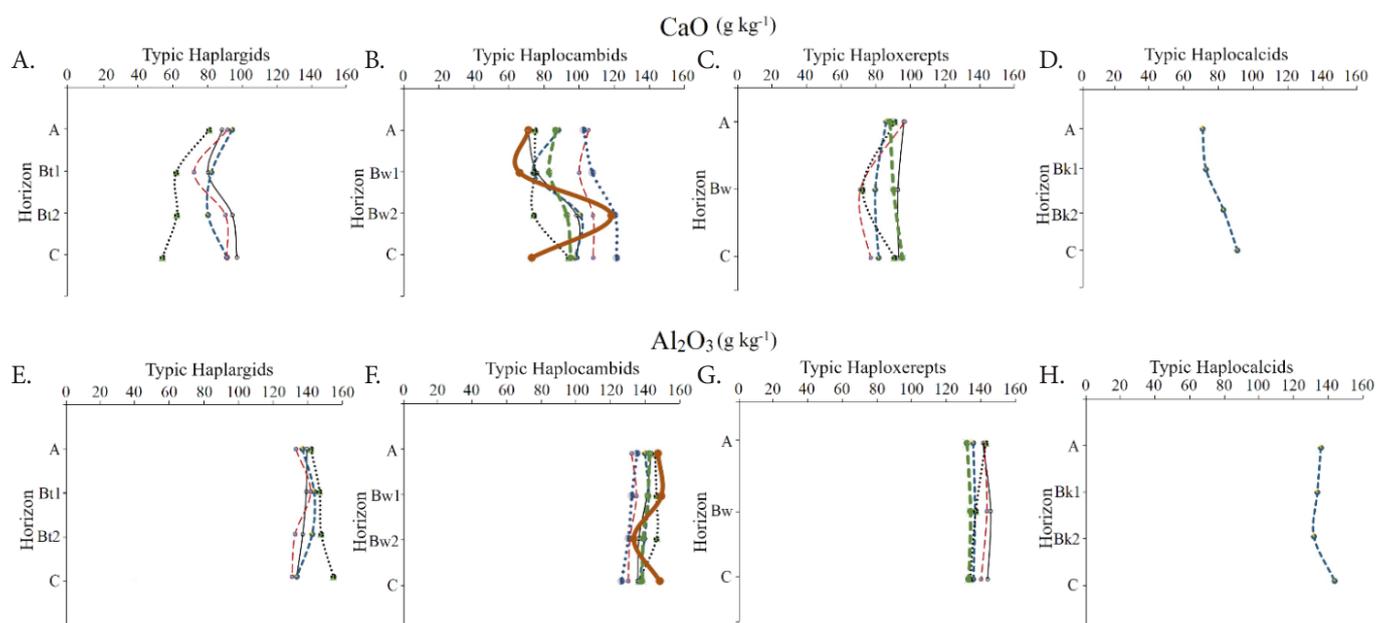
Possible reasons for the weaker performance of the models in this study include the limited amount of samples compared to those of Mohanty et al. (2016). However, PLSR is a statistical method that requires a large number of samples for modelling, and thus its use was not possible for our dataset. Another reason was the narrow range of values for most oxides, which is well known to be a limitation for statistical models. In any case, RF was probably the best option for modelling element contents despite its limited performance.

The oxides contents that were satisfactorily modelled ( $\text{Al}_2\text{O}_3$  and CaO) were predicted using the RF models in the remaining 48 soil samples. The results for the 17 sampled profiles, both measured and predicted, are presented in Figure 2.

Weathering processes in the soils lead to the depletion of elements that are considered mobile, such as Si, Ca or Mg, and an enrichment of those that are not mobile, such as Al, Ti, or Fe. Thus, the comparison of the behaviour of these elements can be used as an indication of the weathering and leaching of the soil (Bahlburg & Dobrzinski, 2011).

The CaO content gradually increased with depth in the typical Haplocalcids (Figure 2D), as expected from the increase of carbonate content and formation of calcic horizons. The formation of these horizons is known to be caused by the release of Ca from the primary minerals and subsequent leaching before precipitating as carbonates. Yousefifard et al. (2012) found that in a semi-arid region of Iran, the percentage of Ca released from primary minerals that precipitates as secondary calcium carbonate to form calcic horizons can reach up to 70-90%. The values of CaO also increased in some profiles of the typical Haplocambids (Figure 2B); although in this case, the migration of CaO was not marked enough to develop calcic horizons yet.

On the contrary, there was no obvious trend in CaO distribution for some profiles of the typical Haploxerepts (Figure



**Figure 2.** CaO and  $\text{Al}_2\text{O}_3$  contents in the 17 sampled profiles, including the measured 5 reference profiles and the 12 predicted by the Random Forest models

2C) and typical Haloargids (Figure 2A). In the typical Haloargids, the most prominent feature is the decrease of CaO from the A horizon to the Bt horizon. In general, CaO values are positively correlated with soil leaching (Wang et al., 2002), and thus it seems that leaching is a more important process in Haplocalcids and Haplocambids, while it is not so relevant in Haploargids and Haploxererts.

In a weathering environment controlled by leaching, a consistent decrease of Al with depth could also be expected. Figure 2E, F, G and H shows that, for each profile, the differences in Al composition among the horizons are not large, and a consistent decrease of  $Al_2O_3$  with depth was only observed in the Haplocambids. The results suggest that these soils were at an early stage of development and weathering was not intense enough to produce significant Al variations with depth.

In fact, in the Haploxererts, and especially in the Haplargids, a decrease in the Al content was observed with the transition from the A horizon to the Bw (Haplocambids) or Bt1 (Haplargids) horizons. This decrease was most likely caused by an increase in clay content with depth, since  $Al_2O_3$  contents in phyllosilicate clay minerals are higher than those that have been found in primary aluminosilicates such as feldspars (Muhs, 2001).

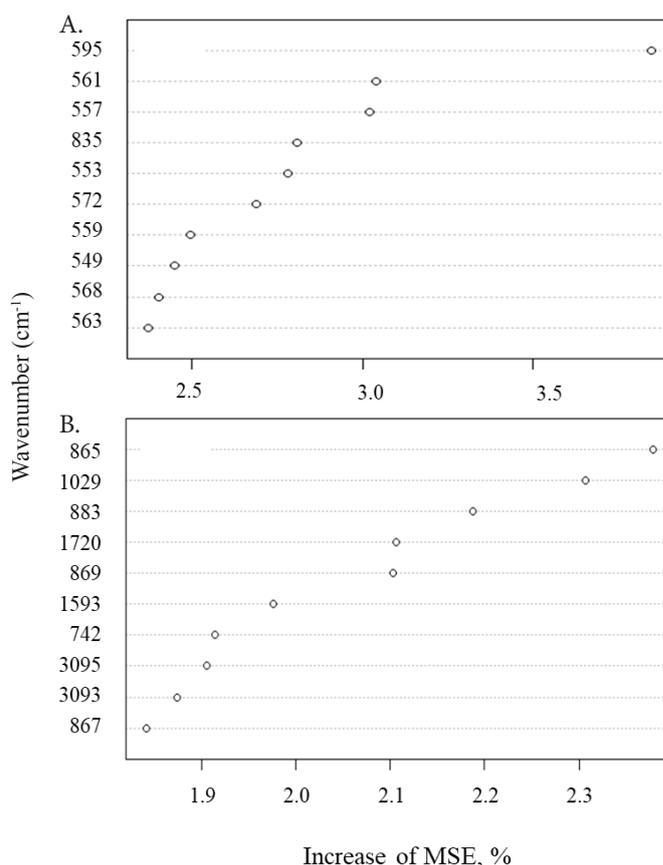
In summary, and based on the distribution of elements in the soil profiles, soils in Miandoab County are young, and some leaching was taking place in the profiles, mobilizing the most soluble elements (Ca), but they were an early stage of development and did not have obvious differences in less mobile elements (Al).

In addition to the goodness of fit parameters, RF provided the most relevant wavenumbers for the prediction of each element. Those wavenumbers are presented in Figure 3 for those indices that showed a good fit when modelling  $Al_2O_3$ , and CaO. The most relevant bands for the prediction of  $Al_2O_3$  were  $595\text{ cm}^{-1}$ , followed by  $561$  and  $557\text{ cm}^{-1}$ . In fact, 9 of the 10 most important bands are in the region between  $500$  and  $650\text{ cm}^{-1}$ , which have been ascribed to a combination of out-of-plane OH bending vibrations and gibbsite and alumino-silicate Al-O vibrations (Ali & Padmanabhan, 2017).

In the case of the CaO (Figure 3B), the interpretation of the important spectral bands was more complex. The 10 most important ones included those at  $865$ ,  $883$ ,  $869$  and  $867\text{ cm}^{-1}$ , which were part of a broad peak with a maximum at  $875\text{ cm}^{-1}$ , characteristic of carbonates. This band has been found useful for carbonates quantification by other authors (Tatzber et al., 2007).

The bands at  $3093$  and  $3095\text{ cm}^{-1}$  could also be related to carbonates have also been ascribed to carbonates, but their meaning is uncertain. The bands at  $1029$  and  $1593\text{ cm}^{-1}$  were related to silicates, and this could be derived from the presence of Ca in the exchangeable complex adsorbed to the clay surfaces.

Thus, for Al and Ca, the combination of FTIR spectra and RF not only allowed the creation of models that can be used to predict these elements contents in other soil samples of the area by simply measuring their FTIR spectra, but it also allowed the



MSE - Mean square error

**Figure 3.** Important bands in the MIR region for modelling A)  $Al_2O_3$ , and B) CaO contents

relation of the amounts and variability of these two elements with the soil components where they are more abundant: carbonates in the case of Ca, and silicates in the case of Al.

## CONCLUSIONS

1. The Fourier transform infrared spectra reflected the composition of these soils, with distinctive signatures, including the presence of calcite and low activity clays, mainly halloysite.
2. The combination of Fourier transform infrared spectra and Random Forest models provided adequate predictions of two of the analysed oxides,  $Al_2O_3$  and CaO. The important spectral bands for the prediction of these oxides were related to sounding functional groups and soil components.
3. The interpretation of the measured and predicted profiles indicated that the soils in the Miandoab County, West Azerbaijan, Iran, are at an early stage of development. In each profile, the elemental composition was very similar for all horizons, and a consistent enrichment of CaO with depth was only observed in the Typic Haplocalcid.

## ACKNOWLEDGMENTS

The XRF and FTIR spectroscopy analyses and part of the soil analysis were performed during a sabbatical leave by the first author at the University of A Coruna. The authors are deeply grateful to the faculty and staff members of the University of A Coruna for all the help and assistance. Financial

support for this work was provided by Tabriz University and University of A Coruna.

### LITERATURE CITED

- Ali, A. M.; Padmanabhan, E. The impact of pH and temperature on gibbsite reactivity with quartz. *International Journal of Petrochemistry and Research*, v.1, p.40-45, 2017. <https://doi.org/10.18689/ijpr-1000108>
- Bahlburg, H.; Dobrzinski, N. A review of the Chemical Index of Alteration (CIA) and its application to the study of Neoproterozoic glacial deposits and climate transitions the geological record of Neoproterozoic glaciations. *Geological Society*, v.36, p.81-92, 2011.
- Bower, C. A.; Reitemeier, R. F.; Fireman, M. Exchangeable cation analysis of saline and alkali soils. *Soil Science*, v.73, p.251-262, 1952. <https://doi.org/10.1097/00010694-195204000-00001>
- Broomandi, P.; Dabir, B.; Bonakdarpour, B.; Rashidi, Y. Identification of dust storm origin in South-West of Iran. *Journal of Environmental Health Science and Engineering*, v.15, p.1-14, 2017. <https://doi.org/10.1186/s40201-017-0280-4>
- Calderón, F. J.; Reeves, J. B.; Collins, H. P.; Paul, E. A. Chemical differences in soil organic matter fractions determined by diffuse-reflectance mid-infrared spectroscopy. *Soil Science Society of America Journal*, v.75, p.568-579, 2011. <https://doi.org/10.2136/sssaj2009.0375>
- Gee, G. W.; Bauder, J. W. Particle-size analysis. In: Klute, A. (ed.). *Methods of soil analysis - Part 1: Physical and mineralogical methods*. Madison: Soil Science Society of America, 1986. Cap.16, p.383-411.
- Hunt, G. R.; Salisbury, J. W. Visible and near infrared spectra of minerals and rocks: I - Silicate minerals. *Modern Geology*, v.1, p.283-300, 1970.
- Jeleńska, M.; Hasso-Agopsowicz, A.; Kadzialko-Hofmokr, M.; Sukhorada, A.; Bondar, K.; Matviushina, Z. H. Magnetic iron oxides occurring in chernozem soil from Ukraine and Poland as indicators of pedogenic processes. *Studia Geophysica et Geodaetica*, v.52, p.255-270, 2008. <https://doi.org/10.1007/s11200-008-0017-z>
- Joussein, E.; Petit, S.; Churchman, J.; Theng, B.; Righi, D.; Delvaux, B. Halloysite clay minerals: A review. *Clay Minerals*, v.40, p.383-426, 2005. <https://doi.org/10.1180/0009855054040180>
- Kohler, A.; Kirschner, C.; Oust, A.; Martens, H. Extended multiplicative signal correction as a tool for separation and characterization of physical and chemical information in Fourier transform infrared microscopy images of cryo-sections of beef loin. *Journal of Applied Spectroscopy*, v.59, p.707-716, 2005. <https://doi.org/10.1366/0003702054280649>
- Liaw, A.; Wiener, M. Classification and regression by random Forest. *R News*, v.2, p.18-22, 2002.
- Liland, K. H. EMSC: Extended Multiplicative Signal Correction - R package version 0.9.0. 2017. Available on: <<https://CRAN.R-project.org/package=EMSC>>. Accessed on: Dec. 2018.
- Madejová, J. FTIR techniques in clay mineral studies. *Vibrational Spectroscopy*, v.31, p.1-10, 2003. [https://doi.org/10.1016/S0924-2031\(02\)00065-6](https://doi.org/10.1016/S0924-2031(02)00065-6)
- Mohanty, B.; Gupta, A.; Das, B. S. Estimation of weathering indices using spectral reflectance over visible to mid-infrared region. *Geoderma*, v.266, p.111-119, 2016. <https://doi.org/10.1016/j.geoderma.2015.11.030>
- Moradi, N.; Sadaghiani, M. H. R.; Sepehr, E.; Mandoulakani, B. A. Effects of low-molecular-weight organic acids on phosphorus sorption characteristics in some calcareous soils. *Turkish Journal of Agriculture Forestry*, v.36, p.459-468, 2012.
- Muhs, D. R. Evolution of soils on quaternary reef terraces of Barbados, West Indies. *Quaternary Research*, v.56, p.66-78, 2001. <https://doi.org/10.1006/qres.2001.2237>
- Müller, C. M.; Pejčić, B.; Esteban, L. el; Piane, C. D.; Raven, M.; Mizaikoff, B. Infrared attenuated total reflectance spectroscopy: An innovative strategy for analyzing mineral components in energy relevant systems. *Scientific Reports*, v.4, p.1-11, 2014.
- Nelson, D. W.; Sommers, L. E. J. M. Total carbon, organic carbon, and organic matter. In: Sparks, D. L.; Page, A. L.; Helmke, P. A.; Loeppert, R. H. (eds.). *Methods of soil analysis - Part 3: Chemical methods*. Madison: Soil Science Society of America, 1996. Cap.33, p.961-1010.
- Osat, M.; Heidari, A.; Eghbal, M. K.; Mahmoodi, S. Impacts of topographic attributes on soil taxonomic classes and weathering indices in a hilly landscape in Northern Iran. *Geoderma*, v.281, p.90-101, 2016. <https://doi.org/10.1016/j.geoderma.2016.06.020>
- R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2017. Available on: <<https://www.R-project.org/>>. Accessed on: Dec. 2018.
- Rial, M.; Cortizas, A. M.; Rodríguez-Lado, L. Mapping soil organic carbon content using spectroscopic and environmental data: A case study in acidic soils from NW Spain. *Science of the Total Environment*, v.539, p.26-35, 2016. <https://doi.org/10.1016/j.scitotenv.2015.08.088>
- Robertson, A. H. J.; Shand, C.; Pérez-Fernández, E. The application of Fourier transform infrared, near infrared and X-ray fluorescence spectroscopy to soil analysis. *Spectroscopy Europe*, v.28, p.9-13, 2016.
- Saraskanroud, S. A.; Zeinali, B.; Mohammadnejad, V. Analysis physical and chemical properties of soil and morphometric impacts on gully erosion. *Desert*, v.22, p.157-166, 2017.
- Soil Survey Staff. *Keys to soil taxonomy*. Washington: United States Department of Agriculture, 2014. 644p.
- Tatzber, M.; Stemmer, M.; Spiegel, H.; Katzberger, C.; Haberhauer, G.; Gerzabek, M. H. An alternative method to measure carbonate in soils by FT-IR spectroscopy. *Environmental Chemistry Letters*, v.5, p.9-12, 2007. <https://doi.org/10.1007/s10311-006-0079-5>
- Wang, W.; Li, R.; Tan, J.; Luo, K.; Yang, L.; Li, H.; Li, Y. Adsorption and leaching of fluoride in soils of China. *Fluoride*, v.35, p.122-129, 2002.
- Wiesmeier, M.; Barthold, F.; Blank, B.; Kögel-Knabner, I. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. *Plant and Soil*, v.340, p.7-24, 2011. <https://doi.org/10.1007/s11104-010-0425-z>
- Youseifard, M.; Ayoubi, S.; Jalalian, A.; Khademi, H.; Makkizadeh, M. A. Mass balance of major elements in relation to weathering in soils developed on igneous rocks in a semiarid region, Northwestern Iran. *Journal of Mountain Science*, v.9, p.41-58, 2012. <https://doi.org/10.1007/s11629-012-2208-x>
- R. Bras. Eng. Agric. Ambiental, v.23, n.6, p.460-466, 2019.