DOI: https://doi.org/10.29298/rmcf.v9i47.158

Article

Organic carbon and total nitrogen in Mexican forest soils through VIS-NIR spectroscopy

Berenice Esquivel-Valenzuela[1], José Antonio Cueto-Wong[2*], Carlos Omar Cruz-Gaistardo[3], Armando Guerrero-Peña[4], Aarón Jarquín-Sánchez[5], David Burgos-Córdova[2]

[1]Unidad Regional Universitaria de Zonas Áridas, Universidad Autónoma Chapingo. México.

[2]Campo Experimental La Laguna, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. México.

[3]Instituto Nacional de Estadística y Geografía. México.

[4]Colegio de Postgraduados Campus Tabasco. México.

[5]Colegio de la Frontera Sur Unidad Villahermosa. México.

*Autor por correspondencia; correo-e: cueto.jose@inifap.gob.mx.

Abstract:

The sustainable use of the soil resource, as well as the correct decision making regarding its management depends on the knowledge of its properties. Conventional soil analysis methods are laborious, costly and generate large amounts of chemical waste. Due to the need for methods that allow a rapid, reliable analysis of soil properties, the objective of this work was to develop a prediction model for the content of organic carbon (OC) and total nitrogen (TN) in the soil through Visible and Near Infrared regions (VIS-NIR) spectroscopy. The conventional analysis of TC and TN of 599 forest soils from different regions of Mexico was carried out by dry combustion, and the content of inorganic carbon (IC) was determined by Bernard's calcimeter method. The prediction models were developed in a FOSS NIR System 6500. The models generated in the calibration process presented R^2 values of 0.93 and 0.88 for OC and TN, respectively. The values of the relationship between the standard error of prediction and the standard deviation of the samples (RPD, for its acronym in English) for both properties were higher than 2. In the case of the validation process, the values of R^2 were higher than 0.9, and the RPD was also above 2 in both properties. The results of this study show that VIS-NIR spectroscopy is an alternative technique to the conventional analysis methods of organic carbon and total nitrogen of the soil.

Key words: Absorbance, spectral analysis, prediction models, chemometrics, reflectance, forest soils.

Fecha de recepción/Reception date: 13 de diciembre de 2017

Fecha de aceptación/Acceptance date: 15 de abril de 201

Introduction

The sustainability of the agricultural and forest production, extensive livestock breeding and the environment in general is broadly related to the management and conservation of the soil. This sustainability can only be reached based on the scientific and technical knowledge of this resource (FAO, 2006). The soil is crucial for life, among other reasons, because of the role that it plays in the cycle of certain elements (Maestre *et al.*, 2008). Knowledge of its physical, chemical and biological characteristics is considered essential for better decision-making regarding its management.

The properties of the soil are generally determined through laboratory methods, with wet or dry chemistry techniques, which in most cases are laborious, time-consuming and costly (Ge *et al.*, 2011). Besides, they generate chemical waste which, when inadequately managed, can cause environmental pollution (Zornoza *et al*., 2008). Furthermore, certain chemical characteristics of the soil have a very dynamic cycle and a great spatial variability. This makes the obtainment of reliable information more difficult and costly, and therefore large amounts of samples must be analyzed in order to attain good knowledge of the behavior of these properties (Plant, 2001).

For this reason, there is a worldwide need to develop rapid, inexpensive, accurate and reliable methods for analyzing the edaphic properties (Shepherd and Walsh, 2007).

Many applications of infrared spectroscopy have been developed in the last few decades in both the agricultural and environmental sciences because of their high sensitivity in the detection of organic and inorganic soil components. This is based on the relationship of existing between intense fundamental molecular vibrations with the soil components in the mid-infrared region (2 500 to 25 000 nm). At the same time, in the near-infrared region (700 to 2 500 nm), overtones and combinations of these fundamental vibrations are generated, due to the lengthening and flexion of bonds N-H, C-H and O-H, as well as electronic transitions in the visible region (400 to 700 nm) of the electromagnetic spectrum (Viscarra *et al*., 2006).

Particularly, near-infrared spectroscopy (NIR, for its acronym in English) is a method that allows indirect quantification of certain properties of the soil based on the interaction of matter with an incidental beam of light in which a portion of photons is absorbed (absorbance) and the rest is reflected (reflectance) (Pérez *et al.*, 2014). Radiation absorption has also been registered in the infrared region by several functional groups, such as C-H, N-H, S-H, C=O and O-H of molecules of the sample, present mainly in organic molecules (Viscarra *et al.*, 2006).

VIS-NIR technology has many advantages compared to conventional analyses: it is quick, effective, non-destructive, low-cost, it requires a minimum analysis time per sample, it is simple and can be an ideal complement of the classic methods, or even replace them, once robust calibrations have been developed (Terhoeven *et al*., 2008; Xie *et al*., 2012). In addition, it has the ability to predict various properties based on a single spectrum.

In the field of agronomy, NIR or VIS-NIR spectroscopy encompasses several fields. One is the determination of soil properties, such as the content of nitrogen (Jarquín *et al*., 2011), carbon and total nitrogen (Fuentes *et al.*, 2012), organic matter (Rodríguez *et al.*, 2015), percentage of sand, silt and clay (Macías *et al.*, 2015), basal respiration (Maestre *et al.*, 2008), ^{13}C (Fuentes *et al*., 2009) and soil classification (Bastidas and Carbonell, 2010). This technology is also used for the indirect detection of plant diseases (Pérez *et al*., 2014), as well as for the evaluation of the quality of certain fodders (Valenciaga and Oliveira, 2006).

In Mexico, the VIS-NIR technique has been little applied in forest soil analyses, to a large extent because of the difficulty of having access to a significant amount of samples in order to carry out the corresponding calibrations. This is particularly difficult in a country like Mexico, which is characterized by its great natural diversity of soil types, climates and vegetation types, and by the ways in which ecosystems have been managed through time. For example, according to INEGI (2007), 26 of the 32 groups recognized by the World Reference Base for Soil Resources of the International Union of Soil Sciences (IUSS, 2007) exist in the country. This requires the development of methods that allow a quick, accurate estimation of the spatial variability of the physical and chemical properties of large amounts of soil samples at low cost and with the least possible impact on the environment.

The purpose of this work was to create a model for predicting the content of organic carbon and total nitrogen in samples of forest soils of Mexico using spectroscopy in the VIS-NIR spectral region.

Materials and Methods

The present research was carried out between 2015 and 2016 at the soils laboratory of the *Campo Experimental La Laguna* (La Laguna Experimental Station) dependent on the *Centro de Investigación Regional Norte-Centro* (North-Central Regional Research Center) of the *Instituto de Investigaciones Forestales, Agrícolas y Pecuarias* (National Institute of Research on Forestry, Agriculture and Livestock), located in *Matamoros*, *Coahuila*, Mexico.

Samples

A total of 599 of forest soils from various regions of Mexico (Figure 1) were analyzed. The samples came from the 2014 National Inventory of Forest and Soils of the *Comisión Nacional Forestal* (National Commission of Forestry) (Conafor). The samples were dried at ambient temperature, ground and sifted using meshes with 2 and 0.5 mm pore openings, in order to obtain fine fractions (used in the conventional analyses) and coarse fractions (utilized to capture spectrums in VIS-NIR), respectively. The samples were subsequently stored at ambient temperature in hermetically closed plastic containers.

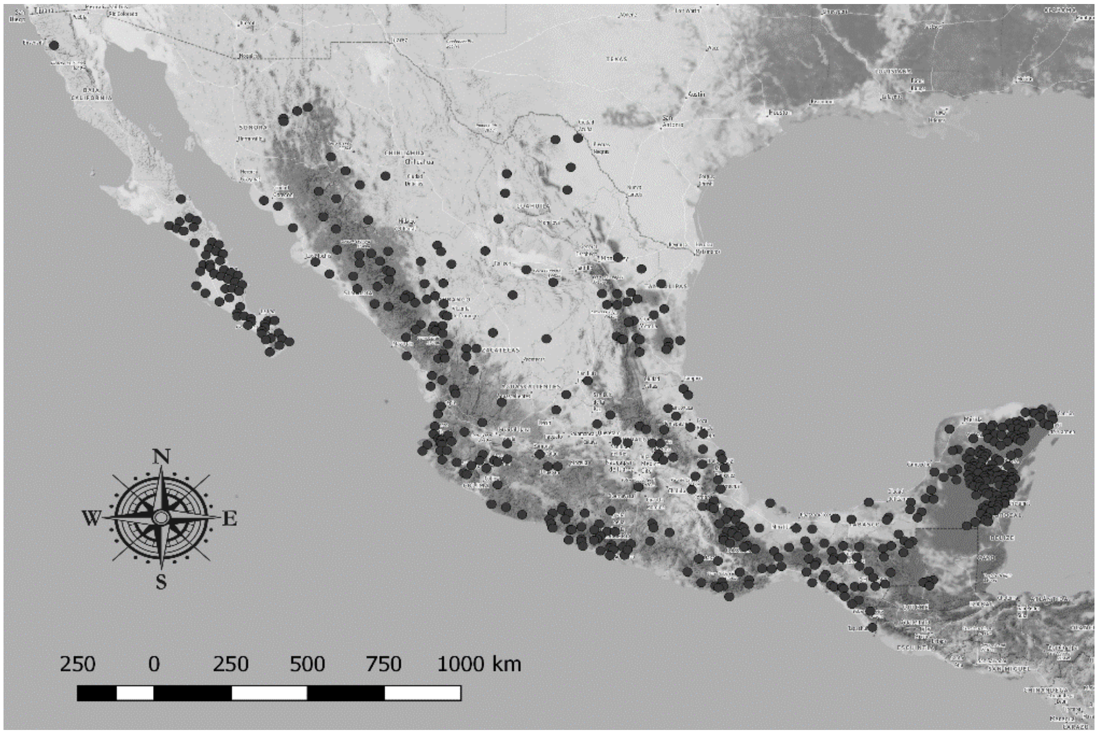

Figure 1. Location of sampling points.

Conventional analyses

The samples were analyzed using conventional methods, based on the coarse fraction of the soil. The values for total carbon (TC) and total nitrogen (TN) were estimated using the Dumas method in a (ThermoScientific) Total Flash 2000 Total Elemental Analyzer. Between 20 and 50 mg of soil previously dried at ambient temperature and sifted 150 µm were weighed (OAHUS PA224C analytical scale). The samples were calcined (Thermo Scientific Flash 2000 elemental analyzer) at 950 °C, using oxygen as oxidizing agent (AOAC, 2005). The content of inorganic carbon (IC) was estimated based on the determination of total carbonates, using the Bernard calcimeter method, which measures the volume of CO_2 emitted by the sample when reacting to the presence of HCl (Gaucher, 1971). The content of organic carbon

(OC) was determined by the difference between the total carbon minus the inorganic carbon obtained from the total carbonates analysis.

Development of the NIR model

This consisted of four basic stages: analysis of all the samples using conventional methods, obtainment of the spectrum of each sample, calibration of the regression equations and validation of the model (Macías *et al.,* 2015). The VIS-NIR model was developed using a (FOSS) NIR 6500 Feed and Forage analyzer. Previously to the capture of the spectrums, the performance tests were run in order to ensure that the equipment was working correctly. The average measured reflectance was 32 scans per simple in a wave-length interval of 400 to 2 500 nm (visible region and near-infrared region), with 2 nm between collected data points, for a total of 1 050 spectral points.

Calibration. 448 samples selected through a discriminative analysis based on Mahalanobis distances (H) were utilized, which allow reducing the number of spectral bands by means of the principal component analysis (PCA). Variables representing a large proportion of the variability of the initial bands are thereby obtained, and the extreme values are those with a distance of more than 3. This procedure is helpful for identifying and eliminating biased data (outliers), which, when incorporated to the model, reduce its reliability (Pell, 2000). The analyses mentioned above have been widely used for managing soil spectral responses (Guerrero *et al*., 2010; Gogé *et al.,* 2012).

The calibration equation was obtained using the winISI v4.20 software (Infrasoft International, 2010), with modified partial least squares regression (MPLS) between the results of the conventional methods and spectral data generated in the region between 400 and 2 500 nm.

The following mathematical treatments were previously applied: SNV (Standard Normal Variate) and Detrend, for the correction of dispersion and particle size problems, and adjustments to the baseline and trend of the data, respectively, as well as a 2,4,4,1 array, where the first number indicates the derivative applied for improving the spectral solution, the second refers to the interval between the calculated derivatives, the third is the length of the segment that will be softened with the purpose of minimizing the spectral noise, and the fourth indicates that the second softened segment was not utilized.

Validation. The validation of the prediction model was carried out with 46 of the total samples received, which were not included in the calibration process. These were used to determine the accuracy and precision of the model developed for predicting the various parameters of interest, through a comparison between the predicted values and the values estimated using the traditional methods. The capture of spectrums was carried out with the ISIscan v3.1 software (Infrasoft International, 2010), with the same parameters that were utilized in the calibration process, while the chemometric operations were performed using the winISI v4.20 software (Infrasoft International, 2010).

The accuracy of the model was evaluated based on the determination coefficient (R^2) and the relationship between the standard prediction error and the standard deviation of the samples (RPD) (Minasny and McBratney, 2013). An R^2 value of 0.66 to 0.81 indicates a good model; an R^2 of 0.82 to 0.90 reveals a good prediction, while an R^2 of more than 0.91 is considered excellent (Williams, 2003). For the RPD values, this study utilized the criteria developed by Chang *et al*. (2001), who define three categories: A= RPD> 2.0: good; B= 1.4≤ RPD ≤ 2.0: acceptable, and C: RPD <1.4: unreliable.

Results

The values predicted by the developed VIS-NIR model for OC ranged between 0.02 % and 18.29 %, with a mean of 5.45 % (Table 1). The values for TN ranged between 0.002 % and 1.445 %, from very low (<0.05 %) to very high (>0.25 %), according to the NOM-021-SEMARNAT-2000, 73 % of the samples had a very high total nitrogen content, while only 3 % were classified as having a very low content.

Table 1. Descriptive statistics of the conventional analysis values and the values predicted with VIS-NIR for OC and TN.

	Organic Carbon (%)				**Total Nitrogen (%)**			
	Medium	**Min**	**Max**	**SD**	**Medium**	**Min**	**Max**	**DS**
Reference	5.42	0.23	21.02	3.90	0.48	0.005	2.004	0.38
VIS-NIR	5.45	0.02	18.29	3.48	0.46	0.002	1.445	0.29

SD = Standard deviation

Figure 2 shows the spectrums generated in the VIS-NIR region. In the VIS region, peaks can be observed between 450 and 600 nm, which are partly related to the OC content of the soil (Viscarra *et al.*, 2006). In the NIR interval, the spectrums had higher absorbance peaks, approximately at 1 400, 1 900 and 2 200 nm. The O-H and aliphatic C-H molecules are located in the 1400 nm band; the amide N-H and the O-H occur at 1 900 nm, while the 2 200 nm band is generally associated with phenolic O-H, amine N-H, and the aliphatic C-H group (Cozzolino and Morón, 2003).

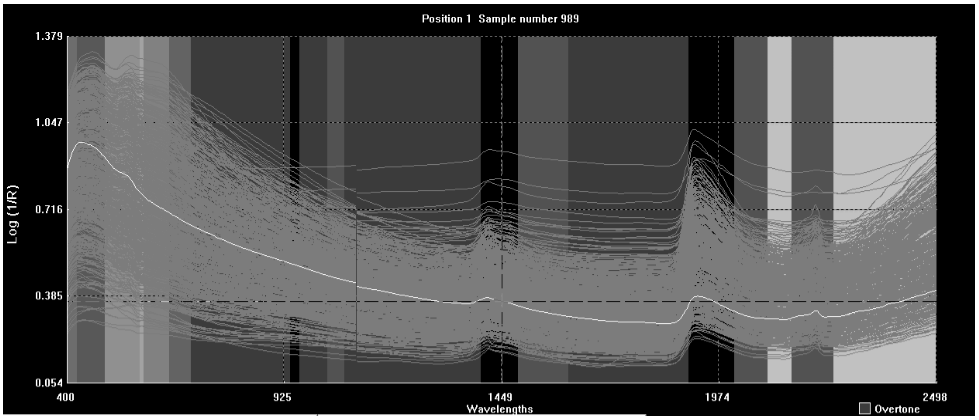

Figure 2. Spectrums of the analyzed soils.

The R^2 values obtained for the calibration models were 0.93 for OC and 0.88 for TN (Figure 3). In the case of the RPD, they were 2.69 for OC and 2.05 for TN; the models are therefore rated as having a good predictive capacity, within category A, as defined by Chang *et al.* (2001).

In the validation, the R^2 values for OC and TN were 0.92 and 0.91, respectively, and those of RPD were above 2 for both.

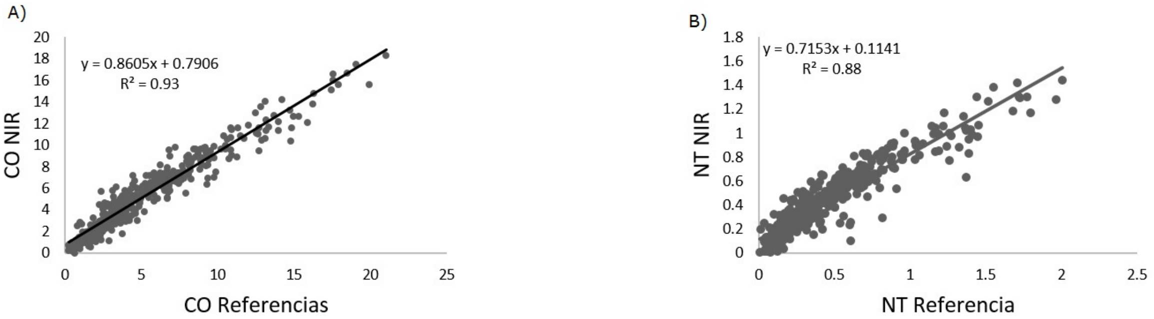

Figure 3. Relationships between the results of the conventional methods (reference) and the results predicted with the VIS-NIR model in the calibration process of parameters A) organic carbon and B) total nitrogen.

Discussion

Today there are 3 999 million hectares of forests in the world (30.6 %) (MacDicken *et al.*, 2016), whose carbon (C) is estimated in 861 Pg (1 Pg=1 × 1 015 g), 383 Pg of which are in the soil (44 %) (Pan *et al.*, 2011). Moreover, edaphic N gives fertility to the soil, allowing the development of forest areas which, in turn, provide various environmental services to society (Ruiz *et al.,* 2007).

While the development of NIR predictive models for OC have focused in forest areas, due to the importance of the carbon sequestration process, several studies on the TN content have been performed in agricultural areas (Wetterlind *et al*., 2008; Zhang *et al*., 2016), due to the importance of this element in the growth and development of the crops, with the purpose of developing optimal fertilization methodologies.

The models generated in this study exhibited a high level of correlation and a good predictive capacity. In the case of OC, the most robust had an R^2 of more than 0.9 and an RPD above 2; these figures coincide with those obtained before (McCarty *et al*., 2002; Sarkhot *et al.*, 2011; Kodaira and Shibusawa, 2013). The average R^2 of various OC models with NIR, according to a study by Viscarra *et al.* (2006), was 0.81.

The models for TN have also exhibited good adjustments (Reeves and McCarty, 2001). The R^2 intervals for TN ranged between 0.68 and 0.98 (Nduwamungu *et al*., 2009).

Figure 3 shows that the OC, which expressed a better fit than TN, has less data dispersion. This may be due to the low values of total soil nitrogen, for, although in the NIR region the radiation is absorbed by chemical bonds such as C-H, N-H, S-H, C = O and O-H of any chemical bond occurring in the sample, it is absorbed in proportion to the concentration of these compounds (Zornoza *et al.,* 2008). Furthermore, as shown in Table 1, the TN values are very low, compared to those of OC, with a C/N ratio of 10:1 in the values of the mean, the maximum reference values and the values predicted by VIS-NIR.

Although the model developed for nitrogen is considered to have good predictive capacity, the statistical metrics can be improved by extending the number of samples, as well as by verifying the assumption of normality of the data, which, though not an indispensable requirement, improves the predictive capacity of the model (Diggle and Ribeiro, 2000).

The spectral bands agree with the data reported by other authors (Fidencio *et al*., 2002; Zhang *et al.,* 2016). The VIS-NIR regions exhibited absorption of O-H (1 400 and 1 900) and C-H (2 200 nm) bonds; both overtones are related to the organic matter (Salgó *et al.*, 1998) and associated with the micro- and macro-elements occurring in the soil fraction in the form of various compounds (Cozzolino and Morón, 2003).

The independent validation group also presented a good correlation with R^2 values above 0.9 and RPD values above 2 for both OC and TN. This suggests that the models developed have a good predictive capacity (Chang *et al.*, 2001) for the analyzed properties.

The soil OC and TN concentrations are properties that possess a broad theoretical base in regard to predictive models (Barthès *et al*., 2010; Macías *et al*., 2015; Terra *et al*., 2015; Sisouane *et al*., 2017). However, once the calibration equations have been developed, they must be continually validated, with independent samples, but within the interval considered in the initial model. Subsequently, the validated samples can be added to the initial database, which allows the development of new models with a wider interval, and therefore with greater predictability. This deserves special attention in the case of Mexico, since, due to its great variety of soils, a large number of samples is required to design models that can be applied at the national level.

Conclusions

The R^2 and RPD values in both the calibration and the validation processes evidence the high predictive value of the VIS-NIR models for OC and TN concentrations in forest soils.

VIS-NIR spectroscopy has proved to be an alternative technique to the conventional soil analysis methods; it has the potential to carry out rapid and accurate predictions regarding the chemical properties of the soil.

Acknowledgements

The authors wish to express their gratitude to the *Comisión Nacional Forestal* (National Commission of Forestry) for the support provided in the collection of the samples, as well as to INIFAP for the facilities provided for the analysis of the samples.

Conflict of interest

The authors declare no conflict of interest.

Contribution by author

Berenice Esquivel Valenzuela: drafting if the abstract and the introduction, results, discussion and conclusions sections of the manuscript; José Antonio Cueto Wong: drafting of the materials and methods, results and discussion sections of the manuscript, general revision and editing; Carlos Omar Cruz Gaistardo: field work, handling and preparation of the samples; Armando Guerrero Peña: revision and suggestions for the drafting and editing of the abstract and the introduction and materials and methods sections of the manuscript; Aarón Jarquín Sánchez: revision and suggestions for the drafting and editing of the results, discussion and conclusions sections of the manuscript; David Burgos Córdova: laboratory work, support in the drafting of the materials and methods section.