Prediction of soil carbon and nitrogen contents using visible and near infrared diffuse reflectance spectroscopy in varying salt-affected soils in Sine Saloum (Senegal)

Aurélie Cambou, Bernard Barthès, Patricia Moulin, Laure Chauvin, El Hadji Faye, Dominique Masse, Tiphaine Chevallier, Lydie Chapuis-Lardy

Prediction of soil carbon and nitrogen contents using visible and near infrared diffuse reflectance spectroscopy in varying salt-affected soils in Sine Saloum (Senegal)

Aurélie Cambou, Bernard Barthès, Patricia Moulin, Laure Chauvin, El Hadji Faye, Dominique Masse, Tiphaine Chevallier, Lydie Chapuis-Lardy

Prediction of soil carbon and nitrogen contents using visible and near infrared diffuse reflectance spectroscopy in varying salt-affected soils in Sine Saloum (Senegal)

Aurélie CAMBOU[a,*], Bernard G. BARTHÈS[a], Patricia MOULIN[a,b], Laure CHAUVIN[c], El Hadji FAYE[d], Dominique MASSE[a,e], Tiphaine CHEVALLIER[a], Lydie CHAPUIS-LARDY[a,e]

[a] Eco&Sols, Université de Montpellier, CIRAD, INRAE, IRD, Montpellier SupAgro, 34060 Montpellier, France
[b] Imago, IRD, Dakar, Senegal
[c] Université Cheikh Anta Diop (UCAD), Dakar, Senegal
[d] Institut Supérieur de Formation Agricole et Rurale (ISFAR), Université de Thiès, BP 54 - Bambey, Senegal
[e] IRD, LMI IESOL, Dakar, Senegal

Email addresses of other authors: bernard.barthes@ird.fr; patricia.moulin@ird.fr; chauvinlaure@gmail.com; elhadji.faye@uadb.edu.sn; dominique.masse@ird.fr; tiphaine.chevallier@ird.fr; lydie.lardy@ird.fr
***Corresponding author:**
Aurélie CAMBOU
Email: aurelie.cambou5@gmail.com
Full postal address: UMR Eco&Sols, Montpellier SupAgro, 2 Place Viala, bât. 12, 34060 Montpellier Cedex 2, France

Type of paper: Full research paper

Abstract

Soil organic carbon (C) and nitrogen (N) contents have an essential role in soil fertility, but they may be affected by salinity, which is especially responsible for land degradation in arid and semiarid regions. The objective of this work was to study the ability of visible and near infrared diffuse reflectance spectroscopy (VNIRS) to predict soil C and N contents and electrical conductivity (EC, a proxy for soil salinity) in variably salt-affected topsoils of the Sine Saloum region (Senegal). Different calibration procedures and spectral pretreatments were compared, and variable log-transformation usefulness was evaluated for prediction optimization.
Predictions involved three calibration procedures: global partial least squares regression (PLSR), which used all calibration samples similarly; locally weighted (local) PLSR, with target samples predicted individually by giving higher weight to closest calibration spectra; and global PLSR per salinity class, after spectral discrimination of these classes. Predictions were performed with possible spectrum pretreatments (e.g., derivatization) and variable decimal log-transformation.
The study was performed on 311 topsoil samples (0-25 cm depth), either unsalted to slightly salty (Salt-, EC ≤ 2 mS cm^{-1}; 262 samples) or medium to highly salty (Salt+, EC > 2 mS cm^{-1}; 49 samples). Soil salinity was accurately discriminated using spectra: in validation, 100 % and 95 % of Salt- and Salt+ samples were correctly assigned on average, respectively. Best C and N content predictions were achieved after log-transformation using calibration by class (R^2_{VAL} = 0.87) and local calibration (R^2_{VAL} = 0.77), respectively; best EC prediction was achieved without log-transformation using global calibration (R^2_{VAL} = 0.90). This suggested C and N content predictions were affected by salinity; logC and logN distributions were almost symmetrical, hence log-transformation usefulness, while logEC distribution was very asymmetrical. No pretreatment yielded systematically good predictions; nevertheless, first-order derivative using 31-point gap often yielded good predictions, and second-order derivatives poor results.

Keywords
electrical conductivity; spectral pretreatment; partial least squares regression; locally weighted regression; discriminant analysis; semiarid West Africa

Highlights

- Soil C and N contents and salinity were accurately predicted using VNIRS (RPD > 2)
- C content was best predicted using PLSR by salinity class on log-transformed values
- N content was best predicted using locally weighted PLSR on log-transformed values
- Salinity was best predicted using global PLSR without log-transformation
- Spectrum pretreatment optimization depended on the variable and PLSR type

Graphical abstract

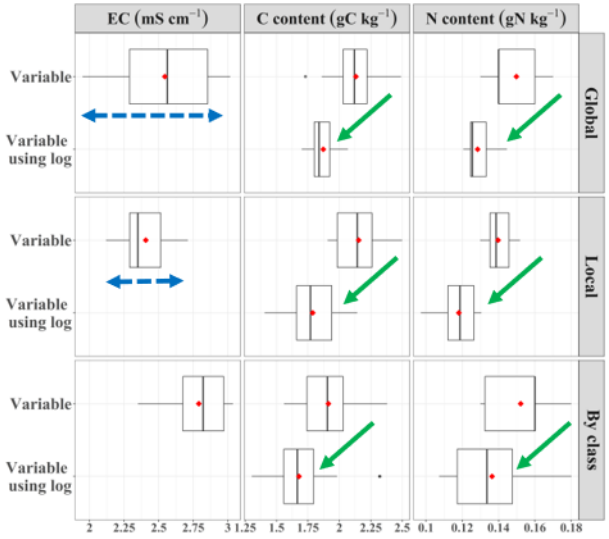
VNIRS prediction of C and N contents and salinity in variably salt-affected soils (Senegal)

Dataset: 311 variably salty topsoil samples from Sine Saloum area, Senegal
Two salinity classes: unsalted-slightly salty (EC ≤ 2 mS cm^{-1}; 262 samples) vs. medium-highly salty soils (EC > 2 mS cm^{-1}; 49 samples)
Spectral pretreatments: 46 VNIR spectrum types (e.g. derivatization)
Prediction: three partial least squares regression (PLSR) types: i. Global; ii. Locally weighted (local); iii. Global by salinity class discriminated using PLSDA

Prediction of
Soil electrical conductivity (EC)
Soil carbon (C) content using or not $\log_{10}$-transformation
Soil nitrogen (N) content using or not $\log_{10}$-transformation

Determination of the root mean square error of prediction (RMSEP) for each model on 62 validation samples
(the 249 calibration samples were selected using the Kennard-Stone algorithm)

Distribution of RMSEP for each prediction model of each variable over the 46 spectrum types (vertical lines inside boxes represent medians, red diamonds means)



The best model depended on the studied variable:
- **Global PLSR for EC**
- **PLSR by salinity class for C content using log**
- **Local PLSR for N content using log**

Spectrum type effect on RMSEP variation depended on model type & studied variable

Systematic gain of accuracy using $\log_{10}$-transformation of C and N contents

1. Introduction

Soil organic matter (SOM), made mainly of carbon (C; 58 % of the SOM) coupled to nitrogen (N), phosphorus and sulfur, plays an essential role in soil physical (e.g., soil aeration, aggregation), chemical (e.g., pH regulation and nutrient reserve) and biological fertility (e.g., mineralization and nutrient recycling by heterotrophic soil organisms; Lal, 2014). Optimizing these functions and processes allows soils to provide ecosystem services, such as food production, climate regulation or even water storage, regulation and supply (Dominati et al., 2010; Lal, 2014). Thus, maintaining SOM stocks is a major issue worldwide in the context of climate change and land degradation (Dignac et al., 2017; Gupta, 2019).

Soil salinization, which is the accumulation of water-soluble salts in the soil, can have natural or anthropic origins (primary or secondary salinization, respectively). The proportion of salt-affected soils, located mainly in arid and semi-arid regions, is estimated to be 7.5 % of the global land surface area (Hossain, 2019). Salinity affects soil properties, particularly C and N cycling. Indeed, a high level of salts reduces soil microbial biomass and activity and thus SOM decomposition (Pletz and Haynes, 2003; Yuan et al., 2007). Moreover, plant growth is limited in salt-affected soils: their biomass production is reduced, leading to a lower amount of fresh organic matter input into soil, often causing a SOM stock decrease (Wong et al., 2010). As in many countries in dry regions, Senegal is impacted by high primary salinization, principally due to climate change, resulting in locally so-called "Tannes" (highly degraded salty plots, literally tanned or burnt out; Diatta et al., 2001). The current surface area of Senegal salt-affected soils is higher than 1.7 million ha, and the Tanne surface area is expected to increase in the future at the expense of agricultural soils, mangroves and soils under natural vegetation (Faye et al., 2019; Sadio, 1991). In this context, to be able to prevent soil degradation, it is necessary to spatially and temporally assess soil salinization processes and estimate their long-term impact on SOM content.

Accurate mapping of spatial and temporal changes in SOM content and soil salinity levels at different scales requires the analysis of many samples. Visible and near infrared diffuse reflectance (VNIR) spectroscopy (VNIRS) has been proposed as an alternative to conventional analytical methods to assess soil properties: its cost- and time-effectiveness has been notably evidenced for SOM content determination (O'Rourke and Holden, 2011; Stenberg et al., 2010). VNIRS has also shown its usefulness for quantifying soil organic C and N contents and soil electrical conductivity (EC), which is used as a proxy for soil salinity (Clairotte et al., 2016; Stenberg et al., 2010; Viscarra Rossel et al., 2006). Several authors also observed a direct effect of salt content on the VNIR absorbance spectrum (Farifteh et al., 2008; Li et al., 2019; Wang et al., 2018).

Some challenges have, however, been identified in VNIRS applications to soils. In regions affected by variable levels of soil salinization, soil spectral libraries may be characterized by important spectrum diversity, possibly leading to less accurate prediction of SOM content (Moura-Bueno et al., 2019). To overcome the issue of soil diversity in spectral libraries, different types of models can be used. For example, Liu et al. (2018), who used a VNIR spectral library including different soil types, obtained the best predictions of soil organic C content when they first discriminated soil types from their VNIR spectra and then performed predictions by soil type. Another strategy consists of performing individual prediction of each target

sample, using only calibration samples that are its spectral neighbours; this is the principle of local partial least squares (PLS) regression (PLSR), which has shown good results for the prediction of soil organic C in large and heterogeneous spectral libraries (Clairotte et al., 2016; Nocita et al., 2014). Spectrum pretreatment (e.g., standardization or derivatization) has also been used to improve prediction accuracy, especially to reduce spectrum noise that disturbs the relationship between spectra and studied sample properties (Boysworth and Booksh, 2007; Stenberg et al., 2010). Few authors have specifically studied the effect of spectral pretreatments on prediction model accuracy. Liu et al. (2019), who predicted SOM content using VNIRS, observed an effect of spectral pretreatment (six pretreatments tested) on the calibration set selection based on spectral representativeness, which thus influenced SOM content prediction performance. Moura-Bueno et al. (2019) combined four types of prediction models with six spectral pretreatments and observed an effect of the combination model × pretreatment on soil organic C content prediction using VNIR spectra. Recently, other authors tested fractional derivatives and reported good VNIRS predictions with 1.5-order derivatives for salt content (Wang et al., 2018) or 0.2–0.8-order derivatives for SOM (Xu et al., 2020). However, most of these studies considered only a few spectral pretreatments and did not consider their possible combination (e.g., standardization then derivatization). In addition, prediction may be hampered for explained variables that do not follow a normal distribution, so several authors have used log-transformation (with either natural or decimal logarithm), root-square-transformation or Box Cox transformation to obtain an approximately normal distribution and thus improve prediction (Vasques et al., 2008; Liu and Chen, 2012; Terra et al., 2015; Lobsey et al., 2017). For instance, Vasques et al. (2008) and Terra et al. (2015) reported better VNIRS predictions of soil organic C with $\log_{10}$ transformation than without $\log_{10}$ transformation. The latter authors compared different variable transformations: the best results were achieved with $\log_{10}$ transformation for several soil properties, such as organic C or exchangeable bases, with square-root transformation for other properties, such as clay activity or most micronutrients, with Box Cox transformation for sand content only, and without transformation for clay content and most oxides.

Therefore, predicting soil properties from spectral libraries that include both unsalted and salty soils might be challenging, but appropriate combinations of model type and pretreatments could help to address this challenge. The objective of this work was to optimize VNIRS predictions of topsoil organic C and total N contents and EC in a spectral library including variably salty soils by identifying the best combinations of the type of prediction model, spectral pretreatment and variable transformation. More specifically, the study was conducted on soils of the Sine Saloum region (Senegal) and three types of PLSR (global, i.e., common; locally weighted; and by EC class) were combined with 46 spectral pretreatments (e.g., centring, standard normal variate, detrending, derivatization, and their possible associations) to test prediction accuracy. In addition, the log-transformation of explained variables was also tested.

2. Materials and methods

2.1. Studied region and soil sampling

The soil samples that were studied originate from the Senegalese administrative region of Fatick, which is 100 to 250 km east-southeast of Dakar and covers 6850 km². The climate is semiarid, with 400- to 600-mm annual rainfall and 28–29 °C mean annual temperature. The region includes the deltaic Sine Saloum estuary in the west (Quaternary sediments); gently rolling plains derived from an ancient dune field overlying a continental sedimentary basin (late Cretaceous) in the east; and residual dissected plateaus of the same continental sedimentary basin in the north (Roger et al., 2009; Tappan et al., 2004). The main soil types are Gleyic Solonchaks (soluble-salt rich and hydromorphic), Ferralic/Sideralic Arenosols (very sandy, strongly weathered, oxide-rich), Umbric Gleysols (waterlogged, dark-coloured, with low base saturation), Ferric Lixisols (with low-activity clays, high base status and oxide concretions), Stagnic Fluvisols (stratified sediments, long waterlogged), and, to a lesser extent, Haplic Arenosols (very sandy) and Dystric Regosols (poorly developed, with low base saturation; IUSS Working Group WRB, 2015). Natural vegetation is mainly tree savannah and shrub savannah but also mangroves. The main crops are peanuts and pearl millet, sometimes maize, and market gardening (on some hydromorphic soils). Commonly, livestock is more or less integrated into agricultural systems, especially cattle and small ruminants.

As part of the study of Chauvin (2013), the sampled sites were chosen to capture the regional variability of land covers and soil types, except mangroves and mudflats on Fluvisols (due to accessibility issues). The sampling design, which involved 312 sites, was based on a classification carried out using two 2010 Landsat 7 images (183 × 170 km² each) provided by the sensor Enhanced Thematic Mapper, which has eight bands from 0.45 to 12.50 µm with 30-m resolution in general (USGS, 2011). Images were analysed using ENVI 4.5 software (ITT Visual Information Solutions, Boulder, CO, USA) for geometric correction, mosaicking and colour composites and ArcGIS 9.3 software (ESRI, Redlands, CA, USA) for digitization and for creating a land cover map. All these sites had soils with texture dominated by sand-size particles. Overall, the site latitude ranged from 13°35′35″ to 14°41′33″ N, and the site longitude ranged from 16°38′04″ to 15°35′28″ W. One soil sample was collected at each site at a depth of 0–25 cm using a manual auger.

2.2. Soil analyses

Before analyses, the soil samples were air-dried and then crushed using a mortar and pestle before 2-mm sieving. All analyses were carried out in ISO9001:2015 certified laboratories of the French National Research Institute for Sustainable Development (IRD) in Dakar. Total C and N contents were determined on 0.2-mm ground, 100-mg aliquots according to ISO 10694:1995 and 13878:1998 procedures, respectively (ISO, 1995 and 1998, respectively) using a CHN elemental analyser (Thermo Finnigan Flash EA1112, Milan, Italy). Soil samples collected in the study area were expected to be carbonate-free; thus, all carbon was considered organic. However, one sample had a very high C-to-N ratio (28.9), suggesting that it might contain carbonates. This sample was considered an outlier and removed from the soil sample set, which thus included 311 samples. The EC was determined according to the ISO 11265:1994 procedure (ISO, 1994) on suspensions of 20 g of 2-mm sieved soil in 100 mL demineralized

water using an ohmmeter (SympHony SB70C, VWR, Mont-Royal, QC, Canada). Two soil salinity classes were distinguished according to Sadio (1991), who studied the Sine Saloum area and considered vegetation and soil texture in addition to EC: unsalted or slightly salty soils (denoted as Salt-), with EC ≤ 2 mS cm^{-1}; and medium to highly salty soils (Salt+), with EC > 2 mS cm^{-1}.

Reflectance spectra in the visible and near infrared regions were acquired between 350 and 2500 nm at 1 nm intervals using a portable LabSpec 4 spectrophotometer (Analytical Spectral Devices, i.e., ASD, Boulder, CO, USA). This instrument is equipped with a contact probe, and the samples are scanned manually with this probe (surface area scanned: 80 mm²). Each VNIR spectrum resulted from the averaging of 32 coadded scans, and absorbance zeroing was carried out every hour using a reference standard (Spectralon, i.e., polytetrafluoroethylene). Each sample spectrum resulted from the averaging of spectra acquired on three aliquots of 2-mm sieved air-dried samples that had been oven-dried at 40 °C for at least 12 h.

2.3. Chemometric analyses

2.3.1. Pretreatments

Before analyses, reflectance spectra were converted into absorbance, which was calculated as the decimal logarithm of the inverse of reflectance (absorbance = $\log_{10}$[1/reflectance]). Several common spectrum pretreatments were used, alone or in conjunction, to reduce baseline variations, enhance spectral features, reduce the particle-size scattering effect, remove linear or curvilinear trends of each spectrum, and/or remove additive or multiplicative signal effects (Boysworth and Booksh, 2007; Stenberg et al., 2010): Savitsky-Golay smoothing (Smoo), centring (Centr), standard normal variate (SNV), 1st- and 2nd-order detrending (D1, D2), 1st- and 2nd-order derivative with 11- or 31-point gaps (Der11, Der13, Der21 and Der23); Centr followed by D1, D2 or the derivatives mentioned (e.g., CentrD1, CentrD2 or CentrDer11); SNV followed by D1 (SNVD1), D2 (SNVD2) or the derivatives mentioned (e.g., SNVDer11); D1 or D2 followed by the derivatives mentioned (e.g., D1Der11); CentrD1 or CentrD2 followed by the derivatives mentioned (e.g. CentrD1Der11); SNVD1 or SNVD2 followed by the derivatives mentioned (e.g. SNVD1Der11); and raw absorbance spectra, with no pretreatment, were also studied (Raw). All pretreatments aimed at amplifying the useful parts of spectra and at reducing irrelevant information. More specifically, most of the pretreatments aim at removing additive and multiplicative effects due to light scattering and at enhancing spectral features (e.g., SNV, detrending and derivatives) but using different approaches and with different results (Rinnan et al., 2009; Stenberg et al., 2010). Therefore, combining spectral pretreatments is a common practice and has often proven useful (Rinnan et al., 2009; Stenberg et al., 2010; Ghozalideh et al., 2013). In total, 46 spectrum types were studied, including 45 types of pretreated spectra and raw spectra (Fig. S1). Other common spectrum pretreatments, such as multiplicative scatter correction (MSC) and continuum removal, were not included in this set of pretreatments. Indeed, the former corrects spectra from trends measured over the sample spectrum set so that the corrected spectrum of a given sample changes according to the sample set to which it belongs, which may complicate the use of MSC (Boysworth and Booksh, 2007). The latter has most generally been used for reflectance but not absorbance spectra, especially in remote sensing studies (Clark and Roush, 1984; Cong et al., 2018).

The distributions of C and N contents and EC were skewed, with skewness coefficients reaching 2.4, 3.1 and 4.2, respectively (Table 1). As non-normal distributions hamper statistical procedures, the decimal logarithmic transformation (log_{10}) of these variables was carried out to try to achieve more normal distributions (Lobsey et al., 2017; Terra et al., 2015). Indeed, the distributions of logC, logN and logEC were much less skewed, but the distribution of logEC was still skewed noticeably (skewness coefficients were 0.5, 0.7 and 1.3 in the calibration set, respectively). To document the effect of log-transformation, regression procedures were performed on both C content and logC, N content and logN, and EC and logEC. Variable transformations using the natural logarithm (ln) and square root were also tested, but in general, these transformations resulted in poorer predictions than using log_{10} and will not be presented.

2.3.2. Calibration and validation sets

The set of 311 samples (after one sample was removed as an outlier, cf. 2.2) was divided into a calibration set for building prediction models and a validation set to test these models. This distinction between the calibration and validation sets was based on principal component analysis (PCA) performed on smoothed (2nd-order Savitzky-Golay filtering, width 11) and then centred spectra using the R package FactoMineR (Lê et al., 2008). PCA condenses the huge but redundant information carried by spectra (here absorbance at 2151 wavelengths) into a small number of latent variables (LVs) that are linear combinations of absorbances, and LVs are built to be orthogonal one to another (no redundancy) and to explain maximum variance (to represent at best spectral variability). Then, the Kennard-Stone algorithm (Kennard and Stone, 1969) was applied to PCA scores to select spectrally representative samples for calibration using the R package soil.spec (Sila et al., 2014), and the remaining samples were used for validation. The Kennard-Stone algorithm has become very popular in recent years for optimizing the spectral representativeness of calibration samples within a spectral library (Nocita et al., 2014; Clairotte et al., 2016; Lobsey et al., 2017; Wang et al., 2018; Liu et al., 2019; Moura-Bueno et al., 2019). This algorithm is particularly relevant when the library reflects the variability of land covers and soil types in a given area (e.g., region, country), as was the case in the present study (cf. section 2.1). Indeed, in such cases, validation results provide realistic approximations of how accurately new samples from this area would be predicted using the library. This algorithm first selects the pair of samples separated by the largest Euclidian distance, then the sample most distant from samples already selected, etc., until the required number of samples, here calibration samples, is reached. Due to presumably high soil variability, the size of the calibration set was set to 249 samples (80%) and the size of the validation set to 62 samples (20%).

2.3.3. Regression procedures

Three regression procedures were carried out, which all involved PLSR: global PLSR; locally weighted PLSR; and PLSR by salinity class, after class discrimination using PLS discriminant analysis (PLSDA). All procedures were performed using the R package rnirs (Lesnoff et al., 2020). The PLS procedure aims at condensing the huge and redundant information carried by spectra into a small number of LVs that are: i) linear combinations of absorbances, ii) orthogonal one with another, and iii) built to maximize their covariance with the explained variable (namely, C content, logC, N content, logN, EC or logEC); the last point differs from

PCA, which aims at describing spectrum set diversity (cf. 2.3.2), while PLS is a step in the regression or discrimination procedure. Then, regression or discriminant analysis is carried out with LVs.

Global PLSR is the common PLSR procedure. One prediction model is built using all calibration samples and is then applied to all validation samples (Boysworth and Booksh, 2007).

Locally weighted PLSR (hereinafter called local PLSR) builds one prediction model for each validation sample individually and weights the contribution of calibration samples to model building based on their spectral similarity with that validation sample (Boysworth and Booksh, 2007). Spectral similarity was calculated according to the R correlation coefficient between validation and calibration samples. The weights assigned to calibration samples were calculated according to Lesnoff et al. (2020) using equation 1:

$$w = \exp[-u / SD(u)] \quad \text{Equation 1}$$

where w is the weight, exp is the exponential function, SD is the standard deviation, and u is a parameter defined according to equation 2:

$$u = d / \max(d) \quad \text{Equation 2}$$

where d measures spectral dissimilarity calculated according to equation 3, and max(d) is its maximum over the calibration set:

$$d = [0.5 \times (1 - R)]^{0.5} \quad \text{Equation 3}$$

where R is the correlation coefficient between a given validation sample and each calibration sample.

In addition, weights assigned to calibration samples were also calculated according to equation 4 instead of equation 1:

$$w = \exp[-u / 2SD(u)] \quad \text{Equation 4}$$

However, the prediction results using both weight functions did not differ much; moreover, those achieved with equation 1 were better in general than their counterparts using equation 4 (which gave relatively lower weight to the closest neighbours when compared to equation 1), so it did not seem useful to present the results achieved using equation 4.

PLSR by class first involved PLSDA to predict the salinity class of validation samples (Salt- vs. Salt+) using the PLSDA-lm procedure of the rnirs package (lm for linear model; Lesnoff et al., 2020). This procedure involved the creation of two dummy variables (e.g., the resulting Salt+ variable obtained the value 1 for Salt+ samples and 0 for Salt- samples); PLSR was then used to predict the class corresponding to the dummy variable for which prediction was the highest (e.g., if Salt+ prediction yielded 0.9 and Salt- prediction 0.1, the sample was classified as Salt+). Then, for each soil property considered (C content, logC, N content, logN, EC and logEC), global PLSR built with all calibration samples from a given salinity class (Salt- or Salt+) was applied to validation samples belonging to this class according to PLSDA.

Therefore, for each variable (C, N and EC), three regression procedures (global, local and by class) were tested with 46 spectrum types and possible variable log-transformation (cf. 2.3.1).

Whatever the procedure (global PLSR, local PLSR or PLSDA), randomly selected fivefold cross-validation with 10 replicates was performed on the calibration set, and the number of LVs that minimized the root mean square error of cross-validation (RMSECV) was considered optimal and selected. When a prediction yielded a negative value, it was replaced by zero. All the steps of the chemometric analyses are summarized in Fig. 1.

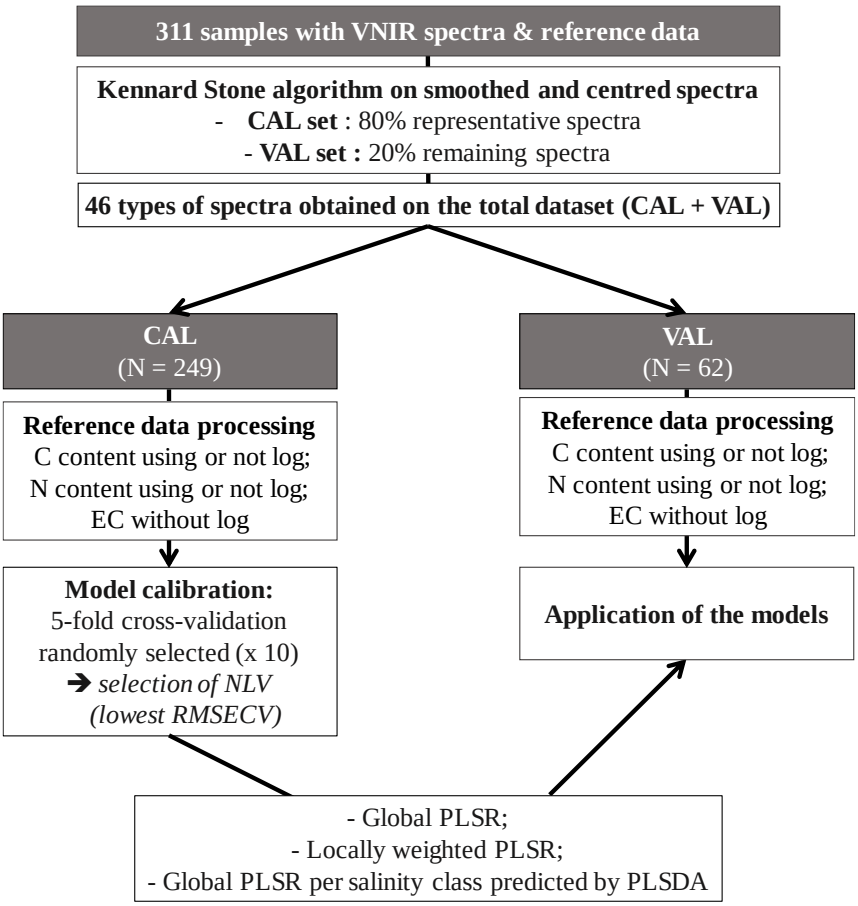


Fig. 1. Diagram of the main steps followed in this study. NLV is the number of latent variables. The dark-background boxes refer to datasets and the white-background ones to processes.

2.3.4. Evaluation of prediction model performance

The parameters used for assessing the goodness of fit of predictions were as follows:

- the root mean square error (RMSE), either calculated over the calibration set in cross-validation (RMSECV) or over the validation set (RMSEP);
- the bias, which is the mean residue over the validation set;
- R^2 between predictions and observations calculated over the calibration set in cross-validation (R^2_{CV}) or over the validation set (R^2_{VAL});
- RPD, which is the ratio of standard deviation to RMSE, on either the calibration set or validation set (RPD_{CV} and RPD_{VAL}, respectively); Chang et al. (2001) considered that RPD > 2 corresponded to accurate NIRS predictions of soil properties;

- RPIQ, which is the ratio of interquartile range (i.e., difference between the third and first quartiles; IQR) to RMSE (RPIQ$_{CV}$, RPIQ$_{VAL}$) and has been recommended instead of RPD for variables that do not follow a normal distribution (Bellon-Maurel et al., 2010).

For PLSDA, other performance parameters were considered:
- sensitivity, which is the proportion of samples correctly assigned to a class (in % of the set);
- specificity, which is the proportion of samples correctly identified as not belonging to the class considered (in %; Altman and Bland, 1994).

2.3.5. Comparisons between sets and between model performances

Similarity between two spectra can be evaluated using the coefficient of determination (R^2), and similarity between two spectrums sets can be evaluated by average R^2 over all possible pairs of spectra made of one spectrum from each set. Therefore, spectral similarity between the calibration and validation sets, which depended on the spectrum type (raw or pretreated spectra), was evaluated using the average R^2 between every calibration and validation spectrum. To evaluate the calibration neighbourhood of validation samples, the number of calibration spectra that were correlated to each validation spectrum with $R^2 \geq 0.95$ was also considered.

For a given variable, the average RMSEP calculated using all 46 types of spectra was compared between the different regression procedures using Student's paired *t*-test when the RMSEP distribution was normal or the Wilcoxon signed rank sum test otherwise after normality was evaluated with the Shapiro-Wilk test. In the same way, for a given regression procedure, the average RMSEP over the 46 types of spectra was compared between predictions using the log-transformation of the considered variable or not.

3. Results and discussion

3.1. Reference data

Table 1 presents the distributions of the three studied variables, namely, C and N contents and EC, without and with log-transformation, in the calibration and validation sets. The observed C and N contents were low, with respective means and SDs amounting to 4.3 ± 3.2 g C kg^{-1} and 0.34 ± 0.24 g N kg^{-1} for the whole dataset; moreover, the observed C and N contents were highly correlated ($R = 0.95$), which confirmed their organic nature (data not shown). When considering both salinity classes, the distribution of observed C and N contents (mean, median, SD and IQR) was quite similar in the calibration and validation sets but with a wider range in the calibration than in the validation set (mostly for N content). However, their distribution and range per salinity class were quite different between the calibration and validation datasets, particularly in Salt+ (mean and SD were 5.6 ± 5.0 vs. 2.3 ± 1.6 g C kg^{-1} and 0.44 ± 0.35 vs. 0.21 ± 0.14 g N kg^{-1}, respectively). For both C and N contents, medians did not differ much in general between Salt- and Salt+ in the calibration set; however, in the validation set, Salt+ samples had twice smaller median C and N contents than Salt- samples. Comparable observations were made for the mean, although a slight difference was also observed in the calibration set (25-29 % lower in Salt- than in Salt+ for both variables). The relative SD (i.e., ratio of SD to mean; RSD, expressed in %) was slightly lower in Salt- than Salt+: for both C and N contents, it was 68 % vs. 80-90 % in the calibration set and 51-63 % vs. 65-70 % in the

validation set for Salt- vs. Salt+, respectively. This indicated higher heterogeneity but without systematically lower values of C and N contents in Salt+ than in Salt-, in accordance with Pankhurst et al. (2001), who studied variably salt-affected soils from different Australian regions and did not find any linear (negative) correlation between soil organic C content and EC for their whole dataset.

Table 1. Distribution of C and N contents and EC before and after log-transformation in calibration and validation sets according to salinity classes. Salt- corresponds to unsalted or slightly salty samples and Salt+ to medium to highly salty ones (EC ≤ vs. > 2 mS cm^{-1}). N_{total}, N_{salt-} and N_{salt+} are the number of samples in total, Salt- and Salt+ sets, respectively. SD is the standard deviation, IQR the interquartile range (difference between third and first quartiles) and Skew the skewness. All parameters have the same unit as the soil property considered, except skewness (unitless).

Soil property	Set	Calibration (N_{total}=249; N_{salt-}=214; N_{salt+}=35)							Validation (N_{total}=62; N_{salt-}=48; N_{salt+}=14)						
		Min	Median	Mean	Max	SD	IQR	Skew	Min	Median	Mean	Max	SD	IQR	Skew
C content	Both	0.8	3.2	4.2	20.7	3.2	2.8	2.4	1.0	3.9	4.6	17.0	3.3	3.3	1.9
(g C kg^{-1})	Salt-	0.8	3.2	4.0	20.7	2.7	2.7	2.7	1.5	4.5	5.3	17.0	3.3	3.2	1.8
	Salt+	0.8	2.9	5.6	17.9	5.0	5.9	1.2	1.0	1.9	2.3	7.3	1.6	1.4	2.1
N content	Both	0.08	0.26	0.34	2.10	0.25	0.20	3.1	0.10	0.31	0.35	1.20	0.20	0.20	1.7
(g N kg^{-1})	Salt-	0.09	0.26	0.33	2.10	0.22	0.19	3.9	0.15	0.35	0.40	1.20	0.20	0.16	1.8
	Salt+	0.08	0.30	0.44	1.43	0.35	0.48	1.2	0.10	0.16	0.21	0.62	0.14	0.14	1.9
EC	Both	0.0	0.0	1.4	29.0	4.1	0.1	4.2	0.0	0.0	2.7	28.6	6.1	0.8	2.8
(mS cm^{-1})	Salt-	0.0	0.0	0.1	1.2	0.2	0.0	4.1	0.0	0.0	0.1	1.3	0.2	0.0	3.6
	Salt+	2.1	6.9	9.3	29.0	6.8	6.4	1.5	3.7	8.1	11.5	28.6	8.0	9.9	1.0
log C	Both	-0.1	0.5	0.5	1.3	0.3	0.4	0.5	0.0	0.6	0.6	1.2	0.3	0.4	0.1
[log$_{10}$(g C kg^{-1})]	Salt-	-0.1	0.5	0.5	1.3	0.2	0.3	0.6	0.2	0.6	0.7	1.2	0.2	0.3	0.4
	Salt+	-0.1	0.5	0.6	1.3	0.4	0.6	0.2	0.0	0.3	0.3	0.9	0.2	0.3	0.8
log N	Both	-1.1	-0.6	-0.5	0.3	0.2	0.3	0.7	-1.0	-0.5	-0.5	0.1	0.2	0.3	0.0
[log$_{10}$(g N kg^{-1})]	Salt-	-1.0	-0.6	-0.5	0.3	0.2	0.3	0.9	-0.8	-0.5	-0.4	0.1	0.2	0.2	0.4
	Salt+	-1.1	-0.5	-0.5	0.2	0.3	0.6	0.1	-1.0	-0.8	-0.7	-0.2	0.2	0.3	0.9
log EC	Both	-2.4	-1.6	-1.2	1.5	1.0	1.0	1.3	-2.1	-1.4	-0.9	1.5	1.1	1.6	0.9
[log$_{10}$(mS cm^{-1})]	Salt-	-2.4	-1.7	-1.6	0.1	0.5	0.6	1.2	-2.1	-1.5	-1.4	0.1	0.5	0.5	1.4
	Salt+	0.3	0.8	0.9	1.5	0.3	0.4	0.2	0.6	0.9	1.0	1.5	0.3	0.4	0.3

The observed EC showed much higher variation in Salt- than Salt+: the RSD of both the calibration and validation sets was 220-230 % for Salt- vs. 70-74 % for Salt+. Moreover, the RSD also differed in the total set between the calibration and validation sets: 298 % vs. 226 %, respectively. Indeed, the EC distribution in Salt+ was quite different between the calibration and validation datasets. In contrast, the EC distribution in Salt- was similar in both the calibration and validation sets: 0.1 ± 0.2 mS cm^{-1} in each set.

EC was measured in a 1:5 soil:water extract, which does not necessarily represent actual field conditions but allows comparison between samples under controlled conditions. Thus, the observed EC values presented here could underestimate sample salinity under field conditions, as their initial water content was not considered (Mavi and Marschner, 2017).

The three variables did not have normal distributions: the skewness value was systematically positive in the calibration and validation sets, with many small values. The log-transformation decreased the skewness parameter close to zero for all three variables so that distributions were more symmetrical; the log-transformation also allowed the distributions per salinity class to be closer between the calibration and validation datasets. However, the distribution of logEC in the calibration and validation sets (except for Salt+) was still very skewed.

3.2. Prediction of the soil electrical conductivity

The log-transformation of EC did not improve its prediction, probably because the distribution of logEC was still very skewed; thus, the corresponding results will not be presented.

3.2.1. Sample discrimination according to EC class

The PLSDA, built to discriminate soil salinity classes using sample VNIR spectra, either raw or pretreated, yielded accurate results: in validation, systematically 100 % of the Salt- samples were correctly discriminated, and on average over all 46 spectrum types (±SD), 95.2 %(±7.5 %) of the Salt+ samples were correctly assigned (Table 2). Some authors also showed accurate discrimination of soil classes when using VNIR spectra. For example, Liu et al. (2018) achieved accurate discrimination of five soil classes (0-20 cm depth) studied in different provinces of China using PLSDA based on VNIR spectra. Vasques et al. (2014) also observed a strong correlation between soil VNIR spectra and taxonomic classes, which was partly explained by soil colour. In the present study, the noticeable PLSDA efficiency could also be explained by the effect of salinity on soil colour and thus on spectrum absorbance (Shahid et al., 2018). Indeed, an effect of soil salinity on VNIR spectra has been reported in the literature: Wang et al. (2018) reported a negative correlation between spectral absorbance and EC, which was also the case in the present study (data not shown); in contrast, Li et al. (2019) found a positive correlation between soil spectral absorbance and salt concentration. These contradictory results could hardly be explained by the predominance of different salts between the studied soils since sodium ions were predominant in the soils studied by Li et al. (2019) and Wang et al. (2018), while sulfate ions were predominant in the soil samples of the present study (Sadio, 1991). However, the presence of salts can variably affect other soil characteristics, such as their composition or structure, and thus indirectly modify soil spectra (Shahid et al., 2018; Stenberg et al., 2010). While the discrimination of Salt- samples was perfect, regardless of the pretreatment, the discrimination of Salt+ samples was slightly affected by the pretreatment used. Indeed, 10.9 % of the 46 types of spectra led to incorrect discrimination for very few Salt+ validation samples (only three or four, depending on the spectrum type considered). In contrast, 89.1 % of the spectrum types led to perfect or nearly perfect (i.e., only one misclassified sample) discrimination for this class. In total, four samples were potentially misclassified, which differed according to spectrum type: they had EC between 3.7 and 6.8 mS cm^{-1}; three of them had the lowest EC in the Salt+ class. We could therefore suppose, as reported by Liu et al. (2018), that depending on spectrum type, these samples could have spectral characteristics closer to the characteristics of Salt- than Salt+ samples.

Table 2. Confusion matrix of PLSDA validation for the two salinity classes (Salt- i.e. EC ≤ 2 mS cm^{-1} vs. Salt+ i.e. EC > 2 mS cm^{-1}) according to the spectrum type (ST) considered. Smoo refers to smoothing; Centr to centring; SNV to standard normal variate; D1 and D2 to 1st- and 2nd-order detrending, respectively; Der111, Der131, Der211 and Der231 to 1st- and 2nd-order derivative with 11- or 31-point gaps, respectively.

ST	%STs with similar result	Predicted salinity	Observed salinity		Total	Sensitivity (%)	Specificity (%)
			Salt- (N_{Salt-}=48)	Salt+ (N_{Salt+}=14)			
Centr, D1, D2, CentrD1, CentrD2, Der111, CentrDer111, D1Der111, D2Der111, CentrD1Der111, CentrD2Der111, Der131, CentrDer131, D1Der131, D2Der131, CentrD1Der131, CentrD2Der131, SNVD2Der131, Der211, CentrDer211, D1Der211, D2Der211, CentrD1Der211, CentrD2Der211, D2Der231, CentrD2Der231	56.5	Salt-	48	0	48	100.0	100.0
		Salt+	0	14	14	100.0	100.0
Raw, Smoo, SNV, SNVD1, SNVD2, SNVD2Der111, SNVDer131, SNVD1Der131, Der231, CentrDer231, D1Der231, CentrD1Der231, SNVDer231, SNVD1Der231, SNVD2Der231	32.6	Salt-	48	1	49	100.0	92.9
		Salt+	0	13	13	92.9	100.0
SNVDer111, SNVD1Der111	4.3	Salt-	48	2	50	100.0	85.7
		Salt+	0	12	12	85.7	100.0
SNVDer211, SNVD1Der211, SNVD2Der211	6.5	Salt-	48	4	52	100.0	71.4
		Salt+	0	10	10	71.4	100.0
Total	100.0						
Mean over all STs (and SD)		Salt-	48.0 (0.0)	0.7 (1.1)	48.7 (1.1)	100.0 (0.0)	95.2 (7.5)
		Salt+	0.0 (0.0)	13.3 (1.1)	13.3 (1.1)	95.2 (7.5)	100.0 (0.0)

3.2.2. Prediction of EC

For each EC calibration procedure without log-transformation (global, local, by class), the results that yielded the lowest RMSEP are presented in Table 3 and Fig. 2, while Fig. 3 compares RMSEP among the three calibration procedures. The most accurate EC prediction was obtained with a global model (using SNVD1; RMSEP = 1.9 mS cm^{-1}), while the best predictions achieved with local PLSR (using SNVD2Der111) and PLSR by class (using Der131) were less accurate (RMSEP = 2.1 and 2.3 mS cm^{-1}, respectively; Table 3 and Fig. 2). Thus, common PLSR (i.e., global PLSR) was appropriate for EC prediction. When averaged over all 46 spectrum types considered, local PLSR yielded the lowest RMSEP (2.4 ± 0.2 mS cm^{-1}), close but significantly lower than global PLSR (2.5 ± 0.3 mS cm^{-1}), which itself was significantly lower than PLSR by class (2.8 ± 0.2 mS cm^{-1}; Table 3 and Fig. 3).

EC in Salt+ samples was more accurately predicted when calibration was carried out on both Salt- and Salt+ samples than on Salt+ samples only (Fig. 2). This could be explained by the small number of Salt+ samples in the calibration set, which limited the prediction accuracy of

the model by class (Kuang and Mouazen, 2012). In contrast, EC in Salt- samples was more accurately predicted when only Salt- samples were used for calibration, probably because of the heterogeneity provided by the addition of Salt+ samples, with a clustering effect (Fig. 2; Stenberg et al., 2010).

To our knowledge, no study has focused on EC (1:5 soil:water) prediction in variably salt-affected soils using VNIRS, but predictions achieved for Salt- samples were comparable to those reported in the literature for the same class. Dunn et al. (2002) studied topsoil samples originating from southern New South Wales (Australia), with variable texture but low EC (< 1.8 mS cm^{-1}); using global calibration, they achieved $RPD_{VAL} = 1.2$, which is comparable to the results of the present study for Salt- samples ($RPD_{VAL} = 1.3$ on average and up to 1.4; Table 3). Islam et al. (2003) studied EC prediction using ultraviolet and VNIRS (250-2500 nm) in unsalted and slightly salty soils (EC < 1.5 mS cm^{-1}) in New South Wales and southeast Queensland (Australia) and achieved $RPD_{VAL} = 1.0$. Concerning the Salt+ class, Weindorf et al. (2016) used VNIRS for predicting EC in salt-rich samples collected at several depths in the Monegros region, Spain (average EC $= 5.8 \pm 3.8$ mS cm^{-1}). They found $1.6 \leq RPD_{VAL} \leq 2.0$ and $2.4 \leq RPIQ_{VAL} \leq 3.0$ using support vector regressions and penalized spline regressions, which is more accurate than in the present study, possibly attributed to the more homogeneous EC distribution in the sample population studied by Weindorf et al. (2016) and to the limited number of Salt+ samples in the present study.

The interpretation of EC prediction accuracy differed according to the performance parameter that was considered. For the total validation dataset, RPD_{VAL} and R^2_{VAL} suggested that the models were accurate in general, whereas $RPIQ_{VAL}$ suggested that the models were not accurate (Table 3). Due to the very skewed distribution of EC (log EC was still far from normal, cf. 2.3.1), the IQR was very low and much smaller than SD, so RPIQ was also much smaller than RPD. This was, however, not the case in the Salt+ class, but the latter included few samples. Contradiction between high RPD and low RPIQ highlighted the difficulty of interpreting prediction results sometimes, and the limits of parameters used for evaluating the performance of prediction models for variables that have very asymmetrical and/or clustered distributions. For non-normally distributed variables, Bellon-Maurel et al. (2010) recommended considering RPIQ instead of RPD and thus referring RMSE to IQR instead of SD for evaluating model performance; however, this approach does not necessarily lead to easily interpretable results for very skewed distributions, as seen here. Moreover, studying EC in both clusters separately (Salt- and Salt+) did not necessarily improve the prediction results. On the one hand, EC was very homogeneous in Salt- samples: SD (0.1 mS cm^{-1}) and IQR (0.0 mS cm^{-1}) were very small, so RMSEP could hardly be lower. On the other hand, the number of Salt+ samples was too small to allow accurate prediction, as mentioned previously.

Table 3. Cross-validation and validation results of VNIRS predictions of electrical conductivity ($mS\ cm^{-1}$) for each PLSR method i) using raw spectra; ii) using the spectrum type that minimized RMSEP; iii) averaged over all 46 spectrum types (with SD). For PLSR by class, validation parameters were calculated for i) the total validation set, ii) the validation samples predicted as Salt- by PLSDA; iii) the validation samples predicted as Salt+ by PLSDA (the size of Salt- and Salt+ validation sets depended on the spectrum type, cf. Table 2).

Spectrum type	VAL set	LV	Calibration (N_{total}=249; N_{salt-}=214; N_{salt+}=35)				Validation (N_{total}=62; unfixed N_{salt-} and N_{salt+})							
			R^2_{CV}	RMSE-CV	RPD_{CV}	$RPIQ_{CV}$	Mean	SD	R^2_{VAL}	$Slope_{VAL}$	$Bias_{VAL}$	RMSEP	RPD_{VAL}	$RPIQ_{VAL}$
Global PLSR														
Raw	Total	11	0.84	1.6	2.5	0.1	2.7	6.1	0.88	0.83	0.1	2.1	2.9	0.4
SNVD1	Total	9	0.85	1.6	2.6	0.1	2.7	6.1	0.90	0.84	0.1	1.9	3.1	0.4
Mean	Total	7	0.82	1.8	2.3	0.1	2.7	6.1	0.83	0.77	0.2	2.5	2.4	0.3
(SD)		(2)	(0.02)	(0.1)	(0.1)	(0.0)			(0.04)	(0.06)	(0.1)	(0.3)	(0.3)	(0.0)
Local PLSR														
Raw	Total	7	0.82	1.8	2.3	0.1	2.7	6.1	0.87	0.88	0.3	2.2	2.8	0.4
SNVD2 Der111	Total	5	0.84	1.7	2.5	0.1	2.7	6.1	0.89	0.79	0.0	2.1	2.9	0.4
Mean	Total	7	0.82	1.8	2.3	0.1	2.7	6.1	0.85	0.81	0.1	2.4	2.5	0.3
(SD)		(2)	(0.02)	(0.1)	(0.1)	(0.0)			(0.03)	(0.04)	(0.1)	(0.2)	(0.2)	(0.0)
PLSR By class*														
Raw	Total	9, 11, 8	0.84	1.7	2.4	0.1	2.7	6.1	0.79	0.8	-0.3	2.8	2.2	0.3
	Salt-		0.55	0.1	1.4	0.3	0.2	0.8	0.62	0.31	0.0	0.5	1.4	0.1
	Salt+		0.58	5.1	1.3	1.3	12.0	8.1	0.49	0.65	-1.3	6.0	1.4	1.8
Der131	Total	7, 15, 3	0.81	1.8	2.2	0.1	2.7	6.1	0.86	0.76	-0.1	2.3	2.6	0.3
	Salt-		0.63	0.1	1.6	0.3	0.1	0.2	0.63	0.99	0.1	0.2	1.2	0.2
	Salt+		0.50	5.1	1.3	1.3	11.5	8.0	0.70	0.43	-0.5	4.9	1.6	2.0
Mean	Total	6, 14, 6	0.83	1.7	2.4	0.1	2.7	6.1	0.79	0.80	-0.2	2.8	2.2	0.3
(SD)		(2, 2, 2)	(0.02)	(0.1)	(0.1)	(0.0)			(0.04)	(0.04)	(0.2)	(0.2)	(0.2)	(0.0)
	Salt-		0.59	0.1	1.5	0.3	0.2	0.6	0.63	0.60	0.0	0.5	1.3	0.1
			(0.05)	(0.0)	(0.1)	(0.0)	(0.1)	(0.5)	(0.10)	(0.30)	(0.1)	(0.4)	(0.1)	(0.1)
	Salt+		0.56	5.0	1.4	1.3	11.8	8.1	0.49	0.59	-0.8	5.9	1.4	1.8
			(0.05)	(0.2)	(0.1)	(0.1)	(0.7)	(0.1)	(0.10)	(0.10)	(0.6)	(0.4)	(0.1)	(0.2)

* For PLSR by class, the number of latent variables is mentioned for PLSDA, PLSR for samples predicted as Salt- and PLSR for samples predicted as Salt+, respectively (with SDs into brackets when averaged over all spectrum types)

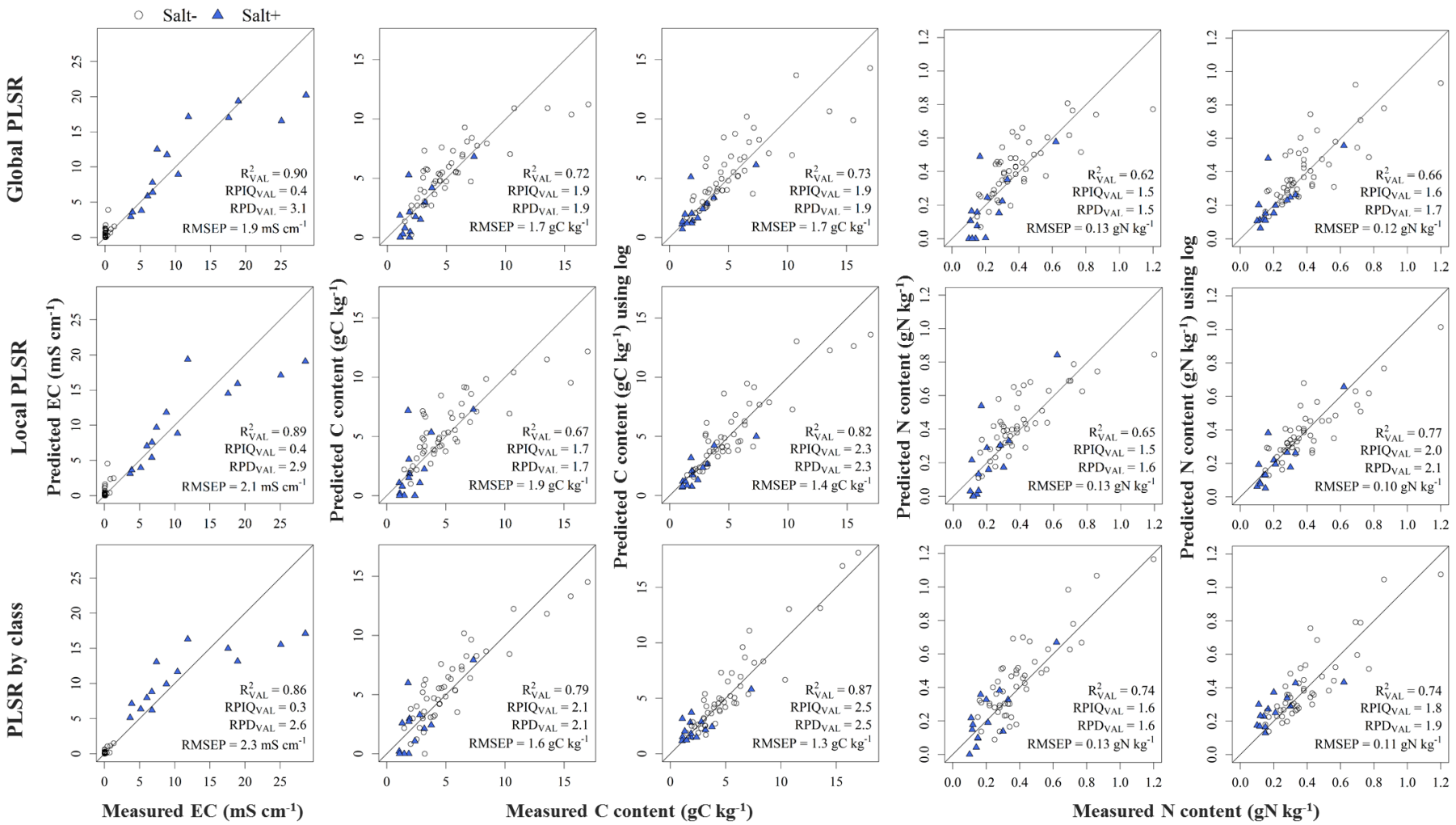


Fig. 2 Measured vs. predicted EC (without log-transformation) and C and N contents (with and without log-transformation) using the spectral type (among the 46 tested) that minimized RMSEP for each PLSR (cf. Table 6). The vertical axis title specifies when C and N contents were predicted after log-transformation ("using log"). Salt- corresponds to unsalted or slightly salty samples and Salt+ to medium to highly salty samples. The line represents $y = x$.

3.3 Prediction of soil organic carbon content

Table 4 presents C content prediction results using raw spectra, using spectrum types that yielded lowest RMSEP, and on average over all spectrum types, for the three calibration procedures, without log-transformation or using log-transformation, while boxplots presented in Fig. 3 compare distributions of RMSEP over all spectrum types according to calibration procedure and possible log-transformation of C content.

Using log-transformation or not, the best validation results were obtained with PLSR by class: without log-transformation, the lowest RMSEP was 1.6, 1.7 and 1.9 gC kg^{-1} with PLSR by class, global and local PLSR, respectively; using log-transformation, the lowest RMSEP was 1.3, 1.7 and 1.4 gC kg^{-1}, respectively. Better validation results with PLSR by class than with global and local PLSR seemed to result mainly from better prediction on C-rich samples, which were Salt- samples in general (Fig. 2). The same result was found when RMSEP was averaged over the 46 spectrum types, using log-transformation or not (Table 4; Fig. 3), which could be attributed to higher sample set homogeneity allowed by calibration by class in both calibration and validation. Indeed, Brunet et al. (2007) reported that NIRS prediction accuracy increased when soil C and N content calibrations were built and applied to texturally homogeneous sample sets. Liu et al. (2018) also reported better soil organic C content prediction with models built and applied by soil type than using the whole spectral library (RMSEP decreased by 11%). Better prediction using calibration by salinity class could additionally be explained by the positive correlation between soil salinity and soil moisture, even in dried samples (Nawar et al., 2015), as heterogeneous soil moisture can impact C content prediction accuracy (Allory et al., 2019).

Moreover, regardless of the calibration procedure, the lowest RMSEP was obtained using variable log-transformation, which in particular allowed low C content values to be better predicted (Table 4; Fig. 2). The mean RMSEP over all 46 spectrum types was also smaller using log-transformation, regardless of the calibration procedure. Indeed, the log-transformation of C content led to less asymmetrical variable distribution and, in particular, to "ungrouping" of C-poor samples, which contributed to higher model accuracy (Table 1; Lucà et al., 2017; Moura-Bueno et al., 2019).

On the whole, the benefit of PLSR by class using log-transformed C content values was noticeable when compared with global PLSR using C content values (i.e., common procedure), either considering best results or averages over all spectrum types. The most accurate prediction of C content, using PLSR by class, log-transformation and SNVDer131, yielded RMSEP = 1.3 gC kg^{-1}, RPD_{VAL} = 2.5 and $RPIQ_{VAL}$ = 2.5, which, according to Chang et al. (2001), was accurate (RPD_{VAL} > 2). Under such conditions, RMSEP was 24% lower than the lowest RMSEP obtained with global PLSR (using D2 with log-transformation or SNV without log-transformation; Table 4), which is a higher difference than that reported by Liu et al. (2018) when comparing PLSR by soil type and global PLSR in Chinese provinces. Other studies that addressed VNIRS or NIRS prediction of sandy topsoil C content at the regional scale used global calibration in general and achieved results comparable to those achieved with this procedure in the present study (e.g., RPD_{VAL} = 2.3 for Barthès et al., 2008, for a sample population originating from four sites in Burkina Faso and Congo-Brazzaville; RPD_{VAL} = 1.9 for Cambule et al., 2012, in a 1000-km² area in Mozambique, while RPD_{VAL} = 1.9 was achieved in the present study).

The model performance could probably be improved by increasing the number of Salt+ samples, which was small. Indeed, the positive effect of increasing the number of calibration samples on validation accuracy has been reported by several authors, either at the local scale (Lucà et al., 2017) or national scale (Clairotte et al., 2016), until an optimum was reached, depending on sample population diversity.

Table 4. Cross-validation and validation results of VNIRS predictions of soil organic carbon content (gC kg^{-1}) using or not log-transformation for each PLSR method i) using raw spectra; ii) using the spectrum type that minimized RMSEP; iii) averaged over all 46 spectrum types (with SD).

Model	Spectrum type	LV	Calibration (N_{total}=249)				Validation (N_{total}=62)					
			R^2_{CV}	RMSE-CV	RPD$_{CV}$	RPIQ$_{CV}$	R^2_{VAL}	Slope	Bias$_{VAL}$	RMSE-P	RPD$_{VAL}$	RPIQ$_{VAL}$
Global – no log	Raw	13	0.68	1.9	1.7	1.5	0.64	0.70	-0.1	2.0	1.7	1.7
	SNV	11	0.57	2.2	1.5	1.3	0.72	0.74	0.1	1.7	1.9	1.9
	Mean	12	0.62	2.0	1.6	1.4	0.59	0.68	0.2	2.1	1.5	1.5
	(SD)	(1)	(0.03)	(0.1)	(0.1)	(0.0)	(0.06)	(0.05)	(0.1)	(0.2)	(0.1)	(0.1)
Global – using log	Raw	14	0.41	4.3	0.7	0.6	0.71	0.69	-0.2	1.8	1.9	1.8
	D2	12	0.66	1.9	1.7	1.5	0.73	0.79	0.1	1.7	1.9	1.9
	Mean	13	0.66	2.0	1.6	1.4	0.68	0.74	0.1	1.9	1.8	1.7
	(SD)	(2)	(0.07)	(0.5)	(0.2)	(0.2)	(0.04)	(0.08)	(0.1)	(0.1)	(0.1)	(0.1)
Local – no log	Raw	9	0.63	2.0	1.6	1.4	0.6	0.69	0.0	2.1	1.6	1.6
	SNV	7	0.56	2.2	1.5	1.3	0.67	0.74	0.2	1.9	1.7	1.7
	Mean	10	0.64	2.0	1.6	1.4	0.59	0.68	0.2	2.2	1.5	1.5
	(SD)	(2)	(0.03)	(0.1)	(0.1)	(0.1)	(0.06)	(0.05)	(0.1)	(0.2)	(0.1)	(0.1)
Local – using log	Raw	11	0.57	2.5	1.3	1.1	0.74	0.79	-0.2	1.7	2.0	2.0
	Centr	11	0.39	3.9	0.8	0.7	0.82	0.85	-0.1	1.4	2.3	2.3
	Mean	11	0.67	1.9	1.7	1.5	0.70	0.75	0.1	1.8	1.8	1.8
	(SD)	(2)	(0.09)	(0.4)	(0.3)	(0.2)	(0.06)	(0.06)	(0.1)	(0.2)	(0.2)	(0.2)
By class – no log	Raw	9, 15, 8*	0.65	1.9	1.6	1.4	0.72	0.81	0.2	1.8	1.9	1.9
	SNVDer131	7, 15, 9*	0.70	1.8	1.8	1.6	0.79	0.91	0.2	1.6	2.1	2.1
	Mean	6, 13, 8	0.60	2.0	1.6	1.4	0.67	0.78	0.1	1.9	1.7	1.7
	(SD)	(2, 3, 4)*	(0.07)	(0.2)	(0.1)	(0.1)	(0.08)	(0.09)	(0.1)	(0.2)	(0.2)	(0.2)
By class – using log	Raw	9, 13, 8*	0.65	2.0	1.6	1.4	0.75	0.80	0.0	1.6	2.0	2.0
	SNVDer131	7, 15, 7*	0.74	1.7	1.9	1.7	0.87	1.02	0.2	1.3	2.5	2.5
	Mean	6, 14, 6	0.64	1.9	1.7	1.5	0.75	0.84	0.2	1.7	2.0	2.0
	(SD)	(2, 2, 2)*	(0.07)	(0.2)	(0.2)	(0.1)	(0.06)	(0.09)	(0.1)	(0.2)	(0.2)	(0.2)

* For PLSR by class, the number of latent variables is mentioned for PLSDA, PLSR for samples predicted as Salt- and PLSR for samples predicted as Salt+, respectively (with SDs into brackets when averaged over all spectrum types)

3.4. Prediction of soil nitrogen content

Predictions of N content were performed with the three calibration procedures, either using log-transformation or not. Table 5 presents the results achieved using raw spectra, using the most appropriate spectrum types (i.e., lowest RMSEP) and on average over all spectrum types. In

addition, Fig. 3 presents the effects of the calibration procedure, spectrum type and log-transformation on the N content prediction accuracy.

Without log-transformation, the three models were not accurate according to Chang et al. (2001; RPD < 2) and provided similar RMSEP (0.13 g N kg^{-1}) when using the most appropriate pretreatment. Nevertheless, PLSR by class tended to predict the N content of Salt+ samples more accurately than global and local PLSR (Fig. 2): despite the small number of Salt+ samples, their N content prediction was thus more accurate using only Salt+ samples for calibration than by using both Salt- and Salt+ samples, which was not observed for C prediction of Salt+ samples. Therefore, the VNIR spectral signature of salinity might have an effect on N content prediction, which has not yet been reported in the literature. Using log-transformation, the most accurate predictions were achieved with local PLSR (and SNVDer1l1), yielding RMSEP = 0.10 g N kg^{-1}, RPD_{VAL} = 2.1 and $RPIQ_{VAL}$ = 2.0 and was thus accurate according to Chang et al. (2001). In contrast, the most accurate predictions achieved with the global model and by class yielded RPD_{VAL} = 1.7 and 1.9, respectively, so they were not accurate according to this reference. Using log-transformation or not, the lowest mean RMSEP over all 46 spectrum types was also obtained with local PLSR (Table 4; Fig. 3).

For each calibration procedure, the lowest RMSEP and lowest mean RMSEP over all spectrum types were obtained with N content log-transformation, as was also the case for C. However, unlike C content, the benefit of log-transformation tended to concern not only N-poor samples but also most samples poorly predicted without log-transformation (Fig. 2). The benefit of log-transformation on RMSEP for a given calibration procedure was not much larger than the standard error of laboratory analysis (≈ ±0.01 g N kg^{-1}; ISO, 1998); although significant, this benefit was thus limited in general.

On the whole, the benefit of local PLSR using log-transformed N content values was noticeable when compared with global PLSR using N content values (i.e., common procedure), either considering best results or averages over all spectrum types.

Because of the very low N content (mean and SD were 0.34 ± 0.24 g C kg^{-1} over the whole dataset), the results obtained in this study were necessarily less accurate than the results mostly reported in the literature (e.g., RPD_{VAL} = 2.8 and 4.0 for Chang and Laird, 2002 and Morón and Cozzolino, 2004, who studied sample sets where the N content averaged 2.8 and 5.8 g N kg^{-1}, respectively). However, for a topsoil sample population originating from four sandy sites in north, central and south Burkina Faso and in Congo-Brazzaville, with N content similar to the N content of the present study (mean and SD were 0.37 ± 0.20 g N kg^{-1} vs. 0.34 ± 0.24 g N kg^{-1} here), Barthès et al. (2008) achieved RPD_{VAL} = 2.2 using global NIRS calibration (also using spectrally representative samples for calibration). This result is comparable to the best predictions achieved in the present study, using local calibration but better than global PLSR predictions achieved here, possibly because Barthès et al. (2008) studied unsalted soils.

Table 5. Cross-validation and validation results of VNIRS predictions of soil nitrogen content (g N kg^{-1}) using or not log-transformation for each PLSR method i) using raw spectra; ii) using the spectrum type that minimized RMSEP; iii) averaged over all 46 spectrum types (with SD).

Model	Spectrum type	LV	Calibration (N_{total}=249)				Validation (N_{total}=62)					
			R^2_{CV}	RMSE-CV	RPD$_{CV}$	RPIQ$_{CV}$	R^2_{VAL}	Slope	Bias$_{VAL}$	RMSE-P	RPD$_{VAL}$	RPIQ$_{VAL}$
Global – no log	Raw	12	0.63	0.15	1.6	1.3	0.57	0.79	-0.01	0.14	1.4	1.4
	SNV	11	0.48	0.19	1.3	1.1	0.62	0.78	0.01	0.13	1.5	1.5
	Mean	12	0.55	0.17	1.4	1.2	0.54	0.76	0.02	0.15	1.4	1.3
	(SD)	(1)	(0.03)	(0.01)	(0.1)	(0.0)	(0.06)	(0.07)	(0.01)	(0.01)	(0.1)	(0.1)
Global – using log	Raw	12	0.50	0.21	1.1	0.9	0.65	0.76	-0.01	0.12	1.7	1.6
	D2	12	0.64	0.15	1.6	1.3	0.66	0.78	0.01	0.12	1.7	1.6
	Mean	13	0.64	0.16	1.6	1.3	0.63	0.78	0.01	0.13	1.6	1.5
	(SD)	(2)	(0.07)	(0.02)	(0.2)	(0.1)	(0.05)	(0.08)	(0.01)	(0.01)	(0.1)	(0.1)
Local – no log	Raw	11	0.63	0.16	1.6	1.3	0.58	0.83	0.00	0.15	1.4	1.3
	SNVD2 Der 1 31	7	0.59	0.16	1.5	1.2	0.65	0.84	0.03	0.13	1.6	1.5
	Mean	10	0.58	0.17	1.5	1.2	0.60	0.80	0.02	0.14	1.4	1.4
	(SD)	(2)	(0.05)	(0.01)	(0.1)	(0.1)	(0.03)	(0.05)	(0.01)	(0.01)	(0.1)	(0.1)
Local – using log	Raw	11	0.65	0.16	1.5	1.2	0.63	0.76	-0.01	0.13	1.6	1.6
	SNVDer 2 11	15	0.50	0.22	1.1	0.9	0.77	0.78	-0.01	0.10	2.1	2.0
	Mean	11	0.65	0.16	1.6	1.3	0.67	0.78	0.01	0.12	1.7	1.7
	(SD)	(2)	(0.09)	(0.03)	(0.3)	(0.2)	(0.06)	(0.08)	(0.01)	(0.01)	(0.2)	(0.2)
By class – no log	Raw	9, 12, 7*	0.61	0.16	1.6	1.3	0.47	0.71	0.01	0.16	1.2	1.2
	Der 1 31	7, 15, 9*	0.59	0.16	1.5	1.2	0.74	1.02	0.04	0.13	1.6	1.6
	Mean	6, 12, 8	0.51	0.18	1.4	1.1	0.55	0.79	0.03	0.15	1.3	1.3
	(SD)	(2, 3, 3)*	(0.07)	(0.01)	(0.1)	(0.1)	(0.09)	(0.11)	(0.02)	(0.02)	(0.1)	(0.1)
By class – using log	Raw	9, 13, 8*	0.65	0.16	1.6	1.3	0.68	0.84	0.01	0.12	1.7	1.6
	SNV	9, 15, 3*	0.57	0.17	1.5	1.2	0.74	0.83	0.02	0.11	1.9	1.8
	Mean	6, 14, 8	0.61	0.16	1.6	1.3	0.66	0.91	0.03	0.14	1.5	1.5
	(SD)	(2, 1, 4)*	(0.07)	(0.01)	(0.1)	(0.1)	(0.09)	(0.12)	(0.01)	(0.02)	(0.2)	(0.2)

* For PLSR by class, the number of latent variables is mentioned for PLSDA, PLSR for samples predicted as Salt- and PLSR for samples predicted as Salt+, respectively (with SDs into brackets when averaged over all spectrum types)

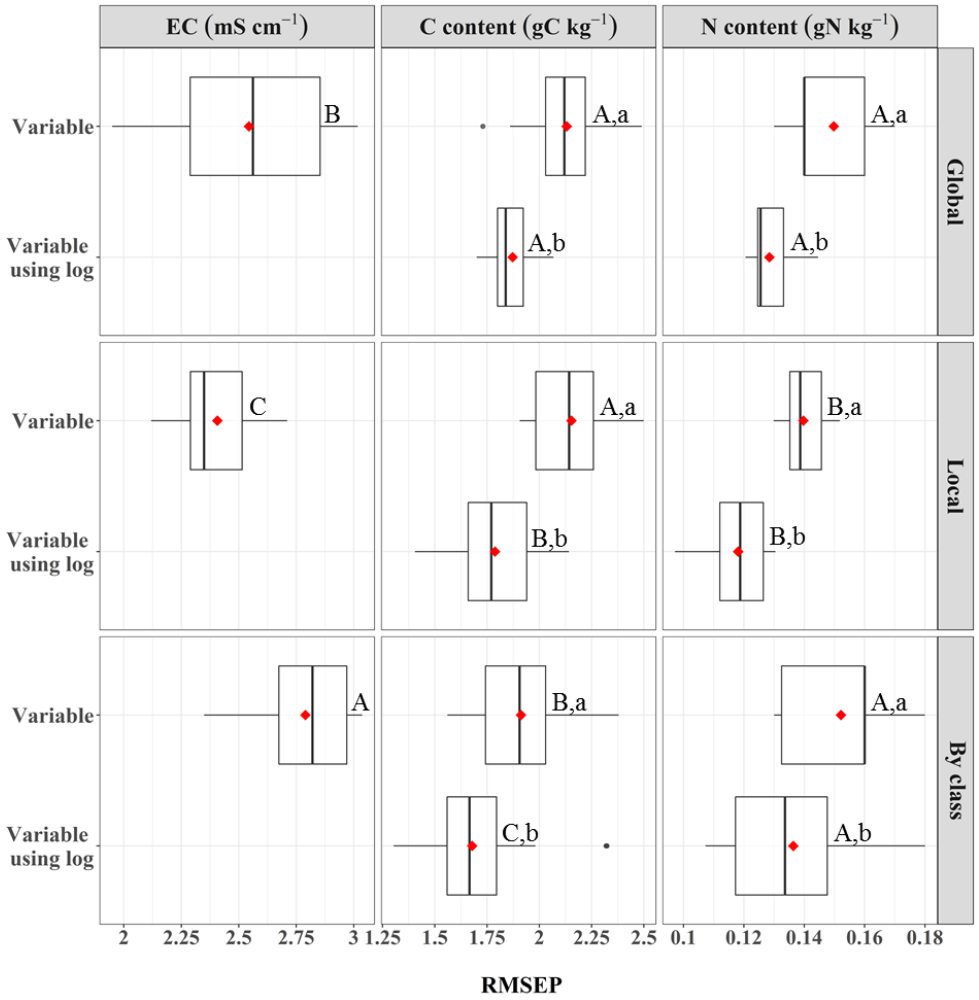


Fig. 3. Distribution of RMSEP over the 46 spectrum types for the three variables and three PLSR procedures, possibly using log-transformation for C and N contents. Vertical lines inside boxes represent medians, red diamonds means. The bottom and top of each box are first and third quartiles. Boxplot whisker ends are either the minimum and maximum values when they were included in 1.5 times the interquartile range; otherwise, the latter value was used (and the minimum and/or maximum points were represented out of the boxplot). Lower-case letters indicate significant effects of log-transformation; and capital letters indicate significant effects of PLSR method, considering procedures without and with log-transformation separately (the significance of differences was estimated using Student's paired *t*-test for normal distributions, Mann-Whitney-Wilcoxon test otherwise; $p < 0.05$); "a" and "A" were assigned to highest values.

3.5. *Effect of spectrum type on spectral similarity and validation results*

3.5.1. Effect of spectrum type on spectral similarity between calibration and validation sets

The R^2 coefficient was used to evaluate the similarity between calibration and validation sample spectra (e.g., Shenk et al., 1997), which depended on spectrum type (raw or pretreated spectrum). Table 6 shows: i) the average R^2 between the calibration and validation spectra, and ii) the average number of calibration neighbours ($R^2 > 0.95$) per validation sample, both considering the spectrum type that yielded the best validation results for each PLSR procedure and variable.

For predictions without log-transformation, the most accurate global PLSR were achieved with pretreatments that resulted in strong spectral similarity between calibration and validation samples: SNV for C and N contents and SNVD1 for EC (average $R^2 = 0.95$; 172 and 154 spectral neighbours on average, respectively). This result seemed logical because all calibration spectra have the same weight when building global PLSR, which is similarly applied to all validation samples. For local PLSR without log-transformation, the best prediction of C content was achieved with SNV, which resulted in strong spectral similarity between calibration and validation samples, while the best predictions for N content and EC were achieved with SNVD2 Der131 and SNVD2 Der111, which resulted in lower similarity (average $R^2 = 0.85$-0.87; 51-52 calibration neighbours on average). As the closest calibration neighbours had the highest weight in local model building, this difference indicated that more calibration neighbours were required for predicting C content than N content and EC, suggesting that N and salinity might have clearer spectral signatures than C. This result is counterintuitive since the C content is much higher than the N content in soils, so that a stronger spectral signature would have been expected for soil C than for soil N. However, fewer calibration neighbours were required to achieve the best N prediction than the best C prediction, which suggested that the soil N signature was easier to catch. Thus, we might hypothesize that the spectral signature of soil N was less dispersed than the spectral signature of soil C, in accordance with the much smaller chemical diversity of N compounds than C compounds in soils. The best PLSR by class without log-transformation were obtained with spectrum types that resulted in intermediate similarity between calibration and validation spectra: Der131 for N content and EC and SNVDer131 for C content (average $R^2 = 0.87$; 59 spectral neighbours on average). This result could be explained by PLSR by class involving three separate models: i) PLSDA to discriminate Salt- and Salt+ samples; ii) specific global PLSR for samples predicted as Salt- (built with Salt- calibration samples); and iii) specific global PLSR for samples predicted as Salt+ (built with Salt+ calibration samples). Thus, we might hypothesize that optimizing PLSR by class required a kind of compromise over the three separate models. The best predictions in PLSR by class were achieved with spectrum types that resulted in intermediate similarity between calibration and validation spectra, so we might assume that such spectrum types helped to reach this compromise.

In contrast, using log-transformation of C and N contents, the most accurate global models were obtained with D2, which split the total set between calibration and validation samples (average $R^2 = 0.73$; 34 calibration neighbours on average). Strong spectral similarity between calibration and validation samples being no longer required suggested that log-transformation of C and N contents would make global calibrations easier. The best local PLSR for C content was obtained using Centr, which yielded high spectral similarity (average $R^2 = 0.95$ and 172 calibration neighbours on average, as was also the case without log-transformation but using SNV), while the best local PLSR for N content was obtained with SNVDer211, which yielded low spectral similarity (average $R^2 = 0.80$ and 10 calibration neighbours only on average, i.e., lower than for its counterpart without log-transformation). In line with previous considerations, the N spectral signature was suggested to be even clearer using log-transformation. The best model by class for C content was obtained with SNVDer131 (as was also the case without log-transformation), which yielded intermediate spectral similarity, while the best model by class for N content was obtained with SNV, which yielded high spectral similarity (higher than for its counterpart

without log-transformation). In line with previous considerations, with log-transformation, optimizing the three steps of PLSR by class was suggested to still require a kind of compromise for C (as was the case without log-transformation: intermediate spectral similarity) but not for N (high similarity).

Overall, the best prediction for C content was achieved with PLSR by class and log-transformation using a pretreatment that yielded intermediate similarity between calibration and validation samples. The best prediction for N content was achieved with local PLSR and log-transformation using a pretreatment that yielded low similarity. Finally, the best EC prediction was achieved with global PLSR without log-transformation using a pretreatment that yielded high similarity. Possible reasons that would explain the pretreatment effect on prediction results have rarely been examined specifically; nevertheless, Liu et al. (2019) also observed an effect of pretreatment on the Kennard Stone selection of calibration samples based on spectral representativeness and thus on prediction results.

Table 6. Effect of the spectrum type (ST) that minimized RMSEP on spectral similarity according to i) average R^2 (+/-SD) between calibration and validation spectra (CAL and VAL, N_{total} = 249 and 62, respectively); and ii) average number of calibration neighbours (+/-SD) per validation spectrum considering $R^2 > 0.95$. RMSEP was calculated for global, local and per-class calibration of C and N contents and EC, possibly using their log-transformation (except for EC). Best predictions are underlined.

Combination of PLSR and ST that minimized RMSEP	Average R^2 between CAL and VAL	Average number of calibration neighbours with $R^2 > 0.95$
Without variable log-transformation		
SNV for global PLSR on C and N and local PLSR on C	0.95+/-0.06	172+/-69
<u>SNVD1 for global PLSR on EC</u>	0.95+/-0.05	154+/-63
Der131 for per-class PLSR on N and EC	0.87+/-0.09	59+/-47
SNVDer131 for per-class PLSR on C	0.87+/-0.09	59+/-47
SNVD2 Der111 for local PLSR on EC	0.87+/-0.09	51+/-45
SNVD2 Der131 for local PLSR on N	0.85+/-0.11	52+/-43
With variable log-transformation		
Centr for local PLSR on C	0.95+/-0.06	172+/-69
SNV for per-class PLSR on N	0.95+/-0.06	172+/-69
<u>SNVDer131 for per-class PLSR on C</u>	0.87+/-0.09	59+/-47
<u>SNVDer211 for local PLSR on N</u>	0.80+/-0.11	10+/-12
D2 for global PLSR on C and N	0.73+/-0.24	34+/-33

3.5.2. Effect of spectrum type on prediction accuracy

The effect of spectrum type on prediction results involved aspects other than spectral similarity between calibration and validation samples. Fig. 4 presents RMSEP for each combination model × variable × spectrum type, and in the case of C and N contents, either using variable log-transformation or not. Twenty-two spectrum types (out of 46) were removed because they always yielded prediction results similar to or slightly worse than those achieved with other

particular pretreatments, regardless of the variable or calibration procedure (e.g., Centr before another pretreatment always yielded exactly the same result as that pretreatment alone). A similar figure with all 46 spectrum types is presented in Fig. S1.

The best predictions of C content required log-transformation and were achieved with calibration by class using SNVDer131, SNVD1Der131, SNVD2Der131 (RMSEP = 1.3-1.4 g C kg^{-1}), SNV and Der131 (1.4-1.5 g C kg^{-1}), but very good predictions were also achieved with local calibration using Centr and SNV (RMSEP = 1.4-1.5 g C kg^{-1}). The best N content predictions also required log-transformation and were achieved with local calibration using SNVDer211 (and SNVD1Der211 and SNVD2Der211), Der211 (and D1Der211 and D2Der211; Fig. S1; RMSEP=0.10 g N kg), SNV, D1 and SNVD1 (0.11 g N kg^{-1}). However, comparable prediction results were also achieved with a model by class using SNV, SNVD1, SNVD2 and SNVD2Der131 (0.11 g N kg^{-1}). The best EC predictions were achieved with the global model using SNVD1, SNV, Raw and Centr (RMSEP = 1.9-2.1 mS cm^{-1}), but comparable results were achieved with the local model using SNVD2Der111 and Centr (2.1 mS cm^{-1}).

No given pretreatment yielded good prediction results over all variables and calibration procedures, but trends could be observed. For C and N contents, either using log-transformation or not, pretreatments with Der131 often yielded some of the best prediction results for a given calibration procedure and never yielded poor results. In contrast, good predictions of C and N contents were rarely achieved with pretreatments that involved Der211 and Der231. For EC, SNV always yielded good predictions, while Der211 (and D1Der211 and D2Der211; Fig. S1) always yielded poor predictions.

Comparing the effects of pretreatments regardless of whether log-transformation was used led to contrasting observations. For global calibration of C content, a given pretreatment most often yielded comparable performances with or without log-transformation, either good (Centr, SNV, D1 and D2), poor (2nd-order derivatives) or intermediate (SNVD1 and SNVD2). The trend was similar but less marked for local calibrations of C and N contents, but the pretreatments that yielded good or poor predictions were not necessarily the same as for global calibration of C content. However, the trend tended to be opposite for the global calibration of N content: a majority of pretreatments that yielded good predictions with log-transformation yielded poor predictions without log-transformation. The situation was intermediate for calibration by class of C and N contents, with comparable proportions of pretreatments that yielded similar kinds of results and opposite kinds of results with and without log-transformation.

Depending on the variable and calibration procedure, the variation of prediction results according to spectrum type might be high or low; thus, pretreatment optimization was sometimes crucial and sometimes not crucial, which has rarely been reported in the literature. The RSD of RMSEP over all 46 spectrum types was the lowest, 5%, for C and N contents in global PLSR with log and for N content in local PLSR without log; in such conditions, pretreatment selection was not truly crucial (e.g., for global C content calibration with log, 1.7 ≤ RMSEP ≤ 2.1 g C kg^{-1}). The RSD of RMSEP was 7-8% for C and N contents in global PLSR and for EC in local and per-class PLSR, all without log. The RSD of RMSEP was 9-10% in local PLSR for C content with or without log and for N content with log, which yielded the best N content predictions. The RSD of RMSEP was 11-13% for EC in global PLSR, which yielded the best EC predictions, and for C and N contents in PLSR by class, all without log;

and for C content in PLSR by class with log, which yielded the best C content predictions. The RSD of RMSEP was 15% for N content in PLSR by class with log; in such conditions, selecting an appropriate pretreatment was decisive. The RSD of RMSEP was even > 60% for EC with log, whatever the PLSR: here, avoiding an inappropriate pretreatment was indispensable (Fig S1). The RSDs of RPD_{VAL} and $RPIQ_{VAL}$ were generally close to the RSD of RMSEP (data not shown). Thus, the best predictions of C and N contents and EC were achieved with calibration procedures that produced rather high variation in prediction results depending on spectrum type (9% for N and 12% for C and EC). Actually, the effect of pretreatment tended to be less decisive in general in global calibration than in calibration by class: the RSD of RMSEP ranged from 5 to 8% for the former, except 12% for EC, vs. 11-15% for the latter, except 7% for EC. The effect of pretreatment was intermediate in local calibration (8-10% except 5% for N content). This result might be due to the larger effect of pretreatments when fewer samples are used for calibration: this was the case in calibration by class (only calibration samples from the class considered were used) and to some extent in local calibration (only calibration neighbours had a noticeable contribution). This effect has been partly reported by Liu et al. (2019), who studied the effect of calibration set size on SOM prediction using six pretreatments: they observed that pretreatment affected prediction accuracy only when fewer than 70-80 samples were used for calibration (global calibration here). This effect might also be inferred from comparisons between studies that involved sample sets of different sizes and tested several pretreatments (e.g., with global PLSR, Clairotte et al., 2016, on a set of > 3800 samples, achieved best VNIRS predictions of organic C content just using a moving average on 10 bands, while Vasques et al., 2008, on a set of ca. 550 samples, achieved the best VNIRS predictions of log C using derivatives). Moreover, in calibration by class, misclassification (which might concern up to four samples; Table 2) could increase the variability of prediction accuracy depending on pretreatment. Moura-Bueno et al. (2019) also observed that the range of C content prediction accuracy according to VNIR spectrum pretreatment varied between the four calibration procedures they tested. The present study additionally shows that the pretreatment effect depended on the studied soil property and its possible log-transformation.

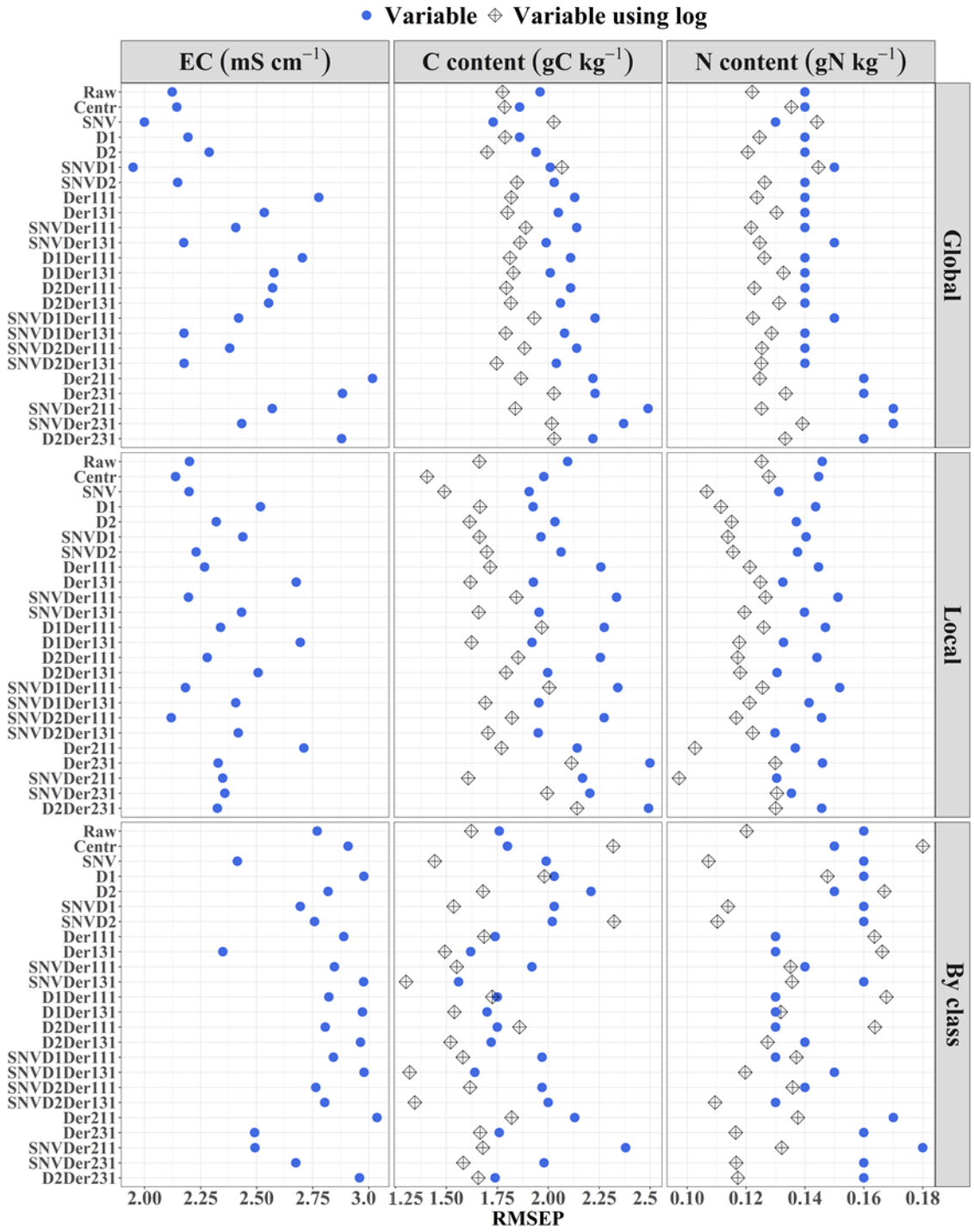


Fig. 4 Variations of RMSEP for EC and C and N content predictions according to PLSR procedure, spectrum type (24 spectrum types, after 22 were removed, cf. subsection 3.5.2) and possible log-transformation (results with log-transformation not presented for EC, cf. subsection 3.2).

Conclusion

The present study evaluated the effects of calibration procedures (global, local or per-salinity class PLSR, i.e., for Salt- ≤ 2 mS cm^{-1} vs. Salt+ > 2 mS cm^{-1}), log-transformation of the explained variable, and spectrum type (46 were tested) on VNIRS predictions of topsoil C and N contents and EC at the regional scale in variably salty soils of the Sine Saloum area of Senegal. The best prediction of C content was achieved with PLSR by class applied to spectrum absorbance pretreated with SNVDer131 and using log-transformation (RMSEP = 1.3 gC kg^{-1}, RPD_{VAL} = 2.5 and $RPIQ_{VAL}$ = 2.5). In contrast, the best prediction of N content was achieved with local PLSR applied to spectrum absorbance pretreated with SNVDer211, also using log-transformation (RMSEP = 0.10 gN kg^{-1}, RPD_{VAL} = 2.1 and $RPIQ_{VAL}$ = 2.0). The best EC prediction was achieved with global PLSR applied to spectrum absorbance pretreated with SNVD1, without log-transformation (RMSEP = 1.9 mS cm^{-1}, RPD_{VAL} = 3.1 and $RPIQ_{VAL}$ = 0.4; the latter was explained by the dominance of very slightly salty samples). We might, however, assume that global calibration would not necessarily be the most appropriate procedure for predicting EC in a sample population with more Salt+ samples. Moreover, the distributions of logC and logN were almost symmetrical, hence the usefulness of log-transformation for predicting these variables; however, the distribution of logEC was still very asymmetrical, so log-transformation of EC did not help its prediction. Spectrum pretreatment affected prediction accuracy, but no pretreatment yielded good prediction results over all variables and calibration procedures; nevertheless, pretreatments with Der131 (1st-order derivative with a 31-point gap) often yielded good predictions, especially for C and N contents, while 2nd-order derivatives yielded poor results in general.

Therefore, no unique procedure would optimize VNIRS prediction of soil properties in a heterogeneous regional spectral library: calibration approach as well as processing of explanatory and explained variables must be tailored depending on the property and its distribution, as highlighted by the results of the present study; but also depending on sample set size and diversity, which was not studied here but has been suggested by other studies. Nevertheless, and importantly, the present study showed that accurate prediction of the soil salinity class could easily be achieved by PLSDA (on average, over all spectrum types, 100% and 95% of Salt- and Salt+ validation samples were correctly assigned, respectively). The present study also showed that accurate VNIRS predictions of C and N contents and EC in variably salt-affected soils could be achieved ($RPD_{VAL} > 2$) using different combinations of calibration procedures and processing (including pretreatment and log-transformation).

Acknowledgements

This work was supported by the SoCa project funded by the Climate Initiative of BNP Paribas Foundation, France. This work also received support from the French National Research Institute for Sustainable Development (IRD, France). The authors acknowledge the joint international laboratory IESOL for technical assistance and the LAMA/IMAGO laboratory for analytical services. Laure Chauvin personally thanks Prof. Bienvenue Sambou (UCAD-ISE), Abdourahmane Tamba (ISRA-CNRF) and Amadou Tahirou Diaw (UCAD-LERG) for their guidance, and the staff in the research groups they lead for kind support during her Master's degree. The authors are grateful to Matthieu Lesnoff (CIRAD), Jean-Michel Roger (INRAE) and Jean-Pierre Montoroi (IRD) for scientific exchanges. We sincerely thank the landowners and authorities of Senegalese villages for granting permission for soil collection. Finally, helpful comments by anonymous referees are gratefully acknowledged.

References

Allory, V., Cambou, A., Moulin, P., Schwartz, C., Cannavo, P., Vidal-Beaudet, L., Barthès, B. G., 2019. Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions. Science of the Total Environment 686, 764–773. https://doi.org/10.1016/j.scitotenv.2019.05.192

Altman, D.G., Bland, J. M., 1994. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. British Medical Journal 308, 1552. https://doi.org/10.1136/bmj.308.6943.1552

Barthès, B. G., Brunet, D., Hien, E., Enjalric, F., Conche, S., Freschet, G.T., d'Annunzio, R., Toucet-Louri, J., 2008. Determining the distributions of soil carbon and nitrogen in particle size fractions using near-infrared reflectance spectrum of bulk soil samples. Soil Biology and Biochemistry 40, 1533–1537. https://doi.org/10.1016/j.soilbio.2007.12.023

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., McBratney, A., 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC Trends in Analytical Chemistry 29, 1073–1081. https://doi.org/10.1016/j.trac.2010.05.006

Boysworth, M.K., Booksh, K.S., 2007. Aspects of multivariate calibration applied to near-infrared spectroscopy, in: Handbook of Near-Infrared Analysis. Burns, D.A., Ciurczak, E.W. (Eds.), CRC Press, Boca Raton, FL, USA, pp. 207–229. https://doi.org/10.1201/9781420007374-15

Brunet, D., Barthès, B. G., Chotte, J.-L., Feller, C., 2007. Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: effects of sample grinding and set heterogeneity. Geoderma 139, 106–117. https://doi.org/10.1016/j.geoderma.2007.01.007

Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., Smaling, E.M.A., 2012. Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique. Geoderma 183–184, 41–48. https://doi.org/10.1016/j.geoderma.2012.03.011

Chang, C.-W., Laird, D.A., 2002. Near-infrared reflectance spectroscopic analysis of soil C and N. Soil Science 167, 110–116. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R. Jr., 2001. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. Soil Science Society of America Journal 65, 480–490. https://doi.org/10.2136/sssaj2001.652480x

Chauvin, L., 2013. La Salinisation des Terres dans la Région de Fatick (Sénégal) : Etendue et Conséquences sur les Services Ecosystémiques du Système de Production Agropastoral (MSc. Thesis). University Cheikh Anta Diop, Dakar, Senegal.

Clairotte, M., Grinand, C., Kouakoua, E., Thébault, A., Saby, N.P.A., Bernoux, M., Barthès, B. G., 2016. National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. Geoderma 276, 41–52. https://doi.org/10.1016/j.geoderma.2016.04.021

Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. Journal of Geophysical Research: Solid Earth 89, 6329–6340. https://doi.org/10.1029/JB089iB07p06329

Cong, L.-X., Huang, M., Liu, X.-L., Qi, Y.-S., 2018. Retrieval of soil organic carbon based on bi-continuum removal combined with orthogonal partial least squares. Spectroscopy and Spectral Analysis 38, 941–947. https://doi.org/10.3964/j.issn.1000-0593(2018)03-0941-07

Diatta, M., Diack, M., Sene, M., Pasternak, D., 2001. Bioreclamation of Saline Soils of the Western Coast of Senegal, in: Combating Desertification with Plants. Pasternak, D., Schlissel, A. Springer, Dordrecht, Germany, pp. 315–324.
Dignac, M.-F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G.T., Garnier, P., Guenet, B., Hedde, M., Klumpp, K., Lashermes, G., Maron, P.-A., Nunan, N., Roumet, C., Basile-Doelsch, I., 2017. Increasing soil carbon storage: mechanisms, effects of agricultural practices and proxies. A review. Agronomy for Sustainable Development 37. https://doi.org/10.1007/s13593-017-0421-2
Dominati, E., Patterson, M., Mackay, A., 2010. A framework for classifying and quantifying the natural capital and ecosystem services of soils. Ecological Economics 69, 1858–1868. https://doi.org/10.1016/j.ecolecon.2010.05.002
Dunn, B.W., Batten, G.D., Beecher, H.G., Ciavarella, S., 2002. The potential of near-infrared reflectance spectroscopy for soil analysis —a case study from the Riverine Plain of south-eastern Australia. Australian Journal of Experimental Agriculture 42, 607–614. https://doi.org/10.1071/EA01172
Farifteh, J., van der Meer, F., van der Meijde, M., Atzberger, C., 2008. Spectral characteristics of salt-affected soils: A laboratory experiment. Geoderma 145, 196–206. https://doi.org/10.1016/j.geoderma.2008.03.011
Faye, E., Touré, M.A., Diouf, Y.K., 2019. Effets du stress salin sur la germination des graines de Jatropha curcas L. VertigO - la revue électronique en sciences de l'environnement. https://doi.org/10.4000/vertigo.25327
Gupta, G.S., 2019. Land degradation and challenges of food security. Review of European Studies 11, 63–72. https://doi.org/10.5539/res.v11n1p63
Hossain, M.S., 2019. Present scenario of global salt affected soils, its management and importance of salinity research. International Research Journal of Biological Sciences 1, 1–3. https://doi.org/irjbs.2019.1.3
Islam, K., Singh, B., McBratney, A., 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. Soil Research 41, 1101–1114. https://doi.org/10.1071/sr02137
ISO (International Organization for Standardisation), 1994. ISO 11265:1994 —Soil Quality — Determination of the Specific Electrical Conductivity. ISO, Geneva.
ISO (International Organization for Standardisation), 1995. ISO 10694:1995 —Soil Quality — Determination of Organic and Total Carbon after Dry Combustion (Elementary Analysis). ISO, Geneva.
ISO (International Organization for Standardisation), 1998. ISO 13878:1998 —Soil Quality — Determination of Total Nitrogen Content by Dry Combustion ("Elemental Analysis"). ISO, Geneva.
IUSS (International Union of Soil Science) Working Group WRB (World Reference Base), 2015. World Reference Base for Soil Resources 2014, Update 2015: International Soil Classification System for Naming Soils and Creating Legends for Soil Maps. World Soil Resources Reports, 106, FAO, Rome.
Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137–148. https://doi.org/10.2307/1266770

Kuang, B., Mouazen, A.M., 2012. Influence of the number of samples on prediction error of visible and near infrared spectroscopy of selected soil properties at the farm scale. European Journal of Soil Science 63, 421–429. https://doi.org/10.1111/j.1365-2389.2012.01456.x
Lal, R., 2014. Soil conservation and ecosystem services. International Soil and Water Conservation Research 2, 36–47. https://doi.org/10.1016/S2095-6339(15)30021-6
Lê, S., Josse, J., Husson, F., 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25, 1–18. https://doi.org/10.18637/jss.v025.i01
Lesnoff, M., Metz, M., Roger, J.-M., 2020. Comparison of locally weighted PLS strategies for regression and discrimination on agronomic NIR data. Journal of Chemometrics 34, e3209. https://doi.org/10.1002/cem.3209
Li, X., Ren, J., Zhao, K., Liang, Z., 2019. Correlation between spectral characteristics and physicochemical parameters of soda-saline soils in different states. Remote Sensing 11, 388. https://doi.org/10.3390/rs11040388
Liu, Y., Chen, Y., 2012. Feasibility of estimating Cu contamination in floodplain soils using VNIR spectroscopy – A case study in the Le'an River floodplain, China. Soil and Sediment Contamination: An International Journal 21, 951–969. https://doi.org/10.1080/15320383.2012.712069
Liu, Y., Liu, Y., Chen, Y., Zhang, Y., Shi, T., Wang, J., Hong, Y., Fei, T., Zhang, Y., 2019. The influence of spectral pretreatment on the selection of representative calibration samples for soil organic matter estimation using Vis-NIR reflectance spectroscopy. Remote Sensing 11, 450. https://doi.org/10.3390/rs11040450
Liu, Y., Shi, Z., Zhang, G., Chen, Y., Li, S., Hong, Y., Shi, T., Wang, J., Liu, Y., 2018. Application of spectrally derived soil type as ancillary data to improve the estimation of soil organic carbon by using the Chinese soil Vis-NIR spectral library. Remote Sensing 10, 1747. https://doi.org/10.3390/rs10111747
Lobsey, C.R., Viscarra Rossel, R.A., Roudier, P., Hedley, C.B., 2017. RS-local data-mines information from spectral libraries to improve local calibrations. European Journal of Soil Science 68, 840–852. https://doi.org/10.1111/ejss.12490
Lucà, F., Conforti, M., Castrignanò, A., Matteucci, G., Buttafuoco, G., 2017. Effect of calibration set size on prediction at local scale of soil carbon by Vis-NIR spectroscopy. Geoderma 288, 175–183. https://doi.org/10.1016/j.geoderma.2016.11.015
Mavi, M.S., Marschner, P., 2017. Impact of salinity on respiration and organic matter dynamics in soils is more closely related to osmotic potential than to electrical conductivity. Pedosphere 27, 949–956. https://doi.org/10.1016/S1002-0160(17)60418-1
Morón, A., Cozzolino, D., 2004. Determination of potentially mineralizable nitrogen and nitrogen in particulate organic matter fractions in soil by visible and near-infrared reflectance spectroscopy. Journal of Agricultural Science 142, 335–343. https://doi.org/10.1017/S0021859604004290
Moura-Bueno, J.M., Dalmolin, R.S.D., ten Caten, A., Dotto, A.C., Demattê, J.A.M., 2019. Stratification of a local VIS-NIR-SWIR spectral library by homogeneity criteria yields more accurate soil organic carbon predictions. Geoderma 337, 565–581. https://doi.org/10.1016/j.geoderma.2018.10.015

Nawar, S., Buddenbaum, H., Hill, J., 2015. Estimation of soil salinity using three quantitative methods based on visible and near-infrared reflectance spectroscopy: a case study from Egypt. Arabian Journal of Geosciences 8, 5127–5140. https://doi.org/10.1007/s12517-014-1580-y
Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B., Montanarella, L., 2014. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. Soil Biology and Biochemistry 68, 337–347. https://doi.org/j.soilbio.2013.10.022
O' Rourke, S. M., Holden, N. M., 2011. Optical sensing and chemometric analysis of soil organic carbon – a cost effective alternative to conventional laboratory methods? Soil Use and Management 27, 143–155. https://doi.org/10.1111/j.1475-2743.2011.00337.x
Pankhurst, C. E., Yu, S., Hawke, B. G., Harch, B. D., 2001. Capacity of fatty acid profiles and substrate utilization patterns to describe differences in soil microbial communities associated with increased salinity or alkalinity at three locations in South Australia. Biology and Fertility of Soils 33, 204–217. https://doi.org/10.1007/s003740000309
Rietz, D. N., Haynes, R. J., 2003. Effects of irrigation-induced salinity and sodicity on soil microbial activity. Soil Biology and Biochemistry 35, 845–854. https://doi.org/10.1016/S0038-0717(03)00125-1
Rinnan, Å., van den Berg, F., Engelsen, S. B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry 28, 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007
Roger, J., Noël, B. J., Barusseau, J. P., Serrano, O., Nehlig, P., Duvail, C., 2009. Notice Explicative de la Carte Géologique du Sénégal à 1/500 000, Feuilles Nord-Ouest, Nord-Est et Sud-Ouest. Ministère des Mines, de l'Industrie et des PME, Direction des Mines et de la Géologie, Dakar, Senegal.
Sadio, S., 1991. Pédogenèse et Potentialités Forestières des Sols Sulfatés Acides Salés des Tannes du Sine Saloum, Sénégal. ORSTOM Editions, Paris.
Shahid, S. A., Zaman, M., Heng, L., 2018. Introduction to Soil Salinity, Sodicity and Diagnostics Techniques, in: Zaman, M., Shahid, S. A., Heng, L. (Eds.), Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques. Springer, Dordrecht, Germany, pp. 1–42. https://doi.org/10.1007/978-3-319-96190-3_1
Shenk, J. S., Westerhaus, M. O., Berzaghi, P., 1997. Investigation of a LOCAL calibration procedure for near infrared instruments. Journal of Near Infrared Spectroscopy 5, 223–232. https://doi.org/10.1255/jnirs.115
Sila, A., Hengl, T., Terhoeven-Urselmans, T., 2014. soil.spec: Soil Spectroscopy Tools and Reference models. R package version 2. http://14.139.160.195/cran/web/packages/soil.spec/index.html (last accessed on July 8, 2020).
Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., Wetterlind, J., 2010. Visible and Near Infrared Spectroscopy in soil science. Advances in Agronomy 107, 163–215. https://doi.org/10.1016/S0065-2113(10)07005-7
Tappan, G. G., Sall, M., Wood, E. C., Cushing, M., 2004. Ecoregions and land cover trends in Senegal. Journal of Arid Environments 59, 427–462. https://doi.org/10.1016/j.jaridenv.2004.03.018

Terra, F.S., Demattê, J.A.M., Viscarra Rossel, R.A., 2015. Spectral libraries for quantitative analyses of tropical Brazilian soils: Comparing vis–NIR and mid-IR reflectance data. Geoderma 255–256, 81–93. https://doi.org/10.1016/j.geoderma.2015.04.017
USGS (United States Geological Service), 2011. Landsat 7 Data Users Handbook. Washington, DC, USA.
Vasques, G.M., Demattê, J.A.M., Viscarra Rossel, R.A., Ramírez-López, L., Terra, F.S., 2014. Soil classification using visible/near-infrared diffuse reflectance spectra from multiple depths. Geoderma 223–225, 73–78. https://doi.org/10.1016/j.geoderma.2014.01.019
Vasques, G.M., Grunwald, S., Sickman, J.O., 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma 146, 14–25. https://doi.org/10.1016/j.geoderma.2008.04.007
Viscarra Rossel, R.A., Walvoort, D.J.J., 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, 59–75. https://doi.org/10.1016/j.geoderma.2005.03.007
Wang, J., Ding, J., Abulimiti, A., Cai, L., 2018. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS–NIR) spectroscopy, Ebinur Lake Wetland, Northwest China. PeerJ 6, e4703. https://doi.org/10.7717/peerj.4703
Weindorf, D.C., Chakraborty, S., Herrero, J., Li, B., Castañeda, C., Choudhury, A., 2016. Simultaneous assessment of key properties of arid soil by combined PXRF and Vis–NIR data. European Journal of Soil Science 67, 173–183. https://doi.org/10.1111/ejss.12320
Wong, V.N.L., Greene, R.S.B., Dalal, R.C., Murphy, B.W., 2010. Soil carbon dynamics in saline and sodic soils: a review. Soil Use and Management 26, 2–11. https://doi.org/10.1111/j.1475-2743.2009.00251.x
Xu, X., Chen, S., Xu, Z., Yu, Y., Zhang, S., Dai, R., 2020. Exploring appropriate preprocessing techniques for hyperspectral soil organic matter content estimation in Black soil area. Remote Sensing 12, 3765. https://doi.org/10.3390/rs12223765
Yuan, B.-C., Li, Z.-Z., Liu, H., Gao, M., Zhang, Y.-Y., 2007. Microbial biomass and activity in salt affected soils under arid conditions. Applied Soil Ecology 35, 319–328. https://doi.org/10.1016/j.apsoil.2006.07.004

Supplementary material

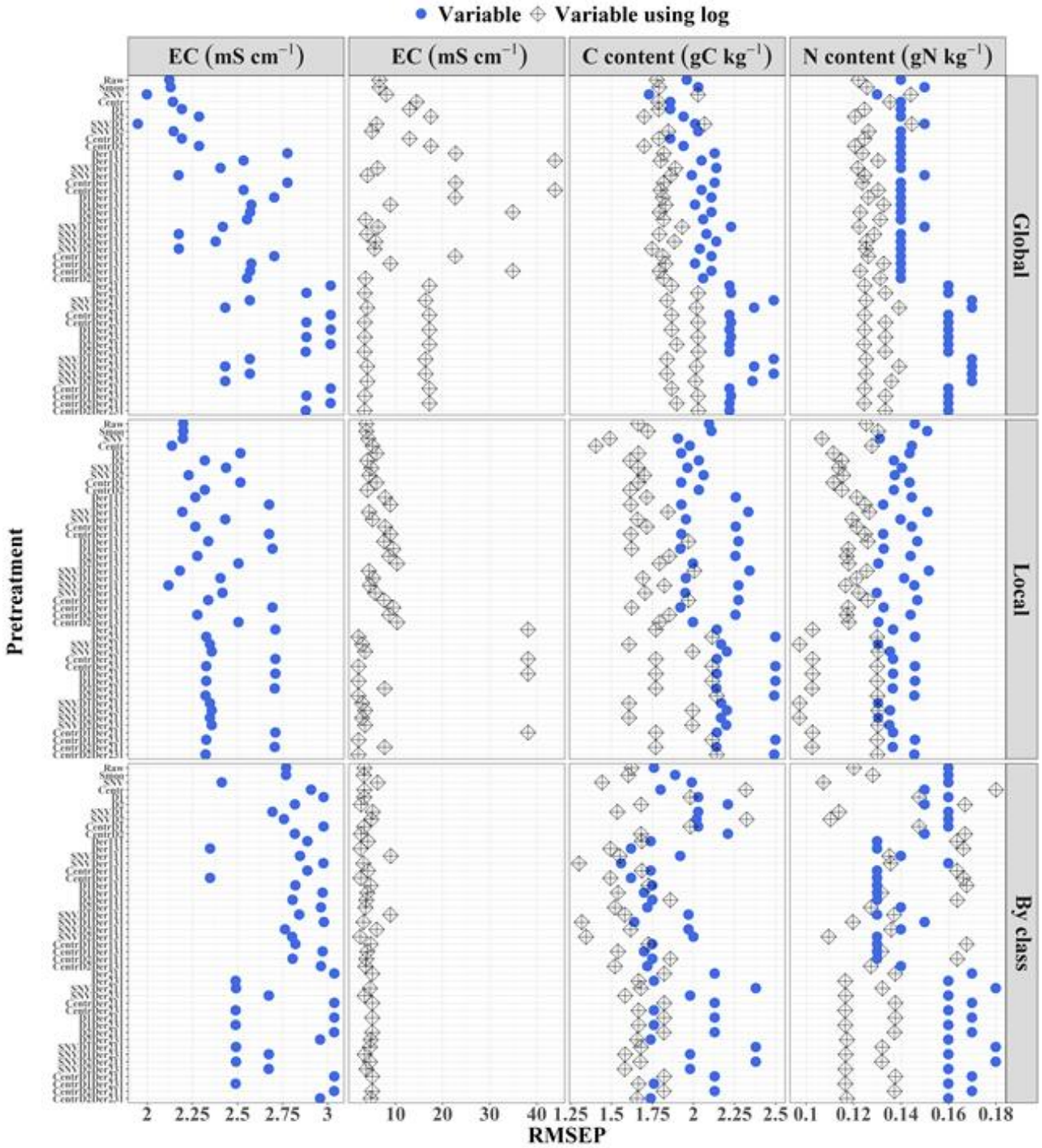


Fig. S1. Variations of RMSEP for EC and C and N content predictions according to PLSR procedure, spectrum type (46 were tested) and possible log-transformation. Very high RMSEP achieved for EC prediction using PLSR by class and log-transformation with Der1, Centr Der1, D1 Der1 and Centr D1 Der1 have not been included in the figure (RMSEP = 341,011 mS cm^{-1}).