Non-saturated soil organic horizon characterization via advanced proximal sensors

NON-SATURATED SOIL ORGANIC HORIZON CHARACTERIZATION VIA ADVANCED PROXIMAL SENSORS

Valeria Cardelli[a*], David C. Weindorf[b*], Somsubhra Chakraborty[c], Bin Li[d], Mauro De Feudis[e], Stefania Cocco[a], Alberto Agnelli[c], Ashok Choudhury[c], Deb Prasad Ray[f], Giuseppe Corti[a]

[a]Department of Agricultural, Food and Environmental Sciences, Università Politecnica delle Marche, Ancona, AN, Italy

[b]Department of Plant and Soil Sciences, Texas Tech University, Lubbock, TX, USA

[c]Uttar Banga Krishi Viswavidyalaya, Cooch Behar, India

[d]Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA, USA

[e]Department of Agricultural, Food and Environmental Sciences, Università degli Studi di Perugia, Perugia, PG, Italy

[f]National Institute of Research on Jute and Allied Fibre Technology, Kolkata, India

*Corresponding author:
David C. Weindorf, Ph.D.
Department of Plant and Soil Science
Box 42122
Lubbock, TX 79409, USA
E-mail: david.weindorf@ttu.edu

Abstract

The organic fraction of soils is critically important to soil health and optimal ecosystem functioning. Traditional analysis of soil organic horizons (O horizons) has been dependent upon laboratory-based instrumentation. Simultaneously, the use of proximal sensors such as portable X-ray fluorescence (PXRF) spectrometry along with visible near infrared diffuse reflectance spectroscopy (VisNIR DRS) have gained popularity for providing rapidly acquired spectral and elemental data useful for soil physicochemical property quantification. However, PXRF and VisNIR DRS have mostly been applied to the assessment of mineral soils. This preliminary study evaluated 136 organic laden soil samples (most aptly described as upland, non-saturated O horizons) using both laboratory based instrumentation (CN analyzer) and proximal sensors to evaluate total carbon (TC) and total nitrogen (TN). Results revealed that combining model outcomes using model fusion improved TC and TN prediction accuracies relative to using an individual instrument (PXRF or VisNIR DRS) or model averaging with improvements in root mean square error (RMSE) on the order of 10-47% and 10-67% for TC and TN, respectively. Partial least squares + random forest (PLS+RF) approaches emerged as the best model for both predicting both TC and TN in organic laden soil samples. These results suggest that the strong predictive applications of proximal sensors extensively documented on mineral soils, may show similar promise for determination of a wide number of physicochemical properties on organic soil matrices, yet further exploration with a larger and more diverse dataset is recommended.

Key words: Spectroscopy; O horizons; proximal sensors

Abbreviations: SOM, soil organic matter; TC, total carbon; TN, total nitrogen; VisNIR-DRS, visible near infrared diffuse reflectance spectroscopy; PXRF, portable X-ray fluorescence

1. Introduction

Organic matter decomposition is a fundamental process for sustaining life on Earth (Gosz et al., 1976). The term soil organic matter (SOM) refers to all organic material in soil, from freshly deposited detritus or litter to highly decomposed, stable forms such as humic and fulvic acid (Stevenson, 1994). Organic matter cycling helps to maintain ecosystem functionality as several ecological functions are correlated to the decay processes of the organic layers of forest soils. Indeed, decomposition and mineralization processes of organic residues affect nutrients cycling and induce the release of elements that represent the principal resources for plants and microbes (Berger et al., 2002; Berg and McClaugherty, 2008), such as macro- and micro-nutrients, and essential molecules for energy metabolism, photosynthesis, and membrane transport (Huttl and Schaaf, 1997). One of the main factors controlling the organic matter decomposition processes is the quality of the litter produced by plants (Ge et al., 2013). The specific chemical proprieties of the plant litter and its decay products, in turn, influence the underlying mineral soil (Wardle et al., 2004; Ball et al., 2014). Six et al. (2004) noted that the decomposition of SOM has an impact on several important soil properties as it improves soil aggregation (Bronick and Lal, 2005), enhances the activity of the soil microbial community (Ball et al., 2014; Carrillo et al., 2012; García-Palacios et al., 2013), and affects mineral weathering (Qafoku, 2015) and soil fertility (Kaiser et al., 2008). Thus, the knowledge of the characteristics and composition of SOM, and in particular of the organic layers, may help to ascertain certain soil ecosystem functions.

Current methods of SOM characterization are well established (Nelson and Sommers, 1996), but are largely laboratory based. Recently, several studies have investigated rapid, inexpensive, and non-destructive methods, such as visible near infrared diffuse reflectance spectroscopy (VisNIR-DRS) and portable x-ray fluorescence spectrometry (PXRF) for soil analysis (Horta et al., 2015; Weindorf et al., 2014). These proximal sensing methods have become increasingly accurate and

widely accepted offering data *in situ* in seconds given virtually no pre-processing requirements (Viscarra Rossel et al., 2006a; 2006b), with substantive advantages over traditional laboratory-based techniques. VisNIR-DRS is a spectrometric method which uses wavelengths across visible and near infrared regions (350-2500 nm) to explore the interaction between incident radiation and reflectance off of the soil surface; absorption is facilitated by C-H, N-H, or O-H bonds within the matrix (Chang et al., 2005). Due to this characteristic, it is highly applicable to C and N determination in soils. However, VisNIR-DRS spectra are generally weak, non-specific, and somewhat broad in their extent because of overlapping spectral signatures arising from variable soil components (Stenberg et al., 2010). As such, the instrument alone does not provide sufficient accuracy for complete soil characterization (Morgan et al., 2009). In fact, others have suggested the application of VisNIR-DRS in tandem with other sensing technologies (Brown et al., 2006; Fajardo et al., 2015). A complementary technique, PXRF, provides a multi-elemental analysis with a large range of quantification from low mg kg^{-1} to 100% for many elements (Hettipathirana, 2004). However, elements with stable electron configuration and low fluorescent energy (e.g., Na, N, H, Li, C) are not detectable (Wang et al., 2015). Nonetheless, several recent studies (e.g., Aldabaa et al., 2015; Chakraborty et al., 2015; Wang et al., 2015) have shown compelling predictive accuracy by combining the spectral signature of VisNIR-DRS with elemental data from PXRF, the latter used as auxiliary input data into the original advanced regression model. Individual or combined use of these two instruments allows for characterization of multiple soil parameters to include SOM (Stenberg et al., 2010), total carbon, total nitrogen (Wang et al., 2015), total phosphorus (Hu, 2013), cation exchange capacity (Sharma et al., 2015), pH (Sharma et al., 2014), salinity, (Swanhart et al., 2014), texture (Zhu et al., 2011), and contaminants (Chakraborty et al., 2015; Horta et al., 2015; Paulette et al., 2015). While the aforementioned studies offer wide-ranging application, most were conducted on mineral soils with limited organic content. Comparatively less information is available on the use of combined proximal sensors for soil organic layer (O horizon)

characterization. Wang et al. (2015) evaluated total carbon and nitrogen via combined PXRF and VisNIR-DRS approaches, but did so on mineral soils, lacking any analysis of true organic horizons. Similarly, Chang at al. (2002) have shown the efficacy of VisNIR-DRS to characterize soil carbon, but again, the soils evaluated were largely mineral soils. McWhirt et al. (2012) used a single sensor approach (VisNIR-DRS) to characterize the organic matter content of compost products. Whilst organic, composted products differ substantively in their physicochemical composition from that of organic soils. By contrast, the present study explicitly aims to evaluate the combined use of both proximal sensors (PXRF and VisNIR-DRS) in characterization of organic soil horizons in variable states of decay. As such, the objectives of this study were to: 1) quantify total carbon and nitrogen in natural organic soils by VisNIR-DRS and PXRF individualistically and, 2) explore if there is a benefit in predictive accuracy from concatenating VisNIR-DRS spectra and PXRF elements. We hypothesize that total carbon and nitrogen of largely organic horizons can be directly predicted from the reported PXRF elements and VisNIR-DRS spectra. We further hypothesize that either a fused model or model averaging approach will produce better predictability than either the VisNIR-DRS or PXRF approach independently.

2. Materials and Methods

2.1. General occurrences and features

In sum, 136 organic laden samples from non-saturated, uplands were collected in Italy and United States of America (Texas and New Mexico) during 2014 and 2015; a few mineral laden soils were also included as part of this dataset as a link to previously established work on mineral soils. The sites differed substantively in their geological composition, soil development, climate, and vegetation.

In Italy, a total of 39 organic horizons were collected from forest soils across three different sites on the Apennines chain (central Italy): Mount Acuto, Mount San Vicino, and Mount

Terminillo. The soils developed from limestone of different geological origin: Mount Acuto is characterized by limestone (Lower Cretaceous - Aptiano), Mount San Vicino is grey limestone with traces of flintstone and marl from the Jurassic (Lias) Pliensbachian Sinemurian and Mount Terminillo is grey limestone with trace amounts of flintstone (Jurassic Toarcian-Sinemurian) (ISPRA, 2015). The soils of these areas are classified as Mollisols or Inceptisols (Soil Survey Staff, 2014), characterized by a mesic soil temperature regime (10°C to 12°C) along with an udic soil moisture regime (from 825 mm to 1430 mm precipitation). In the three areas the cover vegetation was mainly composed of *Fagus sylvatica* from 80 to 99%, with *Carpinus betulus* at Mount Acuto, *Quercus cerris*, *Castanea sativa,* and *Sorbus aria* at Mount San Vicino, *Laburnum anagyroides* and *Acer* spp. at Mount Terminillo.

In Texas, 16 alluvial organic samples were collected in backwater areas along the North Fork of the Brazos River in Lubbock County in major land resource area (MLRA) 77C - Southern High Plains - Southern Part (Soil Survey Staff, 2006). Soils of this MLRA are generally developed by eolian deposits in the Blackwater Draw Formation of Pleistocene age, classified as Alfisols, Inceptisols, Mollisols, and Vertisols, and have a thermic soil temperature regime (13°C to 17°C) and an ustic soil moisture regime (from 405 to 560 mm precipitation). Mostly short and mid prairie grasses and scanty tree and shrubs (e.g., *Bouteloua gracilis*, *Bouteloua dactyloides*, *Bouteloua curtipendula)* are prevalent. Separately, 27 various organic horizons were sampled in forested areas of the George Bush Intercontinental Airport, WG Jones State Forest, San Jacinto River, and Sam Houston National Forest; all generally in the vicinity of Houston, Texas. The WG Jones State Forest, San Jacinto River, and Sam Houston National Forest occur in MLRA 133B - Western Coastal Plain (Soil Survey Staff, 2006), where soils developed from Tertiary and Cretaceous marine sediments consisting of inter bedded sandstone, siltstone, shale and loose primary particles. In particular, the Reklaw and Weches Formations in the Claiborne Group form the Redland area of East Texas. The main soil orders in this MLRA are Alfisols and Ultisols with a thermic soil

147 temperature regime (16°C to 20°C), an udic or aquic soil moisture regime (990 to 1,600 mm
148 precipitation). Vegetation of the area is typified by pine-hardwood species such as *Pinus taeda*,
149 *Pinus echinata*, *Liquidambar styraciflua*, *Quercus falcate,* and *Cornus florida*; *Callicapra*
150 *americana*, and *Smilax* spp. are common in the woody understory. *Schizachyrium scoparium* and
151 *Bothriochloa barbinodis* are the dominant herbaceous species. Of the 27 samples collected in this
152 area, four were collected in densely wooded pine forests adjacent to George Bush Intercontinental
153 Airport, which is a few kilometers beyond the aforementioned MLRA boundary, but quite similar in
154 the organic horizons sampled.

155 In New Mexico, 54 samples were collected near the periphery of the Lincoln National Forest of
156 Lincoln County; the horizons sampled were dominantly organic, but a few transitioned into organic
157 laden mineral soils. The area is in MLRA 39-Arizona and New Mexico Mountains (Soil Survey
158 Staff, 2006). The area is characterized by Cenozoic volcanic rock and various sedimentary sections
159 of the Colorado Plateau. The southern and eastern parts contain Permian and Cretaceous
160 sedimentary rock over a Precambrian granite core. Main soil orders of this MLRA are Inceptisols,
161 Mollisols, Alfisols, and Entisols, with mainly frigid or mesic soil temperature regimes (2°C to 13°C)
162 and an ustic/udic moisture regime (358 mm to 760 mm precipitation). *Pinus ponderosa* occurs is the
163 dominant vegetation in the low and intermediate heights, while at higher altitudes, *Picea abies* and
164 *Pseudotsuga menziesii* are commonplace.

165 *2.2. Field Sampling*

166 For maintaining an extensive variation in terms of physicochemical characteristics, composition
167 and origin, samples were randomly collected within morphologically established, organic laden
168 horizons. Almost all samples were classified as O horizons according to Référentiel pédologique
169 (2008). The different stages of degradation were classified as: i) OL, consisting of leaf debris with
170 little or no degradation (<10% of fine organic matter) whereby the botanic origin was easily
171 recognized; sub-horizons were recognized as OLn, fresh litter with minimal degradation, and OLv,

where the plant material were subject to initial degradation such as changes in color, volume, and inter-particle linkages; and ii) OH, with >70% of fine organic (humic) material following extensive decomposition (Table 1). The latter is homogenous, massive, and reddish-brown to black in color with an abundance of fine roots. These three states of decomposition are roughly akin to fibric, hemic, and sapric materials as defined by the Soil Survey Staff (2014), respectively.

2.3. Laboratory characterization and VisNIR scanning

Soil samples were air dried and ground to < 1-mm prior to chemical analyses and spectral scanning. Total C (TC) and total N (TN) were determined by a TruSpec® dry combustion analyzer (LECO Corp., MI, USA) according to Dumas method combustion (Nelson and Sommers, 1996).

Spectral reflectance values of air dried and ground soil samples were measured proximally over the VisNIR region (350-2500 nm) by a portable PSR-3500® VisNIR spectroradiometer (Spectral Evolution, USA). The reflectance data were resampled to 1 nm output values. Scanning was done with a contact probe containing a 5W halogen lamp, minimizing errors associated with stray light during measurements. Each sample was uniformly tiled in a glass petri plate and scanned four times, physically repositioning the probe prior to each scan. The mean of 10 internal scan over 1.5 seconds produced one individual scan. The spectroradiometer was standardized after every two samples using a NIST certified white reference. The average spectral curve was calculated and further used for spectral preprocessing and subsequent predictive modeling.

Necessary treatments of raw average reflectance spectra were done in R version 2.11.0 (R Development Core Team, 2014) following the spline fitting methods outlined in Wang et al. (2015). The spectral preprocessing involved converting reflectance to absorbance by log(1/R) which was executed in the Unscrambler®X 10.3 software (CAMO Software Inc., Woodbridge, NJ).

2.4. Scanning with PXRF

A DP-6000 Delta Premium portable X-ray fluorescence (PXRF) spectrometer (Olympus, Waltham, MA, USA) facilitated sample scanning. Configured with a Rh x-ray tube, the instrument was operated at 15–40 kV; an integrated ultra-high resolution (165 eV) silicon drift detector quantified each element detected. Before scanning, a 316 alloy clip was used to standardize the instrument. Primary analysis was conducted in Soil Mode (three beams of 30 s each); it can detect the following: Ag, As, Ba, Ca, Cd, Cl, Co, Cr, Cu, Fe, Hg, K, Mn, Mo, Ni, P, Pb, Rb, S, Sb, Se, Sn, Sr, Ti, V, Zn, and Zr. A second scanning was performed with Geochem Mode (two beams of 30 s each) in order to measure Mg, S, Al, and Si. Geochem and Soil Mode scans were done in duplicate, with the spectrometer physically repositioned between each scan. Sample homogeneity was ensured during the grinding step. Data were then averaged between scans to obtain a mean of elemental data for each sample. Data quality was ensured via the scanning of two NIST certified reference samples, with recovery percentage calculated as PXRF reported vs. NIST certified values. Results were as follows (PXRF reported/NIST certified [recovery]): Zn 4,263/4,180 mg kg^{-1} [1.02]; Cu 3,559/3,420 mg kg^{-1} [1.04]; K 27,081/21,700 [1.25]; Ca 9,119/9,640 mg kg^{-1} [0.95]; Ti 3,540/3,110 mg kg^{-1} [1.14]; Mn 2,237/2,140 [1.05]; Fe 50,574/43,200 mg kg^{-1} [1.17]; As 1,639/1,540 mg kg^{-1} [1.06]; Sr 269/255 [1.05]; Pb 5608/5520 [1.02].

2.5. Data mining

We performed all statistical modeling via R version 2.11.0 (R Development Core Team, 2014) software. We also checked the normality of residuals by the Shapiro-Wilk test at a 5% significance level. In the present study, both original TC and TN values were negatively skewed (Pearson skewness coefficient -1.78 and -0.37 for TC and TN, respectively), while Box–Cox conversion (Box and Cox, 1964) using $\lambda=0$ ($\log_{10}$ transformed) was unable to conform normal distribution. Principal component analysis (PCA) was executed via R version 2.11.0 using the 'prcomp' function to visualize spectral behavior of soil samples from different geographical areas and organic horizons. Optimal number of principal components (PC) was decided from a *screeplot*.

In this study, the whole dataset was split into 70% calibration (n=96) and 30% validation set (n=40) using the Kennard-Stone algorithm (Kennard and Stone, 1969), which is an adaptive procedure to select the most representative samples based on Euclidean distance. Calibration samples were used to establish a prediction model, whereas validation samples were used to assess the model's predictive ability (Chen et al., 2015). Initially, we targeted both TC and TN with VisNIR-DRS spectra only via partial least squares regression (PLS), elastic net regression (ENET), penalized spline regression (PSR), and random forest regression (RF) (Breiman, 2001; Guyon et al., 2002; Viscarra Rossel et al., 2006b). Additionally, both TC and TN were predicted with PXRF elemental data via PLS, ENET, and RF models.

Regularized methods, which play an important role in both statistical and data mining problems, can be described as Eq. 1:

$$\hat{\beta}(\lambda) = arg\min_{\beta} L(Y, X\beta) + \lambda J(\beta) \quad (1)$$

where $L(Y, X\beta)$ is a non-negative loss function, $J(\beta)$ is a non-negative penalty on the model complexity and λ is the non-negative tuning parameter. The key idea for regularized methods is to balance between the goodness-of-fit on the training data and the complexity of the model. Usually, a complicated model always shows a good fit to the calibration data. Nonetheless, this commonly leads to problems of overfitting, where the model is too adapted to the training data and often has a poor prediction performance on new samples.

Elastic net, devised by Zou and Hastie (2005) is a regularized regression method that linearly combines the ridge penalty and the least absolute shrinkage and selection operator (LASSO) penalty (Eq. 2)

$$\hat{\beta}(\lambda_1, \lambda_2) = arg\min_{\beta} SSE + \lambda_1 \sum|\beta_j| + \lambda_2 \sum \beta_j^2 \quad (2)$$

where SSE is the sum of squared errors, $\sum|\beta_j|$ is the LASSO penalty, and $\sum \beta_j^2$ is the ridge penalty. In high dimensional data analysis where many predictors are correlated to each other, the LASSO penalty tends to pick a few of the predictors and discard the others (sparse model), while the ridge

245 penalty shrinks the coefficients of correlated predictors towards each other (dense model). The
246 ENET penalty combines these two penalties so that when the predictors are correlated in groups, it
247 will produce a sparse model with good prediction performance, while boosting a grouping effect.
248 Another advantage of ENET is that it can handle the high dimension and low sample size problem.
249 In this study, ENET was run using the 'glmnet' package in R, produced by Friedman et al. (2015).
250 Details of PSR and RF methodology are summarized in Wang et al. (2015).

251 *2.6. Model Fusion and Model Averaging*

252 Model fusion and model averaging (model ensemble) were tested to determine if combination
253 of the predictions for the VisNIR-DRS and PXRF models into a single composite score could
254 improve predictive accuracy of TC and TN. Initially, three fused modeling approaches (PSR+RF),
255 (PLS+RF) and (ENET+RF) were employed where PSR/PLS/ENET were used to fit the training set
256 (containing VisNIR-DRS spectra only). Next, we ran the RF using the PSR/PLS/ENET residual as
257 the response and PXRF elements as the predictors (Chakraborty et al., 2015). Succinctly, the
258 prediction value on the validation set using (PSR+RF), (PLS+RF) and (ENET+RF) contains two
259 additive parts: the prediction from PSR/PLS/ENET model using the VisNIR-DRS spectra plus the
260 prediction from RF model using the PXRF data. An outline of this fused procedure is shown in Fig.
261 1.

262 For model averaging we used Granger-Ramanathan averaging (GRA) (Granger and
263 Ramanathan, 1984) with some modifications. The GRA approach requires fitting a multivariate
264 linear regression model where lab measured soil property values are regressed against the
265 corresponding predictions derived from VisNIR-DRS and PXRF models. Initially, the
266 PSR/PLS/ENET model was fitted on the calibration set (n=96) using the VisNIR-DRS spectra.
267 Subsequently, the RF model was fitted using only PXRF elements as predictors and TC or TN as
268 the response. Note that this RF model was not the same as the RF model in fused approaches

269 (PSR+RF, PLS/RF, or ENET+RF) as the latter used residual as a response. Next, a bivariate linear
270 regression was fitted with Eq. 3:

271 $$Y = a + b*X + c*Z \quad (3)$$

272 where Y is either TC or TN on calibration set; X is the fitted value (on calibration set) from the
273 PSR/PLS/ENET model; Z is the fitted value (on the calibration set) from the RF model; *a* is the
274 intercept; *b* is the estimated model weight for the PSR model; and *c* is the estimated weight for the
275 RF model. To predict the validation set (n=40), we first generated the prediction value, X*, from
276 the PSR/PLS/ENET model, and prediction value, Z*, from the RF model. The GRA prediction is
277 given as Eq. 4:

278 $$Y^* = a + bX^* + cZ^* \quad (4)$$

279 We followed the Kennard-Stone splitting scheme for both fused models and GRA.

280 *2.7. Whole geographical area and organic horizon holdout validation*

281 Further, whole geographical area and organic horizon holdout validations were executed for
282 both TC and TN to determine how area to area and horizon to horizon heterogeneity affected
283 prediction accuracies (Brown el al., 2005). Whole-geographical area holdouts were achieved by
284 calibrating a model using the best performing algorithm with three geographical areas and then
285 validated using the fourth area. Moreover, whole-organic horizon holdouts were achieved by
286 calibrating a model using the best performing algorithm with two organic horizons and then
287 validated using the third organic horizon.

288 In this study, the root mean square error (RMSE), regression coefficient (R^2), residual prediction
289 deviation (RPD) (Eq. 5), bias, and ratio of performance to inter-quartile distance (RPIQ) (Eq. 6)
290 were used to evaluate model performance (Gauch et al., 2003 Bellon-Maurel et al., 2010). RPD

based model accuracy classification scheme devised by Chang et al. (2001) was followed for evaluating model accuracy.

$$RPD = \left[\frac{1/(n-1)\sum_{i=1}^{n}(Yobs-Ymean)^2}{1/n\sum_{i=1}^{n}(Yobs-Ypred)^2}\right]^{0.5}_{Validation} \quad (5)$$

$$Bias = \sum_{i=1}^{n}(Ypred - Ymean)/n \quad (6)$$

where, *Yobs* and *Ypred* are the observed and predicted response variables, respectively, *Ymean* is the average of the *Yobs* values, and *n* denotes the sample number in the validation data set. RPIQ was defined as IQ/SEP, where SEP represents the standard error of prediction, and IQ denotes the interquartile distance of the validation set (IQ=Q3–Q1). In this study, optimal model performance featured low RMSE values along with high RPD, R^2, and RPIQ.

3. Results and Discussion

3.1. Soil properties and PCA

Statistical moments for TC and TN measured on the soil samples used in the regression models are listed in Table 2. Soil TC and TN showed negatively skewed distribution with mean concentrations of 39.05% and 1.10%, respectively for the whole dataset; 39.53% and 1.08%, respectively for the calibration set; and 37.90% and 1.14%, respectively for the validation set. Of the 13 elements obtained by the PXRF soil mode, only nine [Zn, S, K, Ca, Ti, Rb, Mn, Fe, Sr] were suitable for use in the multivariate predictive models with continuous data across all samples scanned. Within-region variability of TC and TN exhibited substantial variation. For instance, in Italy sample TC ranged from 28.06% to 45.94%. Showing similar variation, Lubbock samples showed TC ranging from 25.93% to 41.20%. Comparatively higher variability was observed for Houston samples (11.42-48.54%), while maximum TC variability was observed in samples from New Mexico region (0.06-51.30%). Lubbock samples showed TN ranging from 1.13-1.60%. Comparatively higher variability was observed for Italy (0.80-1.87%) and Houston samples (0.30-1.47%). Following a similar trend of TC, maximum variability in TN was observed from New

315 Mexico soils (0.04-1.87%). Overall, variability in estimated TC and TN was due largely to the
316 variability in the parent material, climate, C source, land use, and vegetation cover. However, the
317 extent of variability appeared related with the range of the mean annual soil temperature of the
318 considered sampling area, the highest in the New Mexico soils (2-13°C), the lowest in the Italy soils
319 (10-12°C), intermediate in the other two subsets (16-20°C in the Houston soils, and 13-17°C in the
320 Lubbock soils).

321 Levene's test yielded equality of variance of TC and TN values among the training and test sets
322 ($p=0.65$ and 0.55 for TC and TN, respectively). Moreover, t-test could not establish a significant
323 difference between mean TC ($p=0.53$) and TN ($p=0.51$) for these two data sets. These results
324 indicated that the validation samples based on the Kennard-Stone method can properly represent the
325 studied population. The first two leading PCs constituted over 95% of the spectral variation (Fig.
326 2a). Although some overlapping among samples from four regions were discernible from the PC1
327 vs. PC2 plot, these regions had different range, shape, and distributions in the spectral space,
328 indicating variable SOM composition from different organic input sources (Fig. 2b). Moreover, the
329 PCA plot indicating samples from three different organic horizons (OH, OLn and OLv) (Fig. 2c)
330 indicated that OH horizon samples were quite different from the samples from OLn and OLv
331 horizons. The OH horizon samples tended to have larger values on the PC1 scores. Contrarywise,
332 the samples from the OLn and OLv horizons were relatively close. Notably, the samples from the
333 OLv horizon tended to have relatively larger PC2 values while there was an obvious outlying
334 sample with the largest PC1 value and smallest PC2 value.

335 *3.2. Validation results and model parsimony*

336 In fused models, both TC and TN were estimated with good RPIQ values, better than using an
337 individual instrument or GRA (Table 3), which has been noted before elsewhere (Wang et al., 2015).
338 Fused (PLS+RF) and (PSR+RF) models yielded the highest RPDs for TC (2.84) and TN (2.62),
339 respectively. While RPD values for both tested soil properties dropped in (ENET+RF) fused models,

they still explained nearly 80% of the TC and TN variability. Overall, all three fused models of TC showed similar accuracies (Table 3), as reported for soil organic C in the review of Soriano-Disla et al. (2014) where the median R^2 of validations was calculated as 0.83 (33 studies).

Wavelength selection not only enhances the stability of the model but also makes the model more parsimonious. Although (PSR+RF) yielded the lowest RMSE for TN (0.160%), both PSR and RF are not parsimonious models since they do not select the subset of important variables. Although for RF it is possible to select different *mtry* values, which is the size of the candidate subset for each splitting, models with different *mtry* values still select all the variables. Despite controlling the smoothness of the neighboring coefficients through λ, PSR is also unable to select important wavelengths. Figure 3 illustrates parsimony features of ENET for TC (Fig. 3a) and TN (Fig. 3b) where cross-validation error, test error (mean squared error or MSE), and the degrees of freedom (DF) (the number of selected wavelengths in the ENET model) were plotted against log of λ which is the parameter to tune the ENET parsimony. Note that the larger the λ, the more parsimonious the model. The vertical dash line shows the optimal λ that minimizes the cross-validation error. The bottom plot shows the ENET coefficient plot against the VisNIR wavelengths. For both soil properties, most of the coefficients were zero, indicating that they were not selected into the model. Additionally, based on the DF plots, it was evident that the optimal TC and TN models selected only ~300-350 out of 2151 wavelengths. Conversely, using a fused (PLS+RF) model on the reduced feature sets (14 and 13 latent factors for TC and TN, respectively) returned the best result for TC and slightly higher RMSE for TN (0.191%) than (PSR+RF) (0.160%) (Table 3). Based on these observations, we conclude that while considering both model precision and parsimony, (PLS+RF) emerged as the best model for both TC and TN.

For both TC and TN, while utilizing VisNIR-DRS or PXRF alone, RF and PLS models yielded the largest validation RMSE values, respectively, and were therefore the least precise (Table 3). In general, PXRF measurements were poorly correlated to both TC and TN, suggesting its limitation

365 to make full geochemical assessment alone due to its inability to handle low concentrations and
366 light elements ($Z \leq 11$) (Weindorf et al., 2008). Among nine elements used in predictive models,
367 only Al (1.48%), Ca (2.53%) and Si (4.83%) were present in high concentrations, while the other
368 six elements were present at low levels (<1%). Thus, it seemed prudent to calibrate using PXRF
369 spectra instead of elements if high accuracy is desired. While GRA substantially worsen the
370 validation results for TC, (PSR+RF) model averaging for TN yielded close RMSE (0.197%) to
371 those produced by fused models. The reasonably good performance of PSR in all three approaches
372 (VisNIR only, fused, and GRA) can be attributed to the fact that in case of 'fat' data (large number
373 of dimensions and small sample size) PSR uses all samples and smooths constraints on the
374 coefficient. Therefore, PSR often works well on signal regression problems which are usually
375 strongly dimensional and have a relatively small sample size.

376 As coefficients of determinations of TN were rather similar to those of TC using VisNIR-DRS
377 in isolation, we conclude that VisNIR-DRS sensed a grouping of soil components comprising
378 organic functional groups that contain organic N fractions (Wang et al., 2015) (Table 3). Moreover,
379 both TC and TN are spectrally active components or chromophores, which absorb incident energy
380 at discrete energy levels and show broad and weak absorption features in the VisNIR region (Ben-
381 Dor, 2011) Although intense VisNIR-DRS spectral bands cannot be directly ascribed to metals or
382 other components, Song et al. (2012) revealed that metals can interact with chromophores such as
383 soil C. However, S was negatively correlated with TC (ρ=-0.48); likely since S is very light with an
384 atomic mass of 16 and was present in low concentrations (max=0.43%).

385 Plots of observed *vs*. model predicted (VisNIR-DRS only and fused) TC and TN values are
386 presented in Figs. 4 and 5, respectively. In general, all models showed overestimation at lower TC
387 or TN values and underestimation at higher values. Several of these overestimations occurred
388 because of the relative scarcity of observations with low values, which was expected since the
389 samples came dominantly from organic horizons. Another possible explanation for TC and TN

prediction errors is undecomposed organic matter and variable C sources, as noted before elsewhere (Henderson et al., 1992). Indeed, the VisNIR-DRS spectra for SOM have not been fully defined yet, because of the complexity or unclear definitions of these materials (Brown et al., 2006). Probably the TC and TN predictability could have been enhanced by using soil management or vegetation specific models (Morgan et al., 2009); nonetheless, exploring this notion was beyond the scope of this study.

3.3. Model fusion vs. model averaging

It was apparent that GRA model averaging by combining the VisNIR-DRS and PXRF predictions for consolidated use could not produce better accuracy than model fusion. In practice, "good" model averaging should contain the models that complement each other. In fused models, RF was fitted based on the PLS/PSR/ENET residuals. The sequential fitting in PLS/PSR/ENET + RF allowed the RF model to complement the former. Conversely, combining PLS/PSR/ENET and RF in parallel fashion through GRA (namely, separately on VisNIR-DRS and PXRF) produced incompatibility. Succinctly, they perhaps made the same mistake on the same sample. To justify this postulation and clearly visualize the prediction improvement by using the fused models, Fig. 6 represents a scatterplot matrix produced in R using the *spm* function in the car library, taking the PLS+RF model as an example. The diagonal elements are the density plots for the four competitors [observed TC values (OBS), RF in (PLS+RF) fused model (RF1), and RF in GRA (RF2)]. For example, the upper left one is the density plot of TC while the tickmarks at the bottom axis show the observed values in the data. The off-diagonal elements are the pairwise scatter plots of four competitors, together with the best linear and nonlinear smoothers. For example, Fig. 6e shows the scatter plot of observed TC (on the vertical axis) and predicted TC values from PLS model (on the horizontal axis) using spectral data. The green and red solid lines are the fitted linear regression line and the loess smoother (a popular nonlinear smoother using local linear regression) fit, respectively. The red dash lines represent one standard error above and below the estimated function. It can be

observed that PLS tended to underestimate the samples with TC values between 30 to 45% and overestimate the samples with TC values < 20% through the bending shape of the red loess smoother (Fig. 6e). Interestingly, while RF1 "corrected" the PLS model by lifting the underestimated area (30-45% TC) and lowering the overestimated region (<20% TC) (Fig. 6i), this was not totally unexpected owing to the non-linear and contingent relationships among VisNIR-DRS reflectance and soil constituents (Brown et al., 2006). However, RF2 made the same "mistake" as PLS by underestimating and overestimating in the same regions (Fig. 6d). These trends clearly explained why sequentially fitting (PLS+RF) in fused model improves prediction accuracy relative to parallel fitting of PLS and RF through GRA. Note that, the same argument is applicable to (PSR+RF) or (ENET+RF) in predicting TC and TN. Model averaging even could not outperform VisNIR-DRS only model performance (Table 3). Thus, the postulations of Abbott (2014), who reported that model averaging nearly always improves model predictive accuracy and rarely predicts worse than single models, may not be generalized.

3.4. Relative improvement of soil property predictions after model fusion

Table 4 demonstrates that fused models had relative improvements on the prediction accuracy of TC and TN from GRA or VisNIR-DRS and PXRF used in isolation. Notably, RMSE improvements of 10-47% for TC and 10-67% for TN were achieved by the PLS+RF and PSR+RF model, respectively, demonstrating the potential of the synergistic use of VisNIR-DRS and PXRF for the estimation of soil attributes. On average, the PLS+RF model for TC and the PSR+RF model for TN produced almost a one fourth reduction in RMSE from using VisNIR-DRS in isolation. While the PLS+RF fused model increased RPIQ values ranging between 11 to 97% for TC, the PSR+RF model for TN yielded three times higher RPIQ than using PXRF in isolation. Also, the PSR+RF approach for TC produced substantively better results (R^2 0.87; RMSE 3.01%; RPD 2.75) than those obtained by McWhirt et al. (2012) (R^2 0.82; RMSE 10.1%; RPD 1.72) who used VisNIR-DRS in isolation when evaluating the organic fraction of composted products.

3.5 Whole geographical area and organic horizon holdout validations

Holding out a whole geographical area or organic horizon reduced prediction accuracies in all cases where the soils of the geographical area or horizon held out were not represented by another area or horizon, creating extrapolation (Tables 5 and 6). Validation with an individual geographical area or horizon exhibited increases in RMSEs and a reduction in RPDs relative to the best performing models (PLS+RF and PSR+RF for TC and TN, respectively) using all 136 samples. As expected, when principal component analysis was implemented on the VisNIR spectrum only, it clearly demonstrated different spectral behaviors of the four geographical areas (Fig. 2b). Further, while executing Principal Component Regression (PCR, data not shown) to calculate the actual number of PCs required to represent VisNIR spectral variability, results indicated the need for at least 12 PCs in PCR based on full cross validation. In other words, although some overlap among the four areas was apparent in the PC1 vs. PC2 plot, that only reflected the first two PCs. Summarily, it can be concluded that since the four geographical areas had different VisNIR spectra, using all samples in a single prediction model is preferred over the area-holdout scheme since the latter involves extrapolation in predicting samples from different areas.

Moreover, based on the whole organic horizon-holdout results, all models predicted poorly on holding out OH and OLn samples. For OH, the main reason of poor model performance was that OH samples were quite different from OLn and OLv samples because of variable rates of decomposition. Hence the fitted models using OLn and OLv samples could not predict OH samples well. Conversely, since the OLn horizon represented most of the samples (i.e 64 out of 136 samples), the sample size for fitting the model using the OLv and OH was the smallest. In addition, the OH samples in the training data made the fitted model poor in predicting OLn samples in the validation set.

3.6 Some practical concerns

464 The fused PSR+RF model had some limitations too. Initially, PSR was applied on VisNIR-DRS
465 spectra with the assumption that the underlying PSR coefficient curve is smooth. Secondly, RF was
466 used to fit the residuals of the PSR model on the PXRF elemental data assuming that both VisNIR-
467 DRS and PXRF data contain some information on the response variable. In particular, PXRF
468 contains some extra information about TC and TN that were not explained (or covered) by the
469 VisNIR-DRS spectra. We have examined the feasibility of the approach in this research and have
470 provided a preliminary contribution to the issue of validation, however, further intensive research is
471 recommended to confirm our results. Nonetheless, the combined use of PXRF and VisNIR data
472 show strong potential for the accurate assessment of soil organic horizons. The spectral and
473 elemental data produced by proximal sensor analysis has previously been shown to be useful in the
474 prediction of a wide number of physicochemical properties in mineral soils. With regard to TC and
475 TN content of organic soil horizons, similar impressive predictability from proximal sensor data
476 was observed.
477
478 **4. Conclusions**
479 Summarily, this preliminary study of 136 organic laden soil samples from non-saturated uplands
480 of Italy and the United States were evaluated by standard laboratory analysis as well as proximal
481 sensor (PXRF + VisNIR DRS) scans. Our results show that:
482 (i) Combining model outcomes using model fusion improved TC and TN prediction accuracies
483 relative to using an individual instrument (PXRF or VisNIR DRS) or model averaging. Overall, the
484 relative improvement in % RMSE ranged from 10-47% and 10-67% for TC and TN, respectively.
485 (ii) Considering both model precision and parsimony, PLS+RF emerged as the best model for
486 predicting both TC and TN in organic laden soil samples. Conversely, while using all spectral
487 variables, PSR+RF yielded the best model results (lowest RMSE and highest RPIQ) for TN.

(iii) Specifically, GRA model averaging by combining the VisNIR-DRS and PXRF predictions could not produce better accuracy than model fusion or individual instrument use, possibly due to model incompatibility while fitting them in parallel fashion.

(iv) The poor predictability with unreliable RPIQs while using PXRF in isolation underlines the need for using PXRF spectra instead of elemental data.

Acknowledgments

The authors are grateful for support from the BL Allen Endowment in Pedology at Texas Tech University in completing this research.

References

Abbott, D. 2014. Applied Predictive Analytics. Principles and Techniques for the professional data analyst. John Wiley & Sons, Inc., Indianapolis, IN. p. 307-326.

Aldabaa, A.A.A., Weindorf D.C., Chakraborty S., Sharma A., Li, B., 2015. Combination of proximal and remote sensing methods for rapid soil salinity quantification. Geoderma 239-240, 34-46.

Association française pour l'Etude du sol. 2008. Référentiel pédologique. Ed. Quae RD 10, 78026 Versailles Cedex, France.

Ball, B.A., Carrillo, Y., Molina, M., 2014. The influence of litter composition across the litter-soil interface on mass loss, nitrogen dynamics and the decomposer community. Soil Biol. Biochem. 69, 71–82.

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. Trends Analyt. Chem. 29, 1073–1081.

Ben-Dor, E., 2011. Characterization of soil properties using reflectance spectroscopy, in Thenkabail, P.S., Lyon, J.G., Huete, A. (Eds.), Hyperspectral Remote Sens. Veg. CRC Press. 513– 558.

Berg, B., McClaugherty, C., 2008. Plant Litter - Decomposition, Humus Formation, Carbon Sequestration. Springer Berlin Heidelberg, Berlin, Heidelberg.

Berger, T.W., Neubauer, C., Glatzel, G., 2002. Factors controlling soil carbon and nitrogen stores in pure stands of Norway spruce (Picea abies) and mixed species stands in Austria. For. Ecol. Manage. 159, 3–14.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc. Series B Stat. Methodol. 26, 211–252.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Bronick, C.J., Lal, R., 2005. Soil structure and management: a review. Geoderma 124, 3–22.

523 Brown, D.J., Bricklemyer, R.S., Miller, P.R., 2005. Validation requirements for diffuse reflectance
524 soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma
525 129, 251-267.
526 Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil
527 characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132, 273–290.
528 Carrillo, Y., Ball, B.A., Strickland, M.S., Bradford, M.A., 2012. Legacies of plant litter on carbon
529 and nitrogen dynamics and the role of the soil community. Pedobiologia (Jena) 55, 185–192.
530 Chakraborty, S., Weindorf, D.C., Li, B., Aldabaa, A.A.A., Ghosh, R.K., Paul, S., Nasim Ali, M.,
531 2015. Development of a hybrid proximal sensing method for rapid identification of petroleum
532 contaminated soils. Sci. Total Environ. 514, 399–408.
533 Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance
534 Spectroscopy: Principal Components Regression Analyses of soil properties. Soil Sci. Soc. Am.
535 J. 65, 480.
536 Chang, C.-W., Laird, D.A., 2002. Near-Infrared Reflectance spectroscopic analysis of soil C and N.
537 Soil Sci. 167, 110–116.
538 Chang, C.-W., You, C.-F., Huang, C.-Y., Lee, T.-Q., 2005. Rapid determination of chemical and
539 physical properties in marine sediments using a Near-Infrared Reflectance spectroscopic
540 technique. Appl. Geochemistry 20, 1637–1647.
541 Chen, T., Chang, Q., Clevers, J.G.P.W., Kooistra, L., 2015. Rapid identification of soil cadmium
542 pollution risk at regional scale based on Visible and Near-Infrared spectroscopy. Environ. Pollut.
543 206, 217–26.
544 Fajardo, M., McBratney, A., Whelan, B., 2015. Fuzzy clustering of Vis–NIR spectra for the
545 objective recognition of soil morphological horizons in soil profiles. Geoderma 263, 244-253.
546 Friedman, A.J., Hastie, T., Simon, N., Tibshirani, R., Hastie, M.T., 2015. Lasso and Elastic-Net
547 Regularized Generalized Linear Models. Available online at: https//cran.r-

548 project.org/web/packages/glmnet/glmnet.pdf. (Verified on July 29, 2015).
549 García-Palacios, P., Maestre, F.T., Kattge, J., Wall, D.H., 2013. Climate and litter quality
550 differently modulate the effects of soil fauna on litter decomposition across biomes. Ecol. Lett.
551 16, 1045–53.
552 Gauch, H.G., Hwang, J.T.G., Fick, G.W., 2003. Model evaluation by comparison of model-based
553 predictions and measured values. Agron. J. 95, 1442-1446.
554 Ge, X., Zeng, L., Xiao, W., Huang, Z., Geng, X., Tan, B., 2013. Effect of litter substrate quality and
555 soil nutrients on forest litter decomposition: A review. Acta Ecol. Sin. 33, 102–108.
556 Gosz, J.R., Likens, G.E., Bormann, F.H., 1976. Organic matter and nutrient dynamics of the forest
557 and forest floor in the Hubbard Brook forest. Oecologia 22, 305–320.
558 Granger, C.W.J., Ramanathan, R., 1984. Improved methods of combining forecasts. J. Forecast. 3,
559 197–204.
560 Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using
561 support vector machines. Mach. Learn. 46, 389–422.
562 Henderson, T.L., Baumgardner, M.F., Franzmeier, D.P., Stott, D.E., Coster, D.C., 1992. High
563 dimensional reflectance analysis of soil organic matter. Soil Sci. Soc. Am. J. 56, 865-872.
564 Hettipathirana, T.D., 2004. Simultaneous determination of parts-per-million level Cr, As, Cd and Pb,
565 and major elements in low level contaminated soils using borate fusion and energy dispersive X-
566 ray fluorescence spectrometry with polarized excitation. Spectrochim. Acta Part B At. Spectrosc.
567 59, 223–229.
568 Horta, A., Malone, B., Stockmann, U., Minasny, B., Bishop, T.F.A., McBratney, A.B., Pallasser, R.,
569 Pozza, L., 2015. Potential of integrated field spectroscopy and spatial analysis for enhanced
570 assessment of soil contamination: A prospective review. Geoderma 241-242, 180–209.
571 Hu, X.-Y., 2013. Application of Visible/Near-Infrared Spectra in modeling of soil total phosphorus.
572 Pedosphere 23, 417–421.

573 Huttl, R.F., Schaaf, W., 1997. Magnesium deficiency in forest ecosystems, first ed. Springer,
574 Netherland.
575 ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale), 2015. Carta geologica d'Italia.
576 URL http://www.isprambiente.gov.it/Media/carg/ (Verified on October 28, 2015).
577 Kaiser, M., Ellerbrock, R.H., Gerke, H.H., 2008. Cation exchange capacity and composition of
578 soluble soil organic matter fractions. Soil Sci. Soc. Am. J. 72, 1278.
579 Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137-
580 148.
581 McWhirt, A.L., Weindorf, D.C., Chakraborty, S., Li, B., 2012. Visible near infrared diffuse
582 reflectance spectoscopy (VisNIR DRS) for rapid measurement of organic matter in compost.
583 Waste Management & Research 30, 1049-1058.
584 Morgan, C.L.S., Waiser, T.H., Brown, D.J., Hallmark, C.T., 2009. Simulated in situ
585 characterization of soil organic and inorganic carbon with Visible Near-Infrared Diffuse
586 Reflectance spectroscopy. Geoderma 151, 249–256.
587 Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter, in Sparks,
588 D.L. (Ed.). Methods of soil analysis - Part 3. Chemical Methods. Soil Sci. Soc. Am., Madison,
589 WI.
590 Paulette, L., Man, T., Weindorf, D.C., Person, T., 2015. Rapid assessment of soil and contaminant
591 variability via Portable X-ray Fluorescence spectroscopy: Copşa Mică, Romania. Geoderma 243-
592 244, 130–140.
593 Qafoku, N.P., 2015. Climate-Change effects on soils: accelerated weathering, soil carbon, and
594 elemental cycling. Chapter Two. Adv. Agron. 131, 111–172.
595 R Development Core Team, 2008. R: a language and environment for statistical computing . R
596 Found. Stat. Comput. Vienna, Austria (Available online with Updat at http//www.cran.r-
597 project.org, Verified on January 6, 2016).

598 Rossel, R.A.V., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance
599 spectra. Geoderma 158, 46–54.
600 Sharma, A., Weindorf, D.C., Man, T., Aldabaa, A.A.A., Chakraborty, S., 2014. Characterizing soils
601 via Portable X-ray Fluorescence spectrometer: 3. Soil reaction (pH). Geoderma 232-234, 141–
602 147.
603 Sharma, A., Weindorf, D.C., Wang, D., Chakraborty, S., 2015. Characterizing soils via Portable X-
604 ray Fluorescence spectrometer: 4. Cation exchange capacity (CEC). Geoderma 239-240, 130–
605 134.
606 Six, J., Bossuyt, H., Degryze, S., Denef, K., 2004. A history of research on the link between
607 (micro)aggregates, soil biota, and soil organic matter dynamics. Soil Tillage Res. 79, 7–31.
608 Soil Survey Staff, 2006. Land resource regions and major land resource areas of the United States,
609 the Caribbean, and the Pacific Basin. USDA Handbook 296. US Gov. Print. Off. Washington,
610 DC. 682.
611 Soil Survey Staff, 2014. Keys to soil taxonomy.12th ed. USDA-Natural Resour. Conserv. Serv.
612 Washington, DC.
613 Song, L., Langfelder, P., Horvath, S., 2012. Comparison of co-expression measures: mutual
614 information, correlation, and model based indices. BMC Bioinformatics 13(Mi), 328.
615 Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., Macdonald, L.M., McLaughlin, M.J., 2014.
616 The performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for prediction of
617 soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49, 139–186.
618 Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and Near Infrared
619 Spectroscopy in soil science. Chapter Five – Adv. Agron., 163–215.
620 Stevenson, F.J., 1994. Humus chemistry: genesis, composition, reactions. Wiley. New York, NY,
621 USA.
622 Swanhart, S., Weindorf, D.C., Chakraborty, S., Bark, N., Zhu, Y., Nelson, C., Shook, K., Acree, A.,

623 2014. Soil salinity measurement via Portable X-ray Fluorescence Spectrometry. Soil Sci. 179,
624 417–423.
625 Viscarra Rossel, R.A., McGlynn, R.N., McBratney, A.B., 2006a. Determining the composition of
626 mineral-organic mixes using UV–vis–NIR diffuse reflectance spectroscopy. Geoderma 137, 70–
627 82.
628 Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006b.
629 Visible, Near Infrared, Mid Infrared or combined diffuse reflectance spectroscopy for
630 simultaneous assessment of various soil properties. Geoderma 131, 59–75.
631 Wang, D., Chakraborty, S., Weindorf, D.C., Li, B., Sharma, A., Paul, S., Ali, M.N., 2015.
632 Synthesized use of VisNIR DRS and PXRF for soil characterization: total carbon and total
633 nitrogen. Geoderma 243-244, 157–167.
634 Wardle, D. a, Bardgett, R.D., Klironomos, J.N., Setälä, H., van der Putten, W.H., Wall, D.H., 2004.
635 Ecological linkages between aboveground and belowground biota. Science 304, 1629–1633.
636 Weindorf, D.C., Sarkar, R., Dia, M., Wang, H., Chang, Q., Haggard, B., McWhirt, A., Wooten, A.,
637 2008. Correlation of X-ray Fluorescence Spectrometry and Inductively Coupled Plasma Atomic
638 Emission Spectroscopy for elemental determination in composted products. Compost Sci. Util.
639 16, 79–82.
640 Weindorf, D.C., Bark, N., Zhu, Y., 2014. Advances in Portable X-ray Fluorescence (PXRF) for
641 environmental, pedological, and agronomic applications. Adv. Agron. 128, 1–45.
642 Zhu, Y., Weindorf, D.C., Zhang, W., 2011. Characterizing soils using a Portable X-ray
643 Fluorescence Spectrometer: 1. Soil texture. Geoderma 167-168, 167–177.
644 Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc.
645 Ser. B Statistical Methodol. 67, 301–320.
646

FIGURE CAPTIONS

Fig. 1. Schematic diagram of fused PSR + RF, PLS+RF and ENET + RF prediction models used in the study.

Fig. 2. Plots showing a) "Screeplot" of the first 10 principal components (PCs), b) pairwise PC plots of the first two components using VisNIR DRS first derivative spectra indicating four different geographical areas and c) pairwise PC plots of the first two components using VisNIR DRS first derivative spectra indicating three different organic horizons. Lines represent convex hulls and colored dots represent centroids of datasets from four different geographic areas. ITA, LBB, HOU and NMX represent samples from Italy, Lubbock, Houston and New Mexico regions, respectively.

Fig. 3. Parsimony features of ENET for a) TC and b) TN. CV error, MSE, DF and λ represent cross-validation error, test error, the number of selected wavelengths and the parameter to tune the parsimony of the model, respectively.

Fig. 4. Plots of observed vs. model predicted (VisNIR DRS only and fused) TC values using the validation set (with dotted 1:1 line).

Fig. 5. Plots of observed vs. model predicted (VisNIR DRS only and fused) TN values using the validation set (with dotted 1:1 line).

Fig.6. Diagnostic scatter plot matrix showing density plots for four competitors: observed TC values (OBS), RF in (PLS+RF) fused model (RF1) and RF in GRA (RF2). Black circles represent validation samples. The green and red solid lines are the fitted linear regression line and the loess smoother fit, respectively. The red dash lines represent one standard error above and below the estimated function. For diagonal plots, the vertical axis shows the density function for its corresponding values. For example, Fig. 6a is the density plot of TC while the tickmarks at the bottom axis show the observed values in the data. In the off-diagonal plots, their axes are all in the % unit. For example, Fig. 6e shows the scatter plot of observed TC % (on the vertical axis) and predicted TC % values from PLS model (on the horizontal axis) using spectral data.

NON-SATURATED SOIL ORGANIC HORIZON CHARACTERIZATION VIA ADVANCED PROXIMAL SENSORS

Valeria Cardelli[a*], David C. Weindorf[b*], Somsubhra Chakraborty[c], Bin Li[d], Mauro De Feudis[e], Stefania Cocco[a], Alberto Agnelli[c], Ashok Choudhury[c], Deb Prasad Ray[f], Giuseppe Corti[a]

[a]Department of Agricultural, Food and Environmental Sciences, Università Politecnica delle Marche, Ancona, AN, Italy

[b]Department of Plant and Soil Sciences, Texas Tech University, Lubbock, TX, USA

[c]Uttar Banga Krishi Viswavidyalaya, Cooch Behar, India

[d]Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA, USA

[e]Department of Agricultural, Food and Environmental Sciences, Università degli Studi di Perugia, Perugia, PG, Italy

[f]National Institute of Research on Jute and Allied Fibre Technology, Kolkata, India

*Corresponding author:
David C. Weindorf, Ph.D.
Department of Plant and Soil Science
Box 42122
Lubbock, TX 79409, USA
E-mail: david.weindorf@ttu.edu

Abstract

The organic fraction of soils is critically important to soil health and optimal ecosystem functioning. Traditional analysis of soil organic horizons (O horizons) has been dependent upon laboratory-based instrumentation. Simultaneously, the use of proximal sensors such as portable X-ray fluorescence (PXRF) spectrometry along with visible near infrared diffuse reflectance spectroscopy (VisNIR DRS) have gained popularity for providing rapidly acquired spectral and elemental data useful for soil physicochemical property quantification. However, PXRF and VisNIR DRS have mostly been applied to the assessment of mineral soils. This preliminary study evaluated 136 organic laden soil samples (most aptly described as upland, non-saturated O horizons) using both laboratory based instrumentation (CN analyzer) and proximal sensors to evaluate total carbon (TC) and total nitrogen (TN). Results revealed that combining model outcomes using model fusion improved TC and TN prediction accuracies relative to using an individual instrument (PXRF or VisNIR DRS) or model averaging with improvements in root mean square error (RMSE) on the order of 10-47% and 10-67% for TC and TN, respectively. Partial least squares + random forest (PLS+RF) approaches emerged as the best model for both predicting both TC and TN in organic laden soil samples. These results suggest that the strong predictive applications of proximal sensors extensively documented on mineral soils, may show similar promise for determination of a wide number of physicochemical properties on organic soil matrices, yet further exploration with a larger and more diverse dataset is recommended.

Key words: Spectroscopy; O horizons; proximal sensors

Abbreviations: SOM, soil organic matter; TC, total carbon; TN, total nitrogen; VisNIR-DRS, visible near infrared diffuse reflectance spectroscopy; PXRF, portable X-ray fluorescence

47

48 **1. Introduction**

49 Organic matter decomposition is a fundamental process for sustaining life on Earth (Gosz et al.,
50 1976). The term soil organic matter (SOM) refers to all organic material in soil, from freshly
51 deposited detritus or litter to highly decomposed, stable forms such as humic and fulvic acid
52 (Stevenson, 1994). Organic matter cycling helps to maintain ecosystem functionality as several
53 ecological functions are correlated to the decay processes of the organic layers of forest soils.
54 Indeed, decomposition and mineralization processes of organic residues affect nutrients cycling and
55 induce the release of elements that represent the principal resources for plants and microbes (Berger
56 et al., 2002; Berg and McClaugherty, 2008), such as macro- and micro-nutrients, and essential
57 molecules for energy metabolism, photosynthesis, and membrane transport (Huttl and Schaaf, 1997).
58 One of the main factors controlling the organic matter decomposition processes is the quality of the
59 litter produced by plants (Ge et al., 2013). The specific chemical proprieties of the plant litter and its
60 decay products, in turn, influence the underlying mineral soil (Wardle et al., 2004; Ball et al., 2014).
61 Six et al. (2004) noted that the decomposition of SOM has an impact on several important soil
62 properties as it improves soil aggregation (Bronick and Lal, 2005), enhances the activity of the soil
63 microbial community (Ball et al., 2014; Carrillo et al., 2012; García-Palacios et al., 2013), and
64 affects mineral weathering (Qafoku, 2015) and soil fertility (Kaiser et al., 2008). Thus, the
65 knowledge of the characteristics and composition of SOM, and in particular of the organic layers,
66 may help to ascertain certain soil ecosystem functions.

67 Current methods of SOM characterization are well established (Nelson and Sommers, 1996),
68 but are largely laboratory based. Recently, several studies have investigated rapid, inexpensive, and
69 non-destructive methods, such as visible near infrared diffuse reflectance spectroscopy (VisNIR-
70 DRS) and portable x-ray fluorescence spectrometry (PXRF) for soil analysis (Horta et al., 2015;
71 Weindorf et al., 2014). These proximal sensing methods have become increasingly accurate and

widely accepted offering data *in situ* in seconds given virtually no pre-processing requirements (Viscarra Rossel et al., 2006a; 2006b), with substantive advantages over traditional laboratory-based techniques. VisNIR-DRS is a spectrometric method which uses wavelengths across visible and near infrared regions (350-2500 nm) to explore the interaction between incident radiation and reflectance off of the soil surface; absorption is facilitated by C-H, N-H, or O-H bonds within the matrix (Chang et al., 2005). Due to this characteristic, it is highly applicable to C and N determination in soils. However, VisNIR-DRS spectra are generally weak, non-specific, and somewhat broad in their extent because of overlapping spectral signatures arising from variable soil components (Stenberg et al., 2010). As such, the instrument alone does not provide sufficient accuracy for complete soil characterization (Morgan et al., 2009). In fact, others have suggested the application of VisNIR-DRS in tandem with other sensing technologies (Brown et al., 2006; Fajardo et al., 2015). A complementary technique, PXRF, provides a multi-elemental analysis with a large range of quantification from low mg kg^{-1} to 100% for many elements (Hettipathirana, 2004). However, elements with stable electron configuration and low fluorescent energy (e.g., Na, N, H, Li, C) are not detectable (Wang et al., 2015). Nonetheless, several recent studies (e.g., Aldabaa et al., 2015; Chakraborty et al., 2015; Wang et al., 2015) have shown compelling predictive accuracy by combining the spectral signature of VisNIR-DRS with elemental data from PXRF, the latter used as auxiliary input data into the original advanced regression model. Individual or combined use of these two instruments allows for characterization of multiple soil parameters to include SOM (Stenberg et al., 2010), total carbon, total nitrogen (Wang et al., 2015), total phosphorus (Hu, 2013), cation exchange capacity (Sharma et al., 2015), pH (Sharma et al., 2014), salinity, (Swanhart et al., 2014), texture (Zhu et al., 2011), and contaminants (Chakraborty et al., 2015; Horta et al., 2015; Paulette et al., 2015). While the aforementioned studies offer wide-ranging application, most were conducted on mineral soils with limited organic content. Comparatively less information is available on the use of combined proximal sensors for soil organic layer (O horizon)

characterization. Wang et al. (2015) evaluated total carbon and nitrogen via combined PXRF and VisNIR-DRS approaches, but did so on mineral soils, lacking any analysis of true organic horizons. Similarly, Chang at al. (2002) have shown the efficacy of VisNIR-DRS to characterize soil carbon, but again, the soils evaluated were largely mineral soils. McWhirt et al. (2012) used a single sensor approach (VisNIR-DRS) to characterize the organic matter content of compost products. Whilst organic, composted products differ substantively in their physicochemical composition from that of organic soils. By contrast, the present study explicitly aims to evaluate the combined use of both proximal sensors (PXRF and VisNIR-DRS) in characterization of organic soil horizons in variable states of decay. As such, the objectives of this study were to: 1) quantify total carbon and nitrogen in natural organic soils by VisNIR-DRS and PXRF individualistically and, 2) explore if there is a benefit in predictive accuracy from concatenating VisNIR-DRS spectra and PXRF elements. We hypothesize that total carbon and nitrogen of largely organic horizons can be directly predicted from the reported PXRF elements and VisNIR-DRS spectra. We further hypothesize that either a fused model or model averaging approach will produce better predictability than either the VisNIR-DRS or PXRF approach independently.

2. Materials and Methods

2.1. General occurrences and features

In sum, 136 organic laden samples from non-saturated, uplands were collected in Italy and United States of America (Texas and New Mexico) during 2014 and 2015; a few mineral laden soils were also included as part of this dataset as a link to previously established work on mineral soils. The sites differed substantively in their geological composition, soil development, climate, and vegetation.

In Italy, a total of 39 organic horizons were collected from forest soils across three different sites on the Apennines chain (central Italy): Mount Acuto, Mount San Vicino, and Mount

Terminillo. The soils developed from limestone of different geological origin: Mount Acuto is characterized by limestone (Lower Cretaceous - Aptiano), Mount San Vicino is grey limestone with traces of flintstone and marl from the Jurassic (Lias) Pliensbachian Sinemurian and Mount Terminillo is grey limestone with trace amounts of flintstone (Jurassic Toarcian-Sinemurian) (ISPRA, 2015). The soils of these areas are classified as Mollisols or Inceptisols (Soil Survey Staff, 2014), characterized by a mesic soil temperature regime (10°C to 12°C) along with an udic soil moisture regime (from 825 mm to 1430 mm precipitation). In the three areas the cover vegetation was mainly composed of *Fagus sylvatica* from 80 to 99%, with *Carpinus betulus* at Mount Acuto, *Quercus cerris*, *Castanea sativa,* and *Sorbus aria* at Mount San Vicino, *Laburnum anagyroides* and *Acer* spp. at Mount Terminillo.

In Texas, 16 alluvial organic samples were collected in backwater areas along the North Fork of the Brazos River in Lubbock County in major land resource area (MLRA) 77C - Southern High Plains - Southern Part (Soil Survey Staff, 2006). Soils of this MLRA are generally developed by eolian deposits in the Blackwater Draw Formation of Pleistocene age, classified as Alfisols, Inceptisols, Mollisols, and Vertisols, and have a thermic soil temperature regime (13°C to 17°C) and an ustic soil moisture regime (from 405 to 560 mm precipitation). Mostly short and mid prairie grasses and scanty tree and shrubs (e.g., *Bouteloua gracilis*, *Bouteloua dactyloides*, *Bouteloua curtipendula)* are prevalent. Separately, 27 various organic horizons were sampled in forested areas of the George Bush Intercontinental Airport, WG Jones State Forest, San Jacinto River, and Sam Houston National Forest; all generally in the vicinity of Houston, Texas. The WG Jones State Forest, San Jacinto River, and Sam Houston National Forest occur in MLRA 133B - Western Coastal Plain (Soil Survey Staff, 2006), where soils developed from Tertiary and Cretaceous marine sediments consisting of inter bedded sandstone, siltstone, shale and loose primary particles. In particular, the Reklaw and Weches Formations in the Claiborne Group form the Redland area of East Texas. The main soil orders in this MLRA are Alfisols and Ultisols with a thermic soil

147 temperature regime (16°C to 20°C), an udic or aquic soil moisture regime (990 to 1,600 mm
148 precipitation). Vegetation of the area is typified by pine-hardwood species such as *Pinus taeda*,
149 *Pinus echinata*, *Liquidambar styraciflua*, *Quercus falcate,* and *Cornus florida*; *Callicapra*
150 *americana*, and *Smilax* spp. are common in the woody understory. *Schizachyrium scoparium* and
151 *Bothriochloa barbinodis* are the dominant herbaceous species. Of the 27 samples collected in this
152 area, four were collected in densely wooded pine forests adjacent to George Bush Intercontinental
153 Airport, which is a few kilometers beyond the aforementioned MLRA boundary, but quite similar in
154 the organic horizons sampled.

155 In New Mexico, 54 samples were collected near the periphery of the Lincoln National Forest of
156 Lincoln County; the horizons sampled were dominantly organic, but a few transitioned into organic
157 laden mineral soils. The area is in MLRA 39-Arizona and New Mexico Mountains (Soil Survey
158 Staff, 2006). The area is characterized by Cenozoic volcanic rock and various sedimentary sections
159 of the Colorado Plateau. The southern and eastern parts contain Permian and Cretaceous
160 sedimentary rock over a Precambrian granite core. Main soil orders of this MLRA are Inceptisols,
161 Mollisols, Alfisols, and Entisols, with mainly frigid or mesic soil temperature regimes (2°C to 13°C)
162 and an ustic/udic moisture regime (358 mm to 760 mm precipitation). *Pinus ponderosa* occurs is the
163 dominant vegetation in the low and intermediate heights, while at higher altitudes, *Picea abies* and
164 *Pseudotsuga menziesii* are commonplace.

165 *2.2. Field Sampling*

166 For maintaining an extensive variation in terms of physicochemical characteristics, composition
167 and origin, samples were randomly collected within morphologically established, organic laden
168 horizons. Almost all samples were classified as O horizons according to Référentiel pédologique
169 (2008). The different stages of degradation were classified as: i) OL, consisting of leaf debris with
170 little or no degradation (<10% of fine organic matter) whereby the botanic origin was easily
171 recognized; sub-horizons were recognized as OLn, fresh litter with minimal degradation, and OLv,

where the plant material were subject to initial degradation such as changes in color, volume, and inter-particle linkages; and ii) OH, with >70% of fine organic (humic) material following extensive decomposition (Table 1). The latter is homogenous, massive, and reddish-brown to black in color with an abundance of fine roots. These three states of decomposition are roughly akin to fibric, hemic, and sapric materials as defined by the Soil Survey Staff (2014), respectively.

2.3. Laboratory characterization and VisNIR scanning

Soil samples were air dried and ground to < 1-mm prior to chemical analyses and spectral scanning. Total C (TC) and total N (TN) were determined by a TruSpec® dry combustion analyzer (LECO Corp., MI, USA) according to Dumas method combustion (Nelson and Sommers, 1996).

Spectral reflectance values of air dried and ground soil samples were measured proximally over the VisNIR region (350-2500 nm) by a portable PSR-3500® VisNIR spectroradiometer (Spectral Evolution, USA). The reflectance data were resampled to 1 nm output values. Scanning was done with a contact probe containing a 5W halogen lamp, minimizing errors associated with stray light during measurements. Each sample was uniformly tiled in a glass petri plate and scanned four times, physically repositioning the probe prior to each scan. The mean of 10 internal scan over 1.5 seconds produced one individual scan. The spectroradiometer was standardized after every two samples using a NIST certified white reference. The average spectral curve was calculated and further used for spectral preprocessing and subsequent predictive modeling.

Necessary treatments of raw average reflectance spectra were done in R version 2.11.0 (R Development Core Team, 2014) following the spline fitting methods outlined in Wang et al. (2015). The spectral preprocessing involved converting reflectance to absorbance by log(1/R) which was executed in the Unscrambler®X 10.3 software (CAMO Software Inc., Woodbridge, NJ).

2.4. Scanning with PXRF

A DP-6000 Delta Premium portable X-ray fluorescence (PXRF) spectrometer (Olympus, Waltham, MA, USA) facilitated sample scanning. Configured with a Rh x-ray tube, the instrument was operated at 15–40 kV; an integrated ultra-high resolution (165 eV) silicon drift detector quantified each element detected. Before scanning, a 316 alloy clip was used to standardize the instrument. Primary analysis was conducted in Soil Mode (three beams of 30 s each); it can detect the following: Ag, As, Ba, Ca, Cd, Cl, Co, Cr, Cu, Fe, Hg, K, Mn, Mo, Ni, P, Pb, Rb, S, Sb, Se, Sn, Sr, Ti, V, Zn, and Zr. A second scanning was performed with Geochem Mode (two beams of 30 s each) in order to measure Mg, S, Al, and Si. Geochem and Soil Mode scans were done in duplicate, with the spectrometer physically repositioned between each scan. Sample homogeneity was ensured during the grinding step. Data were then averaged between scans to obtain a mean of elemental data for each sample. Data quality was ensured via the scanning of two NIST certified reference samples, with recovery percentage calculated as PXRF reported vs. NIST certified values. Results were as follows (PXRF reported/NIST certified [recovery]): Zn 4,263/4,180 mg kg^{-1} [1.02]; Cu 3,559/3,420 mg kg^{-1} [1.04]; K 27,081/21,700 [1.25]; Ca 9,119/9,640 mg kg^{-1} [0.95]; Ti 3,540/3,110 mg kg^{-1} [1.14]; Mn 2,237/2,140 [1.05]; Fe 50,574/43,200 mg kg^{-1} [1.17]; As 1,639/1,540 mg kg^{-1} [1.06]; Sr 269/255 [1.05]; Pb 5608/5520 [1.02].

2.5. Data mining

We performed all statistical modeling via R version 2.11.0 (R Development Core Team, 2014) software. We also checked the normality of residuals by the Shapiro-Wilk test at a 5% significance level. In the present study, both original TC and TN values were negatively skewed (Pearson skewness coefficient -1.78 and -0.37 for TC and TN, respectively), while Box–Cox conversion (Box and Cox, 1964) using $\lambda=0$ ($\log_{10}$ transformed) was unable to conform normal distribution. Principal component analysis (PCA) was executed via R version 2.11.0 using the 'prcomp' function to visualize spectral behavior of soil samples from different geographical areas and organic horizons. Optimal number of principal components (PC) was decided from a *screeplot*.

In this study, the whole dataset was split into 70% calibration (n=96) and 30% validation set (n=40) using the Kennard-Stone algorithm (Kennard and Stone, 1969), which is an adaptive procedure to select the most representative samples based on Euclidean distance. Calibration samples were used to establish a prediction model, whereas validation samples were used to assess the model's predictive ability (Chen et al., 2015). Initially, we targeted both TC and TN with VisNIR-DRS spectra only via partial least squares regression (PLS), elastic net regression (ENET), penalized spline regression (PSR), and random forest regression (RF) (Breiman, 2001; Guyon et al., 2002; Viscarra Rossel et al., 2006b). Additionally, both TC and TN were predicted with PXRF elemental data via PLS, ENET, and RF models.

Regularized methods, which play an important role in both statistical and data mining problems, can be described as Eq. 1:

$$\hat{\beta}(\lambda) = arg\min_{\beta} L(Y, X\beta) + \lambda J(\beta) \quad (1)$$

where $L(Y, X\beta)$ is a non-negative loss function, $J(\beta)$ is a non-negative penalty on the model complexity and λ is the non-negative tuning parameter. The key idea for regularized methods is to balance between the goodness-of-fit on the training data and the complexity of the model. Usually, a complicated model always shows a good fit to the calibration data. Nonetheless, this commonly leads to problems of overfitting, where the model is too adapted to the training data and often has a poor prediction performance on new samples.

Elastic net, devised by Zou and Hastie (2005) is a regularized regression method that linearly combines the ridge penalty and the least absolute shrinkage and selection operator (LASSO) penalty (Eq. 2)

$$\hat{\beta}(\lambda_1, \lambda_2) = arg\min_{\beta} SSE + \lambda_1 \sum|\beta_j| + \lambda_2 \sum \beta_j^2 \quad (2)$$

where SSE is the sum of squared errors, $\sum|\beta_j|$ is the LASSO penalty, and $\sum \beta_j^2$ is the ridge penalty. In high dimensional data analysis where many predictors are correlated to each other, the LASSO penalty tends to pick a few of the predictors and discard the others (sparse model), while the ridge

245 penalty shrinks the coefficients of correlated predictors towards each other (dense model). The
246 ENET penalty combines these two penalties so that when the predictors are correlated in groups, it
247 will produce a sparse model with good prediction performance, while boosting a grouping effect.
248 Another advantage of ENET is that it can handle the high dimension and low sample size problem.
249 In this study, ENET was run using the 'glmnet' package in R, produced by Friedman et al. (2015).
250 Details of PSR and RF methodology are summarized in Wang et al. (2015).

251 *2.6. Model Fusion and Model Averaging*

252 Model fusion and model averaging (model ensemble) were tested to determine if combination
253 of the predictions for the VisNIR-DRS and PXRF models into a single composite score could
254 improve predictive accuracy of TC and TN. Initially, three fused modeling approaches (PSR+RF),
255 (PLS+RF) and (ENET+RF) were employed where PSR/PLS/ENET were used to fit the training set
256 (containing VisNIR-DRS spectra only). Next, we ran the RF using the PSR/PLS/ENET residual as
257 the response and PXRF elements as the predictors (Chakraborty et al., 2015). Succinctly, the
258 prediction value on the validation set using (PSR+RF), (PLS+RF) and (ENET+RF) contains two
259 additive parts: the prediction from PSR/PLS/ENET model using the VisNIR-DRS spectra plus the
260 prediction from RF model using the PXRF data. An outline of this fused procedure is shown in Fig.
261 1.

262 For model averaging we used Granger-Ramanathan averaging (GRA) (Granger and
263 Ramanathan, 1984) with some modifications. The GRA approach requires fitting a multivariate
264 linear regression model where lab measured soil property values are regressed against the
265 corresponding predictions derived from VisNIR-DRS and PXRF models. Initially, the
266 PSR/PLS/ENET model was fitted on the calibration set (n=96) using the VisNIR-DRS spectra.
267 Subsequently, the RF model was fitted using only PXRF elements as predictors and TC or TN as
268 the response. Note that this RF model was not the same as the RF model in fused approaches

269 (PSR+RF, PLS/RF, or ENET+RF) as the latter used residual as a response. Next, a bivariate linear
270 regression was fitted with Eq. 3:

271 $$Y = a + b*X + c*Z \quad (3)$$

272 where Y is either TC or TN on calibration set; X is the fitted value (on calibration set) from the
273 PSR/PLS/ENET model; Z is the fitted value (on the calibration set) from the RF model; *a* is the
274 intercept; *b* is the estimated model weight for the PSR model; and *c* is the estimated weight for the
275 RF model. To predict the validation set (n=40), we first generated the prediction value, X*, from
276 the PSR/PLS/ENET model, and prediction value, Z*, from the RF model. The GRA prediction is
277 given as Eq. 4:

278 $$Y^* = a + bX^* + cZ^* \quad (4)$$

279 We followed the Kennard-Stone splitting scheme for both fused models and GRA.

280 *2.7. Whole geographical area and organic horizon holdout validation*

281 Further, whole geographical area and organic horizon holdout validations were executed for
282 both TC and TN to determine how area to area and horizon to horizon heterogeneity affected
283 prediction accuracies (Brown el al., 2005). Whole-geographical area holdouts were achieved by
284 calibrating a model using the best performing algorithm with three geographical areas and then
285 validated using the fourth area. Moreover, whole-organic horizon holdouts were achieved by
286 calibrating a model using the best performing algorithm with two organic horizons and then
287 validated using the third organic horizon.

288 In this study, the root mean square error (RMSE), regression coefficient (R^2), residual prediction
289 deviation (RPD) (Eq. 5), bias, and ratio of performance to inter-quartile distance (RPIQ) (Eq. 6)
290 were used to evaluate model performance (Gauch et al., 2003 Bellon-Maurel et al., 2010). RPD

based model accuracy classification scheme devised by Chang et al. (2001) was followed for evaluating model accuracy.

$$RPD = \left[\frac{1/(n-1)\sum_{i=1}^{n}(Yobs-Ymean)^2}{1/n\sum_{i=1}^{n}(Yobs-Ypred)^2}\right]^{0.5}_{Validation} \quad (5)$$

$$Bias = \sum_{i=1}^{n}(Ypred - Ymean)/n \quad (6)$$

where, *Yobs* and *Ypred* are the observed and predicted response variables, respectively, *Ymean* is the average of the *Yobs* values, and *n* denotes the sample number in the validation data set. RPIQ was defined as IQ/SEP, where SEP represents the standard error of prediction, and IQ denotes the interquartile distance of the validation set (IQ=Q3–Q1). In this study, optimal model performance featured low RMSE values along with high RPD, R^2, and RPIQ.

3. Results and Discussion

3.1. Soil properties and PCA

Statistical moments for TC and TN measured on the soil samples used in the regression models are listed in Table 2. Soil TC and TN showed negatively skewed distribution with mean concentrations of 39.05% and 1.10%, respectively for the whole dataset; 39.53% and 1.08%, respectively for the calibration set; and 37.90% and 1.14%, respectively for the validation set. Of the 13 elements obtained by the PXRF soil mode, only nine [Zn, S, K, Ca, Ti, Rb, Mn, Fe, Sr] were suitable for use in the multivariate predictive models with continuous data across all samples scanned. Within-region variability of TC and TN exhibited substantial variation. For instance, in Italy sample TC ranged from 28.06% to 45.94%. Showing similar variation, Lubbock samples showed TC ranging from 25.93% to 41.20%. Comparatively higher variability was observed for Houston samples (11.42-48.54%), while maximum TC variability was observed in samples from New Mexico region (0.06-51.30%). Lubbock samples showed TN ranging from 1.13-1.60%. Comparatively higher variability was observed for Italy (0.80-1.87%) and Houston samples (0.30-1.47%). Following a similar trend of TC, maximum variability in TN was observed from New

315 Mexico soils (0.04-1.87%). Overall, variability in estimated TC and TN was due largely to the
316 variability in the parent material, climate, C source, land use, and vegetation cover. However, the
317 extent of variability appeared related with the range of the mean annual soil temperature of the
318 considered sampling area, the highest in the New Mexico soils (2-13°C), the lowest in the Italy soils
319 (10-12°C), intermediate in the other two subsets (16-20°C in the Houston soils, and 13-17°C in the
320 Lubbock soils).

321 Levene's test yielded equality of variance of TC and TN values among the training and test sets
322 ($p=0.65$ and 0.55 for TC and TN, respectively). Moreover, t-test could not establish a significant
323 difference between mean TC ($p=0.53$) and TN ($p=0.51$) for these two data sets. These results
324 indicated that the validation samples based on the Kennard-Stone method can properly represent the
325 studied population. The first two leading PCs constituted over 95% of the spectral variation (Fig.
326 2a). Although some overlapping among samples from four regions were discernible from the PC1
327 vs. PC2 plot, these regions had different range, shape, and distributions in the spectral space,
328 indicating variable SOM composition from different organic input sources (Fig. 2b). Moreover, the
329 PCA plot indicating samples from three different organic horizons (OH, OLn and OLv) (Fig. 2c)
330 indicated that OH horizon samples were quite different from the samples from OLn and OLv
331 horizons. The OH horizon samples tended to have larger values on the PC1 scores. Contrarywise,
332 the samples from the OLn and OLv horizons were relatively close. Notably, the samples from the
333 OLv horizon tended to have relatively larger PC2 values while there was an obvious outlying
334 sample with the largest PC1 value and smallest PC2 value.

335 *3.2. Validation results and model parsimony*

336 In fused models, both TC and TN were estimated with good RPIQ values, better than using an
337 individual instrument or GRA (Table 3), which has been noted before elsewhere (Wang et al., 2015).
338 Fused (PLS+RF) and (PSR+RF) models yielded the highest RPDs for TC (2.84) and TN (2.62),
339 respectively. While RPD values for both tested soil properties dropped in (ENET+RF) fused models,

they still explained nearly 80% of the TC and TN variability. Overall, all three fused models of TC showed similar accuracies (Table 3), as reported for soil organic C in the review of Soriano-Disla et al. (2014) where the median R^2 of validations was calculated as 0.83 (33 studies).

Wavelength selection not only enhances the stability of the model but also makes the model more parsimonious. Although (PSR+RF) yielded the lowest RMSE for TN (0.160%), both PSR and RF are not parsimonious models since they do not select the subset of important variables. Although for RF it is possible to select different *mtry* values, which is the size of the candidate subset for each splitting, models with different *mtry* values still select all the variables. Despite controlling the smoothness of the neighboring coefficients through λ, PSR is also unable to select important wavelengths. Figure 3 illustrates parsimony features of ENET for TC (Fig. 3a) and TN (Fig. 3b) where cross-validation error, test error (mean squared error or MSE), and the degrees of freedom (DF) (the number of selected wavelengths in the ENET model) were plotted against log of λ which is the parameter to tune the ENET parsimony. Note that the larger the λ, the more parsimonious the model. The vertical dash line shows the optimal λ that minimizes the cross-validation error. The bottom plot shows the ENET coefficient plot against the VisNIR wavelengths. For both soil properties, most of the coefficients were zero, indicating that they were not selected into the model. Additionally, based on the DF plots, it was evident that the optimal TC and TN models selected only ~300-350 out of 2151 wavelengths. Conversely, using a fused (PLS+RF) model on the reduced feature sets (14 and 13 latent factors for TC and TN, respectively) returned the best result for TC and slightly higher RMSE for TN (0.191%) than (PSR+RF) (0.160%) (Table 3). Based on these observations, we conclude that while considering both model precision and parsimony, (PLS+RF) emerged as the best model for both TC and TN.

For both TC and TN, while utilizing VisNIR-DRS or PXRF alone, RF and PLS models yielded the largest validation RMSE values, respectively, and were therefore the least precise (Table 3). In general, PXRF measurements were poorly correlated to both TC and TN, suggesting its limitation

365 to make full geochemical assessment alone due to its inability to handle low concentrations and
366 light elements ($Z \leq 11$) (Weindorf et al., 2008). Among nine elements used in predictive models,
367 only Al (1.48%), Ca (2.53%) and Si (4.83%) were present in high concentrations, while the other
368 six elements were present at low levels (<1%). Thus, it seemed prudent to calibrate using PXRF
369 spectra instead of elements if high accuracy is desired. While GRA substantially worsen the
370 validation results for TC, (PSR+RF) model averaging for TN yielded close RMSE (0.197%) to
371 those produced by fused models. The reasonably good performance of PSR in all three approaches
372 (VisNIR only, fused, and GRA) can be attributed to the fact that in case of 'fat' data (large number
373 of dimensions and small sample size) PSR uses all samples and smooths constraints on the
374 coefficient. Therefore, PSR often works well on signal regression problems which are usually
375 strongly dimensional and have a relatively small sample size.

376 As coefficients of determinations of TN were rather similar to those of TC using VisNIR-DRS
377 in isolation, we conclude that VisNIR-DRS sensed a grouping of soil components comprising
378 organic functional groups that contain organic N fractions (Wang et al., 2015) (Table 3). Moreover,
379 both TC and TN are spectrally active components or chromophores, which absorb incident energy
380 at discrete energy levels and show broad and weak absorption features in the VisNIR region (Ben-
381 Dor, 2011) Although intense VisNIR-DRS spectral bands cannot be directly ascribed to metals or
382 other components, Song et al. (2012) revealed that metals can interact with chromophores such as
383 soil C. However, S was negatively correlated with TC (ρ=-0.48); likely since S is very light with an
384 atomic mass of 16 and was present in low concentrations (max=0.43%).

385 Plots of observed *vs*. model predicted (VisNIR-DRS only and fused) TC and TN values are
386 presented in Figs. 4 and 5, respectively. In general, all models showed overestimation at lower TC
387 or TN values and underestimation at higher values. Several of these overestimations occurred
388 because of the relative scarcity of observations with low values, which was expected since the
389 samples came dominantly from organic horizons. Another possible explanation for TC and TN

prediction errors is undecomposed organic matter and variable C sources, as noted before elsewhere (Henderson et al., 1992). Indeed, the VisNIR-DRS spectra for SOM have not been fully defined yet, because of the complexity or unclear definitions of these materials (Brown et al., 2006). Probably the TC and TN predictability could have been enhanced by using soil management or vegetation specific models (Morgan et al., 2009); nonetheless, exploring this notion was beyond the scope of this study.

3.3. Model fusion vs. model averaging

It was apparent that GRA model averaging by combining the VisNIR-DRS and PXRF predictions for consolidated use could not produce better accuracy than model fusion. In practice, "good" model averaging should contain the models that complement each other. In fused models, RF was fitted based on the PLS/PSR/ENET residuals. The sequential fitting in PLS/PSR/ENET + RF allowed the RF model to complement the former. Conversely, combining PLS/PSR/ENET and RF in parallel fashion through GRA (namely, separately on VisNIR-DRS and PXRF) produced incompatibility. Succinctly, they perhaps made the same mistake on the same sample. To justify this postulation and clearly visualize the prediction improvement by using the fused models, Fig. 6 represents a scatterplot matrix produced in R using the *spm* function in the car library, taking the PLS+RF model as an example. The diagonal elements are the density plots for the four competitors [observed TC values (OBS), RF in (PLS+RF) fused model (RF1), and RF in GRA (RF2)]. For example, the upper left one is the density plot of TC while the tickmarks at the bottom axis show the observed values in the data. The off-diagonal elements are the pairwise scatter plots of four competitors, together with the best linear and nonlinear smoothers. For example, Fig. 6e shows the scatter plot of observed TC (on the vertical axis) and predicted TC values from PLS model (on the horizontal axis) using spectral data. The green and red solid lines are the fitted linear regression line and the loess smoother (a popular nonlinear smoother using local linear regression) fit, respectively. The red dash lines represent one standard error above and below the estimated function. It can be

observed that PLS tended to underestimate the samples with TC values between 30 to 45% and overestimate the samples with TC values < 20% through the bending shape of the red loess smoother (Fig. 6e). Interestingly, while RF1 "corrected" the PLS model by lifting the underestimated area (30-45% TC) and lowering the overestimated region (<20% TC) (Fig. 6i), this was not totally unexpected owing to the non-linear and contingent relationships among VisNIR-DRS reflectance and soil constituents (Brown et al., 2006). However, RF2 made the same "mistake" as PLS by underestimating and overestimating in the same regions (Fig. 6d). These trends clearly explained why sequentially fitting (PLS+RF) in fused model improves prediction accuracy relative to parallel fitting of PLS and RF through GRA. Note that, the same argument is applicable to (PSR+RF) or (ENET+RF) in predicting TC and TN. Model averaging even could not outperform VisNIR-DRS only model performance (Table 3). Thus, the postulations of Abbott (2014), who reported that model averaging nearly always improves model predictive accuracy and rarely predicts worse than single models, may not be generalized.

3.4. Relative improvement of soil property predictions after model fusion

Table 4 demonstrates that fused models had relative improvements on the prediction accuracy of TC and TN from GRA or VisNIR-DRS and PXRF used in isolation. Notably, RMSE improvements of 10-47% for TC and 10-67% for TN were achieved by the PLS+RF and PSR+RF model, respectively, demonstrating the potential of the synergistic use of VisNIR-DRS and PXRF for the estimation of soil attributes. On average, the PLS+RF model for TC and the PSR+RF model for TN produced almost a one fourth reduction in RMSE from using VisNIR-DRS in isolation. While the PLS+RF fused model increased RPIQ values ranging between 11 to 97% for TC, the PSR+RF model for TN yielded three times higher RPIQ than using PXRF in isolation. Also, the PSR+RF approach for TC produced substantively better results (R^2 0.87; RMSE 3.01%; RPD 2.75) than those obtained by McWhirt et al. (2012) (R^2 0.82; RMSE 10.1%; RPD 1.72) who used VisNIR-DRS in isolation when evaluating the organic fraction of composted products.

440 *3.5 Whole geographical area and organic horizon holdout validations*

441 Holding out a whole geographical area or organic horizon reduced prediction accuracies in all
442 cases where the soils of the geographical area or horizon held out were not represented by another
443 area or horizon, creating extrapolation (Tables 5 and 6). Validation with an individual geographical
444 area or horizon exhibited increases in RMSEs and a reduction in RPDs relative to the best
445 performing models (PLS+RF and PSR+RF for TC and TN, respectively) using all 136 samples. As
446 expected, when principal component analysis was implemented on the VisNIR spectrum only, it
447 clearly demonstrated different spectral behaviors of the four geographical areas (Fig. 2b). Further,
448 while executing Principal Component Regression (PCR, data not shown) to calculate the actual
449 number of PCs required to represent VisNIR spectral variability, results indicated the need for at
450 least 12 PCs in PCR based on full cross validation. In other words, although some overlap among
451 the four areas was apparent in the PC1 vs. PC2 plot, that only reflected the first two PCs.
452 Summarily, it can be concluded that since the four geographical areas had different VisNIR spectra,
453 using all samples in a single prediction model is preferred over the area-holdout scheme since the
454 latter involves extrapolation in predicting samples from different areas.

455 Moreover, based on the whole organic horizon-holdout results, all models predicted poorly on
456 holding out OH and OLn samples. For OH, the main reason of poor model performance was that
457 OH samples were quite different from OLn and OLv samples because of variable rates of
458 decomposition. Hence the fitted models using OLn and OLv samples could not predict OH samples
459 well. Conversely, since the OLn horizon represented most of the samples (i.e 64 out of 136
460 samples), the sample size for fitting the model using the OLv and OH was the smallest. In addition,
461 the OH samples in the training data made the fitted model poor in predicting OLn samples in the
462 validation set.

463 *3.6 Some practical concerns*

464 The fused PSR+RF model had some limitations too. Initially, PSR was applied on VisNIR-DRS
465 spectra with the assumption that the underlying PSR coefficient curve is smooth. Secondly, RF was
466 used to fit the residuals of the PSR model on the PXRF elemental data assuming that both VisNIR-
467 DRS and PXRF data contain some information on the response variable. In particular, PXRF
468 contains some extra information about TC and TN that were not explained (or covered) by the
469 VisNIR-DRS spectra. We have examined the feasibility of the approach in this research and have
470 provided a preliminary contribution to the issue of validation, however, further intensive research is
471 recommended to confirm our results. Nonetheless, the combined use of PXRF and VisNIR data
472 show strong potential for the accurate assessment of soil organic horizons. The spectral and
473 elemental data produced by proximal sensor analysis has previously been shown to be useful in the
474 prediction of a wide number of physicochemical properties in mineral soils. With regard to TC and
475 TN content of organic soil horizons, similar impressive predictability from proximal sensor data
476 was observed.
477
478 **4. Conclusions**
479 Summarily, this preliminary study of 136 organic laden soil samples from non-saturated uplands
480 of Italy and the United States were evaluated by standard laboratory analysis as well as proximal
481 sensor (PXRF + VisNIR DRS) scans. Our results show that:
482 (i) Combining model outcomes using model fusion improved TC and TN prediction accuracies
483 relative to using an individual instrument (PXRF or VisNIR DRS) or model averaging. Overall, the
484 relative improvement in % RMSE ranged from 10-47% and 10-67% for TC and TN, respectively.
485 (ii) Considering both model precision and parsimony, PLS+RF emerged as the best model for
486 predicting both TC and TN in organic laden soil samples. Conversely, while using all spectral
487 variables, PSR+RF yielded the best model results (lowest RMSE and highest RPIQ) for TN.

(iii) Specifically, GRA model averaging by combining the VisNIR-DRS and PXRF predictions could not produce better accuracy than model fusion or individual instrument use, possibly due to model incompatibility while fitting them in parallel fashion.

(iv) The poor predictability with unreliable RPIQs while using PXRF in isolation underlines the need for using PXRF spectra instead of elemental data.

Acknowledgments

The authors are grateful for support from the BL Allen Endowment in Pedology at Texas Tech University in completing this research.

References

Abbott, D. 2014. Applied Predictive Analytics. Principles and Techniques for the professional data analyst. John Wiley & Sons, Inc., Indianapolis, IN. p. 307-326.

Aldabaa, A.A.A., Weindorf D.C., Chakraborty S., Sharma A., Li, B., 2015. Combination of proximal and remote sensing methods for rapid soil salinity quantification. Geoderma 239-240, 34-46.

Association française pour l'Etude du sol. 2008. Référentiel pédologique. Ed. Quae RD 10, 78026 Versailles Cedex, France.

Ball, B.A., Carrillo, Y., Molina, M., 2014. The influence of litter composition across the litter-soil interface on mass loss, nitrogen dynamics and the decomposer community. Soil Biol. Biochem. 69, 71–82.

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. Trends Analyt. Chem. 29, 1073–1081.

Ben-Dor, E., 2011. Characterization of soil properties using reflectance spectroscopy, in Thenkabail, P.S., Lyon, J.G., Huete, A. (Eds.), Hyperspectral Remote Sens. Veg. CRC Press. 513– 558.

Berg, B., McClaugherty, C., 2008. Plant Litter - Decomposition, Humus Formation, Carbon Sequestration. Springer Berlin Heidelberg, Berlin, Heidelberg.

Berger, T.W., Neubauer, C., Glatzel, G., 2002. Factors controlling soil carbon and nitrogen stores in pure stands of Norway spruce (Picea abies) and mixed species stands in Austria. For. Ecol. Manage. 159, 3–14.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc. Series B Stat. Methodol. 26, 211–252.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Bronick, C.J., Lal, R., 2005. Soil structure and management: a review. Geoderma 124, 3–22.

523 Brown, D.J., Bricklemyer, R.S., Miller, P.R., 2005. Validation requirements for diffuse reflectance
524 soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma
525 129, 251-267.
526 Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil
527 characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132, 273–290.
528 Carrillo, Y., Ball, B.A., Strickland, M.S., Bradford, M.A., 2012. Legacies of plant litter on carbon
529 and nitrogen dynamics and the role of the soil community. Pedobiologia (Jena) 55, 185–192.
530 Chakraborty, S., Weindorf, D.C., Li, B., Aldabaa, A.A.A., Ghosh, R.K., Paul, S., Nasim Ali, M.,
531 2015. Development of a hybrid proximal sensing method for rapid identification of petroleum
532 contaminated soils. Sci. Total Environ. 514, 399–408.
533 Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance
534 Spectroscopy: Principal Components Regression Analyses of soil properties. Soil Sci. Soc. Am.
535 J. 65, 480.
536 Chang, C.-W., Laird, D.A., 2002. Near-Infrared Reflectance spectroscopic analysis of soil C and N.
537 Soil Sci. 167, 110–116.
538 Chang, C.-W., You, C.-F., Huang, C.-Y., Lee, T.-Q., 2005. Rapid determination of chemical and
539 physical properties in marine sediments using a Near-Infrared Reflectance spectroscopic
540 technique. Appl. Geochemistry 20, 1637–1647.
541 Chen, T., Chang, Q., Clevers, J.G.P.W., Kooistra, L., 2015. Rapid identification of soil cadmium
542 pollution risk at regional scale based on Visible and Near-Infrared spectroscopy. Environ. Pollut.
543 206, 217–26.
544 Fajardo, M., McBratney, A., Whelan, B., 2015. Fuzzy clustering of Vis–NIR spectra for the
545 objective recognition of soil morphological horizons in soil profiles. Geoderma 263, 244-253.
546 Friedman, A.J., Hastie, T., Simon, N., Tibshirani, R., Hastie, M.T., 2015. Lasso and Elastic-Net
547 Regularized Generalized Linear Models. Available online at: https//cran.r-

548 project.org/web/packages/glmnet/glmnet.pdf. (Verified on July 29, 2015).
549 García-Palacios, P., Maestre, F.T., Kattge, J., Wall, D.H., 2013. Climate and litter quality
550 differently modulate the effects of soil fauna on litter decomposition across biomes. Ecol. Lett.
551 16, 1045–53.
552 Gauch, H.G., Hwang, J.T.G., Fick, G.W., 2003. Model evaluation by comparison of model-based
553 predictions and measured values. Agron. J. 95, 1442-1446.
554 Ge, X., Zeng, L., Xiao, W., Huang, Z., Geng, X., Tan, B., 2013. Effect of litter substrate quality and
555 soil nutrients on forest litter decomposition: A review. Acta Ecol. Sin. 33, 102–108.
556 Gosz, J.R., Likens, G.E., Bormann, F.H., 1976. Organic matter and nutrient dynamics of the forest
557 and forest floor in the Hubbard Brook forest. Oecologia 22, 305–320.
558 Granger, C.W.J., Ramanathan, R., 1984. Improved methods of combining forecasts. J. Forecast. 3,
559 197–204.
560 Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using
561 support vector machines. Mach. Learn. 46, 389–422.
562 Henderson, T.L., Baumgardner, M.F., Franzmeier, D.P., Stott, D.E., Coster, D.C., 1992. High
563 dimensional reflectance analysis of soil organic matter. Soil Sci. Soc. Am. J. 56, 865-872.
564 Hettipathirana, T.D., 2004. Simultaneous determination of parts-per-million level Cr, As, Cd and Pb,
565 and major elements in low level contaminated soils using borate fusion and energy dispersive X-
566 ray fluorescence spectrometry with polarized excitation. Spectrochim. Acta Part B At. Spectrosc.
567 59, 223–229.
568 Horta, A., Malone, B., Stockmann, U., Minasny, B., Bishop, T.F.A., McBratney, A.B., Pallasser, R.,
569 Pozza, L., 2015. Potential of integrated field spectroscopy and spatial analysis for enhanced
570 assessment of soil contamination: A prospective review. Geoderma 241-242, 180–209.
571 Hu, X.-Y., 2013. Application of Visible/Near-Infrared Spectra in modeling of soil total phosphorus.
572 Pedosphere 23, 417–421.

573 Huttl, R.F., Schaaf, W., 1997. Magnesium deficiency in forest ecosystems, first ed. Springer,
574 Netherland.
575 ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale), 2015. Carta geologica d'Italia.
576 URL http://www.isprambiente.gov.it/Media/carg/ (Verified on October 28, 2015).
577 Kaiser, M., Ellerbrock, R.H., Gerke, H.H., 2008. Cation exchange capacity and composition of
578 soluble soil organic matter fractions. Soil Sci. Soc. Am. J. 72, 1278.
579 Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137-
580 148.
581 McWhirt, A.L., Weindorf, D.C., Chakraborty, S., Li, B., 2012. Visible near infrared diffuse
582 reflectance spectoscopy (VisNIR DRS) for rapid measurement of organic matter in compost.
583 Waste Management & Research 30, 1049-1058.
584 Morgan, C.L.S., Waiser, T.H., Brown, D.J., Hallmark, C.T., 2009. Simulated in situ
585 characterization of soil organic and inorganic carbon with Visible Near-Infrared Diffuse
586 Reflectance spectroscopy. Geoderma 151, 249–256.
587 Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter, in Sparks,
588 D.L. (Ed.). Methods of soil analysis - Part 3. Chemical Methods. Soil Sci. Soc. Am., Madison,
589 WI.
590 Paulette, L., Man, T., Weindorf, D.C., Person, T., 2015. Rapid assessment of soil and contaminant
591 variability via Portable X-ray Fluorescence spectroscopy: Copşa Mică, Romania. Geoderma 243-
592 244, 130–140.
593 Qafoku, N.P., 2015. Climate-Change effects on soils: accelerated weathering, soil carbon, and
594 elemental cycling. Chapter Two. Adv. Agron. 131, 111–172.
595 R Development Core Team, 2008. R: a language and environment for statistical computing . R
596 Found. Stat. Comput. Vienna, Austria (Available online with Updat at http//www.cran.r-
597 project.org, Verified on January 6, 2016).

598 Rossel, R.A.V., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance
599 spectra. Geoderma 158, 46–54.
600 Sharma, A., Weindorf, D.C., Man, T., Aldabaa, A.A.A., Chakraborty, S., 2014. Characterizing soils
601 via Portable X-ray Fluorescence spectrometer: 3. Soil reaction (pH). Geoderma 232-234, 141–
602 147.
603 Sharma, A., Weindorf, D.C., Wang, D., Chakraborty, S., 2015. Characterizing soils via Portable X-
604 ray Fluorescence spectrometer: 4. Cation exchange capacity (CEC). Geoderma 239-240, 130–
605 134.
606 Six, J., Bossuyt, H., Degryze, S., Denef, K., 2004. A history of research on the link between
607 (micro)aggregates, soil biota, and soil organic matter dynamics. Soil Tillage Res. 79, 7–31.
608 Soil Survey Staff, 2006. Land resource regions and major land resource areas of the United States,
609 the Caribbean, and the Pacific Basin. USDA Handbook 296. US Gov. Print. Off. Washington,
610 DC. 682.
611 Soil Survey Staff, 2014. Keys to soil taxonomy.12th ed. USDA-Natural Resour. Conserv. Serv.
612 Washington, DC.
613 Song, L., Langfelder, P., Horvath, S., 2012. Comparison of co-expression measures: mutual
614 information, correlation, and model based indices. BMC Bioinformatics 13(Mi), 328.
615 Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., Macdonald, L.M., McLaughlin, M.J., 2014.
616 The performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for prediction of
617 soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49, 139–186.
618 Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and Near Infrared
619 Spectroscopy in soil science. Chapter Five – Adv. Agron., 163–215.
620 Stevenson, F.J., 1994. Humus chemistry: genesis, composition, reactions. Wiley. New York, NY,
621 USA.
622 Swanhart, S., Weindorf, D.C., Chakraborty, S., Bark, N., Zhu, Y., Nelson, C., Shook, K., Acree, A.,

623 2014. Soil salinity measurement via Portable X-ray Fluorescence Spectrometry. Soil Sci. 179,
624 417–423.
625 Viscarra Rossel, R.A., McGlynn, R.N., McBratney, A.B., 2006a. Determining the composition of
626 mineral-organic mixes using UV–vis–NIR diffuse reflectance spectroscopy. Geoderma 137, 70–
627 82.
628 Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006b.
629 Visible, Near Infrared, Mid Infrared or combined diffuse reflectance spectroscopy for
630 simultaneous assessment of various soil properties. Geoderma 131, 59–75.
631 Wang, D., Chakraborty, S., Weindorf, D.C., Li, B., Sharma, A., Paul, S., Ali, M.N., 2015.
632 Synthesized use of VisNIR DRS and PXRF for soil characterization: total carbon and total
633 nitrogen. Geoderma 243-244, 157–167.
634 Wardle, D. a, Bardgett, R.D., Klironomos, J.N., Setälä, H., van der Putten, W.H., Wall, D.H., 2004.
635 Ecological linkages between aboveground and belowground biota. Science 304, 1629–1633.
636 Weindorf, D.C., Sarkar, R., Dia, M., Wang, H., Chang, Q., Haggard, B., McWhirt, A., Wooten, A.,
637 2008. Correlation of X-ray Fluorescence Spectrometry and Inductively Coupled Plasma Atomic
638 Emission Spectroscopy for elemental determination in composted products. Compost Sci. Util.
639 16, 79–82.
640 Weindorf, D.C., Bark, N., Zhu, Y., 2014. Advances in Portable X-ray Fluorescence (PXRF) for
641 environmental, pedological, and agronomic applications. Adv. Agron. 128, 1–45.
642 Zhu, Y., Weindorf, D.C., Zhang, W., 2011. Characterizing soils using a Portable X-ray
643 Fluorescence Spectrometer: 1. Soil texture. Geoderma 167-168, 167–177.
644 Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc.
645 Ser. B Statistical Methodol. 67, 301–320.
646
647

FIGURE CAPTIONS

Fig. 1. Schematic diagram of fused PSR + RF, PLS+RF and ENET + RF prediction models used in the study.

Fig. 2. Plots showing a) "Screeplot" of the first 10 principal components (PCs), b) pairwise PC plots of the first two components using VisNIR DRS first derivative spectra indicating four different geographical areas and c) pairwise PC plots of the first two components using VisNIR DRS first derivative spectra indicating three different organic horizons. Lines represent convex hulls and colored dots represent centroids of datasets from four different geographic areas. ITA, LBB, HOU and NMX represent samples from Italy, Lubbock, Houston and New Mexico regions, respectively.

Fig. 3. Parsimony features of ENET for a) TC and b) TN. CV error, MSE, DF and λ represent cross-validation error, test error, the number of selected wavelengths and the parameter to tune the parsimony of the model, respectively.

Fig. 4. Plots of observed vs. model predicted (VisNIR DRS only and fused) TC values using the validation set (with dotted 1:1 line).

Fig. 5. Plots of observed vs. model predicted (VisNIR DRS only and fused) TN values using the validation set (with dotted 1:1 line).

Fig.6. Diagnostic scatter plot matrix showing density plots for four competitors: observed TC values (OBS), RF in (PLS+RF) fused model (RF1) and RF in GRA (RF2). Black circles represent validation samples. The green and red solid lines are the fitted linear regression line and the loess smoother fit, respectively. The red dash lines represent one standard error above and below the estimated function. For diagonal plots, the vertical axis shows the density function for its corresponding values. For example, Fig. 6a is the density plot of TC while the tickmarks at the

672 bottom axis show the observed values in the data. In the off-diagonal plots, their axes are all in the
673 % unit. For example, Fig. 6e shows the scatter plot of observed TC % (on the vertical axis) and
674 predicted TC % values from PLS model (on the horizontal axis) using spectral data.
675
676

Table 1. General description of the different type of organic horizons collected divided by study sites. For symbols see legend.

ITALY

Study site: Central Apennines - Mount Acuto, Mount San Vicino, and Mount Terminillo.

	Classification		Description
	Horizon[a]	Horizon[b]	
	OLn	Fibric	Easily recognizable beech cupules, leaves, twigs and bark. Thickness of 2 to 5 cm. Absence of tree roots and micelia. Very few and few presence of small macrofauna and mesofauna.
	OLv	Hemic	Brownish and degraded beech cupules, leaves, twigs and partially degraded bark and beechnuts. Thickness of 2 to 11 cm. General absence of tree roots; where present they are very few or few. Micelia is present from few to plentiful. Presence of small macrofauna and mesofauna from few to abundant.
	OH	Sapric	Extensive decomposition, plant parts are not recognizable. Darkish beechnuts. Reduced thickness of 1 to 4 cm. Dark horizon. Tree roots vary from absent to abundant. Micelia are generally abundant. Small macrofauna and mesofauna are plentiful to abundant.

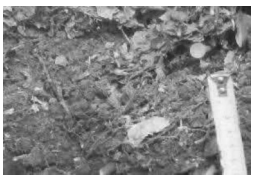
TEXAS

Study site: Lubbock - North Fork of the Brazos River; Houston - George Bush Intercontinental Airport, WG Jones State Forest, San Jacinto River and Sam Houston National Forest.

	Classification		Description
Lubbock			
	Horizon[a]	Horizon[b]	
	OLn	Fibric	Organic layer originated by deposition after an alluvial event. Recognizable branches, leaves, with a predominance of twigs and bark. Spot with accumulation of grass leaves. Thickness of 4 to 6 cm. No roots and micelia. No small macrofauna or mesofauna activities.
	OLv	Hemic	Brownish and darkish degraded vegetal material made by not easily recognizable leaves, twigs, and bark. Thickness from 5 to 10 cm. Roots and micelia are not present. Very few and few presence of small macrofauna and mesofauna.
Houston			
	OLn	Fibric	Non-decomposed pine leaves, pine cones, twigs and bark. Thickness of 1 to 5 cm. Absence of root and micelia. Few presence of small macrofauna and mesofauna

	OLv	Hemic	Brownish and pressed recognizable pine leaves, bark, and twigs. Thickness of 2 to 6 cm. General presence of tree roots from very few to few. Micelia generally goes from very few to abundant, with spots of very abundant presence. Considerable small macrofauna and mesofauna activities.
	OH	Sapric	Extensive decomposition, plant parts are not recognizable. Thickness of 1 to 2 cm. Tree roots vary from absent to few. Micelia is reduced from very few to plentiful. Small macrofauna and mesofauna are plentiful to abundant.

NEW MEXICO
Lincoln County - Lincoln National Forest.

	Classification		Description
	Horizon[a]	Horizon[b]	
	OLn	Fibric	Not decomposed pine and deciduous leaves, pine cones, twigs, and bark. In spots, woody parts of bark from degradation of dead trees. Thickness of 1 to 7 cm. Absence of roots and micelia. The presence of degraded tree parts induce a considerable presence of small macrofauna and mesofauna, in general absent or few.
	OLv	Hemic	Degraded pine leaves and deciduous twigs, presence of pine cones. Degraded barks reduced in fine dust, structure not recognizable. Brownish horizons. Thickness of 1 to 8 cm. Few roots. Micelia generally goes from few to plentiful. Considerable small macrofauna and mesofauna activities.
	OH	Sapric	Vegetal material completely decomposed, only pine cones are still recognizable. Thickness from 2 to 3 cm. Generally few tree roots. Micelia is few to plentiful. Small macrofauna and mesofauna are plentiful to abundant.

[a]horizon designation per Association Française pour l'Etude du sol (2008).
[b]horizon designation per Soil Survey Staff (2014).

.

Table 2. Summary statistics of soil properties.

Property	n	Mean	Std. Dev	Median	Min	Max
Calibration dataset						
TC	96	39.53	8.74	41.64	0.06	51.3
TN	96	1.08	0.38	1.13	0.04	1.87
Validation dataset						
TC	40	37.9	8.33	38.07	11.42	51.1
TN	40	1.14	0.44	1.26	0.32	1.86
Whole dataset						
TC	136	39.05	8.63	40.95	0.06	51.3
TN	136	1.1	0.4	1.13	0.04	1.87

Table 3. Validation statistics of multivariate models based on Kennard-Stone splitting.

Soil Property (%)	Approach	Model	R^2	PLS LF[a]	RMSE (%)	RPD	Bias	RPIQ
Total C	VisNIR DRS	PSR	0.80	-	3.46	2.40	0.15	3.02
		PLS	0.82	14	3.27	2.52	-0.15	3.23
		ENET	0.73	-	4.24	1.96	0.53	2.50
		RF	0.60	-	5.14	1.62	-0.03	2.06
	PXRF	PLS	0.53	5	5.59	1.49	0.89	1.89
		ENET	0.62	-	5.03	1.65	0.04	2.10
		RF	0.58	-	5.30	1.57	0.20	2.00
	VisNIR DRS+PXRF	PSR+RF	0.87	-	3.01	2.75	0.31	3.50
		PLS+RF	0.89	14	2.94	2.84	0.22	3.60
		ENET+RF	0.79	-	3.73	2.23	0.56	2.84
	GRA	PSR.RF	0.59	-	5.27	1.58	-0.13	2.01
		PLS.RF	0.60	14	5.15	1.62	-0.13	2.06
		ENET.RF	0.62	-	5.05	1.65	-0.11	2.10
Total N	VisNIR DRS	PSR	0.81	-	0.178	2.46	0.03	3.80
		PLS	0.75	13	0.214	2.01	0.03	3.25
		ENET	0.74	-	0.221	2.00	0.05	3.14
		RF	0.61	-	0.271	1.62	-0.01	2.56
	PXRF	PLS	0.28	4	0.492	0.89	0.02	1.41
		ENET	0.32	-	0.356	1.23	0.01	1.95
		RF	0.57	-	0.285	1.54	0.01	2.43
	VisNIR DRS+PXRF	PSR+RF	0.85	-	0.160	2.62	0.03	4.12
		PLS+RF	0.82	13	0.191	2.30	0.05	3.59
		ENET+RF	0.78	-	0.203	2.17	0.06	3.42
	GRA	PSR.RF	0.79	-	0.197	2.22	0.03	3.51
		PLS.RF	0.64	13	0.262	1.68	0.02	2.65
		ENET.RF	0.65	-	0.256	1.71	0.03	2.70

[a]PLS LF, partial least squares latent factor.

Table 4. Relative improvement of % RMSE and RPIQ in fused models.

Soil Property (%)	Approach	Model	RMSE (%)	RPIQ	Comparing Approach	Model	% Improvement in RMSE	% Improvement in RPIQ
Total C	VisNIR DRS	PSR	3.46	3.02	VisNIR DRS+PXRF	PLS+RF	15.02	19.20
		PLS	3.27	3.23			10.09	11.45
		ENET	4.24	2.50			30.66	44.00
		RF	5.14	2.06			42.80	74.75
	PXRF	PLS	5.59	1.89			47.40	90.47
		ENET	5.03	2.10			41.55	71.42
		RF	5.30	2.00			44.50	80.00
	GRA	PSR.RF	5.27	2.01			44.21	79.10
		PLS.RF	5.15	2.06			42.90	74.75
		ENET.RF	5.05	2.10			41.78	71.42
Total N	VisNIR DRS	PSR	0.178	3.80	VisNIR DRS+PXRF	PSR+RF	10.11	08.42
		PLS	0.214	3.25			25.23	26.76
		ENET	0.221	3.14			27.60	31.21
		RF	0.271	2.56			40.95	60.93
	PXRF	PLS	0.492	1.41			67.47	192.19
		ENET	0.356	1.95			55.05	111.28
		RF	0.285	2.43			43.85	69.54
	GRA	PSR.RF	0.197	3.51			18.78	17.37
		PLS.RF	0.262	2.65			38.93	55.47
		ENET.RF	0.256	2.70			37.50	52.59

Table 5. Prediction accuracies of Total C and Total N using whole geographical area holdout validation. Models were created using PLS+RF model (for Total C) and PSR+RF model (for Total N) with three geographical areas and then validated using the fourth area. Eight separate models were calibrated and validated.

Validation location	Parameter	RMSE (%)	RPD	Bias	RPIQ
Italy	Total C	5.26	0.92	1.19	1.07
Lubbock		2.84	1.57	-2.04	2.73
Houston		5.16	1.92	-0.41	1.26
New Mexico		12.55	0.80	-0.52	0.70
Italy	Total N	0.28	0.99	-0.11	1.30
Lubbock		0.16	0.72	-0.09	0.93
Houston		0.24	1.15	0.13	1.26
New Mexico		0.99	0.40	0.42	0.58

Table 6. Prediction accuracies of Total C and Total N using whole-organic horizon holdout validation. Models were created using PLS+RF (for Total C) and PSR+RF (for Total N) with two organic horizons and then validated using the third organic horizon. Six separate models were calibrated and validated.

Validation location	Parameter	RMSE (%)	RPD	Bias	RPIQ
OH	Total C	9.22	0.77	5.90	1.14
OLn		10.32	0.90	3.42	1.05
OLv		7.17	1.01	-1.83	0.79
OH	Total N	0.47	0.82	0.23	1.16
OLn		1.19	0.31	0.16	0.40
OLv		0.34	1.25	0.05	2.00

Figure1

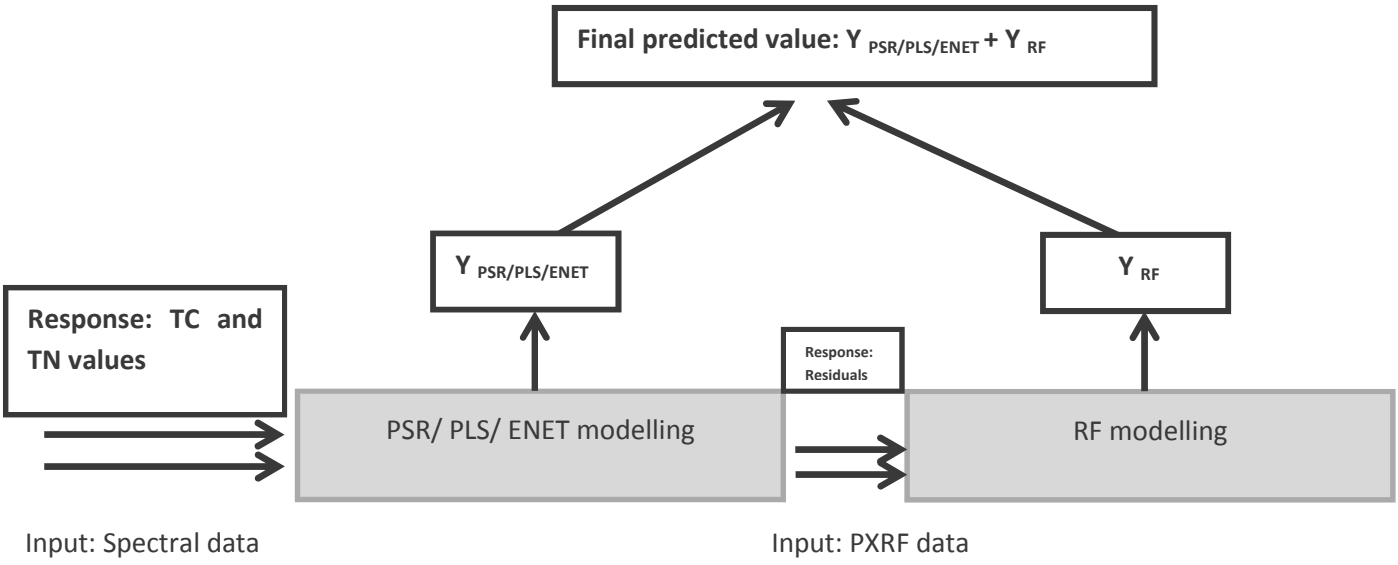


Fig. 1

Figure2

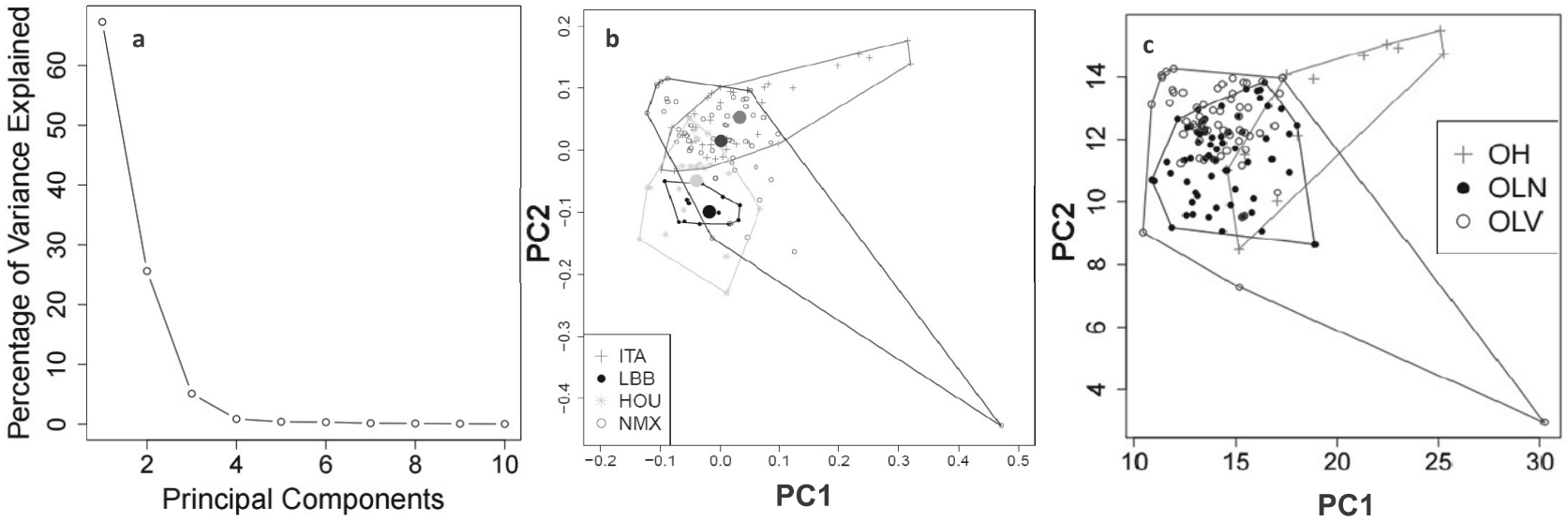


Fig. 2

Figure3

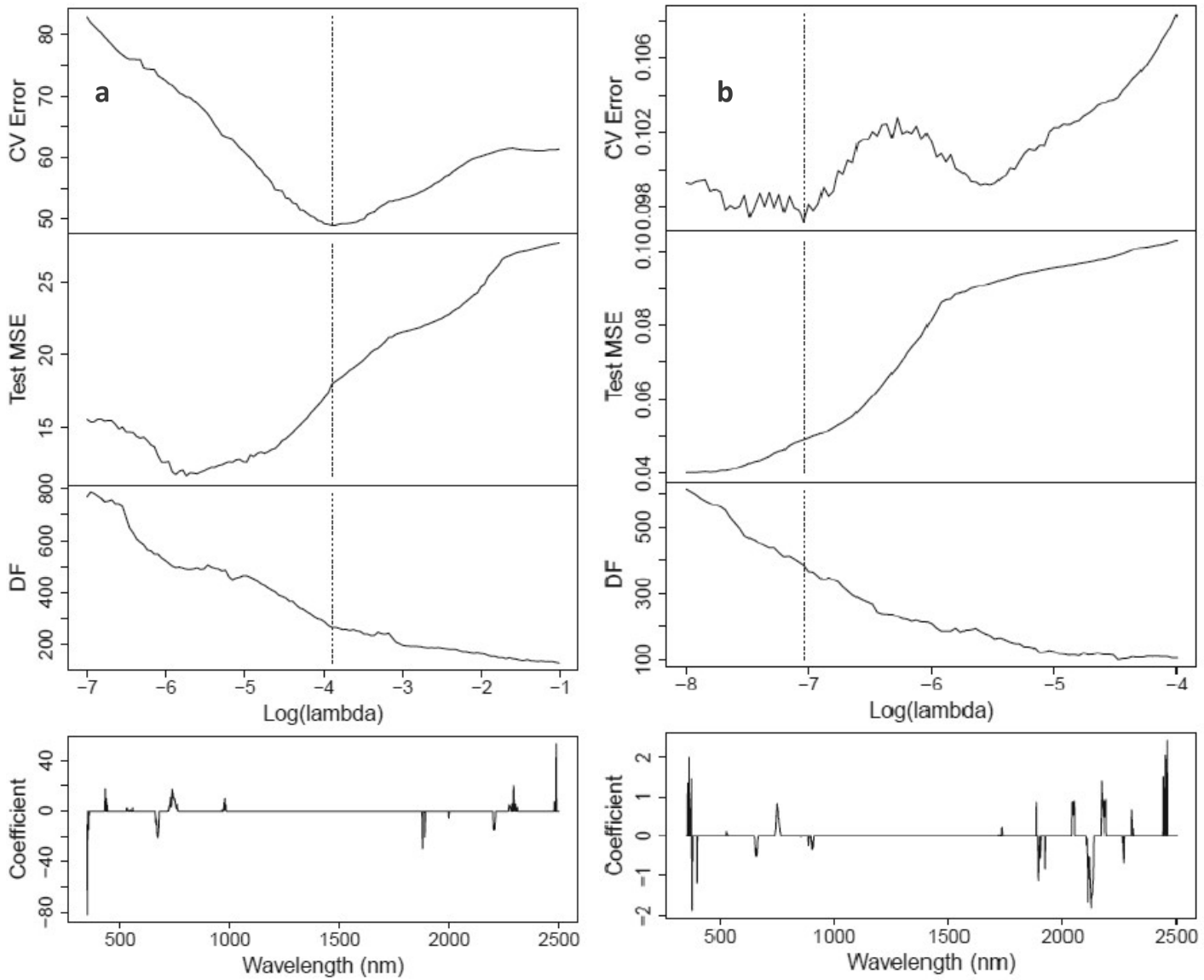


Fig. 3

Figure4

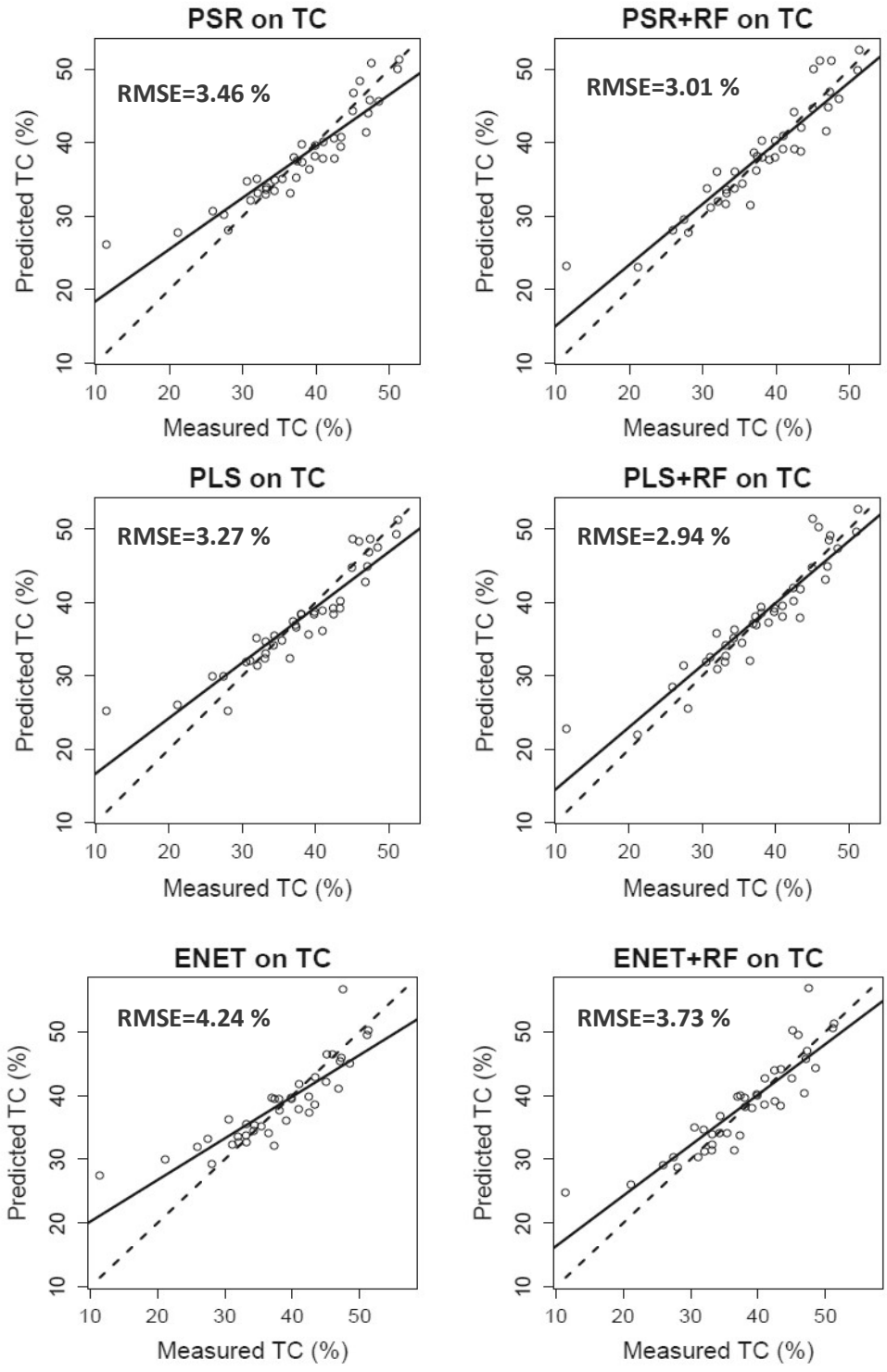


Fig. 4

Figure5

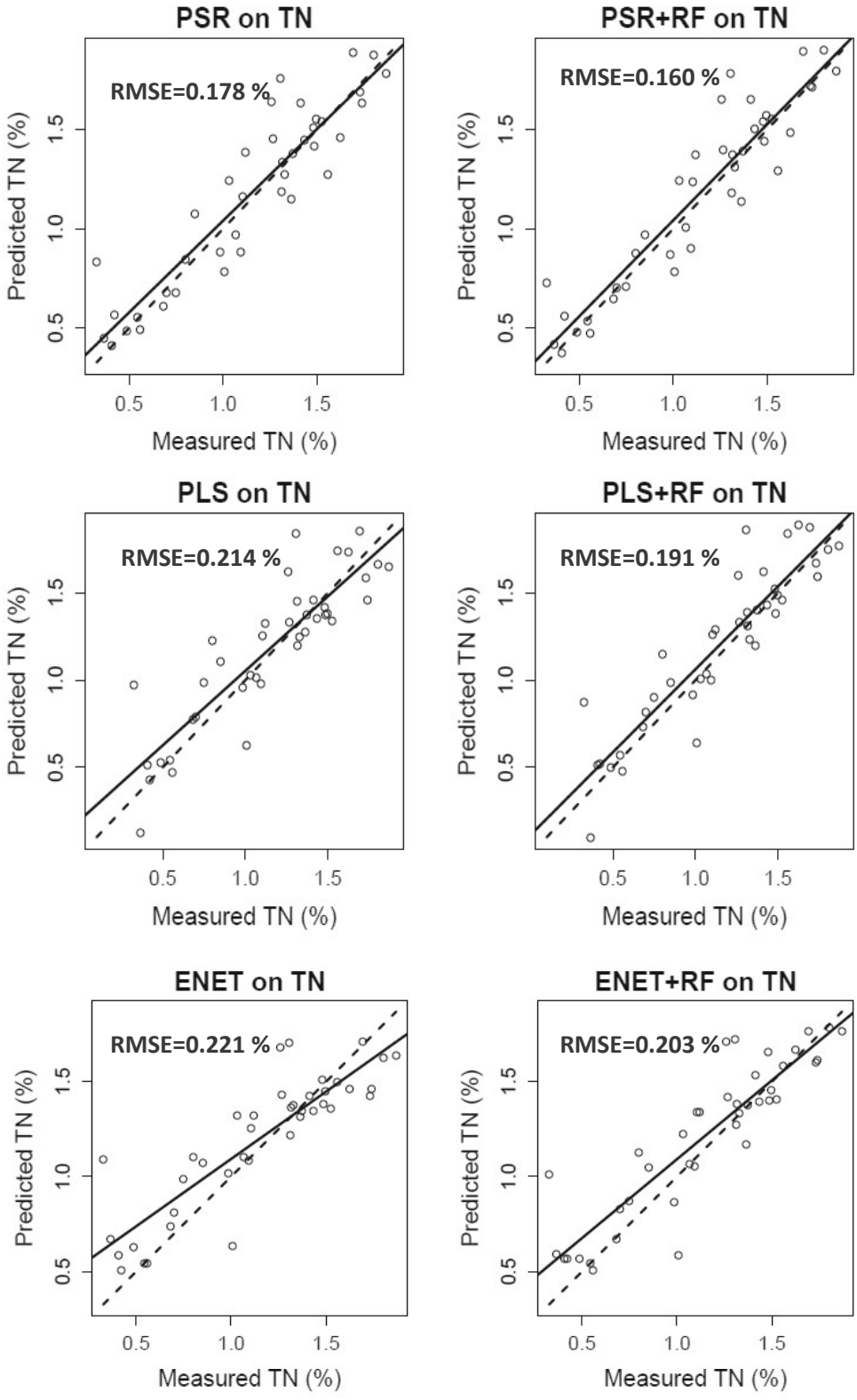


Fig. 5

Figure6

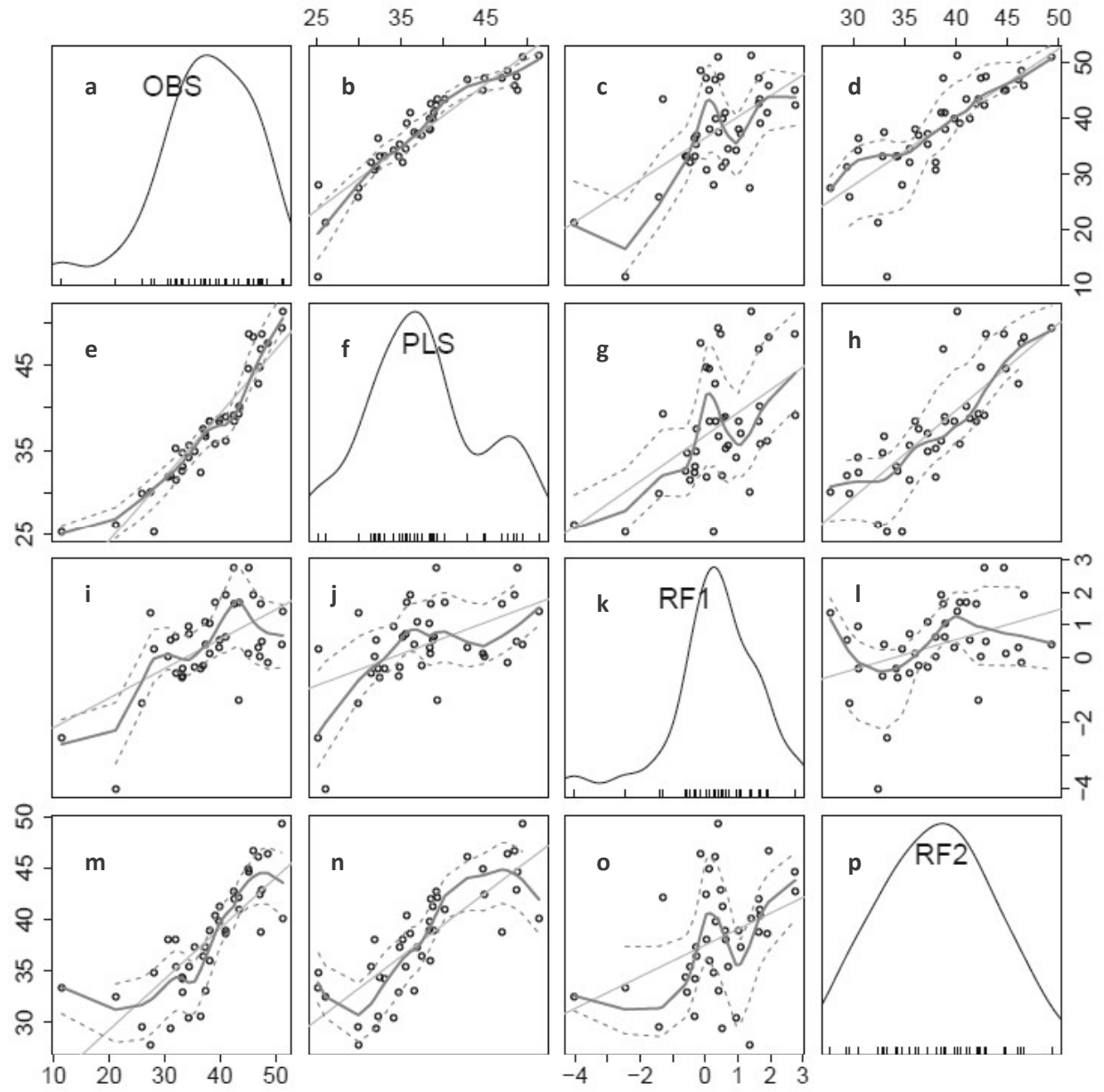


Fig. 6