



Prediction of β -glucosidase and β -glucosaminidase activities, soil organic C, and amino sugar N in a diverse population of soils using near infrared reflectance spectroscopy

Warren A. Dick^{a,*}, Basanthi Thavamani^a, Shannon Conley^a, Robert Blaisdell^b, Aditi Sengupta^a

^aSchool of Environment and Natural Resources, The Ohio State University, The Ohio Agricultural Research and Development Center, Wooster, OH 44691, USA

^bDepartment of Rangeland Ecology and Management, Texas A&M University, College Station, TX 77843, USA

ARTICLE INFO

Article history:

Received 2 September 2011

Received in revised form

2 April 2012

Accepted 5 April 2012

Available online 21 April 2012

Keywords:

Near infrared reflectance spectroscopy

β -Glucosidase

β -Glucosaminidase

Soil organic C

Soil enzymes

Amino sugar N

ABSTRACT

There is a need for methods that can rapidly measure multiple biological properties simultaneously. A near infrared spectroscopy (NIRS) method was used to predict β -glucosidase (EC 3.2.1.21) and β -glucosaminidase (NAGase, EC 3.2.1.52) activities, and soil organic C and amino sugar N concentrations in 184 diverse soils of Ohio. The laboratory-measured values of the variables were calibrated against NIR spectral data with partial least squares regression analysis. Statistical analysis of the spectral data was done using the multivariate analysis software Unscrambler 8.0 (CAMO Inc). The first differential transformation of the spectral data in the NIR region (1100–2498 nm) generally yielded best results for developing multivariate calibration models. The multivariate models developed were validated using the full cross validation method and the test set method with a test set size of approximately 45 samples. The R^2 values, testing variation between concentrations as measured by the NIR method and chemical methods, were 0.91 for organic carbon (OC), 0.92 for amino sugar N, and 0.82 for both soil β -glucosidase and β -glucosaminidase enzyme activities. Our study showed that the NIRS method has the potential to simultaneously, rapidly and accurately predict values of multiple related variables. The equipment needed for the NIRS method is not expensive and can be used where very large numbers of samples need to be rapidly analyzed. Indeed, the prediction equations can be constantly improved as more data points are entered into the correlations between laboratory-measured values and NIRS values.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The measurement of soil properties often requires a separate standard methodology for each individual property. This can be tedious and costly, especially if large numbers of samples are to be analyzed. As an alternative to standard methodologies, the near infrared reflectance spectroscopy (NIRS) approach provides the opportunity for rapid, concurrent, and inexpensive analysis (Ben-Dor and Banin, 1995; Chang et al., 2001). The process is non-destructive and requires a minimal amount of sample (Chang et al., 2001). Extensive sample preparation is typically unnecessary, which eliminates chemical waste and conserves time (Rinnan and Rinnan, 2007). Because large numbers of samples can be analyzed quickly, statistical confidence in the calculations can easily be improved upon (Janik et al., 1998).

To detect the presence of analytes, NIRS measures the diffuse reflectance of electromagnetic radiation created by the vibrational

modes of molecular bonds (Ben-Dor and Banin, 1995; Rinnan and Rinnan, 2007). Physical characteristics, such as particle size and arrangement, also affect NIR spectra. To quantify the raw spectral data, multivariate calibration is used to establish a linear relationship between absorption calculations [A (absorbance) = $\log(1/R)$ (reflectance)] and analyte concentration (Ben-Dor and Banin, 1995; Chang et al., 2001). If an analyte does not absorb light in the NIR range, correlating properties can be measured instead (Terhoeven-Urselmans et al., 2006).

Popular NIRS calibration methods for soil property analysis include Fourier regression, partial least square regression (PLSR), neural networks, principal component regression (PCR), stepwise multiple linear regression (SMLR) and locally weighted regression (LWR) (Chang et al., 2001). To select and verify the calibration model, a validation process is employed that compares the concentration predicted from the calibration equation with sample concentrations measured using a standard analytical technique (Ben-Dor and Banin, 1995) for a subset of the samples. The degree of correlation between measured and predicted results (R^2) and the ratio between measurement error and prediction error (RPD) are commonly used

* Corresponding author. Tel.: +1 330 263 3877; fax: +1 330 263 3658.
E-mail address: dick.5@osu.edu (W.A. Dick).

to assess the accuracy of the NIRS model (Ben-Dor and Banin, 1995; Cohen et al., 2005).

Both soil organic C and amino sugar N are fractions of soil organic matter and major sources of nutrients for plants and soil organisms (Bauer, 1994). Soil organic C fuels the metabolic activity of microorganisms (Bauer, 1994) and is often considered an important indicator of soil productivity (Reeves, 1997; Shukla et al., 2006). β -glucosidase, which releases glucose, also helps provide energy to soil microorganisms and is directly related to soil organic C content (Eivazi and Tabatabai, 1988; Bandick and Dick, 1999). This enzyme's activity has been used to closely monitor rapid changes in soil organic C brought about by soil management effects (Bandick and Dick, 1999).

Despite such advantages, the use of enzyme activities to indicate soil quality or to study soil functions is often restricted to biochemical or microbiological laboratories. However, basic soil properties such as mineral content, total organic C and N, pH, C and N mineralization, biomass C, soil respiration and moisture content have been successfully determined by spectral methods (Chang et al., 2001; Cohen et al., 2005). For example, NIRS models have been tested and developed for addressing regional soil performance issues or "specific soil queries." The NIRS method was used to explore the effects of a wildfire and earthworm activity on soil ecosystem functions (Cécillon et al., 2009; Cassagne et al., 2008) and for the determination of Florida wetland soil quality (Cohen et al., 2005). Moros et al. (2009) created NIRS models to determine toxic metal concentration, lime content, organic matter and electrical conductivity for soils in the Murcia region of Spain. NIRS has also been used to rapidly determine activities for the soil enzymes β -glucosidase, acid and alkaline phosphatases, arylsulfatase, dehydrogenase, peptidase, and urease (Cohen et al., 2005; Mimmo et al., 2002; Reeves et al., 2000; Zornoza et al., 2008).

The measurement of mineralizable soil N is crucial for determining the needed amounts of fertilizer, but a standardized predictive measurement procedure has been difficult to establish. Chemical indicators, such as NO_3^- , often yield inaccurate results (Mulvaney et al., 2001). The NIRS method could improve the feasibility of using biological indicators, such as β -glucosaminidase activity, for rapidly estimating N mineralization. Amino sugar N analysis has emerged as a potentially improved procedure for gauging mineralizable soil N content (Mulvaney et al., 2001). The activity in soil of β -glucosaminidase, which produces amino sugars, has also been associated with the N acquisition process of microorganisms (Parham and Deng, 2000; Sinsabaugh and Moorhead, 1995). In reality, a suite of biological C and N indicators may be much more accurate in assessing N mineralization in soil, and thus native soil N fertility, than a single soil parameter.

There is, thus, a need for methods that can rapidly measure multiple biological properties simultaneously. Measurement of multiple soil properties at one time is clearly feasible using the NIRS method. Our research objective, therefore, was to develop and test a NIRS model for the prediction of organic C content, amino sugar N concentration, and the activity of the soil enzymes β -glucosidase and β -glucosaminidase, in a set of diverse Ohio soils.

2. Materials and methods

2.1. Soil sample collection and storage

The majority of the soil samples used in the development of the NIRS calibration model were chosen at random from the Soil Testing Laboratory at the Ohio Agricultural Research and Development Center (Wooster, OH). The exact origins of sites in Ohio were not known for 169 of the total 184 samples. The total amount of samples (184) collected exceeded the recommended minimum

(150) of soil samples that are needed to develop NIRS predictive equations (Windham et al., 1989). Collected samples were air-dried and sieved through a 2 mm mesh screen. Soil samples were stored for several weeks to several months in cardboard boxes at room temperature.

2.2. Standard soil property methodologies

The organic C content was determined in the STAR Laboratory at the Ohio State University, Wooster, OH (<http://oardc.osu.edu/starlab/>) by correlating loss-on-ignition values (Storer, 1984) with organic C values (International Organization for Standardization, 1995) using a regression equation ($R^2 = 0.9929$) developed after analyzing a very large number of samples in the STAR Laboratory. The amino sugar N concentration, a fraction of total organic C and N, was estimated using the Illinois Nitrogen Soil Test by incubating samples at 55 °C for 5 h in 2 M NaOH as described by Khan et al. (2001). Spectrophotometric assays were used to measure the activity of β -glucosidase (Eivazi and Tabatabai, 1988) by incubating 1 g soil (air-dry basis) for 1 h with *p*-nitrophenyl- β -D-glucoside (Sigma Chemical Co., St. Louis, MO, USA) at pH 6.0 (modified universal buffer). β -glucosaminidase (Parham and Deng, 2000) was measured by incubating 1 g soil for 1 h with *p*-nitrophenyl-N-acetyl- β -D-glucosaminidine (Sigma Chemical Co.) at pH 5.5 (acetate buffer).

The properties measured, along with the references for the methods used to measure these properties, are listed in Table 1. Statistical correlations were performed to determine inter-variable relationships between these soil properties.

2.3. Spectral analysis of the soil samples

A FOSS-TECATOR 6500 spectrometer was used to collect spectral data for the soil samples. Prior to analysis, samples were oven-dried (60 °C) for twelve hours and stored in a desiccator. Using a rotating cup with a quartz lens, reflectance measurements were taken from 400 to 2498 nm that covered both the visible and near infrared regions of electromagnetic radiation. Data were collected every 8 nm at a resolution of 2 nm resulting in 1050 data points for each sample. Prior to preprocessing, the reflectance data was converted to units of $\log(1/R)$ and the measured soil property data were entered into a spreadsheet as X-values and Y-values, respectively.

Unscrambler version 8.0 (CAMO Inc., 2003) software was used to process the spectral data. Two data sets were created using 700 data points from the 1100–2498 nm NIR wavelength region and 1050 data points from the 400–2498 nm visible-NIR wavelength region. The 400–2498 nm, or extended, data set was reduced by averaging every four adjacent spectral data points to produce 263 new data points. Similarly, the 1100–2498 nm data set was reduced to 175 data points from 700 spectral data points.

Table 1

Soil C enzymes and properties measured in 184 soil samples (air-dried).

Carbon enzymes and properties ^a	Mean	SD	Range	Method reference
Soil organic C (g kg ⁻¹ soil)	24.1	18.0	1.1–127	Storer, 1984
Amino sugar N (mg kg ⁻¹ soil)	216	96.9	9.04–614	Khan et al., 2001
β -glucosidase (mg kg ⁻¹ soil h ⁻¹)	87.5	62.7	0.27–328	Eivazi and Tabatabai, 1988
β -glucosaminidase (mg kg ⁻¹ soil h ⁻¹)	34.4	25.2	0.28–116	Parham and Deng, 2000

^a For each C enzyme or property measured, the number of samples (*n*) in the data set is 184. SD = Standard deviation.

2.4. Multivariate model development, calibration and validation

To prepare the raw spectral data for multivariate analysis, the data underwent several mathematical preprocessing methods including standardization, normalization, 1st differential transformation, 2nd differential transformation, and multiple scatter correction (MSC). Standardization, which uses a scaling factor to prevent range-based domination of one variable over another, improved calibration. However, normalization, which is performed to compensate for measurement variances, did not improve calibration. First derivative preprocessing and second derivative preprocessing were performed on the standardized reflectance values. Using Savitzky–Golay derivatives, the differentiation was performed with second-order polynomial and smoothing with window width of 1 nm. First and second derivative transformations differentiate and separate the absorption peaks (Fig. 1B and C). Multiplicative scatter correction (MSC) handles the spectral scatter effects, similar to the second differential transformation and performs the amplification and offset correction. However, MSC did not improve calibration.

Data were calibrated with laboratory-measured values of the soil properties using partial least squares regression (PLS-R) analysis (Table 2). Geometrically, data points can be viewed as swarms of points in variable space and the behavioral dimension of the swarms is described by partial least squares components. In order to detect and model the structural pattern obtained by the interrelations of the variables in space, PLS-R analysis used the Y-data (i.e. organic C, amino sugar N concentrations, and β -glucosidase and β -glucosaminidase activities) structure to decompose the X-matrix data (i.e. the spectral data) to get an optimal regression (Esbensen, 2002) which produced the lowest standard error of calibration (RMSEC) and highest regression coefficient (R^2) of the calibration. The two best models, models 1 and 2, were used for validation purposes; however, in the case of β -glucosaminidase, only one model provided good results and only this model was used (Table 2).

Two types of validation were used in this study; test-set validation and full-cross validation (Table 2). In test-set validation, a separate sample set is used. The test-set method is generally considered the best method of validation (Esbensen, 2002) as the samples in the test-set are not used in the calibration procedure. However, it requires more samples so as to have one set for model development and one set for testing the model. In this study, test-set validation was performed with a randomly selected set of approximately 45 soil samples and this subset of samples was not used in the development of the calibration models. In full-cross validation method, the calibration and the validation sample sets are the same. As many subsets of samples are made as there are samples. Each subset has all the samples except one (Stark, 1988). These subsets are used for validating the model. The squared difference between the NIR-predicted value and the Y-value (organic C, amino sugar N concentrations, and β -glucosidase and β -glucosaminidase activities) for each omitted sample was summed and averaged giving the Y-variance. Full-cross validation method is often preferred over test-set validation when the sample size is limited. The optimum regression equations were determined as those that produced the lowest standard error of prediction (RMSEP) and highest regression coefficient (R^2) of the validation.

3. Results

3.1. Standard soil property measurements

For the full set of 184 samples, the organic C values ranged from 1.1–127 g/kg soil with a narrow concentration range of 10–30 g/kg soil for 73% of the samples. The amino sugar N values ranged from

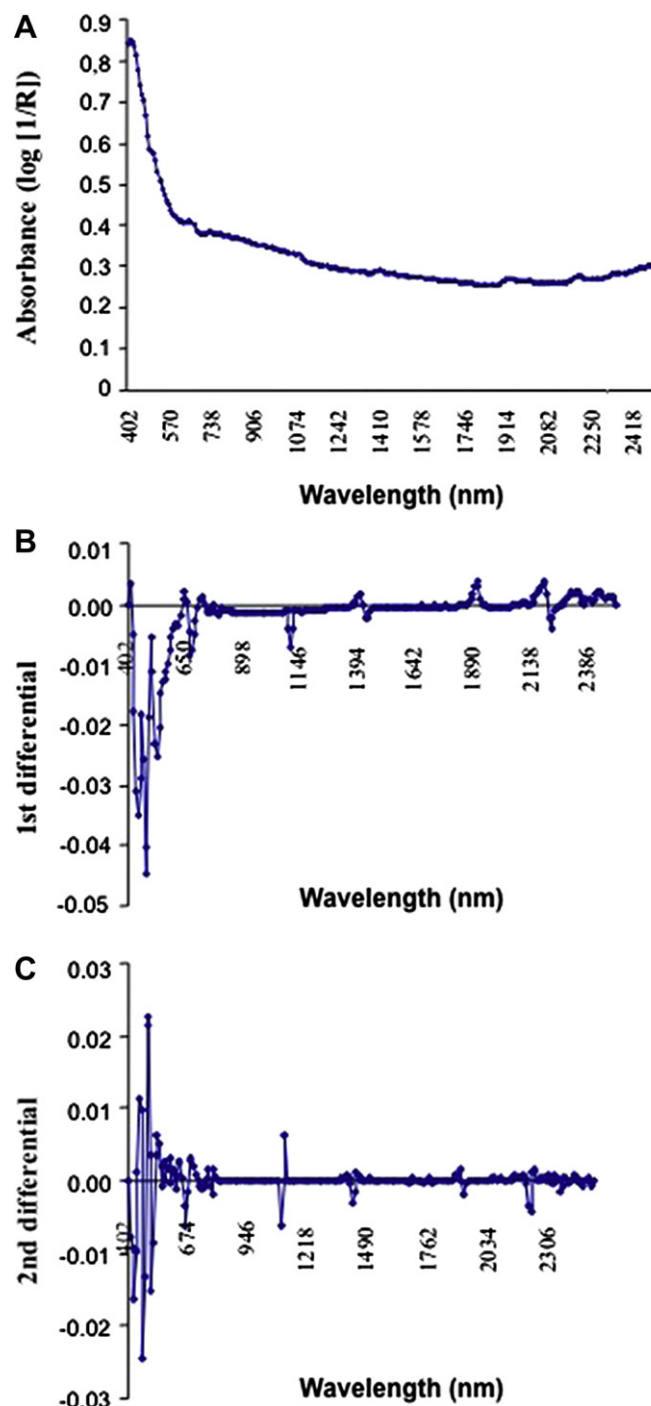


Fig. 1. Spectrum of a soil sample: (A) raw spectral data; (B) first differential of spectral data; and (C) second differential of spectral data.

9.04–614 mg N/kg soil with the concentration range of 100–380 mg N/kg soil representing 92% of the samples. β -glucosidase activity values ranged from 0.27–328 mg/kg soil/h with activity levels of 83% of the samples distributed over a range of 0–120 mg/kg soil/h. β -glucosaminidase activity values ranged from 0.28–116 mg/kg soil/h with activity levels of 90% of the samples distributed over a range of 0–70 mg/kg soil/h. Statistical analyses indicated that organic C was strongly related to amino sugar N ($R^2 = 0.88$), β -glucosidase activity ($R^2 = 0.70$) and β -glucosaminidase activity ($R^2 = 0.57$). The relationship between amino sugar N and β -glucosaminidase was also strongly related ($R^2 = 0.65$).

Table 2
Assessment of the soil properties calibration and validation models.^a

Carbon enzymes and properties ^b	<i>n</i>	C	T	<i>R</i> ²	RMSEC (RMSEP)	CV	Method
<i>Calibration model</i>							
SOC 1 ^c	181	5	1st	0.92	6.6	27.3	n.a.
SOC 2	132	7	2nd	0.96	5.5	22.8	n.a.
Amino sugar 1	168	8	1st	0.92	31.5	14.5	n.a.
Amino sugar 2	130	5	1st	0.88	41.6	19.3	n.a.
β-glucosidase 1	139	11	1st	0.90	27.1	30.9	n.a.
β-glucosidase 2	168	12	1st	0.88	28.2	32.2	n.a.
β-glucosaminidase	175	13	1st	0.89	11.2	32.5	n.a.
<i>Validation model</i>							
SOC 1	181	5	1st	0.91	7.2	29.8	Cross
SOC 2	45	7	2nd	0.83	7.0	29.0	Test
Amino Sugar 1	168	8	1st	0.90	37.0	17.1	Cross
Amino Sugar 2	48	5	1st	0.92	48.7	22.5	Test
β-glucosidase 1	139	11	1st	0.82	35.6	40.6	Cross
β-glucosidase 2	168	12	1st	0.78	37.5	42.8	Cross
β-glucosaminidase	175	13	1st	0.82	14.4	41.8	Cross

^a In all cases reported in this table, the NIR region (1100–2498 nm) data were used to develop both the calibration and validation models.

^b Units are SOC (g kg⁻¹ soil), amino sugar N (mg N kg⁻¹ soil), β-glucosidase (mg *p*-nitrophenol kg soil⁻¹ h⁻¹) and β-glucosaminidase (mg *p*-nitrophenol kg soil⁻¹ h⁻¹).

^c Abbreviations: 1,2 = Models 1 and 2; SOC = soil organic carbon; *n* = number of samples; C = number of components; T = type of differential transformation (1st or 2nd); *R*² = regression coefficient; RMSEC = root mean square error of calibration; RMSEP = Root mean square error of prediction; CV = coefficient of variation (%), and n.a. = not applicable.

3.2. Multivariate calibration model development

The multivariate development process was used to produce different calibration models for soil organic C content (24 models), amino sugar N content (16 models), and activities of β-glucosidase and β-glucosaminidase (12 models each). To achieve the best calibration results, different combinations of data transformation, spectral regions, sample set sizes and validation methods were tested. For both soil organic C and amino sugar N, significantly better values for *R*² and RMSEC were observed using first differential transformation of data points in the NIR wavelength region (1100–2498 nm) than in the extended visible-NIR spectral region (400–2498 nm).

Calibration Models 1 and 2 for organic C had *R*² values of 0.92 and 0.96, respectively (Table 2). Calibration Models 1 and 2 for amino sugar N had *R*² values of 0.92 and 0.88, respectively. However, the RMSEC values were much larger for the amino sugar N values than for the organic C values. This is not surprising as the actual values for amino sugar N were also larger than for organic C. The *R*² values for β-glucosidase Models 1 and 2 for the NIR wavelength region (1100–2498 nm) were 0.90 and 0.88, respectively (Table 2). For the β-glucosaminidase activity values, little difference was observed between first and second differential transformations of both the extended (400–2498 nm) and the NIR (1100–2498 nm) data sets. Thus only the first differential transformation for the NIR data set was used to develop the calibration model (*R*² was 0.89, Table 2).

In an attempt to produce more representative models, outlier values from our random population of Ohio soils that exhibited extremes in soil organic C values or other soil properties such as pH, salt concentrations, etc. were removed. This creates subsets of the samples with characteristics more typical of Ohio soils. For example, removal of 13 samples (or 7.1% of the sample population) resulted in 2.6-fold reduction in the range of organic C contents. Similarly removal of 6.5% of the sample population (i.e. 12 samples) resulted in 1.8-fold reduction in the variation of amino sugar N

content. For enzyme activities, removal of 17% of the sample population (i.e. 32 samples) resulted in 2.8-fold reduction in the variation of β-glucosidase activity values while removal of 16% of the sample population (i.e. 29 samples) resulted in 2.0-fold reduction in the variation of β-glucosaminidase activity values. However, for the model calibrations described, only the most extreme three to six outlier sample values were removed from the datasets, as too rigorous removal of outliers can compromise the overall effectiveness of the NIR approach for estimating soil properties.

3.3. Model validation

The best multivariate model (SOC 1) developed to predict organic C values from NIR spectral data had a validation coefficient value of 0.91 and an RMSEP value of 7.2 g/kg of soil (Table 2; Fig. 2). The spectral data were preprocessed with first differential transformation and used only the NIR region (1100–2498 nm) of the spectrum. This model was validated by the full-cross validation method using 181 soil samples and explained 83.5% of the Y variance (measured organic C values) and >90% of X variance (spectral data). Another promising model (SOC 2) for predicting organic C had an *R*² value of 0.83 for model validation and an RMSEC value of 7.0 g/kg of soil (Table 2; Fig. 2). This model explained 67.5% of Y variance and >95% X variance and was calibrated with 132 soil samples and validated with a test-set of 45 samples.

The best multivariate model developed to predict amino sugar N values from NIR spectral data had an *R*² value of 0.90 and an RMSEP value of 37.0 mg N/kg of soil (Table 2; Fig. 2). The spectral data used for this model were preprocessed with first differentially transformation based on NIR region (1100–2498 nm) of the spectrum using 168 samples. The model was validated with full-cross validation method. The model explained 80% of Y variance (measured amino sugar N values) and >80% of X variance (spectral data). The second most promising validation model had an *R*² value of 0.92 and an RMSEP value of 48.7 mg N/kg soil (Table 2; Fig. 2). This model used 130 samples for calibration and 48 samples for validation with test set method. This model explained 72.5% of the Y variance and >80% of X variance.

The best multivariate model developed for predicting β-glucosidase activity had an *R*² value of 0.82 and an RMSEP value of 35.6 mg/kg soil/h (Table 2; Fig. 2). The spectral data was preprocessed with first differential transformation using the NIR region (1100–2498 nm) of the spectrum. The model was validated using full-cross validation method with 139 soil samples. The next best model had an *R*² value of 0.78 with an RMSEP value of 37.5 mg/kg soil/h (Table 2; Fig. 2). This model was also based on spectral data from the NIR region (1100–2498 nm) and processed with first differential transformation. The model was validated using full-cross validation method with 168 soil samples. The predicted errors in both models were rather large and about half of the mean value of β-glucosidase activity for these samples.

The best multivariate model developed for predicting β-glucosaminidase activity had an *R*² value of 0.82 and an RMSEP value of 14.4 mg/kg soil/h (Table 2; Fig. 2). The spectral data was preprocessed with first differential transformation using the NIR region (1100–2498 nm) of the spectrum. The model was validated using full-cross validation method with 175 soil samples. The model explained 66.5% of Y variance (measured β-glucosaminidase values) and 98% of X variance (spectral data). The test-set validation method was not successful in predicting β-glucosaminidase activity values because the test-set size (27% of the size of calibration set) was not at the minimum level needed to validate the calibrations developed and was not successful in improving the predicted results.

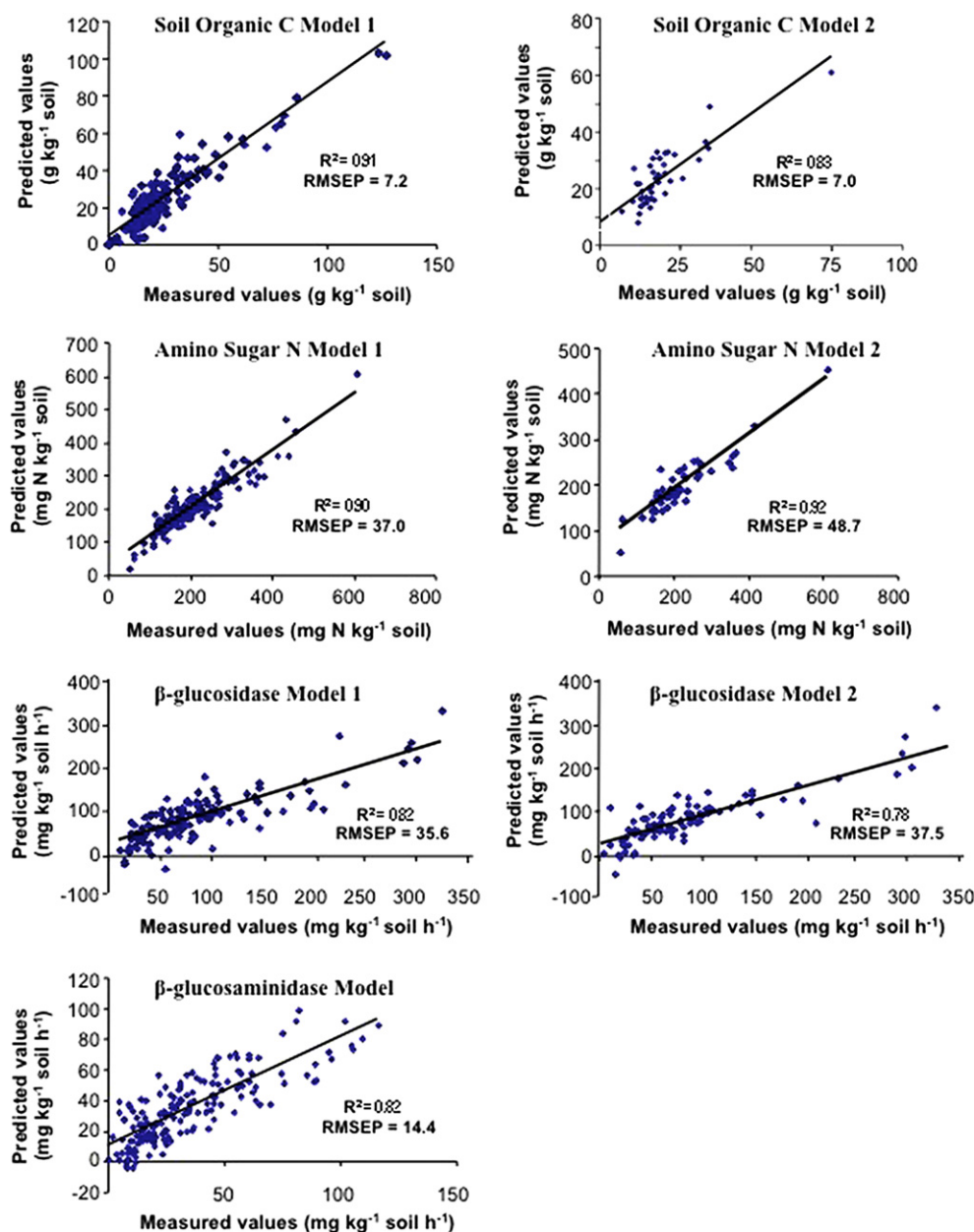


Fig. 2. Plots depicting the relationship of measured and NIRS predicted values. The validation model descriptions are provided in Table 2 but are summarized as follows. The full-cross validation model was used for the soil organic C Model 1, amino sugar N Model 1, β -glucosidase Model 1 and Model 2 and the β -glucosaminidase Model. The cross-test validation model was used for soil organic Model 2 and amino sugar N Model 2.

4. Discussion

Previous studies have reported a strong relationship between soil organic C and β -glucosidase activity (de la Paz Jimenéz, 2002; Eivazi and Tabatabai, 1990). This enzyme's activity has been used to closely monitor rapid changes in soil organic C brought about by soil management effects (Bandick and Dick, 1999). The activity of β -glucosidase is also often used as a component in multiparametric indices estimating soil organic matter content (Bastida et al., 2008) because this enzyme is thought to participate in the degradation of organic matter (de la Paz Jimenéz, 2002). The enzyme, β -glucosaminidase, participates in amino sugar production (Ekenler and Tabatabai, 2003) that eventually leads to production of mineral N in soil.

Our goal was to determine whether we could use the NIRS method to rapidly measure biological parameters in a diverse set of

soils from Ohio, USA. In this study, NIR reflective spectroscopy was found to be a potentially valuable tool to rapidly and simultaneously estimate soil organic C, amino sugar N, β -glucosidase activity and β -glucosaminidase activity. It is non destructive of the samples and does not involve the use of any chemical reagent. For β -glucosidase activity, calibrations using the NIR region (1100–2498 nm) compared to the extended spectral region (400–2498 nm) produced better R^2 values but with higher RMSEC values. Calibration based on the extended spectral region (400–2498 nm) had poor correlations probably due to background noise caused by signals for variables other than the ones tested.

Test-set and full-cross validations were performed using the calibration models. For the test set validations, a subset of samples were randomly chosen from the full set of samples and this subset of samples was not used to produce the calibration models. For soil organic C, the test-set method produced lower RMSEP than the

cross validation method, but for amino sugar N, the test-set method and the cross validation method yielded essentially the same result (Table 2). However, the test-set method gave very poor validation results for β -glucosidase and β -glucosaminidase and only the results of the cross validation method are shown in Table 2 for these enzymes. It is believed that the test set yielded poor results because of an insufficient test set size. Although the test sample size was 27% of the size of the calibration set, the same amount of samples for calibration and test set validation has been recommended (Esbensen, 2002).

The coefficient of variation (CV) values for the different validation models varied from a low of 17.1% for the amino sugar N model 1 to a high of 42.8% for β -glucosidase model 2 (Table 2). The precision of the NIRS estimates for β -glucosidase, β -glucosaminidase, soil organic C, and amino sugar N can be increased, however, by constantly adding data points into the correlations between laboratory-measured values and NIRS values.

To obtain good NIRS predicted values for soil biological properties, it is important that all sample parameters are standardized so that there is consistency in how measurements are made in the laboratory and by the NIRS method. For example there should be consistent and close timing between when the NIRS data and the laboratory-measured data are obtained. In this study, air-dried samples were used to measure the biological properties in the laboratory and the same samples were then oven-dried before spectral data were obtained. This was all completed within a matter of a few weeks. However, the slight difference in time between when air-dried soil samples were analyzed in the laboratory and when samples were oven-dried and spectral data were obtained could have affected how well the NIRS method predicted soil organic C, amino sugar N, and β -glucosidase and β -glucosaminidase activities.

The NIRS method can be fine-tuned to reflect samples from a specific set of fields or soil type. In this study, we used soil samples from across the entire state of Ohio in the USA. This is a rather large region and certainly made it more difficult to develop a calibration model that could accurately reflect the diversity of soils in our population of soils from Ohio. However, as previously mentioned, the equations can be constantly improved by adding more data points into the calibrations models that relate laboratory-measured values and NIRS values.

5. Conclusions

The NIR spectroscopy method was successful in rapidly and accurately predicting soil organic C content, amino sugar N concentrations, β -glucosidase activity, and β -glucosaminidase activity. Indeed, the NIRS method clearly is able to measure multiple soil properties with a single spectral scan. The prediction equations can be constantly improved as more data points are entered into the correlations between laboratory-measured values and NIRS values. The equipment needed is not expensive and the NIRS method can be used where very large numbers of samples need to be rapidly analyzed.

References

Bandick, A., Dick, R., 1999. Field management effects on soil enzyme activities. *Soil Biology & Biochemistry* 31, 1471–1479.

Bastida, F., Zsolnay, A., Hernández, T., García, C., 2008. Past, present and future of soil quality indices: a biological perspective. *Geoderma* 147, 159–171.

Bauer, A., 1994. Quantification of the effect of soil organic matter content on soil productivity. *Soil Science of America Journal* 58, 185–193.

Ben-Dor, E., Banin, A., 1995. Near infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Science of America Journal* 59, 364–372.

CAMO Inc, 2003. The Unscrambler User Manual. 8.0. Woodridge, NJ, USA.

Cassagne, N., Spiegelberger, T., Cécillon, L., Juvy, B., Brun, J.J., 2008. The impact of soil temperature increase on organic matter and faunal properties in a frozen calcareous scree in the French alps. *Geoderma* 146, 239–247.

Cécillon, L., Cassagne, N., Czarnes, S., Gros, R., Vennetier, M., Brun, J.J., 2009. Predicting soil quality indices with near infrared analysis in a wildfire chronosequence. *Science of the Total Environment* 407, 1200–1205.

Chang, C.W., Laird, D.W., Mausbach, M.J., Hurburgh, C.R., 2001. Near-infrared reflectance spectroscopy – principal components regression analyses of soil properties. *Soil Science Society of America Journal* 65, 480–490.

Cohen, M.J., Prenger, J.P., DeBusk, W.F., 2005. Visible-near infrared spectroscopy for rapid, non-destructive assessment of wetland soil quality. *Journal of Environmental Quality* 34, 1422–1434.

de la Paz Jimenez, M., 2002. Soil quality: A new index based on microbiological and biochemical parameters. *Biology & Fertility of Soils* 35, 302–306.

Eivazi, F., Tabatabai, M.A., 1988. Glucosidases and galactosidases in soils. *Soil Biology & Biochemistry* 20, 601–606.

Eivazi, F., Tabatabai, M.A., 1990. Factors affecting glucosidase and galactosidase activities in soils. *Soil Biology & Biochemistry* 22, 891–897.

Ekenler, M., Tabatabai, M.A., 2003. Tillage and residue management effects on β -glucosaminidase activity in soils. *Soil Biology & Biochemistry* 35, 871–874.

Esbensen, K.H., 2002. Multivariate Calibration. In: Anonymous *Multivariate Data Analysis in Practice*, fifth ed. CAMO Process AS, Norway.

International Organization for Standardization, 1995. *Soil Quality – Determination of Organic and Total Carbon after Dry Combustion (Elementary Analysis)*. International Organization for Standardization, Geneva, Switzerland. ISO 10694:1995(E).

Janik, L.J., Merry, R.H., Skjemstad, J.O., 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture* 38, 681–696.

Khan, S.A., Mulvaney, R.L., Hoef, R.G., 2001. A simple test for detecting sites that are nonresponsive to nitrogen fertilization. *Soil Science Society America Journal* 65, 1751–1760.

Mimmo, T., Reeves, J.B., McCarty, G.W., Galletti, G., 2002. Determination of biological measures by mid-infrared diffuse reflectance spectroscopy in soils within a landscape. *Soil Science* 167, 281–287.

Moros, J., de Vallejuelo, S.F.O., Gredilla, A., de Diego, A., Madariaga, J.M., Garrigues, S., de la Guardia, M., 2009. Use of reflectance infrared spectroscopy for monitoring the metal content of the estuaries of the Nerbio-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country). *Environmental Science & Technology* 43, 9314–9320.

Mulvaney, R.L., Khan, S.A., Hoef, R.G., Brown, H.M., 2001. A soil organic nitrogen fraction that reduces the need for nitrogen fertilization. *Soil Science Society of America Journal* 65, 1164–1172.

Parham, J.A., Deng, S.P., 2000. Detection, quantification and characterization of beta-glucosaminidase activity in soil. *Soil Biology & Biochemistry* 32, 1183–1190.

Reeves, D.W., 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil & Tillage Research* 43, 131–167.

Reeves, J.B., McCarty, G.W., Meisinger, J.J., 2000. Near infrared reflectance spectroscopy for the determination of biological activity in agricultural soils. *Journal of Near Infrared Spectroscopy* 8, 161–170.

Rinnan, R., Rinnan, Å., 2007. Application of near infrared reflectance (NIR) and fluorescence spectroscopy to analysis of microbiological and chemical properties of arctic soil. *Soil Biology & Biochemistry* 39, 1664–1673.

Shukla, M.K., Lal, R., Ebinger, M., 2006. Determining soil quality indicators by factor analysis. *Soil & Tillage Research* 87, 194–204.

Sinsabaugh, R.L., Moorhead, D.L., 1995. Resource allocation to extracellular enzyme production: a model for nitrogen and phosphorus control of litter decomposition. *Soil Biology & Biochemistry* 26, 1305–1311.

Stark, E., 1988. Calibration methods for NIRS analysis. In: Creaser, C.S., Davies, A.M.C. (Eds.), *Analytical Applications of Spectroscopy*. Royal Society of Chemistry, England, pp. 21–34.

Storer, D.A., 1984. A simple high sample volume ashing procedure for determining soil organic matter. *Communication in Soil Science & Plant Analysis* 15, 759–772.

Terhoeven-Urselmans, T., Michel, K., Helfrich, M., Flessa, H., Ludwig, B., 2006. Near-infrared spectroscopy can predict the composition of organic matter in soil and litter. *Journal of Plant Nutrition & Soil Science* 169, 168–174.

Windham, W.R., Mertens, D.R., Barton II, F.E., 1989. Protocol for NIRS calibration: sample selection and equation development and validation. In: Marten, G.C., Shenk, J.S., II, F.E. (Eds.), *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality*. USDA, Springfield, VA, pp. 96–103.

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., Mataix-Beneyto, J., 2008. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biology & Biochemistry* 40, 1923–1930.