Rapid estimation of compost enzymatic activity by spectral analysis method combined with machine learning

CrossMark

Somsubhra Chakraborty [a,*], Bhabani S. Das [b], Md. Nasim Ali [a], Bin Li [c], M.C. Sarathjith [b], K. Majumdar [d], D.P. Ray [e]

[a] IRDM Faculty Centre, Ramakrishna Mission Vivekananda University, Kolkata 700103, India
[b] Department of Agricultural and Food Engineering, IIT, Kharagpur 721302, India
[c] Department of Experimental Statistics, Louisiana State University, 61 Agriculture Administration Building, Baton Rouge, LA 70803, USA
[d] Soil Testing Laboratory, Kalimpong 734301, India
[e] National Institute of Research on Jute and Allied Fibre Technology, Kolkata 700040, India

ARTICLE INFO

ABSTRACT

The aim of this study was to investigate the feasibility of using visible near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) as an easy, inexpensive, and rapid method to predict compost enzymatic activity, which traditionally measured by fluorescein diacetate hydrolysis (FDA-HR) assay. Compost samples representative of five different compost facilities were scanned by DRS, and the raw reflectance spectra were preprocessed using seven spectral transformations for predicting compost FDA-HR with six multivariate algorithms. Although principal component analysis for all spectral pretreatments satisfactorily identified the clusters by compost types, it could not separate different FDA contents. Furthermore, the artificial neural network multilayer perceptron (residual prediction deviation = 3.2, validation $r^2 = 0.91$ and RMSE = 13.38 $\mu g\,g^{-1}\,h^{-1}$) outperformed other multivariate models to capture the highly non-linear relationships between compost enzymatic activity and VisNIR reflectance spectra after Savitzky–Golay first derivative pretreatment. This work demonstrates the efficiency of VisNIR DRS for predicting compost enzymatic as well as microbial activity.

1. Introduction

While resource recovery and subsequent composting play a key role in agriculture, horticulture, soil science, and waste management; constant monitoring of compost quality is time consuming and laborious. Microbial enzymatic activity which is an indicator of overall microbial activity of compost, is currently estimated by the non-specific fluorescein diacetate hydrolysis (FDA-HR) assay (Ntougias et al., 2006). Furthermore, FDA-HR assay is also used to predict the suppressive capacity of compost on plant pathogens (Inbar et al., 1991). Nonetheless, the estimation procedure is expensive, time-consuming, and strictly lab based. The visible, near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS), an emerging tool in precision agriculture, may be a promising alternative to estimate compost microbial activity in a rapid and cost-effective manner.

The DRS approach consists of measuring reflectance values for a given sample as a function of wavelength (λ) over 350–2500 nm

region of the electromagnetic spectra. These reflectance values (also, known as spectral signatures) mainly arise from atomic electronic transitions and vibrational stretching and bending of structural groups of molecular atoms. While the fundamental vibrations of most organic molecules occur in mid-infrared (MIR) region, NIR spectra are dominated by weak overtones and combinations of fundamental vibrational bands. The active bonds in organic matter in the VisNIR region (350–2500 nm) are the O—H, C—N, N—H, and C=O groups (Malley et al., 2002). There is a growing body of literature on the use of near infrared spectroscopy (NIRS) for empirical calibrations, to simultaneously predict compost physical, chemical, and biological properties (Michel et al., 2006; Chakraborty et al., 2013).

Machine learning is a data based procedure allowing computers to 'learn' and 'recognize' the patterns of the empirical data (Mitchell, 1997). Several prediction methods have been developed in statistics and machine learning. Specifically, artificial neural network (ANN) based models, tree-learning techniques, rule-learning algorithms, and the traditional modeling approach such as linear regression, are the most widely used (Wu et al., 2013). Chemometric data reduction methods such as partial least squares regression (PLS), principal component regression (PCR), stepwise

multiple linear regression etc. have been used to calibrate VisNIR spectra with several compost properties. The reported predictive accuracy of DRS-compost studies suggested that this hyperspectral technique by means of its rapidity, non-destructive analysis with minimal sample pretreatment might replace standard laboratory methods for some compost applications (McWhirt et al., 2012; Chakraborty et al., 2013). Additionally, VisNIR DRS is field portable and might remove the constraints of quantifying compost FDA-HR in the lab.

Since there is a limited number of studies on the use of NIRS for compost FDA-HR assay, more investigations on the applicability of VisNIR DRS for compost FDA-HR assay involving a sample set having diversified compost samples are warranted. The objectives of this study are (i) to compare 35 different combinations of seven spectral pretreatments and five state-of-the-art machine learning algorithms to unravel the relationship between FDA-HR assay and compost VisNIR spectra, and (ii) to determine whether ANN model could assist in improving the DRS prediction accuracy. Since conventional PLS-based sensors are complicated for field use, the overarching goal of this study was to identify some other options to conventional VisNIR-PLS models, which would aid in designing a practical sensor configuration for a field person.

2. Materials and methods

2.1. Sample collection

This study included one hundred compost samples from five full scale composting facilities (referred to as Dairy, Vermi, PDOM, Dairybio, and Hostel) of Ramakrishna Mission Ashrama, Narendrapur, India (22°26′21″N, 88°23′45″E). To offset some spatial heterogeneity that may occur within a single pile, each sample represented a mixture of three subsamples collected across the same pile. For each subsample, a hole was made with a hand shovel and samples near the top, middle, and bottom were collected and mixed in a polyethylene bucket. Subsequently, samples were placed in low-density polyethylene zip-lock bags and transported to the laboratory inside a cooler box ($-4 °C$). Upon arrival at the lab, the samples were sieved (10 mm), homogenized, air-dried at 45 °C, ball milled, and stored in plastic bags at room temperature for VisNIR scanning and FDA-HR assay. A "fresh" aliquot of each sample was kept in 4 °C and subsequently used for compost physicochemical analyses (EC, water holding capacity, pH, moisture content, organic carbon, and volatile solids) within one week of sample collection. Note that, air-dried samples were preferred for FDA-HR assay over "fresh" aliquot since Mondini et al. (2004) demonstrated that the use of air-dried samples improves the consistency and the applicability of the enzymatic methods for the characterization of the composting process.

2.2. Compost physicochemical characterization

In this study, AR grade (Sigma) chemicals were used without further purification. All solutions were prepared with MilliQ^{TM} (18.2 M Ω) water and sterilized by filtration (0.44 μm pore) or by autoclave at 120 °C. Compost water holding capacity (WHC) was estimated following modified ASTM method D 2980-71 (American Society of Testing Materials, 1971). Electrical conductivity and pH were determined following method 04.11-A 1:5 Slurry (USDA-USCC, 2002) using a Mettler-Toledo pH meter (Mettler-Toledo Inc., Columbus, OH, USA) and Systronics 306 EC meter (Systronics India Ltd., Ahmedabad, India), respectively. Moisture content was estimated via method 03.09-A (USDA-USCC, 2002) while total volatile solid was estimated following Wu and Ma (2001). The organic matter (%) of compost samples was analyzed following TMECC

method 0.50.7-A loss on ignition (LOI) (USDA-USCC, 2002). All characterizations were done in four replicates. Details of WHC and volatile solid estimation are elaborated in Supplementary Material (SM).

2.3. Compost fluorescein diacetate hydrolysis assay

The FDA-HR assay is considered non-specific since it is sensitive to the activity of several enzyme classes including lipases, esterases, and proteases. Activity of these enzymes causes the hydrolytic cleavage of FDA (colorless) into fluorescein (fluorescent yellow-green). In present study, the FDA-HR assay was carried out following the method given by Schnurer and Rosswall (1982). Briefly, 5 g of well-mixed, air-dried compost was mixed with 50 mL potassium phosphate buffer (pH 6.7; 8.1 g KH_2PO_4, and 1.3 g $K_2HPO_4L^{-1}$) and 0.5 mL FDA (2 mg FDA mL acetone^{-1}). The mixture was incubated at 25 °C under continuous horizontal shaking at 120 rpm. The reaction was stopped after 30 min by adding 50 mL acetone and the extracts were filtered through Whatman-42 cellulose filter (GE healthcare, Little Chalfont, UK). The extinction of the extracts was determined against a blank at 492 nm using a Fisher Scientific Evolution 60S UV–Visible Spectrophotometer (Thermo Scientific Barnstead, Dubuque, IA). Moreover, seven standards in the range of 1–6 μg fluorescein mL^{-1} were prepared. Four analytical replicates and triplicate controls were used. These FDA-HR ($μg\ g^{-1}\ h^{-1}$) values were used as dependent variables in the subsequent modeling studies.

2.4. Spectral scanning and pretreatments

In the laboratory, the 100 compost samples were scanned using a field portable ASD FieldSpec® VisNIR spectroradiometer (Analytical Spectral Devices, CO, USA) with a spectral range of 350–2500 nm. The spectroradiometer had a 2-nm sampling interval and a spectral resolution of 3- and 10-nm wavelengths from 350 to 1000 nm and 1000 to 2500 nm, respectively. Samples were allowed to assume room temperature, evenly distributed in a opaque polypropylene sample holder and scanned from top with an ASD contact probe connected to the FieldSpec® with a fiber optic cable, having a 2-cm-diameter circular viewing area and built-in halogen light source (Analytical Spectral Devices, CO, USA). Full contact with the sample was ensured to avoid outside interference. Each sample was scanned four times with a 90° rotation between scans to obtain an average spectral curve. Each individual scan was an average of 10 internal scans over a time of 1.5 s. The detector was white referenced (after each sample) using a white spectralon panel with 99% reflectance, ensuring that fluctuating downwelling irradiance could not saturate the detector. Raw reflectance spectra were processed via a statistical analysis software package, R version 2.11.0 (R Development Core Team, 2008) using custom "R" routines (Chakraborty et al., 2013) for direct extraction of smoothed reflectance at 10 nm intervals.

This study used seven spectral pretreatment methods to prepare the compost smoothed spectra for analysis, and six multivariate algorithms to develop the predictive models. Spectral pretreatments helped in reducing the influence of the side information contained in the spectra. The pretreatment transformations applied were- gap segment derivative (segment size = 7 and gap size = 7) (GSD), Norris gap derivative across a seven-band window (NGD), Savitzky–Golay first derivative using a first-order polynomial across a ten band window (SG), standard normal variate transformation (SNV), normalization by range (NRA), log (1/reflectance) (ABS), and multiplicative scatter correction (MSC). All pretreatment transformations were implemented in the Unscrambler X 10.3 software (CAMO Software Inc., Woodbridge, NJ). All

seven spectra were included as candidate explanatory variables for FDA-HR in subsequent modeling analyses.

2.5. Machine learning

Initially, five multivariate methods tested were partial least squares (PLS), principal component regression (PCR), random forest (RF), support vector regression (SVR), and penalized spline regression (PSR) (Breiman, 2001; Guyon et al., 2002). Two samples were not included in the multivariate models due to missing values and removed *a priori*. Two model evaluation approaches were taken: (i) Initially, the whole dataset was randomly divided only once into a ∼70% training set ($n = 70$) for calibration and a ∼30% independent validation set ($n = 28$) to prevent overfitting and (ii) this step was repeated 50 times for more honest evaluation of the prediction performance. Each time the five methods were applied on the training set and validated by test samples. For the latter approach, the root mean squared error (RMSE) of test set was extracted which resulted in 1750 [7(preprocessing) × 5(model) × 50(iteration)] RMSE on test sets. Additionally, the average and standard deviation (SD) of RMSE from 50 iterations were also extracted to measure the stability of predictive models. For PSR, the cubic B-spline was used via R version 2.14.1 (R Development Core Team, 2008) as the base function with 100 equally spaced knots. The order of the penalty was set to the default value of three. The optimal value for the penalty-tuning parameter was selected by minimizing the leave-one-out-cross-validation (LOOCV) error on the training set. Moreover, the 'randomForest' package was used in R to build the random forest model. The number of trees in random forest was set to the default value of 500. The coefficient of determination (r^2), RMSE, residual prediction deviation (RPD), and bias were used as rubrics for judging model generalizing capability.

Subsequently, the best performing spectral pretreatment that was selected by aforementioned models was combined with ANN to test whether it can further improve DRS predictability (Ripley, 1996) (For more details on ANN, see SM). The principal component analysis (PCA) approach was applied for dimensionality reduction and qualitative VisNIR discrimination of the inherently different compost samples for all spectral pretreatments. Fisher's linear discriminant analysis (LDA) was then applied on the selected leading PCs, assuming equal prior probability for each group. To assess classification results, kappa coefficients were computed (Thompson and Walter, 1988). PCA was performed using R version 2.11.0 sing function 'prcomp'.

3. Results and discussion

3.1. Compost physicochemical and biological properties

Compost samples varied in their physical, chemical, and biological properties since composition of feedstock and compost maturity varied among facilities (Table 1). EC was positively correlated with compost pH ($\rho = 0.88$) and WHC ($\rho = 0.94$) while compost age was negatively correlated with moisture content ($\rho = -0.9$) (Table SM-1). Notably, in compliance with the earlier findings (Michel et al., 2006), the absence of significant linear correlation between compost FDA-HR and other tested compost properties (compost organic matter, volatile solids, WHC, pH, EC, moisture content, and age) justified the exclusion of latter group as auxiliary predictors in the subsequent VisNIR FDA-HR models. Nevertheless, one must use caution while comparing and interpreting the compost properties due to unequal replicates for each of the composts within the experimental design. The authors acknowledge that the variability of number of compost replicates somewhat constrained the global applicability of the dataset. However, this research was intended to investigate the viability of different spectral analysis techniques and ascertain which techniques show the most promise for future investigations. Note that, the variable enzymatic activity could not be explained by the formation of stable humo-enzymatic complex which protects the extracellular enzymes (Mondini et al., 2004), possibly due to lack of correlation between FDA and compost organic matter. The testing of compost humification degree (CEC/Total organic carbon) was, however, beyond the scope of this study.

3.2. Principal component analysis and linear discriminant analysis

Despite the high dimensionality of the spectral data (215 spectral channels from 350 to 2500 nm at 10-nm intervals), first three PCs tended to explain almost 100% of the spectral variance for all seven spectral pretreatments. Separate pairwise PC score plots (PC1 vs. PC2) indicating five different compost types were used to discriminate compost reflectance spectra and identify spectral similarities within a single compost type for each spectral pretreatment (Fig. 1). Any significant variation due to changes in spectral treatment was not discernible and all plots appeared almost identical to each other. Moreover, all plots exhibited clear "compost type-clustered" structure. This could be explained by the heterogeneity which resulted partly from compositional variability of raw materials and partly from variable composting time. Notably, a subtle chronological distribution of samples was observed on PC2. The separation of the VisNIR profiles of 25-days-old PDOM, 45-days-old vermi, and 150–180-days-old Dairybio, Dairy, and Hostel composts on this axis, may be explained by difference in chemical composition with increasing compost age. In particular, the alteration in the number of certain functional groups (OH, aliphatic C—H etc.) and varying contents of water, starch, cellulose, pectin, and other spectrally active components perhaps significantly influenced VisNIR spectra (Malley et al., 2005). Additionally, the close clusters formed by two relatively older composts (Dairy and Hostel) uphold the idea of decreasing compost heterogeneity with maturation (Hsu and Lo, 1999). The first leading component (PC1) distinguished the PDOM samples from vermi and dairy

Table 1
Average physicochemical and biological properties of compost samples.

Compost[a]	Compost method	n	Age (days)	pH	EC (Sm^{-1})	WHC[b] (gk g^{-1})	Moisture content (gk g^{-1})	Volatile solids (gk g^{-1})	Organic matter (%)	FDA-HR (µg g^{-1} h^{-1})
Dairy	Passive windrow	61	180	5.8	0.34	1043	547	659	22.3	603.3
Vermi	Vessel	11	45	6.3	0.42	1254	645	543	20.24	666.3
PDOM	Vessel	16	25	6.8	0.40	1123	754	608	19.85	678.2
Dairybio	Channel	6	150	5.9	0.37	1065	532	641	25.40	705.4
Hostel	Vessel	6	180	7.2	0.46	1267	568	754	19.30	691.2

[a] Dairy, fully decomposed mixture of dairy yard waste and dairy manure for 5–6 months; Vermi, vermicompost with *Eisenia foetida*; PDOM, partially decomposed organic matter containing mixture of chopped plant litter and dairy manure slurry; Dairybio, mixture of biogas plant slurry and dairy manure; Hostel, compost from a mixture of kitchen waste, dairy manure, and virgin soil.
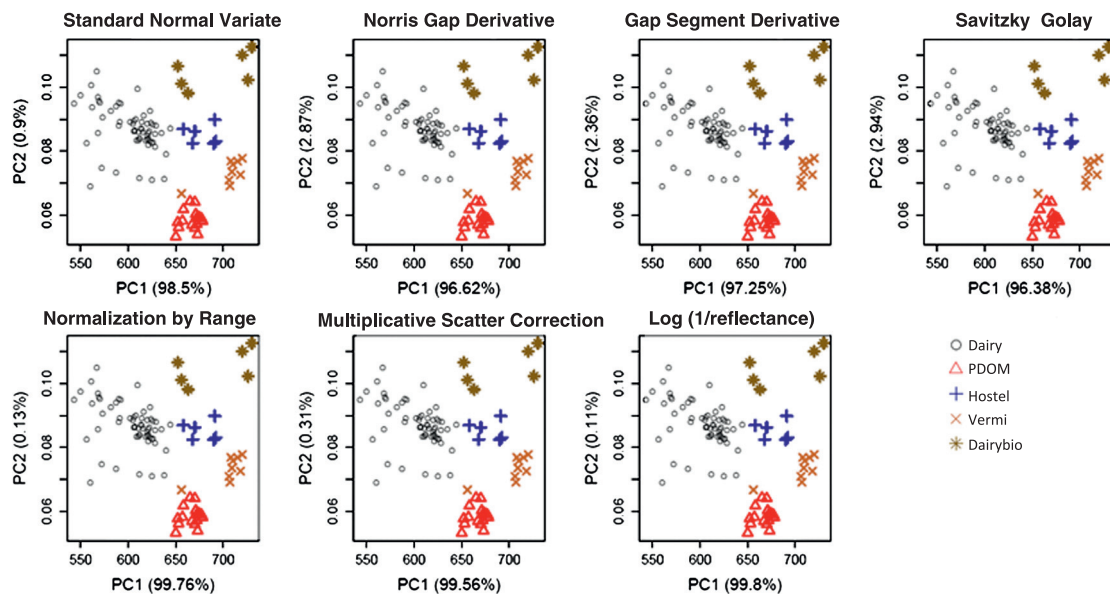
[b] WHC, water holding capacity.

Fig. 1. Pairwise principal component (PC) plot for PC1 vs. PC2 of seven spectral pretreatments.

samples. Conversely, some dairybio samples were mixed with dairy samples due to their compositional resemblance (Table 1). Both PDOM and vermi samples tended to have low scores along PC 2 and spectral diversity was relatively low in these two compost samples. No obvious outlier among the samples was seen. Results of LDA classification closely followed results of visual PC plot inspections (Table 2). Indeed, for compost type classification, LDA was 100% accurate in classifying compost types for all spectral pretreatments except SNV and MSC. For MSC and SNV, LDA correctly classified all but three and five samples, respectively by compost type. Therefore, it can be confirmed that this hyperspectral sensing offers both the power of PCA as a means of discriminating VisNIR

reflectance spectra of heterogeneous compost types and exhibiting a clustering tendency for compositional similarity with great sensitivity. Nevertheless, no clear trend for FDA was obvious from the plot, perhaps due to lack of correlation between compost age and FDA (Table SM-1). Thus, the postulations of Cayuela et al., 2008, who reported that FDA can unambiguously discriminate compost samples with different ages, may not be generalized.

3.3. Qualitative characterization of compost reflectance spectra

Given that, PC loadings can be interpreted as correlations between the variables (wavelength) and the components of interest

Table 2
Confusion matrix showing classification of compost samples using the Fisher's Linear Discriminant Analysis (LDA). The first two principal components (PC) scores of the seven pretreated spectra were used as the explanatory variable. The weighted kappa coefficients are also given. The bold values are correctly classified samples.

	GSD ($\kappa = 1$)					NGD ($\kappa = 1$)					ABS ($\kappa = 1$)				
	Dairy	PDOM	Hostel	Vermi	Dairybio	Dairy	PDOM	Hostel	Vermi	Dairybio	Dairy	PDOM	Hostel	Vermi	Dairybio
Predicted dairy	**61**	0	0	0	0	**61**	0	0	0	0	**61**	0	0	0	0
Predicted PDOM	0	**16**	0	0	0	0	**16**	0	0	0	0	**16**	0	0	0
Predicted hostel	0	0	**6**	0	0	0	0	**6**	0	0	0	0	**6**	0	0
Predicted vermi	0	0	0	**9**	0	0	0	0	**9**	0	0	0	0	**9**	0
Predicted dairybio	0	0	0	0	**6**	0	0	0	0	**6**	0	0	0	0	**6**
	Overall accuracy				100%	Overall accuracy				100%	Overall accuracy				100%
	NRA ($\kappa = 1$)					MSC ($\kappa = 0.94$)					SG ($\kappa = 1$)				
Predicted dairy	**61**	0	0	0	0	**61**	0	3	0	0	**61**	0	0	0	0
Predicted PDOM	0	**16**	0	0	0	0	**16**	0	0	0	0	**16**	0	0	0
Predicted hostel	0	0	**6**	0	0	0	0	**3**	0	0	0	0	**6**	0	0
Predicted vermi	0	0	0	**9**	0	0	0	0	**9**	0	0	0	0	**9**	0
Predicted dairybio	0	0	0	0	**6**	0	0	0	0	**6**	0	0	0	0	**6**
	Overall accuracy				100%	Overall accuracy				96.9%	Overall accuracy				100%

	SNV ($\kappa = 0.91$)				
	Dairy	PDOM	Hostel	Vermi	Dairybio
Predicted dairy	**61**	0	0	0	0
Predicted PDOM	0	**16**	0	0	0
Predicted hostel	0	0	**6**	0	1
Predicted vermi	0	0	0	**7**	2
Predicted dairybio	0	0	0	2	**3**
	Overall accuracy				94.9%

(enzymes) to identify the underlying correlation, the PC1 loadings (Fig. SM-1) qualitatively characterized SG spectra by interpreting negative and positive peaks associated with the component of interest and interfering components, respectively (Chakraborty et al., 2013). Considering that real molecules do not behave totally harmonically, minor positional changes in the spectral features from the exact anticipated position were possible. Briefly, mild positive contribution (negative peak) at 350–550 nm (visible region) indicated the VisNIR sensitivity towards compost color (Ben-Dor et al., 1997). Additionally, strong positive contributions at ~1400 nm ($2\upsilon_2 + \upsilon_3$) and ~1900 nm ($\upsilon_2 + \upsilon_3$) for OH in water molecules with minor positive contribution from near 2250 to 2500 nm for methyl ($3\upsilon_1$) were evident (Viscarra Rossel and Beherens, 2010). Strikingly, the latter region may contain 2337 nm (3rd overtone of COO— stretching of CH_3 of protein) which was previously assigned responsible for hydrolysis of FDA (Michel et al., 2006). Since the compost samples were air-dried prior to the spectral scanning, one can expect that the reflectance values, especially at 1400 and 1900 nm actually translated the adsorbed water molecules on the highly hygroscopic compost. That notwithstanding, it is likely that the water feature around 1400 nm might have overlapped with other OH groups in cellulose molecules (Elvidge, 1990), carboxylic acid ($4\upsilon_1$) (Viscarra Rossel and Beherens, 2010) or with CH_2 groups in lignin molecules (McLellan et al., 1991). Therefore, precise assignment of this feature was complicated.

3.4. Multivariate modeling

Initially, among the five multivariate algorithms (PLS, PSR, PCR, RF, and SVR), FDA-HR was estimated with the greatest accuracy by RF. For a single random split, lab-measured versus RF predicted FDA-HR calibration models showed close out-of-bag prediction r^2 (similar to LOOCV prediction on each of the training points), ranging from 0.73 to 0.8 for all spectral pretreatments except ABS and MSC (0.66 for each) (Fig. 2). In terms of calibration, NRA provided the best results, with an r^2 of 0.8. In general, RF models with all spectral pretreatments showed underestimation at lower FDA-HR values and overestimation at higher values. Model predictions for the compost FDA-HR in validation set are summarized in Table 3.

Not shown, the range of values in the validation set was encompassed by the range of the calibration set. The prediction quality was judged by generalization capability (validation r^2, validation RMSE, bias, and RPD) of the test set. Given that RPD is the ratio of standard deviation and RMSE, model predictability decreases when validation set standard deviation (SD) is comparatively larger than estimation error (RMSE). Chang et al. (2001) categorized the accuracy and stability of their spectroscopy models based on the RPD values of validation set. The RPD >2.0 were considered stable and accurate predictive models; RPD values between 1.4 and 2.0 indicated fair models that could be improved by more accurate predictive techniques; RPD values <1.4 indicated poor predictive capacity. In this study, successful prediction of FDA-HR contents with RPD ranging from 2.01 to 2.30 for RF and SVR models supported the findings of Michel et al. (2006) that NIRS is a promising tool for estimating microbial activity in compost samples. However, the prediction was more accurate with higher validation r^2 ranging from 0.79 to 0.8 and lower RMSE (16.55–17.14 µg g^{-1} h^{-1}) for RF models with all spectral pretreatments except NRA (r^2 = 0.74 and RMSE = 18.65 µg g^{-1} h^{-1}) and SNV (r^2 = 0.70 and RMSE = 20.23 µg g^{-1} h^{-1}). Norris Gap Derivative (which in turn a special case of GSD with segment size s = 1) and SG use a smoothing of the spectra prior to calculating the derivative for decreasing the detrimental effect on the signal-to-noise ratio. This effect perhaps helped in increasing subsequent model's generalization capability. Additionally, MSC corrects differences in the base line and in the trend with an added benefit of producing transformed spectra similar to the original spectra which collectively leads to easier optical interpretation. The wide range found for FDA-HR values (543–730 µg g^{-1} h^{-1}) in the tested samples perhaps contributed to the satisfactory results too, because the consistency of a NIR spectroscopic model is by and large restricted to the range of parameter values. Compost samples from the curing phase were not included in the predictive model. During curing, compost products wait outside the pit for packing and commercialization. Moreover, they experience minimal compositional and spectral changes, in contrast to the significant differences seen in composting days. Consequently, such difference can add to the VisNIR estimations. On the contrary, one reason for the underperformance of RF model using
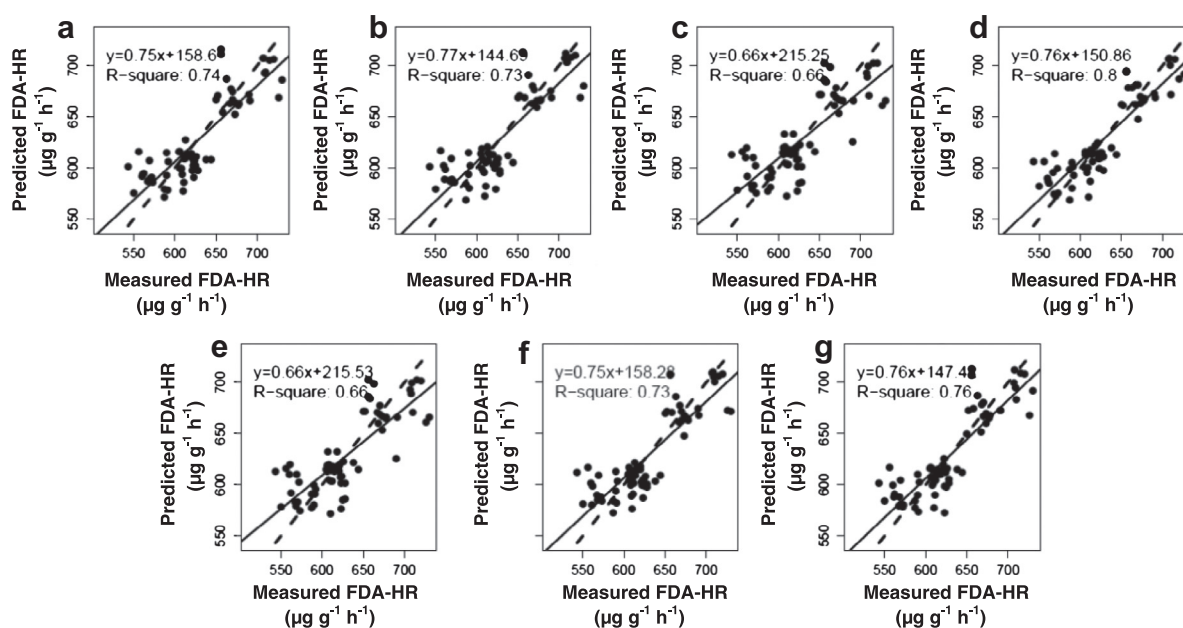

Fig. 2. Lab-measured vs. random forest regression (RF) predicted FDA-HR (µg g^{-1} h^{-1}) using training set for (a) GSD, (b) NGD, (c) ABS, (d) NRA, (e) MSC, (f) SG, and (g) SNV pretreatment (one random split). The solid line is the regression line, and the dashed line is a 1:1 line.

Table 3
Summary validation statistics of the models (one random split) obtained for compost FDA-HR by the different multivariate methods (except ANN) associated with their respective pretreatment transformations.

Model[a]	Pretreatment[b]	Validation			
		r^2	RPD[c]	RMSE ($\mu g\ g^{-1}\ h^{-1}$)	Bias ($\mu g\ g^{-1}\ h^{-1}$)
RF	GSD	0.79	2.24	17.12	−2.73
	NGD	0.79	2.25	17.08	−3.46
	ABS	0.79	2.24	17.12	−2.12
	NRA	0.74	2.01	18.65	−2.80
	MSC	0.79	2.24	17.14	−2.12
	SG	0.80	2.30	16.55	−3.60
	SNV	0.70	1.9	20.23	−4.30
PLS	GSD	0.66	1.75	21.89	−3.89
	NGD	0.68	1.82	21.10	−4.01
	ABS	070	1.85	20.76	−4.50
	NRA	0.71	1.90	20.21	−2.76
	MSC	0.72	1.93	19.90	−3.79
	SG	0.67	1.76	21.74	−3.80
	SNV	0.72	1.95	19.70	−3.32
PCR	GSD	0.67	1.78	21.47	−4.14
	NGD	0.67	1.79	21.36	−3.45
	ABS	0.70	1.85	20.76	−4.62
	NRA	0.72	1.92	20.00	−1.27
	MSC	0.72	1.93	19.81	−3.18
	SG	0.67	1.77	21.62	−4.35
	SNV	0.68	1.81	21.11	−2.43
PSR	GSD	0.69	1.83	20.88	−5.56
	NGD	0.69	1.83	20.87	−5.78
	ABS	0.70	1.86	20.55	−4.83
	NRA	0.71	1.89	20.26	−1.59
	MSC	0.72	1.92	19.95	−4.01
	SG	0.69	1.85	20.63	−4.62
	SNV	0.72	1.94	19.70	−2.80
SVR	GSD	0.80	2.27	16.85	−1.60
	NGD	0.79	2.26	16.94	−2.28
	ABS	0.76	2.09	18.35	−4.42
	NRA	0.77	2.13	17.92	−2.62
	MSC	0.77	2.11	18.11	−4.24
	SG	0.79	2.24	17.06	−2.68
	SNV	0.73	1.99	19.22	−1.45

[a] RF, random forest; PLS, partial least squares regression; PCR, principal component regression; PSR, penalized spline regression; SVR, support vector regression.
[b] GSD, gap segment derivative; NGD, Norris gap-derivative; ABS, log(1/reflectance); NRA, normalization by range; MSC, multiplicative scatter correction; SG, Savitzky–Golay first derivative using a first-order polynomial across a ten band window; SNV, standard normal variate.
[c] RPD, residual prediction deviation.

NRA and SNV data could be the use of only 70 calibration samples. Although model generalization capability for PLS, PSR, and PCR was less than those of RF and SVR models, former models can be classified as moderately successful with RPD values ranging from 1.75 to 1.95 (Chang et al., 2001).

Plots of observed vs. RF predicted FDA-HR for all spectral pretreatments are presented in Fig. 3. In general, RF predictions of FDA-HR for all pretreatments closely approximated the 1:1 line and had less average bias ($-3.01\ \mu g\ g^{-1}\ h^{-1}$) than their PLS ($-3.72\ \mu g\ g^{-1}\ h^{-1}$), PCR ($-3.34\ \mu g\ g^{-1}\ h^{-1}$), and PSR ($-4.17\ \mu g\ g^{-1}\ h^{-1}$) counterparts except for SVR model ($-2.75\ \mu g\ g^{-1}\ h^{-1}$) (Table 3). However, all model biases were negligible than corresponding MSEs and thus accounted for a very trivial part of the overall lack of fit. Hence, model inaccuracy could be primarily attributed to a lack of correlation with regression line near unity.

Fig. 4 shows the training (blue[1] dotted line) and test (red line) errors (RMSE) on different values of *mtry*, which is the size of the candidate subset for each splitting. The vertical dashed line is the

[1] For interpretation of color in Fig. 4, the reader is referred to the web version of this article.

optimal *mtry* value that minimizes the test error. Except for NRA and SNV, the prediction performance was somewhat insensitive to the *mtry* value. The test error was used as an estimate of the prediction performance for the model on a new dataset, while training error was substantially below the test error. The ABS, MSC, and SG based parsimonious models used only 1 *mtry* (model factors) and clearly outperformed NRA (7 *mtry*), NGD (100 *mtry*), GSD (150 *mtry*), and SNV (150 *mtry*) based models. This, together with RPD values (2.24–2.30), may suggest stable, effective VisNIR models with suitable spectral transformations those can distinguish compost FDA-HR for different types of compost and predict the microbial activity.

While comparing the average and standard deviation (SD) of RMSE from 50 iterations of test set, the RF model again outperformed other four multivariate models with lowest average RMSEs for most of the spectral pretreatments; and low range of SDs of RMSE (3.53–3.73 $\mu g\ g^{-1}\ h^{-1}$), indicating greater stability (Fig. SM-2). On contrary, RF (NRA) model showed least stability exhibiting SD of RMSEs of 4.57 $\mu g\ g^{-1}\ h^{-1}$, and thus not recommended.

3.5. Significant wavelengths selected by multivariate algorithms

The significant wavelengths of all four models except SVR are plotted in Fig. SM-3. Since SVR does not select variables and it uses inner product of data points, it was difficult to get the coefficient for each channel. For PLS and PCR, the absolute value of the coefficients were averaged for each channel (over seven spectral treatments), since the size of the coefficients are often regarded as the significance of that variable (wavelength here). For PSR, the scaled absolute values of the coefficient were averaged over seven datasets. Scaling was done by dividing the coefficient by its standard deviation. RF automatically produced the variable importance for each channel which was averaged over seven spectral datasets. Noticeably, visual inspection suggests that the predictive information contained in the compost spectral curves is actually concentrated in a subset of important wavelengths. Additionally, overlapping region of significant wavelengths revealed that almost all algorithms captured the crucial absorbance signatures [1364 nm (C—H combination band of CH_3 groups), 1774 nm (1st overtone of C—H stretching of cellulose), 1896 nm (2nd overtone of C=O stretching of CH_2O groups), 2193 nm (Amide), and 2337 nm (3rd overtone of COO— stretching of CH_3 of protein)]. These signatures were previously assigned responsible for either hydrolysis of FDA (Michel et al., 2006) or for other protein containing compounds (Curran et al., 1992). Further close investigation of the 1360–2340 nm spectral subset region (after truncating from the whole 350–2500 nm spectrum) for two end members (SG and SNV, based on validation performance) of RF models (Fig. SM-4) revealed that RF scores at the abovementioned spectral signatures (~1364 nm, ~1770 nm, ~1896 nm, ~2200 nm, and ~2340 nm) were comparatively higher in SG than SNV, possibly resulting in better generalization capacity of SG based model. Cast in this light, it was evident that complementary to other report (Kooistra et al., 2003), pretreatments of spectral data boosted the accuracy of regression models.

3.6. Artificial neural network for improving prediction accuracy

The ANN models using the SG spectra considerably improved the prediction accuracy and produced the smallest validation RMSE for FDA-HR (Table 4). Interestingly, while using 7 neurons in each hidden layer, MLP 214-7-1 network did not showed overfitting and produced the best prediction (RPD = 3.20, r^2 = 0.91, RMSE = 13.38 $\mu g\ g^{-1}\ h^{-1}$) of all the models tested. The validation r^2 for MLP 214-9-1 and MLP 214-8-1 were similar with values of 0.89 and 0.88, respectively. Notably, MLP 214-4-1 showed worst
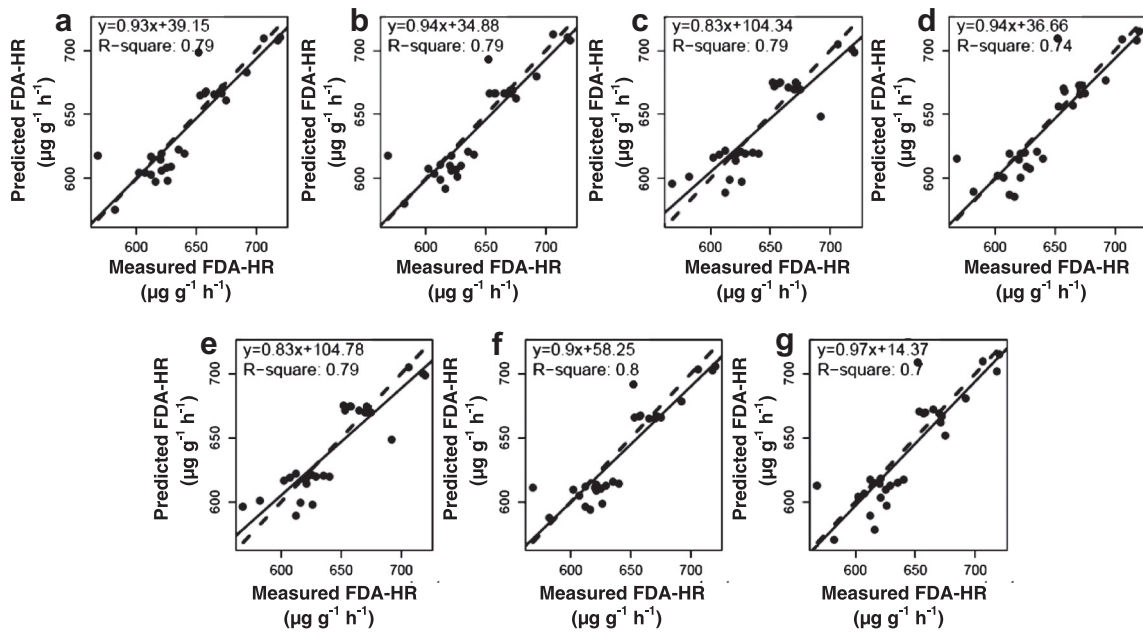
Fig. 3. Lab-measured vs. RF predicted FDA-HR ($\mu g\ g^{-1}\ h^{-1}$) using validation set for (a) GSD, (b) NGD, (c) ABS, (d) NRA, (e) MSC, (f) SG, and (g) SNV pretreatment (one random split). The solid line is the regression line, and the dashed line is a 1:1 line.
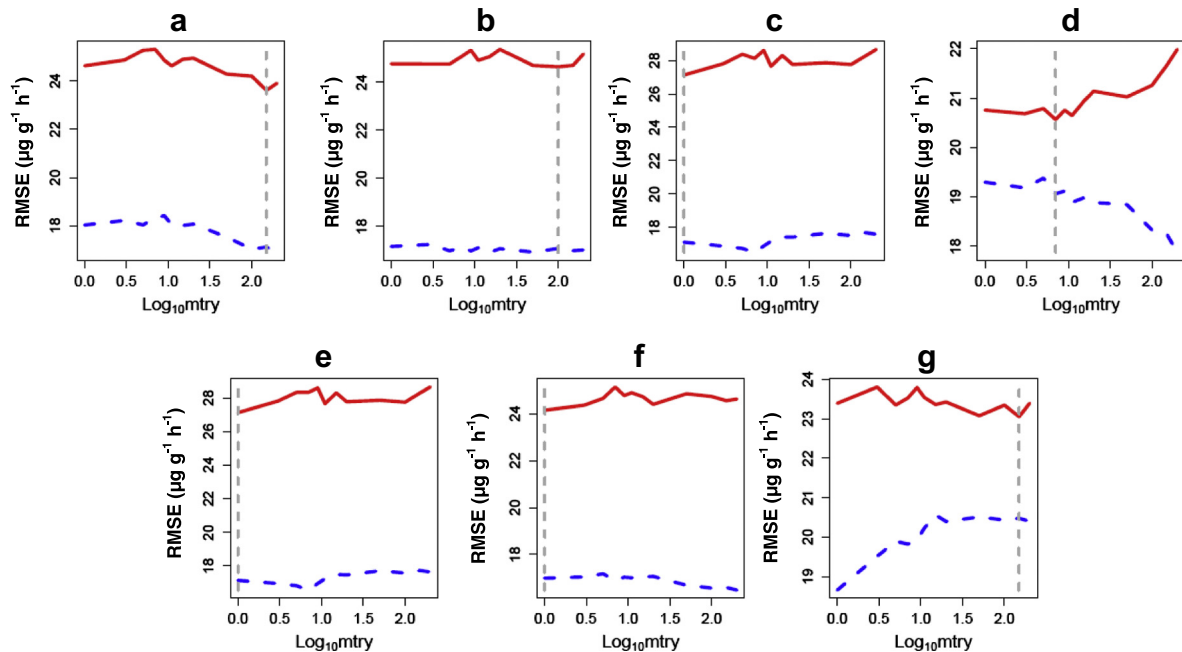

Fig. 4. Changes in RMSE ($\mu g\ g^{-1}\ h^{-1}$) on different values of the tuning parameter ($Log_{10}mtry$) for the RF models for (a) GSD, (b) NGD, (c) ABS, (d) NRA, (e) MSC, (f) SG, and (g) SNV spectral pretreatment.

generalization capacity among MLP models with RPD of 2.10. Considering training performance, the MLP 214-7-1 and MLP 214-9-1 networks produced similar training r^2 (0.92). However, it should be noted that just because the training r^2 were comparatively low in the rest three networks did not essentially implied badly trained networks. In fact it could be an indication of conservative networks, which did their best to avoid fitting (modeling) the noise, which is crucial. Additionally, the scatterplots of predicted vs. measured FDA-HR values for best five networks showed that most points did not lie exactly on the 1:1 line (Fig. 5). Perhaps the networks have recognized some noise on the target values and

have avoided modeling them as true signals, which is a preferred outcome.

3.7. VisNIR DRS and FDA-HR assay

Overall, the example presented here demonstrated that VisNIR DRS could be utilized as a rapid and useful tool for estimating FDA-HR. Spectral pretreatment and subsequent state – of-the-art ANN approach helped to gain insights into the correlation between spectral signatures and compost FDA-HA variability. Practically, it is not easy to visually extract and efficiently scrutinize spectral

Table 4
Summary statistics of the ANN models obtained for compost FDA-HR with SG spectra.

Net name	Training r^2	Validation r^2	Validation RMSE ($\mu g\,g^{-1}\,h^{-1}$)	RPD[b]	Training algorithm	Error function	Hidden activation	Output activation
MLP 214-7-1[a]	0.92	0.91	13.38	3.20	BFGS 18[c]	SOS[d]	Identity[e]	Logistic[g]
MLP 214-9-1	0.92	0.89	15.02	2.90	BFGS 15	SOS	Identity	Identity
MLP 214-8-1	0.91	0.88	15.13	2.88	BFGS 13	SOS	Tanh[f]	Tanh
MLP 214-5-1	0.84	0.83	17.52	2.50	BFGS 15	SOS	Identity	Exponential[h]
MLP 214-4-1	0.83	0.81	18.19	2.10	BFGS 15	SOS	Identity	Logistic

[a] MLP, multilayer perceptron (number of inputs, number of neurons in the hidden layer, and the number of outputs). For example, the model name "MLP 1-2-1" indicates a multilayer perceptron network with 1 input, 2 neurons in each layer, and 1 output.
[b] RPD, residual prediction deviation.
[c] BFGS, Broyden–Fletcher–Goldfarb–Shanno.
[d] SOS, sum of squares.
[e] Identity, identity function. With this function, the activation level is passed on directly as the output of the neurons.
[f] Tanh, hyperbolic tangent function which is a symmetric S-shaped (sigmoid) function, whose output lies in the range (−1, +1).
[g] Logistic, logistic sigmoid function which is an S-shaped (sigmoid) curve, with output in the range (0, 1).
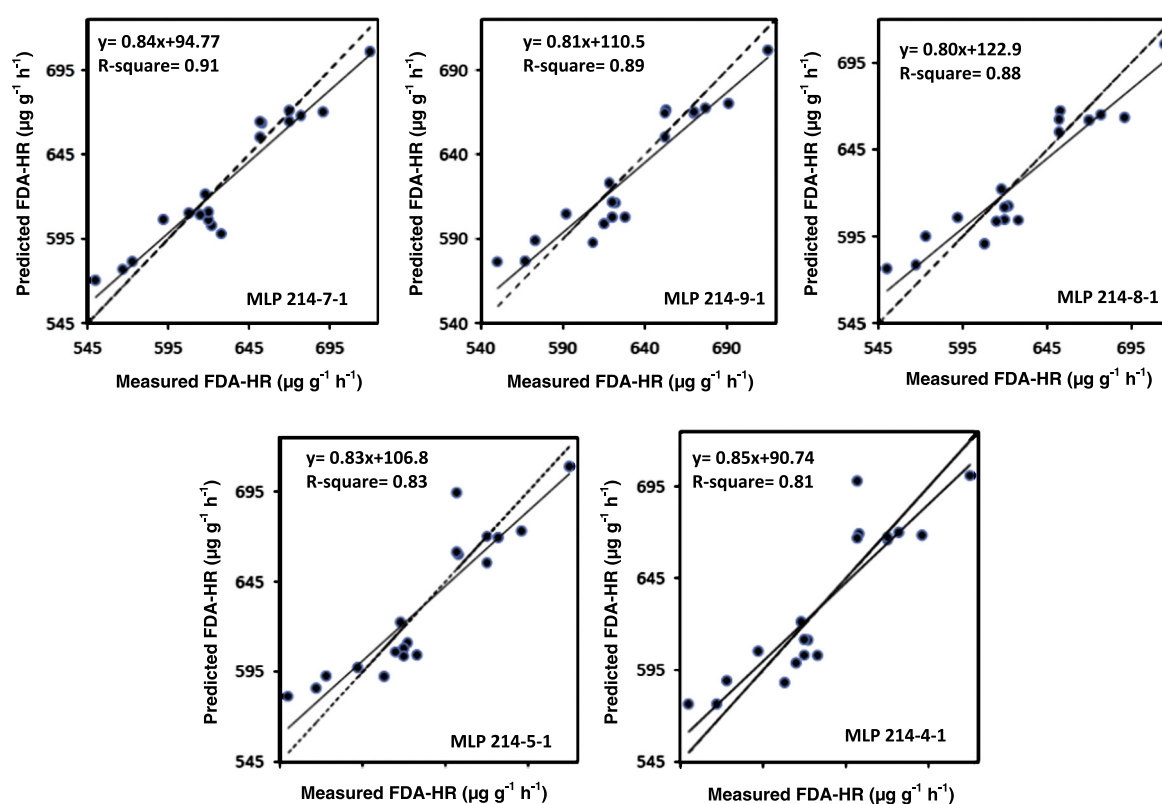[h] Exp, negative exponential activation function.

Fig. 5. Lab-measured vs. ANN predicted FDA-HR ($\mu g\,g^{-1}\,h^{-1}$) using validation set with SG pretreatment for five best MLP networks. The solid line is the regression line, and the dashed line is a 1:1 line.

variations directly from the raw reflectance spectra (Ben-Dor et al., 1997). Moreover, for biological materials such as compost, the spectral scattering properties are highly complex, very soft, or model-free. Therefore, spectral pretreatment is obligatory to remove scatter from the pure VisNIR reflectance spectra as well as to enhance and assign maximum spectral information (strong and weak). Since the popularity of neural network methodology is rapidly growing in environmental science, the well-trained ANN networks could serve wide environmental users. The better results were obtained from SG since it is one of the filters which can smooth out the reflectance signal without much destroying its original properties. Results also demonstrated that the selected FDA-HR related spectral signatures could be selectively used to improve model parsimony and developing on the move VisNIR sensor system for in-situ compost characterization. Although ANN model remained superior to other five models, it was by no means

exhaustive and perhaps requires large dataset before drawing a stronger conclusion. One advantage of this study was to use different composts to predict enzymatic activity, contrary to enzymatic measurements of a same compost windrow. In precision agriculture, the development of field-based electro-optical sensors based on conventional PLS model is always challenging since it involves hundreds of filter-detector pairs with variable central wavelengths and a constant bandwidth of 10 nm and an extra circuit block to combine these detectors' output signals into 18 synthetic signals (Chakraborty et al., 2013). Given the intricacy of the analytical procedures that ANN (SG) can entail, it is recommended that SG which is based on local least-squares polynomial approximation, be integrated into commercial machine learning related softwares. Taking all the above collectively, this study clearly recognized the potential of the VisNIR-ANN model with SG pretreatment as a viable alternative to the VisNIR-PLS for rapid and low-cost estimation of

compost FDA-HR as an addition to the certified methods for compost enzymatic activity and disease suppressive capacity analysis.

4. Conclusion

In this study, 100 compost samples from five full scale composting facilities were collected and scanned by DRS, and the raw reflectance spectra were preprocessed using seven spectral transformations for predicting compost FDA-HR with six multivariate algorithms. Although principal component analysis for all spectral pretreatments satisfactorily identified the clusters by compost types, it could not separate different FDA contents. Visual interpretations from PC plots were quantitatively confirmed by LDA which was 100% accurate in classifying compost types for all spectral pretreatments except SNV (94.9%) and MSC (96.9%). Random Forest models outperformed PLS, PSR, PCR, and SVR models with higher validation r^2 ranging from 0.79-0.8 and lower RMSE (16.55–17.14 $\mu g\,g^{-1}\,h^{-1}$) with all spectral pretreatments except NRA ($r^2 = 0.74$ and RMSE = 18.65 $\mu g\,g^{-1}\,h^{-1}$) and SNV ($r^2 = 0.70$ and RMSE = 20.23 $\mu g\,g^{-1}\,h^{-1}$). Furthermore, the ANN multilayer perceptron after Savitzky–Golay pretreatment was able to increase the DRS prediction accuracy (RPD = 3.2, validation $r^2 = 0.91$ and RMSE = 13.38 $\mu g\,g^{-1}\,h^{-1}$) by capturing the highly non-linear relationships between compost enzymatic activity and VisNIR reflectance spectra. This suggest that the combination of VisNIR DRS and machine learning tools could help reduce biochemical analysis frequency for compost microbial activity analysis and support decision making for quality control of end products, substantially saving labor and cost for this purpose. Obviously, more fundamental investigations as to how compost enzymes influences optical properties are warranted and improvement could be made by increasing sample numbers or with an advanced spectral treatment, such as wavelet decomposition.

Acknowledgements

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.wasman.2013.12.010.

References

ASTM D 2980-71, 1971. American Society of Testing Materials, Philadelphia, PA, 19103.

Ben-Dor, E., Inbar, Y., Chen, Y., 1997. The reflectance spectra of organic matter in the visible near infrared and short wave infrared region (400–2500 nm) during a control decomposition process. Remote Sens. Environ. 61, 1–15.

Breiman, L., 2001. Random forests. Machine Learning 45, 5–32.

Cayuela, M.L., Mondini, C., Sanchez-Monedero, M.A., Roig, A., 2008. Chemical properties and hydrolytic enzyme activities for the characterisation of two-phase olive mill wastes composting. Bioresour. Technol. 99, 4255–4262.

Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., Peterson, D.L., 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. Remote Sens. Environ. 39, 153–166.

Chakraborty, S., Weindorf, D.C., Ali, N., Li, B., Ge, Y., Darilek, J.L., 2013. Spectral data mining for rapid measurement of organic matter in unsieved moist compost. Appl. Opt. 52, B82–B92.

Chang, C., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near infrared reflectance spectroscopy: principal components regression analysis of soil properties. Soil Sci. Soc. Am. J. 65, 480–490.

Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. Int. J. Remote Sens. 11, 1775–1795.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using SVM. Machine Learning 46, 389–422.

Hsu, J.H., Lo, S.L., 1999. Chemical and spectroscopic analysis of organic matter transformations during composting of pig manure. Environ. Pollut. 104, 189–196.

Inbar, Y., Boehm, M.J., Hoitink, H.A.J., 1991. Hydrolysis of fluorescein diacetate in sphagnum peat container media for predicting suppressiveness to damping off caused by Pythium ultimum. Soil Biol. Biochem. 23, 479–483.

Kooistra, L., Wanders, J., Epema, G.F., Leuven, R.S.E.W., Wehrens, R., Buydens, L.M.C., 2003. The potential of field spectroscopy for the assessment of sediment properties in river floodplains. Anal. Chim. Acta 484, 189–200.

Malley, D.F., McClure, C., Martin, P.D., Buckley, K., McCaughey, W.P., 2005. Compositional analysis of cattle manure during composting using a field-portable near-infrared spectrometer. Commun. Soil Sci. Plant Anal. 36, 455–475.

Malley, D.F., Yesmin, L., Eilers, R.G., 2002. Rapid analysis of hog manure and manure amended soils using near-infrared spectroscopy. Soil Sci. Soc. Am. J. 66, 1677–1686.

McLellan, T.M., Aber, J.D., Martin, M.E., Melillo, J.M., Nadelhoffer, K.J., 1991. Determination of nitrogen, and cellulose content of decomposing leaf material by near infrared reflectance spectroscopy. Can. J. For. Res. 21, 1684–1688.

McWhirt, A.L., Weindorf, D.C., Chakraborty, S., Li, B., 2012. Visible near infrared diffuse reflectance spectroscopy (VisNIR DRS) for rapid measurement of organic matter in compost. Waste Manage. Res. 30, 1049–1058.

Michel, K., Bruns, C., Tarhoeven-Urselmans, T., Kleikamp, B., Ludwig, B., 2006. Determination of chemical and biological properties of composts using near infrared spectroscopy. J. Near Infrared Spectrosc. 14, 251–259.

Mitchell, T., 1997. Machine Learning. McGraw Hill.

Mondini, C., Fornasier, F., Sinicco, T., 2004. Enzymatic activity as a parameter for the characterization of the composting process. Soil Biol. Biochem. 36, 1587–1594.

Ntougias, S., Ehaliotis, C., Papadopoulou, K.K., Zervakis, G., 2006. Application of respiration and FDA hydrolysis measurements for estimating microbial activity during composting processes. Biol. Fertil. Soils 42, 330–337.

R Development Core Team, 2008. R: a language and environment for statistical computing. Available online with updates at <http://www.cran.r-project.org>. R Foundation for Statistical Computing, Vienna, Austria. (Verified 29 July 2013).

Ripley, B.D., 1996. Pattern Recognition and Neural Networks. Cambridge University Press.

Schnurer, J., Rosswall, T., 1982. Fluorescein diacetate hydrolysis as a measure of total microbial activity in soil and litter. Appl. Environ. Microbiol. 43, 1256–1261.

Thompson, W.D., Walter, S.D., 1988. A reappraisal of the kappa coefficient. J. Clin. Epidemiol. 41, 949–958.

USDA-USCC, 2002. Test Methods for the Examination of Composts and Composting. Holbrook, NY: Composting Council Research and Education Foundation.

Viscarra Rossel, R.A., Beherens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54.

Wu, G., Kechavarzi, C., Li, X., Wu, S., Pollard, J.T., Sui, H., Coulon, F., 2013. Machine learning models for predicting PAHs bioavailability in compost amended soils. Chem. Eng. J. 223, 747–754.

Wu, L., Ma, L., 2001. Effects of sample storage on biosolids compost stability and maturity evaluation. J. Environ. Qual. 30, 222–228.