JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

J. Agric. Food Chem. 2010, 58, 3093–3100 3093
DOI:10.1021/jf903872r

# Discrimination of Farm Waste Contamination by Fluorescence Spectroscopy Coupled with Multivariate Analysis during a Biodegradation Study

Muhammad Bilal,[†,‡,§] Anne Jaffrezic,*[,†,‡,§] Yves Dudal,[∥] Cedric Le Guillou,[†,‡,§] Safya Menasseri,[†,‡,§] and Christian Walter[†,‡,§]

[†]Agrocampus Ouest, and [‡]INRA, UMR1069, Soil Agro and HydroSystem, F-35000 Rennes, France, [§]Université Européenne de Bretagne, France, and [∥]UMR Eco&Sols - Ecologie fonctionnelle et Biogéochimie des Sols, Montpellier SupAgro-INRA-IRD, F-34060 Montpellier Cedex 2

The persistence of potential tracers of dissolved organic matter (DOM) generated from farm waste-amended soil was investigated by fluorescence spectroscopy coupled with classification and regression tree (CART) and principal component analysis (PCA) during a short-term (8 days) to midterm (60 days) biodegradation study. Pig manure (PM), cow manure (CM), wheat straw (WS), and soil alone (SA) treatment inputs were used. Waste amendments were potential sources of higher DOM concentrations. PCA revealed the DOM quality differences between farm wastes and soil alone as well as a significant shift observed from the biochemical to the geochemical fluorescent fraction in SA and PM treatments. The tryptophan:Humic-Like ratio and tryptophan zone were the potential discriminators of recent and midterm pollution by farm wastes. Integral intensities of the Fulvic-Like zone and region III discriminated the PM from CM and WS during the 60 days. CART analysis showed 90 and 100% potential for farm wastes discrimination from soil during P1 and P2, respectively. The prediction successes were 72 and 57% for PM from other wastes and 60 and 100% for WS during both periods. Fluorescence spectroscopy in combination with CART analysis can be a nondestructive innovative method for monitoring susceptible farm waste contamination.

**KEYWORDS: Farm waste; soil; fluorescence spectroscopy; CART analysis; biodegradation**

## INTRODUCTION

Elevated dissolved organic matter (DOM) concentrations have been reported in fresh water environments across Europe and North America (*1*). This increase has a significant impact on the functioning of aquatic ecosystems (*2*) and leads to the formation of carcinogenic disinfection byproducts such as trihalomethan (*3*) during the chlorination process of water treatment.

Agricultural land spreading of farm wastes for plant nutrient recycling and crop production improves soil quality [organic matter contents, physical properties such as aggregate structural stability, texture, porosity, infiltration, water-holding capacity, and biological activity (*4*)]. However, it also increases the potential for negatively impacting the environmental quality through significantly higher DOM level in soils (*5*), which ultimately reaches rivers that drain these cultivated, amended soils (*6*). Plant biomass, litter leachates, root exudates, soil humus, and microbial degradation products are also considered the main sources of DOM in soil (*5*). Agricultural intensification has a major impact on the increase in DOM concentration through land use change and soil disturbance, farm waste soil amendments (*7*−*9*), and higher mobilization of native soil carbon due to animal waste (*10*, *11*). It is thus essential to gain insight into how DOM issued from these farm wastes changes upon decomposition when it comes in contact with soil after amendments.

Biodegradation kinetics of soluble organic matter highlight two fractions: a rapidly decomposable fraction with a turnover time of less than 1 day (containing 29−36% of the total carbon) and a slowly decomposable fraction with a turnover time of about 80 days (*12*). However, much less research has been done to acknowledge the biodegradation potential of farm wastes DOM after soil amendment. Animal fecal contamination in rivers has been investigated with biomarkers of sterol and bile acids (*13*) and sterol/stenol in pig slurry (PS) (*14*). The characterization of these tracers requires solvent extraction and chromatographic detection. There is a need to develop cheap and nondestructive tools for tracing these heterogeneous sources of DOM as a prerequisite to management actions for river water quality restoration at catchment scale.

In various environmental applications, three-dimensional fluorescence excitation−emission matrix (3D-EEM) spectroscopy has been used for monitoring and discrimination of organic matter in soil and lakes considering fluorescence intensity peaks and their ratios with a peak-picking method (*15*). Humic-Like (HL) peak C and tryptophan (TRY) and tyrosine-like peaks T and B have been used for monitoring DOM in treated effluents, farm wastes, treated sewage wastes, and sewer discharge (*16*−*19*) and in coastal environments subjected to anthropogenic inputs (*20*).

*To whom correspondence should be addressed. Tel: (33)(0)2 23 48 54 20. Fax: (33)(0)2 23 48 54 30. E-mail: anne.jaffrezic@agrocampus-ouest.fr.



Published on Web 02/12/2010

However, in the current study, instead of taking a few data points in the form of peak picking, whole 3D-EEM spectra were analyzed quantitatively with fluorescence regional integration (*21*).

Besides this, machine learning multivariate analysis is an ideal tool for the exercise when large data sets are involved. Recent literature highlights the performance of multivariate techniques (principal component analysis, PCA) in fluorescence fingerprinting of DOM for water treatment EEM (*22*, *23*) and a hierarchical clustering method for DOM sourcing of marine water samples (*24*). Parallel factor analysis (PARAFAC) also helped to characterize the fluorescent landscape of DOM from aqueous extracts of soils and soil amendments by decomposing the fluorescent EEM into different independent fluorescent components (*25*). These methods have the advantage of saving time and being a more accurate analysis over the traditional peak-picking technique. In the current study, we introduced classification and regression tree (CART) analysis, a nonparametric data-mining approach, for the class membership of a categorical-dependent variable without getting any assumption about the distributions of the variables (*26*). The aims of this study are 2-fold: (i) to investigate the potential of 3D fluorescence spectroscopy coupled with CART analysis to identify the optical tracers of DOM released from soil alone and from farm wastes-amended soil [pig manure (PM), cow manure (CM), and wheat straw (WS)] and (ii) to analyze the short-term to midterm persistence of fluorescence indices of farm waste contamination during a biodegradation experiment.

## MATERIALS AND METHODS

The topsoil horizon from an agricultural field was sampled after wheat crop harvest from the experimental station of Kerguéhennec in Morbihan, East Brittany, France. The soil, derived from mica schist, is a Humic Cambisol (FAO) with a loamy texture (17% clay, 42% silt, and 41% sand), an organic matter content of 37 g $kg^{-1}$, and a pH (H20) of 6.0.

**Organic Products Characterization and Experimental Design.** A crop residue (WS) and two farm manures, that is, PS and CM, were used as organic amendments. PS samples were separated into its liquid and solid parts through centrifugation at 3000 rpm for 30 min. A solid PS was used and referred to as PM in this study. Total organic C and N contents of these materials were determined by an elemental analyzer (Flash EA 1112, Thermofinningan, Milan, Italy). The C:N ratios of PM, CM, and WS were 10, 33, and 110, respectively. C:N ratios of PM and CM were comparable to farm wastes studied by Morvan (*27*) in which C:N ratios for PM and CMs were $< 15$ and $> 25$, respectively.

**Experimental Conditions of Biodegradation.** In the laboratory, soil samples were air-dried and crumbled manually by removing the unrefined residues of organic matter. Soil aggregates were chosen after the samples were sieving through 3.15−5 mm mesh size and then stored in the darkness at 4 °C. The aggregates were moistened by capillary action and then subjected to 2.5 pF to attain a water-holding capacity of 21.2%. Soil samples were preincubated at 25 °C for 6 days before the experiment to minimize microbial activity variation due to temperature change. The organic materials were air-dried and crushed to 1 mm particle size and then incorporated homogenously into the moist, sieved, and preincubated soil at a rate of 4 g C $kg^{-1}$ dry soil. The soil mineral N content was adjusted to 75 mg N/kg dry soil by adding a potash fertilizer ($KNO_3$) solution to ensure mineral nitrogen availability for the microorganisms during biodegradation and a follow-up for mineral nitrogen content was done during the whole study time. Samples were incubated at 25 °C in hermetically closed jars in the darkness. A tube containing 40 mL of deionized water was introduced in each jar to minimize sample desiccation. The atmosphere in jars was regularly renewed to maintain an aerobic environment for microbial degradation. All of the treatments were sampled after 0, 3, 7, 15, 30, and 56 days after incubation along with three replicates. We divided the whole data for fluorescent DOM characterization into period P1 (0 −7 days after incubation) and period P2 (08−56 days after incubation). We marked periods P1 as short-term and P2 as midterm farm wastes pollution.

**Extraction of DOM.** DOM was extracted with a 2:1 ultrapure water to soil ratio. Soil−water suspensions were shaken mechanically on an orbital shaker for 2 h and then centrifuged at 3000 rpm for 30 min and filtered through 0.7 and 0.22 $\mu$m nitrocellulose filters. To avoid any contamination, all of the filters were rinsed with ultrapure water and dried overnight before vacuum filtration.

**Chemical Analysis.** The dissolved organic carbon (DOC) in each solution was measured on a Shimadzu TOC 5050 A total carbon analyzer. The accuracy on DOC measurements was ±5%, based on repeated measurements of standard solutions (K-phtalate). The UV−visible absorbance was measured on a Perkin-Elmer Lambda 20 UV−visible spectrophotometer across 200−600 nm excitation wavelengths range with a data interval of 0.5 nm, a slit width of 2 nm, and a scan speed of 120 nm/min.

Fluorescence measurements of DOM were performed using a Perkin-Elmer LS-55B luminescence spectrometer. The spectrophotometer used a xenon excitation source, and slits were set to 5 nm for both excitation and emission. To obtain EEM spectra, excitation wavelengths were incremented from 200 to 425 nm at steps of 5 nm, and emission was detected from 250 to 600 nm with a 0.5 nm step. The scan speed was set at 1500 nm/min, yielding an EEM in 22 min with 45 total scans. To minimize the temperature effect, samples were allowed to equilibrate with room temperature (20 ± 2 °C) prior to fluorescence analysis. The whole fluorescence data set presented in this study was normalized at 5 mg $L^{-1}$ DOC. Linearity was carried out between DOM concentration and fluorescence intensity with dilution of high DOM concentration samples. To eliminate the second-order Raleigh light scattering, excitation and emission cutoff filters were applied at 230−310 and 380−600 nm, respectively, on the lower side of 3D plots (**Figure 1**).

Inner filter effects were removed with the formula (*28*). To maintain the consistency of measurements and standardize the whole fluorescence data set, all of the integrated fluorescence intensities were normalized to average Raman emission intensity units of 19 for ultrapure water samples ($n = 25$) at excitation and emission wavelengths of 350 and 397 nm, respectively. A Raman-normalized integrated EEM spectrum of ultrapure water was subtracted from the data sample to eliminate the water Raman scatter peak.

**Regional Integration of EEM.** An internal program was developed in the laboratory using the R software (http://www.r-project.org) for the integration of fluorescence intensities across the whole EEM landscape. Here, peaks at shorter wavelengths ($< 250$ nm) and shorter emission wavelengths ($< 380$ nm) are related to simple aromatic proteins such as tyrosine and TRY (regions I and II). Peaks at intermediate excitation wavelengths (250−340 nm) and shorter emission wavelengths ($< 380$ nm) are related to soluble microbial byproduct-like material (region IV), while peaks located at the excitation wavelengths (230−300 nm) and the emission wavelengths (380−575 nm) represent Humic-Like substances (region III). Peaks at longer excitation wavelengths ($> 300$ nm) and longer emission wavelengths ($> 380$ nm) are related to Fulvic-Like organics (region V). With this technique, EEM is divided into biochemical (bio) (I, II, and IV) and geochemical (geo) (III and V) fluorescent regions (**Figure 1a**) and three peak intensity zones of TRY, Fulvic-Like (FL), and HL fluorescence (**Figure 1b**).

The quantitative analysis included the integration of fluorescent volume beneath each region and zone. Moreover, ratios TRY:HL, TRY:FL, HL:FL, bio:geo, IV:V, and III:V were also calculated. Forty-five spectral loadings were used to reproduce 3D plots of fluorescence intensity as a function of excitation and emission wavelengths. The humification index (HIX) was determined according to Ohno (*28*).

**Statistical Analysis.** PCA was applied to the spectroscopic data of DOM issued from farm wastes and SA during biodegradation study periods P1 and P2 with R software (package ade4). Significant differences among the temporal shift of treatments were tested by one-way analysis of variance (ANOVA) ($p < 0.05$).

Unlike traditional statistical techniques, we applied a CART tree approach (*26*) as they were adopted to predict a qualitative property by selecting the most discriminant quantitative predictors. It can also handle numerical data that are highly skewed or multimodel with categorical predictors having either ordinal or nonordinal structures. CART used

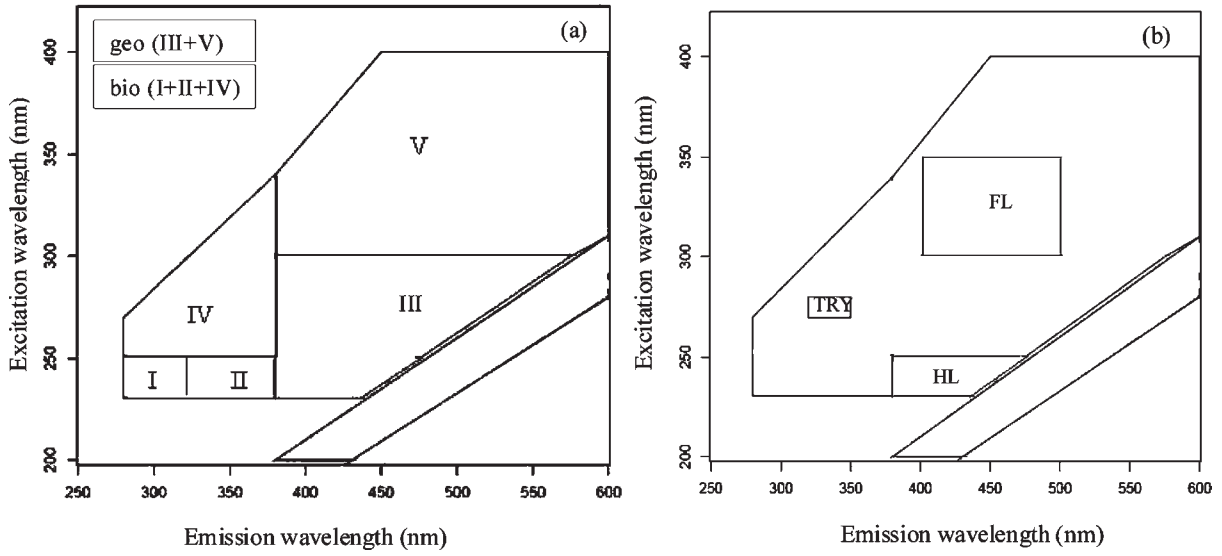


**Figure 1.** Integration of fluorescence intensities across regions (**a**) and maximum peak intensity zones (**b**).

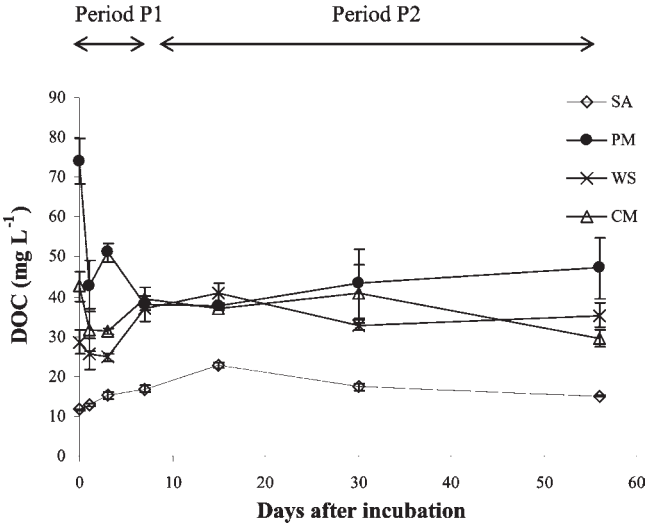


**Figure 2.** Time series of DOC concentrations of four treatments. Abbreviations: SA, soil alone. Bars indicate standard error (SE), and $N = 3$.

optimal univariate splits by carrying out an exhaustive search for all possible splits for each predictor variable and found the best split having higher improvement in the prediction accuracy. The tree structure started with the root node, which contains all of the observations of SA, WS, PM, and CM treatments in the form of histogram plots. The splitting of root node results in the child nodes, which again becomes the parent node if division continues and the nodes where division finishes or homogeneity occurs called terminal nodes. CART analysis was performed with STATISTICA (version 7.1).

## RESULTS

**Temporal Dynamics of DOC.** Period P1was marked by strong variations of DOC concentration in the farm wastes-amended soil. After 3 h, DOCs in PM, CM, and WS treatments were 73.8 ± 5.7, 42.5 ± 3.6, and 28.6 ± 2.9 mg $L^{-1}$, respectively, as compared to soil alone 11.6 ± 0.2 mg $L^{-1}$ (**Figure 2**). Within 24 h, the DOC decrease was more important in PM and CM treatments (31.1 and 10.9 mg $L^{-1}$) than in the WS treatment (2.9 mg $L^{-1}$).

DOC concentrations were almost the same in all of the farm waste treatments (38 mg $L^{-1}$) on the seventh day after incubation. During period P2, DOC concentrations were almost stable in farm wastes treatments. At the end of the study period, the PM treatment showed higher DOC values of 47.2 ± 7.5 mg $L^{-1}$ as compared to WS and CM treatments with 35.3 ± 2.9 and 29.2 ± 2.1 mg $L^{-1}$, respectively. During the whole study period, farm wastes showed higher DOCs as compared to soil alone. The DOC concentration in the SA treatment varied between 11.6 and 16.9 mg $L^{-1}$ during period P1, but during P2, DOC dynamics was stable except for a peak of 22.8 mg $L^{-1}$ on the 15th day after incubation.

**Spectral Differences among the Farm Wastes Treatments and Soil Alone.** PCA was applied to the integrated fluorescence properties of farm wastes and soil alone treatments, to investigate the spectral differences as well as to retrieve the additional information on temporal shift of the observed indices during period P2. A preliminary comparison of average was conducted to select the pertinent spectroscopic indices that discriminate the modalities.

Axes 1 and 2 of the PCA explained 47.5 and 28.6%, respectively, of variability in 14 spectroscopic indices of SA, PM, CM, and WS treatments distribution during both degradation periods P1 and P2 (**Figure 3**). SA treatment was clearly separated from the farm wastes treatments in opposite quadrants with negative scores on axis 1 during period P1 and positive scores on axis 2 in period P2 as shown in **Figure 3**. The average axis 2 score for SA treatments (2.59) was significantly higher in period P2 than during period P1 (0.63) ($p < 0.05$). Geochemical integrated fluorescence intensities across the regions geo (III + V) and the zones HL and FL, ratio HL:FL, and HIX had strong negative weightings on axis 1 (**Table 1**), which separated SA treatment from the farm wastes during period P1. However, during biodegradation period P2, only HIX separated the SA treatment with its positive weightings on axis 2. SA treatment during P2 showed a negative Pearson correlation ($r$) to TRY (−0.68) and to ratios TRY:HL (−0.92), TRY:FL (−0.88), bio:geo (−0.93), and IV:V (−0.96).

Among the farm wastes, ratios bio:geo, IV:V, TRY:HL, and TRY:FL had strong positive weightings on axis 1 (**Table 1**) where CM and WS treatments grouped together and separated from PM treatment in both periods (**Figure 3**). Axis 2 of PCA discriminated the PM treatment from the WS and CM treatments during both periods. There were significantly higher ($p < 0.05$) average scores for PM during P2 (−0.64) as compared to P1 (−2.66). Biochemical integrated fluorescence intensities across

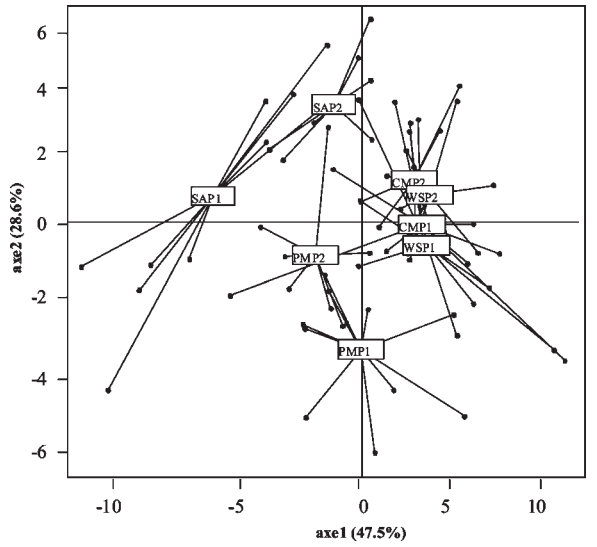


**Figure 3.** PCA analysis of farm wastes biodegradation study for two periods, P1 (0−7 days after incubation) and P2 (8−56 days after incubation). Abbreviations of farm wastes treatments: wheat straw, WSP1 and WSP2; pig manure, PMP1 and PMP2; cow manure, CMP1 and CMP2; and a control treatment, that is, soil alone, SAP1 and SAP2. The PCA run included the distribution of 14 variables (**Table 1**) of the integrated fluorescence properties of DOC and absorbance $A_{(365)}$ on axis 1 and axis 2.

**Table 1.** PCA Weightings for the Spectroscopic Parameters (Variables) during Biodegradation Study Periods P1 and P2

| variables | axis 1 | axis 2 |
|---|---|---|
| region IV | −0.34 | −0.92 |
| region III | −0.89 | −0.44 |
| geo | −0.88 | −0.44 |
| ratio bio:geo | 0.68 | −0.60 |
| ratio IILV | −0.45 | −0.16 |
| ratio IV:V | 0.75 | −0.62 |
| FL | −0.88 | −0.43 |
| HL | −0.91 | −0.27 |
| TRY | 0.25 | −0.93 |
| ratio TRY:HL | 0.78 | −0.59 |
| ratio TRY:FL | 0.70 | −0.52 |
| ratio HL:FL | −0.73 | 0.11 |
| $A_{(365)}$ | −0.13 | 0.05 |
| HIX | −0.79 | 0.54 |

region IV and TRY zone had strong negative weightings on axis 2 (**Table 1**) and separated the PM during P1 from the rest of the farm wastes and SA treatments. However, during P2 in PM treatment, fluorescence indices shifted from biochemical (TRY zone and region IV) to geochemical (geo, FL, and HL) fluorescent fractions.

**CART Analysis.** *Farm Wastes Tracer during Period P1.* Different tree structures for the P1 data set are shown in **Table 2**, and tree number 2 was chosen as an optimal tree (marked "*") with the minimal cost−complexity measures [cross-validation (CV) cost-misclassification costs of test samples and resubstitution cost-misclassification cost of learning sample data set] and node complexity (a penalty for additional terminal nodes). Terminal nodes numbers described the complexity measurement. The tree structure complexity decreased from tree 1 to tree 8. The tree structure with one terminal node showed equal misclassification costs (CV cost and resubstitution cost). The optimum tree structure obtained at the end of pruning is drawn in **Figure 4**.

**Table 2.** Cost Complexity Measures of All Possible Trees for the Period P1 Data Set

| all possible trees | terminal nodes numbers | CV cost | CV SE | resubsititution cost | node complexity |
|---|---|---|---|---|---|
| 1 | 8 | 0.325 | 0.067 | 0.000 | 0.000 |
| **2*** | **6** | **0.302** | **0.067** | **0.045** | **0.023** |
| 3 | 5 | 0.373 | 0.069 | 0.095 | 0.050 |
| 4 | 3 | 0.395 | 0.068 | 0.295 | 0.100 |
| 5 | 2 | 0.500 | 0.031 | 0.500 | 0.205 |
| 6 | 1 | 0.750 | 0.000 | 0.750 | 0.250 |

In this optimal tree constructed, there were five child nodes (dotted line squares) and six terminal nodes (solid black line squares). The integrated fluorescence intensities ratio Try:HL was the first splitter that divided the root node into a terminal node containing all of the observation of SA treatment and a child node separating the farm wastes treatments.

Among the farm waste treatments, integrated fluorescence intensities across the FL zone classified PM treatment from CM and WS at node 3. The second discriminator of farm waste treatments was ratio III:V, which separated the WS from the CM treatment. Finally, TRY zone differentiated the CM from WS treatment and allocated it to terminal node 11. However, confusion remained in the discrimination of CM treatment as it was often misclassified with WS treatment.

The prediction accuracy was assessed by a CV approach as shown in **Table 3**. The overall prediction accuracy of farm wastes treatments as well as soil alone was 62.7% for the period P1 data set. The optimum tree (**Figure 4**) demonstrated a high accuracy (90.9%) in predicting SA treatment and relatively high (72.7%) for PM treatment and fair prediction accuracy (60%) for WS treatment, but CM treatment was poorly predicted (27.3%). Among the farm wastes, there was almost complete discrimination between PM and CM treatments with only a 9% CM misclassification rate with PM treatment. However, the misclassification rate of CM treatment was high (63.6%) with WS treatment.

*Farm Wastes Tracer during Period P2.* All possible trees for period P2 are shown in **Table 4** with tree 3 marked ("*") as an optimal tree after pruning. The optimal tree structure obtained at the end of pruning is shown in **Figure 5**. The first discriminator splitting the root node was the integrated fluorescence intensities across TRY zone, which classified SA treatment from the farm wastes. Among the farm waste treatments, integrated fluorescence intensities across region III discriminated PM treatment from CM and WS treatments. Spectral absorbance $A_{365}$ discriminated WS treatment, but CM was mostly misclassified with WS treatment.

The prediction accuracy assessment of optimum tree for the biodegradation period P2 was 66.7%. This tree had a high accuracy (100%) for predicting SA and WS treatments and fair accuracy (57.1%) for PM prediction, but the prediction accuracy for CM treatment was 0% as it misclassified with WS treatment (**Table 5**). During the biodegradation period P2, discrimination of PM treatment from CM was 100%, but 28.6% was misclassified with WS treatment. The CM treatment was 100% misclassified with WS treatment, while the WS treatment was 100% correctly classified from the rest of the farm wastes treatments.

## DISCUSSION

**Impact of Farm Wastes on DOM Production during Biodegradation.** Significantly higher DOM concentrations in the farm waste treatments throughout the incubation experiment confirm the impact of farm waste manuring on soil DOM concentrations.

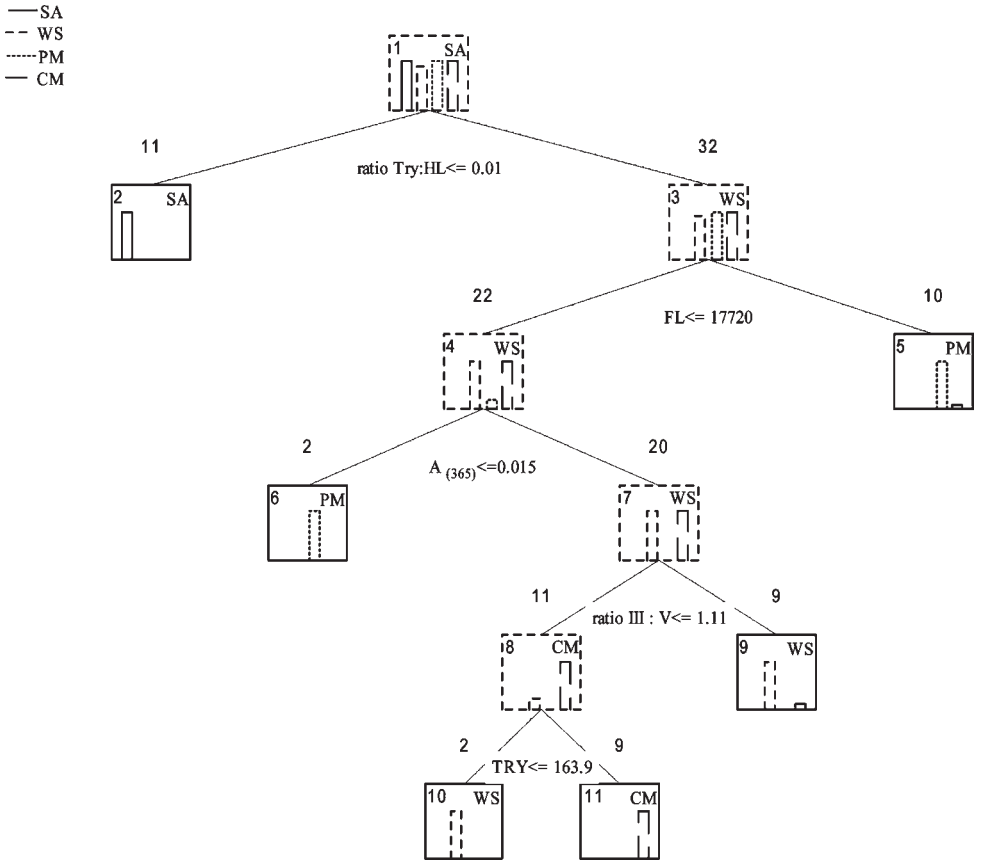


**Figure 4.** Optimum tree for the fluorescence properties of DOC issued from the farm wastes during biodegradation for period P1 (0−7 days after incubation). Treatment abbreviations: SA, soil alone. Predictor variable abbreviations are integrated fluorescence intensities of across zones of FL, HL, and TRY, ratio TRY: HL and regional ratio III:V of integral intensities across regions III and V, and spectral absorbance $A_{(365)}$.

**Table 3.** Confusion Matrix of Predicted versus Observed Treatment Resulting from a Cross-Validation Procedure Applied on the Optimum Tree for Period P1

| | observed | | | |
|---|---|---|---|---|
| | SA | WS | PM | CM |
| predicted | *n* = 11 | *n* = 10 | *n* = 11 | *n* = 11 |
| SA (*n* = 10) | **90.9%** | 0% | 0% | 0% |
| WS (*n* = 16) | 0% | **60%** | 30% | 63.63% |
| PM (*n* = 11) | 0% | 20% | **72.7%** | 9.09% |
| CM (*n* = 6) | 10% | 20% | 0% | **27.27%** |
| | total accuracy rate (*n* = 43) | | | 62.79% |

**Table 4.** Cost Complexity Measures of All Possible Trees for the Period P2 Data Set

| all possible tree | terminal nodes numbers | CV cost | CV SE | resubsititution cost | node complexity |
|---|---|---|---|---|---|
| 1 | 6 | 0.277 | 0.071 | 0.000 | 0.000 |
| 2 | 5 | 0.277 | 0.071 | 0.035 | 0.035 |
| **3*** | **4** | **0.305** | **0.074** | **0.107** | **0.071** |
| 4 | 3 | 0.357 | 0.046 | 0.250 | 0.143 |
| 5 | 1 | 0.750 | 0.000 | 0.750 | 0.250 |

Previous studies had recognized similar trends of DOM in cultivated soil (*5*, *29*) as well as in the rivers draining farm waste-fertilized catchments (*6*). In the soil alone treatment, the DOM peak 15 days after incubation indicated the possible DOM release from dead microbial biomass that starved from the depletion of substrate.

A strong decrease in DOM concentrations in PM and CM treatments within 24 h suggested the presence of a rapidly biodegradable fraction of DOM (23−41% decomposable soluble carbon in CM and PM, respectively, in our study), and this decrease could also be related to the preferential consumption of simple carbohydrate monomers, organic acids, and protein fractions of DOM during the initial phase of decomposition (*30*). DOM dynamics during both P1 and P2 periods suggested a more biodegradable DOM fraction in farm waste treatments as compared to soil alone (*12*). The DOM pool demonstrated stability against biodegradation up to 30 days in CM treatment, and subsequent decline reflected its higher susceptibility to biodegradation as compared to PM-amended soil treatment after 30 days. At the end of the experiment, a significantly higher DOM in PM treatment as compared to CM treatment ($p < 0.05$) indicated a higher DOM production potential of PMs, whereas others (*31*) found an opposite trend of increasing DOM concentration in the CM and decreasing in PM after decomposition. This reflects the variability of diet fiber contents that can have a great effect on wastes composition for a given type of animal (*32*). Using only the DOM parameter, farm wastes were discriminated from soil alone with higher DOM concentrations and also PM discriminated from CM and WS in the start and only from CM in the end of the biodegradation period.

**Persistence of Spectral Indices of Soil and Farm Waste Using PCA Analysis.** Temporal variability of fluorescence properties of DOM released from PM and soil alone was detected using PCA analysis (**Figure 3**). Therefore, spectral indices are not persistent in PM and soil alone treatments. From a qualitative point of view, strong similarities were observed in DOM fluorescence indices from CM and WS soil extracts, which reflects the same spectral composition of DOM. As a consequence, certain persistence of

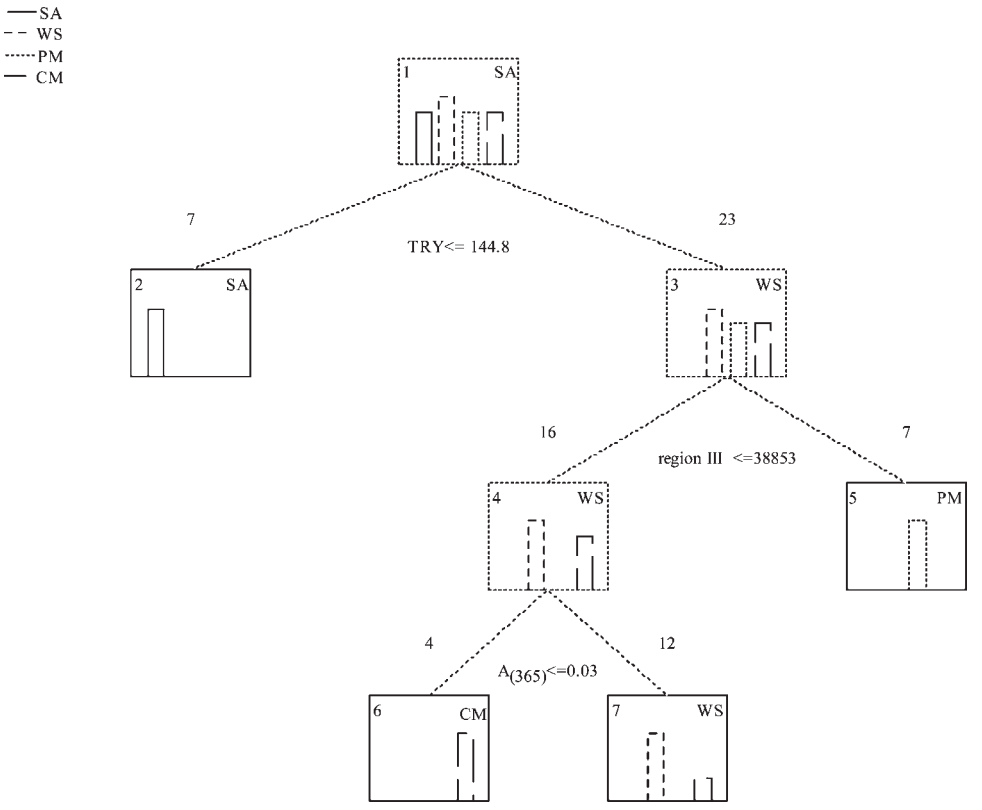


**Figure 5.** Optimum tree for the fluorescence properties of DOC issued from the farm wastes during biodegradation for period P2 (8−56 days after incubation). Abbreviations are the integrated fluorescence intensities across TRY zone and region III and spectral absorbance, $A_{(365)}$.

**Table 5.** Confusion Matrix of Predicted versus Observed Treatment Resulting from a Cross-Validation Procedure Applied on the Optimum Tree for Period P2

| | observed | | | |
|---|---|---|---|---|
| | SA | WS | PM | CM |
| predicted | $n = 7$ | $n = 9$ | $n = 7$ | $n = 7$ |
| SA ($n = 8$) | **100%** | 0% | 14.28% | 0% |
| WS ($n = 18$) | 0% | **100%** | 28.57% | 100% |
| PM ($n = 4$) | 0% | 0% | **57.1%** | 0% |
| CM ($n = 0$) | 0% | 0% | 0% | 0% |
| | total accuracy rate ($n = 30$) | | | 66.67% |

fluorescence signature is observed (**Figure 3**). After canceling out the carbon rate differences among the farm waste input (4 g C/kg dry soil) and DOM differences among all of the treatments during fluorescence measurements (fluorescence intensities normalized at 5 mg $L^{-1}$), the distinction between soil and farm wastes along the PCA axes during both study periods reflected the DOM quality differences. PM could be discriminated from WS and CM treatments by biochemical integrated fluorescence across region IV and the TRY zone during P1. This suggests heterogeneity in DOM quality among the farm wastes. The data illustrate the wide variation and dissimilar effects of decomposition on the TRY zone and region IV among the farm wastes during period P1. A temporal shift of PM treatment from biochemical (region IV and TRY zone) to geochemical fluorescence (HL, FL, and geochemical region) properties from period P1 to P2 confirms the biodegradation of biochemical fluorescence indices in P1. However, during period P2, the presence of more condensed aromatic structures and humified fluorescent fraction of DOM in the PM treatment indicates the persistence in the biodegradation environment and can be related to high organic matter degradation. For SA treatment during P2, a strong negative correlation between HIX and ratios TRY, TRY:HL, TRY:FL, bio:geo, and IV:V suggests that it can be discriminated with higher HIX and lower ratios of TRY:HL, TRY:FL, bio:geo, and IV:V during midterm biodegradation from farm wastes. Strong structural changes of DOM must have occurred during the degradation process, leading to a higher increase in carboxylic groups in soil and preferential consumption of protein contents that result in higher humification, and as a consequence, HIX discriminates soil from the farm wastes. Zsolnay (*33*) also calculated HIX to differentiate the microbial cell lysis products and more humified DOM. Biodegradation effects on DOM are not coherent among farm wastes studied as we observe an evolutionary trend in fluorescence indices of PM but lack of significant evolution of DOM fluorescence properties in CM. This reflects the variation in the chemical properties of feed materials as well as the different digestive processes of the animals (*31*).

**Potential of CART Analysis for Discriminating the Farm Wastes during Biodegradation.** The CART tree approach (*26*) enabled us to find the best predictor/tracer of various farm waste treatments during two biodegradation study periods, P1 and P2. We hypothesize that farm wastes contamination can be short-term (recent contact of farm wastes with water, 0−7 days) or midterm (through runoff from farm waste spreading on cultivated hillslopes after 1 or 2 months). Our results suggest that short-term farm wastes pollution can be traced with higher ratio TRY:HL values (split value ≥0.013 RU) and average farm wastes pollution with higher TRY zone values (split value ≥144.8 RU) and qualify as potential tracers of farm wastes. Among the farm wastes treatments, the FL zone is ranked as the most discriminant predictor of PM during period P1, and the FL zone shows its positive correlation with biochemical region IV ($r$, 0.77), which suggests that the FL zone and region IV can trace the fluorescent fraction of PM during P1. Region III during period P2 is the best predictor of PM, and its weaker correlation ($r$, 0.17) with TRY

confirms the degradation of biochemical fluorescent fraction of PM. It also suggests that the fluorescent fraction of PM treatment gets more humified as we observe that region III discriminates the PM treatment during period P2. Spectral absorbance $A_{365}$ qualifies as a potential tracer of WS during period P2, which identifies the increasing chromomorphic fraction of DOM during WS biodegradation. The ratio III:V is suggested as the only discriminator of WS that separates from CM treatment with 60% classification success. The misclassification rate of CM with WS during both periods of biodegradation indicates the presence of a common substrate quality, that is, residues of WS in CM treatment. The potential of CART analysis success for predicting the farm wastes treatments as well as soil alone was estimated by CV to be globally of 63 and 66% for both periods P1 and P2, respectively. We also tested the performance of CART analysis by using the same fluorescence properties of DOM from three incubated soil samples (test sample) (similar type of soil as used in current study) along with the data set of period P1. The CART tree correctly classified the test samples with soil with the same variable of TRY:HL. During period P2, we obtained globally the same tree structure, but the tree was less complex and easier to interpret as compared to the tree in period P1. Classification success for SA treatment (91 and 100% for P1 and P2, respectively) suggests the compositional differences in soil DOM as compared to farm wastes.

Fluorescence spectroscopic characterization in combination with PCA analysis reflected the degradation of biochemical fluorescence indices during short-term contamination in PM and shifted toward geochemical integral intensities in midterm pollution with more condensed and humified geochemical structures of FL and HL substances, which could persist in the degradation environment. CART analysis enabled us to trace farm waste contamination by considering stepwise the most discriminant variable selection and complexity reduction. Farm wastes were discriminated from soil alone with the ratio TRY:HL and TRY zone during short and midterm pollution with prediction successes of 90 and 100%, respectively. PM waste was discriminated from CM and WS by FL zone and region III with prediction accuracies of 72.7 and 57.1%, respectively. WS was classified from CM by $A_{365}$ with 100% accuracy rate during P2. However, CM was generally found to be misclassified with WS due to common substrate quality. This investigation underlines the potential of 3D excitation−emission fluorescence spectroscopy in combination with CART analysis as a nondestructive innovative method for monitoring farm waste DOM contamination.

This method was tested as an alternative method to PARAFAC with a simple fluorescence index based on a regional integration procedure. PARAFAC is a robust method that is very efficient in obtaining spectral images of DOM components and accounts for physical phenomena, that is, the lack of distinctly separated spectral areas and often-observed overlapping of emission peaks components isolated via PARAFAC.

CART analysis is found useful as it extracts the most salient information from the large data set and also gives a misclassification probability for the classifier. The CART tree procedure also gives easily interpreted information regarding the predictive structure of the data. However, the potential of the CART approach for discrimination of DOM has to be tested by another data set with different types of soils and animal wastes.

## ACKNOWLEDGMENT

We acknowledge Armelle Racape and Patrice Petitjean for their technical assistance in experimentation and analytical analysis. We also thank Nicolas Bottinellli for his comments and suggestions in data analysis.

## LITERATURE CITED

(1) Worrall, F.; Burt, T.; Shedden, R. Long term records of riverine dissolved organic matter. *Biogeochemistry* **2003**, *64* (2), 165–178.
(2) Kalbitz, K.; Wennrich, R. Mobilization of heavy metals and arsenic in polluted wetland soils and its dependence on dissolved organic matter. *Sci. Total Environ.* **1998**, *209* (1), 27–39.
(3) Sirivedhin, T.; Gray, K. A. Comparison of the disinfection by-product formation potentials between a wastewater effluent and surface waters. *Water Res.* **2005**, *39* (6), 1025–1036.
(4) Barzegar, A. R.; Yousefi, A.; Daryashenas, A. The effect of addition of different amounts and types of organic materials on soil physical properties and yield of wheat. *Plant Soil* **2002**, *247* (2), 295–301.
(5) Kalbitz, K.; Solinger, S.; Park, J.-H.; Michalzik, B.; Matzner, E. Controls on the dynamics of dissolved organic matter in soils: A review. *Soil Sci.* **2000**, *165* (4), 277–304.
(6) Jardé, E.; Gruau, G.; Mansuy-Huault, L. Detection of manure-derived organic compounds in rivers draining agricultural areas of intensive manure spreading. *Appl. Geochem.* **2007**, *22* (8), 1814–1824.
(7) Molinero, J.; Burke, R. Effects of land use on dissolved organic matter biogeochemistry in piedmont headwater streams of the Southeastern United States. *Hydrobiologia* **2009**, *635*, 289–308.
(8) Royer, I.; Angers, D. A.; Chantigny, M. H.; Simard, R. R.; Cluis, D. Dissolved organic carbon in runoff and tile-drain water under corn and forage fertilized with hog manure. *J. Environ. Qual.* **2007**, *36* (3), 855–863.
(9) Chantigny, M. H. Dissolved and water-extractable organic matter in soils: A review on the influence of land use and management practices. *Geoderma* **2003**, *113* (3−4), 357–380.
(10) Bol, R.; Ostle, N. J.; Friedrich, C.; Amelung, W.; Sanders, I. The Influence of Dung Amendments on Dissolved Organic Matter in Grassland Soil Leachates—Preliminary Results from a Lysimeter Study. *Isot. Environ. Health Stud.* **1999**, *35* (1), 97–109.
(11) Shand, C. A.; Williams, B. L.; Smith, S.; Young, M. E. Temporal changes in C, P and N concentrations in soil solution following application of synthetic sheep urine to a soil under grass. *Plant Soil* **2000**, *222* (1), 1–13.
(12) Gregorich, E. G.; Beare, M. H.; Stoklas, U.; St-Georges, P. Biodegradability of soluble organic matter in maize-cropped soils. *Geoderma* **2003**, *113* (3−4), 237–252.
(13) Tyagi, P.; Edwards, D.; Coyne, M. Fecal Sterol and Bile Acid Biomarkers: Runoff Concentrations in Animal Waste-Amended Pastures. *Water, Air, Soil Pollut.* **2009**, *198* (1), 45–54.
(14) Jardé, E.; Gruau, G.; Mansuy-Huault, L.; Peu, P.; Martinez, J. Using sterols to detect pig slurry contribution to soil organic matter. *Water, Air, Soil Pollut.* **2007**, *178* (1), 169–178.
(15) McKnight, D. M.; Boyer, E. W.; Westerhoff, P. K.; Doran, P. T.; Kulbe, T.; Andersen, D. T. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnol. Oceanogr.* **2001**, *46* (1), 38–48.
(16) Lee, S.; Ahn, K. H. Monitoring of COD as an organic indicator in waste water and treated effluent by fluorescence excitation-emission (FEEM) matrix characterization. *Water Sci. Technol.* **2004**, *50* (8), 57–63.
(17) Baker, A. Fluorescence properties of some farm wastes: implications for water quality monitoring. *Water Res.* **2002**, *36* (1), 189–195.
(18) Baker, A.; Inverarity, R. Protein-like fluorescence intensity as a possible tool for determining river water quality. *Hydrol. Processes* **2004**, *18* (15), 2927–2945.
(19) Saadi, I.; Borisover, M.; Armon, R.; Laor, Y. Monitoring of effluent DOM biodegradation using fluorescence, UV and DOC measurements. *Chemosphere* **2006**, *63* (3), 530–539.
(20) Parlanti, E.; Worz, K.; Geoffroy, L.; Lamotte, M. Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. *Org. Geochem.* **2000**, *31* (12), 1765–1781.
(21) Chen, W.; Westerhoff, P.; Leenheer, J. A.; Booksh, K. Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. *Environ. Sci. Technol.* **2003**, *37* (24), 5701–5710.

(22) Tartakovsky, B.; Lishman, L. A.; Legge, R. L. Application of multi-wavelength fluorometry for monitoring wastewater treatment process dynamics. *Water Res*. **1996**, *30* (12), 2941–2948.
(23) Peiris, R. H.; Hallé, C.; Budman, H.; Moresoli, C.; Peldszus, S.; Huck, P. M.; Legge, R. L. Identifying fouling events in a membrane-based drinking water treatment process using principal component analysis of fluorescence excitation-emission matrices. *Water Res*. **2010**, *44* (1), 185–194.
(24) Jiang, F.; Lee, F. S.-C.; Wang, X.; Dai, D. The application of excitation/emission matrix spectroscopy combined with multivariate analysis for the characterization and source identification of dissolved organic matter in seawater of Bohai Sea, China. *Mar*. *Chem*. **2008**, *110* (1−2), 109–119.
(25) Ohno, T.; Bro, R. Dissolved organic matter characterization using multiway spectral decomposition of fluorescence landscapes. *Soil Sci. Soc. Am. J*. **2006**, *70* (6), 2028–2037.
(26) Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification and Regression Trees*; Wadsworth: Belmont, CA, 1984.
(27) Morvan, T.; Nicolardot, B.; Péan, L. Biochemical composition and kinetics of C and N mineralization of animal wastes: A typological approach. *Biol. Fertil. Soils* **2006**, *42* (6), 513–522.
(28) Ohno, T. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. *Environ*. *Sci. Technol*. **2002**, *36* (4), 742–746.
(29) Shand, C.; Coutts, G. The effects of sheep faeces on soil solution composition. *Plant Soil* **2006**, *285* (1), 135–148.
(30) Marschner, B.; Kalbitz, K. Controls of bioavailability and biodegradability of dissolved organic matter in soils. *Geoderma* **2003**, *113* (3−4), 211–235.
(31) Hunt, J. F.; Ohno, T. Characterization of fresh and decomposed dissolved organic matter using excitation-emission matrix fluorescence spectroscopy and multiway analysis. *J*. *Agric*. *Food Chem*. **2007**, *55* (6), 2121–2128.
(32) Shriver, J. A.; Carter, S. D.; Sutton, A. L.; Richert, B. T.; Senne, B. W.; Pettey, L. A. Effects of adding fiber sources to reduced-crude protein, amino acid-supplemented diets on nitrogen excretion, growth performance, and carcass traits of finishing pigs. *J*. *Anim*. *Sci*. **2003**, *81* (2), 492–502.
(33) Zsolnay, A.; Baigar, E.; Jimenez, M.; Steinweg, B.; Saccomandi, F. Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. *Chemosphere* **1999**, *38* (1), 45–50.

**Received for review November 4, 2009. Revised manuscript received January 25, 2010. Accepted January 25, 2010. We gratefully acknowledge the Higher Education Commission of Pakistan for financing a Ph.D. scholarship.**