Prediction by near infrared spectroscopy of the composition of plant raw materials from the organic fertiliser industry and of crop residues from tropical agrosystems

L. Thuriès,[a,§,*] D. Bastianelli,[b] F. Davrieux,[c] L. Bonnal,[b] R. Oliver,[d] M. Pansu[e] and C. Feller[e]
[a]*Phalippou-Frayssinet SAS, Organic Fertilisers, La Mothe, F-81240 Rouairoux, France. E-mail: thuries@cirad.fr*

[b]*CIRAD-EMVT, Laboratory of Animal Nutrition, TA 30/A, F-34398 Montpellier Cedex 5, France*

[c]*CIRAD-CP, UPR33, TA 80/16, 73 Av. JF Breton, F-34398 Montpellier Cedex 5, France*

[d]*CIRAD-CA, UPR78, Organic Matter in Tropical Soils (MOST) Laboratory, TA 70/01, F-34398 Montpellier Cedex 5, France*

[e]*IRD, Organic Matter in Tropical Soils (MOST) Laboratory, BP64501, F-34394 Montpellier Cedex 5, France*

The dynamics of carbon (C) and nitrogen (N) of plant residues and organic fertilisers are of great interest for agricultural and global warming studies. The proportion of the fractions obtained from biochemical analyses (fibres by sequential Van Soest analysis) can be used for predicting both C and N transformation of organic materials in soils. Considering the expensive and time-consuming Van Soest method, the principal aim of this study was to elaborate near infrared (NIR) calibrations for fibres, in order to use them for consecutive studies (for example, our works on transformation of added organics or TAO model). A wide set of organic fertilisers and their raw materials was sampled, including plant materials originating from temperate (especially Mediterranean) and tropical regions. The particular objective of this work was to build NIR calibrations for fibre fractions, along with C and N content, in plant materials used in the organic fertiliser industry and green house gases mitigating strategies. The second particular objective was to test for two levels of validation of the equations previously elaborated: (1) validation with a set of randomly chosen samples that was not considered during the calibration step, (2) extrapolation of the predictive capacity of the equations when applying them to outliers that were previously discarded. The fibres were the best predicted parameters, as R^2=0.95, 0.91, 0.97, 0.97 for neutral detergent soluble, hemicelluloses, cellulose and lignin, respectively, whereas the characteristics of total organic matter had R^2 varying from 0.87 (N Kjeldahl) to 0.94 (C Dumas). The accuracy of the calibrations developed for fibres was confirmed by the first level of validation, since the standard errors of prediction were close to the corresponding standard errors of cross-validation and the standard errors of calibration. Nevertheless, the calibrations developed for ash and C Dumas were not so good. Surprisingly, at the second level of validation, some outliers were not so badly predicted. This can illustrate the robustness of the calibrations for cellulose, lignin and, to a lesser extent, N Dumas which are key parameters for our modelling works on C and N transformation of added organics in soils.

Keywords: plant raw materials organic fertiliser, Van Soest fibre analysis, lignin, NIR spectroscopy

Introduction

In both tropical and temperate climates, the added organic materials (AOM) take part in soil organic matter (SOM) building and represent an important source of nutrients for plants and soil organisms or a factor of soil property improvement. The AOM can be classified into two main groups: T1, farm manures or products from the organic fertiliser industry (including composts made from agri-food residues) and T2, plant residues from crop recycling or fallow. Utilising organic fertilisers, and especially composts, may lead to global benefits such as recycling and reducing the

[§]*Current address: IRD-CIRAD Organic Matter in Tropical Soils' Joint Laboratory (MOST), TA 70/01, F-34398 Montpellier Cedex 5, France.*

volume of agro-food wastes, a solution for waste management. From an environmental point of view, their utilisation can reduce the emission of greenhouse gases (GHG) due to fossil-fuel consumption linked to the synthesis, transport and application of chemical fertilisers. In tropical areas that are confronted with an increase in population, a decrease in available agricultural land and the limited funds available to small farmers to buy chemical fertilisers, different strategies may be involved to improve soil fertility while, if possible, mitigating GHG emissions by increasing carbon (C) sequestration. Most of the strategies must improve the management of organic residues from fallow production, crop recycling, direct drill with cover crops and agro-forestry systems.[1] Direct use of organic residues and use after transformation such as composting must be compared. It is, therefore, essential to compare the transformation of the AOM of both types (1) and (2) in soils and link them to the characteristics of AOM.

As far back as a century ago,[2] AOM composition was thought to have a direct influence on AOM decomposition/transformation. Along with the interaction of the physical–chemical environment, resource quality was believed to have a major influence on the decomposition process, through the action of decomposers.[3] Intense work is still devoted to investigating the relationships between the characteristics and transformation of AOM.[4] Among AOM characteristics, the C : N ratio and its fibre content, as measured by sequential chemical digestion by Van Soest,[5] have been found to be valuable descriptors of the potential capacity of C and N transformation.[6–9] Biochemical composition of AOM can thus be used as input data in various C and N models.[8–12]

Near infrared reflectance (NIR) spectroscopy has proven its usefulness (compared with classical time-consuming biochemical methods) for the estimation of a variety of parameters in diverse natural/agricultural products. Among them, nitrogen and fibres have been estimated by NIR for assessment of silage digestibility.[13–16] This has a potential interest for AOM decomposition studies, since plant degradation by ruminants presents some analogies with litter transformation in soils.[17] Total organic matter, carbon, nitrogen and fibres have also been estimated by NIR for the evaluation of litter quality and decomposition rates.[18–21]

NIR has also been used to predict N content in crop residues from Northern Europe[22,23] and raw materials such as poultry litter[24] utilised as is or for composting purposes. Precise C, N or fibre content has been estimated using NIR in several composts made from various raw materials (for example, swine excreta,[25] municipal solid waste,[26] manure and vegetables,[27] wheat straw and poultry litter[28] or tofu refuse[29]).

Other parameters of common AOM applied to soils have also been estimated by NIR. In soil and manured soil,[30–32] potential mineralisable C and N have been determined, along with the fibre content in manure.[31,33] Dynamics of C and N have also been studied by NIR through potential mineralisable N in soil from Northern[34–36] and Southern environments[37] and soil biomass and enzymatic activities[38,39] as important descriptors of their dynamics.

Our main objective was to test the ability of NIR to estimate the biochemical composition of a wide range of plant raw materials in order to provide input data for the transformation of added organics (TAO) model. The materials have been selected either because they were used as bases for the elaboration of composts and organic fertilisers or because they represented common situations in the tropics.

Material and methods

Organic materials

Our large database (> 2400 samples), consisting of plant materials, manures and composts, shows a huge variability. Since it was unrealistic to obtain good calibration equations on such a diverse set, it was thus decided to consider only the T1 and T2 types of plant materials for calibration ($n = 146$ samples). Their characteristics are presented in Table 1.T1 materials were plant raw materials originating from the agri-food industry which are used in the fertiliser industry: wet and dry grape skins, de-oiled grape pips, coffee cake, de-fatted cocoa bean, cocoa shell, olive pulp, maize cob, barley

Table 1. Chemical and biochemical characteristics of the total set ($n = 146$) of plant samples (g 100 g^{-1} DM).

(g 100 g^{-1} DM)	*n*	Mean	*SD*	Min.	Max.
Ash	146	7.5	4.3	0.5	32.7
C Dumas	143	46.2	2.2	33.5	49.8
N Dumas	137	2.1	1.1	0.26	7.9
N Kjeldahl	86	2.4	1.0	0.50	8.2
NDSol	146	26.9	12.2	2.2	63.4
Hem	146	14.0	9.4	0.04	48.9
Cel	146	25.7	12.3	4.2	64.7
Lig	146	25.9	16.3	3.4	63.9

straw, rice hull, rape seed cake, sunflower husks and soybean cake. T2 materials were plant materials collected in Brazil and Kenya originating from research studies about carbon sequestration in several agricultural systems. The samples originated from different species and nature: (i) trees (*Bactris gasipaes, Eritrina fusca, Gliricidia sepium, Inga edulis, Theobroma grandiflorum*), (ii) shrubs (*Flemingia macrophylla, Sesbania sesban, Tephrosia candida, Tephrosia vogelii*), (iii) crops (*Glycine max, Oryza sativa, Pennisetum tiphoides, Zea mays*) and (iv) cover crops (*Arachis pintoi, Brachiaria sp.*, *Canavalia ensiformis, Crotalaria grahamiana, Crotalaria paulina, Desmodium ovalifolium, Eleusine sp.*, *Mucuna edulis, Pueraria phaseoloides*). The samples of these tropical plants covered several plant parts: total aboveground material, roots, stems, twigs, pods and leaves. Their litters (pure/mixed with weeds) were also collected.

The raw materials from the agri-food industry are of a peculiar nature; they cannot be considered as native since they are processed. For example, the grape residues (wet/dry skins, pips) have been submitted to intense industrial processing methods (for example, wine elaboration, separation of the skins and the pips, oil extraction from pips). The same applies to cocoa residues (removal of the shell, butter extraction) or the olive pulp (grinding, oil extraction). Coffee cakes are also the result of several industrial processes (roasting, grinding and extraction of the soluble compounds). Coffee and grape skins often contain high amounts of water and they quickly begin to compost. These materials differ from those entering the agro-food industry or those leaving the field and, since they are more or less transformed, they resemble fresh manure or compost in their early stages of processing. Therefore, NIR predictions developed here differ from those dedicated to crop residues alone,[22–23] plant parts and litters from tropical[40] or Mediterranean shrubs and tree origins alone.[19–21] They encompass this variability. For these reasons, it seemed useful to widen the comparisons by comparing our results with other studies which involve not only different plant materials "as they were" (i.e. not transformed) but also works on compost and manure which has been more or less composted.

Preparation and organic matter characterisation

The AOM originating from the agri-food industry and used as raw materials in the organic fertiliser industry were sampled by the operators in the factory's QC laboratory. The sampling was done according to the factory's ISO 9001 procedure: when received and discharged from a truck, an AOM was immediately pre-sampled (about 10 kg) at different locations in the pile; it was then homogenised and reduced to a sample of about 1 kg and stored in a sealed plastic bag. Each sample was then analysed in duplicate for its moisture content (105°C in an oven until constant weight) organic matter and ash (loss on ignition, 525°C) and total nitrogen (Kjeldahl, NKj) content. These last two parameters and associated methods were recommended by Palm and Rowland[41] to be determined in order to provide some basic information and make comparisons of plant quality possible in larger datasets.

The AOM from tropical origins were collected on-field in Brazil and Kenya and were chopped into pieces about 2 cm long before further analysis. The roots were carefully cleaned with a diluted NaCl solution for the dispersion of the attached clay materials.

An aliquot of each AOM was rapidly dried in an aerated oven at 40°C to prevent N volatilisation and Maillard reactions.[42] It was then ground to pass a 1 mm sieve for a sequential analysis of fibres as reported by Van Soest *et al*.[43] Each ground AOM sample was successively extracted for NDF (neutral detergent fibre), ADF (acid detergent fibre) and ADL (acid detergent lignin). At each step of extraction, the products obtained were filtered, dried at 40°C, weighed and one replicate was dried at 105°C for determining residual moisture and then ignited gradually at 525°C for ash content, resulting in more replicates for NDF (six reps) than for ADF (four reps) and ADL (two reps). The different organic fractions (ash free) were calculated as follows:

$\text{NDSol (neutral detergent soluble)} = (\text{OM} - \text{Ash}_{\text{OM}}) - (\text{NDF} - \text{Ash}_{\text{NDF}})$;

$\text{Hem (hemicelluloses-like)} = (\text{NDF} - \text{Ash}_{NDF}) - (\text{ADF} - \text{Ash}_{\text{ADF}})$

$\text{Cel (cellulose-like)} = (\text{ADF} - \text{Ash}_{\text{ADF}}) - (\text{ADL} - \text{Ash}_{ADL})$

$\text{Lig (lignin-like)} = (\text{ADL} - \text{Ash}_{\text{ADL}})$.

A more complete description of the analytical procedure is given in Thuriès *et al.*[11] An aliquot of each 1 mm ground AOM was further ground to 100 μm and analysed for its C and N contents (C and N Dumas) with an elemental analyser (Fisons NA2000). It could be pointed out that some AOM were, therefore, analysed for nitrogen on both raw sample and dried sample, resulting in two distinct descriptions of N content. This is interesting, since most studies only involve N measurement on dried samples which have possibly been submitted to N losses during drying.

NIR spectra acquisition

Two replicates of each sample (dried at 40°C, < 1 mm ground) were packed in circular cups (50 mm diameter) sealed with paperback and then scanned (area scanned about 20 mm diameter) in reflectance mode on a monochromator spectrometer (NIRS 6500; Foss NIRSystems, Silver Spring, MD, USA). Spectral data were collected every 2 nm from 400 to 2498 nm. The spectra (average of 32 scans) were recorded as log (1/reflectance). Each sample was scanned twice (two different cup fillings) and the spectra were averaged.

Spectral pre-treatments, calibration and cross-validation

Four spectral pre-treatments were explored: standard normal variate and detrend (SNVD),[44] detrend, multiplicative scatter correction (MSC) and weighted multiplicative scatter correction (WMSC). The spectra were transformed by mathematical treatments according to the WIN-ISI software (Infrasoft International, Port Matilda, PA, USA): 1st or

2nd derivative calculated on five datapoints and smoothed (Savitzky and Golay smoothing) on five datapoints.

Calibration of the parameters studied here was performed using the modified partial least square regression (mPLS) of WIN-ISI (Infrasoft International, Port Matilda, PA, USA).[45] Calibration statistics include the standard error of calibration (*SEC*), the coefficient of determination (R^2) and the standard error of cross-validation (*SECV*). In order to minimise overfitting of the equations, cross-validation was used during calibration development. Cross-validation is a Jack-knife method based on the prediction of sub-groups (in this case, four subgroups were used) by equations developed on the rest of the database. It results in the calculation of a residual standard deviation (*SECV)*. In addition to the R^2, the *RPD* ratio ($RPDcv = SD/SECV$) was used to evaluate the general quality of the calibration developed for each parameter. Unlike *SEC* and *SECV*, *RPD* has the advantage of being independent of the parameter units[13] and can, therefore, be compared between parameters.

Outliers, validation and extrapolation

Some samples were removed from the calibration set either for being compositional or spectral outliers. Compositional outliers were identified as having extreme concentrations outside the normal range for their category. They consisted of one de-oiled grape pip sample (low NDSol, high Lig, due to high lignin + cutin contents in the pip wall); one maize stalk (low NKj and high Hem content); two de-fatted soybeans and one rape seed (high NDSol, due to a high level of soluble proteins, being proteaginous grains) and, finally, one *Brachiaria* and one *Eleusine* root (high mineral content, obviously due to clay contamination of the roots, although washed). The removal of spectral outliers was based on Mahalanobis distance $H > 3$ in the principal component analysis (PCA) from the average spectrum of the file.[45] They consisted of four grape skins, one cocoa shell, one sunflower's husks and seven tropical plants. Removing these outliers avoided calibrations stretched by the leverage of few points. All the 13 coffee cake samples had $H > 3$. The projection of these samples in the principal components space of the spectral variables (PCA space) was atypical; they were clearly isolated from the others. It was thus decided to distinguish two sets of outliers: (i) the coffee samples ($n = 13$) and (ii) the other outliers ($n = 20$).

The resulting 113 [i.e. 146 samples – (13 + 20) outliers] set of samples was further split into a calibration subset (CAL) and a validation subset (VAL). Among the 113 samples, 20 were randomly chosen, then discarded in the new calibration procedure. CAL thus comprised 93 samples and the remaining 20 samples were considered as VAL. The standard error of prediction (*SEP*) was estimated with the VAL subset and the *RPDp* was also calculated as $RPDp = SD_{VAL} / SEP$. The equations elaborated on CAL were also tested for their robustness on an extrapolation subset (EXT), which included the 13 coffee samples and the 20 other outliers described above. The interest of this procedure is to evaluate the ability of models to predict the composition of samples of the same nature (plants used in practice) but non-typical (here, outliers), which is extremely useful since, in practice, all types of sample are likely to arrive in the factory.

Results and discussion

The major characteristics of the 146 plant samples are presented in Table 1. As a preliminary remark, as several parameters were obtained by the difference from others, it could be thought that they were probably correlated (for example, the fibre parameters). Nevertheless, among fibres, there was only one high correlation ($r^2 = 0.54$, $p < 0.01$; data not shown) between Hcel and Lig. Also, N Dumas and Cel were significantly correlated ($r^2 = 0.55$, $p < 0.01$; data not shown). All the measured variables showed a wide range of values. Indeed, although all samples were of plant origin, their nature covered a high diversity of species from very different geographical and climatic origins. Although very different, the organic materials from plant origin were homogeneously located in the PCA space of spectral variables. The values of the measured parameters were continuously dispersed and they did not constitute discrete classes. Nitrogen (NKj*)* range and average content measured in our study were intermediate to those reported for temperate crop residues (1.85; 0.22–6.03%)[23] and for tropical plants (2.8; 0.45–5%),[40] whereas Lig range and average content were much higher in our materials (27; 3.4–63.9%) compared with temperate crop residues (0.1–23%),[46] tropical plants (11; 4–28%)[40] or Mediterranean plant parts and litters (5-36%).[19] Even in mushroom compost with similar mean Lig content (20.5%), the range was smaller (15.8–25.8%).[47] Mean Cel content (25.7%) was similar to and the range was higher than those reported for Mediterranean plant parts and litters (6–35%),[19] transformed plants as silage (27.4; 23.8–30.5%),[13] (2.0–5.5%)[16] or dairy manures (28.8; 17.3–49.5%).[31] The range of Cel values of our dataset was close to that of the temperate crop residues (7–58%).[46] As mentioned earlier, our set of materials covered both climatic origins encountered in Stenberg *et al.*[23] and Shepherd *et al.*[40] studies, as well as the Mediterranean climate. The sampling concerned several crop residues as in Stenberg *et al.*[23] and parts of trees, shrubs, herbs and litters as in Shepherd *et al*.,[40] but our Mediterranean samples largely differed from those reported in Gillon *et al.*[19] and Gillon and David.[20] Indeed, the agri-food residues from the Mediterranean regions used in the organic fertiliser industry comprised parts of fruits (i.e. more or less processed and extracted) with pips (for example, grape) or kernels (for example, olive). In addition, some of the agri-food residues from the tropical regions used in the organic fertiliser industry comprised parts of beans (for example, coffee) and bean shells (for example, cocoa). Since they were parts of storage organs, they contained high amounts of lignin in order to protect the seed (particularly for olive and grape).

Our samples (i) were from several geographical and climatic origins, (ii) had a large botanical diversity, (iii) originated from different plant parts, (iv) have been submitted to different industrial processes for several of them, (v) could be used in the fertiliser industry or directly as OM input in soil and, in addition, (vi) had wider ranges in chemical composition than previously published NIR calibration studies. With all these factors combined, this conferred a particular originality to this database.

Spectral treatments

A preliminary study (data not shown) was made upon the selection of the spectral segment, the mathematical treatment and scatter correction of the data. Although spectra were collected between 400–2498 nm, the 400–1100 nm region was discarded as it introduced instability in models (smaller *SEC* but higher *SECV* values). This may be due to a wide colour range of the raw materials (for example, from light yellow to dark brown) leading to a confusion between colour and chemical composition (Figure 1). This was also observed for mushroom compost.[47] In total, eight pre-treatments were tested (four scatter corrections and two levels of derivation) in order to find the pre-treatment providing the best calibration. For example, the *SECV* varied within the following ranges: 3.04–3.70, 3.10–3.51, 3.23–3.66, 3.79–4.47, 1.24–1.45 and 0.16–0.26 for NDSol, Hem, Cel, Lig, Ash and NKj, respectively (detailed data not shown). Dardenne *et al.*[48] have shown that different regression methods gave almost equal standard errors of prediction and the maximum effort should be devoted to obtaining numerous and accurate reference data. Since the treatments tested here did not give a wide variation among the *SECV*, the models were thus considered as relatively robust. The SNVD 2,5,5 pre-treatment—which induced slightly better performance of calibration models—was chosen.

Calibrations and cross-validation statistics

Calibration and cross-validation statistics for chemical and biochemical parameters of the plant samples ($n = 113$), using second derivative and SNVD, are presented in Table 2 and NIR-predicted versus laboratory-measured parameters are presented in Figures 2–5.

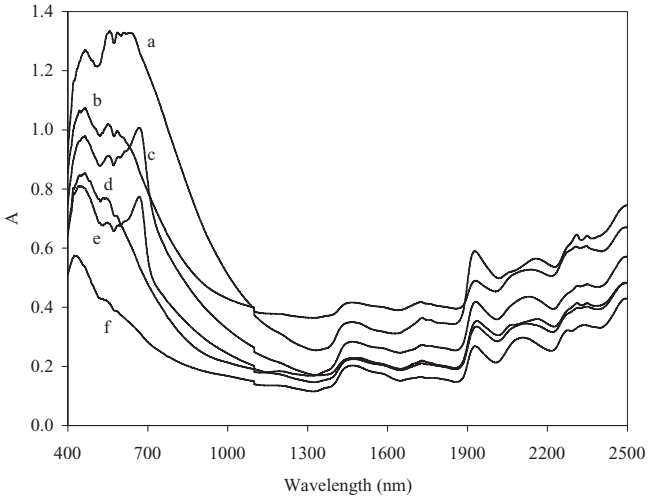

Figure 1. Absorbance spectra of some contrasting materials among the total plant samples set (a = grape skins; b = olive pulps; c = Mucuna leaves; d = cocoa shells; e = Crotalaria leaves and stems; f = maize leaves and stover).

Total organic matter characteristics: ash, C Dumas, N Dumas and N Kjeldahl

The calibration equations for chemical composition of OM (Ash, C Dumas, N Dumas, N Kjeldahl) are presented in Table 2 and Figures 2(a)–(d).

The relatively low variability in ash content (lack of samples under 5 g 100 g^{-1} and/or samples above 10 g 100 g^{-1}, Figure 2(a)] lead to a low *RPD* value, similar to those reported for silage (0.83 and 2.4).[13] In this latter reference, most samples were litters or roots, thus more or less contaminated by soil. Also, problems of homogeneity due to a possible segregation of particles may have occurred when pouring the sample into the NIR sample cells. Being heavier than the organic

Table 2. Calibration and cross-validation statistics for chemical and biochemical characteristics (g 100 g^{-1} DM) of the plant samples ($n = 113$, without outliers) using second derivative and SNVD. See text and Table 1 for parameter abbreviations.

(g 100 g^{-1} DM)	*n*	Mean	*SD*	*SEC*	R^2	*SECV*	*RPDcv*
Ash	109	7.4	2.9	1.01	0.88	1.31	2.2
C Dumas	109	46.2	1.5	0.50	0.89	0.71	2.1
N Dumas	106	2.0	0.8	0.15	0.97	0.19	4.4
N Kjeldahl	58	2.3	0.6	0.15	0.93	0.22	2.6
NDSol	110	26.3	10.8	2.53	0.95	3.43	3.2
Hem	109	13.8	8.5	2.42	0.92	3.20	2.7
Cel	110	25.1	12.8	2.47	0.96	3.16	4.1
Lig	111	27.0	17.0	3.01	0.97	3.57	4.8

matter particles, the soil particles may possibly have covered the glass window of the cell, leading to misinterpretations of not so representative spectra. However, the *SECV* value for Ash could be considered as satisfactory as it was less here than half of that reported for dairy manure (3.36 g 100 g^{-1}),[31] similar to that of mushroom composts (0.99 g 100 g^{-1})[47] and twice as high as that of silage (0.58 g 100 g^{-1}).[13]

Similarly, the coefficient of variation of C Dumas was less than 5% which is very low. This explains the modest performance of the C Dumas model in term of R^2 and *RPD* [Table 2 and Figure 2(b)]. Therefore, improving the performance of this model should be possible by analysing more samples (set size extension and range extension).

Nitrogen measured on dry samples (N Dumas) lead to a better equation, with *SEC* as low as 0.15%. Equations developed for nitrogen measured on fresh samples (N Kjeldahl) involved less samples than N Dumas ($n = 58$, Table 2). This is particularly clear when comparing Figure 2(c) with Figure 2(d). The *SEC* value of NKj obtained in this study was about twice that reported for Mediterranean leaves and litters (0.08 g 100 g^{-1})[20] and compost (0.07 and 0.03 g 100 g^{-1}),[28,47] but was lower than that reported for other heterogeneous materials such as bovine manure (0.74 g 100 g^{-1}).[33] The *SECV* of NKj can be compared with other studies: dairy manures (4.42, 0.23 and 0.05 g 100 g^{-1}),[30–31,49] or compost (0.08 g 100 g^{-}1 DM and 0.03 g 100 g^{-1} bulk weight).[28,47]

RPD helps in comparing the calibrations from different studies (different population abundance and material homogeneity). Values of $RPD > 3$ and $R^2 > 0.90$ are considered to indicate successful calibrations for variable databases.[32,50] *RPDcv* obtained on dry samples (N Dumas) was high and similar to those reported for Northern European crop residues (4.4 to 5.0).[22,23] *RPDcv* for NKj was relatively low (Table 2) in comparison with other transformed OMs which, in general, exceed three (for example, References 27 and 28), Future improvement seems possible; for example, consider expanding the database with more samples with < 2 g 100 g^{-1} NKj content. Comparing the N measured on dried (N Dumas*)* and raw (N Kjeldahl*)* samples allowed their different meanings to be explored. NKj is, in principle, closer to the "real" N content in raw samples. However, since the spectra are measured after drying and grinding, they are understandably more distant from NKj than from N Dumas, analysed on these dried samples. Moreover, NKj was analysed on raw samples which are heterogeneous and particularly wet—which sometimes interferes with dry matter estimation. This could have some repercussions on the final result for NKj expressed on a dry weight basis. Furthermore, some losses of inorganic N may have occurred during the course of drying, even using a

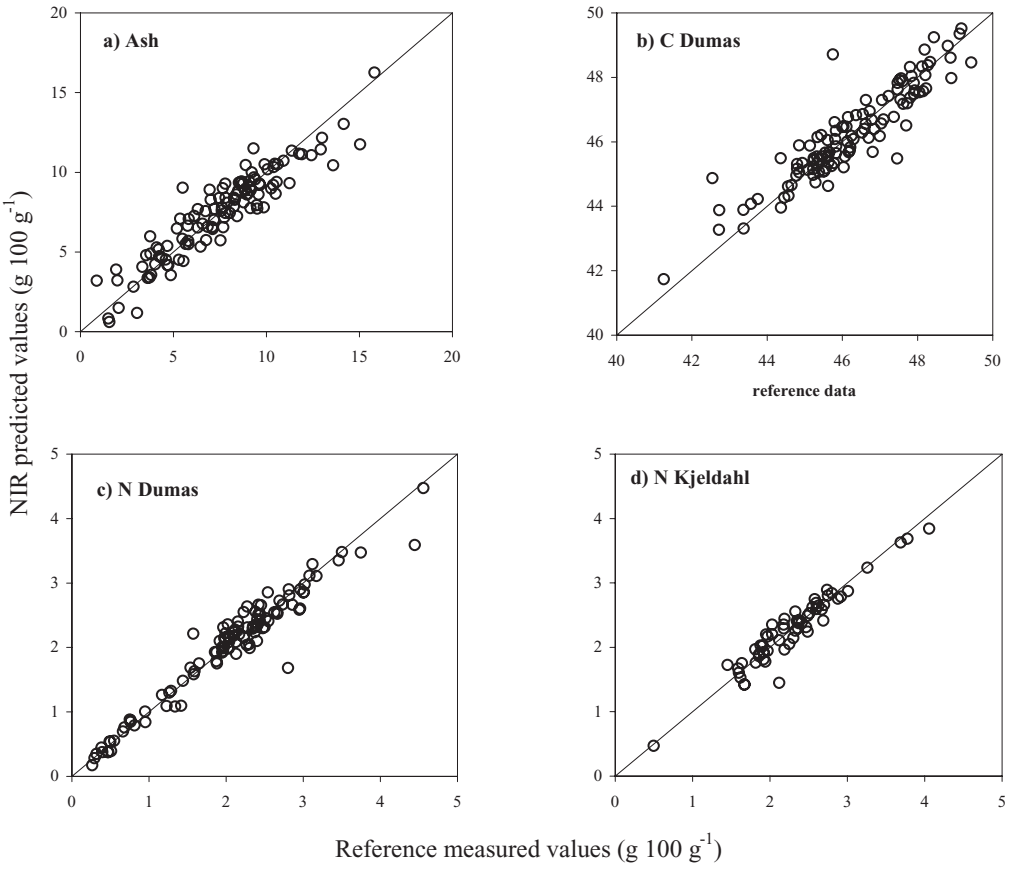

Figure 2. NIR predicted versus measured values (g 100 g^{-1} DM) for the whole dataset ($n = 113$): (a) ash, (b) C Dumas, (c) N Dumas and (d) N Kjeldahl.

limited temperature (40°C). All in all, one can consider that NKj calibrations, although appearing less precise than N Dumas, are more useful to describe raw samples.

It must be noted that the fact that one to three samples are outliers for Ash, C Dumas, N Dumas or NKj during the calibration process (Figure 2) should lead to a re-run of the analysis, which was not possible in the present study.

Fibre characteristics: NDSoluble, hemicelluloses, cellulose and lignin

All biochemical parameters were quite well adjusted as all R^2 were above 0.9 [Table 2 and Figures 3(a)–(d)] and had less than ten calibration outliers (student *t* test, $p<0.05$). *RPDcv* were particularly good for Cel and Lig (>4) but less satisfactory for Hem. The highest differences between *SEC* and *SECV* were observed for NDSol and Hem, due to outliers. It is interesting to note that the *SECV* values were almost identical for all the variables (Table 2), which shows that the uncertainty (combination of the analytical error and the ability of NIR to account for the component) is the same for all nature of fibre fractions.

The NDSol content was fairly well adjusted ($R^2=0.95$; *RPDcv*>3). The *SEC* associated with NDSol represented 9.7% of the mean and was lower than the one reported for temperate plant materials (3.82 g 100 g^{-1}).[23] Even if NDSol is obtained by subtracting the organic part of NDF to the total amount of organic matter, the differences in *SEC* seemed quite normal as our set had NDSol mean and *SD* values of about half those reported in that reference (mean NDSol = 100–51.5 = 48.5 g 100 g^{-1}).[23] The same type of remark can be made for published NDF calibrations for bovine manure and mushroom compost. For these materials, the NDSol mean values (100–48.6 = 51.4 g 100 g^{-1}) were about twice our values, whereas the *SD* associated with NDF were much lower than ours. These figures explain why the *SEC* associated with our NDSol was almost twice that reported for bovine manure (1.35 g 100 g^{-1})[33] or mushroom compost (1.33 g 100 g^{-1}).[47] In our study, *SECV* for NDSol was similar to that of temperate plant materials (3.82 g 100 g^{-1})[23] or dairy manure (2.8 g 100 g^{-1}),[31] but higher than that of mushroom compost (1.14 g 100 g^{-1})[47] and (0.57 g 100 g^{-1} b.w.).[28] *RPDcv* for NDSol in our study was close to that calculated for NDF of mushroom compost (3.07).[28] The materials considered in this study were intermediate in both terms of nature/origin (tropical and Mediterranean plants more or less transformed) and NDF content between the temperate plant materials from Northern Europe and composts and manures.

The model developed for Hem was satisfactory ($R^2=0.92$) though having the lowest *RPDcv* (<3, Table 2) among the other fibre parameters. *SEC* associated to Hem was twice as high as

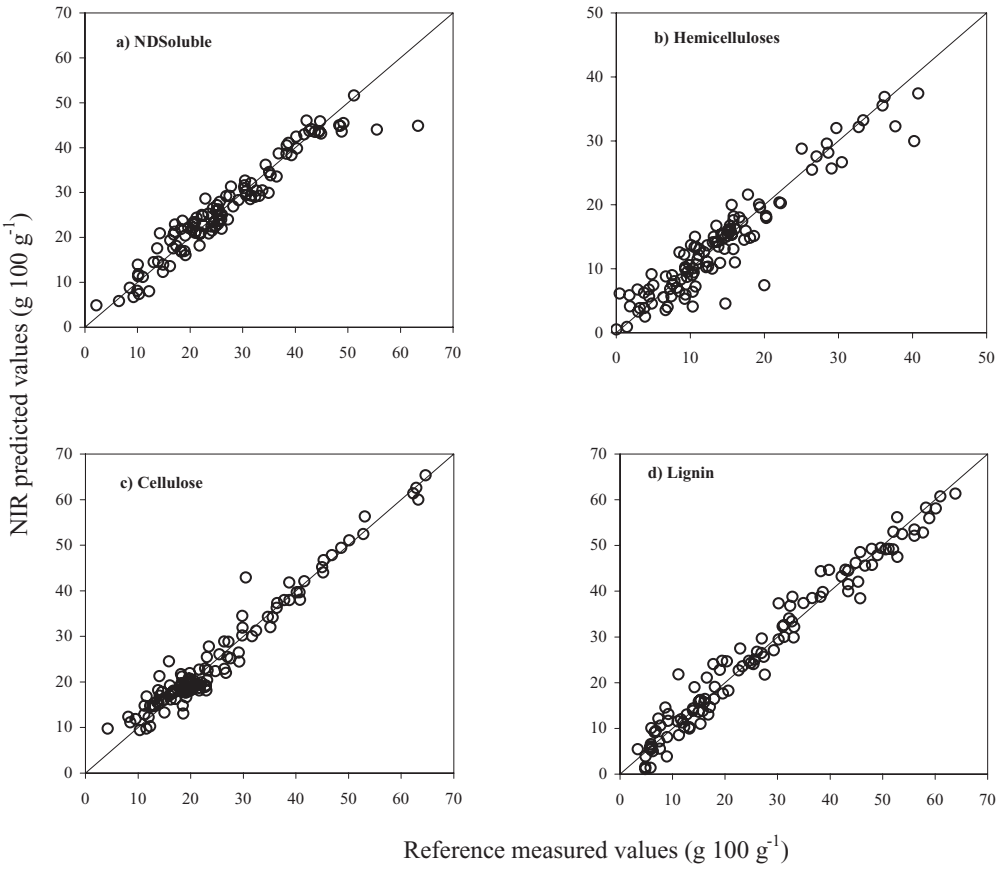

Figure 3. NIR predicted versus measured values (g 100 g^{-1} DM) for the whole dataset ($n=113$): (a) NDSoluble, (b) hemicelluloses, (c) cellulose and (d) lignin.

that reported for mushroom compost (1.01 g 100 g^{-1}).[47] *SECV* for Hem was similar to that reported for dairy manures (2.83 g 100 g^{-1})[31] but higher than that of mushroom compost (1.1 g 100 g^{-1}).[47] Among the fibres measured in compost,[47] hemicelluloses also gave the lowest R^2. This could be explained partly by the reference measure of this fraction at the laboratory; it is sometimes very sensitive to small experimental errors (such as extraction time, filtration procedure or ash estimation). Furthermore, hemicelluloses are complex polysaccharides made of different types of monomers, whose nature and occurrence can vary among species.[51]

The Cel content was particularly well adjusted ($R^2 = 0.96$; $RPDcv > 4$). Despite the much higher variability of our database, *SEC* associated with Cel was less than twice as high as that reported for holm oak leaves and litters (1.7 g 100 g^{-1})[20] and mushroom compost (1.36 g 100 g^{-1})[47] but much higher than that of silage (0.46 g 100 g^{-1}).[13] Similar mean values were reported for mushroom compost[47] and silage[13] and *SECV* were also lower (1.45 and 0.58 g 100 g^{-1}) than in the present study—the ranges being much smaller in compost and silage. Although from a heterogeneous dataset in terms of nature of products and origins, the *RPDcv* for Cel in our study was higher than that of a dataset comprising several mixtures of legumes and grasses (2.1; García-Ciudad *et al*.[52]) and also higher than that of a homogeneous dataset of silage (3.1).[13] The *RPD* for Cel of our dataset was comparable with that of Cel C of plant materials from Northern Europe (around 4, Table 2 and Stenberg *et al*.[23]) as was the range of Cel content (7–58 g 100 g^{-1}).[46] These particularly good performances were probably partly due to the nature of the cellulose which is a polymer of a unique type of sugar, whereas the other fibres are made of several types of monomers. Another hypothesis is the large range of values in our study, as in that of Stenberg *et al*.[23]

The model developed for the Lig content was very satisfactory ($R^2 = 0.97$; $RPDcv > 4$). Lig was, therefore, the best adjusted parameter. *SEC* associated with Lig was close to that reported for holm oak leaves and litters (2.81 g 100 g^{-1}).[20] *SECV* for Lig in our study was higher than the one reported for a heterogeneous set of plants from sub-alpine meadows and shrublands (1.37 g 100 g^{-1})[53] or more transformed materials such as dairy manures (1.11 g 100 g^{-1})[31] and mushroom compost (0.76 g 100 g^{-1}).[47] In comparison with less transformed plant materials, our model performance was close to that of grasses ($RPD = 3.3$; García-Ciudad *et al*.[52]) and was way above that of crops from Northern Europe ($RPD = 1.6$ for Lig C; Stenberg *et al*.[23]). As for Cel, the different performances could be due to the ranges. Indeed, it is particularly wide in our study (3.4–64 g 100 g^{-1}; Table 1) compared with that of crops from Northern Europe (0.1–23 g 100 g^{-1}).[46]

The samples which were located far away from the median in Figure 3 were underestimated for NDSol (two samples: *S. sesban* leaves and cocoa), as for Hem (maize stalks, *G. sepium* leaves and grape skin), whereas they were overestimated for Cel (*P. tiphoides* and cocoa) and Lig (maize roots). Some samples were mispredicted for several parameters and since NDSol, Hem, Cel and Lig values are calculated by difference, this is not surprising. The highest discrepancy between the reference and the predicted content was observed for NDSol and Hem. This is not surprising, as NDSol and Hem were obtained during the first stages of the sequential extraction. These early stages are the most susceptible to misleading results, due to filtration and rinsing difficulties, especially for AOM with high starch or protein content which can interfere with fibre analysis. This was possibly the case for two samples, rape seed and cocoa as seeds.

Validation

Validation with samples not included in the calibration set is an essential step in NIR calibration building process. However, in relatively small datasets such as the one used in this study, the number of validation samples is necessarily limited. Cross-validation, as used in former sections, is therefore complementary as it uses successively all samples in calibration and validation processes. In this work, devoted to real applications in practice, it was also decided to perform an "extrapolation validation", using samples identified as outliers in the database, in order to test the ability of our calibrations to estimate the characteristics of non-typical raw materials as those we sometimes have to cope with in real situations.

The Validation (VAL) sub-set (selection of 20 samples at random among the 113 samples) gave a homogeneously-formed set with samples covering all natures (grape skins, grape pips, cocoa, olive pulps and tropical plants). The major characteristics of VAL are presented in Table 3. The means and *SD* of this sub-set are quite similar to those of the whole set of plant samples. The VAL sub-set thus encompasses the variability of the whole set of plant samples.

New equations were computed on the CAL sub-set ($n = 93$), then they were used to predict the VAL sub-set ($n = 20$) parameters and to calculate the *SEP*. Calibration and cross-validation statistics (on CAL), bias-corrected *SEP* (on VAL) for each parameter, using the same mathematical treatment and calibration procedures, are presented in Table 4. As expected, the *SEC, SECV* and *RPDcv* from CAL were extremely close to those of the whole set of plant samples. Indeed, these sets only differed by 20 samples randomly chosen. This shows that the calibrations were relatively robust despite the limited number of samples. *RPDcv* varied from 2.4 for NKj to 4.8 for N Dumas. As hypothesised above, the relatively poor performance of the Hem or NKj models could be due to (i) some problems during the reference analyses, (ii) a lack in certain values (a way of improvement can be to complete a range, for example with values < 1.5 g 100 g^{-1} for NKj) or (iii) a lack of size extension of the set (for example, $n = 46$ for NKj compared with $n = 92$ for Lig).

On most parameters, the bias-corrected *SEPc* of VAL were close to the *SECV* of CAL, especially for NKj, N Dumas and Cel, indicating that these models were not exaggeratedly influenced by a few samples. However, *SEPc* were

Table 3. Chemical and biochemical characteristics (in g 100 g^{-1} DM) of the sub-sets of plant samples (CAL, n=93; VAL, n=20). See text and Table 1 for parameter abbreviations.

	Calibration samples (n=93)				Validation samples (n=20)			
(g 100 g^{-1} DM)	Mean	*SD*	Min.	Max.	Mean	*SD*	Min.	Max.
Ash	7.5	3.1	0.9	15.8	7.5	2.5	1.5	11.2
C Dumas	46.2	1.6	41.3	49.4	46.2	1.3	44.4	49.1
N Dumas	2.03	0.89	0.26	4.56	1.99	0.79	0.50	3.17
N Kjeldahl	2.29	0.62	0.50	4.06	2.31	0.42	1.59	2.88
NDSol	27.1	11.5	8.5	63.4	25.0	12.2	2.2	44.5
Hem	14.2	8.9	0.04	40.8	13.2	8.0	1.47	32.8
Cel	24.9	12.8	4.2	64.7	25.4	13.1	11.2	62.3
Lig	26.3	17.1	4.8	61.0	28.8	17.1	3.4	63.9

Table 4. Calibration and cross-validation statistics (CAL sub-set), bias-corrected standard error of prediction (VAL sub-set) and associated RPD for chemical and biochemical characteristics (in g 100 g^{-1} DM) of the plant samples using second derivative and SNVD. See text and Table 1 for parameter abbreviations.

(g 100 g^{-1} DM)	n	Mean	*SD*	*SEC*	R^2	*SECV*	*RPDcv*	*SEPc*	*RPDp*
Ash	91	7.6	3.1	0.80	0.93	1.24	2.5	2.14	1.2
C Dumas	88	46.3	1.5	0.36	0.94	0.54	2.8	1.13	1.1
N Dumas	84	2.00	0.87	0.11	0.98	0.18	4.8	0.23	3.4
N Kjeldahl	46	2.23	0.56	0.20	0.87	0.23	2.4	0.23	1.8
NDSol	89	26.7	10.6	2.42	0.95	3.37	3.1	3.02	4.1
Hem	89	13.3	8.0	2.43	0.91	3.16	2.5	3.86	2.1
Cel	92	24.8	12.8	2.31	0.97	3.16	4.1	2.92	4.5
Lig	90	26.4	17.1	3.17	0.97	3.95	4.3	2.80	6.1

1.7 to 2.1 times higher than *SECV* for Ash and C Dumas, respectively. The *SEPc* of NDSol in this study was higher than that reported for bovine manure (0.82 g 100 g^{-1}).[33] The *SEPc* of NKj in our study was twice that of bovine manure (0.1 g 100 g^{-1}),[33] swine excreta compost (0.09 g 100 g^{-1})[25] and compost (0.08 g 100 g^{-1}),[29] all kinds of material being unique in nature. This tends to indicate that the variability of our materials is a little too large to expect very good prediction capacities. Nevertheless, our *SEPc* for NKj was less than half the *SEPc* reported for various Northern European crop residues (0.62 g 100 g^{-1}).[22]

RPDp (= *SD* of VaL/*SEPc*) ranged from 2.1 for Hem to 6.1 for Lig. The predictive power of the equations developed here was, therefore, relatively good for the fibre parameters. Nevertheless, *RPDp* were not satisfactory for Ash and C Dumas, as was reported for organic matter in swine excreta compost (1.46).[25]

The "extrapolation validation" was performed on the 20 outliers (OUT sub-set) originally discarded from the database (spectral or compositional outliers) with the equations elaborated on CAL. Furthermore, attemps were also made to predict the composition of the 13 coffee by-products samples being discarded from the beginning of the calibration procedure. These results are presented in Figures 4 and 5.

Among the VAL sub-set, the fibre parameters were reasonably well predicted for most samples (Hem was the poorest). For the chemical parameters, one can point out the relatively good predictive capacity for the N Dumas equation, whereas in the case of Ash and C Dumas it was not as good.

Figure 6 presents a comparison between the standard errors of the reference laboratory data (laboratory repeatability, *SEref*) and the *SEP(c)*. The *SEref* for Ash and NKj were estimated on 12500 samples analysed in duplicate in the organic fertilier company's certified ISO9001-2000 control laboratory. The *SEref* for fibres were estimated on 350 samples analysed in duplicate for Lig and Cel, in triplicate for Hem and four replicates for NDSol. For Ash, the *SEP* was twice the *SEref*; for NKj, the ratio *SEP/SEref* was

3.7. This last figure cannot be considered as being a very high error, as it only represents a coefficient of variation of 2.7% for NKj. The *SEP* values were under the normative tolerances for organic amendments (NFU#44051)[54] (max 3.0 g 100 g^{-1} bulk weight for Ash and min–max 0.2–0.3 g 100 g^{-1} b.w. for NKj). For the fibre parameters, the *SEP* was less than twice the *SEref*. This difference was acceptable, as this has helped meet a major objective of this work, which is an estimation of the biochemical fractions of very variable plant materials. For example, even estimated with a possible error of 5 g 100 g^{-1} DM, two plant materials with predicted Lig contents of 45 and 10 g 100 g^{-1} DM would obviously be candidates for entering a composting process for the first and entering an organic fertiliser composition or a direct use on field for the latter. Considering the time-consuming and costly procedure required for the reference analyses of fibre, the *SEP* are quite encouraging.

The predictions of most parameters were acceptable for a majority of the samples from the OUT sub-set [Figures (4) and (5)]. Although discarded before the calibration procedure, these samples generally entered the relationship predicted/measured. This can illustrate the rather good extrapolation capacity of the models developed here. Furthermore, a hypothesis for the bad prediction could be found for some of the outliers among the OUT sub-set. Indeed, the cocoa outlier for NDSol was also an outlier for Cel and its reference data differed from the mean reference data for this kind of material (for example, mean NDSol > 40 and mean Cel < 20 g 100 g^{-1} for the other cocoas). The poor prediction for Lig in *Eleusine* roots may be due to clay contamination. The other outliers were less explainable. Nevertheless, along with the first level of validation, this attempted extrapolation was rather successful.

The coffee by-product samples were poorly predicted for several parameters [Figures (4) and (5)]. This was particularly true for Ash and NKj [Figures 4(a) and (d)]. These poor predictions could be attributed to the peculiar nature of the coffee cakes, since they were heavily processed. Since coffee cakes generally contain high amounts of lipids (some measured in this study were above 21 g 100 g^{-1} DM, data not shown), this fraction could have caused special problems such as spectral misinterpretation. It could also be hypothesised that the reference data were possibly questionable. Indeed, coffee generally has very high amounts of organic matter (> 98%), thus very low Ash content, particularly susceptible to experimental errors. The major exception was for NDSol and Cel [Figures 5(a) and (c)], where the coffee by-products were rather well located upon the predicted/meas-

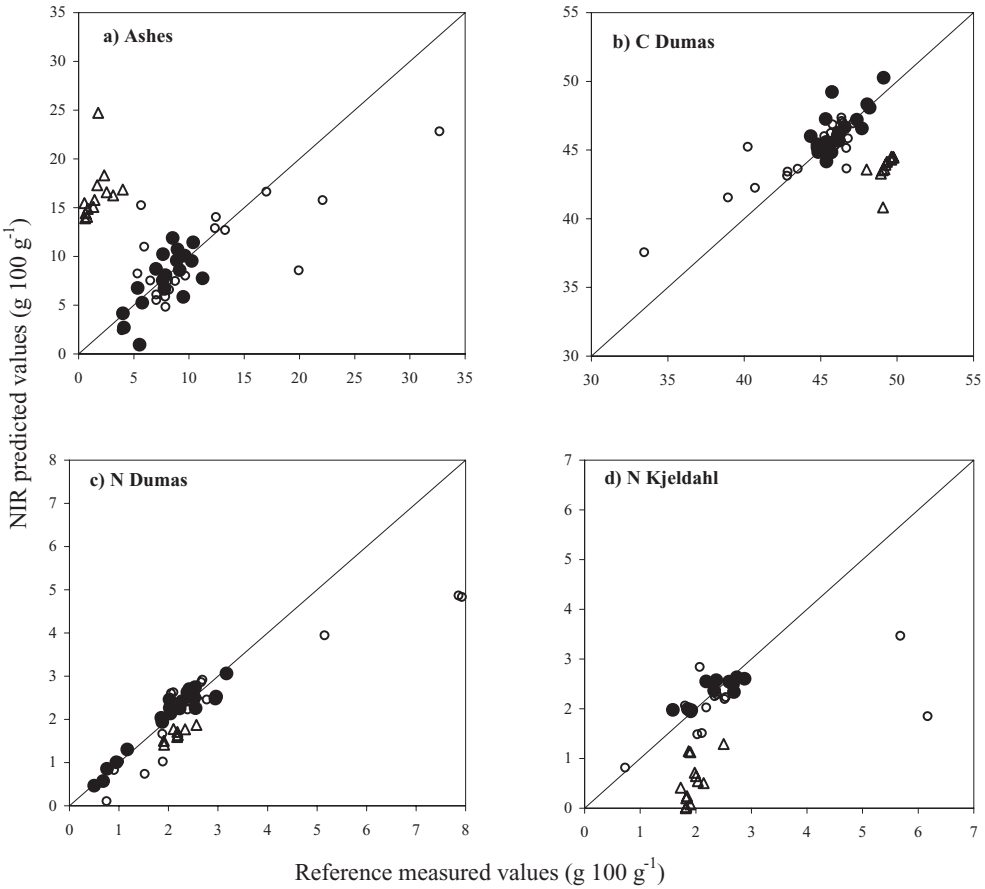

Figure 4. NIR predicted versus measured values (g 100 g^{-1} DM) for the VAL (dots) and EXT (circles) sub-sets, coffee samples (triangles), (a) ash, (b) C Dumas, (c) N Dumas and (d) N Kjeldahl.

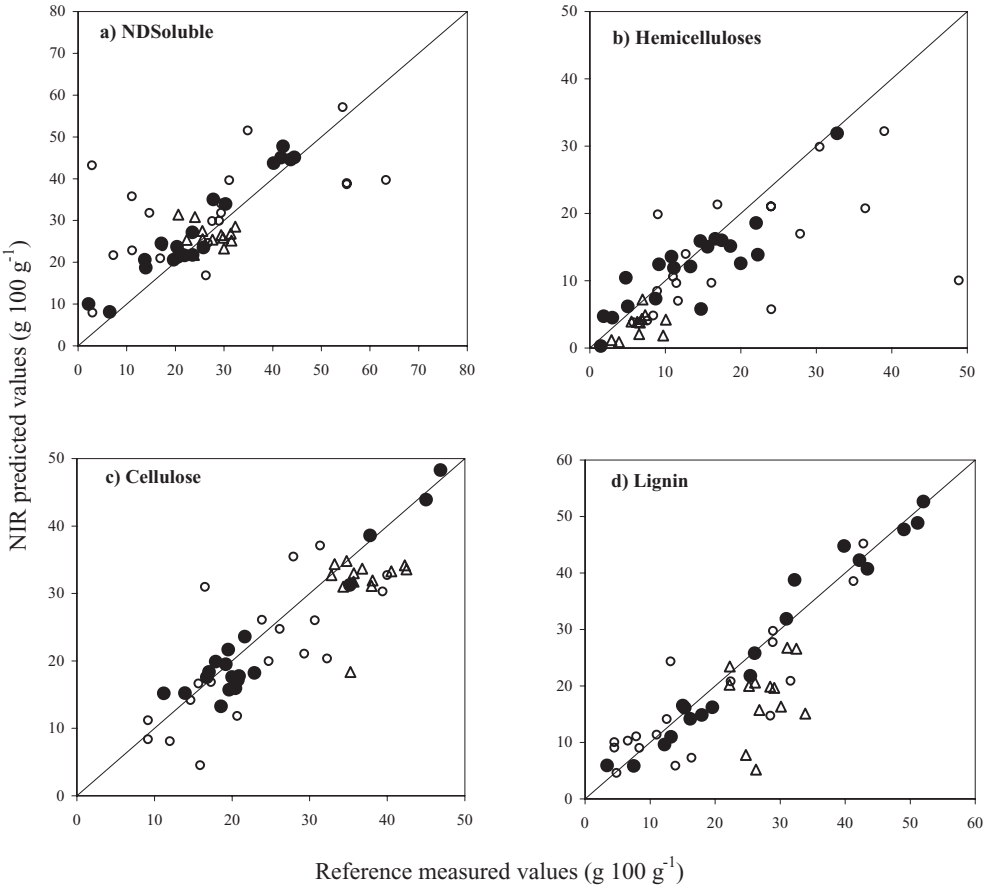

Figure 5. NIR predicted versus measured values (g 100 g^{-1} DM) for the VAL (dots) and EXT (circles) sub-sets, coffee samples (triangles), (a) NDSoluble, (b) hemicelluloses, (c) cellulose and (d) lignin.

ured median. Except for this aspect, coffees were always isolated from the other samples and consisted of distinct sets [for example, Figures 4(a), (b) and (d)]. There is a real qualitative gap between coffee by-products and the rest of the database. More work is needed to predict the biochemical characteristics of coffee cakes; (i) multiply the reference data and attempt new calibrations including the coffee samples among the other plant samples or (ii) develop separate calibrations dedicated to coffee by-products.

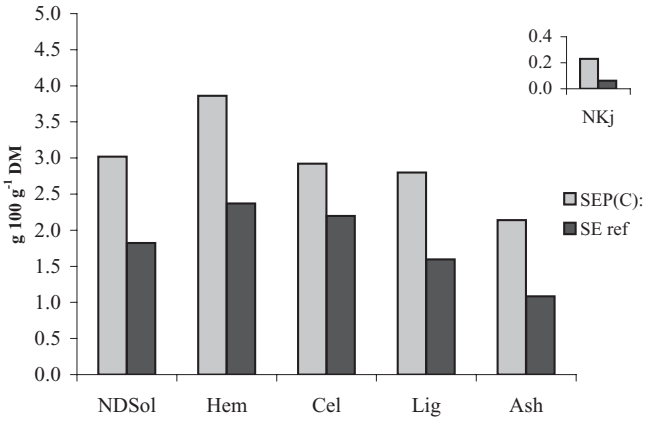

Figure 6. *SEref*, standard errors of reference data and *SEP*(*C*), standard errors of prediction (bias corrected), in g 100 g^{-1} DM.

Conclusion

This study showed that it was possible to develop useful calibrations for the major biochemical characteristics of a very wide range of plant materials. The same overall models were suitable for such variable as raw materials utilised in the organic fertiliser industry (grape skins, olive pulps, cocoa shells…) or plants used in improved agricultural systems, namely direct drilling and agroforestry (*Tephrosia*, *Gliricidia*…). The repartition of the chemical and biochemical parameters was continuous, hence models are generally suitable for all kinds of plant materials and are not just suitable for a single type of material. Coffee cakes were an exception; no satisfactory model was found and more effort is needed to build a particular calibration. The models developed in the general database were particularly promising for the prediction of cellulose and lignin content. The

associated *SECV* and *SEP* were acceptable for the fibre parameters, compared with the standard errors of reference data. Furthermore, considering the costly and time-consuming procedure required to obtain the corresponding reference data, their NIR estimations seem sufficiently precise and very profitable for our purpose of raw materials quality control. For ash, C Dumas and Kjeldahl nitrogen content, the model performances were not so good, although their associated *SECV* and *SEP* were far less than the normal tolerances. Several levels of validation and extrapolation were tested. Predictions on a validation set ($n=20$ samples randomly chosen among 113) confirmed the quality of cellulose and lignin models. An originality of this study was to attempt to predict the parameters of the spectral and compositional outliers (removed before calibration), with the equations developed during calibration, as a test for extrapolation capacity of equations. Surprisingly, a majority of these outliers were acceptably predicted for most of the parameters, although they were known to be very different from the samples present in the calibration sub-set. Models developed here can thus be considered as robust. Future work on our larger database (>2400 samples) are expected to improve the predictions and diminish the *SECV* and *SEP* to a comparable level to that of reference data and make them of possible direct use in the TAO model.[8,9,11,12] However, the prediction accuracy is already largely sufficient for usage recommendations in the factory or on field.

Acknowledgements

The authors would like to thank Dr S. Sena-Alfaia (INPA) for providing the Brazilian plant samples (FFEM project "Brésil"), Dr A. Albrecht (IRD) for providing the Kenyan plant samples (INCO/DEV project "IMPALA"), the Phalippou-Frayssinet SAS company for providing the raw materials used for organic fertilisers, J.L. Marger (CIRAD), S. Carayon, N. Guibaud and P. Viste (Control Laboratory, Phalippou-Frayssinet SAS) for technical assistance and D. Beunard, L. Cayeux, E. Ginestet, T. Mazoyer and F. Petit who took part in this project during their BSc or MSc studies.

References

1. A. Albrecht and S.T. Kandji, *Agric. Ecos. Envir.* **99,** 15 (2003).
2. E. Wollny, *La décomposition des matières organiques et les formes d'humus dans leurs rapports avec l'agriculture*. Berger-Levrault, Paris, France (1902).
3. M.J. Swift, O.W. Heal and J.M. Anderson, *Decomposition in terrestrial ecosystems*. Blackwell Scientific Publications, Oxford, UK (1979).
4. G. Cadish and K.E. Giller, *Driven by nature. Plant litter quality and decomposition.* CAB International, Wallingford, UK (1997).
5. P.J. Van Soest, *J. Assoc. Off. Agr. Chem.* **46,** 829 (1963).
6. T.M. Henriksen and T.A. Breland, *Soil Biol. Biochem.* **31,** 1135 (1999).
7. I. Trinsoutrot, S. Recous, B. Bentz, M. Linères, D. Chèneby and B. Nicolardot, *Soil Sci. Soc. Am. J.* **64,** 918 (2000).
8. L. Thuriès, M. Pansu, C. Feller, P. Herrmann and J.-C. Rémy, *Soil Biol. Biochem.* **33,** 997 (2001).
9. M. Pansu, L. Thuriès, M.-C. Larré-Larrouy and P. Bottner, *Soil Biol. Biochem.* **35,** 353 (2003).
10. B. Nicolardot, S. Recous and B. Mary, *Plant Soil* **83,** 83 (2001).
11. L. Thuriès, M. Pansu, M.-C. Larré-Larrouy and C. Feller, *Soil Biol. Biochem.* **34,** 239 (2002).
12. M. Pansu and L. Thuriès, *Soil Biol. Biochem.* **35,** 37 (2003).
13. G. Sinnaeve, P. Dardenne, R. Agneessens and R. Biston, *J. Near Infrared Spectrosc.* **2,** 79 (1994).
14. J.B. Reeves, III, *J. Near Infrared Spectrosc.* **2,** 49 (1994).
15. B. De La Roza, A. Martinez, B. Santos, J. Gonzalez and G. Gomez, *J. Near Infrared Spectrosc.* **6,** 145 (1998).
16. M. Hellämäki, M. Rinne, P. Huhtanen and J. Nousiainen, in *9th International Conference on Near Infrared Spectroscopy,* Verona, Italy, Ed by A.M.C. Davies and R. Giangiacomo. NIR publications, Chichester, UK, p. 741 (1999).
17. A. Chesson in *Driven by nature, Plant litter quality and decomposition*, Ed by G. Cadish and K.E. Giller. CAB International, Wallingford, UK, p. 47 (1997).
18. T.M. McLellan, J.D. Aber, M.E. Martin, J.M. Melillo and K.J. Nadelhoffer, *Can. J. Forest Res.* **21,** 1684 (1991).
19. D. Gillon, R. Joffre and A. Ibrahima, *Ecology* **80,** 175 (1999).
20. D. Gillon and J.F. David, *Soil Biol. Biochem.* **33,** 2159 (2001).
21. R. Joffre, G.I. Agren, D. Gillon and E. Bosatta, *Oikos* **93,** 451 (2001).
22. L.S. Jensen, A. Pedersen, B. Stenberg, C. Lundström, T.A. Breland, T.M. Henriksen, F. Palmason, J. Gudmunson and T. Salo, in *17th World Congress of Soil Science. Soil science: confronting new realities in the 21st century*, Ed by ISSS. ISSS, Bangkok, Thailand, p. 758 (2002).
23. B. Stenberg, L.S. Jensen, E. Nordkvist, T.A. Breland, A. Pedersen, J. Gudmundsson, S. Bruun, T. Salo, F. Palmason, T.M. Henriksen and A. Korsaeth. *J. Near Infrared Spectrosc.* **12,** 331 (2004).
24. A.S. Tasistro, S. Shaaban, D.E. Kissel and P.F. Vendrell, *Commun. Soil Sci. Plant Anal.* **34,** 1367 (2003).
25. J.J. Nam, K.Y. Jung and S.H. Lee, in *9th International Conference on Near Infrared Spectroscopy,* Verona, Italy, Ed by A.M.C. Davies and R. Giangiacomo. NIR Publications,Chichester, UK, p. 613 (1999).
26. E. Johansson and A. Brundin in International Symposium Composting and Compost Utilization, Colombus, USA (2002).

27. J.F. Collard, R. Agneessens, D. Stilmant and P. Dardenne, in *Proceedings of the 11th International Conference on Near Infrared Spectroscopy,* Cordoba, Spain, Ed by A.M.C. Davies and A. Garrido-Varo. NIR Publications, Chichester, UK, p. 769 (2003).
28. J.W. Resink and M. van den Hurk in *Science and Cultivation of Edible Fungi*, Ed by V. Griensven. Balkema, Rotterdam, The Netherlands, p. 349 (2000).
29. K. Suehara, Y. Nakano and T. Yano, *J. Near Infrared Spectrosc.* **9,** 35 (2001).
30. J.B. Reeves, III and J.A.S. Van Kessel, *J. Near Infrared Spectrosc.* **7,** 195 (1999).
31. J.B. Reeves, III and J.A.S. Van Kessel, *J. Near Infrared Spectrosc.* **10,** 93 (2002).
32. D.F. Malley, L. Yesmin and R.G. Eilers, *Soil Sci. Soc. Am. J.* **66,** 1677 (2002).
33. D. Cozzolino, A. La Manna and D. Vaz Martins, *J. Near Infrared Spectrosc.* **10,** 309 (2002).
34. G. Fystro, *Plant Soil* **246,** 139 (2002).
35. M.M. Coûteaux, B. Berg and P. Rovira, *Soil Biol. Biochem.* **35,** 1587 (2003).
36. C.A. Russell, *Commun. Soil Sci. Plant Anal.* **34,** 1557 (2003).
37. K.D. Shepherd and M.G. Walsh, *Soil Sci. Soc. Am. J.* **66,** 988 (2002).
38. J.B. Reeves, III, G.W. MacCarty and J.J. Meisinger, *J. Near Infrared Spectrosc.* **8,** 161 (2000).
39. J.B. Reeves, III and G.W. MacCarty, *J. Near Infrared Spectrosc.* **9,** 25 (2001).
40. K.D. Shepherd, C.A. Palm, C.N. Gachengo and B. Vanlauwe, *Agron. J.* **95,** 1314 (2003).
41. C.A. Palm and A.P. Rowland in *Driven by Nature, Plant Litter Quality and Decomposition*, Ed by G. Cadish and K.E. Giller. CAB International, Wallingford, UK, p. 379 (1997).
42. P.J. Van Soest and V.C. Mason, *Anim. Feed Sci. Tech.* **32,** 45 (1991).
43. P.J. Van Soest, J.B. Robertson and B.A. Lewis, *J. Dairy Sci.* **74,** 3583 (1991).
44. R.J. Barnes, M.S. Dhanoa and S.J. Lister, *Appl. Spectrosc.* **43,** 772 (1989).
45. J.S. Shenk and M.O. Westerhaus, *Crop Sci.* **31,** 469 (1991).
46. L. Jensen, T. Salo, F. Palmason, T. Breland, T. Henriksen, B. Stenberg, A. Pedersen, C. Lundström and M. Esala, *Plant Soil* **273,** 307 (2005).
47. H.S.S. Sharma, M. Kilpatrick and L. Burns, *J. Near Infrared Spectrosc.* **8,** 11 (2000).
48. P. Dardenne, G. Sinnaeve and V. Baeten, *J. Near Infrared Spectrosc.* **8,** 229 (2000).
49. J.B. Reeves III and J.A.S. Van Kessel, *J. Near Infrared Spectrosc.* **8,** 151 (2000).
50. P. Williams, *NIR news* **4(6),** 13 (1993).
51. H. Titeux, *Transfert de carbone organique dissous et de métaux dans des sols forestiers acides: influence du fonctionnement de la litière et des reserves minérales du sol*, PhD Thesis. UCL, Louvain la Neuve, Belgium, p. 390. (2005).
52. A. García-Ciudad, B. García-Criado, M.E. Pérez-Corona, B. Vazquez de Aldana and A.M. Ruano-Ramos, *J. Sci. Food Agric.* **63,** 419 (1993).
53. M. Lembke, D. Bastianelli, J.-F. Dobremez, C. Hervouet, L. Bonnal and D. Maillard, in *12th International Conference on Near-infrared Spectroscopy*. Aukland, New Zealand. Abstract p. 155 (2005).
54. NFU#44051 *Amendements organiques*, Ed by AFNOR, 1999 Edition. AFNOR, Paris, France, p. 694. (1999).

Received: 16 February 2005
Revised: 30 September 2005
Accepted: 13 October 2005
Web Publication: 2 December 2005