Bioresource Technology 175 (2015) 382–390

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Biochemical methane potential prediction of plant biomasses: Comparing chemical composition versus near infrared methods and linear versus non-linear models

Bruno Godin [a,*], Frédéric Mayer [b,c], Richard Agneessens [a], Patrick Gerin [c], Pierre Dardenne [a], Philippe Delfosse [b], Jérôme Delcarte [a]

[a] Walloon Agricultural Research Center – CRA-W, Valorisation of Agricultural Products Department – Biomass, Bioproducts and Energy Unit, Chaussée de Namur, 146, B-5030 Gembloux, Belgium
[b] Centre de Recherche Public – Gabriel Lippmann, Environment and Agro-Biotechnologies Department, Rue du Brill, 41, L-4422 Belvaux, Luxembourg
[c] Université catholique de Louvain, Earth & Life Institute – Bioengineering Group, Croix du Sud, 2 Box L7.05.19, B-1348 Louvain-la-Neuve, Belgium

HIGHLIGHTS

- Predictions based on the NIR spectrum were most reliable to estimate the BMP.
- NIR predictions of the BMP made by local models were reliable and quantitative.
- Non-linear models gave more reliable predictions than linear models.
- Biomass presentation form did not influence the model's prediction performances.

ARTICLE INFO

Article history:
Received 4 September 2014
Received in revised form 21 October 2014
Accepted 23 October 2014
Available online 29 October 2014

Keywords:
Anaerobic digestion
Biogas
Multivariate data analysis
Chemometrics
Prediction

ABSTRACT

The reliability of different models to predict the biochemical methane potential (BMP) of various plant biomasses using a multispecies dataset was compared. The most reliable prediction models of the BMP were those based on the near infrared (NIR) spectrum compared to those based on the chemical composition. The NIR predictions of local (specific regression and non-linear) models were able to estimate quantitatively, rapidly, cheaply and easily the BMP. Such a model could be further used for biomethanation plant management and optimization. The predictions of non-linear models were more reliable compared to those of linear models. The presentation form (green-dried, silage-dried and silage-wet form) of biomasses to the NIR spectrometer did not influence the performances of the NIR prediction models. The accuracy of the BMP method should be improved to enhance further the BMP prediction models.

1. Introduction

The production of biogas by anaerobic digestion (biomethanation) of plant biomasses is of growing importance in the context of renewable energy production (Amon et al., 2007; Triolo et al., 2012). The anaerobic digestion process consists in the anaerobic conversion of the organic matter into biogas by microorganisms. The produced biogas is a mixture mainly made of methane and carbon dioxide (Duncan and Nigel, 2003). Plant biomasses such as corn (*Zea mays* L.) and meadow (e.g. *Festuca arundinacea* Schreb.) silages are commonly used feedstocks for biomethanation (Amon et al., 2007). This type of renewable energy production is socio-economically cost-efficient and environmentally efficient (e.g. reduction of greenhouse gas emissions) (Amon et al., 2007; Triolo et al., 2012). It is also a convenient source of renewable energy as it offers the possibility to use multiple feedstocks, and to meet different types of energy needs (heat, electricity, and fuel) and fertilizers for agriculture (Ward et al., 2008; Triolo et al., 2012).

The biochemical methane potential (BMP expressed as m^3 of methane per kg of organic matter) is the most relevant method

Abbreviations: asl, above sea level; BMP, biochemical methane potential; C.-V., cross-validation; CV, coefficient of variation; DM, dry matter; eDOM, enzymatically digestible organic matter; MedRE, median standard residual error of prediction; MLR, multiple linear regression; *n*, number of samples; PLS, partial least square; NIR, near infrared; R^2Med, coefficient of determination of prediction based on median variables; RPDMed, ratio of the median standard deviation of the variable to MedRE; SD, mean standard deviation; SDMed, median standard deviation; SEL, standard error of laboratory; Val., validation; VS, organic matter (volatile solids); VST, Van Soest.

* Corresponding author. Tel.: +32 81 62 71 49; fax: +32 81 61 58 47.
E-mail address: b.godin@cra.wallonie.be (B. Godin).

http://dx.doi.org/10.1016/j.biortech.2014.10.115
0960-8524/

used to determine the biogas production potential of biomasses (Grieder et al., 2011; Raju et al., 2011). It is a batch assay of 30–100 days of the sample's anaerobic biodegradation. This assay reproduces the biomethanation process conditions in a small biogas fermenter and measures the methane production (Grieder et al., 2011; Raju et al., 2011). The main drawbacks of the BMP method are that it is generally time and resource consuming. These are important burdens for industrial plant management and optimization (Grieder et al., 2011; Doublet et al., 2013). Therefore, there is a need to develop simple, fast and reliable models to predict reliably the BMP.

The gas production by anaerobic digestion depends on the feedstock's chemical characteristics such as chemical composition (contents of lignin, cellulose, hemicelluloses, starch, total soluble sugars, proteins and lipids) and anaerobic organic matter digestibility (Grieder et al., 2011, 2012; Triolo et al., 2011). The chemical composition can be determined in a cost-effective way in only a few days (Raju et al., 2011). It can be used to predict the BMP in a faster and cheaper way, as compared to the BMP measurement, if suitable prediction models can be developed. It has been shown for corn that the chemical composition (contents of lignin, hemicelluloses, total soluble sugars and lipids) enables a reliable prediction of the BMP (Rath et al., 2013). Such a model needs to be built with more than one variable to be reliable (multivariate models; Rath et al., 2013).

The anaerobic organic matter digestibility can be assessed by the enzymatically digestible organic matter (eDOM) determined by the De Boever method (De Boever et al., 1986). This relatively simple and fast method can be used to assess the suitability of plant biomasses to be converted by anaerobic digestion (Godin et al., 2013a,b,c). The eDOM can be considered as the minimum level of anaerobic digestibility of the plant biomass. Indeed, the microorganisms of the anaerobic digestion are expected to produce more enzymes in-situ and for a longer period of time compared to the enzyme cocktail used in the analysis. The eDOM can be predicted from the near infrared (NIR) spectrum of the organic matter with suitably developed models. The prediction performances of the eDOM by models based on the NIR spectrum are known to be excellent: coefficient of determination of 0.95 and ratio of the mean standard deviation of the predicted variable to the mean standard residual error of prediction of 4.6 (Decruyenaere et al., 2009).

Prediction models based on the NIR spectrum can also be very useful to rapidly, cheaply and easily predict the BMP of feedstocks. The NIR-based prediction models have been shown to be reliable for the BMP prediction of meadow grasses (Raju et al., 2011), fibrous plant biomasses, (Triolo et al., 2014) and a wide range of organic substrates (Lesteur et al., 2011; Doublet et al., 2013). The prediction models were based on the linear regression of the partial least square (PLS). However, the NIR predictions can also be made by non-linear models. The local (specific regression and non-linear) model is one of those models which are able to improve the prediction performances of a variable based on the NIR spectrum (Shenk et al., 1997). The local model builds a regression for each sample separately by selecting its most similar spectral neighbors in the used dataset. This selection is then used to develop a specific PLS model for the predicted sample (Shenk et al., 1997). While linear models are commonly used to develop prediction models from the NIR spectrum, to our knowledge, non-linear models have not been used to predict the BMP of plant feedstocks.

To have a reliable prediction model of a secondary method (NIR spectrum and chemical composition in the present study) based on the reference method (BMP and eDOM in the present study), it is important to have: (1) an accurate reference method; (2) a large variability for the values of the reference and secondary methods. The use of a multiproduct (multispecies) dataset with similar plant species helps to enlarge this variability (Berzaghi et al., 2000). The presentation form (eg. water content and particle size) of a sample to the NIR spectrometer is known to affect the prediction performances of a NIR prediction model. A higher water content and/or a bigger particle size tend to decrease the reliability of the NIR prediction model because they tend to hide the NIR spectral information (Bertrand and Dufour, 2006).

The aim of this paper is to compare the reliability of the BMP of plant biomasses predicted by models using the chemical composition or the NIR spectrum as predictor and also to use a multiproduct (multispecies) dataset, to test the influence of the state of the biomass when recording NIR spectrum (green-dried, silage-wet or silage-dried), and to assess the influence of the model used to make the prediction (linear models: PLS models and multivariate linear regression models with a linear matrix; non-linear models: local models and multivariate linear regression models with a non-linear matrix).

2. Methods

2.1. Biomass material

Miscanthus giganteus (*Miscanthus x giganteus* J.M. Greef & Deuter ex Hodk. & Renvoize), switchgrass (*Panicum virgatum* L.), spelt straw (*Triticum aestivum* L. ssp. *spelta* (L.) Thell.), fiber sorghum (*Sorghum bicolor* (L.) Moench), tall fescue (*F. arundinacea* Schreb.) with 3 harvests per year, immature rye (*Secale cereale* L.), and fiber corn (*Z. mays* L.) came from randomized block designed crop trials performed in 2008, 2009, 2010 and/or 2011 at Libramont (498 m above sea level (asl); average annual temperature: 8.6 °C; average annual precipitation: 1260 mm; 49°55′N, 05°24′E; Belgium), Gembloux (161 m asl; average annual temperature: 9.8 °C; average annual precipitation: 856 mm; 50°33′N, 04°43′E), Tinlot (255 m asl; average annual temperature: 9.7 °C; average annual precipitation: 871 mm; 50°28′N, 05°23′E; Belgium), Mötsch (330 m asl; average annual temperature: 8.4 °C; average annual precipitation: 675 mm; 49°57′N, 06°33′E; Germany) or Gerbéviller (260 m asl; average annual temperature: 9.9 °C; average annual precipitation: 1022 mm; 48°29′N, 06°31′E; France). Depending on the crop, trials were performed with different harvest periods (details in Table 1), cultivars (details in Table 1) and/or nitrogen fertilization levels (from 0 to 240 kg of nitrogen per hm^2). From plots between 9 and 24 m^2, the whole above ground biomass was harvested at 10 cm from the ground and chopped (particle size 1–2 cm). Details about the investigated plant biomasses are presented in Table 1.

For each biomass analyzed under its green-dried form, two representative subsamples were dried at 60 °C for 72 h in a forced air oven immediately after the harvest. After the drying process, the two subsamples were milled first with a 4 mm screen BOA hammer mill (Waterleau, Herent, Belgium) and then with a 1 mm screen Cyclotec cyclone mill (FOSS, Hillerød, Denmark). For the storage of the samples, airtight bags were used. They were kept at room temperature and were protected from light in a dark box.

For each biomass analyzed under its silage-wet form, one representative sample was packed in a plastic bag under vacuum. This enabled silaging of the sample. The vacuum sealed plastic bag was stored at room temperature for at least 3 weeks before the laboratory analysis. If gas was produced during silaging, the plastic bag was opened and put again under vacuum.

For each biomass analyzed under its silage-dried, one representative sub-sample of its silage-wet form was taken just before the laboratory analysis. It was dried at 70 °C for 48 h in a forced air oven. The dried sub-sample was milled with a 1 mm screen Cyclotec cyclone mill (FOSS, Hillerød, Denmark). It was then stored at room temperature in airtight bags.

Table 1
Details about the investigated plant biomasses.

Plant species		Cultivar	Harvest period	Near infrared spectrum			Chemical composition
				Green-dried form $n = 588$	Silage-dried form $n = 428$	Silage-wet form $n = 465$	Green-dried form $n = 569$
Miscanthus giganteus	*Miscanthus x giganteus* J.M. Greef & Deuter ex Hodk. & Renvoize	Bical	Early autumn	24	23	24	17
Switchgrass	*Panicum virgatum* L.	Cave-in-Rock	Early autumn	16	8	8	16
Spelt straw	*Triticum aestivum* L. ssp. *spelta* (L.) Thell.	Badengold and Cosmos	Late summer	37	36	37	36
Fiber sorghum	*Sorghum bicolor* (L.) Moench	H133	Late winter	8	0	0	8
Fiber sorghum	*Sorghum bicolor* (L.) Moench	H133	Early autumn	32	28	28	29
Tall fescue	*Festuca arundinacea* Schreb.	Kora and Soni	Spring	118	80	90	117
Tall fescue	*Festuca arundinacea* Schreb.	Kora and Soni	Summer	125	112	123	122
Tall fescue	*Festuca arundinacea* Schreb.	Kora and Soni	Autumn	133	88	118	130
Immature rye	*Secale cereale* L.	Protector and Vitalio	Early spring	22	2	2	21
Fiber corn	*Zea mays* L.	LG Azelo and Ronaldinio	Late winter	14	0	0	9
Fiber corn	*Zea mays* L.	LG Azelo and Ronaldinio	Early autumn	59	51	35	64

n, number of samples

2.2. Chemical reagents and analyses

All chemicals regents were of the analytical grade or equivalent.

The main chemical components (lignin, cellulose, hemicelluloses, starch, total soluble sugars, proteins, mineral compounds) of the assessed biomasses were determined. The Van Soest (VST) method was used to determine lignin VST, cellulose VST and hemicelluloses VST (Van Soest and Wine, 1967; Van Soest, 1973; Godin et al., 2011a,b, 2014). The Van Soest method was achieved with a reflux apparatus using crucibles. The neutral detergent extraction step of the Van Soest method was achieved with the use of α-amylase and sodium sulfite. The acid detergent extraction step of the Van Soest method was realized with the first preliminary extraction step by the Van Soest neutral detergent with the use of α-amylase but without sodium sulfite. The concentrated acid extraction of the Van Soest method was accomplished with sulfuric acid. The fiber residues of the Van Soest method were corrected by their mineral compound content. The starch content was determined by the Ewers method (European Union, 2009). The starch content was only measured for the corn biomasses. It was considered as zero for other analyzed biomasses based on preliminary analyses. The total soluble sugars were determined by the Luff-Schoorl method (European Union, 2009). The proteins' content was determined by the Kjeldahl method using 6.25 as a conversion factor of nitrogen to proteins (AOAC, 1990). The mineral content was determined by the use of a muffle furnace set at 550 °C for 3 h. The organic matter (VS) content was the complement of the mineral compounds (1-content of mineral compounds). The dry matter (DM) content was measured at 103 °C for 4 h to express the results relatively to the dry matter.

The biochemical methane potential (BMP) was determined according to the VDI 4630 standard described by Mayer et al. (2014). Briefly, samples were digested by a mesophilic anaerobic inoculum coming from the wastewater treatment plant of Schifflange (SIVEC, Luxembourg). This inoculum was first stabilized by incubating it at 37 °C for 4 days. An inoculum to substrate ratio of 2 (VS basis) was used to start the anaerobic digestion of the sample. The sample (30 g of wet weight) was digested at 37 °C for a period of 42 days in 2 L digester with a working volume of 1.6 L. The produced biogas was regularly collected and cooled down to 6 °C (to condense water vapor). Then, the volume and methane content of the biogas were measured with a wet drum-type gas meter (TG05, Ritter, Germany) and near infrared sensor (Dynament, UK), respectively. The volumes of biomethane were normalized at 0 °C and 1013 hPa according to the temperature and pressure conditions of each measurement. For each batch (series) of analyses, a triplicate assay was carried out on the inoculum alone and on microcrystalline cellulose. The inoculum alone assays of each batch were carried out to measure its biomethane produced solely. The microcrystalline cellulose assays of all the batches of the present study were used as a control substrate and to determine the standard error of the laboratory (SEL) of the BMP. The BMP value corresponds to the sum of all the partial measurements. The biomethane volumes produced by an inoculum solely were subtracted from this sum.

The enzymatically digestible organic matter (eDOM) was measured by the De Boever method (De Boever et al., 1986). Briefly, samples were incubated, in chronological order, with pepsin in 0.1 mol L^{-1} HCl for 24 h at 40 °C, with 0.1 mol L^{-1} HCl for 45 min at 80 °C, and with cellulase in an acetate buffer at pH 4.8 for 24 h at 40 °C.

All the measurements were carried out on the green-dried (and milled through a 1 mm screen) form of the biomass; except for the BMP analysis which was carried out on the silage-wet (and ground to a particle size of 1–2 cm) form of the biomass because this form of the biomass is closer to the industrial biomethanation process. Duplicate aliquots were measured on each sample, except for the BMP analysis for which one aliquot was measured on each sample.

2.3. Near infrared analysis

For the biomass analyzed under its green-dried form, the near infrared (NIR) reflectance spectra were taken by a NIRSystems 5000 (FOSS, Hillerød, Denmark) NIR spectrometer. Each spectrum was collected in the range of 1100–2498 nm and was the average of 32 scans.

For the biomass analyzed under its silage-wet and silage-dried form, the NIR reflectance spectra were taken by a MPA (Bruker, Billerica, USA) NIR spectrometer. Each spectrum was collected in the range of 1000–2778 nm (10,000–4000 cm^{-1}) with a resolution of 16 cm^{-1} and was the average of 64 scans.

For each presentation form (green-dried; silage-dried; silage-wet) of the biomass, the spectra were first trimmed to

1100–2498 nm. Then, they were normalized by a standard normal variate (SNV) transformation followed by a first order derivation (1, 4, 4, 1; 1st derivative, 4 nm gap, 4 points of first smoothing, 1 point of second smoothing). These treatments were performed using WinISI 4.6.8 (FOSS, Hillerød, Denmark).

2.4. Statistical analysis

Descriptive statistics, Pearson correlations and Tukey–Kramer multiple mean comparison tests (with $\alpha = 0.05$) based on the ANOVA of these means of the data were performed using JMP 11 (SAS Institute, Cary, USA). The differences (with $\alpha = 0.05$) between two MedRE were assessed using an *F*-test of variances using JMP 11 (SAS Institute, Cary, USA).

The non-linear matrix of the chemical composition corresponds to the full factorial design with the square terms of the chemical composition variables (lignin VST, cellulose VST, hemicelluloses VST, starch, total soluble sugars, proteins and mineral compounds). The matrix with the 7 chemical composition variables is referred as the linear matrix. Each independent variable of the linear and non-linear matrix was normalized by its mean and standard deviation (Eq. (1)) to have predictors with the similar weight.

$$\frac{X-\mu}{\sigma} \quad (1)$$

where X = the independent variable, μ = the mean of the independent variable, σ = the standard deviation of the independent variable.

The local (specific regression and non-linear procedure; Shenk et al., 1997), partial least square (PLS) (modified-PLS algorithm; linear regression model) and multiple linear regression (MLR; linear regression model) techniques were used to develop prediction models. The local and PLS models were performed with WinISI 4.6.8 (Infrasoft International LLC, State College, PA) and the MLR models were built with the Unscrambler X 10.3 (CAMO Software, Oslo, Norway).

The specific parameters to each local and PLS model were optimized according to the used software. The aim of this optimization was to build models with the best prediction performances (standard residual error of prediction of leave-one-out full cross-validation) that minimizes the under and over fitting of the model. These parameters were the optimum number of selected samples, and the minimum and maximum PLS components for the local prediction models. The parameter to optimize for the PLS prediction models was the number of PLS components. The MLR models were built with the most statistically significant independent variables. To select them, a mixed stepwise regression was performed where only the independent variables with a p-value < 0.001 were kept.

To prevent overestimation of the prediction performances of the models, each of them was evaluated using an independent validation (Val.) dataset in addition to a leave-one-out full cross-validation (C.-V.). The independent validation dataset was built by splitting the whole dataset into a calibration sub-set (used to build the model) and a validation sub-set (which is predicted). The validation sub-set contained approximately 20% of total samples (approximately 20% per biomass group) and was made out of representative and independent samples in regards to the calibration sub-set. These two sub-sets are considered as independent because, for each species, the samples of these two sub-sets never came from the same cropping site-year-harvest period.

To estimate the prediction performances of the models, the following parameters were determined for C.-V. and Val.: the coefficient of determination of prediction based on medians (R^2Med) (Eq. (2)); the median standard residual error of prediction (MedRE) (Hampel, 1974) (Eq. (3)); the ratio of the median standard deviation (SDMed) of the variable to MedRE (RPDMed = SDMed $*$ MedRE^{-1}); the ratio of MedRE to the standard error of laboratory (SEL) (MedRE $*$ SEL^{-1}). These parameters were calculated based on medians to be robust and to avoid deleting subjectively outlier samples (which have high residual values) of one dataset without deleting them of the other datasets or to have to also delete them of all the other datasets. Therefore, the R^2Med, MedRE, RPDMed and MedRE $*$ SEL^{-1} were determined based on medians.

$$R^2\text{Med} = \frac{\text{SDMed}^2 - \text{MedRE}^2}{\text{SDMed}^2} \quad (2)$$

where SDMed = median standard deviation of the variable, MedRE = median standard residual error of prediction.

$$\text{MedRE} = \text{MAD}^{*}1.4826 \text{ (Hampel,1974)} \quad (3)$$

where MAD = median of the absolute deviation of the residues.

In order to evaluate R^2Med and RDPMed of each prediction model, the guidelines suggested by Malley et al. (2005) were followed: excellent prediction model, $R^2\text{Med} \geqslant 0.95$ and RPDMed $\geqslant 4.0$; successful prediction model, $R^2\text{Med} \geqslant 0.90$ and RPDMed $\geqslant 3.0$; moderately successful prediction model, $R^2\text{Med} \geqslant 0.80$ and RPDMed $\geqslant 2.3$; moderately useful prediction model (semi-quantitative for screening purpose), $R^2\text{Med} \geqslant 0.70$ and RPDMed $\geqslant 1.8$.

3. Results and discussion

3.1. Chemical characteristics

The chemical characteristics of the analyzed biomasses are shown in Fig. 1a (chemical composition: lignin VST, cellulose VST, hemicelluloses VST, starch, total soluble sugars, proteins, mineral compounds) and Fig. 1a (eDOM and BMP). The eDOM has been considered because it is a relatively simple and fast method to assess the suitability of the plant biomasses to be converted by anaerobic digestion (Godin et al., 2013a,b,c). Furthermore, the eDOM will be used to have a better understanding and to compare the prediction performances of the BMP.

The chemical composition and the eDOM of the analyzed biomasses have already been reviewed by Godin et al. (2013a,b,c). Briefly, in the context of the present study, 3 distinctive types of plant biomass profiles can be observed based on the chemical composition (Fig. 1a) and the Tukey–Kramer multiple mean comparison tests: (1) fibrous plant biomasses (Miscanthus giganteus, switchgrass, spelt straw and fiber sorghum late winter) which have higher cellulose VST, hemicelluloses VST and lignin VST contents and lower total soluble sugars, proteins and mineral compounds contents; (2) less fibrous plant biomasses (fiber sorghum early autumn, tall fescue and immature rye) which have lower cellulose VST, hemicelluloses VST and lignin VST contents and higher total soluble sugars, proteins and mineral compounds contents; (3) rich in starch plant biomasses (fiber corn) which have a similar chemical composition to the less fibrous plant biomasses except that they have a high content of starch. The non-identified fraction (Fig. 1a) of the analyzed biomasses is most probably composed of soluble polysaccharides (such as pectins), acetyl groups of structural carbohydrates, acid soluble lignin, organic acids, alcohols, pigments and lipids, as also suggested by Hames (2009).

The fibrous crops which are more lignified (important proportion of stems) have a lower eDOM and the less fibrous or rich in starch crops which are less lignified (important proportion of leaves or of starch storage organs) and have more cells rich in cytoplasm (containing mineral salts) have a higher eDOM (Fig. 1b). The same trend can be observed for the BMP, except for the late winter

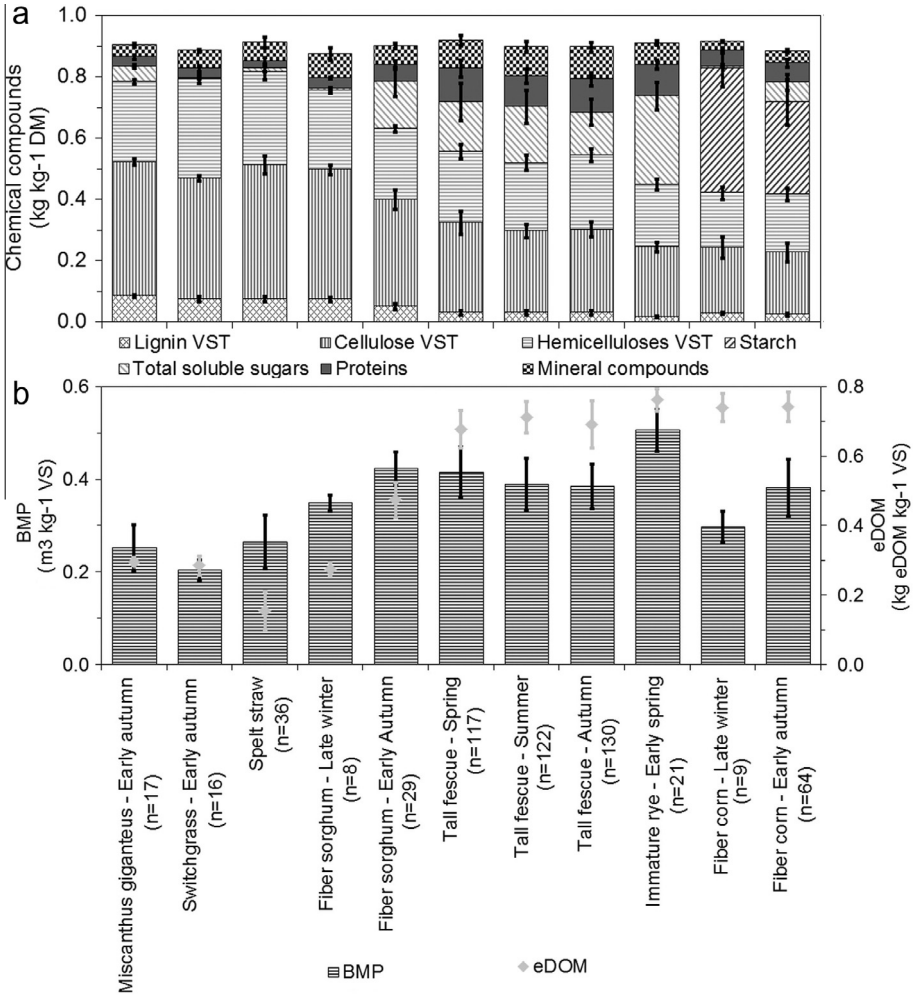

Fig. 1. (a) Chemical composition of the analyzed biomasses. Error bars correspond to standard deviation, *n* to the number of samples; (b) biochemical methane potential (BMP) and enzymatically digestible organic matter (eDOM) of the analyzed biomasses. Error bars correspond to standard deviation, *n* to the number of samples.

fiber corn which has a quite low BMP whereas the eDOM has a high value. This could be explained by less accessible carbohydrates to hydrolysis in the late winter fiber corn because of a more advanced stage of lignification, as compared to early autumn fiber corn. It is only observed for the BMP and not for the eDOM because the grinding of the samples for the eDOM gives a much smaller particle size (dried and milled with a 1 mm screen) compared to the samples for the BMP (ensiled with a particle size of 1–2 cm). The BMP has been determined on the silage (wet) samples with a particle size of 1–2 cm because it enables to be closer to the industrial biomethanation process.

Nearly all the correlations of the chemical components with the BMP and eDOM are significant (p-value < 0.05) (Table 2). Indeed, the BMP and eDOM are known to depend on the chemical composition of the plant biomasses (De Boever et al., 1986; Grieder et al., 2011, 2012; Triolo et al., 2011; Godin et al., 2010, 2013a,b,c). The fibrous components (cellulose VST, hemicelluloses VST and lignin VST) are correlated negatively to the BMP and eDOM whereas the non-fibrous components (starch, total soluble sugars and proteins) are correlated positively to the BMP and eDOM. This can be explained by the distinction of biomasses mainly made of stems, of leaves or of starch storage organs, as previously mentioned.

The content of lignin has the highest correlation with both BMP and eDOM, and it is negative (Table 2). It can be explained by the fact that lignin is not biodegradable under anaerobic conditions and that an increasing degree of lignification increases the inhibition of the anaerobic digestion (Klimiuk et al., 2010; Rath et al., 2013). Therefore, lignin is obviously a key parameter of the chemical composition to explain the inhibition of the anaerobic digestion.

There is only one correlation which is not significant (p-value $\geqslant$ 0.05). It is the correlation between starch and BMP

Table 2
Correlation of the biochemical methane potential (BMP) and the enzymatically digestible organic matter (eDOM) with the chemical composition (n = 569, except for starch n = 73).

	Lignin VST	Cellulose VST	Hemicelluloses VST	Starch	Total soluble sugars	Proteins	Mineral compounds
BMP	−0.58	−0.46	−0.47	0.19*	0.54	0.39	0.32
eDOM	−0.92	−0.92	−0.73	0.77	0.62	0.64	0.24

n, number of samples.
* Not significant (p-value $\geqslant$ 0.05).

(Table 2). This exception can be explained by the fact the late winter fiber corn has a high content of starch but not a high value of BMP. It is due to the less accessible carbohydrates to hydrolysis for BMP method, as previously mentioned.

To enhance the understanding and to improve the correlations between the chemical composition and the BMP, the BMP should be determined on dry samples milled through a 1 mm screen. However, this new presentation form of the biomass would be less close to industrial conditions of biomethanation

3.2. Chemical composition as a predictor

The descriptive statistics for the BMP and eDOM datasets of the analyzed biomasses used to build prediction models with the chemical composition as a predictor are shown in Table 3. The ranges of the BMP and eDOM are quite wide 0.488 m^3 kg^{-1} VS and 0.733 kg eDOM kg^{-1} VS, respectively (Table 3). The inter biomass group SD of the BMP and eDOM is approximately 4 and 2.5 times higher than the intra biomass group SD, respectively (Table 3). Having such a wide range of values with data well spread around the range is important to develop reliable prediction models. This is due to the use of a multispecies (multiproduct) dataset with similar plant species coming from different harvest periods and cropping conditions (year, area, cultivar, nitrogen fertilization level). Another major point to have reliable prediction models is to have a low standard error of laboratory (SEL) of the analytical method for the predicted variable. This is the case for the SEL of the present study's BMP and eDOM measurements of which the relative SEL of 5.1% and 1.1% (Table 3), respectively, are comparable with the best relative SEL of the BMP and eDOM observed in the literature (Decruyenaere et al., 2009; Doublet et al., 2013; Pham et al., 2013). Usually, the BMP method is performed on dry samples milled through a 1 mm screen. However, the present study intended to be closer to the industrial biomethanation process thus the BMP has been performed on silage (wet) samples with a particle size of 1–2 cm. Considering the sample's higher degree of heterogeneity, the present study's SEL is remarkably performant.

Owing to their prediction performances (R^2Med $\geqslant$ 0.70 and RPDMed $\geqslant$ 1.8 of cross-validation and validation) (Table 4), the local (specific regression and non-linear) models of the BMP with the chemical composition as a predictor are reliable enough to be used for semi-quantitative screening using a multispecies (multiproduct) dataset. The median standard residual errors of prediction (MedRE) of these local models are between 0.038 and 0.044 m^3 kg^{-1} VS (Table 4). Owing to their R^2Med and RPDMed of cross-validation and validation values of R^2Med < 0.70 and/or RPDMed < 1.8, the PLS (linear regression) and the MLR (linear regression) models give unsuccessful estimates of the BMP (Table 4). The ratio of MedRE to SEL is lower for the local models as compared to the MLR and PLS models (Table 4). Therefore, MedRE of the local models are closer to SEL compared to the PLS and MLR models. The higher prediction performances (R^2Med, RPDMed and MedRE) of the local models compared to the MLR and PLS models can be explained by the fact that with the local models a specific regression is built for each sample by selecting its most similar spectral neighbors in the used dataset (Shenk et al., 1997). Another advantage of non-linear models such as the local model is that they are able to take into account non-linearity present in the data and non-homogeneity of samples, as compared to linear models such as the MLR and the PLS model.

Table 3
Whole dataset summary for the analysis of the biochemical methane potential (BMP) and the enzymatically digestible organic matter (eDOM).

Biomass form	Near infrared spectrum			Chemical composition	
	Green-dried n = 588	Silage-dried n = 428	Silage-wet n = 465	Green-dried n = 569	
Predicted variable	BMP (m^3 kg^{-1} VS)			BMP (m^3 kg^{-1} VS)	eDOM (kg eDOM kg^{-1} VS)
Minimum	0.147	0.147	0.147	0.147	0.088
Maximum	0.635	0.635	0.635	0.589	0.821
Median	0.387	0.379	0.389	0.389	0.698
SDMed	0.080	0.076	0.073	0.076	0.083
Inter biomass group SD	0.139	0.114	0.125	0.406	0.135
Intra biomass group SD	0.056	0.055	0.055	0.052	0.052
SEL	0.020	0.020	0.020	0.020	0.0075
CV SEL	5.2%	5.3%	5.1%	5.1%	1.1%

n, number of samples.

Table 4
Prediction performances for the cross-validation (C.-V.; n = 569) and the validation (Val.; n = 112) of the biochemical methane potential (BMP) and the enzymatically digestible organic matter (eDOM) based on the chemical composition.

	Prediction of the BMP (m^3 kg^{-1} VS) by the chemical composition						Prediction of the eDOM (kg eDOM kg^{-1} VS) by the chemical composition					
	MLR		PLS		Local		MLR		PLS		Local	
	C.-V.	Val.	C.-V.	Val.	C.-V.	Val.	C.-V.	Val.	C.-V.	Val.	C.-V.	Val.
Linear matrix												
R^2Med	0.441	0.557	0.402	0.650	0.710	0.731	0.763	0.874	0.740	0.821	0.947	0.937
MedRE	0.057	0.055	0.058	0.049	0.041	0.044	0.041	0.030	0.042	0.036	0.019	0.021
Relative MedRE	15%	14%	15%	12%	10%	11%	5.8%	4.3%	6.1%	5.1%	2.7%	3.0%
RPDMed	1.3	1.5	1.3	1.7	1.9	1.9	2.0	2.8	2.0	2.4	4.4	4.0
MedRE $*$ SEL^{-1}	2.9	2.7	2.9	2.4	2.0	2.2	5.5	4.0	5.6	4.8	2.5	2.8
Non-linear matrix												
R^2Med	0.618	0.497	0.721	0.527	0.758	0.719	0.900	0.924	0.929	0.909	0.948	0.901
MedRE	0.047	0.058	0.040	0.057	0.038	0.044	0.026	0.023	0.022	0.025	0.018	0.026
Relative MedRE	12%	15%	10%	14%	9.8%	11%	3.8%	3.3%	3.2%	3.6%	2.5%	3.7%
RPDMed	1.6	1.4	1.9	1.5	2.0	1.9	3.2	3.6	3.7	3.3	4.4	3.2
MedRE $*$ SEL^{-1}	2.3	2.9	2.0	2.8	1.9	2.2	3.5	3.1	3.0	3.4	2.4	3.5

n, number of samples.

The reliability of the prediction models of the eDOM with the chemical composition as a predictor using a multispecies (multiproduct) dataset are shown in Table 4. The generally higher prediction performances (R^2Med and RPDMed) of the local models compared to the MLR and PLS models can be explained by the reasons mentioned above. Owing to their prediction performances (R^2Med and RPDMed), the prediction models of the eDOM are always more reliable compared to those of the BMP (Table 4). It can be explained by: (1) the higher accuracy of the eDOM method (relative SEL of 1.1%) compared to the BMP method (relative SEL of 5.1%) (Table 3); (2) the fact that the eDOM has been determined on more homogenous samples (dried and milled with a 1 mm screen) compared to the BMP (ensiled with a particle size of 1–2 cm). To improve the prediction models of the BMP by the chemical composition, the BMP should be determined on the sample particle size as the chemical composition, on dry samples milled through a 1 mm screen. However, in the present study a particle size of 1–2 cm has been used for the BMP assays to be closer to the industrial biomethanation process.

The correlation between the BMP and the eDOM is of 0.59 (p-value < 0.05). This is not high enough to build a reliable regression between these methods. It can be explained by different particle sizes of these two assays, as explained above, and by difference of degradation of the chemical components of the anaerobic digestion process for which each of these methods has been optimized. The eDOM method has been developed to predict the anaerobic digestion of the plant biomasses in the rumen. Whereas, the BMP method has developed to predict an anaerobic digestion that lasts for a much longer time in a biogas fermenter compared to a rumen. Therefore, chemical components which are not or slowly degraded in the rumen are partially degraded in the biogas process (Grieder et al., 2011, 2012).

The higher correlation of cellulose with the eDOM than with the BMP (Table 2) can be explained by this. Therefore, cellulose is one example of a component which seems to be much less degraded in the rumen compared to the biogas process. Further, cellulose has a negative effect on the MLR prediction model of the eDOM with the linear matrix of the chemical composition as a predictor whereas this effect is positive for the same type of MLR prediction model of the BMP.

For prediction models which are at least semi-quantitative (R^2Med $\geqslant$ 0.70 and RPDMed $\geqslant$ 1.8 of cross-validation and validation), the non-linear matrix of the chemical composition generally improves the prediction performances for the BMP and eDOM (Table 4). It can be explained by the fact that there are also non-linear effects to be taken into account to build a prediction model of the BMP and eDOM. Indeed, most of the parameters of the non-linear matrix for both of these parameters are made of non-linear predictors. These predictors can be found in Supplementary information. Nevertheless, there is no additional improvement of the prediction performances when the non-linear matrix is combined with the local model (Table 4). This means that local model already takes into account the non-linear effects because the local models are specific regressions built for each sample by selecting its most similar neighbors in terms of chemical composition in the used dataset. For the MLR and PLS models which have unsuccessful performances of predictions (R^2Med < 0.70 and/or RPDMed < 1.8 of cross-validation or validation) (Table 4), the use of the non-linear matrix induces a much better model's cross-validation reliability (R^2Med, RPDMed and MedRE) compared to the model's validation reliability (Table 4). This means that the use of the non-linear matrix in non-reliable models induces an over fitting of the predicted variable in cross-validation. Therefore, these models are not robust enough to predict independent samples similar to the analyzed biomasses. It is not the case of the prediction models which are at least semi-quantitative. They are robust enough to predict independent samples similar to the analyzed biomasses owing to their prediction performances of validation which are close to those of cross-validation (Table 4). For future prediction of the BMP and eDOM with the build models, the MedRE of validation should be considered, except if it is lower than the MedRE of cross-validation. In this case, the lower and therefore less performant MedRE of cross-validation should be considered. This can happen because the validation set has a smaller number of samples compared to the cross-validation set.

3.3. Near infrared spectrum as a predictor

The descriptive statistics for the BMP datasets of the analyzed biomasses used to build prediction models with the NIR spectrum as a predictor are shown in Table 3. The datasets corresponding to each presentation form (green-dried, silage-dried and silage-wet) have different numbers of samples because for some samples there was not enough substrate to take the NIR spectrum of the silage-dried and/or the silage-wet presentation form of the biomass. As for the dataset of the BMP prediction models by the chemical composition, the 3 NIR spectrum datasets (green-dried, silage-dried and silage-wet) have a wide range of BMP values with data well spread around the range (Table 3). This is important to develop reliable prediction models, as explained in Section 3.2.

Owing to their prediction performances (R^2Med $\geqslant$ 0.80 and RPDMed $\geqslant$ 2.3 of cross-validation and validation) (Table 5), the

Table 5
Prediction performances for the cross-validation (C.-V.) and the validation (Val.) of the biochemical methane potential (BMP) based on the near infrared (NIR) spectrum.

	Prediction of the BMP (m^3 kg^{-1} VS) by the NIR spectrum					
	Green-dried		Silage-dried		Silage-wet	
	C.-V. n = 588	Val. n = 112	C.-V. n = 428	Val. n = 89	C.-V. n = 465	Val. n = 90
Local						
R^2Med	0.851	0.808	0.863	0.850	0.848	0.849
MedRE	0.031	0.032	0.028	0.031	0.028	0.030
Relative MedRE	7.9%	8.4%	7.5%	8.0%	7.3%	7.8%
RPDMed	2.6	2.3	2.7	2.6	2.6	2.6
MedRE * SEL^{-1}	1.5	1.6	1.4	1.6	1.4	1.5
PLS						
R^2Med	0.796	0.832	0.802	0.771	0.781	0.778
MedRE	0.036	0.030	0.034	0.038	0.034	0.037
Relative MedRE	9.3%	7.9%	8.9%	9.9%	8.7%	9.5%
RPDMed	2.2	2.4	2.2	2.1	2.1	2.1
MedRE * SEL^{-1}	1.8	1.5	1.7	1.9	1.7	1.8

n, number of samples.

local (specific regression and non-linear) models of the BMP with the NIR spectrum as a predictor are reliable enough to be used for quantitative purposes using a multispecies (multiproduct) dataset. The MedRE of these local models are between 0.028 and 0.031 $m^3\ kg^{-1}$ VS (Table 5). These local models are reliable enough to estimate rapidly, cheaply, easily and quantitatively the BMP by NIR. Such models can possibly be further used for biomethanation plant management and optimization. Owing to their R^2Med and RPDMed of cross-validation and validation values of $R^2\text{Med} \geqslant 0.70$ and RPDMed $\geqslant 1.8$, the PLS (linear regression) models are only reliable enough to be used for semi-quantitative screening purposes using a multispecies (multiproduct) dataset. Therefore, their MedRE are globally less performant compared to the local models. The MedRE of these PLS models is between 0.030 and 0.038 $m^3\ kg^{-1}$ VS (Table 5). The ratio of the MedRE to the SEL is lower for the local models as compared to the PLS models (Table 5). Therefore, the MedRE of the local models is closer to SEL compared to the PLS models. The local models have higher prediction performances (R^2Med, RPDMed and MedRE) compared to the PLS models, as explained in Section 3.2.

To increase further the prediction performances of the BMP by the NIR models, the BMP method of the present study could be adapted by getting a better SEL. To achieve this enhancement, the samples could be dried and milled with a 1 mm screen instead of being ensiled (wet) with a particle size of 1–2 cm. This new presentation form of the biomass would decrease the degree of heterogeneity of the sample but it would be less close to industrial conditions of biomethanation. It would most probably increase the laboratory BMP but not the industrial biomethanation process BMP. The SEL could also be improved by increasing the number of replicates but this means that the method will get more expensive.

There is generally no significant difference (p-value $\geqslant 0.05$) between the MedRE of each pair of cross-validation and validation. This means that the built NIR models are robust enough to predict the BMP of independent samples similar to the analyzed biomasses with a comparable MedRE of validation. However, this difference between the MedRE is significant (p-value < 0.05) for the PLS model of biomasses presented under its green-dried form to the NIR spectrometer. For this pair of MedRE, the validation MedRE has a significantly lower value compared to the cross-validation MedRE. It can be explained by the fact that the validation set has a smaller number of samples compared to the cross-validation set. For this specific pair of MedRE, the less performant MedRE of the cross-validation should be considered for future BMP predictions with that model, whereas the MedRE of validation should be taken into account for the other pairs of MedRE of the other models.

The presentation form (green-dried, silage-dried and silage-wet form; different water contents and particle sizes) of biomasses to the NIR spectrometer did not have an obvious influence on the prediction performances (R^2Med, RPDMed and MedRE) of the BMP (Table 5). Thus, there is no presentation form (water content, ensiled or not, particle size) of the biomass to clearly prefer for the NIR models to predict the BMP. This means that these models mostly depend on the main NIR peaks of the infrared spectrum because when a high content of water is present in a sample (like the silage-wet form) then smaller NIR peaks will be hidden by the water NIR peaks. It suggests that NIR can potentially be used for online BMP analysis of wet plant biomasses as process control.

Based on the prediction model performances (R^2Med, RPDMed and MedRE) of the BMP, the local NIR models are much more reliable and the PLS NIR models are slightly more reliable, as compared to the prediction models with the chemical composition (Tables 4 and 5). This can be explained by the fact that with the NIR spectrum all the main chemical components and their interactions are taken into account whereas the chemical composition is only made of a few chemical components.

To compare the prediction performances of NIR prediction models of the same variable but of other studies, the relative prediction error (relative MedRE in the present study) can also be used in addition to the coefficient of determination of prediction (R^2Med in the present study) and the RPD of prediction (RPDMed in the present study). In comparison to other studies of NIR PLS models to predict the BMP (Lesteur et al., 2011; Raju et al., 2011; Doublet et al., 2013; Triolo et al., 2014), the NIR local and PLS models of the present study have higher and similar prediction performances, respectively (Table 5). The relative prediction error of the present study NIR local and PLS models are always below 8.5% and 10% (Table 5), respectively, whereas in these other studies it is always above 11% (Lesteur et al., 2011; Raju et al., 2011; Doublet et al., 2013; Triolo et al., 2014). This means that these other NIR models always have a higher relative prediction error, as compared to the present study. It can be explained by the relative SEL of approximately 5% (Table 3) which is comparable with the best SEL of the BMP observed in the literature (Doublet et al., 2013; Pham et al., 2013). It is also important to have a calibration dataset with a wide range and high variability in the parameter to be predicted, and high spectral variability. This was achieved in the present study by the use of a multispecies (multiproduct) dataset with similar plant species coming from different harvest periods and different cropping conditions (year, area, cultivar, nitrogen fertilization level). When comparing the coefficient of determination of prediction and RPD of prediction between previous studies (Lesteur et al., 2011; Raju et al., 2011; Doublet et al., 2013; Triolo et al., 2014) and the present study, the NIR models of only two of them (Doublet et al., 2013; Triolo et al., 2014) reaches similar performances to prediction quantitatively the BMP (coefficient of determination $\geqslant 0.80$ and RPD $\geqslant 2.3$). However, this can be explained by: (1) the sample size of the independent set to predict the BMP which is much smaller in these two studies (approximately 2–5 times smaller) (Doublet et al., 2013; Triolo et al., 2014), as compared to the present study (Table 5); (2) they determined the BMP on more homogenous samples (dried and milled with a 1 mm screen) compared to the present study (ensiled with a particle size of 1–2 cm), as explained in Section 3.2.

In comparison to the NIR prediction models of the BMP on the basis of their coefficient of determination of prediction (R^2Med in the present study) and RPD of prediction (RPDMed in the present study) (Table 5), the NIR prediction model of the eDOM is much more successful. Its coefficient of determination and RDP are 0.95 and 4.6, respectively (Decruyenaere et al., 2009). The explanation about the reason of this difference of the prediction performances of models to estimate the BMP and eDOM is given in Section 3.2.

4. Conclusions

The most reliable prediction models of the biochemical methane potential (BMP) of various plant biomasses using a multispecies dataset were those based on the near infrared (NIR) spectrum compared to those based on the chemical composition. The NIR predictions of the local models were able to estimate quantitatively, rapidly, cheaply and easily the BMP. Such a model could be further used for biomethanation plant management and optimization. The predictions of non-linear models were more reliable compared to those of linear models. The accuracy of the BMP method should be improved to enhance further the BMP prediction models.

Acknowledgements

This research was funded by the Walloon Agricultural Research Center (CRA-W) with the support of the Belgian Science Policy and

by the ENERBIOM project (ENERBIOM project n°14GR23024 of the European territorial co-operation program in the context of INTERREG IV A « Grande Région » 2007-2013 n°CCI2007CB163PO064, co-financed by the FEDER-EU funds) and the National Funds for Research, Luxembourg (FNR Core - C08/SR/13 - Ability of NIR Spectroscopy to Determine the Biomethane Potential from Organic Wastes and Agricultural Products - BIONIR). The authors are grateful to the technicians of the BIOETHA2 and the ENERBIOM and BIONIR projects for their technical support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biortech.2014.10.115.

References

Amon, T., Amon, B., Kryvoruchko, V., Zollitsch, W., Mayer, K., Gruber, L., 2007. Biogas production from maize and dairy cattle manure – influence of biomass composition on the methane yield. Agric. Ecosyst. Environ. 118, 173–182.

AOAC, 1990. Official Methods of Analysis, 15th ed. Association of Official Analytical Chemist, Washington, DC.

Bertrand, D., Dufour, E., 2006. La spectroscopie infrarouge et ses applications analytiques, 2ième édition. Lavoisier, Paris, France.

Berzaghi, P., Shenk, J., Westerhaus, M., 2000. Local prediction with near infrared multi-product databases. J. Near Infrared Spectrosc. 8, 1–9.

De Boever, J., Cottyn, B., Buysse, F., Vanacker, J., 1986. The use of an enzymatic technique to predict digestibility, metabolisable and net energy of compound feedstuffs for ruminants. Anim. Feed Sci. Technol. 14, 203–214.

Decruyenaere, V., Lecomte, Ph., Demarquilly, C., Aufrere, J., Dardenne, P., Stilmant, D., Buldgen, A., 2009. Evaluation of green forage intake and digestibility in ruminants using near infrared reflectance spectroscopy (NIRS): developing a global calibration. Anim. Feed Sci. Technol. 148, 138–156.

Doublet, J., Boulanger, A., Ponthieux, A., Laroche, C., Poitrenaud, M., Cacho Rivero, J., 2013. Predicting the biochemical methane potential of wide range of organic substrates by near infrared spectroscopy. Bioresour. Technol. 128, 252–258.

Duncan, M., Nigel, H., 2003. The Handbook of Water and Wastewater Microbiology. Academic Press, London, UK.

European Union, 2009. Commission Regulation No 152/2009. Off. J. Eur. Union L54, 1–130.

Godin, B., Ghysel, F., Agneessens, R., Schmit, T., Gofflot, S., Lamaudière, S., Sinnaeve, G., Goffart, J.-P., Gerin, P., Stilmant, D., Delcarte, J., 2010. Détermination de la cellulose, des hémicelluloses, de la lignine et des cendres dans diverses cultures lignocellulosiques dédiées à la production de bioéthanol de deuxième génération. Biotechnol. Agron. Soc. Environ. 14, 549–560.

Godin, B., Agneessens, R., Gerin, P., Delcarte, J., 2011a. Composition of structural carbohydrates in biomass: precision of a liquid chromatography method using a neutral detergent extraction and a charged aerosol detector. Talanta 85, 2014–2026.

Godin, B., Agneessens, R., Gofflot, S., Lamaudière, S., Sinnaeve, G., Gerin, P., Delcarte, J., 2011b. Revue sur les méthodes de caractérisation des polysaccharides structuraux des biomasses lignocellulosiques. Biotechnol. Agron. Soc. Environ. 15, 165–182.

Godin, B., Lamaudière, S., Agneessens, R., Schmit, T., Goffart, J.-P., Stilmants, D., Gerin, P., Delcarte, J., 2013a. Chemical composition and biofuel potentials of a wide diversity of plant biomasses. Energy Fuels 27, 2588–2598.

Godin, B., Lamaudière, S., Agneessens, R., Schmit, T., Goffart, J.-P., Stilmant, D., Gerin, P., Delcarte, J., 2013b. Chemical characteristics and biofuel potential of several vegetal biomasses grown under a wide range of environmental conditions. Ind. Crops Prod. 46, 1–12.

Godin, B., Lamaudière, S., Agneessens, R., Schmit, T., Goffart, J.-P., Stilmant, D., Gerin, P., Delcarte, J., 2013c. Chemical characteristics and biofuels potentials of various plant biomasses: influence of the harvesting date. J. Sci. Food Agric. 93, 3216–3224.

Godin, B., Agneessens, R., Gerin, P., Delcarte, J., 2014. Structural carbohydrates in plant biomasses: correlations between the detergent fiber and the dietary fiber methods. J. Agric. Food Chem. 62, 5609–5616.

Grieder, C., Mittweg, G., Dhillon, B., Montes, J., Orsini, E., Melchinger, A., 2011. Determination of methane fermentation yield and its kinetics by near infrared spectroscopy and chemical composition in maize. J. Near Infrared Spectrosc. 19, 463–477.

Grieder, C., Mittweg, G., Dhillon, B., Montes, J., Orsini, E., Melchinger, A., 2012. Kinetics of methane fermentation yield in biogas reactors: genetic variation and association with chemical composition in maize. Biomass Bioenergy 37, 134–141.

Hames, B., 2009. Biomass Compositional Analysis for Energy Applications. In: Mielenz, J. (Ed.), Biofuels: Methods and Protocols, Methods in Molecular Biology, vol. 581. Humana Press, New York, NY, pp. 145–183.

Hampel, F., 1974. The influence curve and its role in robust estimation. J. Am. Stat. Assoc. 69, 383–393.

Klimiuk, E., Pokoj, T., Budzynski, W., Dubis, B., 2010. Theoretical and observed biogas production from plant biomass of different fibre contents. Bioresour. Technol. 101, 9527–9535.

Lesteur, M., Latrille, E., Bellon-Maurel, V., Roger, J., Gonzalez, C., Junqua, G., Steyer, J., 2011. First step towards a fast analytical method for the determination of biochemical methane potential of solid waste by near infrared spectroscopy. Bioresour. Technol. 102, 2280–2288.

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

Mayer, F., Gerin, P., Noo, A., Foucart, G., Flammang, J., Lemaigre, S., Sinnaeve, G., Dardenne, P., Delfosse, P., 2014. Assessment of factors influencing the biomethane yield of maize silages. Bioresour. Technol. 153, 260–268.

Pham, C., Triolo, J., Cu, T., Pedersen, L., Sommer, S., 2013. Validation and recommendation of methods to measure biogas production potential of animal manure. Asian Aust. J. Anim. Sci. 26, 864–873.

Raju, C., Ward, A., Nielsen, L., Moller, H., 2011. Comparison of near infra-red spectroscopy, neutral detergent fibre assay and in-vitro organic matter digestibility assay for rapid determination of the biochemical methane potential of meadow grasses. Bioresour. Technol. 102, 7835–7839.

Rath, J., Heuwinkel, H., Herrmann, A., 2013. Specific biogas yield of maize can be predicted by the interaction of four biochemical constituents. Bioenergy Res. 6, 939–952.

Shenk, J., Westerhaus, M., Berzaghi, P., 1997. Investigation of a local calibration procedure for near infrared instruments. J. Near Infrared Spectrosc. 5, 223–232.

Triolo, J., Sommer, S., Moller, H., Weisbjerg, M., Jiang, X., 2011. A new algorithm to characterize biodegradability of biomass during anaerobic digestion: influence of lignin concentration on methane production potential. Bioresour. Technol. 102, 9395–9402.

Triolo, J., Pedersen, L., Qu, H., Sommer, S., 2012. Biochemical methane potential and anaerobic biodegradability of non-herbaceous and herbaceous phytomass in biogas production. Bioresour. Technol. 125, 226–232.

Triolo, J., Ward, A., Pedersen, L., Lokke, M., Qu, H., Sommer, S., 2014. Near Infrared Reflectance Spectroscopy (NIRS) for rapid determination of biochemical methane potential of plant biomass. Appl. Energy 116, 52–57.

Van Soest, P., Wine, R., 1967. Use of detergents in the analysis of fibrous feeds. IV. Determination of plant Cell Vall constituents. J. AOAC 50 (1), 50–55.

Van Soest, P., 1973. Collaborative study of acid-detergent fiber and lignin. J. AOAC 56 (4), 781–784.

Ward, A., Hobbs, P., Holliman, P., Jones, D., 2008. Optimisation of the anaerobic digestion of agricultural resources. Bioresour. Technol. 99, 7928–7940.