



Prediction of biochemical methane potential of urban organic waste using Fourier transform mid-infrared photoacoustic spectroscopy and multivariate analysis

Huang, Jing; Bekiaris, Georgios; Fitamo, Temesgen Mathewos ; Scheutz, Charlotte; Bruun, Sander

Published in:
Science of the Total Environment

Link to article, DOI:
[10.1016/j.scitotenv.2021.147959](https://doi.org/10.1016/j.scitotenv.2021.147959)

Publication date:
2021

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Huang, J., Bekiaris, G., Fitamo, T. M., Scheutz, C., & Bruun, S. (2021). Prediction of biochemical methane potential of urban organic waste using Fourier transform mid-infrared photoacoustic spectroscopy and multivariate analysis. *Science of the Total Environment*, 790, Article 147959. <https://doi.org/10.1016/j.scitotenv.2021.147959>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Journal Pre-proof

Prediction of biochemical methane potential of urban organic waste using Fourier transform mid-infrared photoacoustic spectroscopy and multivariate analysis

Jing Huang, Georgios Bekiaris, Temesgen Fitamo, Charlotte Scheutz, Sander Bruun



PII: S0048-9697(21)03030-8

DOI: <https://doi.org/10.1016/j.scitotenv.2021.147959>

Reference: STOTEN 147959

To appear in: *Science of the Total Environment*

Received date: 6 March 2021

Revised date: 10 May 2021

Accepted date: 17 May 2021

Please cite this article as: J. Huang, G. Bekiaris, T. Fitamo, et al., Prediction of biochemical methane potential of urban organic waste using Fourier transform mid-infrared photoacoustic spectroscopy and multivariate analysis, *Science of the Total Environment* (2021), <https://doi.org/10.1016/j.scitotenv.2021.147959>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Elsevier B.V. All rights reserved.

Prediction of biochemical methane potential of urban organic waste using Fourier transform mid-infrared photoacoustic spectroscopy and multivariate analysis

Jing Huang^a, Georgios Bekiaris^b, Temesgen Fitamo^c, Charlotte Scheutz^c, Sander Bruun^{a*}

^a *Department of Plant and Environmental Sciences, University of Copenhagen,*

Thorvaldsensvej 40, 1871 Frederiksberg, Denmark

^b *Department of crop science, Agricultural University of Athens, Pira Odos 75, 11855 Athens, Greece*

^c *Department of Environmental Engineering, Technical University of Denmark, Bygningstorvet, Building 115, DK-2800 Kongens Lyngby, Denmark*

***Corresponding Author:** Sander Bruun; Email: sab@plen.ku.dk

Abstract

Biochemical methane potential (BMP) assays are widely used to control the process of biogas production. However, the continuous evaluation of feedstocks using a BMP assay is expensive, time-consuming and challenging to optimize the composition of feedstocks in biogas plants. In this study, Fourier transform mid-infrared photoacoustic spectroscopy (FTIR-PAS) was used to predict the BMP values of 87 urban organic waste (UOW) samples derived from different sources in Denmark. The developed model of BMP prediction showed a coefficient of determination (R^2) of 0.86 and a root mean square error (RMSE) of 59.3 mL $\text{CH}_4/\text{g VS}$ in prediction. The interpretation of the regression coefficients used in the calibration showed a positive correlation with BMP for relatively easily degradable compounds, such as aliphatics, most likely lipids and amides most likely in proteins, while a negative correlation was found with lignin and hemicellulose.

Keywords: Anaerobic digestion; Biochemical methane potential (BMP); FTIR-PAS; Urban organic waste; PLSR modeling

1. Introduction

Anaerobic digestion (AD) is a well-established waste management method that can be used to recover energy in the form of biogas from a wide variety of organic substrates usually considered waste. An additional advantage is that the nutrients in the organic substrates can subsequently be recycled through land application of the digestates. There are many factors that can affect biogas production during the AD process, including feedstock quality, reactor design and operational conditions (Jingura and Kamusoko, 2017; Wu et al., 2019).

Since the quality of a given organic substrate or feedstock can have a significant effect on biogas production, it is crucial to know its biogas production potential in order to optimize the AD process (Fitamo et al., 2017). Thus biochemical methane potential (BMP) assays are commonly used to quantify the ultimate biogas potential. BMP expresses the maximum amount of biogas that can be produced from a certain feedstock given optimal conditions and sufficient time (Angelidaki et al., 2009). However, these assays are relatively laborious and time-consuming (typically in the range of 30-100 days), and therefore faster methods for estimating BMP would be very useful (Strömberg et al., 2015). Several biochemical models linking spectroscopic measures or other proxies to BMP have been developed in order to be able to predict BMP quickly, as described in Lesteur et al. (2010), such as predicting BMP by determining the concentration of components relevant to biodegradation processes. Triolo et al. (2011) developed a BMP prediction model for animal manure and energy crops based on methods for the extraction of fiber fractions. Furthermore, to speed up the BMP assessment, spectroscopic techniques have proved useful in the rapid characterization of BMP values of various samples that can be used to predict BMP using multivariate analysis. Several studies have investigated the feasibility of infrared (IR) spectroscopy for predicting the BMP values of different feedstocks, including the use of near infrared (NIR) spectroscopy to predict the BMP of municipal solid waste (Lesteur et al., 2011), meadow grasses (Raju et al., 2011), a

broad range of organic substrates (Doublet et al., 2013), plant biomasses (Godin et al., 2015; Triolo et al., 2014) and urban organic waste (Fitamo et al., 2017). Moreover, the prediction of BMP using NIR showed a better result compared with other methods related to elemental analysis, organic fraction composition and chemical oxygen demand based on a comparative analysis (Rodrigues et al., 2019).

Infrared spectroscopic methods are rapid, non-destructive and inexpensive techniques that require no or minor sample preparation. NIR spectroscopy has been used extensively for predictive purposes. However, the NIR region provides limited information about the chemical components related to the predicted values, since it is based on overlapping overtones and combination bands (mostly for C-H and O-H bonds). Mid-infrared (MIR) spectroscopy, however, has the advantage of containing primary vibrational bands that overlap much less and therefore lend themselves to easier interpretation. Furthermore, the development of different detectors and sampling techniques has led to easier sample preparation and increased recording accuracy, which have made MIR spectroscopy increasingly useful for the prediction of various sample properties. However, each detector offers different advantages for the various types of environmental samples (Bekiaris et al., 2020). Recently developed photoacoustic detectors use highly sensitive microphones to record the acoustic signal produced when infrared radiation interacts with the sample. This technique has proven suitable for many especially dark and opaque samples, and has the additional advantage of easy sample preparation and fast spectrum recording, making it ideal for relatively large sample sets.

Bekiaris et al. (2015b) applied Fourier transform mid-infrared photoacoustic spectroscopy (FTIR-PAS) to predict the BMP values of garden waste and showed that the method has the potential not only to predict BMP, but also identify the chemical components that contribute to it. This indicates that there is considerable potential in applying FTIR-PAS to evaluate and

predict the BMP values of organic wastes. However, the method has not been used to characterize a larger variety of urban organic wastes which are commonly used for biogas production and the potential for the techniques in that context therefore remains unexplored. Consequently, the main objectives of this study were: (i) to evaluate the predictive power of FTIR-PAS for BMP values of urban organic waste (UOW) derived from various sources, and (ii) to identify the chemical components of UOW that contribute to predicted BMP based on the recorded mid-infrared spectra.

2. Materials and methods

2.1. Sample collection

A set of 87 UOW samples for which BMP had been already measured by Fitamo et al., (2017) was retrieved from a sample archive and measured on FTIR-PAS instrument. These samples had been collected from various sources in Denmark, including private households, municipalities, canteens, private gardens, recycling stations, waste processing companies and wastewater treatment plants. A detailed description of these UOW samples is given in Fitamo et al. (2017). Briefly, all the UOW samples were divided into six groups based on the type of waste (Table 1). For example food waste samples contained vegetable and animal food waste, residual household waste and bio-waste. Samples of bio-pulp consisted of food and green waste mixtures. The samples of fiber fractions contained paper fractions including kitchen tissue, wrapping paper, egg trays etc. Plant materials were collected from grass clippings, garden waste, flowers, animal straw and vegetation. As for industrial waste, several different kinds of organic waste were collected, such as various types of sludge, cooked straw, low-grade milk, Brie cheese and a blend of dairy products and meat. For the mixture of food waste, plant and industrial samples, a blend of food waste, plant material and sludge was obtained.

2.2. BMP assay

The BMP values of waste samples were determined through batch assays (Fitamo et al., 2017). The BMP of UOW samples was measured in triplicate at a temperature of 37 °C (mesophilic conditions) in batch reactors of 1 L and an effective volume of 0.3 L. The inoculum-to-substrate ratio (ISR) on a volatile solid (VS) basis was 2 and the loading rate was 2.7 g VS/L. The substrate in the reactors was inoculated with inoculum collected from the Va Syd Sjölanda wastewater treatment plant in Malmö, Sweden (SE). The assay bottles containing substrates and inoculum were purged with N₂ to maintain anaerobic conditions. The BMP assay included blanks (inoculum only), test samples (UOW) and a positive control (Avicel®, microcrystalline cellulose). The actual substrate biogas production was corrected for background biogas production from inoculum using the data obtained from the blanks, and the quality of the BMP was validated with the positive controls. The average methane production of the controls was 392 mL CH₄/g VS and the gas production from the inoculum was 10% of the controls.

The concentration of methane in the produced biogas was determined by gas chromatography (TRACE 1310 GC) using a flame ionization detector (TRACE 1310 GC, Thermo Scientific). Methane production was then calculated by multiplying the volume of biogas produced in the headspace by the concentration of methane in the biogas. BMP (mL CH₄/g VS) was determined by dividing the net substrate methane production by the amount of VS of the substrates used (Fitamo et al., 2017).

2.3. FTIR-PAS analysis

The collected UOW samples were freeze-dried and ground by a hammer mill with a 1 mm screen size and packed in small cups (10 mm diameter). The FTIR-PAS spectra of UOW samples were then recorded using a PA301 photoacoustic detector (Gasera Ltd. Turku,

Finland) connected to a Nicolet 6700 FTIR spectrometer (Thermo Scientific, USA). A helium gas purging flow was applied in the detector's recording chamber to reduce the noise produced by moisture. For each sample, 32 scans were recorded in the mid-infrared region between 4000 and 600 cm^{-1} at a resolution of 4 cm^{-1} , and then averaged.

2.4. Data processing and statistical analysis

Prior to any analyses, the FTIR spectra were mean-centered. Principal component analysis (PCA) was then performed to identify potential outliers and provide an overview of the relationship between the samples. Prior to PCA analysis, the spectra were pre-processed using standard normal variate (SNV), de-trending (DT) and Savitzky-Golay (SG) smoothing (four points on each side with a first-order polynomial).

In order to construct a predictive model of BMP, the dataset was divided into a calibration set to develop the model (61 samples) and a test set for validation (20 samples). The test set was constructed by selecting every fourth sample after sorting the samples in the ascending order of their BMP values. Several spectral transformations and combinations were performed, including SG smoothing, normalization, baseline correction, SNV, DT, and first and second derivatives. Subsequently, partial least square regression (PLSR) models were performed for BMP prediction of the UOW samples on the calibration set using different spectra transformations. The optimal component number was selected according to a cross-validation using 20 segments.

In addition, variable reduction was carried out using a jack-knife technique in order to improve the predictions and identify significant sections of the spectra (Martens and Martens, 2000). This procedure was repeated until the model prediction no longer improved.

All the spectra transformations and calibrations were performed by the software Unscrambler 10.3 (CAMO software, Oslo, Norway).

The number of PLSR components were selected based on the procedure implemented in the software where one more component is added if the residual variance increase is more than 6% in a cross validation. The performance of the PLSR models was assessed by the coefficient of determination (R^2) and the root mean square error (RMSE). RMSE represents the average error either from the cross-validation ($RMSE_{CV}$) or from the prediction of the test set validation ($RMSE_P$). Similarly, R^2 denotes the fraction of explained variance of either the cross-validation (R^2_{CV}) or the test set validation (R^2_P).

3. Results and discussion

3.1. BMP values of the UOW samples

The BMP values of UOW samples are summarized by different classes based on the type of waste shown in Table 1. In addition, the distribution of the BMP values for all the UOW samples can be found in Supplementary information (Figure S1). Among the various UOW samples, food waste obtained the highest average BMP value (593 mL CH_4/g VS). Mixed waste and biopulp slurry showed similar average BMP values of 535 mL CH_4/g VS and 528 mL CH_4/g VS respectively. Industrial organic waste displayed the largest variation in BMP, reflecting the wide diversity of samples in this group. For example, milk and dairy products blended with meat byproducts had BMP values of 818 and 762 mL CH_4/g VS respectively, while digested sludge from a biogas plant had a BMP of 148 mL CH_4/g VS. Fiber fraction and plant material samples showed relatively low BMP values, with an average BMP of 335 and 277 mL CH_4/g VS respectively.

3.2. PCA of UOW samples

Before PCA, the averaged spectra of different types of UOW samples were given in supplementary Figure S2. For the PCA, six samples were indicated to be potential outliers by Hotelling- T^2 and Q-residuals and removed. After this, the first three principle components

(PCs) explained 65 %, 10 % and 6 % of the spectral variation respectively. Furthermore, when plotting PC1 vs. PC3, a clear separation of the different types of waste was achieved, with food waste and fibrous materials in particular being separate (Fig. 1). The interpretation of PC1 loadings provided in Fig. 2 revealed a positive correlation between sample components, with regions at 3600 cm^{-1} (O-H stretching in hydroxyl groups and water), 1431 cm^{-1} (carbonate), $1170\text{-}1060\text{ cm}^{-1}$ (C-O stretching in polysaccharides) (Bekiaris et al., 2015a; Bolyard et al., 2019), and 1032 cm^{-1} (xylose or other hemicelluloses) (Sills and Gossett, 2012).

A negative correlation between PC1 and BMP was revealed in the region of 1741 cm^{-1} (C=O stretching vibrations in esters, aldehydes and ketones). This region was also found to be negatively associated with humic acid in samples of composted biowaste (Meissl et al., 2007). Therefore, the negative correlation with 1741 cm^{-1} may also have reflected the degree of humification of the materials, which could reflect the degradation of organic waste. A negative correlation with PC1 was also found in the region of 1554 cm^{-1} (N-H or C-N stretching in amide II), and the region between 1620 and 1675 cm^{-1} , with negative correlations at 1664 cm^{-1} and 1630 cm^{-1} (C=O stretching in carboxylates and amides) (Carballo et al., 2008). Moreover, these regions, especially at 1640 cm^{-1} , could also be associated with aromatic C=C vibrations, C=C in alkenes and OH from water, which makes their interpretation equivocal (Meissl et al., 2007).

Furthermore, PC1 was also negatively correlated with the regions at 2929 and 2854 cm^{-1} (C-H of aliphatic methylene), 1223 cm^{-1} (C-O stretching and CH deformation in polysaccharides and lignin, N-H rocking and C-O stretching in amino acids, Amide III band) (Bolyard et al., 2019), and 850 and 775 cm^{-1} (primary amine group; C-H deformations and ring vibration of α -pyranose ring respectively) (Smidt and Meissl, 2007).

3.3. PLSR models of BMP

The performance of the PLSR models based on the various spectral transformations is given in Table 2. In general, spectral transformations improved the performance of the BMP prediction models, except for Savitzky-Golay smoothing (Model b) and DT (Model f). In addition, normalization, baseline correction and SNV (Model c, d and e) improved the predictive capability of the BMP prediction model in the test set, as assessed by a higher R^2_P and lower $RMSE_P$, but did not improve the performance in the cross-validation, as assessed by R^2_{CV} and $RMSE_{CV}$. For these transformations, the performance in the test set validation was better than in the cross-validation. This was slightly surprising, but most likely could be explained by the small test set, which means that the selected samples were particularly easy to predict with this transformation.

Model j based on a SNV, DT and second derivative transformation achieved the best performance in the cross-validation, with R^2_{CV} of 0.90 and $RMSE_{CV}$ of 51.6 mL CH₄/g VS. Unfortunately, this model did not demonstrate the best performance in the test set validation. The best overall performance with a good prediction performance in the cross-validation of the test set as well was achieved when the spectra were pre-processed with a combination of SNV, DT and first derivative, with R^2_P of 0.86 and $RMSE_P$ of 59.3 mL CH₄/g VS (Model i). This was achieved with only two components (Supplementary Figure S3).

Table 3 provides a summary of studies predicting the BMP values of a variety of waste types using IR spectroscopy. Fitamo et al. (2017) used the same UOW sample set as this research work, and developed NIR spectroscopy-based BMP prediction models. For NIR predictions, the combination of SNV, DT and first derivative also provided the best performance, with values of R^2_P of 0.88 and $RMSE_P$ of 44 mL CH₄/g VS, which was slightly better than the current FTIR-PAS-based models. However, the performance of their models in the cross-

validation was considerably lower with R^2_{CV} of 0.83, and the number of nine PLSR components was considerably higher than in the FTIR-PAS models.

Bekiaris et al. (2015b) used FTIR-PAS to predict the BMP values of plant biomass samples. The best BMP model was obtained by using FTIR-PAS with pre-processing of SNV, DT and second derivative, with R^2_p of 0.81 and $RMSEP$ of 40 mL CH_4/g VS, which was approximately the same as predicted.

The studies of Lesteur et al. (2011) and Bekiaris et al. (2015b) showed that the BMP-predicting models could be improved by using variable selection and reduction methods. This is in agreement with what was found in the present study. The number of PLSR components used here (2-6, Table 2) was in the lower range of the previous studies. This indicated that the variable selection methods combined with the pre-processing methods could reduce the number of PLSR components required and thus also improve the robustness of the models.

3.4. Analysis of regression coefficients

To understand the chemical components reflected in the FTIR-PAS spectra that affect the BMP of the UOW samples, the PLSR model with pre-processing of smoothing and normalization (Model c, Table 2) was chosen for spectral interpretation, which was also the same transformation used for interpretations in Bekiaris et al. (2015b). Fig. 3 shows the regression coefficients of the PLSR model based on Model c in Table 2.

Positive regression coefficients were observed in the spectral regions at 2925 cm^{-1} and 2854 cm^{-1} , representing the C–H stretch of aliphatic structures (Carballo et al., 2008), which are relatively easily degradable. Similar observations were also made by Bekiaris et al. (2015b), who found a positive correlation between the regions at 2910 and 2800 cm^{-1} and predicted the BMP values of plant biomass samples. Positive regression coefficients were also found in the region of 2665 cm^{-1} , associated with the NH_2 stretch of amides (Smidt and Meissl, 2007), but

could also be associated with an overtone of the CH deformation vibration (Socrates, 2001). The interpretation of a spectral region at around 3200 cm^{-1} could be ambiguous due to the overlap of various compounds such as carbonyl, hydroxyl groups, amides etc. (Peltre et al., 2014). The positive correlation between BMP and the spectral region at around 714 cm^{-1} could be attributed to amine groups (Smidt and Meissl, 2007). It may also be assigned to CH_2 deformation vibrations of simple aromatic compounds of low molecular weight and the CH_2 rocking vibration of ethyl groups of alkanes or lipids (Socrates, 2001).

In the present study, the spectral regions with the largest regression coefficients were at 2925 cm^{-1} and 2854 cm^{-1} , corresponding to aliphatic components, while Bekiaris et al. (2015b) found the largest regression coefficients in the regions 980 and 900 cm^{-1} , attributed to amorphous cellulose. This difference is likely to stem from the different origin of the samples. The UOW waste samples analyzed in the present study are much more diverse than the plant biomass samples analyzed in Bekiaris et al. (2015b). A large proportion of the samples were food waste, which contains more aliphatic substances mainly derived from fat, which is rarely found in plant biomass. In contrast, plant biomass mainly consists of cellulose, hemicellulose and lignin, with amorphous cellulose as the easily hydrolysable polymer form of cellulose. UOW samples have more easily degradable compounds than cellulose, which in crystalline form is quite resistant to hydrolysis. The other compounds that positively correlated with BMP were relatively easily degradable, such as amides (2665 cm^{-1}) or amine in proteins and lipids (714 cm^{-1}).

However, several spectral regions displayed negative regression coefficients. For example, the spectral region at $3668\text{-}3402\text{ cm}^{-1}$ is dominated by hydroxyl groups, which are present in most organic compounds, and it is therefore difficult to pinpoint any that are giving rise to the negative regression coefficients (Bekiaris et al., 2015a). Negative regression coefficients were also observed around 1510 cm^{-1} , corresponding to the C=C stretching in the aromatic

ring of lignin and a characteristic “lignin” peak (Laureano-Perez et al., 2005). Moreover, there were negative regression coefficients around 962 cm^{-1} , which could be assigned to C-O-C stretching vibrations that are likely to correspond to cellulose (Bekiaris et al., 2015b). In contrast with the present results, Bekiaris et al. (2015b) showed positive regression coefficients in this region. Again this difference is likely to stem from the different origin of the samples. In the plant biomass samples of Bekiaris et al. (2015b), the cellulose was probably relatively ubiquitous, but in the more diverse UOW samples of the present study, the presence of cellulose was likely to be correlated with lignocellulose that are probably mainly present in large amount in specific samples.

The spectral region at around 1225 cm^{-1} displayed negative regression coefficients and is usually attributed to polysaccharides (Bolyard et al., 2019). In this study, it was most probably dominated by hemicellulose, which together with lignin hinder the degradation of cellulose (Bekiaris et al., 2015b). Additionally, the negative coefficient peak at around 1261 cm^{-1} could be assigned to carboxylic acids (C-O vibration) and amides (C-N vibration) (Smidt and Meissl, 2007). It could also be assigned to aromatic C (Provenzano et al., 2014), which may explain the recalcitrant compounds. The spectral region at around 812 cm^{-1} could be assigned to the vibration of Si-O of quartz (Peltre et al., 2014). However, in this case, it was most probably assigned to the C-H out-of-plane vibration in aromatic compounds (Mochidzuki et al., 2003).

In the study of Meissl et al. (2007), the peak at around 1261 cm^{-1} was also found to be useful for characterizing the biowaste composting process. Meissl et al. (2007) showed that the peak was getting weaker during composting, indicating maturation of the composted material. Therefore this could explain the negatively correlated regression coefficient in the spectral region of 1261 cm^{-1} . For example, organic waste with easily degradable materials can show a

considerable BMP, while materials with less degradability (more humification) may show lower BMP values (Naroznova et al., 2016).

4. Conclusions

To conclude, FTIR-PAS was capable of efficiently predicting the BMP values of UOW (R^2 of 0.86 and RMSE of 59.3 mL $\text{CH}_4/\text{g VS}$ for test set validation). This is comparable to the prediction based on NIR spectroscopy obtained on the same samples in a former study. BMP prediction models could be improved with the use of proper spectral transformation methods. Moreover, several spectral regions related to BMP were identified using PLSR models of FTIR-PAS spectra. In general, the BMP models of the UOW samples displayed positive regression coefficients with compounds such as aliphatics and amides, probably from lipids and proteins that are also considered easily degradable, while compounds such as lignin and hemicellulose, which are usually considered resistant to degradation, displayed a negative correlation.

Acknowledgements

We gratefully acknowledge the financial support from the China Scholarship Council for Jing Huang (201606350206).

References

Angelidaki, I., Alves, M., Bolzonella, D., Borzacconi, L., Campos, J.L., Guwy, A.J.,

Kalyuzhnyi, S., Jenicek, P., Van Lier, J.B., 2009. Defining the biomethane potential (BMP) of solid organic wastes and energy crops: A proposed protocol for batch assays. *Water Sci. Technol.* 59, 927–934.

Bekiaris, G., Bruun, S., Peltre, C., Houot, S., Jensen, L.S., 2015a. FTIR-PAS: A powerful tool for characterising the chemical composition and predicting the labile C fraction of various organic waste products. *Waste Manag.* 39, 45–56.

- Bekiaris, G., Peltre, C., Barsberg, S.T., Bruun, S., Sørensen, K.M., Engelsen, S.B., Magid, J., Hansen, M., Jensen, L.S., 2020. Three different Fourier-transform mid-infrared sampling techniques to characterize bio-organic samples. *J. Environ. Qual.* 49, 1310–1321.
- Bekiaris, G., Triolo, J.M., Peltre, C., Pedersen, L., Jensen, L.S., Bruun, S., 2015b. Rapid estimation of the biochemical methane potential of plant biomasses using Fourier transform mid-infrared photoacoustic spectroscopy. *Bioresour. Technol.* 197, 475–481.
- Bolyard, S.C., Reinhart, D.R., Richardson, D., 2019. Conventional and fourier transform infrared characterization of waste and leachate during municipal solid waste stabilization. *Chemosphere* 227, 34–42.
- Carballo, T., Gil, M.V., Gómez, X., González-Andrés, F., Morán, A., 2008. Characterization of different compost extracts using fourier-transform infrared spectroscopy (FTIR) and thermal analysis. *Biodegradation* 19, 815–830.
- Doublet, J., Boulanger, A., Ponthieux, A., Laroche, C., Poitrenaud, M., Cacho Rivero, J.A., 2013. Predicting the biochemical methane potential of wide range of organic substrates by near infrared spectroscopy. *Bioresour. Technol.* 128, 252–258.
- Fitamo, T., Triolo, J.M., Boldrin, A., Scheutz, C., 2017. Rapid biochemical methane potential prediction of urban organic waste with near-infrared reflectance spectroscopy. *Water Res.* 119, 242–251.
- Godin, B., Mayer, F., Agneessens, R., Gerin, P., Dardenne, P., Delfosse, P., Delcarte, J., 2015. Biochemical methane potential prediction of plant biomasses: Comparing chemical composition versus near infrared methods and linear versus non-linear models. *Bioresour. Technol.* 175, 382–390.

- Jingura, R.M., Kamasoko, R., 2017. Methods for determination of biomethane potential of feedstocks: A review. *Biofuel Res. J.* 4, 573-586.
- Laureano-Perez, L., Teymouri, F., Alizadeh, H., Dale, B.E., 2005. Understanding factors that limit enzymatic hydrolysis of biomass. *Appl. Biochem. Biotechnol.* 124, 1081–1099.
- Lesteur, M., Bellon-Maurel, V., Gonzalez, C., Latrille, E., Roger, J.M., Junqua, G., Steyer, J.P., 2010. Alternative methods for determining anaerobic biodegradability: A review. *Process Biochem.* 45, 431-440.
- Lesteur, M., Latrille, E., Maurel, V.B., Roger, J.M., Gonzalez, C., Junqua, G., Steyer, J.P., 2011. First step towards a fast analytical method for the determination of Biochemical Methane Potential of solid wastes by near infrared spectroscopy. *Bioresour. Technol.* 102, 2280–2288.
- Martens, H., Martens, M., 2000. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual. Prefer.* 11, 5–16.
- Meissl, K., Smidt, E., Schwanninger, M., 2007. Prediction of humic acid content and respiration activity of biogenic waste by means of Fourier transform infrared (FTIR) spectra and partial least squares regression (PLS-R) models. *Talanta* 72, 791–799.
- Mochidzuki, K., Soutric, F., Tadokoro, K., Antal, M.J., Tóth, M., Zelei, B., Várhegyi, G., 2003. Electrical and physical properties of carbonized charcoals. *Ind. Eng. Chem. Res.* 42, 5140–5151.
- Naroznova, I., Møller, J., Scheutz, C., 2016. Characterisation of the biochemical methane potential (BMP) of individual material fractions in Danish source-separated organic household waste. *Waste Manag.* 50, 39–48.

- Peltre, C., Bruun, S., Du, C., Thomsen, I.K., Jensen, L.S., 2014. Assessing soil constituents and labile soil organic carbon by mid-infrared photoacoustic spectroscopy. *Soil Biol. Biochem.* 77, 41–50.
- Provenzano, M.R., Malerba, A.D., Pezzolla, D., Gigliotti, G., 2014. Chemical and spectroscopic characterization of organic matter during the anaerobic digestion and successive composting of pig slurry. *Waste Manag.* 34, 653–660.
- Raju, C.S., Ward, A.J., Nielsen, L., Møller, H.B., 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.
- Rodrigues, R.P., Rodrigues, D.P., Klepacz-Smolva, A., Martins, R.C., Quina, M.J., 2019. Comparative analysis of methods and models for predicting biochemical methane potential of various organic substrates. *Sci. Total Environ.* 649, 1599–1608.
- Sills, D.L., Gossett, J.M., 2012. Using FTIR to predict saccharification from enzymatic hydrolysis of alkali-pretreated biomasses. *Biotechnol. Bioeng.* 109, 353–362.
- Smidt, E., Meissl, K., 2007. The applicability of Fourier transform infrared (FT-IR) spectroscopy in waste management. *Waste Manag.* 27, 268–276.
- Socrates, G. *Infrared and Raman characteristic group frequencies: Tables and charts*. 3rd ed. Chichester, U.K.: John Wiley & Sons; 2001.
- Strömberg, S., Nistor, M., Liu, J., 2015. Early prediction of Biochemical Methane Potential through statistical and kinetic modelling of initial gas production. *Bioresour. Technol.* 176, 233–241.
- Triolo, J.M., Sommer, S.G., Møller, H.B., Weisbjerg, M.R., Jiang, X.Y., 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.M., Ward, A.J., Pedersen, L., Løkke, M.M., Qu, H., Sommer, S.G., 2014. Near

Infrared Reflectance Spectroscopy (NIRS) for rapid determination of biochemical

methane potential of plant biomass. *Appl. Energy* 116, 52–57.

Wu, D., Li, L., Zhao, X., Peng, Y., Yang, P., Peng, X., 2019. Anaerobic digestion: A review

on process monitoring. *Renew. Sustain. Energy Rev.* 103, 1–12.

Journal Pre-proof

Tables**Table 1.**

Maximum, minimum, mean and standard deviation of BMP values (mL CH₄/g VS) of urban organic waste samples. The samples are divided into six different classes based on the type of waste (n: number of samples)

	Food waste	Biopulp	Fiber fraction	Plant materials	Industrial	Mixture of food waste, plant and industrial
<i>n</i>	37	14	18	6	10	2
Max.	906	672	544	383	818	573
Min.	260	292	136	119	148	496
Mean	593	528	335	277	492	535
SD	141	105	100	108	236	54

Table 2.

FTIR-PAS BMP-predicting PLSR models of UOW samples. The number of variable reductions indicates the number of times a jack-knife technique was used to reduce the number of variables included in the final model (Cal: calibration; F: PLSR component number; R^2 : coefficient of determination; RMSE: root mean square value in ml $\text{CH}_4/\text{g VS}$; CV: cross validation; P: prediction for test validation)

Mode l No.	Transformations	Number of variable reductions	Cal set	Test set	F	R^2_{CAL}	J.M.S.E _{CA}	R^2_{CV}	RMSE _C v	R^2_{P}	RMSE _P
a	None	1	61	20	4	0.86	62.6	0.81	72.2	0.78	74.2
b	Smoothing ¹	1	61	20	4	0.86	62.8	0.81	72.4	0.78	74.4
c	Smoothing ¹ +normalization ²	3	61	20	4	0.86	62.0	0.81	72.8	0.86	58.3
d	Smoothing ¹ +baseline ³ +normalization ²	2	61	20	6	0.85	64.1	0.79	76.9	0.87	56.6
e	SNV ⁴	2	61	20	3	0.86	63.5	0.80	74.6	0.85	61.0
f	DT ⁵	1	61	20	5	0.88	58.5	0.80	73.6	0.72	84.6
g	SNV ⁴ +DT ⁵	1	61	20	3	0.85	64.9	0.79	75.4	0.80	70.1
i	SNV+DT+1 st derivative	1	61	20	2	0.89	56.6	0.85	63.7	0.86	59.3
j	SNV+DT+2 nd derivative	1	61	20	3	0.94	40.2	0.90	51.6	0.81	68.9

Note: All the spectra were mean-centered

¹Smoothing: Savitzky-Golay smoothing, 0 polynomial, 9 smoothing points

²Normalization: normalization by mean

³Baseline: linear baseline correction

⁴SNV: standard normal variate

⁵DT: de-trending, 2nd polynomial

1st derivative: Savitzky-Golay 1st derivative, second polynomial, 11 smoothing points

2nd derivative: Savitzky-Golay 2nd derivative, second polynomial, 11 smoothing points

Journal Pre-proof

Table 3.

Prediction of BMP values of different varieties of waste using infrared spectroscopy (IR) reported in previous studies (N: total number of sample set; F: PLSR component number; Cal: calibration; Val: validation; R^2_p : coefficient of determination of prediction; RMSEP: root mean square error of prediction in mL CH_4/g VS; R^2_{Med} : the coefficient of determination of prediction based on medians; MedRE: the median standard residual error of prediction)

Materials	Spectroscopy	Variable range	N	F	Transformations	Cal set	Val set	R^2_p	RMSEP	References
Urban organic waste (UOW)	FTIR-PAS	4000-600 cm^{-1}	87	2	SNV + 1 st derivative	61	20	0.86	59.3	This study
Municipal solid wastes, raw wastes and lignocellulosic green wastes	NIR	1668-2500 nm	102	7	SNV + DT	51	23	0.76	28	(Lesteur et al., 2011)
Meadow grasses	NIR	832-2500 nm	95	13	Mean normalization	95	95 leave one out cross validation	0.69	37	(Raju et al., 2011)
Organic substrates	NIR	999-2500 nm	296	7	SNV + DT + 2 nd derivative	173	53	0.85	40	(Doublet et al., 2013)
Plant biomasses	NIR	1000-2500 nm	88	6	SNV + DT + 2 nd derivative	66	22	0.84	37	(Triolo et al., 2014)
Plant biomasses (green-dried)	NIR	1100-2498 nm	588	11	SNV + 1 st derivative	569	112	0.83 (R^2_{Med})	30 (MedRE)	(Godin et al., 2015)
Plant biomasses	FTIR-PAS	4000-600 cm^{-1}	87	5	SNV + DT + 2 nd	64	21	0.81	40	(Bekiaris et

				derivative				al., 2015b)	
UOW			87	9		63	21	0.88	44
UOW+Plant biomasses	NIR	1000-2500 nm	175	9	SNV + DT + 1 st derivative	108	59	0.89	50

Journal Pre-proof

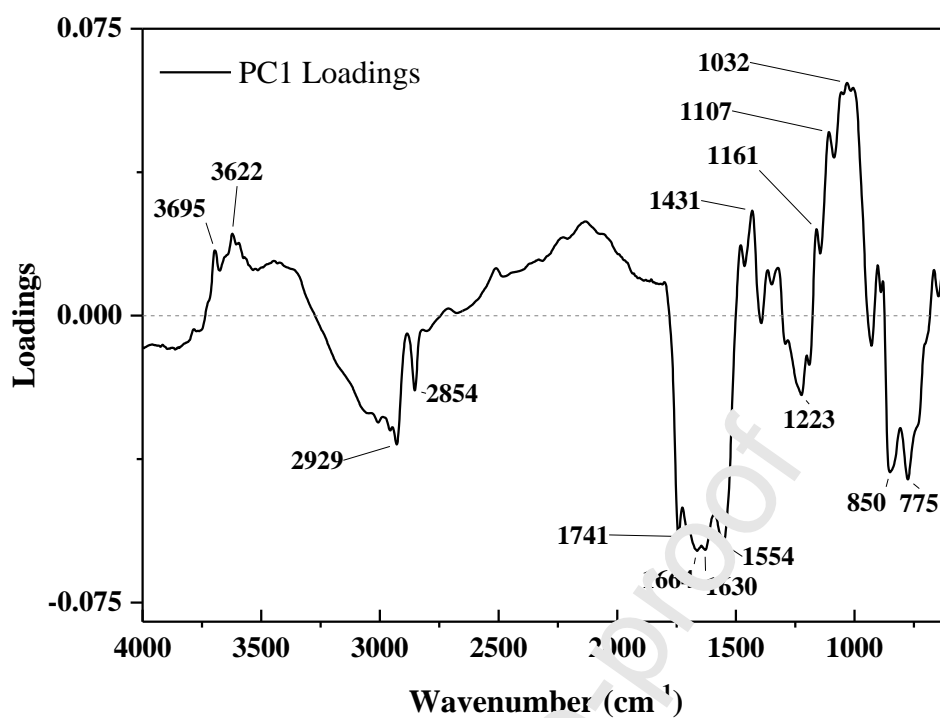


Fig. 2. Loadings of the first principal component (PC1) (explaining 65 % of the spectral variance) in a PCA of UOW samples (see Table 1 for details). Peaks of interest are marked with wavenumbers.

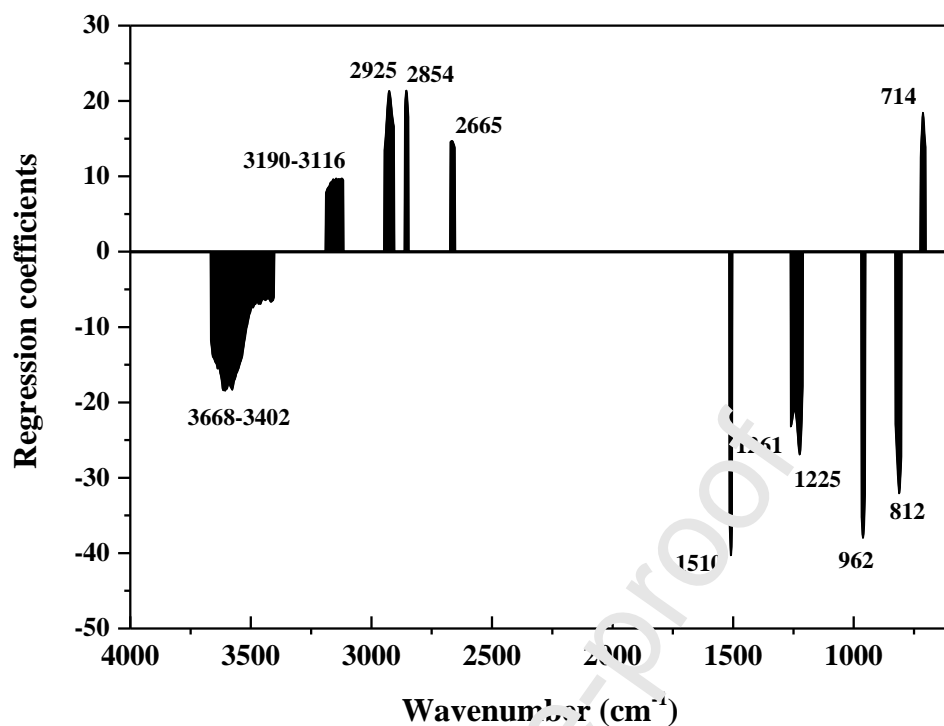


Fig. 3. Regression coefficients selected using a jack-knife variable reduction method of a PLSR model predicting BMP value of COW samples from FTIR-PAS with a pre-processing method of smoothing and normalization (Model c).

Credit author statement

Jing Huang: Conceptualization, Methodology, Writing - Original Draft, Visualization.

Georgios Bekiaris: Writing - Review & Editing. **Temesgen Fitamo:** Resources, Writing -

Review & Editing. **Charlotte Scheutz:** Conceptualization, Writing - Review & Editing.

Sander Bruun: Conceptualization, Supervision, Project administration, Writing - Review & Editing.

Journal Pre-proof

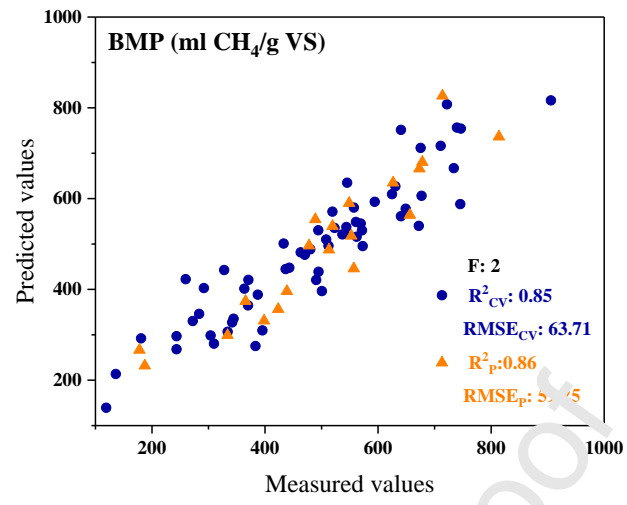
Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof

Graphical abstract



Highlights

FTIR-photoacoustic spectroscopy was able to predict the BMP of urban organic waste.

R^2 and $RMSEP$ of the urban organic waste model were 0.86 and 59.3 mL CH_4/g VS.

A positive correlation was observed between BMP and aliphatic compounds.

BMP was found to be negatively correlated with lignin and hemicellulose.

Journal Pre-proof