Prediction of organic matter accessibility and complexity in anaerobic digestates

David Fernández-Domínguez, Dominique Patureau, Sabine Houot, Nicolas Sertillanges, Bastien Zennaro, Julie Jimenez

Prediction of organic matter accessibility and complexity in anaerobic digestates

David Fernández-Domínguez[a,*], Dominique Patureau[a], Sabine Houot[b], Nicolas Sertillanges[a], Bastien Zennaro[a], Julie Jimenez[a]

[a] INRAE, Univ. Montpellier, LBE, 102 Avenue des étangs, 11100, Narbonne, France

[b] UMR ECOSYS, AgroParisTech, INRAE, Université Paris-Saclay, 78850, Thiverval-Grignon, France

* Corresponding author: david.fernandez-dominguez@inrae.fr

Abstract

Further characterization to properly assess the fate of organic matter quality during anaerobic digestion and organic carbon mineralization in soils is required. Organic matter quality based on its accessibility and complexity was employed to successfully classify 28 substrate/digestate pairs through principal components and hierarchical clustering analysis. The two first components explained 58.02% of the variability and four main groups were separated according to the feedstock type. A decrease in the accessibility (16-66%) and an increase in the complexity (34-98%) of the most accessible fractions was noticed. Besides, an increase of non-biodegradable compounds (17-66%) was globally observed after anaerobic digestion. The observed trends in the conversion of organic matter during anaerobic digestion have allowed to fill the gap in the modeling of the anaerobic digestion process chain. Indeed, partial least squares regressions have accurately predicted the organic matter quality of digestates from their inputs ($R^2 = 0.831$, $Q^2 = 0.593$) although the digester operational conditions

(temperature and hydraulic retention time) were non-explicative enough. As a novel approach, the predicted digestate quality was used to feed a partial least squares regression model previously developed to predict organic carbon mineralization in soil. The combined models have predicted experimental organic carbon mineralization in soil ($R^2 = 0.697$) with a model quality similar to the model for organic carbon mineralization in soil ($R^2 = 0.894$). This is the first study that has successfully conceived an additional step in the prediction of organic matter fate from raw substrate before anaerobic digestion to soil carbon mineralization.

Abbreviations

3D, three dimension; AD, anaerobic digestion; BMP, biochemical methane potential; C, carbon; C_bio, biodegradable carbon; COD, chemical oxygen demand; DOM, Dissolved Organic Matter; EPS, extracellular polymeric substances; FCI, fluorescence complexity index; HRT, Hydraulic Retention Time; HCA, Hierarchical Clustering Analysis; NEOM, Non-Extractable Organic Matter; NMR, nuclear magnetic resonance; OM, organic matter; PCA, Principal Components Analysis; PEOM, Poorly Extractable Organic Matter; PLS, Partial least squares; Q^2, Percent of variation of Y predicted by model in cross-validation, R^2, Correlation coefficient from PLS; REOM, Readily Extractable Organic Matter; RMSE, Root Mean Square Error, RMSE_CV, Root Mean Square Error for Cross Validation; RMSEP, RMSE calculated on validation dataset; PC1, Principal Component 1; PC2, Principal Component 2; Pf(i), fluorescence proportion for a zone (i); SEOM, Slowly Extractable Organic Matter; SPOM, Extractable Soluble from Particulate Organic Matter; T, Temperature; TS, total solids; Vf (i), fluorescence volume for a zone (i); VS, volatile solids

Keywords

biogas effluent, stability, fluorescence, waste characterization, soil

1. Introduction

The current waste management model has started to evolve towards more sustainable and resource recovery strategies (Fonoll et al., 2016; Vidal-Antich et al., 2021). Anaerobic Digestion (AD) is a biological process widely used to convert the organic

matter (OM) present in different wastes into methane (Fernández-Domínguez et al., 2020; Vinardell et al., 2021), along with the production of both OM and nutrient-rich by-product called digestate (Fernandez-Bayo et al., 2018; Guo et al., 2018). Nowadays, digestates represent alternative fertilizers used in agriculture either as organic amendment or fertilizer depending on the process, post-treatment and substrate type (Akhiar et al., 2017; Guilayn et al., 2020). However, digestate efficiency as organic amendment mainly depends on their OM stability (Kögel-Knabner, 2002), which remains a topic of ongoing research.

A need for accurate OM characterization added to the strict limitations by legislation on contaminants, such as heavy metals for the agricultural reuse of specific feedstock, are points of main consideration to enhance digestate management (Khakbaz et al., 2020). Digestate stability has to be properly assessed before land application (Tambone et al., 2013; Maynaud et al., 2017). Digestates often acquire higher biological stability and nutrient availability (particularly nitrogen) than raw substrates, enhancing the interest for land use (Provenzano et al., 2011). During AD, recalcitrant compounds are concentrated since the labile organic structures are preferentially degraded (Insam et al., 2015). Nonetheless, this stabilization could occur due to the interaction of different factors, including (i) degradation and solubilization of simple compounds, (ii) molecular complexification, (iii) complex microbial-related products release from biomass growth and decay (Aemig et al., 2016).

For a suitable agronomic valorization of digestates, evaluating the OM fate during AD and soil carbon (C) mineralization is a meaningful aspect to reduce associated environmental impacts (e.g. C loss, greenhouse gases emission) or favor ecosystem service like soil C sequestration that contributes to climate change mitigation (Minasny et al., 2017). Thus, wider efforts are needed to better understand how OM from non-

digested substrate is transformed into digestate OM (Shakeri Yekta et al., 2019; Tambone et al., 2015).

Standard methods to measure the biodegradability of organic wastes through biodegradability assays are laborious and time-consuming. At least, 30 and 90 days are required for anaerobic biodegradation potential tests and aerobic biodegradation during soil incubations, respectively (Jimenez et al., 2017). Hence, biochemical fractionation is a much more time-saving method to characterize the structural nature of organic wastes and assess waste biodegradability using successive chemical extractions (Teglia et al., 2011a). The Soest (1963) fractionation method was widely applied to associate OM composition with biodegradability (Fernandes et al., 2009; Gunaseelan, 2007; Triolo et al., 2012). Based on this fractionation method, the Iroc indicator (i.e. indicator of residual organic C in soils) has been developed as a proxy of the potentially remaining OM after application in soils for various organic amendments (Lashermes et al., 2009). Moreover, it was successfully used to predict the long term evolution of OM in soils after repeated applications of these organic amendments (Levavasseur et al., 2021). However, biodegradable substrates and digestates were not included in the panel of OM used for the development of the Iroc indicator. In addition, the Soest fractionation method presented limitations in predicting the biodegradability for a wide range of organic residues using biomethane potential tests (BMP) (Mottet et al., 2010). Therefore, less aggressive methods were requested as a more representative way for anaerobic biodegradability prediction (Bareha et al., 2018; Mottet et al., 2013).

With this purpose, Jimenez et al. (2015) suggested a less aggressive fractionation method coupled with 3D fluorescence on the extracted fractions. These authors successfully classified sixty organic residues based on their OM quality, which is defined by the OM accessibility (i.e. compounds availability to microorganisms for intra

or extra-cellular degradation) and complexity (i.e. molecules structure) of organic waste. This methodology is a promising approach compared with single spectroscopic techniques. Indeed, it has allowed an accurate assessment of the OM quality (Muller et al., 2014; Zhang et al., 2019) and has defined new inputs for AD modeling approaches (Jimenez et al., 2020). Interestingly, in Jimenez et al. (2017), the OM quality of 82 samples, comprising organic waste of different origin, was used to accurately predict both BMP and potential C mineralization in soil using Partial Least Square (PLS) regression. Similarly, Bareha et al. (2018) have properly evaluated the correlation between substrates organic nitrogen accessibility indicators with C biodegradability using an extracellular polymeric substances (EPS) fractionation method modified from Jimenez et al. (2015) in PLS regressions. The accessibility and complexity of a treated substrate will probably vary during AD. However, no studies have been focused before on the prediction of digestate OM quality from their corresponding inputs. This approach would make it possible to further predict the potential C mineralization of digestates in soil as described in Jimenez et al. (2017). Therefore, models could be an interesting alternative to fill this gap, saving time and providing new tools to improve AD performance.

Feedstock typology stands as a key factor to understand the final OM quality of digestates (Rocamora et al., 2020). Indeed, Fourier-transform infrared spectroscopy (FTIR) analysis showed that the main spectroscopic features in digestates depend on the composition of the initial biomass (Provenzano et al., 2011). Nevertheless, structural changes of labile and recalcitrant fractions of OM were reported comparing the nuclear magnetic resonance (NMR) spectra of different substrates and their subsequent digestates (Laera et al., 2019; Shakeri Yekta et al., 2019). However, the OM quality evolution due to AD has not been evaluated for a broad range of feedstock. Nonetheless,

it could provide relevant information for different sectors implementing AD (e.g. farms, wastewater treatment plants, solid waste treatment plants).

This study aims to assess the OM accessibility and complexity evolution during AD for a wide variety of feedstocks. To this end, 28 substrate/digestate pairs including different feedstock nature were classified using statistical analyses to evaluate how both feedstock type and AD parameters influenced OM quality in the final digestate. Finally, a model based on linear regression was proposed to predict digestate OM quality from their inputs. The predicted quality was used to feed the PLS model for C biodegradability in soil proposed by Jimenez et al. (2017) to determine the digestates organic C mineralization in soil after land spreading.

2. Material and methods

2.1. Substrate/digestate pairs

Twenty-eight substrate/digestate pairs (a total of 56 samples) were collected from different waste treatment plants in France. The selected organic substrates, including diverse types of feedstocks and origins, were anaerobically digested in laboratory or full-scale reactors under different digester operational conditions (Table 1). The studied samples comprised 6 digestate types from the digestate fertilizing-value typology reported in Guilayn et al. (2019). Each sample was characterized in terms of accessibility and complexity as assessed by the biochemical fractionation method coupled with 3D fluorescence developed in Jimenez et al. (2015) and described below. Biochemical fractionation was conducted on the freeze-dried particular matter and fluorescence on the aqueous phase.

2.2. Fractionation method

Following the reported methodology in Jimenez et al. (2017), all the fresh samples were centrifuged (ca. 18,600 g for 30 min) at 4ºC as a first step. The aqueous phase, namely dissolved organic matter (DOM), was separated. Total Solids (TS) and Volatile Solids (VS) on the raw sample as in the resulting particulate phase were determined (APHA, 2005). Then, the particulate phase was freeze-dried and ground (1 mm) and a quantity of 0.5 g of freeze-dried sample was subjected to the successive chemical extractions (30 mL of each extractant). After every extraction stage, the sample was centrifuged (18,600 g, 20 min, 4 ºC) and the liquid phase was filtered (0.45 µm) and kept for analyses. All tests were carried out in duplicate. The resulting fractions from the fractionation method, ordered from more to less accessible, were defined as: (1) Extractable Soluble from Particulate Organic Matter (SPOM), (2) Readily Extractable Organic Matter (REOM), (3) Slowly Extractable Organic Matter (SEOM), (4) Poorly Extractable Organic Matter (PEOM), (5) Non-Extractable Organic Matter (NEOM). Chemical oxygen demand (COD) and 3D fluorescence spectroscopy were performed on the four liquid extracted fractions (DOM, SPOM, REOM, PEOM). Regarding the fractionation method reproducibility, the measurement error is below 5% (data not shown).

COD was determined on the extracts in duplicate using Aqualytic® Vario COD kits (0-1500 mg O_2/L) analyzed using an ultraviolet (UV) spectrophotometer MultiDirect from Aqualytic®. Samples were diluted with deionized water if required. Units were in mgO_2/L and mg O_2/g TS for the liquid and solid phases, respectively.

2.3. Fluorescence spectroscopy analysis

The 3D fluorescence spectroscopy analyses were conducted on the obtained liquid fractions using a spectrofluorimeter Perkin Elmer LS55. The excitation wavelengths

varied from 200 to 600 nm and the values were recorded every 0.5 nm between 200 and 600 nm, with increments of 10 nm. The resulting spectra were divided into seven different zones (I-VII) corresponding to a biochemical family with common complexity level for spectra interpretation, as described by Jimenez et al. (2014) and He et al. (2011): I, protein-like (Tyrosine); II, protein-like (Tryptophane), III, protein-like (Tyrosine, Tryptophane and microbial products); IV, fulvic acid-like; V, inner filter, glycolated protein-like; VI, melanoidin-like and lignocellulose-like; VII, humic acid-like. Afterward, the obtained data were processed to calculate the fluorescence volume of each zone $V_f(i)$ and then the fluorescence proportion of each zone P_f (i) as described in Eqs. (1) and (2), respectively.

$$V_f(i)(U.A./mg\ COD \cdot L^{-1}) = V_{f\_raw}(i)/COD_{sample} \times 1/\frac{S(i)}{\sum_{i=1}^{7} S(i)} \quad \text{(Eq.1)}$$

$$P_f(i)(\%) = V_f(i)/\sum_{i=1}^{7} V_f(i) \times 100 \quad \text{(Eq.2)}$$

Where:

$V_f(i)$ is the normalized volume of a zone i (U.A./ mg $O_2 \cdot L^{-1}$)

V_{f_raw} is the raw fluorescence volume of a zone i (U.A./ mg $O_2 \cdot L^{-1}$)

COD_{sample} is the COD concentration of the sample (mg $O_2 \cdot L^{-1}$)

S(i) is the area of a zone i (nm^2)

$P_f(i)$ is the fluorescence proportion of a zone i (%)

Moreover, a fluorescence complexity index (FCI) was calculated between the proportions of fluorescence volumes of the most complex molecules zones (IV-VII) and the less complex molecules zones (I-III), as stated in Eq. (3) (Aemig et al., 2016).

$$FCI = \frac{\sum_{i=4}^{7} V_f(i)}{\sum_{j=1}^{3} V_f(i)} \quad \text{(Eq.3)}$$

2.4. Statistical analysis

Principal component analysis (PCA), Hierarchical Clustering Analysis (HCA) and their corresponding plots were carried out in R (4.0.3) (R Development Core Team, 2021). The displayed groups for the PCA individuals were determined by HCA clustering. Both analyses were conducted on a total of 56 samples (28 substrate/digestate pairs) considered individually. A Euclidean distance matrix with center-scaled variables was calculated before HCA analysis and the number of groups was determined heuristically. PCA was performed on center-scaled variables. A total of 34 variables for each sample was evaluated. Six variables were related to the accessibility of the sample and corresponded to the biochemical fractions (DOM, SPOM, REOM, SEOM, PEOM, NEOM) and 28 variables were linked to the complexity of the sample (seven fluorescence variables for each extracted fraction). Tukey's methodology was followed for the displayed boxplots.

In order to determine a linear prediction model of digestate quality from substrates quality, PLS regressions were performed using SIMCA software from UMETRICS. In its simplest form, a linear model specifies the linear relationship between a dependent (response) variable Y, and a set of X predictor variables, the X's. Cross-validation was then performed to test the model's quality. K-fold cross-validation was used. The database was divided into 7 blocks and some samples were selected as validation samples. This step was repeated. The mean and standard deviation of the scores were calculated to estimate the bias and the variance of validation performance. The validation step was performed on 5 samples not used in the calibration step among the 28 samples, chosen as representative of each type of digestate: Sludge_2_D, PS + Cow Food_D, PS + Cow Manure_D, MW4_D and CM3_D.

The parameters from the PLS models used to assess model quality are the following:

Correlation coefficient: R^2.

Root Mean Square Error (RMSE): used as an accuracy measurement of differences between predicted values and measured model values.

RMSE_CV: which is the RMSE for the cross-validation and was applied as prediction model error.

RMSEP: RMSE calculated on the validation dataset.

Q^2: percentage of variation of Y predicted by the PLS model according to cross-validation. This parameter indicates how well the model predicts the data. A large Q^2 (>0.5) indicates good predictivity.

The PLS model was based on 33 variables: the fractionation percentage of COD (DOM+SPOM, REOM, SEOM, PEOM, NEOM) and the fluorescence percentage of the seven zones coming from each fraction. The PLS model coefficients are presented in Supplementary Material. The same X-Variables as Jimenez et al. (2017) model have been used. Thus, DOM and SPOM fractions have been pooled and only SPOM fluorescence percentage has been considered. To combine the PLS model for C biodegradability in soil from Jimenez et al. (2017) and the digestate quality PLS model, experimental data obtained from soil C mineralization tests as described in Jimenez et al. (2017) have been used. Fourteen digestates have been tested among the 28 samples previously defined. Table 2 presents the digestates considered and their respective biodegradable C percentage.

3. Results and Discussion

3.1. Substrates and digestates classification

PCA and HCA analyses were conducted on 33 variables describing the accessibility and complexity of the OM for the 56 samples considered individually. Scores and loadings from PCA are presented in Figure 1a and 1b, respectively. HCA clusters are illustrated in Figure 1c. The PCA analysis has showed that the first two components explained 58.02% of the total variance, meaning that the samples were rather well described by the used characterization data. These results are in accord with previous studies focusing on organic waste classification (Jimenez et al., 2015) or the characterization of post-treated digestates (Maynaud et al., 2017). From the HCA, four main groups were identified by different colors (A, B, C, D) in PCA and HCA plots (Figure 1a and 1c).

The Component 1 (PC1) was described by the complexity of the molecules, from the simplest samples on the right part of the loadings plot (mainly fluorescences zones II-III, related to simple sugars/proteins and microbial products with low complexity) to the most complex samples on the left part (mainly fluorescences zones VI-VII, linked to humic acids, lignocellulose and melanoidin). Hence, PC1 was significantly (p-value < 0.05) correlated to simple and intermediate fluorescence zones, such as REOM_I_II, SEOM_I_II and PEOM_I_IV variables (R^2 = 0.77, 0.67 and 0.66, respectively). Meanwhile, it was anti-correlated (p-value < 0.05) to complex fluorescence zones (VI-VII), particularly REOM_I_VII, REOM_I_VI and SEOM_I_VI variables (R^2 = -0.94, -0.90 and -0.87, respectively).

The second component (PC2) was explained by the complexity added to the accessibility of the sample. A significant (p-value < 0.05) and positive correlation was observed between REOM_C (R^2 = 0.74) and fluorescence zones from I to III (SPOM_I_III, REOM_I_I and SEOM_I_I, with R^2 = 0.78, 0.67 and 0.65, respectively). Contrarily, a significant (p-value < 0.05) and negative correlation was defined between PEOM_C (R^2 = -0.79) and fluorescence zones of intermediate complexity as

278 REOM_I_IV, SEOM_I_IV, SEOM_I_V (R^2 = -0.95, -0.85, and -0.85, respectively).
279 Biochemical fractionation variables were among the main variance contributors for PC2
280 which allowed to identify the most accessible samples on the top (essentially due to
281 REOM) from the less accessible samples (mainly PEOM) on the bottom. DOM fraction
282 was not correlated with SPOM/REOM/SEOM nor to PC1 but to a lesser extent (R^2 = -
283 0.59 and p-value <0.05). SEOM was positively correlated with the most accessible
284 fractions (SPOM and REOM) while PEOM was negatively correlated with REOM and
285 SEOM, as previously stated by Jimenez et al. (2015) and Aemig et al. (2016). Zhang et
286 al. (2019) also showed that SEOM fraction shared protein-like compounds fluorescent
287 peaks with SPOM and REOM fractions whereas PEOM presented a humic-like
288 compound peak as the main peak. Indeed, NMR spectroscopy on the solid fraction
289 showed that SEOM was mainly composed of proteins (Laera et al., 2019). NEOM
290 fraction was non-explicative enough in this study.

291 Figure 1c shows the four groups established from the HCA analysis based on the
292 extracted fractions and their complexity of each sample: (A) pig slurry and slurry
293 mixtures with primary sludge, agro-industrial waste or biowaste, (B) manure, fibers and
294 municipal solid waste, (C) pig slurry mixtures with fiber or food wastes, and (D) sludge.
295 All the groups were strictly related to the feedstock type, highlighting its influence on
296 organic waste classification, which is supported by other authors (Akhiar et al., 2021;
297 Bareha et al., 2018; Guilayn et al., 2019). Groups B, C, and D came from a different
298 main cluster than Group A, probably due to a higher DOM content related to complex
299 fluorescence zones (VI-VII) that could arise from the degradation of refractory
300 compounds of other fractions (Zhang et al., 2019). The high reported complexity in
301 animal slurries could be related to particular recalcitrant alkyl-C (e.g. sterols, lipids,
302 cutin) (Tambone et al., 2019). Group B was defined by PEOM (poorly accessible C)

303 and fluorescence zones IV and V. Indeed, these types of substrates are characterized to
304 present high C/N and TS besides complex proteins and humic/fulvic acids (Akhiar et
305 al., 2017). Finally, Groups C and D were related to low complexity zones (I-III) while
306 sludge showed higher fluorescence proportions on zones III than pig slurry mixtures. In
307 fact, zone III corresponds to protein content and microbial by-products, typically related
308 to activated sludge metabolism and growth/decay (Fang et al., 2015). These results have
309 been also reported by Zhang et al. (2019), who showed that the characteristic presence
310 of protein-like organics (zones II and III) was in the SPOM and REOM fractions of
311 sludge.

312 Interestingly, a broad variance on pig slurry mixtures clustering from fluorescence
313 zones I-III to VI-VII was found depending on the co-substrate added (e.g. pairs of pig
314 slurry + fiber co-substrate were clustered in Group C but spatially distributed close to
315 Group B, which is mainly composed of fiber-rich samples). Notwithstanding that the
316 percentage of co-substrate in raw mass was mainly below 20%, the addition of a co-
317 substrate seems to influence their classification and should be considered to properly
318 classify organic wastes (see Table 1).

319 Most of the digestates were clustered together with their substrates, confirming that the
320 main OM complexity and accessibility prevailed after AD (Provenzano et al., 2014).
321 Nevertheless, within the same cluster, variations in the classification between substrates
322 and digestates have been observed. Since the OM accessibility and complexity
323 conversion after AD remains unclear, it is discussed in the following section.

324 **3.2. OM accessibility and complexity: revealed groups characterization**

325 ***3.2.1. OM accessibility and fluorescence complexity index of substrates***

326 The proportion of total COD in the biochemical fractions and the FCI (defined in
327 Section 2.3) of substrates profiles have been evaluated to define each group and are
328 summarised in Figure 2.

329 The proportions of COD for each extracted fraction varied between groups according to
330 the different origins of the sample (Figure 2). The largest proportion of COD for Group
331 A was in the DOM fraction (30-55%), indicating a notable content of water-soluble
332 organic substances in slurries, such as simple sugars (e.g. sucrose, glucose and
333 fructose), proteins (mainly globular protein), volatile fatty acids or soluble recalcitrant
334 compounds. High DOM fraction content was also reported by Laera et al. (2019) (76%
335 of total COD) for a substrate from a household, slaughterhouse and industrial waste co-
336 digestion plant. For Groups B and C, PEOM fraction was the main COD fraction
337 extracted and ranged between 22 and 58% of total COD. Indeed, these results are
338 consistent since the PEOM fraction is rich in hemicellulose, cellulose-like, starch and
339 certain proteins (Laera et al., 2019), which characterizes the fibrous feedstocks of
340 Groups B and C. Finally, Group D extracted COD was dominated by the NEOM
341 fraction (23-33%). Similar values of NEOM (30-40%) have been reported by different
342 authors for sludge samples (Jimenez et al., 2015; Maynaud et al., 2017).

343 Comparing the accessibility between groups, Group D had the highest SEOM (27%)
344 and SPOM (22%) fractions probably due to high complex protein content (mainly
345 fibrous proteins) and simple sugars/proteins in sludge samples. Group C presented the
346 highest REOM (25%) and NEOM (39%) fractions probably because of the protein and
347 lipid content provided by food wastes. Group B presented the highest PEOM fraction
348 (58%), while the NEOM fraction was the second extracted fraction for all the groups
349 (except for Group D), indicating that all wastes had a considerable amount of non-

extractable OM. Whilst, SPOM and REOM fractions showed low percentages of the total extracted COD for all the groups.

Regarding the FCI of the substrates, Groups A and B had higher FCI in all fractions compared with Groups C and D (Figure 2). Indeed, Group C and D were mainly characterized by fluorescence regions I to III likely they have less complex proteins/lignified compounds than Groups A and B. Only the FCI of the SPOM fraction of Group C (1.03-1.50) was higher than the one of the other Groups. The highest FCI for all groups was for Groups A and B and corresponded to the less accessible fractions (SEOM and PEOM), as reported by Muller et al. (2014). Concerning the FCI of each group, the FCI of the SEOM fraction was the highest for the Group A (3.60). In fact, this fraction targets recalcitrant compounds such as humic-like acids, fulvic-like acids and complex proteins, which could explain that Group A was previously defined by complex fluorescence zones (VI-VII) (see Figure 1). The FCI of the PEOM fraction showed the highest values for Group B and D (2.94 and 1.97, respectively) whereas Group C displayed slight differences in the FCI for all the extracted fractions.

Fluorescence spectroscopy was also performed on some DOM fractions samples (PS, CM/S and HM/S + Slu, BW + CS + AI, Sludge_6 and Sludge_7). However, this data was not considered in this study because (i) all the DOM fractions in the dataset were not analyzed and (ii) Jimenez et al. (2017) did not consider it in their model. Nonetheless, the fluorescence percentage for zone I to VII of SPOM and DOM of these 6 samples were similar for the evaluated substrates. Zhang et al. (2019) reported similar results for sewage sludge. Therefore, DOM fluorescence was not relevant for an accurate prediction of the final C mineralization in soil concerning the present study (see Supplementary Material).

374 ***3.2.2. OM accessibility and fluorescence complexity index conversion after AD***

375 The anaerobic biodegradability of OM also depends on the chemical nature of the
376 compounds, therefore, the more accessible fractions will not strictly be the more
377 biodegradable fractions (Bareha et al., 2019; Mottet et al., 2010). However, simpler
378 compounds (e.g. soluble sugars or proteins) will be extracted in the most easily
379 extractable fractions, while lignocellulose-like compounds are normally present in the
380 last extracted fractions. To evaluate these considerations, Table 3 shows the proportions
381 of COD for each extracted fraction and the FCI evolution between the substrates and
382 their corresponding digestates after AD. Meanwhile, Figure 3 displays the proportions
383 of COD for each extracted fraction (described in Section 2.2) and FCI profiles for the
384 digestates samples of each group.

385 From the four groups evaluated, general assumptions can be drawn regarding the
386 influence of AD in the OM quality of digestates. The SPOM and REOM fractions have
387 decreased after AD whereas the FCI increased for all feedstock groups (except SPOM
388 and REOM complexity of Group C which decreased). Aemig et al. (2019) have also
389 reported the highest biodegradation yields for the most accessible fractions (74% and
390 69% for SPOM and REOM, respectively) after sewage sludge AD. Thus, SPOM and
391 REOM have shown the lowest percentage of extracted COD (less than 10% of total
392 COD) for all digestates. Indeed, the sum of SPOM+REOM in digestates ranged between
393 3 and 27% of the total COD samples (data not shown). Low extracted COD proportions
394 for the most accessible fractions were also found in Bareha et al. (2018) and Laera et al.
395 (2019), where the sum of SPOM, REOM and SEOM was below the 10% of total COD.
396 Similarly, the sum of SPOM + REOM only accounted for the 2.7-10.6% of the total
397 COD of digestates OM by Maynaud et al. (2017). In accordance with the present study,
398 Aemig et al. (2019) and Zhang et al. (2019) have also reported an increase of the FCI

399 for SPOM and REOM fractions since simple fluorescence peaks remarkably decreased
400 after anaerobic digestion.

401 Regarding the DOM fraction, Group A kept the highest COD content in DOM (26-38%)
402 after AD. This could suggest that, DOM fraction could contain non-biodegradable
403 compounds or that AD could produce accessible but complex molecules (Jimenez et al.,
404 2015; Lashermes et al., 2009)**,** which has also been shown during composting (Peltre et
405 al., 2011). Indeed, Zhang et al. (2019) showed that refractory organic compounds (i.e.
406 complex compounds) could be accumulated in DOM due to solubilization. In the
407 studied dataset, the DOM fraction has been displayed to be variable among the groups.
408 Actually, SPOM and DOM digestates spectra have been observed to present
409 differences, although without a significative general trend (see Supplementary
410 Material).

411 The SEOM fraction was kept similar or increased depending on the group. Laera et al.
412 (2019) have also noticed a SEOM fraction increase from 7 to 15% after AD. These
413 authors have associated it with the concentration of protein from the growth/decay of
414 microorganisms and to a preferable consumption of simple soluble compounds.
415 Similarly, the SEOM reduction was smaller compared with SPOM and REOM fractions
416 for sewage sludge and cow manure digestion, as they are expected to have more
417 complex and therefore less degradable compounds (Somers et al., 2021; Zhang et al.,
418 2019). The FCI of SEOM has tended to increase in the digestate of Groups A and C.
419 Meanwhile, the FCI of SEOM of Group B and D remained constant, whilst the SEOM
420 accessibility decreased, probably due to the biodegradation of non-fluorescent
421 compounds, as previously reported by Aemig et al. (2019).

PEOM fraction has also varied among feedstock types. The Groups with high initial PEOM fraction (B and C) has shown a decrease while an increase was observed for the Groups with low initial PEOM fraction (A and D). Similar findings were previously reported in Bareha et al. (2018) and Aemig et al. (2016). These authors have reported a PEOM fraction decrease for cow manure (i.e. rich in PEOM) and an increase for sewage sludge (i.e. poor in PEOM), respectively. Therefore, higher degradation rates of PEOM fraction (cellulose and hemicellulose) during AD could be expected when other simpler molecules are missing (Tambone et al., 2013). Biodegradable cellulose/hemicellulose in the PEOM fraction of Groups B and C have probably contributed to higher biodegradability of the PEOM fraction. However, the PEOM fraction of Groups A and D could contain non-identified complex compounds (e.g. complex proteins) that increase the recalcitrance to AD. The FCI of PEOM has increased for all groups (except for Group C). In contrast, Aemig et al. (2016) have reported no evolution on the FCI for the less accessible fractions during sewage sludge digestion. Nevertheless, a slight increase of the fluorescence intensity of humic-like compounds peaks in PEOM fraction after AD (suggesting poor biodegradability and high complexity of this fraction), was reported as a possible re-polymerization of humic acid (Tang et al., 2018; Zhang et al., 2019).

The NEOM fraction has increased for all groups after AD and represented between 25 and 63% of the total COD in digestates. This lignocellulose-type fraction is concentrated after AD due to its recalcitrance and poor biodegradability in anaerobic conditions (Usman Khan and Kiaer Ahring, 2021). Thus, digestates are expected to be enriched in recalcitrant compounds, enhancing their suitability as soil amendment compared to the raw substrates (Jimenez et al., 2017; Pognani et al., 2010; Shakeri Yekta et al., 2019; Teglia et al., 2011b).

447 The obtained results have shown that the OM conversions shared a similar pattern
448 regardless of the feedstock type: (i) an increase of complexity for the majority of the
449 fractions, (ii) a global decrease of accessibility of the most accessible fractions (SPOM
450 and REOM), and (iii) an increase of the non-extractable organic matter (NEOM)
451 fraction. Nonetheless, it should be noted that the discussed evolution of the OM quality
452 does not distinguish between the contribution due to the exogenous OM (i.e. substrate)
453 and the endogenous OM (i.e. microbial-related products). The contribution of each
454 compartment will be prioritized in forthcoming studies.

455 The accessibility conversion pattern is in agreement with Aemig et al. (2016), who have
456 also shown a decrease for the most accessible fractions (named DOM, S-EPS and RE-
457 EPS) during AD (22-65%, 49-50%, 23-35%, respectively), whereas NEOM relatively
458 increased. Similarly, cow manure AD was evaluated by Somers et al. (2021) and a
459 significant decrease in DOM, SPOM and REOM fractions of 24%, 62%, and 61%, was
460 associated with methane production and hydrolysis of organic matter. Laera et al.
461 (2019) have also noted a remarkable decrease from 76 to 28% for the DOM fraction
462 between raw and digested mixtures of household, slaughterhouse and industrial wastes.
463 Besides, an increase from 9% to 47% for PEOM + NEOM was also reported by these
464 authors. Moreover, Zhang et al. (2019) have also showed a decrease after AD from
465 46.74% to 39.42% for DOM, SPOM and REOM fractions and the FCI increased for all
466 the extracted fractions. Furthermore, the FCI of the REOM fraction was increased by
467 34% during cow manure AD (Somers et al., 2021) whereas a non-significant increase in
468 the FCI of SEOM and PEOM during sludge and cow manure AD was stated by Aemig
469 et al. (2019, 2016) and Somers et al. (2021), respectively. Therefore, general trends on
470 OM accessibility and complexity evolution during AD have been observed.

Additionally, the prediction of OM accessibility and complexity of digestates from their inputs have been further assessed.

3.3. Prediction of digestate OM quality from substrate OM quality

3.3.1. PLS model for the prediction of digestate OM quality

PLS regression was applied on 28 observations (substrate/digestate pair) split in two datasets: a calibration dataset (23 samples) and a validation dataset (5 samples). Seven models were tested (Table 4). Model n°1 used the substrate fractions DOM+SPOM, REOM, SEOM and PEOM as X-variables to predict the digestate fractions (i.e. DOM+SPOM, REOM, SEOM and PEOM). NEOM fraction was not included as this fraction is calculated by difference. Model n°2 was similar to Model n°1 with the addition of the reactor temperature (T) as X-Variable. Model n°3 was the Model n°1 with the addition of Hydraulic Retention Time (HRT) as X-Variables and Model n°4 merged Models n°2 and 3.

Model n°5 used both fractions and fluorescence percentage in each fraction measured in the substrates (32 variables) as X-variables to predict the same 32 Y-variables in the digestates. Models n°6, 7 and 8 were based on Model n°5 with the addition of the T, HRT and both variables as X-variables, respectively. Table 4 presents the quality parameters of each model. The errors of calibration, cross-validation and prediction using cross-validation methods (RMSE, RMSE_CV, and RSMEP) of each Y-variable in each model are specified in the Supplementary Material.

According to Table 4, all the models have shown good quality performances with $Q^2 >$ 0.5 and correlation coefficients of prediction were between 0.691 and 0.832. Models n°5 to 8 had better correlation coefficients than Models n°1 to 4 and smaller calibration

errors (cf. RMSE in Supplementary Material). Indeed, the addition of fluorescence variables from Model nº1 to nº5 improved the prediction of the 32 variables with 7 components. Overall, for Model nº5, R^2 was 0.831 and specific R^2 for each fraction prediction was 0.869, 0.910, 0.736 and 0.882 for respectively DOM+SPOM, REOM, SEOM and PEOM. Furthermore, the prediction error RMSEP was lower in Model n°5 for the fractions prediction than Model n°1 (see Supplementary Material).

The Model n°2 and 4 had the best performances among the models using only fractions and operational parameters as X-Variables. Indeed, the addition of T (Model n°2) and the addition of HRT (Model n°3) as X-variables increased R^2Y but needed a component addition. However, HRT addition did not improve the quality parameters R^2Y and Q^2 (Model n°4 in Table 4). This trend was not the same in the case of Models n°6 to 8 in comparison with Model n°5 without operational parameters. Indeed, despite a little increase of R^2Y when T was added as X-variables (Model N°6), the prediction accuracy (Q^2) was decreased when HRT and T were added. The results obtained in Model n°2 have shown that T was anti-correlated (p-value < 0.05) with SEOM (R^2 = -0.593) and, less significantly with DOM+SPOM (R^2 = -0.313). However, reactors T were mainly mesophilic except for MW digestates (55 °C). These digestates were associated with lower DOM+SPOM and SEOM fractions than the others. Consequently, relying on our dataset, the impact of T was related to the substrate type. To test the impact of T alone, it would have been interesting to get substrate/digestate pairs from various feedstocks at both mesophilic and thermophilic temperature.

Similarly, the obtained results in Model n°3 have shown that HRT was negatively correlated with REOM (R^2 = -0.628) and, positively correlated with PEOM (R^2 = 0.397). However, low HRT values were mainly found for sludge digestion whereas high HRT values were associated with cow manure digestates. Sludge digestate groups

were related to a high value of REOM and SEOM fractions whereas cow manure digestates were characterized by high values of PEOM and low values of DOM+SPOM. Again, it was not possible to distinguish the effect of HRT from substrate impact on digestate quality prediction. Moreover, in the Model n°4, weight coefficients showed that HRT and T were considered as the least important variables for quality prediction (Supplementary Material).

As the operational conditions have been determined as non-explicative enough in this study, and considering the high quality of the prediction, Model n°5 was selected. Besides, the predicted variables of digestates quality are necessary for the PLS model for C biodegradability in soil. To go further, the impact of the most significant X-variables affecting the prediction of Y-variables derived from the calculation of the weight of each variable for Model n°5 (without T and HRT) was analyzed (data not shown).

The characteristics of input OM remarkably influenced the prediction of the digestate OM characteristics, meaning that the main pattern of accessibility present in feedstock input remained in digestate OM after AD. The variables of the most complex fluorescence zones from substrate impacted significantly the digestate quality. The recalcitrant compounds contained in the substrate were preserved in their subsequent digestate, as possible humus precursors (Guilayn et al., 2020; Tambone et al., 2010). Nonetheless, for the prediction of the simpler fluorescence zones, other factors such as: (i) solubilisation/complexification of biodegradable/non-biodegradable compounds, (ii) preferential compounds degradation, (iii) hydrolysis, (iv) prevalence of recalcitrant compounds, (iv) compounds contribution from other fractions are possible hypothesis that contribute to the explanation of the resulting prediction. Therefore, this is the first approach to understand how OM quality varies with AD based on accessibility and

545 complexity for a wide range of feedstock. The addition of the T and the HRT as X-
546 variables to predict OM quality and accessibility confirmed that the operational
547 conditions were not informative enough for the studied dataset. To properly evaluate the
548 impact of the operational conditions on the prediction of digestate quality, the digestion
549 of the same substrate subjected to different T or HRT should be conceived in future
550 investigations.

551 ***3.3.2. Coupling digestate quality prediction with PLS model for carbon***
552 ***biodegradability in soil prediction***

553 To validate the digestate quality PLS model found, the PLS model for C
554 biodegradability in soil was applied on the digestate quality predicted by Model n°5.
555 Among all the samples that were used for the digestate quality model, 14 samples were
556 incubated in soil and biodegradable carbon (C_bio) was obtained after 91 days (Table
557 2).

558 First, the PLS model for C biodegradability in soil was tested. A comparison between
559 the 14 predicted values of proportion of biodegradable C in the digestates and
560 experimental data obtained through soil incubation was plotted in Figure 4a. Results
561 have shown that the PLS model for C biodegradability in soil was successfully able to
562 predict the biodegradable organic C of the 14 digestates (R^2 = 0.739) with low bias.
563 Then, Model n°5 was combined with the PLS model for C biodegradability in soil to
564 predict C_bio. Figure 4b shows that the C_bio prediction was not altered by the models'
565 combination. Indeed, the combined models were able to predict the experimental data
566 (R^2 = 0.697) with a similar bias as in Jimenez et al. (2017). The 5 validation samples
567 that were used for Model n°5 were plotted in black in Figure 4a and b. Prediction error
568 of C_bio obtained by the combined models ranged between 1% and 7%, with R^2 =

569 0.828, which represents a high quality of prediction. Therefore, the reproducibility of
570 the model was confirmed. Moreover, the model validation was performed with external
571 data not included in the dataset used for model calibration. Finally, the 28 pair samples
572 were used to compare both models' predictions as presented in Figure 4c. Results
573 showed that C_bio prediction by the combined models is quite similar to C_bio
574 prediction by Jimenez et al. (2017) (R^2 = 0.894).

575 Thus, Model nº5 provided an extra step on OM fate prediction from raw substrate to soil
576 C mineralization to better understand AD influence on OM accessibility and complexity
577 before land disposal. Additionally, the present study related multiple aspects of
578 scientific interest such as waste characterization, anaerobic transformation processes of
579 OM and soil C mineralization and supposes an innovative approach to enhance the
580 modeling of the AD process chain. Nonetheless, future actions could be addressed to
581 improve the models' application when specific substrates are subjected to different AD
582 operational conditions and soil typologies.

583

584 **4. Conclusions**

585 The prediction of digestate OM quality from their input was evaluated. PCA and HCA
586 analysis have allowed to classify 28 substrate/digestate pairs covering a wide diversity
587 of OM. This classification was based on the extracted fractions from the OM
588 (accessibility) and their complexity assessed by fluorescence. Substrates and their
589 respective digestates were clustered together according to the feedstock type.
590 Nonetheless, common trends on the conversions of OM quality were observed,
591 indicating potential for the prediction of digestate quality of the entire dataset regardless
592 of the feedstock type. Thus, this study proposed a digestate quality PLS model that

accurately predicted ($Q^2 = 0.593$) the digestates OM quality from the substrate OM characteristics. However, future investigations should be focused on subjecting the same substrate to different T or HRT to properly evaluate the impact of operational conditions on the prediction of digestate quality. The predicted digestate OM characteristics validated the prediction of their biodegradability in soils using the PLS model for C biodegradability in soil previously developed. This work performed the combination of both models. OM conversion during AD and soil C mineralization was precisely predicted using a rapid analysis indicator (biochemical fractionation and 3D fluorescence). Such combined models brought a major contribution in the modeling of the AD process chain favoring the development of decision-making tools to properly manage the digestates.

Acknowledgements

This research was financially supported by the French Occitanie Region (No. 00004786-24001447) and the French National Research Institute for Agriculture, Food and Environment (INRAE). The authors are thankful for some of the data retrieved from the DIGESTATE project, funded by ANR (Agence Nationale de la Recherche, France) under the Grant ANR-15- CE34-0003-01. The authors also wish to thank Lucía Braga for her useful remarks.

References

Aemig, Q., Chéron, C., Delgenès, N., Jimenez, J., Houot, S., Steyer, J.-P., Patureau, D., 2016. Distribution of Polycyclic Aromatic Hydrocarbons (PAHs) in sludge organic matter pools as a driving force of their fate during anaerobic digestion. Waste Manag. 48, 389–396. https://doi.org/10.1016/j.wasman.2015.11.045

618 Aemig, Q., Doussiet, N., Danel, A., Delgenès, N., Jimenez, J., Houot, S., Patureau, D., 2019.
619 Organic micropollutants' distribution within sludge organic matter fractions explains
620 their dynamic during sewage sludge anaerobic digestion followed by composting.
621 Environ. Sci. Pollut. Res. 26, 5820–5830. https://doi.org/10.1007/s11356-018-4014-7
622 Akhiar, A., Battimelli, A., Torrijos, M., Carrere, H., 2017. Comprehensive characterization of
623 the liquid fraction of digestates from full-scale anaerobic co-digestion. Waste Manag.
624 59, 118–128. https://doi.org/10.1016/j.wasman.2016.11.005
625 Akhiar, A., Guilayn, F., Torrijos, M., Battimelli, A., Shamsuddin, A.H., Carrère, H., 2021.
626 Correlations between the Composition of Liquid Fraction of Full-Scale Digestates and
627 Process Conditions. Energies 14, 971. https://doi.org/10.3390/en14040971
628 APHA, 2005. Standard Methods for the Examination of Water and Waste-Water, 22nd ed.
629 Washington, DC.
630 Bareha, Y., Girault, R., Guezel, S., Chaker, J., Trémier, A., 2019. Modeling the fate of organic
631 nitrogen during anaerobic digestion: Development of a bioaccessibility based ADM1.
632 Water Res. 154, 298–315. https://doi.org/10.1016/j.watres.2019.02.011
633 Bareha, Y., Girault, R., Jimenez, J., Trémier, A., 2018. Characterization and prediction of
634 organic nitrogen biodegradability during anaerobic digestion: A bioaccessibility
635 approach. Bioresour. Technol. 263, 425–436.
636 https://doi.org/10.1016/j.biortech.2018.04.085
637 Fang, F., Hu, H.-L., Qin, M.-M., Xue, Z.-X., Cao, J.-S., Hu, Z.-R., 2015. Effects of metabolic
638 uncouplers on excess sludge reduction and microbial products of activated sludge.
639 Bioresour. Technol. 185, 1–6. https://doi.org/10.1016/j.biortech.2015.02.054
640 Fernandes, T.V., Klaasse Bos, G.J., Zeeman, G., Sanders, J.P.M., van Lier, J.B., 2009. Effects
641 of thermo-chemical pre-treatment on anaerobic biodegradability and hydrolysis of
642 lignocellulosic biomass. Bioresour. Technol. 100, 2575–2579.
643 https://doi.org/10.1016/j.biortech.2008.12.012
644 Fernandez-Bayo, J.D., Yazdani, R., Simmons, C.W., VanderGheynst, J.S., 2018. Comparison of
645 thermophilic anaerobic and aerobic treatment processes for stabilization of green and
646 food wastes and production of soil amendments. Waste Manag. 77, 555–564.
647 https://doi.org/10.1016/j.wasman.2018.05.006
648 Fernández-Domínguez, D., Astals, S., Peces, M., Frison, N., Bolzonella, D., Mata-Alvarez, J.,
649 Dosta, J., 2020. Volatile fatty acids production from biowaste at mechanical-biological
650 treatment plants: Focusing on fermentation temperature. Bioresour. Technol. 314,
651 123729. https://doi.org/10.1016/j.biortech.2020.123729
652 Fonoll, X., Astals, S., Dosta, J., Mata-Alvarez, J., 2016. Impact of paper and cardboard
653 suppression on OFMSW anaerobic digestion. Waste Manag. 56, 100–105.
654 https://doi.org/10.1016/j.wasman.2016.05.023
655 Guilayn, F., Jimenez, J., Martel, J.-L., Rouez, M., Crest, M., Patureau, D., 2019. First
656 fertilizing-value typology of digestates: A decision-making tool for regulation. Waste
657 Manag. 86, 67–79. https://doi.org/10.1016/j.wasman.2019.01.032
658 Guilayn, F., Rouez, M., Crest, M., Patureau, D., Jimenez, J., 2020. Valorization of digestates
659 from urban or centralized biogas plants: a critical review. Rev. Environ. Sci.
660 Biotechnol. 19. https://doi.org/10.1007/s11157-020-09531-3
661 Gunaseelan, V.N., 2007. Regression models of ultimate methane yields of fruits and vegetable
662 solid wastes, sorghum and napiergrass on chemical composition. Bioresour. Technol.
663 98, 1270–1277. https://doi.org/10.1016/j.biortech.2006.05.014
664 Guo, X., Li, C., Zhu, Q., Huang, T., Cai, Y., Li, N., Liu, J., Tan, X., 2018. Characterization of
665 dissolved organic matter from biogas residue composting using spectroscopic
666 techniques. Waste Manag. 78, 301–309. https://doi.org/10.1016/j.wasman.2018.06.001
667 He, X.-S., Xi, B.-D., Wei, Z.-M., Jiang, Y.-H., Yang, Y., An, D., Cao, J.-L., Liu, H.-L., 2011.
668 Fluorescence excitation–emission matrix spectroscopy with regional integration
669 analysis for characterizing composition and transformation of dissolved organic matter
670 in landfill leachates. J. Hazard. Mater. 190, 293–299.
671 https://doi.org/10.1016/j.jhazmat.2011.03.047

672 Insam, H., Gómez-Brandón, M., Ascher, J., 2015. Manure-based biogas fermentation residues –
673 Friend or foe of soil fertility? Soil Biol. Biochem. 84, 1–14.
674 https://doi.org/10.1016/j.soilbio.2015.02.006
675 Jimenez, J., Gonidec, E., Cacho Rivero, J.A., Latrille, E., Vedrenne, F., Steyer, J.-P., 2014.
676 Prediction of anaerobic biodegradability and bioaccessibility of municipal sludge by
677 coupling sequential extractions with fluorescence spectroscopy: Towards ADM1
678 variables characterization. Water Res. 50, 359–372.
679 https://doi.org/10.1016/j.watres.2013.10.048
680 Jimenez, J., Aemig, Q., Doussiet, N., Steyer, J.-P., Houot, S., Patureau, D., 2015. A new organic
681 matter fractionation methodology for organic wastes: Bioaccessibility and complexity
682 characterization for treatment optimization. Bioresour. Technol. 194, 344–353.
683 https://doi.org/10.1016/j.biortech.2015.07.037
684 Jimenez, J., Lei, H., Steyer, J.-P., Houot, S., Patureau, D., 2017. Methane production and
685 fertilizing value of organic waste: Organic matter characterization for a better prediction
686 of valorization pathways. Bioresour. Technol. 241, 1012–1021.
687 https://doi.org/10.1016/j.biortech.2017.05.176
688 Jimenez, J., Charnier, C., Kouas, M., Latrille, E., Torrijos, M., Harmand, J., Patureau, D.,
689 Spérandio, M., Morgenroth, E., Béline, F., Ekama, G., Vanrolleghem, P.A., Robles, A.,
690 Seco, A., Batstone, D.J., Steyer, J.-P., 2020. Modelling hydrolysis: Simultaneous versus
691 sequential biodegradation of the hydrolysable fractions. Waste Manag. 101, 150–160.
692 https://doi.org/10.1016/j.wasman.2019.10.004
693 Khakbaz, A., Nobili, M.D., Mainardis, M., Contin, M., Aneggi, E., Mattiussi, M., Cabras, I.,
694 Busut, M., Goi, D., 2020. Monitoring of heavy metals, eox and las in sewage sludge for
695 agricultural use: a case study. Detritus 160. https://doi.org/10.31025/2611-
696 4135/2020.13993
697 Kögel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial
698 residues as inputs to soil organic matter. Soil Biol. Biochem. 34, 139–162.
699 https://doi.org/10.1016/S0038-0717(01)00158-4
700 Laera, A., Shakeri Yekta, S., Hedenström, M., Buzier, R., Guibaud, G., Dario, M., Esposito, G.,
701 van Hullebusch, E.D., 2019. A simultaneous assessment of organic matter and trace
702 elements bio-accessibility in substrate and digestate from an anaerobic digestion plant.
703 Bioresour. Technol. 288, 121587. https://doi.org/10.1016/j.biortech.2019.121587
704 Lashermes, G., Nicolardot, B., Parnaudeau, V., Thuriès, L., Chaussod, R., Guillotin, M.L.,
705 Linères, M., Mary, B., Metzger, L., Morvan, T., Tricaud, A., Villette, C., Houot, S.,
706 2009. Indicator of potential residual carbon in soils after exogenous organic matter
707 application. Eur. J. Soil Sci. 60, 297–310. https://doi.org/10.1111/j.1365-
708 2389.2008.01110.x
709 Levavasseur, F., Mary, B., Houot, S., 2021. C and N dynamics with repeated organic
710 amendments can be simulated with the STICS model. Nutr. Cycl. Agroecosystems 119,
711 103–121. https://doi.org/10.1007/s10705-020-10106-5
712 Maynaud, G., Druilhe, C., Daumoin, M., Jimenez, J., Patureau, D., Torrijos, M., Pourcher, A.-
713 M., Wéry, N., 2017. Characterisation of the biodegradability of post-treated digestates
714 via the chemical accessibility and complexity of organic matter. Bioresour. Technol.
715 231, 65–74. https://doi.org/10.1016/j.biortech.2017.01.057
716 Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A.,
717 Chaplot, V., Chen, Z.-S., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B.,
718 Hong, S.Y., Mandal, B., Marchant, B.P., Martin, M., McConkey, B.G., Mulder, V.L.,
719 O'Rourke, S., Richer-de-Forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G.,
720 Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C.-C., Vågen,
721 T.-G., van Wesemael, B., Winowiecki, L., 2017. Soil carbon 4 per mille. Geoderma
722 292, 59–86. https://doi.org/10.1016/j.geoderma.2017.01.002
723 Mottet, A., François, E., Latrille, E., Steyer, J.P., Déléris, S., Vedrenne, F., Carrère, H., 2010.
724 Estimating anaerobic biodegradability indicators for waste activated sludge. Chem. Eng.
725 J. 160, 488–496. https://doi.org/10.1016/j.cej.2010.03.059

Mottet, A., Ramirez, I., Carrère, H., Déléris, S., Vedrenne, F., Jimenez, J., Steyer, J.P., 2013. New fractionation for a better bioaccessibility description of particulate organic matter in a modified ADM1 model. Chem. Eng. J. 228, 871–881. https://doi.org/10.1016/j.cej.2013.05.082

Muller, M., Jimenez, J., Antonini, M., Dudal, Y., Latrille, E., Vedrenne, F., Steyer, J.-P., Patureau, D., 2014. Combining chemical sequential extractions with 3D fluorescence spectroscopy to characterize sludge organic matter. Waste Manag. 34, 2572–2580. https://doi.org/10.1016/j.wasman.2014.07.028

Peltre, C., Thuriès, L., Barthès, B., Brunet, D., Morvan, T., Nicolardot, B., Parnaudeau, V., Houot, S., 2011. Near infrared reflectance spectroscopy: A tool to characterize the composition of different types of exogenous organic matter and their behaviour in soil. Soil Biol. Biochem. 43, 197–205. https://doi.org/10.1016/j.soilbio.2010.09.036

Pognani, M., Barrena, R., Font, X., Scaglia, B., Adani, F., Sánchez, A., 2010. Monitoring the organic matter properties in a combined anaerobic/aerobic full-scale municipal source-separated waste treatment plant. Bioresour. Technol. 101, 6873–6877. https://doi.org/10.1016/j.biortech.2010.03.110

Provenzano, M.R., Iannuzzi, G., Fabbri, C., Senesi, N., 2011. Qualitative Characterization and Differentiation of Digestates from Different Biowastes Using FTIR and Fluorescence Spectroscopies. J. Environ. Prot. 2, 83–89. https://doi.org/10.4236/jep.2011.21009

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. https://doi.org/10.1016/j.wasman.2013.12.001

Rocamora, I., Wagland, S.T., Villa, R., Simpson, E.W., Fernández, O., Bajón-Fernández, Y., 2020. Dry anaerobic digestion of organic waste: A review of operational parameters and their impact on process performance. Bioresour. Technol. 299, 122681. https://doi.org/10.1016/j.biortech.2019.122681

Shakeri Yekta, S., Hedenström, M., Svensson, B.H., Sundgren, I., Dario, M., Enrich-Prast, A., Hertkorn, N., Björn, A., 2019. Molecular characterization of particulate organic matter in full scale anaerobic digesters: An NMR spectroscopy study. Sci. Total Environ. 685, 1107–1115. https://doi.org/10.1016/j.scitotenv.2019.06.264

Soest, V., J, P., 1963. Use of detergents in the analysis of fibrous feeds. 2. A rapid method for the determination of fiber and lignin. J. Assoc. Off. Agric. Chem. 46, 829–835.

Somers, M.H., Jimenez, J., Azman, S., Steyer, J.-P., Baeyens, J., Appels, L., 2021. Ultrasonication affects the bio-accessibility of primary dairy cow manure digestate for secondary post-digestion. Fuel 291, 120140. https://doi.org/10.1016/j.fuel.2021.120140

Tambone, F., Adani, F., Gigliotti, G., Volpe, D., Fabbri, C., Provenzano, M.R., 2013. Organic matter characterization during the anaerobic digestion of different biomasses by means of CPMAS 13C NMR spectroscopy. Biomass Bioenergy 48, 111–120. https://doi.org/10.1016/j.biombioe.2012.11.006

Tambone, F., Terruzzi, L., Scaglia, B., Adani, F., 2015. Composting of the solid fraction of digestate derived from pig slurry: Biological processes and compost properties. Waste Manag. 35, 55–61. https://doi.org/10.1016/j.wasman.2014.10.014

Tambone, F., Orzi, V., Zilio, M., Adani, F., 2019. Measuring the organic amendment properties of the liquid fraction of digestate. Waste Manag. 88, 21–27. https://doi.org/10.1016/j.wasman.2019.03.024

Tang, Y., Li, X., Dong, B., Huang, J., Wei, Y., Dai, X., Dai, L., 2018. Effect of aromatic repolymerization of humic acid-like fraction on digestate phytotoxicity reduction during high-solid anaerobic digestion for stabilization treatment of sewage sludge. Water Res. 143, 436–444. https://doi.org/10.1016/j.watres.2018.07.003

Development Core Team, R., 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Teglia, C., Tremier, A., Martel, J.-L., 2011a. Characterization of Solid Digestates: Part 1, Review of Existing Indicators to Assess Solid Digestates Agricultural Use. Waste Biomass Valorization 2, 43–58. https://doi.org/10.1007/s12649-010-9051-5

781 Teglia, C., Tremier, A., Martel, J.-L., 2011b. Characterization of Solid Digestates: Part 2,
782 Assessment of the Quality and Suitability for Composting of Six Digested Products.
783 Waste Biomass Valorization 2, 113–126. https://doi.org/10.1007/s12649-010-9059-x
784 Triolo, J.M., Pedersen, L., Qu, H., Sommer, S.G., 2012. Biochemical methane potential and
785 anaerobic biodegradability of non-herbaceous and herbaceous phytomass in biogas
786 production. Bioresour. Technol. 125, 226–232.
787 https://doi.org/10.1016/j.biortech.2012.08.079
788 Usman Khan, M., Kiaer Ahring, B., 2021. Improving the biogas yield of manure: Effect of
789 pretreatment on anaerobic digestion of the recalcitrant fraction of manure. Bioresour.
790 Technol. 321, 124427. https://doi.org/10.1016/j.biortech.2020.124427
791 Vidal-Antich, C., Perez-Esteban, N., Astals, S., Peces, M., Mata-Alvarez, J., Dosta, J., 2021.
792 Assessing the potential of waste activated sludge and food waste co-fermentation for
793 carboxylic acids production. Sci. Total Environ. 757, 143763.
794 https://doi.org/10.1016/j.scitotenv.2020.143763
795 Vinardell, S., Astals, S., Koch, K., Mata-Alvarez, J., Dosta, J., 2021. Co-digestion of sewage
796 sludge and food waste in a wastewater treatment plant based on mainstream anaerobic
797 membrane bioreactor technology: A techno-economic evaluation. Bioresour. Technol.
798 330, 124978. https://doi.org/10.1016/j.biortech.2021.124978
799 Zhang, Y., Xu, S., Cui, M., Wong, J.W.C., 2019. Effects of different thermal pretreatments on
800 the biodegradability and bioaccessibility of sewage sludge. Waste Manag. 94, 68–76.
801 https://doi.org/10.1016/j.wasman.2019.05.047
802
803
804
805
806
807
808
809
810
811
812
813
814
815

Table 1. Summary of the raw substrate/digestate pairs type, origin and digester operational conditions

Type of feedstock[a]	Digestate type[b]	Number of substrate/digestate pairs	Scale	Origin	T (°C)	HRT (d)
Cow manure	4	5	Industrial	Farms	35-41	56-75
Cow manure mixtures:		2				
CM + Straw		1	Industrial	Farms	35-37	56
50% CM/S + 23% HM/S + 27% Slu		1	Industrial	Farms	35-37	56
Straw	1	1	Lab scale	Farms	-	-
Biowaste	2	1	Industrial	Municipal solid waste plant	55	21
Biowaste mixtures:		1				
50% BW + 20% CS + 30% AIW	5	1	Industrial	Municipal solid waste plant	37	90
Municipal Waste	5	4	Industrial	Municipal solid waste plant	55	20-28
Pig Slurry	3	1	Industrial	Farms	38	60
Pig Slurry mixtures:		6				
45% PS + 40% PsluAI + 15% others	3	1	Lab scale	Farms	38	60
93% PS + 7% Cow Food	3	1	Lab scale	Farms	38	24
93% PS + 7% Horse Food	3	1	Lab scale	Farms	38	24
80% PS + 20% Maize Silage	1	1	Lab scale	Farms	38	25
62% PS + 38% CM	1	1	Lab scale	Farms	38	24
80% PS + 20% BW	3	1	Lab scale	Farms	38	21
Sludge	2	7	Industrial	Wastewater treatment plant	37	15-25

[a]Type of feedstock: AIW: agroindustrial waste; BW: biowaste; CM: cow manure; CS: cow slurry; HM: horse manure; PS: pig slurry; PSluAI: primary sludge from agroindustry; S: slurry; Slu: Sludge. [b]Digestate type based on Guilayn et al. (2019): (1) Fibrous feedstock (2) Sewage sludge, Biowaste, food agroindustrial residues (FAI) mono/co-digestion; (3) Organic fraction of municipal solid waste (OFMSW), Food waste (FW), FAI, PS mono/co-digestion; (4) Manure/other co-digestion; (5) OFMSW and BW mono/co-digestion; (6) Fibrous feedstock: Cattle manure, green waste, silage.

839 **Table 2.** Biodegradable carbon percentage after soil incubation of several studied
840 digestates coming from Jimenez et al. (2017) data

Digestate name	C_bio (%C)
Sludge2_D	39%
Sludge7_D	31%
CM3_D	15%
MW1_D	25%
MW3_D	32%
MW4_D	19%
MW5_D	27%
PS +BW_D	46%
PS + Cow Food_D	43%
PS + Cow Manure_D	24%
PS + Horse Food_D	35%
PS + Maize Silage_D	44%
BW_D	24%
PS + PrimSluAI_D	31%

841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858

Table 3. Proportions of COD for each extracted fraction and fluorescence complexity index evolution after AD in the different groups. The results display the relative percentage of the increase or decrease as expressed in Tambone et al. (2013).

	Accessibility (% in COD)						Fluorescence complexity index (-)			
Group	**DOM**	**SPOM**	**REOM**	**SEOM**	**PEOM**	**NEOM**	**SPOM**	**REOM**	**SEOM**	**PEOM**
A	--	**--**	**--**	+	++	++	++++	++	+	++
B	N.A.	**--**	**-**	-	-	++	++	++	0	+
C	++++	**---**	**--**	++	--	+	--	-	+++	-
D	--	**--**	**--**	-	+	++	++	++	0	+++

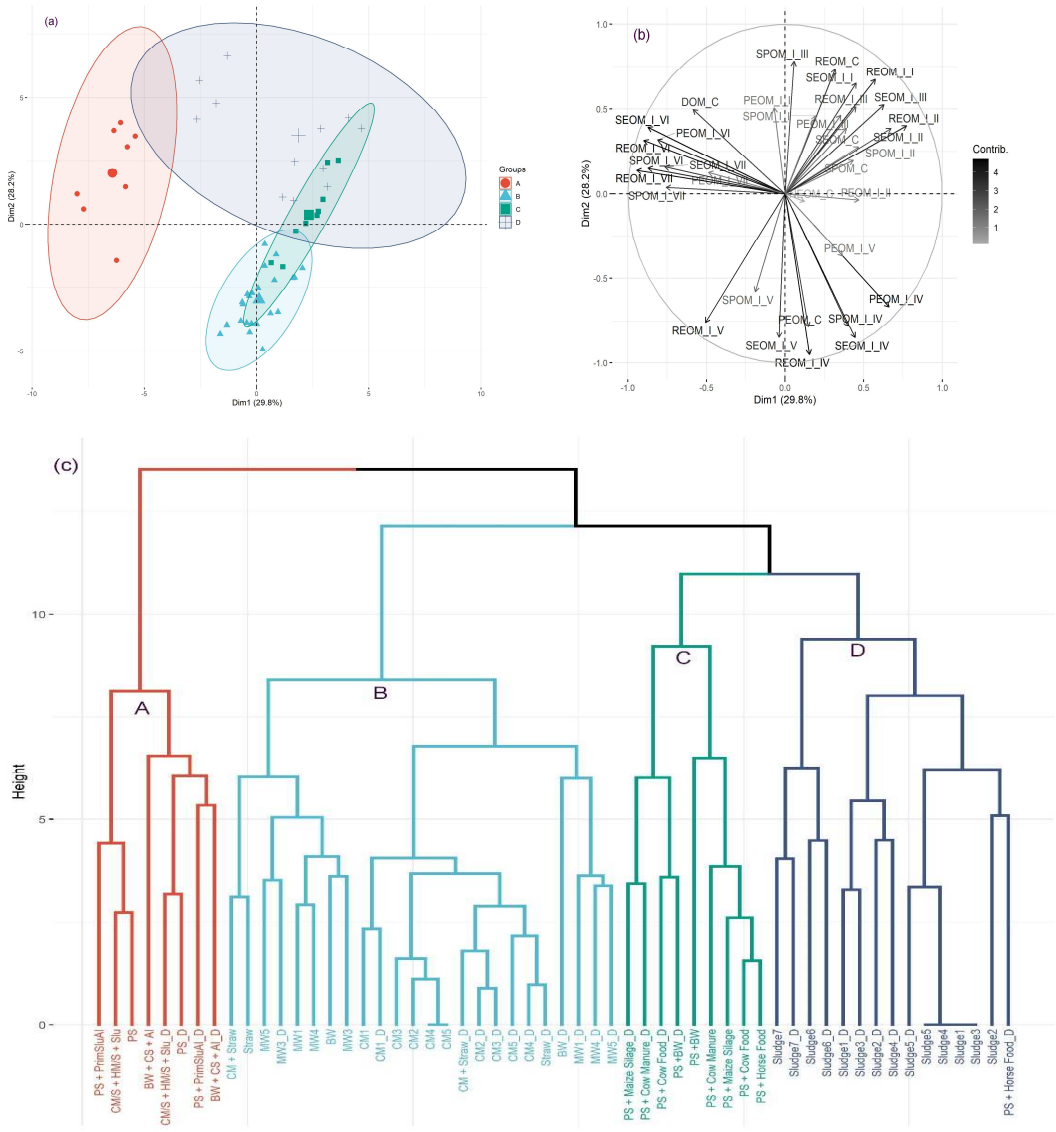
Relative conversion ranges*:* +: 0 to 25%; ++: 25 to 50%; +++: 50 to 75%; ++++: 75 to 100%; -: 0 to -25%; --: -25 to -50%; ---: -50 to -75%; ----: -75 to -100%. Relative percentage = (final value in the digestates - initial value in the substrates)/initial value in the substrates) × 100. N.A. = no presence of DOM for Group B (solid digestates)

887 **Table 4.** Quality parameters of the PLS models

Model	Variables		Model quality			
	X	Y	Components number	R^2X	R^2Y	Q^2
Model n°1	DOM+SPOM; REOM; SEOM; PEOM	DOM+SPOM; REOM; SEOM; PEOM	2	0.892	0.691	0.621
Model n°2	DOM+SPOM; REOM; SEOM; PEOM;T	DOM+SPOM; REOM; SEOM; PEOM	3	0.919	0.763	0.681
Model n°3	DOM+SPOM; REOM; SEOM; PEOM;HRT	DOM+SPOM; REOM; SEOM; PEOM	3	0.922	0.757	0.639
Model n°4	DOM+SPOM; REOM; SEOM; PEOM;HRT;T	DOM+SPOM; REOM; SEOM; PEOM	4	0.938	0.775	0.587
Model n°5	DOM+SPOM; REOM; SEOM; PEOM;Pf_i	DOM+SPOM; REOM; SEOM; PEOM;Pf_i	7	0.953	0.831	0.593
Model n°6	DOM+SPOM; REOM; SEOM; PEOM;Pf_i;T	DOM+SPOM; REOM; SEOM; PEOM;Pf_i	7	0.950	0.832	0.588
Model n°7	DOM+SPOM; REOM; SEOM; PEOM;Pf_i; HRT	DOM+SPOM; REOM; SEOM; PEOM;Pf_i	7	0.951	0.829	0.58
Model n°8	DOM+SPOM; REOM; SEOM; PEOM;Pf_i;T;HRT	DOM+SPOM; REOM; SEOM; PEOM;Pf_i	7	0.949	0.830	0.576

888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906

907



908

909 **Figure 1.** Scores plot (a) and loadings (b) obtained from the PCA analysis, and (c) HCA
910 analysis for the 56 samples studied. PCA individuals are distinguished by shape and
911 color according to the HCA revealed groups (A-D). Ellipses show 95% confidence
912 intervals. Loadings intensity color (plot b) is related to the variables contribution, from
913 low (1) to high (4). Groups: (A) pig slurry and slurry mixtures with primary sludge,
914 agro-industrial waste or biowaste, (B) manure, fibers and municipal solid waste, (C) pig
915 slurry mixtures with fiber or food wastes, and (D) sludge

916

917

918

919

920

921

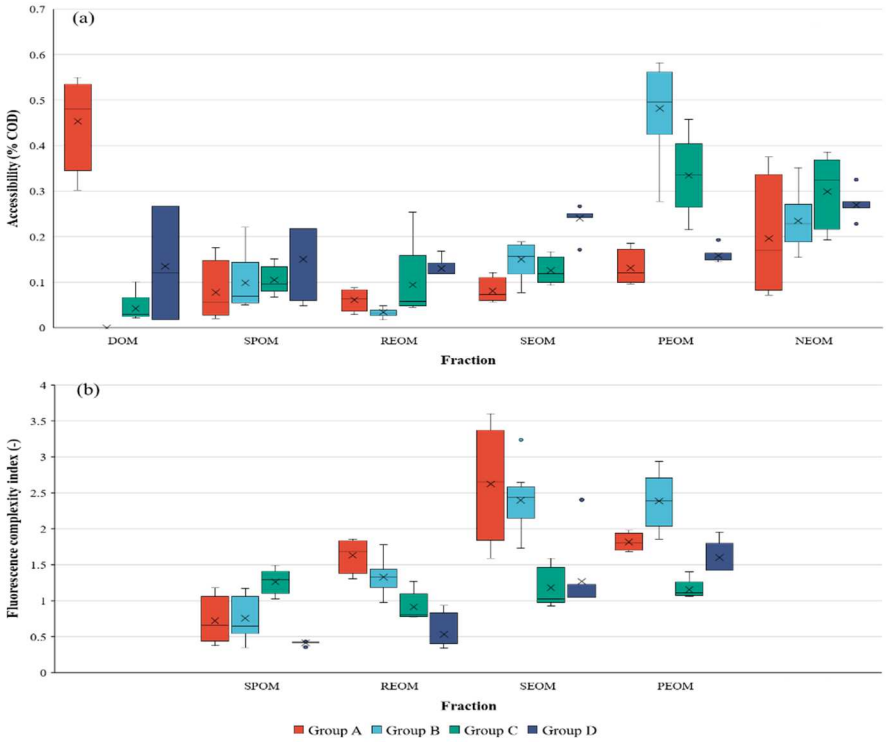


Figure 2. Accessibility (a) and fluorescence complexity index (b) boxplots of substrates for the HCA revealed groups. Groups: (A) pig slurry and slurry mixtures with primary sludge, agro-industrial waste or biowaste, (B) manure, fibers and municipal solid waste, (C) pig slurry mixtures with fiber or food wastes, and (D) sludge

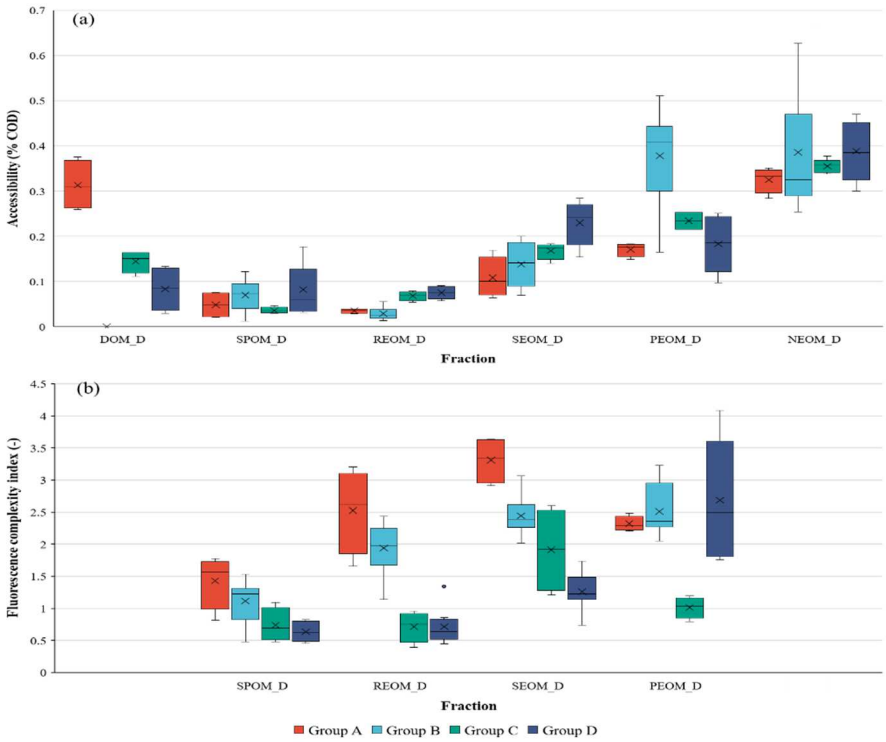
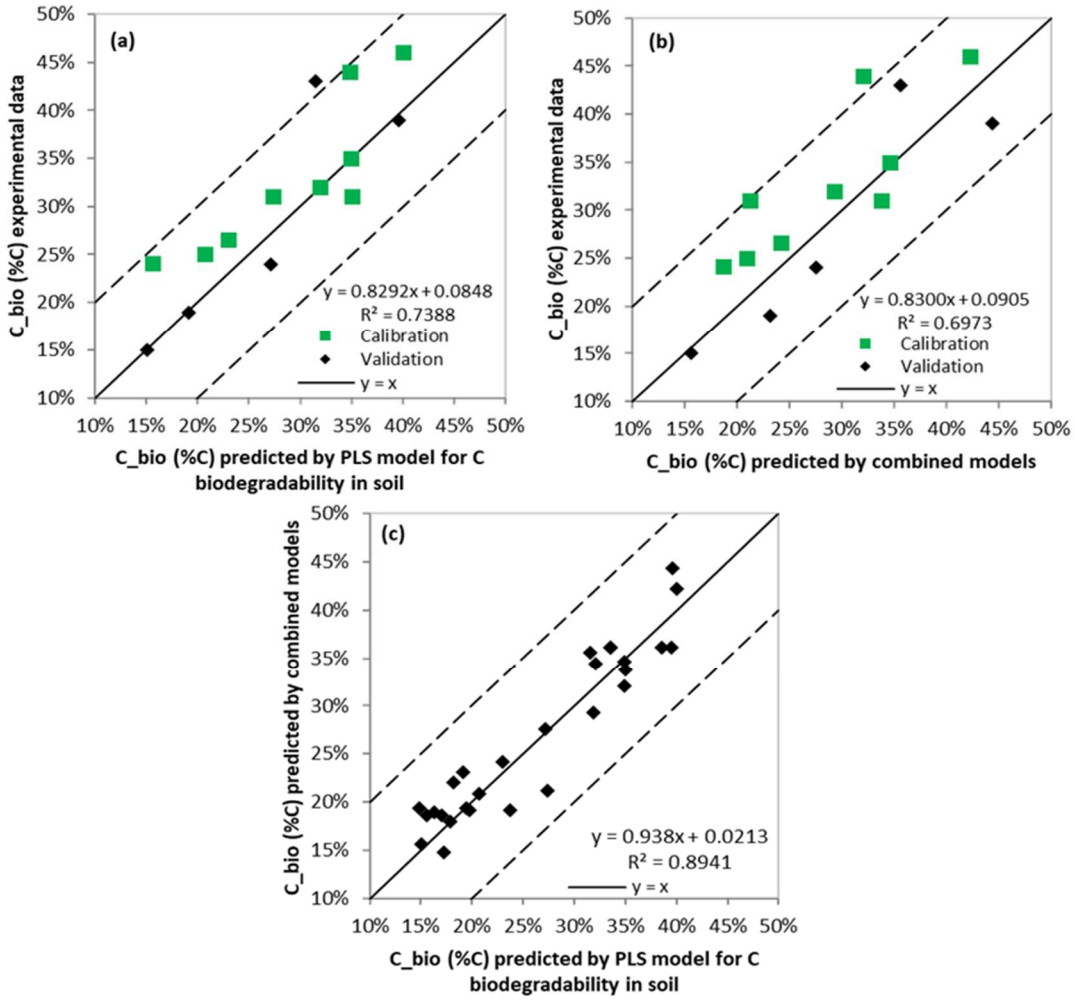


Figure 3. Accessibility (a) and fluorescence complexity index (b) boxplots of digestates for the HCA revealed groups. Groups: (A) pig slurry and slurry mixtures with primary sludge, agro-industrial waste or biowaste, (B) manure, fibers and municipal solid waste, (C) pig slurry mixtures with fiber or food wastes, and (D) sludge

971



972

973 **Figure 4.** Validation of model PLS: comparison between experimental data from
974 biodegradable organic carbon tests on soil with PLS model for C biodegradability in
975 soil (a) and combined models (b); and comparison between combined models and PLS
976 model for C biodegradability in soil (c)

977

978