Comprehensive determination of input state variables dataset required for anaerobic digestion modelling (ADM1) based on characterisation of organic substrates
H. Fisgativa, B. Zennaro, C. Charnier, C. Richard, G. Accarion, F. Béline

To cite this version:
H. Fisgativa, B. Zennaro, C. Charnier, C. Richard, G. Accarion, et al.. Comprehensive determination of input state variables dataset required for anaerobic digestion modelling (ADM1) based on characterisation of organic substrates. Data in Brief, Elsevier, 2020, 29, 10.1016/j.dib.2020.105212. hal-02495855

HAL Id: hal-02495855
https://hal.archives-ouvertes.fr/hal-02495855
Submitted on 2 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Comprehensive determination of input state variables dataset required for anaerobic digestion modelling (ADM1) based on characterisation of organic substrates

Henry Fisgativa a, Bastien Zennaro b, Cyrille Charnier c, Charlotte Richard d, Guillaume Accarion e, Fabrice Béline a,*

a IRSTEA, Rennes, France
b INRA-LBE, Narbonne, France
c BioEnTech, Lyon, France
d ENGIE, Saint-Denis, France
e Akajoule, Saint-Nazaire, France

This article contains the data of 11 organic substrates including physicochemical, biochemical and nutritional characterisations. Additionally, it includes for all substrates the data of organic matter fractionation into easily biodegradable, slowly biodegradable and inert fractions performed with anaerobic respirometry method. Finally, based on physicochemical characterisations and organic matter fractionation, a detailed methodology for the determination of input state variables required for the anaerobic digestion model N°1 (ADM1) was presented and the dataset for all substrates is provided. An example of calculation for one substrate illustrates the methodology for the determination of these variables. Data provided in this article could be useful to any person interested in modelling anaerobic digestion and particularly co-digestion. Data could be also used for implementation of a database for anaerobic digestion modelling.

© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

* Corresponding author.
E-mail address: fabrice.beline@irstea.fr (F. Béline).

https://doi.org/10.1016/j.dib.2020.105212
2352-3409/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Data

The dataset contains the characterisation and fractionation of organic substrates required as input state variable for anaerobic digestion model N°1 (ADM1). The list and the description of substrates studied in this article are presented in Table 1. Table 2 shows the main physicochemical, biochemical and nutritional characteristics of the organic substrates. These characteristics are used to determine input state variables for ADM1. Table 3 shows the fractionation of the organic matter determined using a simplified anaerobic digestion model combined with the methane production rate curves obtained from the anaerobic respirometry test. The simplified model used to generate data provided in Table 3 is described in the subsection “Organic matter fractionation” of the “Experimental Design, Materials, and Methods” section. Table 4 shows the results of the calculation of nitrogen content of inert fractions (Ni) to obtain an accurate value of input organic nitrogen (Norg). Table 5 and Table 6 provide the dataset of input state variables required for the ADM1. Table 5 provides the soluble fractions while Table 6 provides the particulates fractions. These variables were determined using the detailed methodology described in the Experimental Design, Materials, and Methods section.
2. Experimental design, materials, and methods

2.1. Physicochemical and biochemical characterisation of organic substrates

Preparation of substrates and main physicochemical and biochemical characterisation methods used in this article were described by Fisgativa et al. [1]. Briefly, pH was measured on liquid phases or extractions of substrates using a pH probe. TS, VS and COD were determined following standard methods (EN12880-12879, NF T90-101). Total TKN, total NH4⁺, total phosphorus and total potassium were determined with the standard methods (NF EN 13342, NF EN ISO 11885). VFA were determined by high pressure liquid chromatography (HPLC) on liquid phases or liquid extractions of substrates as explained by Girault et al. [2].

Biochemical methane potential (BMP) was performed using raw substrates. Inoculum was collected from a well-established CSTR running with a hydraulic retention time of 24 days and fed with horse feed and centrifuged slurry. BMP bottles were filled using a ratio of 1gVSinoculum:1gVSsubstrate. Before incubation, the head space of the BMP bottles was renewed with N2. The bottles were incubated at 38 °C for about 40 days. Gas production was recorded and gas samples were taken daily the first week and afterwards two to three times per week until the end of the experiment. Gas samples were analysed to determine CH4 and CO2 concentrations by gas chromatography using the method described by Lucas et al. [3].

Lipids content was determined on dry substrates using nuclear magnetic resonance (NMR) as described by Picard et al. [4]. Protein was determined considering the organic nitrogen content (Norg) (Norg content = TKN content – NH4⁺ content) and a ratio protein to Norg of 6.25. The calculation of COD proportion of proteins and lipids was based on methodology described by Girault et al. [2], using the following equations:

$$proteins \ (\%COD) = 100 \frac{\text{proteins} \left( \frac{g}{kgVS} \right) \cdot 1.42 \left( \frac{gO_2}{gprotein} \right)}{\text{COD} \left( \frac{gO_2}{kgVS} \right)}$$

$$lipids \ (\%COD) = 100 \frac{\text{lipids} \left( \frac{g}{kgVS} \right) \cdot 2.86 \left( \frac{gO_2}{gprotein} \right)}{\text{COD} \left( \frac{gO_2}{kgVS} \right)}$$

### Table 1

| Substrate   | Description                                                                 |
|-------------|-----------------------------------------------------------------------------|
| Slurry      | Centrifuged pig slurry                                                     |
| Greases     | Slaughterhouse flotation greases                                            |
| Sludge I    | Primary sewage sludge from municipal wastewater treatment plant             |
| Sludge II   | Secondary sewage sludge from municipal wastewater treatment plant           |
| Vinasses    | Liquid vinasses from wine distilleries                                     |
| Biowaste    | Kitchen waste and leftovers of the dishes served from a collective catering establishment |
| Feed residues | Livestock feed residues                                                   |
| Silage      | Corn silage                                                                |
| Horse feed  | Commercial horse feed                                                      |
| Manure      | Cattle manure with low straw content                                        |
| Grape marc  | Solid remains of grapes pressing from wine distilleries                    |

H. Fisgativa et al. / Data in brief 29 (2020) 105212
Table 2
Physicochemical, biochemical and nutritional characterisation mainly used to determine input state variables for ADM1.

| Substrates | pH  | TS<sup>a</sup> gTS·kgWW<sup>-1</sup> | VS  gVS·kgWW<sup>-1</sup> | CODtot gC·kgWW<sup>-1</sup> | TC  gC·kgWW<sup>-1</sup> | TIC gC·kgWW<sup>-1</sup> | TKN gN·kgWW<sup>-1</sup> | NH4<sup>+</sup> gN·kgWW<sup>-1</sup> | Ptot gP·kgWW<sup>-1</sup> | Ktot gK·kgWW<sup>-1</sup> | Lipids %COD | Proteins %COD | Carbohydrates %COD | BMP NLCH4·kgWW<sup>-1</sup> | BNP gN·kgWW<sup>-1</sup> | VFA gVFA·kgWW<sup>-1</sup> |
|------------|-----|---------------------------------|----------------|----------------|----------------|----------------|--------------------|----------------|----------------|----------------|------------|-------------|--------------|----------------|----------------|----------------|
| Slurry     | 8.60| 10.0                           | 4.4            | 5.0            | 4.1            | 1.1            | 1.9                | 1.5            | 0.10           | 1.74           | 10.7%      | 68.4%       | 21.0%        | 0.4            | 0.28           | 0.01           |
| Greases    | 5.50| 146.3                          | 128.2          | 342.9          | 84.0           | 0.4            | 2.8                | 0.3            | 0.54           | 0.25           | 64.0%      | 6.4%        | 29.5%        | 107.3          | 1.02           | 1.76           |
| I Sludge   | 5.44| 57.3                           | 46.0           | 79.1           | 25.7           | 0.7            | 1.8                | 0.1            | 0.40           | 0.20           | 26.0%      | 19.1%       | 54.9%        | 15.9           | 1.30           | 0.99           |
| II Sludge  | 5.94| 53.7                           | 42.2           | 70.4           | 21.1           | 0.3            | 3.9                | 0.2            | 1.39           | 0.58           | 13.7%      | 46.1%       | 40.2%        | 7.8            | 2.61           | 0.09           |
| Vinasases  | 4.63| 171.3                          | 126.6          | 221.0          | 73.8           | 3.6            | 5.9                | 0.1            | 0.75           | 5.97           | 15.2%      | 23.4%       | 61.4%        | 24.9           | 0.57           | 2.17           |
| Biowaste   | 5.28| 190.2                          | 172.8          | 287.7          | 93.9           | 1.1            | 6.4                | 0.1            | 1.43           | 2.74           | 40.2%      | 19.4%       | 40.5%        | 81.6           | 5.86           | 0.25           |
| Feed residues | 6.19| 802.0                          | 761.1          | 993.8          | 353.2          | 1.0            | 21.2               | 0.8            | 3.65           | 5.16           | 9.6%       | 18.3%       | 72.1%        | 283.3          | 16.65          | 0.53           |
| Silage     | 4.09| 303.0                          | 293.3          | 428.9          | 138.9          | 0.2            | 4.0                | 0.3            | 0.49           | 2.53           | 7.1%       | 7.8%        | 85.2%        | 118.6          | 1.44           | 6.98           |
| Horse feed | 5.78| 891.1                          | 801.1          | 1034           | 381.0          | 4.5            | 20.2               | 0.1            | 5.20           | 11.1           | 10.7%      | 17.3%       | 72.0%        | 288.4          | 16.14          | 0.40           |
| Manure     | 8.54| 191.3                          | 155.7          | 243.1          | 79.6           | 1.2            | 5.2                | 0.8            | 1.22           | 9.02           | 4.7%       | 16.2%       | 79.1%        | 39.6           | 0.43           | 0.00           |
| Grape marc | 4.35| 419.2                          | 395.5          | 549.4          | 211.0          | 11.1           | 8.9                | 0.0            | 4.32           | 25.0           | 22.5%      | 14.2%       | 63.2%        | 44.0           | 0.43           | 0.42           |

<sup>a</sup> Acronyms used on this table are: TS - Total solids, VS- Volatile solids, COD<sub>tot</sub> - Total chemical oxygen demand, TC - Total carbon content, TIC - Total inorganic carbon content, TKN - Total Kjeldahl nitrogen, BMP - Biochemical methane potential, BNP - Biological nitrogen potential, VFA- Volatile fatty acids, WW - Wet weight.
Finally, the COD proportion of carbohydrates was determined as the remaining COD fraction after removal of the lipids and proteins from total COD.

\[
\begin{align*}
\text{carbohydrates} \; (\% \text{COD}) &= 100 - \text{proteins} \; (\% \text{COD}) - \text{lipids} \; (\% \text{COD})
\end{align*}
\]

Table 3
Organic matter fractionation determined by anaerobic respirometry test.

| Substrates          | \(X_i\) \(\text{gO}_2\; \text{kgWW}^{-1}\) | \(S_i\) \(\text{gO}_2\; \text{kgWW}^{-1}\) | \(X_i\) \(\text{gO}_2\; \text{kgWW}^{-1}\) | \(k_b\) \(j^{-1}\) | \(\mu \cdot X_B\) \(j^{-1}\) |
|---------------------|------------------------------------------|------------------------------------------|------------------------------------------|----------------|----------------|
| Slurry              | 3.82                                     | 0.40                                     | 0.81                                     | 0.25           | 0.08           |
| Greases             | 0.00                                     | 92.7                                     | 250                                      | 0.45           | 0.15           |
| I Sludge            | 0.33                                     | 2.49                                     | 76.3                                     | 0.44           | 0.17           |
| II Sludge           | 29.8                                     | 16.2                                     | 24.5                                     | 0.27           | 0.13           |
| Vinasses            | 115                                      | 28.5                                     | 77.0                                     | 0.54           | 0.27           |
| Biowaste            | 41.0                                     | 108                                      | 139                                      | 0.18           | 0.13           |
| Feed residues       | 157                                      | 184                                      | 653                                      | 0.21           | 0.11           |
| Silage              | 75.0                                     | 19.7                                     | 334                                      | 0.37           | 0.16           |
| Horse feed          | 247                                      | 476                                      | 311                                      | 0.19           | 0.17           |
| Manure              | 146                                      | 3.27                                     | 93.4                                     | 0.08           | 0.10           |
| Grape marc          | 366                                      | 181                                      | 2.08                                     | 0.003          | 0.24           |

Table 4
Calculated values of nitrogen content of inert fractions.

| Substrates     | \(N_i\) \(\text{kmoleN kgO}_2^{-1}\) |
|----------------|--------------------------------------|
| Slurry         | 0.007                                |
| Greases        | 0.004\(^a\)                         |
| I Sludge       | 0.090                                |
| II Sludge      | 0.003                                |
| Vinasses       | 0.003                                |
| Biowaste       | 0.001                                |
| Feed residues  | 0.002                                |
| Silage         | 0.002                                |
| Horse feed     | 0.001                                |
| Manure         | 0.002                                |
| Grape marc     | 0.002                                |

\(^a\) Default value.

**2.2. Anaerobic respirometry test**

Anaerobic respirometry test was performed as described by Girault et al. [2]. Briefly, the same inoculum than for BMP was used. Cells of 1.2 L were filled with 1 L of inoculum and 5 gCOD of
Table 5
Set of ADM1 model soluble input state variables for all substrates.

| Substrates | $S_{su}$ | $S_{sa}$ | $S_{aglc}$ | $S_{va}$ | $S_{bu}$ | $S_{pro}$ | $S_{ac}$ | $S_{hn}$ | $S_{ch4}$ | $S_{tc}$ | $S_{tn}$ | $S_{t}$ | $S_{ct}$ | $S_{am}$ |
|------------|----------|----------|------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|--------|
|            | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kgO$_2$ m$^{-3}$ | kM C m$^{-3}$ | kM N m$^{-3}$ | kgO$_2$ m$^{-3}$ | kM m$^{-3}$ | kM m$^{-3}$ |
| Slurry     | 0.26     | 0.00     | 0.13       | 0.00     | 0.00     | 0.00     | 0.01     | 0.00     | 0.00     | 0.09     | 0.11     | 0.00     | 0.04     | 0.02   |
| Greases    | 27.6     | 2.80     | 59.8       | 0.00     | 0.00     | 2.07     | 0.42     | 0.00     | 0.00     | 0.04     | 0.02     | 0.00     | 0.04     | 0.02   |
| I Sludge   | 0.55     | 0.42     | 0.26       | 0.00     | 0.00     | 0.70     | 0.57     | 0.00     | 0.00     | 0.05     | 0.01     | 0.00     | 0.04     | 0.02   |
| II Sludge  | 4.08     | 10.6     | 1.39       | 0.00     | 0.00     | 0.06     | 0.05     | 0.00     | 0.00     | 0.03     | 0.02     | 0.00     | 0.02     | 0.02   |
| Vinasses   | 19.7     | 1.57     | 4.89       | 0.00     | 0.00     | 2.31     | 0.00     | 0.00     | 0.00     | 0.30     | 0.01     | 0.00     | 0.03     | 0.02   |
| Biowaste   | 40.9     | 26.2     | 40.6       | 0.00     | 0.00     | 0.00     | 0.27     | 0.00     | 0.00     | 0.09     | 0.01     | 0.00     | 0.04     | 0.02   |
| Feed residues | 129    | 37.3     | 17.2       | 0.00     | 0.00     | 0.00     | 0.57     | 0.00     | 0.00     | 0.09     | 0.06     | 0.00     | 0.05     | 0.02   |
| Silage     | 9.84     | 0.82     | 0.82       | 0.00     | 0.00     | 2.51     | 5.67     | 0.00     | 0.00     | 0.02     | 0.02     | 0.00     | 0.03     | 0.02   |
| Horse feed | 32.7     | 99.7     | 48.8       | 0.00     | 0.00     | 0.00     | 0.43     | 0.00     | 0.00     | 0.38     | 0.01     | 0.00     | 0.23     | 0.02   |
| Manure     | 2.95     | 0.15     | 0.17       | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.10     | 0.05     | 0.00     | 0.10     | 0.02   |
| Grape marc | 130      | 4.38     | 46.3       | 0.00     | 0.00     | 0.00     | 0.45     | 0.00     | 0.00     | 0.92     | 0.00     | 0.00     | 0.03     | 0.02   |
Table 6
Set of ADM1 model particulate input state variables for all substrates.

| Substrates          | $X_c$ kgO$_2$ m$^{-3}$ | $X_{ch}$ kgO$_2$ m$^{-3}$ | $X_{pr}$ kgO$_2$ m$^{-3}$ | $X_{li}$ kgO$_2$ m$^{-3}$ | $X_{su}$ kgO$_2$ m$^{-3}$ | $X_{aa}$ kgO$_2$ m$^{-3}$ | $X_{glc}$ kgO$_2$ m$^{-3}$ | $X_{c4}$ kgO$_2$ m$^{-3}$ | $X_{pro}$ kgO$_2$ m$^{-3}$ | $X_{sc}$ kgO$_2$ m$^{-3}$ | $X_{h2}$ kgO$_2$ m$^{-3}$ | $X_{i}$ kgO$_2$ m$^{-3}$ |
|---------------------|------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Slurry              | 0.00                   | 0.50                      | 0.00                      | 0.30                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.03                      |
| Greases             | 0.00                   | 76.6                      | 7.60                      | 166                       | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.11                      |
| I Sludge            | 0.00                   | 43.1                      | 12.8                      | 20.4                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.03                      |
| II Sludge           | 0.00                   | 6.3                       | 16.0                      | 2.10                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.08                      |
| Vinasses            | 0.00                   | 58.3                      | 4.20                      | 14.5                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.37                      |
| Biowaste            | 0.00                   | 52.7                      | 33.6                      | 52.3                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.03                      |
| Feed residues       | 0.00                   | 459                       | 132                       | 61.2                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.27                      |
| Silage              | 0.00                   | 296                       | 13.9                      | 24.6                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.17                      |
| Horse feed          | 0.00                   | 214                       | 65.0                      | 31.9                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.28                      |
| Manure              | 0.00                   | 84.2                      | 4.30                      | 4.90                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.29                      |
| Grape marc          | 0.00                   | 1.5                       | 0.10                      | 0.50                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.00                      | 0.60                      |

*a* $X_i$ in this table correspond to the calculate particulate inert content using Ni value from Table 4.
substrate. The renewal of the atmosphere of the cells head space was carried out with a mixture of N2 and CO2 (70/30) to avoid decarbonation of the inoculum. The cells were continuously mixed at 1200 rpm and kept at 38 °C for 17 days. Gas production rate was continuously monitored and recorded by means of pressure sensors. Gas samples were taken twice a day during the first week and once a day during the second week of the experiment. Gas samples were analyses with the same methodology than for BMP.

2.3. Organic matter fractionation

To determine the fractionation of organic matter from substrates, a simplified anaerobic digestion model was used. For this purpose, three compartments were considered:

- An easily biodegradable and directly assimilable fraction (S_s)
- A slowly biodegradable fraction requiring previous hydrolysis before assimilation (X_s)
- An inert fraction (X_i)

According to the simplified anaerobic digestion model (Fig. 1), the inert fraction is not transformed while the X_s fraction is hydrolysed to S_s and the S_s fraction is consumed by anaerobic bacteria (X_B) producing CH_4. As the inoculum to substrate ratio is high, the growth and the decay of bacteria during the trial could be neglected leading to X_B constant during the experiment.

Then the processes can be simulated using the following equations:

$$\frac{dCH_4}{dt} = -(1 - Y) \frac{dS_s}{dt}$$

$$\frac{dX_s}{dt} = - k_h X_s$$

$$\frac{dS_s}{dt} = - \frac{1}{Y (S_s + K_s)} \mu X_B initial + \frac{dX_s}{dt}$$

With K_s: half-saturation constant of S_s.

\( \mu \): bacterial growth rate.

k_h: hydrolysis rate coefficient.

Y: yield of biomass.

Using these equations and the methane production rates (MPR), corresponding to \( \frac{dCH_4}{dt} \), obtained from the anaerobic respirometry tests, the fractions X_s and S_s as well as k_h were determined for each substrate by mathematical optimization (Minimizing last-squares between experimental MPR and \( \frac{dCH_4}{dt} \) from the model combined with Monte Carlo method for parameters and fractions). For this optimization, we considered that:

- The sum of the biodegradable fractions, i.e. S_s + X_s, is equal to the BMP calculated on COD basis. Consequently, the inert fraction X_i is equal to the total COD minus the BMP.
- The fraction S_s must be equal to or greater than the VFAs.

Fig. 1. Description of the anaerobic digestion processes of the model used to fractionate organic matter: 1) Hydrolysis of X_s to S_s, 2) Consumption of S_s by bacteria (X_B) and production of CH_4.
There is still recalcitrant organic matter left in the inoculum. As a result, an additional hydrolysis process was added to the inoculum alone. This hydrolysis process is calculated from similar experiment without substrate addition.

As $X_B$ is considered constant during the experiment and equal to $X_{B,\text{initial}}$, a new parameter called $\mu X_B$ corresponding to $\mu X_{B,\text{initial}}$ was considered as each parameter was not identifiable alone.

The initial $K_s$ and $\mu X_B$ parameters used for optimization were the result of a previous optimization using kinetic MPR from similar experiments with acetate addition. The variation ranges used for optimization for these parameters were 1–10 and 1 to 3, respectively.

2.4. Determination of input variables required for ADM1 modelling

Biodegradable fractionation of organic matter obtained from anaerobic respirometry experiments Table 3 and lipid/carbohydrate/protein fractionation (Table 2) are the basis for the determination of input variables. Indeed, the fractions $S_s$ and $X_s$ must be split to serve as an input to the ADM1 model. In addition, other physicochemical parameters must be adapted for use in the model. The input variables of the model are as follows:

- The $S_s$ fraction was decomposed into the soluble materials considered by the model, namely: sugars ($S_{su}$), amino acids ($S_{aa}$), long chain fatty acids ($S_{lfca}$), acetic acid ($S_{ac}$), propionic acid ($S_{pr}$), butyric acid ($S_{bu}$) and valeric acid ($S_{va}$).
- The $X_s$ fraction was also decomposed into the slowly degradable fractions considered by the model, namely: carbohydrates ($X_{ch}$), proteins ($X_{pr}$) and lipids ($X_{li}$) slowly biodegradable.
- The remaining soluble variables were determined by characterization or by following recommended values, namely: soluble hydrogen ($S_{h2}$), soluble methane ($S_{ch4}$), inorganic carbon ($S_{ic}$), inorganic nitrogen ($S_{in}$), soluble inert particles ($S_i$), cations ($S_{cat}$) and anions ($S_{an}$).
- Also, the remaining particulate variables were determined by characterization or by following recommended values, namely: composites ($X_c$), particulate inerts ($X_{li}$) and biomass responsible for the degradation of: sugars ($X_{ssu}$), amino acids ($X_{aa}$), acetic acid ($X_{ac}$), propionic acid ($X_{pr}$), acids of 4 or more carbons ($X_{c4}$), LCFA ($X_{lfca}$) and hydrogen ($X_{h2}$).

The decomposition of the $S_s$ and $X_s$ fractions was carried out taking into account the characterisation of the substrates (Table 2). In addition, the BNP value was used in the determination of the input variables to include the notion of nitrogen biodegradability. With these elements in hand, the progressive determination of the input variables was carried according to the following steps. The example of the calculation of the input variables for the “Horse Feed” substrate is given for greater clarity:

1. Biodegradable nitrogen was calculated as follows:

$$ (X_{pr} + S_{aa}) = \frac{\text{BNP}}{N_{aa} \times M_N} $$

Where, $X_{pr}$ and $S_{aa}$ were expressed in kgO₂.m⁻³, BNP in kgN.m⁻³, $N_{aa}$ is the nitrogen content of amino acids and proteins expressed in kmoleN.kgO₂⁻¹ (0.007 according to Batstone et al. [6]) and $M_N$ is the molecular weight of nitrogen (14 kgN.kmole⁻¹). For the Horse Feed:

$$ (X_{pr} + S_{aa}) = \frac{16.1}{0.007 \times 14} = 164.7 \text{ kgO}_2 . \text{m}^{-3} $$

H. Fisgativa et al. / Data in brief 29 (2020) 105212
2. This sum \((X_{pr} + S_{aa})\) was split into \(X_{pr}\) and \(S_{aa}\) proportionally to quantity of \(X_s\) and \(S_s\) in the biodegradable fraction as follow:

\[
(X_{pr} + S_{aa}) \times \frac{X_s}{(X_s + S_s)} = X_{pr}
\]

\[
(X_{pr} + S_{aa}) \times \frac{S_s}{(X_s + S_s)} = S_{aa}
\]

In the case of Horse Feed, \(X_s\) represents 39.5\% of the organic biodegradable fraction and \(S_s\) represents 60.5\%, obtaining:

\[
X_{pr} = 164.7 \times 39.5\% = 65.0 \text{ kgO}_2 \cdot \text{m}^{-3}
\]

\[
S_{aa} = 164.7 \times 60.5\% = 99.7 \text{ kgO}_2 \cdot \text{m}^{-3}
\]

3. Then, by subtracting the value of the \(X_{pr}\) from \(X_s\), the fraction of \(X_{s-pr}\) is obtained, corresponding only to carbohydrates and fats. To obtain the fractions \(X_{ch}\) and \(X_{li}\), the distribution of fats and carbohydrates in relation to the COD without taking into account proteins, as follow:

\[
\% \text{Carbohydrates} (% \text{COD}_{\text{proteins}}) = \frac{\text{Carbohydrates} (% \text{COD})}{\text{Carbohydrates} (% \text{COD}) + \text{Lipids} (% \text{COD})}
\]

\[
\% \text{Lipids} (% \text{COD}_{\text{proteins}}) = \frac{\text{Lipids} (% \text{COD})}{\text{Carbohydrates} (% \text{COD}) + \text{Lipids} (% \text{COD})}
\]

Then:

\[
X_s - X_{pr} = (X_{ch} + X_{li})
\]

\[
(X_{ch} + X_{li}) \times \% \text{Carbohydrates} (% \text{COD}_{\text{proteins}}) = X_{ch}
\]

\[
(X_{ch} + X_{li}) \times \% \text{Lipids} (% \text{COD}_{\text{proteins}}) = X_{li}
\]

In the case of the Horse Feed, the percentage of carbohydrate COD (excluding protein) is 87.0\% and the percentage for fat is 13.0\%, therefore:

\[
X_s - X_{pr} = 310.7 - 65.0 = 245.7 \text{ kgO}_2 \cdot \text{m}^{-3} = (X_{ch} + X_{li})
\]
\[ X_{ch} = 245.7 \times 87.0\% = 213.8 \text{ kgO}_2 \text{.m}^{-3} \]

\[ X_h = 245.7 \times 13.0\% = 31.9 \text{ kgO}_2 \text{.m}^{-3} \]

At this stage, the repartition of \( X_s \) was completed.

4. The remaining fractioning of the \( S_s \) follows a similar logic. Hence, subtracting the value of \( S_{aa} \) from the \( S_s \) gives the fraction \( S_{s-aa} \) corresponding only to VFA, LCFA and sugars. The VFA content was based on the chemical characterisation. The remaining \( S_{s-aa-VFA} \) (after subtraction of VFAs) was fractionated with the same percentages of carbohydrates and lipids (% always in COD and without taking into account proteins) to obtain the \( S_{su} \) and \( S_{fa} \) fractions:

\[ S_s - S_{aa} - S_{ac} - S_{pr} - S_{bu} - S_{va} = \left( S_{su} + S_{fa} \right) \]

\[
\left( S_{su} + S_{fa} \right) \times \% \text{Carbohydrates (\%COD)} = S_{su}
\]

\[
\left( S_{su} + S_{fa} \right) \times \% \text{Lipids (\% COD)} = S_{fa}
\]

For Horse Feed, the acetic acid content is 0.4 kgO2. m\(^{-3}\). Propionic, butyric and valeric acids were not detected. The calculation would therefore be:

\[ 476.4 - 99.7 - 0.4 - 0.0 - 0.0 - 0.0 = 376.3 \text{ kgO}_2 \text{.m}^{-3} = \left( S_{su} + S_{fa} \right) \]

\[ S_{su} = 376.3 \times 87.0\% = 327.4 \text{ kgO}_2 \text{.m}^{-3} \]

\[ S_{fa} = 376.3 \times 13.0\% = 48.8 \text{ kgO}_2 \text{.m}^{-3} \]

The repartition of \( S_s \) was completed.

5. The remaining soluble variables were determined as follows:
- The variables \( S_{h2}, S_{ch4} \) and \( S_i \) were considered nil.
- The variable \( S_{ic} \) corresponds to the inorganic carbon content (0.38 kmoleC.m\(^{-3}\) for horse feed).
- The variable \( S_{in} \) corresponds to the NH\(_4^+\) content (0.01 kmoleN.m\(^{-3}\) for horse feed).
- The variable \( S_{an} \) was set at 0.02 kmole.m\(^{-3}\)
- The variable \( S_{cat} \) was calculated according to \( S_{an}, \text{pH and ions of VFA, carbonates and inorganic nitrogen as follows:} \]

\[ S_{cat} = S_{hco3} + \frac{S_{ac\_ion}}{64} + \frac{S_{pro\_ion}}{112} + \frac{S_{h4\_ion}}{160} + \frac{S_{va\_ion}}{208} + \frac{K_W}{S_{h^+}} + S_{an} - S_{in} + S_{nh3} - S_{h^+} \]
Where 64, 112, 160 and 208 are the grams of oxygen needed to degrade acetic, propionic, butyric and valeric acid respectively.

The different components of the equation have been calculated as follows:

\[ S_{h+} = 10^{-pH} \]

\[ S_{\text{hco}3} = \frac{K_{a_{\text{co2}}} \cdot S_{ic}}{(K_{a_{\text{co2}}} + S_{h+})} \]

\[ S_{\text{ac}_{\text{ion}}} = \frac{10^{-pK_{a_{\text{ac}}} \cdot S_{\text{ac}}}}{(10^{-pK_{a_{\text{ac}}} + S_{h+}})} \]

\[ S_{\text{pro}_{\text{ion}}} = \frac{10^{-pK_{a_{\text{pro}}} \cdot S_{\text{pro}}}}{(10^{-pK_{a_{\text{pro}}} + S_{h+}})} \]

\[ S_{\text{bu}_{\text{ion}}} = \frac{10^{-pK_{a_{\text{bu}}} \cdot S_{\text{bu}}}}{(10^{-pK_{a_{\text{bu}}} + S_{h+}})} \]

\[ S_{\text{va}_{\text{ion}}} = \frac{10^{-pK_{a_{\text{va}}} \cdot S_{\text{va}}}}{(10^{-pK_{a_{\text{va}}} + S_{h+}})} \]

\[ S_{\text{nh3}} = S_{\text{in}} - \frac{S_{h+} \cdot S_{\text{in}}}{(K_{a_{\text{in}}} + S_{h+})} \]

Where, the pK\textsubscript{a} values used were: acetic = 4.76 M, propionic = 4.88 M, butyric = 4.82 and valeric = 4.86 M. The value of K\textsubscript{a,co2} was 4.94E-07 and the value of K\textsubscript{a, in} was 1.11E-09.

Making these calculations, the S\textsubscript{cat} for Horse Feed corresponds to 0.23 kmole.m\textsuperscript{-3}.

6. The remaining particulate variables were determined as follows: composites (X\textsubscript{c}) and all biomasses (X\textsubscript{su}, X\textsubscript{aa}, X\textsubscript{ac}, X\textsubscript{pro}, X\textsubscript{c4}, X\textsubscript{lcf}, and X\textsubscript{h2}) were considered nil.

7. As previously mentioned for the simplified model, the inert fraction X\textsubscript{i} is equal to the total COD minus the BMP, i.e. S\textsubscript{s} + X\textsubscript{s}.

8. A nitrogen content of inert (N\textsubscript{i}) was calculated from X\textsubscript{i}. This value of the stoichiometry parameter of the N\textsubscript{i} was calculated for each substrate to be consistent with the measured N\textsubscript{org} content. Indeed, the measured N\textsubscript{org} content includes inert nitrogen in different proportions depending on the substrate, which is not taken into account by the default N\textsubscript{i} parameter (0.004). The process of calculating N\textsubscript{i} for each substrate and the variations it produces are therefore as follows:

- The biodegradable nitrogen (N\textsubscript{org, bio}) of the model was calculated without taking into account X\textsubscript{i} and S\textsubscript{i}.
- The N\textsubscript{org, bio} would therefore be calculated as the sum of the nitrogen content of the variables: S\textsubscript{aa}, X\textsubscript{c}, X\textsubscript{pr}, X\textsubscript{su}, X\textsubscript{aa}, X\textsubscript{lcf}, X\textsubscript{c4}, X\textsubscript{pro}, X\textsubscript{ac}, X\textsubscript{h2}.
- The nitrogen content for composites fraction \( X_c \) was 0.003 kmoleN.kgO\(_2\)\(^{-1}\), for amino acids and proteins fractions \( S_{aa}, X_{pr} \) was 0.007 kmoleN.kgO\(_2\)\(^{-1}\) and for bacteria \( X_{su}, X_{aa}, X_{lf}, X_{ac}, X_{pro}, X_{ac}, X_{h2} \) was 0.006 kmoleN.kgO\(_2\)\(^{-1}\).
- This \( N_{org,bio} \) without \( X_i \) or \( S_i \), was subtracted from the \( N_{org} \) measured at substrates and divided by the sum of the value of \( X_i \) of the substrates calculates by anaerobic respirometry. This relationship allows us to obtain \( N_i \) from each substrate:

\[
N_{org, measured} - N_{org, bio} = N_{org, inert}
\]

\[
\frac{N_{org, inert}}{X_i} = N_i \text{ modified}
\]

By performing these processes, the determination of the input variables of the ADM1 model was complete and could be used for modelling.

**Acknowledgments**

This study is part of the MAPPED project (2017–2019) N°1782C0076, which is supported by the French Environment & Energy Management Agency (ADEME).

**Conflicts of Interest**

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.

**Appendix A. Supplementary data**

Supplementary data related to this article can be found at https://doi.org/10.1016/j.dib.2020.105212.

**References**

[1] H. Fisgativa, C. Marcilhac, R. Girault, M.-L. Daumer, A. Trémier, P. Dabert, F. Béline, Physico-chemical, biochemical and nutritional characterisation of 42 organic wastes and residues from France, Data Brief 19 (2018) 1953–1962.

[2] R. Girault, G. Bridoux, F. Nauleau, C. Poullain, J. Buffet, J.P. Steyer, A.G. Sadowski, F. Beline, A waste characterisation procedure for ADM1 implementation based on degradation kinetics, Water Res. 46 (2012) 4099–4110.

[3] T. Lucas, D. Le Ray, P. Peu, M. Wagner, S. Picard, A new method for continuous assessment of CO2 released from dough baked in ventilated ovens, J. Food Eng. 81 (2007) 1–11.

[4] S. Picard, F. Beline, R. Girault, G. Bridoux, C. Cambert, A. Davenel, Determination of lipid fraction from organic wastes using Nuclear Magnetic Resonance (NMR): comparison to the soxhlet method. 13th World Congress on Anaerobic Digestion, Santiago de Compostela, Spain, 2014.

[5] Y. Bareha, R. Girault, J. Jimenez, A. Trémier, Characterization and prediction of organic nitrogen biodegradability during anaerobic digestion: a bioaccessibility approach, Bioresour. Technol. 263 (2018) 425–436.

[6] D.J. Batstone, J. Keller, I. Angelidaki, S.V. Kalyuzhnyi, S.G. Pavlostathis, A. Rozzi, W.T.M. Sanders, H. Siegrist, V.A. Vavilin, Anaerobic Digestion Model No.1 (ADM1), IWA Publishing, 2002.