

Methane yield optimization using mix response design and bootstrapping: application to solid-state anaerobic co-digestion process of cattle manure and damp grass

A. Coutu, S. Mottelet, S. Guérin, V. Rocher, A. Pauss, T. Ribeiro

▶ To cite this version:

A. Coutu, S. Mottelet, S. Guérin, V. Rocher, A. Pauss, et al.. Methane yield optimization using mix response design and bootstrapping: application to solid-state anaerobic co-digestion process of cattle manure and damp grass. Bioresource Technology Reports, 2022, 17, pp.100883. 10.1016/j.biteb.2021.100883 . hal-03594369

HAL Id: hal-03594369 https://hal.science/hal-03594369

Submitted on 5 Jan 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S2589014X21002619 Manuscript_29ee7b376d2040626d5be22eccf0ab2d

1 Methane yield optimization using mix response design and bootstrapping:

2 application to solid-state anaerobic co-digestion process of cattle manure

3 and damp grass

- 4 A. Coutu^a, S. Mottelet^b, S. Guerin^c, V. Rocher^c, A. Pauss^b, T. Ribeiro^{a*}
- ⁵ Institut Polytechnique UniLaSalle, Université d'Artois, ULR 7519, 19 Rue Pierre Waguet, BP 30313, 60026 Beauvais, France.
- ^bUniversité de Technologie de Compiègne, ESCOM, TIMR (Integrated Transformations of Renewable Matter), Centre de
 recherche Royallieu CS 60 319 60 203 Compiègne Cedex
- 8 ^cDirection Innovation SIAAP Service public pour l'assainissement francilien, 82 avenue Kléber 92 700 Colombes, France
- 9 *Corresponding author: Thierry Ribeiro; Tel.: +33 (0) 344 06 76 11; E-mail: thierry.ribeiro@unilasalle.fr

10 Abstract

11 Optimization of solid-state anaerobic digestion on cattle manure and damp grass were performed simultaneously and combined to a bootstrapping tool to significantly decrease the 12 13 number of experiments for a methane yield optimization. 15 batch reactors at lab scale were 14 launched two times respecting a mix surface response methodology. A numerical method 15 called Bootstrapping was used to verify results significance. Results have shown a significant influence with a p-value between 10⁻⁸ and 10⁻¹¹: the optimal parameter values depend on the 16 substrate composition with a maximal p-value of 5.60.10⁻². The methane yield reached 156.19 17 NL.kgVS-1 for a mixture of cattle manure and damp grass, and 142.92 NL.kgVS-1 for cattle 18 19 manure only. The bootstrapping were validated with a standard error lower than 3% in 20 comparison with ANOVA method, which confirms that the mix surface response methodology combined to bootstrapping is an innovative and efficient way to optimize solid-state anaerobic 21 22 digestion process.

23 Keywords

24 ANOVA; Agricultural waste; Biogas; Experimental design

25 **1** Introduction

26 The share of renewable energy in the European Union in gross final energy consumption 27 reached 18 % in 2018, twice the share of 2004, which was 8.5 %. The European Union's target is to reach 20 % of its energy from renewable sources by 2020 and 32 % by 2030. 12 countries 28 have already reached a share equal or above their 2020 binding targets but many countries 29 30 have to increase their contribution, such as France's share from 16.6 % to 23 % (Eurostat, 31 2020). Anaerobic digestion (AD) of organic wastes is part of this handle. Two types of AD 32 process exist, depending on the total solid content (TS): the liquid AD and the solid-state 33 anaerobic digestion (SS-AD). The SS-AD is defined by a solid content higher than 15 % and is 34 more appropriate for degradation of substrates with varying composition and high solid content (Rocamora et al., 2020; André et al., 2018; Degueurce et al., 2016). This process has 35 increased in recent years, but SSAD is less widespread compared to liquid AD due to the lack of 36 37 knowledge and the operational complexity to sustain a stable production (Rocamora et al., 2020). In fact, the scientific and technological hurdles of the SS-AD are numerous as non-38 39 representativeness of samples, biological process drifts, lack of monitoring indicators in solid 40 media, no hydrodynamic optimization, no comprehension of the process, nor inhibitions of the process (André et al., 2018; André et al., 2019). 41

42 Some of these obstacles as inhibition phenomena and hydrodynamic comprehension can be 43 studied using a type of design of experiments which is a Response Surface Methodology (RSM). This statistical method allows to describe and explain the effect of several independent 44 45 variables on responses with multiple linear regression rules. The ultimate goal can be the 46 determination of variables impact on the response or the optimization of the response (Goupy 47 and Creighton, 2007; Goupy, 2012). In literature, a couple of papers on the AD phenomenon use this tool with surface response designs (SRD) e.g. for the impact of pretreatment (Gunes et 48 49 al., 2021; Jin et al., 2018; Jackowiak et al., 2011), for the recirculation parameters on the SS-AD (Degueurce et al., 2016), for the average velocity optimization on account reactor 50 51 configuration parameters (Leonzio et al., 2019), for the methane yield optimization regarding 52 the Carbon to Nitrogen ratio (C/N), Food to Microbe ratio (F/M) and pH (Kainthola et al., 2020; 53 Kainthola et al., 2019; Mortezaei et al., 2018) on different reactors or other parameters 54 (Jiménez et al., 2014; Kumar et al., 2020; Lee et al., 2020). The mixture designs (MD) work in 55 the same way but consider compositions respecting the mixture properties. The main goal of 56 MD in the AD process is to study the substrate composition. Some papers use this last tool in 57 AD process (Kashi et al., 2017; Wang et al., 2013; Bassard, 2015; Rakotoniaina, 2012; Rao et al., 58 2011). These RSM studies are very interesting and allow deepening AD knowledge. Although a 59 parameter optimization is only valuable for a specific substrate composition and a composition 60 optimization is only valuable for a specific set of parameters. Some adaptations of factorial design are done in the literature (Wang et al., 2013) to consider mixture composition and 61 62 operational parameters at the same time in a SRD but there is no study about a mix design 63 between these two solutions. Moreover, these methods require a compromise between minimizing the number of experiments and to obtain results significance. To determine 64 65 experimental repeatability, several reactors must be monitored under the same conditions and the results must be processed by an analysis of variance (ANOVA). The number of runs 66 67 depends on the type of design used and the number of parameters studied. No information is available in the literature concerning the number of experiments reduction by alternative 68 69 methods as numerical methods. The bootstrapping is a method which were invented by 70 Bradley Efron (Efron, 1979; Efron and Tibshirani, 1986) and involves determining the 71 significance of the model coefficients and response by resampling residues and redefine the 72 model coefficients from a large number of iterations. A parametric bootstrapping method had already been used to successfully estimate the distribution of the anaerobic digestion kinetic
parameter estimators (Ruiz et al., 2005) but there is still no RSM applications.

75 The novelty of this work is to combine a new surface response methodology for SS-AD process, 76 which is a mix design combining the response surface design and the mixture design with a 77 numerical tool to reduce the number of experiments by iterative resampling: the bootstrapping. This study has been conducted on two substrates and two key parameters of 78 79 the SS-AD to maintain time and cost constraints on batch reactors. The substrates used were cattle manure and damp grass. In fact, grass is a substrate with risk of inhibition by acidosis as 80 cattle manure may offset this effect. The chosen parameters have also been selected for their 81 82 impact on the acidosis and these parameters are immersion rate and recirculation frequency. 83 The chosen SRD is a full-factorial design and the chosen MD is a simplex-centroid design.

84

85 2 Materials and methods

86 2.1 Characteristics of substrates and inoculum used

87 Cattle manure (CM) was sampled from the farm of the UniLaSalle Polytechnic Institute (Beauvais, 88 France). CM has been sampled for each set of experiments to maintain its raw properties and avoid 89 any preliminary treatment. The liquid part of the manure or liquid bovin manure (LBM) was used as 90 inoculum to bring the microbial consortium. Damp grass (DG) was mowed from household gardens 91 in Les Clayes-sous-Bois (France) in September 2019 and has been aliquoted and frozen to be used 92 at each set of experiments to avoid any physicochemical differences with DG used. Two sets of 93 experiments were conducted in October 2019 and January 2020. The physicochemical analyses 94 were made on CM and DG for each data set.

95 For all experiments, all analyses were carried out in triplicate. The total solid content (TS) and the 96 volatile solid (VS) of the inoculum, CM and DG were determined by a drying at 105°C for 24h and a 97 combustion at 550°C for 2h (APHA, 1998). The pH of the inoculum has been determined using a pH 98 meter (Mettler Toledo, Switzerland). The total volatile fatty acid content (FOS) and the buffer 99 capacity (TAC) were determined by two titration using sulfuric acid. The first acidification at pH =100 5.0 provides the TAC and the second one at pH = 4.4 provides the FOS. An automatic titrator 101 (Mettler Toledo, Switzerland) has been used to realize these analyses. These tests were done at the 102 beginning of each set of experiments and the results are presented in Table 1. The FOS/TAC value 103 is an indicator to recognize stable AD operation or acidosis risks, where an excessive accumulation 104 of acids leads to AD inhibition (Mcghee, 1968; Nordmann, 1977; Voß et al., 2009).

105 2.2 Experimental set-up

106 For this study, one of the fixed objectives was to denature substrates the less possible. This 107 means that no grinding is done on substrates and the CM is not frozen but sampled and

108 analyzed for each experiment. This choice is done to be as representative as possible of the 109 real experimental conditions of an industrial digester. Thus, used reactors must respect space 110 and cost constraints but it must be sufficiently large for samples studied. Moreover, 111 constraints must be respected: the experimental parameters tested must be adjustable. 112 Fifteen leach bed reactors (LBR) made of Plexiglas with a total volume of 2.65 L were used 113 enabling a compromise between lab scale and substrate representativeness. These reactors 114 were divided into two parts: the upper part containing the solid phase and the bottom part 115 holding the liquid phase. These parts were separated by a perforated grid with a pore size of 5 116 mm diameter. The grid height was adjustable along a central axis of 10 mm diameter by a 117 manual tightening ring screwed, allowing imposing the immersion rate. A manual peristaltic 118 pump was used on each LBR at constant flow speed range and allowed to adjust the 119 recirculation frequency depending on the experiment. Thus, the liquid phase was spread out 120 across the top of the solid phase in the LBR upper part and percolated to drop back into the 121 bottom part. According to the experiment, recirculation frequency was between once a day 122 and once a week. The top of the reactor was directly connected with manual biogas counters. 123 These bottles were graduated and operated according to the Mariotte's law through volume 124 displacement, allowing to measure the cumulative biogas production each day. A temperature 125 reading allowed to convert this volume into a normalized quantity. Biogas composition was 126 analyzed five times a week by a gas chromatography (MicroGC, SRA Agilent 3000A) to determine the biogas composition on every reactor (weekends excepted) to measure 127 128 methane, carbon dioxide, nitrogen and hydrogen production rates. The system is represented 129 Fig 1.

130 At the beginning of the experiment, all the introduced inoculum was recirculated once. For 131 every other recirculation, the spread volume was 0.62 L, representing arbitrarily 2:3 of the 132 total liquid volume. The spreading period was 1 min, because 0.62 L.min⁻¹ is the minimal flow to insure a good liquid distribution over the whole surface of the solid matter. Every reactor 133 134 was filled with the same volatile solid content. The chosen quantity of substrates inserted in 135 each reactor contained the same volatile solid content than 250 g of cattle manure for the first 136 and second run and 300 g of cattle manure for the third run. This was done to have a sufficient 137 percolation. Previous investigations on the LBR permitted to determine this reference (data 138 not shown). DG was introduced on CM, imposing a two-stored solid matter. This successive stratification was chosen to reduce the contact between grass and liquid phase and thus 139 140 acidosis risks (André et al., 2019). At the end of each run of experiments, mass balances were 141 calculated on each reactor.

The ideal gas law was used to obtain the biogas mass flow expression. Then biogas flow was normalized and integrated throughout the experiment to obtain the total mass produced (g). Equations (1) and (2) describe this method.

$$PQ(t)M(t) = m(t)RT(t)$$
(1)

$$\int_{m_0}^{m_f} dm = \frac{P}{R} \int_{t_0}^{t_f} \frac{M(t)Q(t)}{T(t)} dt$$
 (2)

145 2.3 Mix Response Surface Methodology

A mix design was chosen to optimize the response variable. This RSM was composed by coupling a response surface design and a mixture design. The first hypothesis made was the proportionality of the responses on the mixture composition. Concerning MD, a simplex-centroid design was used to determine the effect of mixture composition of the studied response. For this study, the response was the total methane yield Y_{CH4} (NL.kg_{VS}⁻¹). The mixture factors γ_i were quoted as the ratio between the volatile solid content of the component i and the total volatile solid content (% VS_i.VS⁻¹). This unit ensures to obtain comparable results between experimentations and not to be dependent on experimental conditions as component density or moisture. The final equation (3) is the model used.

$$Y_{CH4} = (\alpha_{1,0} + \alpha_{1,1}\delta_1 + \alpha_{1,2}\delta_2 + \alpha_{1,12}\delta_1\delta_2)\gamma_1$$

$$+ (\alpha_{2,0} + \alpha_{2,1}\delta_1 + \alpha_{2,2}\delta_2 + \alpha_{2,12}\delta_1\delta_2)\gamma_2$$

$$+ (\alpha_{3,0} + \alpha_{3,1}\delta_1 + \alpha_{3,2}\delta_2 + \alpha_{3,12}\delta_1\delta_2)\gamma_1\gamma_2$$
(3)

This design permits a better flexibility than classical designs and considers the interactions between SSAD process parameters variations and composition variations on the methane yield. Process factor values must be normalized between -1 and +1 to compare the effects of process factors with each other.

Each run is achieved with 15 LBR experiments. Five replications of the center point were done to estimate the experimental error and confirm the model significance. They are included into the first run to compare the bootstrapping method with the statistical method. For the second run, only the bootstrapping will be used to determine the confidence interval of methane yield. All the experiments are summarized in the **Table 2**.

The statistical analysis was established by a Fisher test for the model with a 95 % confidence level and a Student test for the coefficients value with a 90 % confidence level. The analysis was done with the data analysis tool from Excel, allowing to determine the coefficients value and confidence intervals associated. The adequacy of the model was checked by adjusted regression coefficients (R_{adj}²). The regression models were determined by analysis of variance Page **9** sur **22** (ANOVA) (Miller, 1974 ; O'Connor, 1991) and a regressor selection was done (Lind & Ljung, 2005). After this selection, only 4 to 5 parameters among 12 were maintained for each run. This observation resulted from the high confidence level established. Only 14 experiments were used to realize the analysis on each run because some LBR presented recirculation failures and values were rejected. These LBR were the experiments R₁₅ and R₁₉ for the first run and R₁₃ for the second one. Finally, the assumptions dues to ANOVA method were proved by plotting a residual plot and a residual normality plot.

176 2.4 Repeatability analysis

177 Residual error includes model adjustment error and experimental error. To identify each error 178 and verify the model significance, sum squares and degrees of freedom were calculated. 179 Dedicated experiments R₁₆ to R₂₀ were launched during the first run of experiments. These 180 reactors were loaded with the same substrate's mixture and in the same operating conditions 181 (Table 2). The experimental sum square was obtained by following the final methane yield obtained y_{0k} for these k reactors and was compared to the mean y_0 of these 5 experiments. 182 183 This is how experimental variability was determined. Once the model is fitted with experimental points (R1 to R15, R13 excluded), the difference between experimental methane 184 yield y_{ik} and predicted methane yield $\hat{y_i}$ was done by the residual sum squares. The difference 185 186 between these two calculations allowed determining the adjustment error. The total number 187 of experiments N_{exp} was 18 because two experiments were deleted before (R₁₃ and R₁₉) and P is the standardized regressor number, which value was 5 after regressor selection. An F-value 188 is determined for the model significance and allow creating the confidence interval for each 189 190 value included in the field of study. In numerical statistics, bootstrapping is a resampling 191 method that operates by random sampling with replacement. The main idea is to model the Page **10** sur **22**

192 inference of samples by resampling data and performing conclusions about a sample from 193 resampled results. In inferential statistical analysis, the real population is unknown but using 194 bootstrapping, the population is the sample and is known. The original data set is resampled and create a new sample from the original. The process is repeated during a large number of 195 196 iterations N and each resampled result provides an estimate of the real population. This 197 method was published by Efron (1979) and was inspired by the jackknife method (Miller, 198 1974). It is a very practical tool concerning the determination of samples accuracy and allows 199 obtaining confidence intervals and estimating properties. The main advantage of this tool is 200 the implementation simplicity to determine complex parameters and also helps to check the 201 results stability. This method is more accurate than the standard confidence intervals obtained 202 using classic assumptions of inferential statistics and allows to reduce the number of 203 experiments. By contrast, the resampled results depend on the sample representativeness. 204 Concerning this work, bootstrapping is used for building new samples to offset the limited 205 number of experiments and to assess the significance of experimental results.

206 Bootstrapping was used on two major points of this study: the confidence intervals of 207 model coefficients and the confidence interval of predicted methane yield. The approach used 208 was the residual resampling. For the first point, the supposed model was fitted with 209 experimental data and fitted parameters and residues were retained. Then, randomly 210 resampled residues were added to the fitted response in order to create a new sample. The 211 model was fitted again with experimental data including the resampled results and all samples 212 and new samples were retained. These last two steps were repeated in a large number of 213 times to approach the empiric probability distribution. For the second point, once the model 214 was fitted with experimental data, the sample residues were resampled over the full field of study. In addition to this attention point, the operating method was the same. The residues
were studentized to consider the standard deviation difference between residues. The
complete method is sketched Fig 2. Every calculation and plot were programmed with Scilab
6.0.0.

219 3 Results and discussion

220 3.1 Batch reactor performance

221 For the first run of experiments with the damp grass composition lower than 45%_{VS}, the 222 immersion rate and the recirculation frequency became higher as well as the methane yield. 223 The difference between immersion rates was clearly visible in the LBR when the immersion 224 rate was high, the methane yield was higher. However, the recirculation frequency induced 225 different comportments depending on the immersion rate and the mixture composition. For 226 the second run, when the damp grass composition was higher than 50%_{VS}, two main optimal 227 methane yield zones appeared: when the immersion rate was low and the recirculation 228 frequency was high and vice versa. Moreover, when the damp grass composition was higher 229 than 50%_{vs}, it was very difficult to construe the biogas flow and methane curve shapes. This is 230 a typical example justifying the use of response surface methodology. Three experiments were 231 removed from these two runs because there was some experimental bias like recirculation 232 blockages or sealing problems: R13 for each run of experiments and R19. All the 233 characteristics of the liquid phase and the solid phase are summarized in the Table given in 234 supplementary material.

For each run of experiments, mass balances were calculated on each reactor. Biogas flow has
 been normalized and included in the mass balance to obtain the total mass produced (g). The
 minimum and maximum differences are 0.20 % and 8.66 %. LBR are therefore representing
 Page 12 sur 22

238 mass balance between 91.34 % and 99.80 %, attesting the absence of biogas leakage or reactor 239 failures. This difference may be explained by the mass losses during the reactors emptying and 240 mass weighing. Depending on experimental conditions, the VS removal ranged from 63% to 87% for the first run and from 62.9% to 83.5% for the second run, attesting a great 241 242 biodegradation of substrates by the microbial consortium. These results are illustrated in Fig 3. 243 This conclusion is echoed by the methane yield observed. These consumptions represent between 51% and 95% of the BMP for the first run and between 20% and 57% for the second 244 245 run. These results show that the experimental conditions are not optimal, probably due to the 246 substrates accessibility. Generally, the experiments with a low recirculation frequency and with 247 more damp grass produce less biogas. Some experimental parameters could explain this 248 phenomenon as preferential path and dead zones during percolation of the liquid phase. 249 Moreover, the pH value is considerably higher than the beginning of experiments, between 250 7.79 and 8.05, indicating a good balance between produced and consumed volatile fatty acids 251 (VFA). For the first run, the FOS/TAC ratio is between 0.098 and 0.128 and between 0.112 and 252 0.296 for the second run. This result suggests an impact of damp grass on FOS concentration 253 and buffer capacity.

254 *3.2 Statistical analysis of the results*

Table 3 gives the final model of methane yield coefficients after regressors selection (after removing all nonsignificant regressors, with a p-value > 0.1). Two models were determined according to the run of the experiment, therefore according to experimental conditions. This model presents an adjusted R^2 of 88 ± 1 %, which attesting a great accuracy to describe the reality. The residual variability could be explained by the uncertainty of measurement and by the impact of external parameters variations which were not considered for this study. But these results still remain good results in terms of accuracy. The model p-value is very low,proving the model significance.

263

264 Table 4 provides the regressors value. Each run of experiments presents different models and 265 thus different regressors according to the experimental conditions. In fact, the significant 266 coefficients are not the same depending on the inhibition phenomenon, the substrate mixture, 267 the *inoculum*/substrate ratio and the recirculation frequency. A positive sign of the coefficient 268 means that the coefficient has a favorable impact on the methane yield. Therefore, the cattle 269 manure γ_1 and damp grass γ_2 presence are beneficial to the methane production. The 270 immersion δ_1 and the recirculation frequency δ_2 are beneficial to the methane yield when 271 associated with the substrates mixture, according to the interaction results $\delta_1 \gamma_1 \gamma_2$ and $\delta_2 \gamma_1 \gamma_2$. 272 In contrast, when the damp grass composition is higher (during run 2), the interaction 273 coefficient between substrates $\gamma_1\gamma_2$ is negative. This means that a high composition rate of 274 damp grass induced a negative impact on the methane yield, probably due to inhibition 275 phenomenon. Finally, an interaction coefficient between process parameters and cattle 276 manure composition $\delta_1 \delta_2 \gamma_1$ is negative in the first run of experiments. This could be explained 277 by the substrate's stratification, inducting that cattle manure is easier to be submerged than 278 damp grass. As the damp grass has a higher methanogenic power, the process parameters may 279 not improve the methane yield of a mixture with a low damp grass composition. The residues 280 of each model were plotted and all the postulates may be graphically verified. There is still a 281 slight difference of residual variability between the experimental values and models and a few 282 points which present fractiles slightly apart the normality line. This is why all the residues were normalized by dividing residues by the residual standard deviation. This point will also beconsidered when the bootstrapping is used.

285 3.3 Optimization of the methane yield

The regression model obtained and the coefficients value are attained as illustrated in equations (4) to (7).

$$Y_{CH4,Run1} = (\alpha_{1,0} + \alpha_{1,12}\delta_1\delta_2)\gamma_1 + \alpha_{2,0}\gamma_2 + (\alpha_{3,1}\delta_1 + \alpha_{3,2}\delta_2)\gamma_1\gamma_2$$
(4)

$$Y_{CH4,Run2} = \alpha_{1,0}\gamma_1 + \alpha_{2,0}\gamma_2 + \alpha_{3,0}\gamma_1\gamma_2 + \alpha_{3,2}\delta_2\gamma_1\gamma_2$$
(5)

$$Y_{CH4,Run1} = (147.95 - 8.31\delta_1\delta_2)\gamma_1 + 113.59\gamma_2 + (63.69\delta_1 + 27.48\delta_2)\gamma_1\gamma_2$$
 (6)

$$Y_{CH4,Run2} = 129.41\gamma_1 + 143.45\gamma_2 - 171.86\gamma_1\gamma_2 + 77.79\delta_2\gamma_1\gamma_2$$
(7)

288 These equations were numerically solved in order to find the optimum coefficient values to 289 maximize the methane yield. The maximal production value obtained for each run is respectively 156.19 NL.kgvs⁻¹ and 143.45 NL.kgvs⁻¹. This difference is due to the *inoculum* 290 properties and the differences between recirculation frequencies. Thus, this value could not be 291 292 compared. Therefore, the methane yield comportment can be analyzed and compared. These 293 conclusions are consistent with the result found in the literature (André et al., 2019; Feng et 294 al., 2017) considering the inoculum, cattle manure and damp grass TS, VS and composition 295 differences. The response surface obtained for the first run is illustrated in Fig 4. The predicted 296 methane yield depends on the process parameters and mixture parameters. In this way, when 297 the cattle manure composition increase (figures from the top to the bottom), the predicted methane yield observed depending on the immersion rate and the recirculation frequency. 298 299 These observations indicate a rise of the methane yield when the damp grass and process 300 parameters rise until damp grass composition reaches 45 %_{vs}. If extrapolation is done, beyond Page **15** sur **22** 301 70 $\%_{VS}$ of damp grass, the predicted methane yield curve has a significantly different behavior. 302 This probably represents the impact of contact period between liquid phase and substrates. 303 Concerning the second run of experiments, the immersion rate had not been effectively brought into focus. predicted methane yield is not depending on this parameter. In this way, 304 305 methane yield can be based only on damp grass composition and recirculation frequency as 306 illustrated in Fig 5. A red area means a higher methane yield and a blue area means a lower methane yield. The methane yield surfaces are on the right and methane yield isovalues are 307 308 represented on the left. The predicted methane yield curve has a different behavior than the 309 first experiment. This is probably because the high damp grass percentage induced inhibitions, potentially acidosis, that were not present during the first set of experiments and induce a 310 311 different behavior. This difference of behaviors highlights that extrapolation presents strong 312 uncertainties for RSM methods and should not be used for SS-AD with risks of inhibitions.

313 In these different experimental conditions, the same conclusion is done: in the mixture considered, when the damp grass composition is upper than 70 %_{VS}, inhibitions by acidosis 314 315 could appear. In this case, immersion rate and recirculation frequency can limit this 316 phenomenon. This study highlighted that two process behaviors permit controlling acidosis 317 risks: high immersion rates and low recirculation frequency, or low immersion rates and high 318 recirculation frequency. This could be explained by an equilibrium between foster the contact 319 period between microbial consortium and the higher biodegradable DG substrate and limiting 320 the acidosis risks.

321 3.4 Repeatability study

322 Once the predicted methane yield was optimized and the response surface modeled, it is very 323 important to determine the prediction accuracy to know how these experiments are Page **16** sur **22**

repeatable. For this, two main tools were used: the classical analysis of variance and the 324 325 bootstrapping. ANOVA needs to respect some postulates done on the model residues: variance 326 is equal for all residues, each residue is independent, normally distributed, and the statistical 327 average of residues is zero. The postulate acceptation causes low bias and need a sufficient number of experiments to be correctly verified. The confidence intervals were determined and 328 329 the bootstrapping tool were applied to the experimental results concerning two main points: 330 the confidence intervals on the model coefficients determination and the confidence intervals 331 on the predicted methane produced.

332 Concerning the model coefficients, the bootstrapping was used 200 times with 2000 iterations 333 to approach the real population. This represents 400,000 confidence interval simulations of 334 each coefficient or between 1,600,000 and 2,000,000 calculations. The results are shown Table 335 5. Beyond 100 uses of the bootstrapping tool, the simulations stay in the same domain. When 336 the simulations are over, the maximum upper confidence interval and the minimum lower 337 confidence interval are retained and compared with the ANOVA confidence intervals. The results obtained provide the same confidence intervals with a maximal error of 2.9 %. This 338 339 error could be explained by the absence of postulates and the residual variability awareness by 340 the studentized residues. In this study, residue postulates were verified and enough 341 experimental data were used to obtain a robust analysis of variance. But for other design 342 experimentations where there is a lack of samples, bootstrapping could be a solution to determine the repeatability of the model and may be useful to decrease the necessary number 343 344 of experiments without losing accuracy.

The same study has been done on the confidence intervals of predicted methane yield. The confidence intervals on the first run have been determined for a damp grass composition 347 between 0.15 and 0.45. The extrapolation has been done for a damp grass composition between 0.45 and 0.85. The same work has been done for the second run of experiments for a 348 349 damp grass composition between 0.15 and 0.85. The confidence intervals provided by the ANOVA method and by the bootstrapping method are still comparable. Results are illustrated 350 351 in Fig 6 for the second run. The green surface represents the predicted methane yield. The blue 352 surface represents the ANOVA confidence intervals and the red surfaces represent the bootstrapping confidence intervals. It can be noticed that bootstrapping provides higher 353 354 confidence intervals than ANOVA. It could be explained by the consideration of residue 355 variability. Bootstrapping is a method that analyzes and reproduces a global error between experimental points and predicted points, containing experimental errors and adjustment 356 357 errors. This is why experimental repetitions are not necessary to determine confidence intervals. The predicted methane surface has a very simple shape, due to the predicted 358 359 methane yield non-dependency with the recirculation frequency. The ANOVA confidence 360 intervals are represented in blue and bootstrapping intervals are represented in red. This is 361 why the response surface had also been studied depending on the damp grass composition 362 and the recirculation frequency.

Fig 5. provides a response surface shape close to the studies already present in the literature (Degueurce et al., 2016) and highlights the significance of recirculation frequency on the methane yield. Moreover, for each experiment, bootstrapping provides larger confidence intervals. This difference is probably due to the absence of study hypotheses, confirming its use. The larger intervals allow a reduced risk of experimental errors and adjustment errors.

368 4 Conclusions

The results suggested that interactions between process parameters and mixture parameters are significant and plays an important role in methane production, which validate the usefulness of mix response designs. The repeatability could be studied with numerical tools as bootstrapping for greater accuracy and minimizing the number of experiments. This work proposes an optimal method in terms of time and accuracy to realize response surface studies. However, it is only a first step and it is possible to go far beyond these results, notably using more complex mixtures, more process parameters and different response designs.

376 **5** Aknowledgments

The authors gratefully thank the program MOCOPEE (www.mocopee.fr) and the FEDER fund (FEDER PO Picardie / MOCOPEE PI0012581) for the support provided for this work and the financing of the thesis of Arnaud COUTU. The authors want also to thank Laura André for her technical help in the lab and Joseph Fayolle for the carefully reading and language correction of the manuscript.

382 6 References

- André, L., Pauss, A., Ribeiro, T., 2018. Solid anaerobic digestion: State-of-art, scientific and technological hurdles.
 Bioresource Technology 247, 1027–1037. https://doi.org/10.1016/j.biortech.2017.09.003
- André, L., Zdanevitch, I., Pineau, C., Lencauchez, J., Damiano, A., Pauss, A., Ribeiro, T., 2019. Dry anaerobic co digestion of roadside grass and cattle manure at a 60 L batch pilot scale. Bioresource Technology 289,
 121737. https://doi.org/10.1016/j.biortech.2019.121737
- APHA (1998): Standard Methods for the Examination of Water and Wastewater. American Public Health
 Association, 20th ed. American water works association and water environment federation,
 Washington, USA
- Bassard, D., 2015. Méthodologie de prédiction et d'optimisation du potentiel méthane de mélanges complexes en
 co-digestion (PhD Thesis). UTC de Compiègne.

- 393 Degueurce A., Tremier A., Peu P., 2016. Dynamic effect of leachate recirculation on batch mode solid state
 anaerobic digestion: Influence of recirculated volume, leachate to substrate ratio and recirculation
 395 periodicity. Bioresource Technology 216, 553–561. https://doi.org/10.1016/j.biortech.2016.05.113
- Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife. Ann. Statist. 7, 1–26.
 https://doi.org/10.1214/aos/1176344552
- Efron, B., Tibshirani, R., 1986. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures
 of Statistical Accuracy. Statist. Sci. 1, 54–75. https://doi.org/10.1214/ss/1177013815
- 400 Eurostat, 2020. Indicateurs clés. https://ec.europa.eu/eurostat/fr/web/europe-2020-indicators/europe-2020 401 strategy/headline-indicators-scoreboard (accessed 10.30.20).
- 402 Feng, L., Gao, Y., Kou, W., Lang, X., Liu, Y., Li, R., Yu, M., Shao, L., Wang, X., 2017. Application of the Initial Rate
 403 Method in Anaerobic Digestion of Kitchen Waste. Biomed Res Int 2017.
 404 https://doi.org/10.1155/2017/3808521
- 405 Goupy, J., Creighton, L., 2006. Introduction aux plans d'expériences, 5th ed. Dunod: l'usine nouvelle, Paris.
- Goupy, J.L., 2012. Methods for Experimental Design: Principles and Applications for Physicists and Chemists.
 Elsevier Science.
- Gunes, B., Stokes, J., Davis, P., Connolly, C., Lawler, J., 2021. Optimisation of anaerobic digestion of pot ale after
 thermochemical pre-treatment through Response Surface Methodology. Biomass and Bioenergy 144,
 105902. https://doi.org/10.1016/j.biombioe.2020.105902
- Jackowiak, D., Bassard, D., Pauss, A., Ribeiro, T., 2011. Optimisation of a microwave pretreatment of wheat straw
 for methane production. Bioresource Technology 102, 6750–6756.
- 413 https://doi.org/10.1016/j.biortech.2011.03.107
- Jiménez, J., Guardia-Puebla, Y., Romero-Romero, O., Cisneros-Ortiz, M.E., Guerra, G., Morgan-Sagastume, J.M.,
 Noyola, A., 2014. Methanogenic activity optimization using the response surface methodology, during
 the anaerobic co-digestion of agriculture and industrial wastes. Microbial community diversity. Biomass
 and Bioenergy 71, 84–97. https://doi.org/10.1016/j.biombioe.2014.10.023
- Jin, W., Xu, X., Yang, F., Li, C., Zhou, M., 2018. Performance enhancement by rumen cultures in anaerobic co digestion of corn straw with pig manure. Biomass and Bioenergy 115, 120–129.
- 420 https://doi.org/10.1016/j.biombioe.2018.05.001
- Kainthola, J., Kalamdhad, A.S., Goud, V.V., 2019. Optimization of methane production during anaerobic codigestion of rice straw and hydrilla verticillata using response surface methodology. Fuel 235, 92–99.
 https://doi.org/10.1016/j.fuel.2018.07.094

- Kashi, S., Satari, B., Lundin, M., Horváth, I.S., Othman, M., 2017. Application of a mixture design to identify the
 effects of substrates ratios and interactions on anaerobic co-digestion of municipal sludge, grease trap
 waste, and meat processing waste. Journal of Environmental Chemical Engineering 5, 6156–6164.
 https://doi.org/10.1016/j.jece.2017.11.045
- Kumar, V., Kumar, P., Kumar, P., Singh, J., 2020. Anaerobic digestion of Azolla pinnata biomass grown in integrated
 industrial effluent for enhanced biogas production and COD reduction: Optimization and kinetics
 studies. Environmental Technology & Innovation 17, 100627. https://doi.org/10.1016/j.eti.2020.100627
- Lee, J.T.E., Wang, Q., Lim, E.Y., Liu, Z., He, J., Tong, Y.W., 2020. Optimization of bioaugmentation of the anaerobic
 digestion of Axonopus compressus cowgrass for the production of biomethane. Journal of Cleaner
 Production 258, 120932. https://doi.org/10.1016/j.jclepro.2020.120932
- 434 Leonzio, G., 2019. Fluid dynamic study of anaerobic digester: optimization of mixing and geometric configuration
 435 by using response surface methodology and factorial design. Renewable Energy 136, 769–780.
 436 https://doi.org/10.1016/j.renene.2018.12.115
- Lind, I., Ljung, L., 2005. Regressor selection with the analysis of variance method. Automatica 41, 693–700.
 https://doi.org/10.1016/j.automatica.2004.11.017
- 439 Mcghee, T. J., 1968. A Method for Approximation of the Volatile Acid Concentrations in Anaerobic Digesters,
 440 Water and Sewage Works, April 1968, 115, 162-166
- 441 Miller, R.G., 1974. The Jackknife--A Review. Biometrika 61, 1–15. https://doi.org/10.2307/2334280
- 442 Mortezaei, Y., Amani, T., Elyasi, Sh., 2018. High-rate anaerobic digestion of yogurt wastewater in a hybrid EGSB
 443 and fixed-bed reactor: Optimizing through response surface methodology. Process Safety and
 444 Environmental Protection 113, 255–263. https://doi.org/10.1016/j.psep.2017.10.012
- 445 Nordmann, W. (1977). Die Überwachung der Schlammfaulung. KA-Informationen für das Betriebspersonal,
 446 Beilage zur Korrespondenz Abwasser 3/77
- O'Connor, P.D.T., 1991. Introduction to Statistical Quality Control (2nd edition), D. C. Montgomery, Wiley, 1991.
 Number of pages: 702. £49.35, Paperback £17.50. Quality and Reliability Engineering International 7,
 201–201. https://doi.org/10.1002/qre.4680070316
- 450 Rakotoniaina, V.A., 2012. Co-méthanisation des déchets fermiers et alimentaires : expérimentation et
 451 modélisation (PhD Thesis). La Réunion.
- 452 Rao, P.V., Baral, S.S., 2011. Experimental design of mixture for the anaerobic co-digestion of sewage sludge.
 453 Chemical Engineering Journal 172, 977–986. https://doi.org/10.1016/j.cej.2011.07.010
- 454 Rocamora, I., Wagland, S.T., Villa, R., Simpson, E.W., Fernández, O., Bajón-Fernández, Y., 2020. Dry anaerobic
 455 digestion of organic waste: A review of operational parameters and their impact on process
 456 performance. Bioresource Technology 299, 122681. https://doi.org/10.1016/j.biortech.2019.122681

457	Ruiz, G., Castellano, M., González, W., Roca, E., Lema, J.M., 2005. Anaerobic digestion process parameter
458	identification and marginal confidence intervals by multivariate steady state analysis and bootstrap, in:
459	Puigjaner, L., Espuña, A. (Eds.), Computer Aided Chemical Engineering, European Symposium on
460	Computer-Aided Process Engineering-15, 38 European Symposium of the Working Party on Computer
461	Aided Process Engineering. Elsevier, pp. 1327–1332. https://doi.org/10.1016/S1570-7946(05)80063-X

- Voß, E., Weichgrebe D. Rosenwinkel, K.-H., 2009. FOS/TAC–Deduction, Methods, Application and Significance,
 InternationaleWissenschaftskonferenz. Biogas Science 2009 Science meets Practice", LfL-Bayern, 2-4.
 12.09, Erding. (http://www.ve-gmbh.de/_cms/images/stories/vegmbh_documents/ FOS-TAC DeductionMethodsApplicationSignificance-EVoss. Pdf)
- Wang, Z., Xu, F., Li, Y., 2013. Effects of total ammonia nitrogen concentration on solid-state anaerobic digestion of
 corn stover. Bioresource Technology 144, 281–287. https://doi.org/10.1016/j.biortech.2013.06.106
- 468 Figure captions
- 469 Fig 1. Schematic representation of the lab scale experimentation set up
- 470 **Fig 2.** Global schema of bootstrapping method
- 471 **Fig 3.** TS and VS content for solid and liquid phases for (a) : Run1 and (b) : Run 2
- 472 **Fig 4.** Predicted methane yield for Run 1, according to the recirculation frequency and the
- 473 immersion for a fixed composition
- 474 **Fig 5.** Predicted methane yield for Run 2, according to the recirculation frequency and the damp
- 475 grass composition
- 476 **Fig 6.** Predicted methane yield confidence intervals for the second run (Methane production:
- 477 green, ANOVA: blue, bootstrapping : yellow), depending on recirculation frequency and damp grass
- 478 composition













		TS	VS	рН	FOS	TAC	BMP
	Unit	%	% _{TS}	-	mg.L ⁻¹	mg.L ⁻¹	NL.kgvs ⁻¹
	Standard error	± 0.03	± 0.01	± 0.01	± 5.5	± 62	-
Run 1	Initial CM	22.0	87.7	7.51	-	-	142.92 ± 12.42
	Initial DG	42.2	71.7	5.81	-	-	240.39 ± 8.95
	Initial LM	1.61	62.7	7.92	96.81	3460	10.42 ± 0.01
Run 2	Initial CM	18.5	87.6	7.46	-	-	142.68 ± 16.93
	Initial DG	42.2	71.7	5.82	-	-	240.39 ± 8.95
	Initial LM	1.80	46.3	8.77	1021.21	7081	10.42 ± 0.01

Table 1 Chemical characteristics of inoculum and initial substrates used

Experiment	Standardized factor values		Compositions Run 1		Compositions Run 2		Methane yield Run 1	Methane yield Run 2
	δ_1	δ_2	γ_1	γ2	γ1	γ2	(NL.kgvs ⁻¹)	(NL.kg _{VS} ⁻¹)
R ₁	+1	-1	0.55	0.45	0.15	0.85	144.77	120.51
R_2	+1	-1	0.70	0.30	0.5	0.5	154.86	85.41
R ₃	+1	-1	0.85	0.15	0.85	0.15	150.11	97.51
R ₄	+1	+1	0.55	0.45	0.15	0.85	148.28	126.01
R5	+1	+1	0.70	0.30	0.5	0.5	152.97	124.04
R ₆	+1	+1	0.85	0.15	0.85	0.15	143.42	130.04
R ₇	-1	-1	0.55	0.45	0.15	0.85	95.72	119.80
R ₈	-1	-1	0.70	0.30	0.5	0.5	121.27	47.63
R 9	-1	-1	0.85	0.15	0.85	0.15	123.72	109.45
R ₁₀	-1	+1	0.55	0.45	0.15	0.85	129.99	111.42
R11	-1	+1	0.70	0.30	0.5	0.5	133.70	123.07
R ₁₂	-1	+1	0.85	0.15	0.85	0.15	144.56	88.00
R ₁₃	0	0	0.55	0.45	0.15	0.85	158.91	137.63
R ₁₄	0	0	0.70	0.30	0.5	0.5	147.8	87.19
R15	0	0	0.85	0.15	0.85	0.15	140.98	123.03
R ₁₆	0	0	0.70	0.30	-	-	151.93	-
R17	0	0	0.70	0.30	-	-	142.52	-
R ₁₈	0	0	0.70	0.30	-	-	136.70	-
R ₁₉	0	0	0.70	0.30	-	-	124.00	-
R20	0	0	0.70	0.30	-	-	138.66	-

Table 2 Mix design of experiment used for runs 1 & 2	

Regression	DF	Sum of squares	Mean squares	F-value	p-value
Run 1					
Model	5	2.70.10 ⁵	5.39.10 ⁴	1421.22	1.44.10-11
Residual	9	3.41.10 ²	3.79.10 ¹	-	-
Total	14	$2.70.10^{5}$	-	-	-
R ²	-	0.999	-	-	-
R_{adj}^2	-	0.887	-	-	-
Standard error	-	6.16	-	-	-
Run 2					
Model	4	$1.63.10^{5}$	$4.08.10^4$	134.29	5.13.10-8
Residual	10	3.04.10 ³	$3.04.10^{2}$	-	-
Total	14	$1.66.10^{5}$	-	-	-
R ²	-	0.981	-	-	-
R_{adj}^{2}	-	0.876	-	-	-
Standard error	-	17.42	-	-	-

 Table 3 Analysis of variance on the final model (90 % confidence level)

Regressor	Coefficient value	Standard error	p-value	Lower boundary	Upper boundary
Run 1					
γ1	147.95	4.90	7.35.10-11	140.06	155.84
γ2	113.59	9.91	1.13.10-6	95.43	131.75
$\delta_1\delta_2\gamma_1$	-8.31	2.50	8.95.10-3	-12.89	-3.72
$\delta_1\gamma_1\gamma_2$	63.69	8.83	5.02.10-5	47.50	79.88
$\delta_2\gamma_1\gamma_2$	27.48	8.83	1.25.10-2	11.29	43.67
Run 2					
γ1	129.41	15.99	1.06.10-5	100.43	158.39
γ2	143.45	17.32	8.67.10-6	112.07	174.84
γ1 γ2	-171.86	79.52	5.60.10-2	-315.98	-27.73
$\delta_2\gamma_1\gamma_2$	77.79	28.27	2.04.10-2	26.56	129.02

 Table 4 Analysis of variance on the regressors (90 % confidence level)

Regressor	Coefficien t value	Lower ANOVA	Upper ANOVA	Lower Bootstrapping	Upper Bootstrapping	Relative error (%)
Run 1						
γ1	147.95	140.06	155.84	140.03	155.51	1.91
γ2	113.59	95.43	131.75	95.15	131.19	0.76
$\delta_1\delta_2\gamma_1$	-8.31	-12.89	-3.72	-12.76	-3.85	2.90
$\delta_1\gamma_1\gamma_2$	63.69	47.50	79.88	47.76	79.91	0.74
$\delta_2\gamma_1\gamma_2$	27.48	11.29	43.67	11.43	43.46	1.07
Run 2						
γ1	129.41	100.43	158.39	98.13	157.26	2.01
γ2	143.45	112.07	174.84	110.31	173.38	0.47
γ1 γ2	-171.86	-315.98	-27.73	-323.06	-30.90	1.36
$\delta_2 \; \gamma_1 \; \gamma_2$	77.79	26.56	129.02	25.24	130.54	2.77

Table 5 Analysis of variance and bootstrapping on the coefficients confidence intervals (90 % confidence level)

Methane yield optimization with mix response Surface methodology

+

Repeatability study using Bootstrapping methods

