Agronomic characterization of anaerobic digestates with near-infrared spectroscopy

Bastien Zennaro, Paul Marchand, Eric Latrille, J.-C. Thoisy-Dur, Sabine Houot, Cyril Girardin, Jean-Philippe J.-P. Steyer, Fabrice Béline, Cyrille Charnier, Charlotte Richard, et al.

1 # Agronomic characterization of anaerobic digestates with near-
2 # infrared spectroscopy

3

4 **Authors :** Bastien Zennaro[a]*, Paul Marchand[b], Eric Latrille[a], Jeanne-Chantal Thoisy[b], Sabine

5 Houot[b], Cyril Girardin[b], Jean-Philippe Steyer[a], Fabrice Béline [c], Cyrille Charnier[d], Charlotte

6 Richard[e], Guillaume Accarion [f], Julie Jimenez[a]

7 [a] *INRAE, Univ Montpellier, LBE, 102 Avenue des Etangs, 11100 Narbonne, France*

8 [b] *INRAE, EcoSys, Route de la Ferme, 78850 Thiverval-Grignon, France*

9 [c] *INRAE, UR OPAALE, 35044 Rennes, France*

10 [d] *BioEnTech, 74 Av. Paul Sabatier, 11100 Narbonne, France*

11 [e] *ENGIE, Lab CRIGEN, 361 Avenue du Président Wilson, 93210 Saint-Denis, France*

12 [f] *Akajoule, 18 Boulevard Paul Perrin, 44600 Saint-Nazaire, France*

13

14 **Email address:** bastien.zennaro@inrae.fr

15 **Postal address:** INRAE, Univ Montpellier, LBE, 102 Avenue des Etangs, 11100 Narbonne,

16 France

17

18 **Keywords:** near infrared spectroscopy, anaerobic digestion, digestate, characterization,

19 agronomic value, chemometrics

20

21   **Abbreviations:** AD, anaerobic digestion; NIR, Near infrared; PLSR, partial least squares

22   regression; OM, organic matter; DM, dry matter; TOC, total organic carbon; N, nitrogen;

23   N_{org}, organic nitrogen; P, phosphorus; K, potassium; TKN, total Kjeldahl nitrogen; SNV,

24   standard normal variate; RPD, ratio of performance to deviation; RMSE, Root Mean Square

25   Error; SEL, Standard Error of Laboratory;

26

27   **Abstract:**

28   Anaerobic digestion is an increasingly widespread process for organic waste treatment and

29   renewable energy production due to the methane content of the biogas. This biological

30   process also produces a digestate (i.e., the remaining content of the waste after treatment)

31   with a high fertilizing potential. The digestate composition is highly variable due to the

32   various organic wastes used as feedstock, the different plant configurations, and the post-

33   treatment processes used. In order to optimize digestate spreading on agricultural soils by

34   optimizing the fertilizer dose and, thus, reducing environmental impacts associated to

35   digestate application, the agronomic characterization of digestate is essential.

36   This study investigates the use of near infrared spectroscopy for predicting the most

37   important agronomic parameters from freeze-dried digestates. A data set of 193 digestates

38   was created to calibrate partial least squares regression models predicting organic matter,

39   total organic carbon, organic nitrogen, phosphorus, and potassium contents. The calibration

40   range of the models were between 249.8 and 878.6 $gOM.kgDM^{-1}$, 171.9 and 499.5 $gC.kgDM^{-}$

41   ^{1}, 5.3 and 74.1 $gN.kgDM^{-1}$, 2.7 and 44.9 $gP.kgDM^{-1}$ and between 0.5 and 171.8 $gK.kgDM^{-1}$,

42   respectively. The calibrated models reliably predicted organic matter, total organic carbon,

43   and phosphorus contents for the whole diversity of digestates with root mean square errors

44   of prediction of 70.51 gOM.kgDM^{-1}, 34.84 gC.kgDM^{-1} and 4.08 gP.kgDM^{-1}, respectively. On

45   the other hand, the model prediction of the organic nitrogen content had a root mean

46   square error of 7.55 gN.kgDM^{-1} and was considered as acceptable. Lastly, the results did not

47   demonstrate the feasibility of predicting the potassium content in digestates with near

48   infrared spectroscopy.

49   These results show that near infrared spectroscopy is a very promising analytical method for

50   the characterization of the fertilizing value of digestates, which could provide large benefits

51   in terms of analysis time and cost.

52

1. Introduction

54   Anaerobic digestion (AD) is recognized as an efficient technology in terms of renewable

55   energy production and environmental protection according to the EU 2020 Renewable

56   Energy Directive (Commission of the European Communities, 2009). AD is a microbial

57   process that transforms organic matter into biogas, mainly composed of CH_4 and CO_2. The

58   biogas produced is either converted into electricity and heat by cogeneration, or directly

59   injected into the gas grid after purification. AD is very often considered as a process for

60   organic waste treatment and biogas production, but it also produces a digestion residue

61   called digestate, which can be used as an organic fertilizer on agricultural soils (Houot et al.,

62   2016). It is essential for the biogas plants to make the best use of their digestate so that it is

63   not a cost, but a gain for the plant operators and the environmental impact is minimized

64   when the digestate returns to the soil (Nkoa, 2014).

65   The agronomic and environmental interests of digestate application to agricultural land have

66   been evaluated in several studies (Walsh et al., 2012; Alburquerque et al., 2012; Nkoa, 2014 ;

67   Guilayn et al., 2020 ; Jimenez et al., 2020). It has been demonstrated that digestates are

68   good organic fertilizers and amendments, allowing reducing the use of chemical fertilizers

69   and striving for a circular economy at farm scale.

70   The chemical composition of the digestates is directly linked to the feedstocks used and the

71   AD plant operating conditions (Guilayn et al., 2019). Consequently, digestate compositions

72   show a great variability, and significant characterization needs have been highlighted

73   (Alburquerque et al., 2012; Kataki et al., 2017). The composition of a chemical fertilizer is

74   detailed by the supplier, and the dose to be applied to the field can be adjusted according to

75   the soil and the crops needs. The same information should be available for organic fertilizers

76   and soil improvers coming from digestates. This should lead to a better management of the

77   nitrogen inputs and ammonia volatilization risk. In addition, Guilayn et al. (2019) have

78   highlighted the almost systematic obligation to post-treat digestates (phase separation,

79   composting, stripping etc.) prior to their sale and use as organic fertilizers in order to comply

80   with European standards (European Parliament and Council of the European Union, 2016;

81   European Commission, 2003). Despite continuous improvement in post-treatment

82   technologies and the development of new recovery methods (Monlau et al., 2015; Guilayn

83   et al., 2020), the cost of digestate treatment remains high. Some studies have shown that it

84   is difficult to make a profit by valuing the digestate (KTBL, 2008) and that direct spreading of

85   the digestate remains very often the most profitable option (Fuchs and Drosg, 2013). Regular

86   characterization of the digestates, with at least a determination of the C, N, P, and K

87   contents, for monitoring and optimizing the post-treatment processes is, therefore, essential

88   (Teglia et al., 2011).

89   The physical and chemical laboratory analyses for characterizing the digestates are time and

90   cost expensive. Near infrared (NIR) spectroscopy could be an alternative technology for

91   rapid characterization of digestates. This non-destructive analytical method makes it

92   possible to analyze a sample by observing the harmonic oscillations of the various bonds of

93   the molecules during excitation of these by a source emitting in the wavelengths interval of

94   700 and 2500 nm (Burns and Ciurczak, 2007). From the spectra acquired in near infrared, it is

95   then possible to calibrate a model allowing a spectrum to be linked to one or more reference

96   values, qualitative or quantitative. Partial least squares regression (PLSR) is a model

97   calibration method very commonly used in NIR spectroscopy.

98   NIR spectroscopy has long been used in the chemical and food industry for reaction

99   monitoring or process control. However, many studies have also shown the interest of NIR

100   spectroscopy for the prediction of organic matter (OM) (Hummel et al., 2011), total organic

101   carbon (TOC) (Dalal & Henry, 1986; Barthès et al., 2019), nitrogen (N), phosphorus (P), and

102   potassium (K) from soils (Nduwamungu et al., 2009; He et al., 2007). The OM and the N

103   contents are predicted with good accuracy, while the models for P and K are much less

104   efficient. For the prediction of N, P, and K contents of different plants, NIR spectroscopy is

105   also used with very good results for the prediction of N (Gislum et al., 2004; Petisco et al.,

106   2005; Ward et al ., 2011), but often bad results for P and K (Tremblay et al., 2009; Ward et

107   al., 2011; Petisco et al., 2005). The difficulty in predicting the K content comes from the fact

108   that potassium is only present in ionic form and is, therefore, not directly visible with NIR

109   spectroscopy. Regarding the P content, the chemical bonds involving phosphorus are not

110   active in the range of near infrared wavelength (Mouazen et al., 2016). The models for

111   predicting the K and P contents are models based on indirect correlations with other

112   properties that have a direct spectral response. For the determination of the P content in

113   soils (Mouazen et al., 2016) or in plants (Menesatti et al., 2010), several studies have

114   demonstrated good performances of models based on spectra acquired at the same time in

115   the visible and in the near infrared.

116   NIR spectroscopy is already used in the field of anaerobic digestion for characterizing the

117   substrates to be treated with regard to their biological methane potential (Lesteur et al.,

118   2011), their sugar content, their chemical oxygen demand, their lipid content, and their

119   nitrogen content (Charnier et al., 2017). It is important to note that the complexity of the

120   anaerobic digestion substrates studied has not yet made it possible to calibrate these

121   models on fresh samples, but on freeze-dried and finely ground samples. On the other hand,

122   it has been shown that NIR spectroscopy could be used as a tool for characterizing raw

123   materials and digestates to assess the performance of sewage sludge digesters by only

124   determining volatile fatty acids, alkalinity, and solid and volatile matter (Reed et al., 2011;

125   Jacobi et al., 2011; Awhangbo et al., 2020).

126   Considering the great variability in the composition of the digestates, their chemical

127   characterization for monitoring and optimizing the AD process, the post-treatment of the

128   digestate and their return to the soil appears to be essential. Based on the potential of NIR

129   spectroscopy already demonstrated for the characterization of soils, plants, and biowastes,

130   this study aims to assess, for the first time, the potential of NIR spectroscopy for predicting

131   the most important agronomic characteristics of digestates. This new method could allow a

132   fast and more regular characterization of digestates, thus improving their management.

133

134   **2. Materials and Methods**

135   **2.1. Samples**

136   For this study, 193 partially characterized digestates were combined in one sample data base

137   in order to represent the great diversity of digestates, which can be produced. These

138   digestates came from farm, centralized, and urban anaerobic digesters plants, except for

139   twelve digestates generated by laboratory pilots (Table 1). Some of the digestates collected

140   had been treated by phase separation or composted directly on the production site. The

141   sample base consisted of 46 solid phase digestates, 28 liquid phase digestates, 43

142   composted solid phase digestates, and 76 raw digestates . Apart from four dry anaerobic

143   digestion plants (three piston flow systems and one batch system), units used for collecting

144   digestate samples were wet anaerobic digestion plants (Total Solids content < 20%).

145   **Table 1: Description of the anaerobic digestion plant types used in the digestates data base and**
146   **their respective percentages**

Type	Feedstock	Number	Percentage
Farm plant	Animal manures	110	57%
	Energy crops	5	3%
	Animal manures and agro-industrial wastes	3	2%
	Crop residues	2	1%
	Animal manures and biowastes	2	1%
Centralized plant	Agro-industrial wastes, sewage sludges, animal manures, green wastes, biowastes etc.	23	12%
Urban plant	Sewage sludges	23	12%
	Biowastes	14	7%
	Municipal wastes	11	6%

147

148   **2.2. Analytical methods for reference data**

149   The analytical methods for the determination of the dry and organic matter, total organic

150   carbon, organic nitrogen, phosphorus, and potassium are presented below. The analyses

151   were carried out in triplicate.

152   **2.2.1 Dry matter and organic matter**

153 The dry matter (DM) of a sample was determined by weight difference after 24 hours at 105

154 °C and the organic matter (OM), after at least two hours at 550 °C. The OM was obtained by

155 subtraction of the residual mineral matter obtained after passing at 550 °C from the dry

156 matter.

157 **2.2.2 Total organic carbon**

158 TOC was measured on the freeze-dried and grinded samples by catalytic combustion at 900

159 °C with a Shimadzu TOC-V-SSM-500A. Pure glucose samples were used as references.

160 **2.2.3 Organic nitrogen**

161 Organic nitrogen was calculated by subtracting the ammonium from the total Kjeldahl

162 nitrogen (TKN). The TKN was determined on the raw samples to measure both mineral and

163 organic nitrogen. 1.0 g of sample is added to 5 mL of distilled water. Then, 10 mL of sulfuric

164 acid (98% in mass) containing 0.366 g.L^{-1} of catalyst (copper selenite) are added. The sample

165 is then heated to 420 °C for approximately three hours until complete hydrolysis. The TKN is

166 then measured with a BUCHI 370-K distiller-titrator. Ammonium was measured with a BUCHI

167 370-K distiller-titrator using the liquid fraction of the digestates collected after 20 minutes of

168 centrifugation at 18 600 g and then filtered at 0.45 µm.

169 **2.2.4 Total phosphorus**

170 The total phosphorus content of the substrate samples was analysed using the freeze-dried

171 and 1 mm grinded sample. The sample is first mineralized with a BUCHI digestion unit K-438.

172 Then 0.05 g of sample is added to a flask containing 5 mL of distilled water and 10 mL of

173 sulfuric acid (98% in mass). The flask is then put into the mineralizer for six hours at 420 °C.

174 The organic phosphate is then transformed into ortho-phosphate. The acidity of the sample

175 is neutralized by addition of 20-25 mL sodium hydroxide (32% in mass). The phosphorus

176 assay is carried out with Hack LCK350 kits (2.0 - 20.0 mg.L^{-1} [PO^{4-}-P]) according to the

177 standard EN ISO 6878, and a Hack DR3900 spectrophotometer. The method used in these

178 kits is a colorimetric method known as molybdenum blue described by Murphy and Riley

179 (1962).

180 **2.2.5 Total potassium**

181 The total potassium was measured in the digestate supernatant recovered after

182 centrifugation (17700 g during 20 minutes at 4 °C). After filtration at 0.45 µm, 1 mL of

183 supernatant is introduced into a Hack LCK328 assay kit (8.0 - 50.0 mg.L^{-1} [K^{+}]) and the

184 potassium is measured with a Hack DR3900 spectrophotometer.

185 **2.3 Near infrared analyses**

186 The digestate samples were freeze-dried and then grinded to 200 µm with a ball mill to

187 remove as much as possible the spectral signal of water and to homogenize the samples

188 (Lesteur et al., 2011; Charnier et al., 2017). The samples were then scanned in reflectance

189 from 4000 to 10000 cm^{-1} with a resolution of 4 cm^{-1} on a BUCHI NIR-Flex N-500

190 spectrophotometer. The samples were scanned (96 scans) in vials supplied by the

191 manufacturer and adapted to homogeneous solids. Each sample was analysed three times

192 with a shaking step between each measurement to consider its heterogeneity. For each

193 sample, an average spectrum was calculated from the three scans and used for model

194 calibration.

195

	OM	TOC	N$_{org}$	P	K
	(g OM.kg DM^{-1})	(g C.kg DM^{-1})	(g N.kg DM^{-1})	(g P.kg DM^{-1})	(g K.kg DM^{-1})
Calibration					

Table 2: Dataset features for models' calibration and validation

Samples	118	99	75	65	62
Mean	599.5	326.5	26.7	14.8	41.9
SD	156.5	79.9	16.0	8.4	36.2
Max	878.6	499.5	74.1	44.9	171.8
Min	249.8	171.9	5.3	2.7	0.5
Validation					
Samples	49	39	31	22	17
Mean	638.0	343.3	30.5	14.3	42.9
SD	120.9	80.3	12.3	7.4	38.5
Max	880.0	450.3	62.3	39.8	154.4
Min	291.6	162.8	13.2	2.2	1.9

196

197 Before calibration, the sample base was randomly separated into a calibration set and an

198 external validation set, for each of the different models (Dardenne, 2010). The calibration

199 and validation set features are presented in Table 2.

200 **2.4 Partial Least Squares regression model**

201 The Partial Least Squares regression method was used for the calibration of the models

202 based on NIR spectra. As presented in the introduction, many studies have proven that this

203 method allows obtaining precise and robust models. The best spectral pre-treatment was

204 selected during the calibration and was based on cross-validation. The spectra were first

205 converted to pseudo-absorbance (log (1 / Reflectance)). The spectral range was then

206 reduced to 1300 to 2500 nm. A first derivative was obtained using a Savitzky-Golay filter

207 (derivative: 1, window: 15, polynomial: 2) with the savgol_filter function of scipy.signal in

208 Python 3.7. Finally, the spectra were standardized with a standard normal variate (SNV).

209 The different models were calibrated with a PLSR1 using the Python 3.7 PLSRegression

210 function (sklearn.cross_decomposition) with the NIPALS algorithm implemented from

211 Wegelin (2000) and Tenenhaus (1998). A RepeatedKFold Cross validation was performed for

212 each model with 10 blocks and 30 repetitions.

213 **2.5 Statistical parameters for model evaluation**

214 Different parameters describing the quality of the models were calculated with the

215 equations defined by Dardenne (2010) and Bellon-Maurel et al. (2010).

216 **2.5.1 Ratio of Performance to Deviation**

217 The Ratio of Performance to Deviation (RPD) is used in many studies as a quality criterion for

218 a model (Williams et al., 1987; Malley et al., 2004; Saeys et al., 2005). This parameter is only

219 linked to the R^2 coefficient of determination of the model and can therefore be calculated

220 for calibration, cross-validation, and validation. The RPD most often put forward to

221 demonstrate the quality and usefulness of a model is the validation RPD (RPD_{val}). However, it

222 is essential to note that the RPD_{val} is very dependent on the validation set chosen. The

223 chosen limit values of RPD, defining an accurate model, are very different from one field of

224 study to another (Bellon-Maurel and al., 2010).

225 This parameter was, thus, calculated for calibration (RPD_{cal}), cross-validation (RPD_{cv}), and for

226 validation (RPD_{val}), respectively, with the corresponding coefficients of determination R^2

227 calculated according to Eq. 1 for calibration (R^2_{cal}), cross-validation (R^2_{cv}), and for external

228 validation (R^2_{val}).

229
$$RPD = \frac{1}{\sqrt{1 - R^2}} \quad Eq.\,1$$

230 **2.5.2 Root Mean Square Error**

231 The Root Mean Square Error (RMSE) was calculated respectively for calibration (RMSEC),

232 cross-validation (RMSECV), and validation (RMSEP) according to Eq. 2.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)}{n}} \quad Eq.\,2$$

With $\hat{y}_i$ being the model value predicted from the spectrum of the sample, y_i being the reference value of the sample, and n being the number of samples considered.

2.5.3 Bias

This parameter was calculated for the external validation set following Eq. 3. With $\hat{y}_i$ being the model value predicted from the spectrum of the sample, y_i being the reference value of the sample, and m being the number of samples in the external validation set.

$$Biais = \sum_{i=1}^{m} \frac{\hat{y}_i}{m} - \sum_{i=1}^{m} \frac{y_i}{m} \quad Eq.\,3$$

2.5.4 Standard Error of Laboratory

The Standard Error of Laboratory (SEL) was calculated for the N_{org}, TOC, OM, P and K reference values with n being the number of samples, and k being the number of replications (Eq. 4).

$$SEL = \sqrt{\sum_{1}^{n} \frac{Var_{k-1}(k)}{n}} \quad Eq.\,4$$

The 95% confidence interval of the different analytical methods was calculated following Eq. 5.

$$IC\ 95\% = \pm 1.96 * SEL \quad Eq.\,5$$

2.5.5 Standard deviation of repeatability

250   The standard deviation of repeatability (SDr) of the PLSR model was evaluated by measuring

251   the variance of the spectral triplicates (Eq. 6).

252    $$SDr = \sqrt{\frac{1}{N}\sum Variance(triplicate)} \quad Eq.6$$

253

254   **3. Results and discussion**

255   **3.1 Model calibration**

256   The data treatment, which has been applied to the spectra, was the same for the different

257   models calibrated (1300-2500nm selected, pseudo-absorbance, Savitzky-Golay (derivative: 1,

258   window: 15, polynomial: 2) and SNV), because it gave the best results in terms of RMSECV

259   and RPD_{cv}. The aim of this treatment was to remove as much as possible additive and

260   multiplicative effects on the spectra caused by photon scattering and measurement noise .

261   The different PLSR models were calibrated with four to ten latent variables. The b-

262   coefficients of the regression were not noisy, which suggests a low risk of over-fitting. The

263   number of latent variables selected for each model was chosen to minimize the RMSECV and

264   to maximise the RPD_{cv}. The results are presented in Table 3.

265   The SDr of the models predicting OM, P, and N_{org} were lower than the SEL of the reference

266   method. For the TOC and K models, SDr and SEL were very close. Overall, the prediction of

267   these chemical parameters by NIR spectroscopy and PLSR modelling resulted in better or

268   similar repeatability compared to the reference laboratory analyses.

269   The model predicting the OM of the digestates was calibrated with 10 latent variables and a

270   presented good quality of calibration with a R^2 of 0.88. In addition, RPD_{cal} and RPD_{cv} were

271 very similar, which showed that the calibration base of the model was wide enough. The

272 RPD_{cal} and RPD_{cv} values equal to 2.86 and 2.04 were interpreted as indicators of a satisfying

273 calibration. The RMSEC and RMSECV were also similar (54.68 and 78.09 $gOM.kgDM^{-1}$) and

274 close to the SEL of the reference laboratory analysis equal to 26.30 $gOM.kgDM^{-1}$.

275 The TOC model provided satisfying calibration results with eight latent variables, a R^2 equal

276 to 0.85, and a RPD_{cal} of 2.58. The RMSEC and RMSECV of the model were similar but three

277 and four times higher than the SEL of the reference laboratory analysis equal to 10.61

278 $gC.kgDM^{-1}$. This result highlighted that the accuracy of the model could be theoretically

279 improved, but the predictions were already satisfactory.

280 **Table 3: Models' performances including calibration and validation with an independent data set**

	OM (g OM.kg DM^{-1})	TOC (g C.kg DM^{-1})	N_{org} (g N.kg DM^{-1})	P (g P.kg DM^{-1})	K (g K.kg DM^{-1})
Latent Variables	10	8	4	5	5
$R2_{cal}$	0.88	0.85	0.69	0.71	0.57
RMSEC	54.68	30.92	8.92	4.53	23.80
RPD_{cal}	2.86	2.58	1.79	1.86	1.52
RMSECV	78.09	45.84	11.01	6.21	32.25
RPD_{CV}	2.04	1.78	1.45	1.33	1.11
$R2_{val}$	0.69	0.82	0.67	0.71	0.41
RMSEP	70.51	34.84	7.55	4.08	29.72
Bias	13.47	-8.55	2.03	0.64	-1.06
RPD_{val}	1.72	2.31	1.63	1.81	1.29
SDr	10.01	10.61	1.57	0.89	6.17
SEL	26.30	10.30	2.66	2.86	6.58

281 *$R2_{cal}$ and $R2_{val}$, R2 of calibration and validation dataset ; RPD, ratio of performance to deviation;*
282 *RMSEP, RMSEC and RMSECV, roots mean square error of prediction, calibration and cross-validation;*
283 *SDr, standard deviation of repeatability ; SEL, standard error of laboratory*
284
285
286 For the N_{org} model, four latent variables were selected, and the calibration gave a R^2 of 0.69,

287 a RPD_{cal} of 1.79, and a RPD_{cv} of 1.45. The RMSEC and RMSECV were similar (8.92 and 11.01

288 $gN.kgDM^{-1}$) but four to five times higher than the SEL of the reference laboratory analyse

289   equal to 2.66 gN.kgDM^{-1}. The model was less accurate for high values of N_{org}. The samples

290   with high concentrations of N_{org} were mostly liquid phases of centrifuged digestates.

291   The P model presented was calibrated with five latent variables selected and provided a R^2

292   of 0.71, a RPD_{cal} of 1.86, and a RPD_{cv} of 1.33. The RMSEC and RMSECV are close to the SEL

293   equal to 2.86 gP.kgDM^{-1}. Two predictions of the calibration set were negative values, but

294   they corresponded to digestates with very low P concentrations and predictions remained in

295   the 95% confidence interval of the reference analyse.

296   Five latent variables were selected for the model predicting the K content in digestates

297   resulting in a RPD_{cv} of 1.11 and a RMSECV of 32.25 gK.kgDM^{-1}, almost eight times the SEL

298   (6.58 gK.kgDM^{-1}). The model gave also negative values for low concentrations of K and was

299   less accurate for high concentrations of K.

3.2 Model validation with independent datasets

301   The calibrated models were then tested on independent data sets made up of samples

302   randomly chosen among the whole data set in order to evaluate the models' robustness.

303   Between 41% and 27% of the total data set were used for validation depending on the total

304   number of samples available for each model with the aim of keeping enough samples for the

305   calibration. This selection method for the validation set enabled us to cover the whole range

306   of values for the different models as can be seen in Table 2. It can also be considered that

307   the validation sets contained representative samples of the diversity of digestates.

308   The validation results are presented in Table 3. All the RMSEP of the models were lower than

309   the RMSECV, which means that the models were robust, because the calibration results

310   could be extended over an independent sample set. Furthermore, there was no significant

311   bias to underline. It can be seen in Figure 1 that 95% and 70% of the validation predictions of

312   the P and the OM model, respectively, were inside the 95% confidence interval of the

313   analytical methods. Validation results were excellent for these two models. With reference

314   to the N_{org} and TOC model validation, 55% and 47% of the predictions were within the 95%

315   confidence interval. This result is certainly less satisfactory compared to the P and OM

316   models, but it is important to point out that the analytical methods for measuring N_{org} and

317   TOC were very repeatable and that, consequently, their SEL were extremely low. Lastly, only

318   24% of the validation predictions of the K model were inside the 95% confidence interval of

319   the analytical methods.
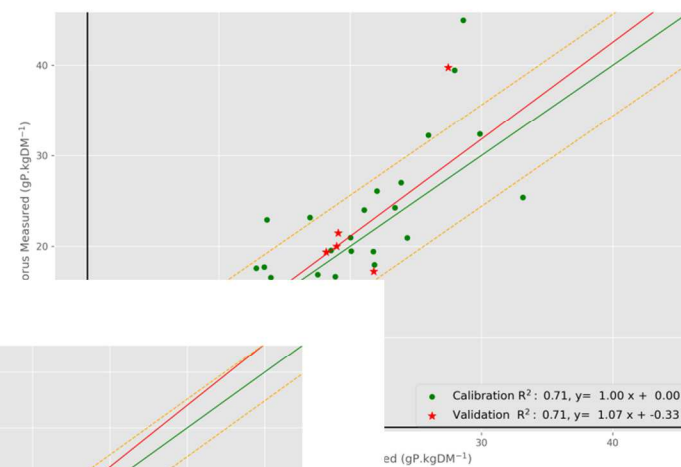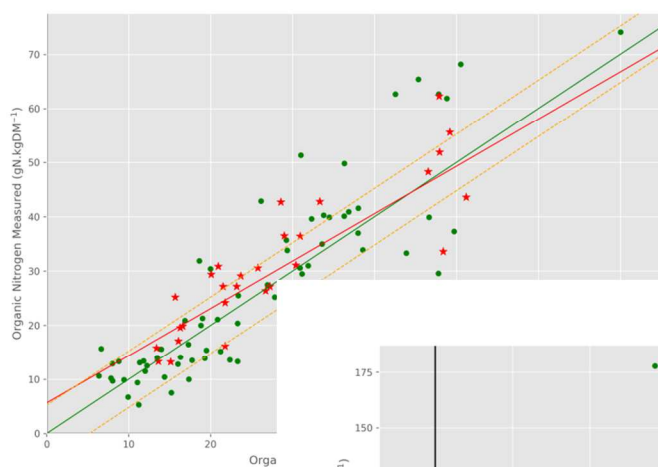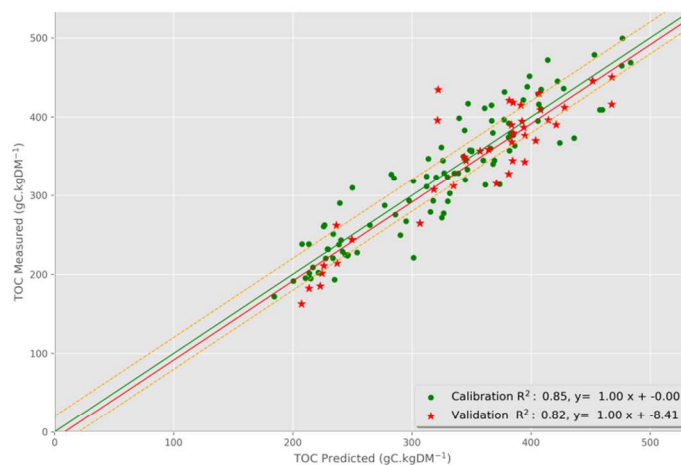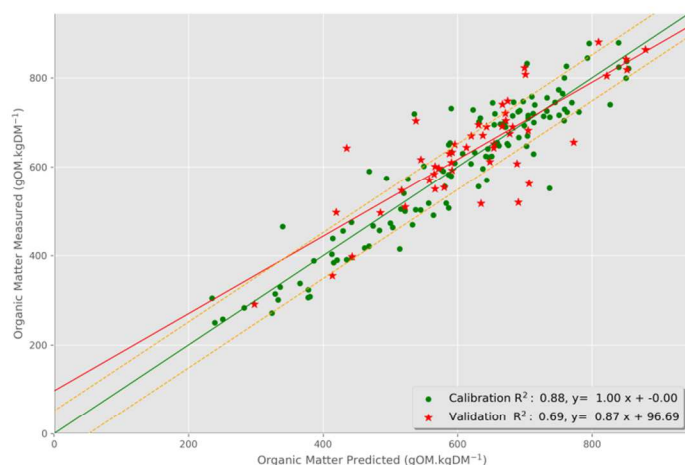
320

321

322

323

324

325
326
327 **Figure 1: Prediction performances of OM, TOC, N_{org}, P and K contents, respectively on the top left,**
328 **top right, middle left, middle right and bottom. The red stars represent the prediction of the**
329 **independent validation set whereas the green circles represent the estimation of the calibration**
330 **set. The green line and the red line represent the regression between reference and predicted**
331 **values for the calibration set and validation set respectively. The dashed yellow lines represent the**
332 **IC 95% of the analytical method.**
333 **3.3 Discussion**

334 The whole data set of digestates used in this study gathered samples from different AD

335 feedstocks and processes, which had created a large variety of digestate compositions. The

336 dry matter content of the digestates data set ranged from 1% (liquid phase of centrifuged

337 digestates) to 80% (composted solid phase of centrifuged digestates). As a result, the

338 distribution ranges of the chemical parameters predicted by the models presented were

339 very wide for most of them. The OM content of calibration samples varied between 250 and

340 879 gOM.kgDM^{-1} and between 292 and 880 gOM.kgDM^{-1} for the validation samples. The

341 residual organic matter in the dry matter of a digestate depends on substrates

342 (biodegradability, composition, etc.), the process removal efficiency, and the post-processing

343 used. It was difficult to find a strong correlation between the type of digestate and its OM

344 concentration, but it was possible to notice a trend. The samples with the lowest OM

345 contents have been reported to be mostly raw digestates and the liquid phase of centrifuged

346 digestates. In contrast, the samples with the highest OM contents were the solid phase of

347 centrifuged digestates and raw digestates. Except for OM contents between 0 and 250 g

348 OM.kgDM^{-1}, which did not correspond to any AD digestates, the OM model range covered

349 the OM contents observed in AD digestates with a homogeneous distribution of the

350 reference values and with a high density of samples. The TOC content represented also a

351 wide range of values with a balanced repartition between 163 and 500 gC.kgDM^{-1}. Because

352 OM and TOC contents are correlated, the same digestates were reported for the lowest and

353 the highest values of both TOC and OM content. Just like the OM content range, no

354 digestate with a TOC content lower than 163 gC.kgDM^{-1} was reported. Even though the SEL

355 of the OM and TOC measurements were low, the RMSEP of the predictions were for both

356 models acceptable. Therefore, the calibrated PLS models predicting OM content and TOC

357 contents can be expected to produce satisfying predictions over the whole diversity of AD

358 digestates.

359 The models predicting N$_{org}$, P, and K were calibrated with only 75, 65, and 62 samples,

360 respectively (Table 2). The N$_{org}$ content of the calibration samples varied between 5 and 74

361 gN.kgDM^{-1} and between 13 and 62 gN.kgDM^{-1} for the validation samples. The distribution of

362 the N$_{org}$ reference values was homogeneous and with an adequate density between 5 to 40

363 gN.kgDM^{-1} (Figure 2). However, there was a lack of digestates with high N$_{org}$ content

364 (between 40 and 80 gN.kgDM^{-1}) like digestates produced by AD plants fed with mainly food

365 wastes or slurry. This was the major reason for the lower prediction accuracy for digestates

366 with high N$_{org}$ contents compared to digestates with lower N$_{org}$ contents. Another reason

367 could be the residual ammonium present in the freeze-dried samples used for NIR

368 measurements, which could lead to disturbances for the N$_{org}$ predictions. Indeed, digestates

369 with high N$_{org}$ contents were mainly digestates also containing the most ammonium, and the

370 absorption areas caused by ammonium molecules (1500-1550 nm and 1950-2050 nm,

371 respectively) are similar to those produced by organic nitrogen compounds (Charnier et al.,

372 2018). Despite the need of digestates with higher N_{org} content to improve the prediction

373 accuracy, the RMSEP of the model was less than three times higher than the SEL, which

374 means that NIR predictions were less precise than the laboratory analyses, but could give a

375 usable estimation for routine analyses.

376 The P content of the digestates used in this study varied between 2 and 45 gP.kgDM^{-1}. The

377 repartition was balanced between 2 and 25 gP.kgDM^{-1}, but only seven calibration samples

378 had a P content between 25 and 45 gP.kgDM^{-1} (Figure 2). Digestates with a high P content

379 came from AD units mostly fed with sewage sludge or agri-food industry waste. These kinds

380 of samples are under-represented in the calibration base and including them would certainly

381 improve the model calibration. Nonetheless, the model gave predictions as accurate as the

382 laboratory analyses; these results are, therefore, very satisfactory.

383 The distribution of the K reference values was homogeneous between 0.5 and 75 gK.kgDM^{-1},

384 but there was an important lack of K contents between 75 and 172 gK.kgDM^{-1} (Figure 2).

385 Both, calibration and validation, gave poor results with a RMSEP five times higher than the

386 SEL. The NIR model predicting the K content will not give results accurate enough for further

387 use in practice.

388 NIR models predicting P and K contents in digestates were both based on indirect

389 correlations due to the lack of absorption zones in NIR wavelengths for these two elements

390 (Mouazen et al., 2016). However, these two models showed very different results. This

391 difference could be explained by the fact that P is partly integrated into organic matter,

392 while K is an element contained only in ionic form in the different organic matrices.

393 The different models were calibrated using NIR-spectra measured on freeze-dried samples

394 to avoid water-related disturbances, which implied that predicted parameters must not be

395   altered by the freeze-drying process. The OM and TOC laboratory analyses were performed

396   on freeze-dried samples, because there was no risk of organic matter and carbon loss during

397   the drying process. Indeed, most of volatile organic compounds like volatile fatty acids are

398   consumed during AD. However, it would have been difficult to predict total nitrogen or

399   ammonium contents from freeze-dried samples because part of the nitrogen can be lost

400   during the freeze-drying process due to ammonia volatilization (Morris et al., 2019) and

401   would no longer be visible on the NIR spectra. For that reason, only N_{org} was predicted in this

402   study. Nevertheless, the determination of ammonium is essential for the assessment of the

403   agronomic potential of digestates and could be determined with a laboratory method in

404   addition to the NIR analysis.

405   Comparing the results of the models calibrated on digestates with those obtained with

406   models calibrated on other matrices with a similar methodology, several observations can be

407   made. Charnier et al. (2017) proposed a model predicting total nitrogen from anaerobic

408   digestion feedstocks with an R^2_{val} (0.77) and RMSEP (8.6 gN.kgDM^{-1}) similar to those

409   proposed in this study. Gislum et al. (2004) found superior results for the prediction of the

410   total nitrogen content in two grasses species with specific models for each. In contrast, Ward

411   et al. (2011) proposed a common model for several meadow grasses and obtained much

412   poorer results. Due to the low mineral nitrogen content of the matrices used in these three

413   studies, the authors were able to predict the total nitrogen content instead of the organic

414   nitrogen content predicted in our study. Concerning the prediction of P and K contents,

415   Ward et al. (2011) also found lower performances than for the total nitrogen model.

416   Contrary to our study, the authors obtained more accurate predictions for the K model than

417   for the P model, which could be explained by stronger indirect correlations with NIR spectra

418   for the K content than for the P content according to the authors.

419   In full-scale AD plants with variable feedstocks (depending on the seasons, the feed-in

420   contracts established etc.), it is difficult to anticipate the digestates that will be produced.

421   The results obtained in this study demonstrate the feasibility of using NIR spectroscopy for

422   the determination of OM, TOC, N_{org}, and P contents in digestates. The calibration bases still

423   need to be complemented with other samples to achieve the performance required at

424   industrial level, but the NIR-PLS method could allow for fast and inexpensive agronomic

425   characterization of the digestates in line with the needs and opportunities of the AD sector.

426   It could thus enable more efficient management of digestate spreading in order to fit into a

427   chemical fertilizer reducing approach and to achieve an agricultural and industrial symbiosis.

428   This new method would also allow for a more frequent or even systematic characterization

429   of digestates, making their use in agriculture much more precise and limiting the

430   environmental impacts that can be caused by an uncontrolled return to the soil.

431   **3.4 Perspectives**

432   The NIR-PLS model prediction method provides many benefits in terms of analysis time and

433   cost. Indeed, a single NIR spectroscopy analysis can determine several parameters

434   simultaneously. The most time consuming step is to freeze-dry the samples. This

435   characterization method can still be optimized by finding a way to remove water-related

436   disturbances and, thus, performing the NIR measurements directly on fresh samples. This

437   improvement path is a current research subject and requires first identifying the effects of

438   water (Mallet et al., 2021) on the NIR signal in order to subtract them afterwards from the

439   NIR signal. This could also allow the calibration of an ammonium prediction model in

440   addition to organic nitrogen, as ammonium is at the very heart of spreading strategies

441   reducing the risk of acidification and eutrophication of natural environments. For instance,

442   previous studies, such as Saeys et al. (2005) and Maleki et al. (2006), have demonstrated

443   that accurate models could be obtained using both NIR and visible spectroscopy for

444   predicting OM, nitrogen, NH_4, and P on fresh samples. As NIR spectroscopy is becoming a

445   more and more financially accessible technology (Mouazen et al., 2016), a NIR

446   characterization method on fresh samples could allow the development of on-line and on-

447   site sensors.

448

449   **4. Conclusion**

450   In this study, a data set containing OM, TOC, N_{org}, P, and K contents of 193 digestates has

451   been created. The freeze-dried digestates were analyzed with NIR spectroscopy with the aim

452   of calibrating NIR-PLS models. Cross-validation and validation with an independent data set

453   were used to evaluate these models. The predictive models showed good prediction

454   accuracy for the estimation of OM, TOC, and P contents in digestates. The predictions of N_{org}

455   were less accurate than laboratory analyses, but usable for some applications. The P and N_{org}

456   models should be improved in future developments by adding more samples to the

457   calibration data set. Finally, calibration of the model predicting the K content was considered

458   unsuccessful, as it did not produce satisfactory results as it has also been shown in other

459   studies. This paper proposes the first steps towards an analytical method for the

460   characterization of the fertilizing value of digestates using NIR spectroscopy.

461

462   **Acknowledgments**

467

468   **References**

469   Alburquerque, J. A., de la Fuente, C., Ferrer-Costa, A., Carrasco, L., Cegarra, J., Abad, M., &

470   Bernal, M. P. (2012). Assessment of the fertiliser potential of digestates from farm and

471   agroindustrial residues. Biomass and bioenergy, 40, 181-189.

472   https://doi.org/10.1016/j.biombioe.2012.02.018

473   Awhangbo, L., Bendoula, R., Roger, J. M., & Béline, F. (2020). Multi-block SO-PLS approach

474   based on infrared spectroscopy for anaerobic digestion process monitoring. Chemometrics

475   and Intelligent Laboratory Systems, 196, 103905.

476   https://doi.org/10.1016/j.chemolab.2019.103905

477   Barthès, B. G., Kouakoua, E., Clairotte, M., Lallemand, J., Chapuis-Lardy, L., Rabenarivo, M., &

478   Roussel, S. (2019). Performance comparison between a miniaturized and a conventional

479   near infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen.

480   Geoderma, 338, 422-429. https://doi.org/10.1016/j.geoderma.2018.12.031

481   Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., & McBratney, A. (2010).

482   Critical review of chemometric indicators commonly used for assessing the quality of the

483   prediction of soil attributes by NIR spectroscopy. TrAC Trends in Analytical Chemistry, 29(9),

484   1073-1081. https://doi.org/10.1016/j.trac.2010.05.006

485   Charnier, C., Latrille, E., Jimenez, J., Lemoine, M., Boulet, J. C., Miroux, J., & Steyer, J. P.

486   (2017). Fast characterization of solid organic waste content with near infrared spectroscopy

487   in anaerobic digestion. Waste management, 59, 140-148.

488   https://doi.org/10.1016/j.wasman.2016.10.029

489   Charnier, C., Latrille, E., Roger, J. M., Miroux, J., & Steyer, J. P. (2018). Near-Infrared

490   Spectrum Analysis to Determine Relationships between Biochemical Composition and

491   Anaerobic Digestion Performances. Chemical Engineering & Technology, 41(4), 727-738.

492   https://doi.org/10.1002/ceat.201700581

493   Ciurczak, E. W., & Burns, D. A. (2007). Handbook of near-infrared analysis. CRC Press.Dalal, R.

494   C., & Henry, R. J. (1986). Simultaneous Determination of Moisture, Organic Carbon, and Total

495   Nitrogen by Near Infrared Reflectance Spectrophotometry 1. Soil Science Society of America

496   Journal, 50(1), 120-123. https://doi.org/10.2136/sssaj1986.03615995005000010023x

497   Cozzolino, D., & Moron, A. (2004). Exploring the use of near infrared reflectance

498   spectroscopy (NIRS) to predict trace minerals in legumes. Animal Feed Science and

499   Technology, 111(1-4), 161-173. https://doi.org/10.1016/j.anifeedsci.2003.08.001

500   Dardenne, P. (2010). Some considerations about NIR spectroscopy: Closing speech at NIR-

501   2009. NIR news, 21(1), 8-14. https://doi.org/10.1255/nirn.1165

502   European Parliament, Council of the European Union, 2016. Procedure 2016/0084/ COD.

503   COM (2016) 157: Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE

504   COUNCIL laying down rules on the making available on the market of CE marked fertilising

505   products and amending Regulations (EC) No 1069/2009 and (EC).

506   European Commission, 2003. Regulation (EC) No 2003/2003 of the European Parliament and

507   of the Council of 13 October 2003 relating to fertilizers, Off J Eur Commun.

508   Fuchs, W., & Drosg, B. (2013). Assessment of the state of the art of technologies for the

509   processing of digestate residue from anaerobic digesters. Water Science and Technology,

510   67(9), 1984–1993. https://doi.org/10.2166/wst.2013.075

511   Gislum, R., Micklander, E., & Nielsen, J. P. (2004). Quantification of nitrogen concentration in

512   perennial ryegrass and red fescue using near-infrared reflectance spectroscopy (NIRS) and

513   chemometrics. Field Crops Research, 88(2-3), 269-277.

514   https://doi.org/10.1016/j.fcr.2004.01.021

515   González-Martín, I., Hernández-Hierro, J. M., & González-Cabrera, J. M. (2007). Use of NIRS

516   technology with a remote reflectance fibre-optic probe for predicting mineral composition

517   (Ca, K, P, Fe, Mn, Na, Zn), protein and moisture in alfalfa. Analytical and bioanalytical

518   chemistry, 387(6), 2199-2205. https://doi.org/10.1007/s00216-006-1039-4

519   Guilayn, F., Jimenez, J., Martel, J. L., Rouez, M., Crest, M., & Patureau, D. (2019). First

520   fertilizing-value typology of digestates: A decision-making tool for regulation. Waste

521   Management, 86, 67-79. https://doi.org/10.1016/j.wasman.2019.01.032

522   Guilayn, F., Rouez, M., Crest, M., Patureau, D., & Jimenez, J. (2020). Valorization of

523   digestates from urban or centralized biogas plants: a critical review. Reviews in

524   Environmental Science and Bio/technology. https://doi.org/10.1007/s11157-020-09531-3

525   He, Y., Huang, M., García, A., Hernández, A., & Song, H. (2007). Prediction of soil

526   macronutrients content using near-infrared spectroscopy. Computers and Electronics in

527   Agriculture, 58(2), 144-153. https://doi.org/10.1016/j.compag.2007.03.011

528   Houot, S., Pons, M. N., Pradel, M., & Tibi, A. (2016). Recyclage de déchets organiques en

529   agriculture: Effets agronomiques et environnementaux de leur épandage. Editions Quae.

530   Hummel, J. W., Sudduth, K. A., & Hollinger, S. E. (2001). Soil moisture and organic matter

531   prediction of surface and subsurface soils using an NIR soil sensor. Computers and

532   electronics in agriculture, 32(2), 149-165. https://doi.org/10.1016/S0168-1699(01)00163-6

533   Jacobi, H. F., Moschner, C. R., & Hartung, E. (2011). Use of near infrared spectroscopy in

534   online-monitoring of feeding substrate quality in anaerobic digestion. Bioresource

535   technology, 102(7), 4688-4696. https://doi.org/10.1016/j.biortech.2011.01.035

536   Jimenez, J., Grigatti, M., Boanini, E., Patureau, D., & Bernet, N. (2020). The impact of biogas

537   digestate typology on nutrient recovery for plant growth: Accessibility indicators for first

538   fertilization prediction. Waste Management, 117, 18-31.

539   https://doi.org/10.1016/j.wasman.2020.07.052

540   Kataki, S., Hazarika, S., & Baruah, D. C. (2017). Investigation on by-products of bioenergy

541   systems (anaerobic digestion and gasification) as potential crop nutrient using FTIR, XRD,

542   SEM analysis and phyto-toxicity test. Journal of Environmental Management, 196, 201-216.

543   https://doi.org/10.1016/j.jenvman.2017.02.058

544   KTBL, (2008). Umweltgerechte, innovative Verfahren zur Abtrennung von Nährstoffen aus

545   Gülle und Gärrückständen - Technologischer Stand, Perspektiven und

546   Entwicklungsmöglichkeiten. Studie im Auftrag der Deutschen Bundesstiftung Umwelt,

547   erstellt durch das Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL),

548   Darmstadt, D, in Zusammenarbeit mit dem Institut für Technologie und Biosystemtechnik

549   der Bundesforschungsanstalt für Landwirtschaft (FAL), Braunschweig, D (only in German).

550 Lesteur, M., Latrille, E., Maurel, V. B., Roger, J. M., Gonzalez, C., Junqua, G., & Steyer, J. P.

551 (2011). First step towards a fast analytical method for the determination of biochemical

552 methane potential of solid wastes by near infrared spectroscopy. Bioresource

553 technology, 102(3), 2280-2288. https://doi.org/10.1016/j.biortech.2010.10.044

554 Maleki, M. R., Van Holm, L., Ramon, H., Merckx, R., De Baerdemaeker, J., & Mouazen, A. M.

555 (2006). Phosphorus sensing for fresh soils using visible and near infrared spectroscopy.

556 Biosystems Engineering, 95(3), 425-

557 436.https://doi.org/10.1016/j.biosystemseng.2006.07.015

558 Mallet, A., Charnier, C., Latrille, E., Bendoula, R., Steyer, J. P., Roger J.M. (2021). Unveiling

559 non-linear water effects in near infrared spectroscopy: A study on organic wastes during

560 drying using chemometrics. Waste Management, 122, 36-48.

561 https://doi.org/10.1016/j.wasman.2020.12.019.

562 Menesatti, P., Antonucci, F., Pallottino, F., Roccuzzo, G., Allegra, M., Stagno, F., & Intrigliolo,

563 F. (2010). Estimation of plant nutritional status by Vis–NIR spectrophotometric analysis on

564 orange leaves [Citrus sinensis (L) Osbeck cv Tarocco]. biosystems engineering, 105(4), 448-

565 454. https://doi.org/10.1016/j.biosystemseng.2010.01.003

566 Monlau, F., Sambusiti, C., Ficara, E., Aboulkas, A., Barakat, A., & Carrere, H. (2015). New

567 opportunities for agricultural digestate valorization: current situation and

568 perspectives. Energy & Environmental Science, 8(9), 2600-2621.

569 https://doi.org/10.1039/C5EE01633A

570 Morris, D. L., Tebbe, A. W., Weiss, W. P., & Lee, C. (2019). Short communication: Effects of

571 drying and analytical methods on nitrogen concentrations of feeds, feces, milk, and urine of

572 dairy cows. Journal of Dairy Science. https://doi.org/10.3168/jds.2019-16256

573   Mouazen, A. M., & Kuang, B. (2016). On-line visible and near infrared spectroscopy for in-

574   field phosphorous management. Soil and Tillage Research, 155, 471-477.

575   https://doi.org/10.1016/j.still.2015.04.003

576   Murphy, J., Riley, J.P. 1962. A modified single solution method for the determination of

577   phosphate in natural waters. Analytica Chimica Acta, 27, 31-36.

578   https://doi.org/10.1016/S0003-2670(00)88444-5

579   Nduwamungu, C., Ziadi, N., Parent, L. É., & Tremblay, G. F. (2009). Mehlich 3 extractable

580   nutrients as determined by near-infrared reflectance spectroscopy. Canadian journal of soil

581   science, 89(5), 579-587. https://doi.org/10.4141/CJSS09018

582   Nkoa, R. (2014). Agricultural benefits and environmental risks of soil fertilization with

583   anaerobic digestates: a review. Agronomy for Sustainable Development, 34(2), 473-492.

584   https://doi.org/10.1007/s13593-013-0196-z

585   Petisco, C., García-Criado, B., De Aldana, B. V., Zabalgogeazcoa, I., & Mediavilla, S. (2005).

586   Use of near-infrared reflectance spectroscopy in predicting nitrogen, phosphorus and

587   calcium contents in heterogeneous woody plant species. Analytical and bioanalytical

588   chemistry, 382(2), 458-465. https://doi.org/10.1007/s00216-004-3046-7

589   Reed, J. P., Devlin, D., Esteves, S. R. R., Dinsdale, R., & Guwy, A. J. (2011). Performance

590   parameter prediction for sewage sludge digesters using reflectance FT-NIR

591   spectroscopy. Water research, 45(8), 2463-2472.

592   https://doi.org/10.1016/j.watres.2011.01.027

593   Saeys, W., Xing, J., De Baerdemaeker, J., Ramon, H., 2005. Comparison of transflectance and

594   reflectance to analyse hog manures. Journal of Near Infrared Spectroscopy 13, 99e107.

595   https://doi.org/10.1255/jnirs.462

596   Teglia, C., Tremier, A., & Martel, J. L. (2011). Characterization of solid digestates: part 1,

597   review of existing indicators to assess solid digestates agricultural use. Waste and Biomass

598   Valorization, 2(1), 43-58. https://doi.org/10.1007/s12649-010-9051-5

599   Tenenhaus, M. (1998). La régression PLS: théorie et pratique. Editions technip.

600   The Commission of the European Communities, 2009. Directive 2009/28/EC of 23 April 2009.

601   Off. J. Eur. Union 140, 47.

602   Tremblay, G. F., Nie, Z., Belanger, G., Pelletier, S., & Allard, G. (2009). Predicting timothy

603   mineral concentrations, dietary cation-anion difference, and grass tetany index by near-

604   infrared reflectance spectroscopy. Journal of dairy science, 92(9), 4499-4506.

605   https://doi.org/10.3168/jds.2008-1973

606   Walsh, J. J., Jones, D. L., Edwards-Jones, G., & Williams, A. P. (2012). Replacing inorganic

607   fertilizer with anaerobic digestate may maintain agricultural productivity at less

608   environmental cost. Journal of Plant Nutrition and Soil Science, 175(6), 840–845.

609   https://doi.org/10.1002/jpln.201200214

610   WANG, K. (2009). Application and prospect of near infrared reflectance spectroscopy in

611   forage analysis. Spectroscopy and Spectral Analysis, 29(3), 635-640.

612   https://doi.org/10.3964/j.issn.1000-0593

613   Ward, A., Nielsen, A. L., & Møller, H. (2011). Rapid assessment of mineral concentration in

614   meadow grasses by near infrared reflectance spectroscopy. Sensors, 11(5), 4830-4839.

615   https://doi.org/10.3390/s110504830

616   Wegelin, J. A. (2000). A survey of Partial Least Squares (PLS) methods, with emphasis on the

617   two-block case. University of Washington, Tech. Rep.

618   Williams, P., & Norris, K. (1987). Near-infrared technology in the agricultural and food

619   industries. American Association of Cereal Chemists, Inc..Malley, D. F., Martin, P. D., & Ben-

620   Dor, E. (2004). Application in analysis of soils. Near-infrared spectroscopy in agriculture,

621   (nearinfraredspe), 729-784.

622