Nutrient quantification of livestock slurry and digestate using a low-cost handheld NIR spectrometer and multi-biomass calibration models

Francesco Tangorra [a], Eleonora Buoio [a], Alberto Finzi [b]*, Annamaria Costa [a], Omar Ferrari [b], Giorgio Provolo [b]

[a] *Department of Veterinary medicine and Animal Sciences. University of Milan, Via dell'Università 6, 26900 Lodi, Italy.*

[b] *Department of Agricultural and Environmental Sciences. University of Milan, Via Celoria 2, 20133 Milan, Italy.*

[*] *Corresponding author: alberto.finzi@unimi.it*

Abstract

Real-time analysis during slurry field application can potentially ensure higher efficiency in crop fertilisation and reduce related environmental problems. Near Infrared Spectroscopy (NIRS) can predict quickly and adequately the nutrient content of the slurry but the devices are still costly. In this study, a low-cost portable device has been tested to predict total solids, total Kjeldahl nitrogen, total ammonia nitrogen, and phosphorus contents in cattle and pig slurries and digestate. Multivariate calibration models were developed for different biomass types (multi-biomass) by comparing support vector machine (SVM) and partial least square (PLS) regressions. The developed SVM and PLS models showed satisfactory and similar results, with performance to deviation ratio (RPD) values ranging between 2.22 and 2.80. The use of the handheld device with the obtained models also meets the requirements of a certification protocol for commercial NIR sensors and therefore can ensure a better slurry management at farm level.

Key Words. Near infrared spectroscopy (NIRS), support vector machine (SVM), slurry, digestate, nutrient prediction, field device

1. INTRODUCTION

In the practice of using manure as crop fertiliser, knowing its nutrient content through real-time analysis during slurry spreading, could reduce potential environmental problems. Thereby also ensuring higher efficiency in crop fertilisation (Burton and Turner 2003; Xing et al., 2008).

Standard laboratory methods for livestock slurry analysis are expensive and time consuming making the precise nutrient field application difficult to realise (Saeys et al., 2005; Feng et al., 2022). Near-Infrared Spectroscopy (NIRS) provides an alternative technique to the standard analysis for quickly determining the composition of

livestock slurry samples (Albrecht et al., 2009). NIRS measures the spectral absorption band of electromagnetic radiation in the range 750-2500 nm, which can be related to the analytical properties of the sample through chemometrics and predictive models can be developed (Peltre et al., 2011). The most commonly used method is the partial least squares (PLS) regression. It is a linear multivariate calibration method, efficient on spectral data and easy to implement (Pasquini, 2018), but it reaches its limits when the signal is disturbed (e.g. variation in particle size or temperature), the concentrations are close to the detection threshold or the correlations are non-linear (Croguennoc et al., 2019). Recently, support vector machine (SVM) regression resulted to be an alternative to PLS for non-destructive C, N, P, and K determination in poultry litter samples within an excellent or acceptable range of predictability (Bedin et al., 2021). The SVM is flexible and powerful computational algorithm for solving machine learning problems, initially created for the development of classification models but quickly adapted to solve regression problems in application with real data, performing a minimization of the effects caused by outliers (Mammone et al., 2009).

Laboratory benchtop NIR instruments, both with dispersive and Fourier Transform (FT) technology, usually operating in the range 1000-2500 nm, provide good results for dry matter (DM), total carbon (TC) and, to a lesser extent, for ash, total nitrogen (TKN) and organic nitrogen (ON) (Cabassi et al., 2015). In recent years, scientific and technological developments have led to portable NIR spectrometers based on diode arrays, CCD (charge-coupled device) detectors, MEMS (micro electro-mechanical systems) interferometers equipped with InGaAs (Indium Gallium Arsenide) detectors which offer significant advantage in terms of price and size (Cabassi et al., 2015; Pasquini 2018; Bec et al. 2020). Nevertheless, handheld NIR spectrometers usually operate between 960 and 1800 nm and have narrower spectral range and lower resolution compared to benchtop NIR instruments, precluding their practical use for slurry nutrient content estimation. The 1800-2498 nm region has been identified as essential for determining dairy manure nutrients, especially for the Ammonium-N which constitutes a large fraction of total N (Reeves and Van Kessel, 2000). The N-H bonds such as the N-H stretching vibration at 2050 to 2060 nm and the N-H deformation second overtone at 2168 nm to 2180 nm mainly contribute to NIRS predictions of N-related constituents (Chen et al., 2013).

The objective of this study was to develop and validate a multivariate and multi-biomass calibration method based on a low-cost handheld FT-NIR spectrometer (NeoSpectra - Si-Ware, Egypt), operating over the 1350-2558 nm wavelength range, which is not easy to implement on this type of spectrometers. Furthermore, the PLS regression models that are commonly adopted for directly determining nutrients in livestock slurries and digestates, were compared with the SVM which are machine learning algorithms still rarely applied in this area. Moreover, an evaluation of its performance through a commercial certification protocol was carried out with a perspective of future implementation at farm level.

2. MATERIALS AND METHODS

2.1 Sampling and reference analysis

A total of 53 samples, including 16 cattle slurry, 21 pig slurry and 16 digestate were collected from 36 farms in Lombardy (Italy). Samples of 10 L were taken from slurry reception pits on dairy and pig farms, while digestate was collected directly from digesters. The farms characteristics are reported in Table S1.

After collection, samples were refrigerated and delivered to the laboratory on the day of sampling. A fraction of each sample was stored at +4°C and reference analyses were performed within 24 hours. The samples were analysed for total solids (TS), total Kjeldahl nitrogen (TKN), total ammoniacal nitrogen (TAN), phosphorus (P) and potassium (K) using standard methods (APHA, 2012). The remaining portion was frozen at – 20 °C until NIR measurements were started.

2.2 NIR spectra acquisition

Slurry and digestate samples were scanned using a NeoSpectra portable spectrometer (Si-Ware, Cairo, Egypt). The NeoSpectra Micro Development Kit consists of three tungsten halogen lamps, a monolithic micro-electro-mechanical system (MEMS) Michelson interferometer and a single InGaAs photodetector. The NeoSpectra Micro is connected to a Raspberry Pi that acts as a host and allows connection via a universal serial bus (USB) to a laptop. The software (Windows and Linux) allows setting the following parameters: scan time, run mode (single or continuous) and data interpolation in each spectrum collected. The wavelength range is from 1350 to 2558 nm and the resolution was set to 16 nm (measured at 1550 nm).

Prior to the first measurement, a background measurement was collected using a Spectralon (99% reflectance). Scanning time was set at 2 s. Frozen slurry samples were thawed at room temperature and carefully stirred manually. Each sample was scanned from the bottom of a Petri dish, filled with 40 ml, in reflectance mode. Forty spectra were collected for each sample (10 scans for each quarter rotation of the Petri dish) and the average spectrum was used for data analysis.

Different spectral pre-processing techniques were tested (Savitzky-Golay smoothing, Standard Normal Variate, Savitzky-Golay first and second derivative) to remove any irrelevant information which can negatively affect the regression models. Considering the inhomogeneous physical structure of the samples, the Savitzky and Golay first derivative with a second-degree polynomial order and 10 smoothing side points was the most suitable pre-treatment able to enhance the resolution, minimise the offsets and the global intensity effects typically arising from unwanted light scattering (Oliveri et al., 2019).

2.3 Multivariate calibration models

PLS and SVM regressions were performed to obtain multi-biomass (cattle slurry, pig slurry, digestate) calibration models using the statistical software JMP Pro 16.2 (SAS Institute, Cary, USA) with the spectral tools developed by Worley (2021). The data set was randomly split into two subsets: the calibration set consisting of about two thirds of the samples (35) and the validation set with the remaining third of samples (18).

Regarding PLS, for each chemical constituent, an individual model was developed using the number of PLS factors that determined the lowest error in cross-validation, considering the results of the predicted residual error sum square (PRESS). To test the robustness of the PLS models, the PLS regression was k-fold cross-validated (k = 7). SVM regression was applied using the linear kernel function, taking into account the linear relationship between absorbance and concentration exhibited by most biological and agricultural materials (Pasquini, 2018). The model parameters were tuned according to the JMP default setting and the procedure was repeated 20 times. For each chemical constituent the best performing regression model was identified based on the highest coefficient of determination (R^2) and the lowest root average squared error (RASE).

The performance of the PLS and SVM regression models for TS, TKN, TAN, P and K were compared. For each model, the coefficient of determination (R^2_{cal} and R^2_{val}), the root mean square error of calibration ($RMSE_C$) and prediction ($RMSE_P$) were obtained. Goodness and accuracy of the models were tested using the ratio of performance to deviation (RPD), calculated as the ratio of the standard deviation (SD) of the reference values to the RMSE (Williams and Sobering, 1996). Based on the RPD values, six quality-classes of calibration model were identified (Cabassi et al., 2015): three suitable for predicting slurry properties (Excellent, Successful and Useful, with RPD > 4, between 3 and 4, and between 2.2 and 3, respectively); two suitable for screening purposes, for distinguishing among low-high values or selecting samples for costly conventional analysis (Moderately useful and Screening, with RPD between 1.7 and 2.2, and between 1.5 and 1.7, respectively); one class (Poor, RPD ≤ 1.5) not acceptable for practical use.

2.4 Potential practical application

Any commercial NIR sensor installable on slurry tankers can be certified to prove its accuracy in predicting TS, TKN, TAN, P, and K content in cattle manures, pig manures, and digestates. This certification sealed by the German Agricultural Society (DLG) assures that at least 60% of the measured samples differ less than 25% from the reference lab value and no measured sample exceeds the 35% of relative deviation (Horf et al., 2022). The performance of NeoSpectra was evaluated using the validation set for both PLS and SVM models, distinguishing its predictive capacity on TS, TKN, TAN, P for each type of slurry but also for all the slurries together.

3. RESULTS AND DISCUSSION

128

129 **3.1 Descriptive statistics**

130 Descriptive statistics for the chemical constituents of all samples and each biomass separately are reported in
131 Fig. 1. The datasets show good variability for the parameters analysed and are representative of livestock slurries
132 produced in Lombardy (Martinez-Suller, et al., 2008; Cabassi et al., 2015; Finzi et al., 2015).

133 Table 1 shows the Pearson correlation coefficient among the 5 parameters. TKN and TAN were highly and
134 positively correlated with TS in cattle and pig slurries as reported in literature (Piccinini and Bortone 1991;
135 Martinez-Suller et al., 2008). The correlation of TAN with TS content was lower compared to TKN, as expected
136 considering that ammonium N is in the liquid phase of manures (Williams et al., 1996; Martinez-Suller et al.,
137 2008). In agreement with the findings of other researchers (Scotford et al., 1998; Moral et al., 2005) P content
138 was well correlated with TS content both in cattle and pig slurries. A positive significant correlation was found
139 also between K and TS content, in accordance with Martinez-Suller et al. (2008). This has been confirmed also
140 for digestate where the TS content resulted moderately correlate with the K content. Contrarily the correlations
141 among TS, TKN, TAN and P were not significant for digestates. This result can be explained by the difference in
142 the composition of the substances used to feed the digesters.

143

144 **3.2 Spectra acquisition**

145 The average NIR raw spectra of the cattle and pig slurry and digestate samples are shown in Fig. 2. The main
146 broad peak at around 1450 nm for the three biomass sets results from the absorption by the O-H bonds of the
147 water. Cattle slurry and digestate samples, although having a higher total solids content than the pig slurry
148 samples (respectively 7.72 ± 1.48 % and 5.19 ± 1.15 % vs. 2.81 ± 1.83 %) show higher water peaks. As reported
149 by Saeys et al. (2005) it is likely that when more solid particles are present more light is scattered and reflected
150 back to the detector resulting in lower overall absorption.

151

152 3.3 Predictive models

153 The results obtained for the multi-biomass (cattle slurry, pig slurry, digestate) PLS and SVM models are
154 summarised in Tables 2 and 3 respectively. The average RMSE and R^2 results of PLS and SVM models are similar
155 with a slight advantage for the latter as highlighted by the model classification. Based on the RPD values, all SVM
156 prediction models can be classified as useful for slurry properties prediction, whereas PLS prediction models for
157 TKN and TAN, having RPD values equal respectively to 2.0 and 2.1, can be classified only as moderately useful for
158 screening purposes. However, the PLS models needs a low number of factors, generally lower than 6, suggesting
159 that the models are not complex. For the variable TS both prediction models showed R^2_C and R^2_V respectively

greater than 0.80 and 0.70 and RPD values greater than 2.2 making the Neospectra portable spectrometer useful for TS estimation of livestock slurries. Because the O-H vibration of water has high absorption intensity in the NIR region, due to the inverse correlation between TS and moisture, good predictions on total solids content of livestock manure are usually expected by NIR spectroscopy (Chen et al., 2013). Our results are comparable with other studies. Mouazen et al., (2005), using a FT-spectrometer (1000-2500 nm) and PLS regression on fresh swine manure obtained R^2 in cross-validation of 0.79 and RPD values equal to 2.19. Better results (R^2_V = 0.91 and RPD = 3.22) were obtained on fresh swine manure by Saeys et al. (2005) using an on-site NIR instrument (Zeiss Corona 45 visNIR 1.7) but in the region 426-1683 nm.

NIR spectroscopy prediction of N-related constituents are mainly contributed to by N-H bonds (N-H stretch first overtone at 1500-1530 nm, N-H stretching vibration at 2050-2060 nm, N-H deformation second overtone at 2168-2180 nm) (Chen et al., 2013). Model statistics for TKN and TAN parameters are similar to those obtained by other researchers. For example, Saeys et al. (2005) on fresh swine manure obtained R^2_V and RPD values respectively of 0.86 and 2.63 for total-N and of 0.76 and 2.00 for ammonia-N using a Zeiss Corona 45 vis-NIR 1.7 in the region 426-1683 nm. R^2_{CV} and RPD for TN in the range 0.80 to 0.92 and 3.32 to 4.87 respectively for various types of manure samples including beef cattle manures, swine solid manures, swine slurry samples, etc. were obtained by Ye et al. (2005) but using a benchtop instrument (FOSS-NIRSystem model 6500 spectrometer).

Some studies (Dagnew et al., 2004; Saeys et al., 2004) obtained unreliable predictions on P content in swine manure with R^2_V and RPD respectively lower than 0.80 and 2.00. The main reason for the inability to predict P content was attributed to the lack of a direct relationship between P and C-H-O-N bonds (Reeves, 2001). Nevertheless, other studies on swine manure (Saeys et al., 2005; Mouazen et al., 2005) achieved a good prediction of P content with R^2_{CV} and RPD respectively ranging between 0.81 and 0.85 and between 2.33 and 2.66. The successful predictions of P in livestock manure could be related to the good relationship existing between P and TS in these matrices, as observed by different researchers (Marino et al., 2008; Martinez-Suller et al., 2008). Considering that TS can be accurately predicted by NIR spectroscopy due to the presence of O-H and C-H bonds, the prediction of P content in livestock manure would be indirect (Chen, et al., 2013). Our results with R^2_V of 0.78 and RPD_V greater than two seem to support this theory.

For the variable K, only a SVM calibration model was obtained (R^2_C=0.59 and RPD=1.92). Theoretically, no absorption bands for metal species are present in the NIR region (Chen et al., 2013) and only a limited number of studies with animal manure compost (Huang et al., 2008) have obtained satisfactory results.

Scatter plots of NIR predicted vs. measured values of TS, TKN, TAN, and P using PLS and SVM models are presented in Fig. 3. The solid black line in the graphs highlights an individual correspondence between predicted

and measured values: the closer to this line, the more sample points, the more accurate the model. SVM performs slightly better than the PLS (Fig. 3). The reason could be that SVM is less prone to overfitting and less affected by outliers compared to other algorithms. Overall, results suggest that the SVM method may be an alternative for NIR calibration and TS, TKN, TAN and P quantification in cattle and pig slurry and digestate.

3.4 Potential practical application

As can be seen in Table 4, Neospectra meets the requirements for DLG certification showing a promising predictive accuracy, although not in all conditions tested, in view of its potential use on slurry tankers. In the simulation carried out, different performance were achieved based on the analytical parameter, the type of slurry and the predictive model used, SVM or PLS. NeoSpectra fulfilled the requirements for DLG approval in the accuracy of measurements of TS in cattle slurry using SVM model and in cattle slurry and digestate using PLS model; TKN in all conditions tested; TAN in all slurries together and in pig slurry using SVM model, while using PLS model all conditions except for cattle slurry; P in cattle slurry and digestate with both models. Considering all the slurries together, the SVM and PLS models are equivalent, confirming the comparable performance reported in Tables 2 and 3. Examining the single effluents, NeoSpectra shows better performance in predicting TS and TAN for digestate using the PLS model, while for pig slurry and cattle slurry no differences emerge between PLS and SVM models. These results demonstrate that NeoSpectra is eligible for DLG certification, meeting the criteria for the entry level of certification. The achievement of this threshold should be considered positively as commercial NIR sensors that have obtained DLG certification still present a high variability in the performance they can guarantee (Horf et al., 2022). Finally, it is important to point out that the scanning of samples in lab performs comparably with the on-field real-time analyses of slurry, because both measurement systems are based on scans in reflectance; in addition, during the field distribution the machines normally scan a flow of slurry, making available several scans that compensate a possible lower stability of the measurement.

4. CONCLUSIONS

The use of a single calibration model to predict TS, TKN, TAN and P content of multiple biomasses (cattle slurry, pig slurry, and digestate) has shown satisfactory performance. This approach simplifies the use of a handheld NIR spectrometer (Neospectra) and reduces the costs of both using and updating the models. The combination of the handheld low cost NIR spectrometer with PLS and SVM resulted useful for prediction of TS, TKN, TAN and P in cattle and pig slurries and digestate. This potentially enables real-time prediction of slurry and digestate composition at farm level and making the precise nutrient application in field possible. The test of Neospectra's

potential to fulfil the requirements for certification of commercial NIR sensors showed a slightly better performance of PLS than SVM model, but only for digestate.

References

Albrecht, R., Joffre, R., Le Petit, J., Terrom, G., & Périssol, C. (2009). Calibration of chemical and biological changes in cocomposting of biowastes using near-infrared spectroscopy. *Environmental Science and Technology*, 43(3), 804–811. https://doi.org/10.1021/es802064u

APHA. (2012). Standard Methods for the Examination of Water and Wastewater. American Public Health Association, American Water Works Association and Water Environment Federation: Washington, DC, USA.

Beć, K.B., Grabska, J., Siesler, H.W., Huck CW. (2020). Handheld near-infrared spectrometers: Where are we heading? *NIR news,* 31(3-4):28-35. DOI:10.1177/0960336020916815

Bedin, F. C. B., Faust, M.V., Guarneri, G.A., Assmann, T. S., Lafay, C. B. B., Soares, L. F., de Oliveira, P.A.V., dos Santos-Tonial, L.M. (2021). NIR associated to PLS and SVM for fast and non-destructive determination of C, N, P, and K contents in poultry litter. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 245, 118834. https://doi.org/10.1016/j.saa.2020.118834.

Burton, C.H., Turner, C., 2003. Manure Management: Treatment Strategies for Sustainable Agriculture, second ed. Silsoe Research Institute, Bedford, UK.

Cabassi, G., Cavalli, D., Fuccella, R., & Gallina, M. P. (2015). Evaluation of four NIR spectrometers in the analysis of cattle slurry. *Biosystems Engineering*, 133, 1–13. https://doi.org/10.1016/j.biosystemseng.2015.02.011

Chen, L., Xing, L., Han, L. (2013). Review of the Application of Near-Infrared Spectroscopy Technology to Determine the Chemical Composition of Animal Manure. *Journal of Environmental Quality,* 42, 1015–1028. DOI: 10.2134/jeq2013.01.0014

Croguennoc, A., Lallemand,J., Roussel S. (2019). Some aspects of SVM Regression: an example for spectroscopic quantitative predictions. Conference Chimiométrie, Montpellier, France. http://eigenvector.com/Docs/Chimiometrie_2019/A_Groguennoc_SVM_chimiometrie_2019.pdf

Dagnew, M. D., Crowe, T. G., & Schoenau, J. J. (2004). Measurement of nutrients in Saskatchewan hog manures using near-infrared spectroscopy. *Canadian Biosystems Engineering*, 46, 33-37. Retrieved from http://www.csbe-scgab.ca/docs/journal/46/c0308.pdf

Evangelista, C., Basiricò, L., Bernabucci, U. (2021). An Overview on the Use of Near Infrared Spectroscopy (NIRS) on Farms for the Management of Dairy Cows. *Agriculture,* 11(4):296. https://doi.org/10.3390/agriculture11040296

Feng, X., Larson, R.A., Digman, M.F. (2022). Evaluation of Near-Infrared Reflectance and Transflectance Sensing System for Predicting Manure Nutrients. *Remote Sensing*, 14, 963. https://doi.org/10.3390/rs14040963

Finzi, A., Oberti, R., Negri, A.S., Perazzolo, F., Cocolo, G., Tambone, F., Cabassi, G., Provolo, G. (2015). Effects of measurement technique and sample preparation on NIR spectroscopy analysis of livestock slurry and digestates. *Biosystems Engineering,* 134. https://doi.org/10.1016/j.biosystemseng.2015.03.015

Horf, M., Vogel, S., Drücker, H., Gebbers, R., Olfs, H.W. (2022). Optical Spectrometry to Determine Nutrient Concentrations and other Physicochemical Parameters in Liquid Organic Manures: A Review. *Agronomy*, 12. https://doi.org/10.3390/agronomy12020514

Huang, G., Han, L., Yang, Z., Wang, X. (2008). Evaluation of the nutrient metal content in Chinese animal manure compost using near infrared spectroscopy (NIRS). *Bioresource Technology*, 99:8164–9. DOI: 10.1016/j.biortech.2008.03.025

Mammone, A., Turchi, M., Cristianini, N. (2009). Support vector machines. *Wiley Interdisciplinary Reviews Computational* Statistics, 1, 283-289. https://doi.org/10.1002/wics.49

Marino, P., De Ferrari, G., Bechini, L. (2008). Description of a sample of liquid dairy manures and relationships between analytical variables. *Biosystems Engineering,* 100 (2) 256–265. https://doi.org/10.1016/j.biosystemseng.2008.02.012

Martínez-Suller, L., Azzellino, A., Provolo, G. (2008). Analysis of livestock slurries from farms across Northern Italy: relationship between indicators and nutrient content. *Biosystems Engineering,* 99 (4) 540–552. https://doi.org/10.1016/j.biosystemseng.2007.12.002

Moral, R., Perez-Murcia, M.D., Perez-Espinosa, A., Moreno-Caselles, J., Paredes, C., Rufete, B. (2005). Estimation of nutrient values of pig slurries in Southeast Spain using easily determined properties. *Waste Management*, 25,719–725. DOI: 10.1016/j.wasman.2004.09.010.

277 Mouazen, A.M., De Baerdemaeker, J., Ramon, H. (2005). Towards development of on-line soil moisture
278 content sensor using a fibre-type NIR spectrophotometer. *Soil & Tillage*, 488 Research, 80, 171-183.

279 Oliveri, P., Malegori, C., Simonetti, R., Casale, M. (2019). The impact of signal pre-processing on the final
280 interpretation of analytical outcomes- A tutorial. *Analytica Chimica Acta*, 1058, 9–17.
281 https://doi.org/10.1016/j.aca.2018.10.055

282 Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives – A
283 review. *Analytica Chimica Acta*, 1026, 8–36. https://doi.org/10.1016/j.aca.2018.04.004

284 Peltre, C., Thuriès, L., Barthès, B., Brunet, D., Morvan, T., Nicolardot, B., Parnaudeau, V., & Houot, S. (2011).
285 Near infrared reflectance spectroscopy: A tool to characterize the composition of different types of
286 exogenous organic matter and their behaviour in soil. *Soil Biology and Biochemistry*, 43(1), 197–205.
287 https://doi.org/10.1016/j.soilbio.2010.09.036

288 Piccinini, S. and Bortone, G. (1991). The fertilizer value of agricultural manure: Simple rapid methods of
289 assessment. *Journal of agricultural engineering research*, 49:197-208.

290 Reeves, J. B., & Van Kessel, J. S. (2000). Near-infrared spectroscopic determination of carbon, total
291 nitrogen, and ammonium-N in dairy manures. *Journal of Dairy Science*, 83(8), 1829–1836.
292 https://doi.org/10.3168/jds.S0022-0302(00)75053-3

293 Reeves, J.B. (2001). Near-infrared diffuse reflectance spectroscopy for the analysis of poultry manures.
294 *Journal of Agricultural and Food Chemistry*, 49 (5), 2193-2197. https://doi.org/10.1021/jf0013961

295 Saeys W., Xing J., De Baerdemaeker J. Ramon H. (2005). Comparison of transflectance and reflectance to
296 analyse hog manures. *Journal of Near Infrared Spectroscopy,* 13, 99–107.

297 Saeys, W., Darius, P., Ramon, H. (2004). Potential for on-site analysis of hog manure using a visual and
298 near-infrared diode array reflectance spectrometer. *Journal of Near Infrared Spectroscopy*, 12, 299–309.
299 https://doi.org/10.1255/jnirs.438

300 Savitzky, A. and Golay, M. (1964). Smoothing and Differentiation of Data by Simplified Least Squares
301 Procedures. *Analytical Chemistry*, 36, 1627-1639. https://doi.org/10.1021/ac60214a047

302 Scotford, I.M., Cumby, T.R., White, R.P. Carton, O., Lorenz, F., Hatterman, U. & Provolo, G. (1998).
303 Estimation of the nutrient value of agricultural slurries by measurement of physical and chemical
304 properties. *Journal of Agricultural Engineering Research*, 71: 291-305.

Van Kessel J. S., Thompson, R. B., and Reeves, J. B. (1999). Rapid on-farm analysis of manure nutrients using quick tests. *Journal Of Production Agriculture*, 12:215-224.

Williams, J.R., Chambers, B.J., Smith, K.A., Brookman, S., Chadwick, D., and Pai, B.F. (1996). The development of user-friendly systems for on-farm estimation of the available nitrogen content in solid manures and slurries. MAFF Open Contract CSA 2849. ADAS and IGER, UK.

Williams, P. C., & Sobering, D. (1996). How do we do it: a brief summary of the methods we use in developing near infrared calibrations. In: A. M. C. Davies & P. C. Williams (Eds.), *Near infrared spectroscopy: the future waves* (pp. 185-188). Chichester: NIR Publications.

Worley, B. (2021). Spectral data analysis: pre-processing and beyond. JMP statistical discovery. Accessed: Oct. 06, 2023. [Online]. Available: https://community.jmp.com/t5/Discovery-Summit-Americas-2021/Analyzing-Spectral-Data-Preprocessing-and-Beyond-2021-US-45MP/ta-p/398718

Xing, L., Chen, L. J., & Han, L. J. (2008). Rapid analysis of layer manure using near-infrared reflectance spectroscopy. *Poultry science*, 87(7), 1281-1286. https://doi.org/10.3382/ps.2007-00464

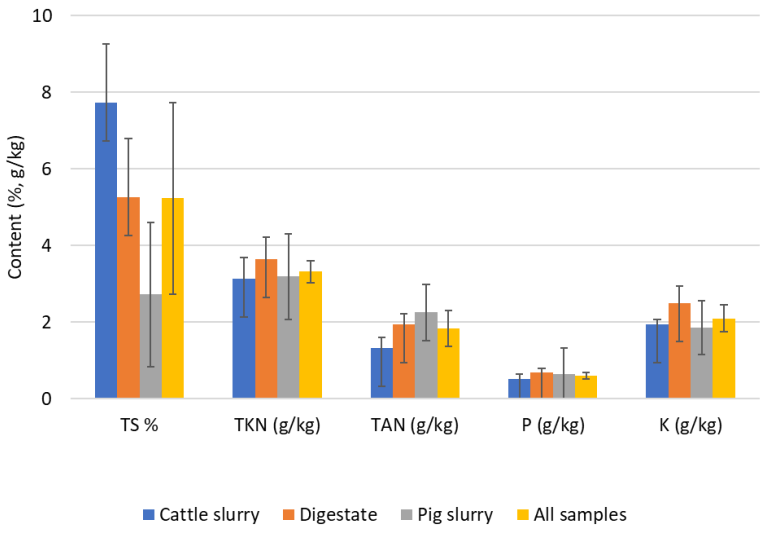
Ye, W., Lorimor, J.C., Hurburgh, C., Zhang, H. & Hattey, J. (2005). Application of near-infrared reflectance spectroscopy for determination of nutrient contents in liquid and solid manures. *Transactions of the ASAE* 48(5): 1911-1918. DOI:10.13031/2013.20000.

Figure Captions

Fig. 1. Mean and standard deviation of reference analysis for all samples and each biomass separately.

Fig. 2. Average NIR raw spectra of the cattle and pig slurry and digestate samples.

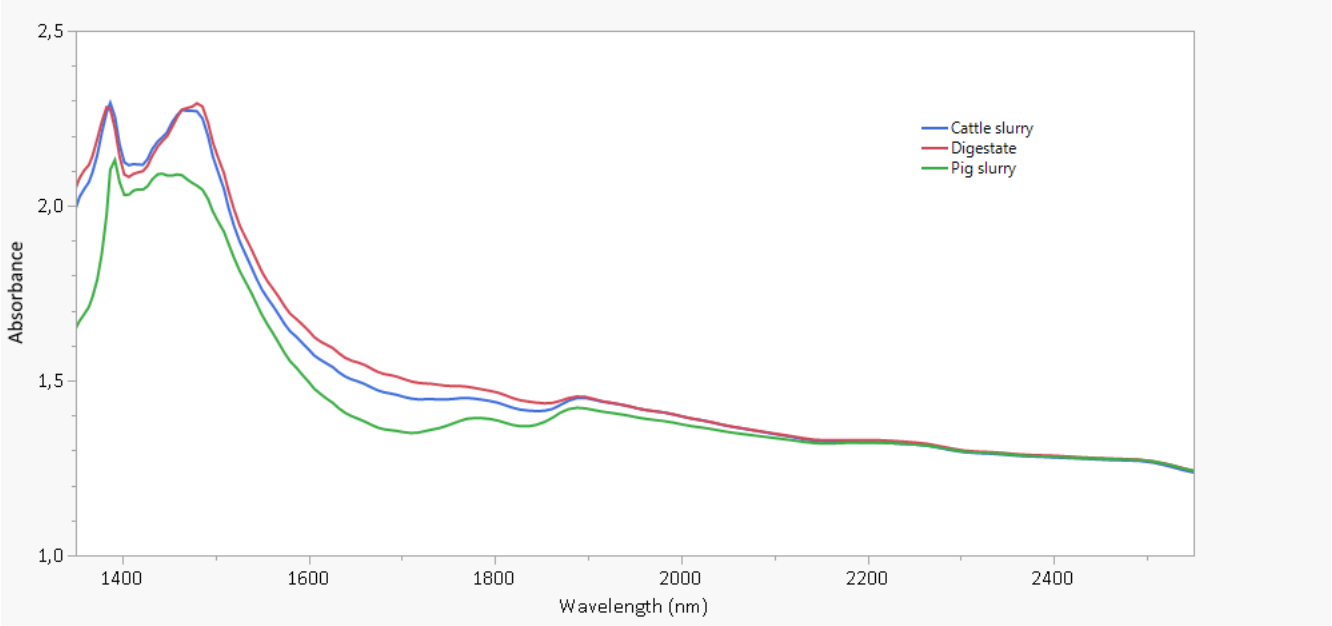
Fig. 3. Plot of reference versus predicted values from the calibration (○) and external validation (►) PLS and SVM models for TS, TKN, TAN, and P. Empty circles indicate calibration samples and full triangles indicate validation samples.



331

332 Fig. 1

333



334

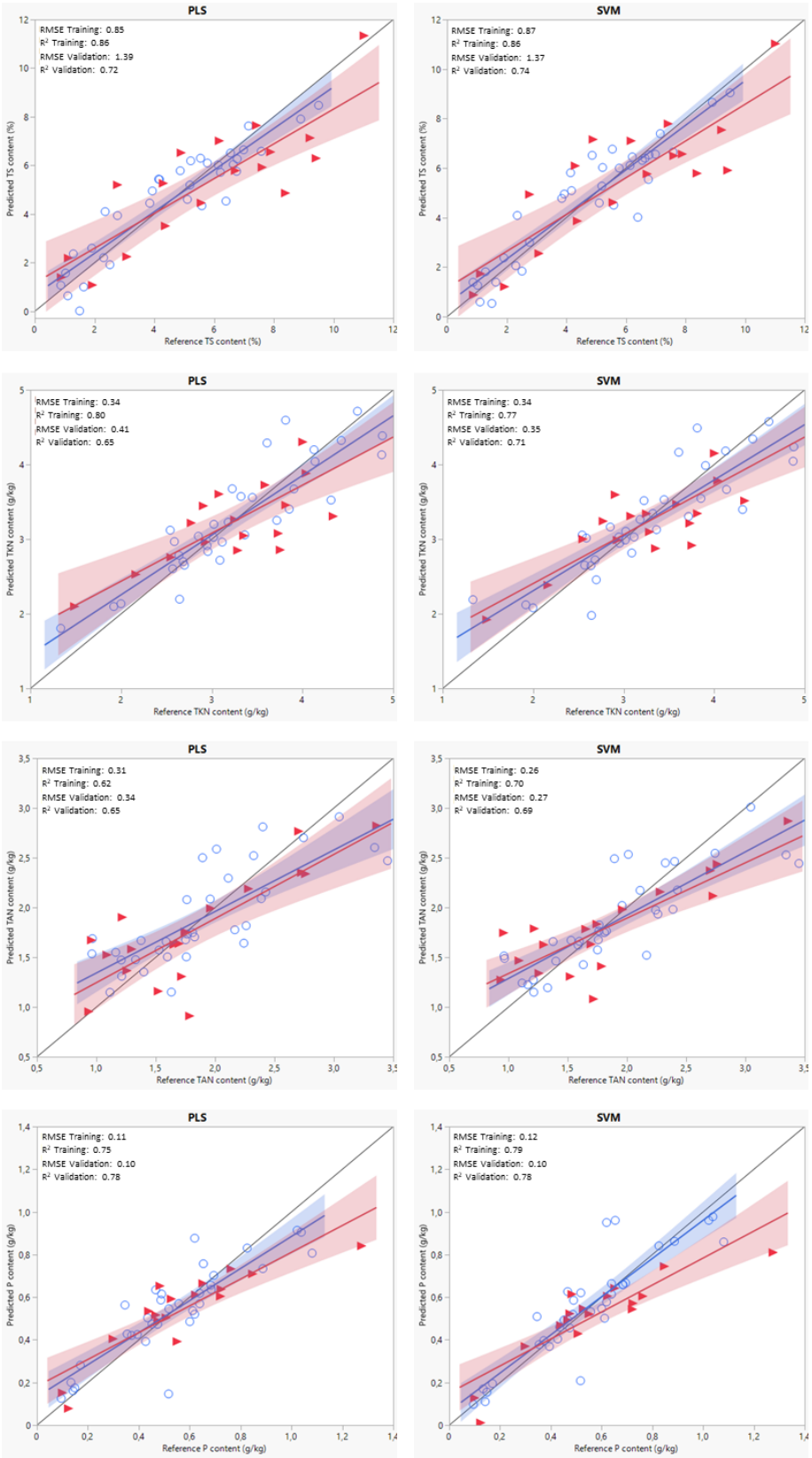
335 Fig. 2

336

337

338

339


PLS
RMSE Training: 0.85
R² Training: 0.86
RMSE Validation: 1.39
R² Validation: 0.72
Predicted TS content (%)
Reference TS content (%)
SVM
RMSE Training: 0.87
R² Training: 0.86
RMSE Validation: 1.37
R² Validation: 0.74
Predicted TS content (%)
Reference TS content (%)
PLS
RMSE Training: 0.34
R² Training: 0.80
RMSE Validation: 0.41
R² Validation: 0.65
Predicted TKN content (g/kg)
Reference TKN content (g/kg)
SVM
RMSE Training: 0.34
R² Training: 0.77
RMSE Validation: 0.35
R² Validation: 0.71
Predicted TKN content (g/kg)
Reference TKN content (g/kg)
PLS
RMSE Training: 0.31
R² Training: 0.62
RMSE Validation: 0.34
R² Validation: 0.65
Predicted TAN content (g/kg)
Reference TAN content (g/kg)
SVM
RMSE Training: 0.26
R² Training: 0.70
RMSE Validation: 0.27
R² Validation: 0.69
Predicted TAN content (g/kg)
Reference TAN content (g/kg)
PLS
RMSE Training: 0.11
R² Training: 0.75
RMSE Validation: 0.10
R² Validation: 0.78
Predicted P content (g/kg)
Reference P content (g/kg)
SVM
RMSE Training: 0.12
R² Training: 0.79
RMSE Validation: 0.10
R² Validation: 0.78
Predicted P content (g/kg)
Reference P content (g/kg)

340

341 Fig. 3

Table 1. Correlations among the chemical-physical parameters for cattle slurry, pig slurry and digestate. The values in parenthesis indicate p-values (bold text, significant at p < 0.05)

	Cattle Slurry					Pig slurry					Digestate				
	TS	TKN	TAN	P	K	TS	TKN	TAN	P	K	TS	TKN	TAN	P	K
TS															
TKN	0.90 (**0.001**)					0.83 (**0.001**)					0.27				
TAN	0.70 (**0.003**)	0.92 (**0.001**)				0.67 (**0.001**)	0.95 (**0.001**)				-0.28	0.71 (**0.002**)			
P	0.77 (**0.001**)	0.84 (**0.001**)	0.80 (**0.001**)			0.94 (**0.001**)	0.72 (**0.001**)	0.60 (**0.004**)			-0.42	0.44	0.61 (**0.01**)		
K	0.71 (**0.002**)	0.81 (**0.002**)	0.77 (**0.001**)	0.62 (**0.01**)		0.56 (**0.009**)	0.78 (**0.001**)	0.84 (**0.001**)	0.49		0.58 (**0.02**)	0.05	-0.35	-0.47	

Table 2. Summary statistics for PLS models

Chemical parameter	Pre-processing[a]	Calibration performance[b]				Model classification[d]	Validation performance[c]			Model classification[d]
		N_F	R^2_C	$RMSE_C$	RPD_C		R^2_V	$RMSE_V$	RPD_V	
TS (%)	SNV, SG1(2,10)	5	0.86	0.85	2.81	U	0.72	1.39	2.16	U
TKN (g/kg)		6	0.80	0.34	2.44	U	0.65	0.41	2.02	MU
TAN (g/kg)		4	0.62	0.31	2.03	MU	0.65	0,34	2.06	MU
P (g/kg)		4	0.75	0.11	2.27	U	0.78	0.10	2.80	U
K (g/kg)		-	-	-	-	-				

[a]Spectral data were pre-processed by mean standard normal variate (SNV) and Savitzky-Golay first (SG1) derivative algorithm (Savitzky and Golay, 1964) using 2nd order polynomial and 10 smoothing side points (software default settings). [b]N_F, number of factors; R^2_C, coefficient of determination of calibration; $RMSE_C$, root mean square error of calibration; RPD_C, ratio of prediction to deviation for the calibration. [c]R^2_V, coefficient of determination of validation; $RMSE_V$, root mean square error of validation; RPD_V, ratio of prediction to deviation for the validation. [d] U=useful; MU=moderately useful.

Table 3. Summary statistics for the SVM models

Chemical parameter	Pre-processing[a]	Calibration performance[b]				Model classification[d]	Validation performance[c]			Model classification[d]
		$RASE_C$	R^2_C	$RMSE_C$	RPD_C		R^2_V	$RMSE_V$	RPD_V	
TS (%)	SNV, SG1(2,10)	0.88	0.86	0.87	2.74	U	0.74	1.37	2.19	U
TKN (g/kg)		0.39	0.77	0.34	2.44	U	0.71	0.35	2.37	U
TAN (g/kg)		0.34	0.70	0.26	2.42	U	0.69	0.27	2.59	U
P (g/kg)		0.12	0.79	0.12	2.08	MU	0.78	0.10	2.80	U
K (g/kg)		0.43	0.59	0.36	1.92	MU	-	-	-	-

[a]Spectral data were pre-processed by mean standard normal variate (SNV) and Savitzky-Golay first (SG1) derivative algorithm (Savitzky and Golay, 1964) using 2nd order polynomial and 10 smoothing side points (software default settings). [b]$RASE_C$, root average squared error of calibration; R^2_C, coefficient of determination of calibration; $RMSE_C$, root mean square error of calibration; RPD_C, ratio of prediction to deviation for the calibration. [c]$RASE_V$, root average squared error of validation; R^2_V, coefficient of determination of validation; $RMSE_V$, root mean square error of validation; RPD_V, ratio of prediction to deviation for the validation [d]U=useful; MU=moderately useful.

Table 4. Predictive accuracy of NeoSpectra for TS, TKN, TAN and P with PLS and SVM models, based on DLG certification thresholds.

	SVM				PLS			
	TS (%)	TKN (g/kg)	TAN (g/kg)	P (g/kg)	TS (%)	TKN (g/kg)	TAN (g/kg)	P (g/kg)
All	-	*	*	-	-	*	*	-
Cattle slurry	*	*	-	*	*	*	-	*
Digestate	-	*	-	*	*	*	*	*
Pig slurry	-	*	*	-	-	*	*	-

*: passed as 60% of the tested sample concentrations differ less than 25% from the reference and no sample concentration higher than 35% relative deviation; -: not passed.