TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Pflanzenernährung

Evaluation of near infrared spectroscopy to estimate process parameters in anaerobic digestion of agricultural feedstocks

Lutz Christian Krapf

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. H. Bernhardt

Prüfer der Dissertation:

1. Univ.-Prof. Dr. U. Schmidhalter
2. Univ.-Prof. Dr. A. Gronauer
   (Universität für Bodenkultur Wien/Österreich)
3. Prof. Dr. H. Heuwinkel
   (Hochschule Weihenstephan-Triesdorf)

Die Dissertation wurde am 02.01.2013 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 11.04.2013 angenommen.

Danksagung

Die Versuche zu dieser Arbeit erfolgten an der Bayerischen Landesanstalt für Landwirtschaft in Freising in den Jahren 2009 bis 2011. Während dieser Zeit erfuhr ich Unterstützung von vielen Seiten. Insbesondere danke ich Urs Schmidhalter für das mir entgegengebrachte Vertrauen zur Anfertigung dieser Dissertation. Meinen Kollegen am Institut für Landtechnik und Tierhaltung, insbesondere den Kollegen der Arbeitsgruppe Biogas, danke ich für die erfahrene Unterstützung. Ich danke Andreas Gronauer für die Betreuung gerade zu Beginn der Versuche. Hauke, unsere Gespräche wären schwer zu ersetzen gewesen und nicht nur dafür danke ich Dir herzlich. Mein Dank gilt Dieter Nast für die gemeinsamen Diskussionen aus denen zahlreiche Ideen in diese Arbeit einflossen. Meine Wertschätzung an die Industriepartner die diese Arbeit begleiteten. Thomas Schneider und Tomas Qvarfort von der NIR-Online GmbH danke ich für die kontinuierliche Unterstützung über den gesamten Projektzeitraum. Der Högemann GmbH danke ich für die apparative Ausstattung. Andreas Niemöller und Jörg Schück von der Bruker Optik GmbH danke ich für die Unterstützung mit dem Ziel der Erweiterung der Fragestellungen dieser Arbeit. Möglich wurde diese Arbeit durch die finanzielle Förderung seitens der Fachagentur Nachwachsende Rohstoffe e.V. (FNR) sowie durch das Bayerische Staatsministerium für Ernährung, Landwirtschaft und Forsten.

Table of content

List of abbreviations

AD	Anaerobic digestion
Cal	Calibration
(E)MSC	(Extended) multiplicative scatter correction
FM	Fresh matter
MD	Mahalanobis distance
MLR	Multilinear regression
NH_4-N	Ammonium
NIR	Near infrared
OLR	Organic loading rate (measured as VS in kg m^{-3} d^{-1})
PC	Principal component
PCA	Principal component analysis
PCR	Principal component regression
PDS	Piecewise direct standardisation
PLS	Partial least square
PRESS	Predicted residual error sum of square
RMSEC	Root mean square error of calibration
RMSECV	Root mean square error of cross-validation
RMSEP	Root mean square error of prediction
RPD	Ratio of performance to deviation
SED	Standard error of differences
SEL	Standard error of the laboratory (analyses for wet chemistry)
SEP	Standard error of prediction
SNV	Standard normal variate
TIC	Total inorganic carbon
TS	Total solids
Val	Validation
VFA_{GC}	Volatile fatty acids analysed by gas chromatography
VFA_{tit}	Volatile fatty acids analysed by titration
VS	Volatile solids

Further abbreviations and symbols are introduced in the text.

Summary

Agricultural biogas plants are often operated in a suboptimal range with respect to their organic loading rate (OLR) and retention time because of insufficient characterisation of the anaerobic digestion (AD) process in real-time. Therefore, their operation cannot be considered efficient, and process imbalances may result in the overall performance loss of the biogas process. The supervision of AD is usually performed by wet chemistry analysis of physicochemical parameters that describe the state of the process. Since this method is time consuming, it does not allow for real-time assessment of the process stage. Furthermore, the analytical results only reflect a given moment in time and, hence, do not cover the dynamics of the process over an extended period. Continuous measurements of the gas composition (CH_4, CO_2, and H_2S) can only provide a rough indication of the acute changes of the AD process and do not allow for the detection of process imbalances at early stages. Optical measuring instruments that can be directly employed in biogas plants to provide continuous real-time information about the actual process conditions constitute a promising approach. In the near infrared (NIR) region, overtone and combination bands of strong fundamental vibrations in the mid-infrared region, foremost for NH-, OH-, and CH-groups, can be detected. Because of the absorption of light by different functional groups along the spectra, it is possible to determine the chemical composition of the sample by multivariate regression. Since sample preparation usually is not required for measurement in the NIR region, in principal, this method allows for an in situ application.

Important parameters for the characterisation of the AD process in the liquid phase are as follows: volatile solids (VS), ammonium (NH_4-N), the pH-buffer capacity (total inorganic carbon, TIC) and the volatile fatty acids (VFA); the foremost of these are acetic and propionic acid. For the indication of instabilities, monitoring of the AD process is important especially after feeding of the digester since these time periods are sensitive because of the formation of acids by the degradation of easily hydrolysable material. The advantage of real-time monitoring is that it has the potential to visualise these process dynamics directly after feeding events. It seems possible that monitoring by NIR spectroscopy supports the operation of the biogas plant at a high OLR while concurrently ensuring stabile conditions without a loss in the degradation performance. This is because the intensity of feeding may be adapted to the current AD conditions.

Future strategies are focussed on using plant materials other than maize, which is currently the dominating species used for biogas processes in Germany. However, an increase in the variation of feedstock may affect the biology of the process; thus, real-time monitoring of AD will be even more necessary in order to secure a stable process. Currently, there are no efficient tools available for this purpose and it is the aim of the present work to support the development of NIR-based monitoring of the AD process. Here, the development of robust calibrations is of importance for future practical applications in order to gain reliable data regarding the process conditions, even in cases of variations in the feedstock mixture applied to the digester. However, the requirements of a feedstock-robust calibration are usually higher when compared to an NIR model based on a constant feedstock composition. This is because spectral interactions between the changing feedstock matrices and the process stages and temperature variations must be included in the calibration process.

Initially, NIR models of the different parameters were developed in the laboratory to assess systematically the effect of increased feedstock variation on model performance. A data set that was based on slurry samples of digesters that had been treated with a mixture of renewable energy crops was first used for developing the models. The spectral recordings were performed offline and calibration was carried out by partial least square regression with wet chemistry methods being used as a reference. Subsequently, the data set was extended by using samples that had been removed from the digesters that were additionally fed with livestock waste, and the models were recalibrated using these samples. A comparison of the results from a cross-validation study revealed that the extension of the data sets by using samples that contained manure did not affect the accuracy of the estimates to a significant degree and similar mean estimation errors (*RMSECV*) were reported for the parameters. With an error of 4.0 g kg^{-1} fresh matter (FM) and 0.16 g kg^{-1} FM, respectively, VS and NH_4-N exhibited the highest potential for estimation via NIR spectroscopy. The TIC was also satisfactorily estimated with a *RMSECV* of near 0.8 g kg^{-1} FM. With an estimation error of 0.9 g kg^{-1} FM, the total VFA model supported the use of NIR-based rapid screening of the dynamics of the acidity level in the digester. However, a comparison of the accuracy of the NIR estimates for TIC and total VFA with the repeatability (*SEL*) of the reference method revealed comparably poor accuracy, which restricted this lab application to screening purposes only.

In the next step, the models were validated using digester slurry that was composed of 26 commercial biogas plants whose material had not been utilised earlier for calibration. The relationships between the spectral readings and the chemical parameters as found for cross-validating the models were confirmed for the most part. With a mean error (*RMSEP*) of 5.0 g kg^{-1} FM for VS and 0.2 g kg^{-1} FM for NH_4-N, the error estimates were only slightly higher. The precision (*SEP* = 1.0 g kg^{-1} FM) of the TIC-model proved to be robust when tested on an external slurry material. Theoretically, there were not any absorption bands for mineral components and their ionic forms, such as the carbonate ion, within the NIR spectrum. Therefore, the satisfactory estimations of this parameter most likely could be explained, since for the tested samples, the alkalinity was found to be a function of the concentration of NH_4-N and total VFA (*RMSEP* = 0.9 g kg^{-1} FM). A comparison of the repeatability of the reference analysis with the repeatability of the NIR method demonstrated similar results, each of which were obtained for the VS and NH_4-N. This suggests that the two parameters could be specifically estimated by NIR spectroscopy even with varying sample matrices. This high precision of the NIR measurements ensured that the overall error of the estimates could be kept low. For the TIC and total VFA, comparably poor repeatability of the NIR method for a single sample was observed, which also explains the limited accuracy of the lab approach for these parameters. Overall, the results that were obtained by testing the models on an open population suggest that a calibration data set that is limited in sample number and feedstock variability already allows for its robust application toward screening across different digesters and varying feedstock mixtures. The calibration data sets with <200 samples proved to be sufficient for reflecting the spectral variability of slurry that was composed of a wider range of feedstock mixtures. High co-linearity was found between the individual acids and the total VFA for the calibration data set as well as among the validation samples. Hence, demonstration of a specific estimation of the individual acids was limited. However, through the additional validation of the models using samples that were artificially spiked with acids, the specific estimation of propionic acid in digester slurry was demonstrated.

A third step was the in situ use of the NIR technique directly on a technical scale digester. Therefore, the spectra that were recorded with the lab instrument were transferred to an online process analyser by applying piecewise direct standardisation. The models were applied to a spectral time series from a digester that was operated using slurry of a different origin to permit independent validation after recalibration under these new environmental conditions. The results of this spectra transfer confirmed the suitability of feedstock-robust modelling also for application directly to the digester and highlighted the possibility for cross-linking the different instruments. Therefore, the advantage of this online application over the lab approach was demonstrated. This advantage was mainly due to the possibility of visualising the long-term trend as well as the short-term dynamics of the process over an extended period, and this was proven for different temperature levels within the digester. However, a comparison of the estimates with the reference values showed that the accuracy of the NIR method was affected by a moderate sample matrix specific error, which has to be considered in future applications.

A second approach that was aimed at studying the ability of the NIR technology to be used for identifying short-term AD process dynamics after feeding was pursued. Spectra were recorded online using a down-scaled digester that was fed with maize silage. A bypass was used to enable continuous spectral recording. Feeding was carried out over a 240-day period at a temporarily high OLR to induce perturbations into the process. By doing this, it was possible to investigate whether or not the NIR method was able to visualise the induced dynamics in real-time. The time series estimates for VS (*RMSEP* = 3.0 g kg^{-1} FM) and total VFA (*RMSEP* = 0.9 g kg^{-1} FM) demonstrated the suitability of this method for discerning changes in concentration. Even small changes in the concentration of both parameters could be visualised due to the high repeatability of the NIR analyses, and this was carried out at high temporal resolution. The total VFA model indicated concentration changes at an early stage of accumulation near 2–3 g kg^{-1} FM. Since the high sensitivity for the total VFA estimates proved to be independent of the VS content, the potential of the NIR technique as a tool for optimising the operation of a digester was shown.

The data set was extended by incorporating samples of a more diverse feedstock mixture and the models were recalibrated with the aim of increasing their robustness. These models were validated by using a second time series, and the accuracies for VS (*RMSEP* = 3.8 g kg^{-1} FM) and total VFA (*RMSEP* = 1.1 g kg^{-1} FM) were mostly kept constant. However, compared to the first spectral time series where pure maize silage was used as feed, the estimates were influenced to a greater degree by noisy structures owing to an increase in the noise in the updated model. This resulted in a moderate loss in the overall validity of the estimates. The results of this second time series demonstrated that the accumulation of VFA due to a low feedstock quality could also be observed via NIR spectroscopy at an early stage. The same was true for detecting an increase in the NH_4-N concentration in the digester after feeding changes towards a feedstock mixture that is rich in protein. During short-term monitoring, it was also possible to determine the propionic acid concentration, which showed a different concentration course than that of the acetic acid and total VFA parameters. This suggests that the NIR analysis also permits the differentiated assessment of AD process stability with respect to the VFA composition.

Collectively, the results of this work confirmed the robustness of the NIR technique for characterising real-time the dynamics of the AD process even for varying feedstock composition and temperature changes within the digester. Thus, the NIR technique is suited to contribute to improvement of future operation of agricultural biogas plants.

Zusammenfassung

Landwirtschaftliche Biogasanlagen (BGA) werden häufig aufgrund unzureichender Prozesscharakterisierung während der Biogasproduktion nicht im optimalen Bereich bezüglich Raumbelastung und Verweilzeit betrieben. Dies hat zur Folge, dass die Betreibung der Anlage ineffizient geregelt ist und im ungünstigsten Fall deren Prozessstabilität nicht mehr gewährleistet werden kann. Laboruntersuchungen des Gärmaterials gestalten sich aufwendig und geben zudem nur eine verzögerte Momentaufnahme des Anlagenprozesses wieder. Kontinuierliche Messungen der Gaszusammensetzung (CH_4, CO_2, H_2S) dienen lediglich der groben „Sichtkontrolle“. Sie zeigen auf, ob sich akute Veränderungen im Prozess anbahnen. Eine interessante Alternative zur Charakterisierung des aktuellen Prozesszustandes stellen optische Messinstrumente dar, welche direkt an die BGA angeschlossen werden können. Im Nahinfrarot- (NIR) Bereich können die Ober- sowie Kombinationsschwingungen starker Grundschwingungen des mittleren Infrarotbereiches detektiert werden, wobei es sich zumeist um Schwingungen der NH-, OH- und CH-Gruppen handelt. Aufgrund der Absorption von Strahlung an einzelnen funktionellen Gruppen entlang des Spektrums kann mittels multivariater Regressionsverfahren auf die chemische Zusammensetzung des Materiales geschlossen werden. Da bei einer Messung im NIR-Bereich auf eine Probenvorbereitung verzichtet werden kann, ist prinzipiell eine In-Situ Anwendung möglich.

Zur Kontrolle des anaeroben Abbaus in der flüssigen Phase werden regelmäßig folgende Prozessparameter verwendet: die organische Trockensubstanz (oTS), Ammonium (NH_4-N), die Pufferkapazität (TAC) sowie die flüchtigen Fettsäuren (FFS), insbesondere die Propion- und Essigsäure. Gerade Fütterungsereignisse sind potentiell kritische Zeitpunkte, da leicht hydrolysierbares Material einen raschen Säureanstieg im Fermenter induzieren kann. Der Vorteil einer Echtzeitmessung liegt darin begründet, die Prozessdynamik direkt nach einer Substratzugabe abzubilden. Die Echtzeit-Kontrolle der Säureparameter mittels NIR kann dabei prinzipiell eine situationsangepasste Fütterung ermöglichen. Dies lässt es denkbar erscheinen, den Prozess in eine maximale Raumbelastung ohne Einbußen an Abbaueffizienz zu fahren.

Unter den Energiepflanzen stellt Mais derzeit das für die landwirtschaftliche Biogasproduktion in Deutschland am häufigsten verwendete Ausgangssubstrat dar. Erklärtes Ziel zukünftiger Anwendungen ist eine vermehrte Einbindung alternativer Einsatzstoffe. Eine erhöhte Variation der Substrate macht jedoch eine zeitnahe Überwachung des Gärprozesses notwendig, um einen stabilen Anlagenbetrieb zu sichern. Hierfür stehen bisher keine leistungsfähigen Werkzeuge zur Verfügung. Das grundlegende Ziel dieser Arbeit besteht daher in der Entwicklung einer solchen Überwachung auf Basis der NIR-Technologie. Die Erstellung robuster Modelle ist dabei für einen zukünftigen Praxiseinsatz eine zentrale Voraussetzung, um bei veränderten Substratmischungen zuverlässige Daten zum Prozesszustand zu gewinnen. Die Anforderungen an eine solche Kalibration sind dabei in der Regel höher als bei einer Entwicklung spezifischer Modelle auf Basis konstanter Substratzusammensetzung. So müssen spektrale Interaktionen zwischen Probenmatrix, Prozesszustand und Temperatur, wie sie in zu analysierenden Gärproben auftreten, bereits im Kalibrierdatensatz abgebildet sein.

Im ersten Arbeitsschritt wurde daher der Einfluss einer Erhöhung der Substratvariation des Kalibrierdatensatzes auf die Schätzgüte der NIR-Modelle überprüft. Dabei diente ein Datensatz mit Gärproben reiner Energiepflanzenmischungen als Ausgangspunkt für die Modellbildung. Die Spektren wurden offline im Labor erhoben, die Modellbildung erfolgte mittels Partial Least Square Regression auf Basis nasschemischer Referenzmethoden. Dieser Datensatz wurde um Proben aus Fermentern erweitert, die zusätzlich mit Wirtschaftsdüngern betriebenen wurden bei anschließender Neukalibration der Modelle. Der Vergleich der Ergebnisse einer Kreuzvalidierung zeigte auf, dass diese Erweiterung um Wirtschaftsdüngerproben nicht zu einer relevanten Abnahme der Schätzgenauigkeit der Modelle führte. So konnten für die einzelnen Parameter jeweils ähnliche mittlere Fehler (*RMSECV*) aufgezeigt werden. Ein Fehler für oTS von 4.0 g kg^{-1} Frischmasse (FM) und für NH_4-N von 0.16 g kg^{-1} FM bestätigte die Möglichkeit für eine erfolgreiche Vorhersage dieser Größen. Ebenfalls konnte der TAC mit einem Wert um 0.8 g kg^{-1} FM befriedigend geschätzt werden. Ein *RMSECV* von 0.9 g kg^{-1} FM für die FFS gesamt unterstreicht die Eignung der NIR-Methodik für ein rasches Aufzeigen der aktuellen Prozessdynamik. Ein Vergleich der Schätzgenauigkeit für TAC und FFS gesamt mit der Wiederholpräzision (*SEL*) der Referenzanalytik zeigte jedoch eine geringe Genauigkeit der NIR-Laboranwendung und somit deren eingeschränkte Eignung über einen Schnelltest hinaus.

Die Modelle wurden im zweiten Schritt an Gärmaterial von 26 praxistypischen BGA validiert, deren Material nicht zuvor für die Modellbildung verwendet wurde. Im Wesentlichen bestätigten sich dabei die für die Kreuzvalidierung aufgezeigten Zusammenhänge zwischen spektraler Signatur und chemischen Größen. So konnte für oTS ein mittlerer Schätzfehler (*RMSEP*) von 5.0 g kg^{-1} FM erzielt werden, für NH_4-N fiel der Fehler mit 0.2 g kg^{-1} FM ebenfalls nur geringfügig höher aus. Mit einer Präzision (*SEP*) von 1.0 g kg^{-1} FM erwies sich das TAC-Modell ebenfalls als robust für den Einsatz an externem Material. Als mineralische Komponente besitzt das Carbonation im NIR-Bereich keine spezifische Signatur. Die befriedigende Schätzung des TAC-Wertes erklärt sich vermutlich dadurch, dass sich dieser primär als eine Funktion der Konzentrationen von NH_4-N und FFS gesamt (*RMSEP* = 0.9 g kg^{-1} FM) ergab. Ein Vergleich der Wiederholpräzision der Referenzanalytik mit jener der NIR-Methodik zeigte ähnliche Ergebnisse für die Parameter oTS und NH_4-N, was die spezifische Messung beider Parameter mittels NIR-Spektroskopie auch bei variierender Probenmatrix unterstreicht. Diese hohe Präzision der NIR-Analyse sicherte einen insgesamt niedrigen Schätzfehler. Für die Parameter TAC und FFS gesamt zeigte sich eine vergleichsweise geringe Wiederholbarkeit der NIR-Messung an einer Probe, was letztlich auch die deutlich eingeschränkte Genauigkeit der Laboranwendung erklärt. Insgesamt legen diese Validierergebnisse an einer offenen Population den Schluss nahe, dass ein in seiner Probenanzahl und Substratzusammensetzung reduzierter Kalibrierdatensatz bereits eine robuste Anwendung für ein anlagen- und substratübergreifendes Screening ermöglicht. Probensätze mit <200 Proben erwiesen sich dabei als ausreichend, um die spektrale Variation von Gärmaterial auf Basis unterschiedlicher Futtermischungen abzubilden. Aufgrund der in den Datensätzen vorliegenden hohen Korrelation zwischen Einzelsäuren und FFS gesamt, war die spezifische Überprüfung der Einzelsäuremodelle lediglich eingeschränkt möglich. Jedoch konnte durch künstliche Zugabe von Einzelsäuren in die Probe die prinzipielle Eignung der NIR-Spektroskopie für eine spezifische Messung der Propionsäure in Gärmaterial aufgezeigt werden.

Der dritte Schritt war ein In-Situ Einsatz an einem Technikumsfermenter. Die mit dem Laborgerät erhobenen Spektren wurden dazu mittels der Piecewise Direct Standardisation auf eine Online-Anwendung übertragen. Nach Neukalibration der Modelle mit den auf die neue Arbeitsumgebung angepassten Spektren, erfolgte deren Validierung an einer spektralen Zeitreihe unter Einsatz von Gärmaterial unterschiedlicher Herkunft. Die Ergebnisse dieses Spektrentransfers bestätigten die Eignung robuster NIR-Modelle für einen direkten Einsatz am Fermenter und demonstrierten die Möglichkeit einer Vernetzung von Geräten untereinander. Dabei konnte der enorme Vorteil einer Online-Applikation gegenüber einem Laboreinsatz aufgezeigt werden. Dieser liegt darin begründet, sowohl den mehrwöchigen Verlaufstrend als auch die kurzzeitige Dynamik der Prozessparameter in einer geschlossenen Messabfolge zu visualisieren. Gleichzeitig zeigte der Vergleich zwischen Schätzwerten und Referenzanalytik, dass die Genauigkeit der NIR-Methode einem moderaten, probenmatrixspezifischen Fehler unterliegt der auch bei zukünftigen Anwendungen zu berücksichtigen ist.

In einem weiteren Arbeitspaket wurde gezielt geprüft, ob sich kurzfristige Prozessänderungen nach Fütterungen mittels NIR-Technik aufzeigen lassen. Die Kalibration erfolgte alternativ zum ersten Ansatz direkt an einem mit Maissilage betriebenen Fermenter. Die kontinuierliche Spektrenmessung erfolgte dabei über einen Pump-Bypass. Über einen Zeitraum von 240 Tagen wurde der Prozess durch stoßweise Fütterung bewusst in kritische Zustände geführt. Es wurde untersucht, ob sich die dadurch einsetzende Prozessdynamik mittels NIR-Spektroskopie in Echtzeit darstellen lässt. Die an einer ersten Zeitreihe erhobenen Schätzungen für oTS (*RMSEP* = 3.0 g kg^{-1} FM) und FFS gesamt (*RMSEP* = 0.9 g kg^{-1} FM) bestätigten die Eignung dieser Technik zum Aufzeigen von Konzentrationsänderungen beider Parameter. Diese konnten zeitlich hochaufgelöst und mit einer hohen Wiederholbarkeit beschrieben werden. Dabei reagierte das Säuremodell bereits auf Änderungen im Bereich 2-3 g kg^{-1} FM. Diese sensible und vom Wassergehalt unabhängige Schätzung unterstreicht das Potenzial dieser Messtechnik als leistungsfähiges Kontrollwerkzeug zur optimierten Führung eines Fermenters.

Der Kalibrierdatensatz wurde danach um Proben mit variierenden Futtermischungen erweitert und es erfolgte eine Neukalibration mit dem Ziel einer Erhöhung der Modellrobustheit. Die Modelle wurden an einer zweiten Zeitreihe validiert wobei die Genauigkeit für oTS (*RMSEP* = 3.8 g kg^{-1} FM) und FFS gesamt (*RMSEP* = 1.1 g kg^{-1} FM) im Wesentlichen stabil blieb. Jedoch zeigte sich ein erhöhter Störanteil im Schätzverlauf verglichen mit der ersten Zeitreihe welcher vermutlich eine Erhöhung des spektralen Rauschanteiles im erweiterten Modell wiederspiegelt. Dies führte insgesamt zu einem moderaten Verlust an Aussagekraft der Modelle. Die Ergebnisse dieser Zeitreihe unterstreichen, dass eine Akkumulation der FFS auch bei schwankender Substratqualität mittels NIR-Technik frühzeitig sichtbar gemacht werden kann; ebenso ein Anstieg der NH_4-N Konzentration nach Einsatz N-reicher Mischungen. Bei der Betrachtung der Kurzzeitdynamik der Säureparameter konnte ein von der Essigsäure bzw. der FFS gesamt abweichender Verlaufstrend der Propionsäure korrekt geschätzt werden. Dies zeigt die Möglichkeit einer differenzierten Bewertung der Prozessstabilität mittels NIR-Analytik auf.

In ihrer Gesamtheit unterstützen die Ergebnisse die Eignung der NIR-Spektroskopie zur Beschreibung der Prozessdynamik des anaeroben Abbaus auch bei veränderten Substrat-mischungen und variierenden Temperaturniveaus. Dies unterstreicht, dass die NIR-Technik zukünftig einen wichtigen Beitrag für eine optimierte Betriebsweise einer BGA leisten kann.

1. Introduction

1.1. Problem

Anaerobic digestion (AD) is the natural breakdown of organic material to biogas, which consists primarily of methane and carbon dioxide (Weiland, 2010). The conversion of biomass to biogas is a complex biochemical process that occurs through a series of reactions in which several groups of microorganisms are involved (Gerardi, 2003). Under balanced operation, the production rate of the intermediates generated during this multistep process equals the consumption rate and there is little accumulation of these compounds (Pullammanappallil et al., 2001). However, the acid- and methane-forming microbial groups involved in this process differ in terms of physiology, nutritional needs, growth kinetics, and sensitivity to environmental conditions (Chen et al., 2008; Pohland and Ghosh, 1971). Thus, AD is sensitive to disturbances, i.e., imbalances in the activity of the different microorganisms (Reed and Davies, 2011). Factors identified as most disturbing include organic overload, the presence of inhibitory compounds, and fluctuations in process conditions including changes in the temperature and feedstock composition (Switzenbaum et al., 1990; Ahring et al., 1995).

The AD process can be used in agricultural biogas plants for the combined generation of heat and power, thereby constituting a source for locally generated renewable energy (Lehtomäki et al., 2007). Feedstock from the agricultural sector includes livestock waste, crop residues, and energy crops, i.e., plants grown specifically for use as energy sources (Weiland, 2010). Feeding fluctuations can induce process instabilities in AD, particularly when using easily hydrolysable feedstock that is rich in starch, which results in the rapid formation of volatile fatty acids (VFA) (Myint et al., 2007; Ward et al., 2008). The accumulation of these intermediates reflects a kinetic uncoupling between acid producers and consumers, and leads to a decrease in the buffering capacity as the bicarbonate is released as carbon dioxide (Jantsch and Mattiasson, 2004). In the case of progressive VFA accumulation, a pH drop will be the consequence, which may significantly slow the rate of methane formation, inducing a risk of AD process arrest (Boe et al., 2010). As a consequence, commercial biogas plants are often operated at a less-than-optimum loading rate in order to prevent critical process conditions (Ward et al., 2008). However, from an economical point of view, full-load operation is demanded as long as the requirements of high conversion efficiencies are met (Boe, 2006).

To achieve optimum loading rates and to prevent AD instabilities, process parameters including VFA and buffer capacity must be monitored on a real-time basis. Although wet chemistry analysis can regularly provide accurate information about the physicochemical conditions in the digester, the time delay between sampling and acquiring the analytical results is a limiting factor as it does not provide real-time assessment of the biogas process (Holm-Nielsen et al., 2006). Considering that acid accumulation can occur within days, if not hours, the AD process may fail before the operator can take action, resulting in extended recovery times, loss of production, and restart expenses (Boe, 2006). Thus, it follows that wet chemistry cannot be employed as an efficient tool for the monitoring of AD. Therefore, alternative technologies for real-time monitoring and control are necessary in order to improve the process efficiency of biogas plants (Weiland, 2010).

1.2. Near infrared spectroscopy for the monitoring of anaerobic digestion

Of the different technologies tested for improving AD monitoring (see references in Ward et al., 2008), near infrared (NIR) spectroscopy constitutes a promising method which has already shown its potential for process control in the food sector, animal nutrition, and the chemical industry (Roberts et al., 2004; Burns and Ciurczak, 2008). NIR spectroscopy is based on the interaction of electromagnetic radiation with material in the range of 780 to 2500 nm (12820 to 4000 cm^{-1}) (Stuart, 2004). This region contains the absorption bands that correspond to overtones and combinations of the fundamental vibrations of the mid-infrared region (MIR, 4000 to 400 cm^{-1}) (Bokobza, 2002). Organic matter selectively absorbs NIR radiation to yield information about the molecular bonds, foremost of which are the CH, OH and NH bonds. Because these chemical functional groups each have a specific vibrational frequency, in principal, their absorption information can be described by three parameters: the location of the absorption information along the spectrum, the relative intensity of the absorption peak, and the bandwidth of the peak (Workman and Shenk, 2004). The intensities of the absorption usually decrease by a factor of 10 to 100 for each step from the fundamental to the next overtone as a consequence of the lower transition probabilities. Therefore, measurements can be performed with a comparatively large sample thickness when working in the NIR region. This constitutes an important aspect for the implementation of the NIR technique in process monitoring and control since it allows for in situ application (Siesler, 2008).

A common assumption regarding spectroscopic measurements in the MIR and NIR regions is that the amount of light absorbed by the sample at different wavelengths is directly related to the concentration of chemical functional groups, i.e., the contribution of an absorber to the absorbance is proportional to the concentration of the absorber (Griffiths and Dahm, 2008). As the concentrations of the functional groups are in turn related to concentrations of different physicochemical parameters (e.g., water content, VFA in digester slurry), property values can be determined. This linear Lambert-Beer's law relationship can be given as (Tillmann, 1996):

$$A_\lambda = \log\left(\frac{1}{R_\lambda}\right) = c * \epsilon_\lambda * l$$

where A_λ = absorbance, R_λ = reflection, c = molar concentration of the absorber,
ϵ_λ = molar absorptivity of the absorbent at wavelength λ, and
l = sample path length

Deviations from this ideal case may occur, i.e., because of changes in sample and instrument temperature, and interferences caused by overlapping of absorption bands (Workman and Shenk, 2004). Furthermore, Lambert-Beer's law is derived for transmission spectroscopy of clear liquids (Kessler, 2007), whereas no definite reflectance theory exists because an 'infinite number of integrals would be required to describe all the combined light interaction effects at all surfaces under varying conditions' (Workman Jr, 2008).

In diffuse reflectance, which is the measurement mode used in this work, incident light penetrates the surface of the sample and is scattered in all spatial directions. The light that is diffusely reflected from the surface, or from the portion near the surface, can be detected (Kawano, 2002). In addition to the absorption characteristics due to chemical bonds, the spectrum also contains information about the physical-optical composition of a material as a consequence of radiation scatter, and surface and diffuse reflection of the sample (Workman and Shenk, 2004). Therefore, the spectrum obtained from diffuse reflectance measurements is a composite of all chemical and physical-optical information of the outer surface layer of the sample (Næs et al., 2002). In order to reveal specific information from NIR signals, e.g., for quantitative analysis, the spectrum usually cannot be used directly and chemometric methods (Danzer et al., 2001) are necessary in order to describe the mathematical relationships between spectral data and certain characteristics of the sample.

The invention of modern NIR spectroscopy is strongly connected to the use of multivariate data analysis methods for solving basic problems, foremost of which are non-selectivity and multicollinearity (Næs et al., 2002). The first problem describes the low structural selectivity of NIR spectra, e.g., compared to MIR spectra, owing to the superposition of many overtone and combination bands in the NIR region (Siesler, 2008). Because of these interferences, no single wavelength usually provides sufficient information, i.e., for derivation of a calibration equation for the estimation of a chemical constituent. By using multivariate calibration, i.e., using several wavelengths in the calibration equation, signal selectivity is not required and calibration is possible, even in the presence of multiple spectral interferences (Boysworth and Booksh, 2008). For multilinear regression (MLR), the optical data is used directly for development of a calibration to establish a transfer function between the spectrometer response (X) and the reference data (Y) that can be given as (Tillmann, 1996):

$$c = b_0 + b_1 * A_{\lambda 1} + b_2 * A_{\lambda 2} + \cdots + b_i * A_{\lambda \mathrm{i}} + \cdots + b_n * A_{\lambda \mathrm{n}}$$

where c = concentration of the ingredient,
$A_{\lambda i}$ = absorbance at wavelength λ_i,
b_i = regression coefficient for wavelength λ_i, and
i = 1…n wavelength

However, when working in the NIR region, the second problem of multicollinearity is usually observed, which relates to situations where the spectra are highly intercorrelated (Kessler, 2007). If ordinary MLR are applied to a highly intercorrelated data set, unstable and unreliable regression coefficients are obtained because MLR requires each X variable to have unique information about Y (Næs et al., 2002). In principal, this problem can be solved by search strategies such as stepwise multiple linear regression (SMLR) where the collinearity is handled by selection of a subset of individual, distinct X variables from all the available X variables (Bjørsvik and Martens, 2008). However, there is often lack of a sound basis for such variable selection in the case of NIR spectra where a large number of data points are usually recorded (McClure, 2008; Workman Jr, 2008).

The principal component regression (PCR) can be used as an alternative calibration technique designed to handle intercorrelated regressors such as variables of NIR spectra (Martens and Næs, 1991). In the first step, a principal component analysis (PCA) is performed with the X variables to reduce the amount of data in the case of high collinearity (Miller and Miller, 2005). According to Næs et al. (2002), a PCA that is applied to a data matrix of samples by variables constructs new variables that are defined as principal components (PCs). These variables are linear combinations of the original variables with the feature of being uncorrelated (orthogonal) to each other. The first PC (the first factor) includes as much variability as possible of all the original variables, and it is constructed to have maximum variance amongst all such linear combinations. Each successive new variable accounts for as much of the remaining variability as possible (Næs et al., 2002). In a next step, the PC score values that describe the location of the samples along each PC calculated from the PCA are used as predictors in an MLR instead of using the original optical data directly (Esbensen, 2009).

The PCs can also be used for qualitative data analysis, e.g., they can be plotted for visual interpretation of the inter-relationships between the different variables. Because the scores show the location of the samples along each PC, they can be used for the detection and interpretation of sample patterns such as groupings, similarities, and differences (Esbensen, 2009). Thus, the score values can also be used for identification of outliers, i.e., samples whose spectral characteristics are different from the rest (Kessler, 2007). The loading values indicate the relationship between the scores and the original X variables, and can help determine which of the variables are relevant for each of the PCs in the data set (Næs et al., 2002).

Similar to PCR, partial least squares (PLS) regression uses linear combinations of the predictor variables for quantitative analysis rather than the optical data directly (Workman Jr, 2008). However, differences between the methods exist in terms of how the factors are extracted. PCR extracts factors in a way that maximizes the variability of the sample score, while being independent of previously extracted factors (Westerhaus et al., 2004). In PLS regression, the Y data are actively used for extraction of the factors, which implies that the original X variables that indicate a high correlation with the response variable (Y data) are given extra weight because they are of particular importance for estimation (Miller and Miller, 2005).

In this work, the method of PLS regression was used for the development of the NIR calibrations and for solving the problems of non-selectivity and multicollinearity. Here, the reason that PLS is preferred over PCR is because it was assumed that most of the AD process parameters investigated in this work contribute little to the overall spectral variability since their concentrations in the slurry are very low, except for volatile solids (VS). Therefore, most of the spectral variation can be considered irrelevant for modelling since it is related to other physicochemical phenomena. Hence, PLS regression may offer a better chance to extract the spectral structures that are particularly relevant for the target values used in this work (Bjørsvik and Martens, 2008). Because the PLS algorithm reduces the impact of irrelevant X variations in the calibration modelling, models simpler than PCR may also be obtained. However, in case of noisy Y data, the PLS regression tends to overfit owing to inclusion of instrumental artefacts (Martens and Næs, 1991). Therefore, for realistic assessment of the model performance, including determination of the optimal number of factors, the independent validation of a PLS model constitutes a critical aspect (Esbensen and Geladi, 2010).

Since PLS regression is designed to handle highly intercorrelated regressors, the full spectra can be used for model development (Bjørsvik and Martens, 2008). However, improvements in the robustness of a PLS model, i.e., against external variations, can be obtained by excluding superfluous or noisy X variables from calibration (Swierenga et al., 2000). In this work, two different methods were used for variable selection, with the aim of excluding data irrelevant for calibration. An optimization routine was used for automatically testing different combinations of spectral regions for calibration development. In addition, selection of individual spectral data points was performed based on uncertainty testing conducted by applying jack-knifing (Martens and Martens, 2000), and both methods will be described in more detail in the methods sections (chpt. 4.1.1.3 and 4.3.1.4).

For calibration purposes, spectra should contain as little irrelevant information as possible in order to improve estimation performance and interpretability of the spectra (Næs et al., 2002). Therefore, data pre-treatment methods are frequently used prior to calibration in spectroscopic applications. The aim is to remove light scattering effects caused by non-homogeneous distribution of particles and differences in refractive indices influencing the light's interaction with the analyte (Duckworth, 2004). Derivatives of spectral data are frequently used to remove baseline offsets and to enhance visual resolution (Heise and Winzen, 2002). Because spectra are not a continuous function, but a series of equally spaced data points, computation of a derivative can be obtained, e.g., by fitting a polynomial to the data points, which is the method applied in this study (Næs et al, 2002). For this, a low-order polynomial is fitted to the data points in a localized segment of the spectra using least squares (Savitzky and Golay, 1964). The size of the spectral segment has a direct influence on the degree of smoothing, which also has the risk of smoothing out chemical information from the spectra. Although some general rule for the selection of the segment size has been stated (Duckworth, 2004), the selection of an optimal segment size can be best determined on an individual basis (Næs et al., 2002).

Further spectra pre-treatment used in this work includes multiplicative scatter correction (MSC) (Heise and Winzen, 2002). Scattering theory assumes that scattering has a multiplicative (amplification) effect and an additive (offset) effect on reflection spectra, which implies that the spectra contain a broad, changing background from differential scattering at each wavelength (Boysworth and Booksh, 2008). The aim of MSC is to remove these effects from the data by shifting and scaling each spectrum to a target spectrum, which is usually the mean spectrum of the calibration set (Heise and Winzen, 2002). Therefore, correction coefficients are computed from a regression of each individual spectrum onto the target spectrum. For correction of future samples, the mean of the calibration set can then be employed as the scatter standard (Boysworth and Booksh, 2008). As an alternative to MSC, the standard normal variate (SNV) transformation method works in a way that each spectrum is centred around zero by subtraction of the mean and then each spectral data point is divided by the standard deviation of the complete spectrum (Heise and Winzen, 2002). Although these methods differ in how the scattering effects are removed from the spectra, the results obtained from MSC and SNV are nearly equivalent in most cases (Kessler, 2007).

In spectroscopic applications, several situations can arise in which an NIR model can become invalid. One scenario involves the occurrence of changes in the physical or chemical constitution of the sample to be analysed. In the case of AD monitoring, variations in the pH and

the temperature of the slurry can change the hydrogen bonding within the sample, affecting band position and width (Workman, Jr, 2008). Furthermore, the main challenge of a feedstock-robust calibration is the complex nature of the NIR spectra, which are characterised by interconstituent interactions and the strong influence of the sample's scattering properties on the spectral signature (Osborne and Fearn, 1986). One strategy to avoid erroneous estimations is to incorporate these expected variations into the calibration (Swierenga et al., 2000), a strategy that was pursued in the present work. This correction for new spectral variation can be achieved by inclusion of new factors in the PLS model. As a consequence, fewer sources of variation need to be controlled in future applications. However, in the case of strong non-linear relationships, this strategy is limited and the inclusion of this spectral variation may result in considerable performance loss of the estimates (Swierenga et al., 2000).

Another scenario refers to a practical limitation for the use of an NIR model when the samples to be estimated are measured on an instrument that was not used to develop the calibration (Feudale et al., 2002). One example is the situation where the calibration is performed using a lab instrument and then transferred to an online process analyser (Heise and Winzen, 2002; Næs et al., 2002). In such a case, the calibration developed on the lab instrument is not applicable to the estimation of samples recorded under the new conditions, and strongly biased estimations can result (Feudale et al., 2002; Despagne et al., 2000). This error occurs because the interactions between light and the sample will generally be different among different optical spectrometer systems, causing variations in the shape and intensity of the spectra (Shenk, 2004). To avoid a full recalibration of the model under these new conditions, standardisation methods can be applied to modify the response function of one instrument (or application) so that it is similar to that of another instrument. Principally, different strategies can be used: the regression coefficients of the prediction model can be standardised, or the values or the spectral responses can be predicted by mathematical manipulation (Feudale et al., 2002). The latter strategy is preferred in practice (Næs et al., 2002) and of the different methods available for this strategy, piecewise direct standardisation (PDS) is probably used most frequently (Feudale et al., 2002). The goal of PDS is to generate a transfer function, which is developed from samples measured in parallel on different instruments. For the derivation of the transformation matrix, each data point on the one instrument is reconstructed from data within a small window on the other instrument (Heise and Winzen, 2002). This is accomplished by performing a local multivariate transfer model between the spectral windows of the host and the central point in each corresponding spectral window on the master (Despagne et al., 2000). The requirements for such standard samples have been extensively discussed, as have methods for their selection from a larger set of samples (see references in Bouveresse and Campbell, 2008). In principal, the samples used for building the transfer function should optimally represent the spectral variability in the sample set that is to be transferred (Forina et al., 2007). Furthermore, standard samples should be similar when measured on the different instruments; i.e., they should be physically and chemically stable (Bouveresse and Campbell, 2008). These requirements can be limited by practical considerations and the success of PDS also depends on the extent to which the standard samples meet these criteria.

1.3. Previous research on the use of NIR spectroscopy in AD

In an earlier study reporting a method for NIR monitoring of the agricultural biogas process, Holm-Nielsen et al. (2006) focused on the representative sampling for process analytical characterisation of heterogeneous bioslurry. Samples removed from a down-scaled bioreactor were generated according to the theory of sampling (Esbensen and Geladi, 2010) and subsequent PLS models were created for total solids (TS), VS, and the chemical oxygen demand. From the modelling results, it was concluded that the mass reduction of the sample from the initial field sampling to the minute sampling for the ultimate analytical volume plays a central role for a representative referencing and for avoiding biased analytical results for calibration. In a subsequent study by Holm-Nielsen et al. (2007), the slurry from two commercial biogas plants was used for offline calibration of VS, ammonium (NH_4-N), organic nitrogen, total VFA, and single acids. The results of this study emphasized the good potential of NIR spectroscopy for the measurement of total VFA and, to some extent, the single acids by PLS regression. High accuracy of the calibration for VS and NH_4-N was also demonstrated. The monitoring of the latter parameter is particularly important in the case of protein-rich feedstock, which may result in inhibitory effects on the AD process due to the formation of free ammonia (Hansen et al., 1998). Holm-Nielsen et al. (2008) further performed an online monitoring of a glycerol-boosted anaerobic digestion process where a high performance of the NIR estimates for the concentration of total VFA and single acids could be verified. By using slurry from a biogas plant fed with manure and organic food industrial waste, this study confirmed the applicability of NIR to multicomponent mixtures. A biogas test plant was further used by Holm-Nielsen and Esbensen (2011), and was operated under dynamic conditions with respect to temperature and feeding. The NIR calibration results for the total VFA and individual acids reported satisfactory prediction performances under such conditions. This study also confirmed the suitability of VFA as a key indicator for assessing the AD process stability.

An online approach under laboratory conditions using cattle manure was reported by Lomborg et al. (2009) for the development of PLS models for VS and VFA. In order to increase gas yield, the digester was additionally fed with maize silage resulting in the formation of high VFA concentrations. The authors argued that real-time monitoring via NIR may be the key to enabling full-load operation, while concurrently balancing the feeding regime in order to avoid organic overload. The long-term applicability of NIR spectroscopy to the online estimation of the levels of total VFA, acetic acid, and propionic acid under large-scale conditions was first examined by Jacobi et al. (2009). An NIR sensor was integrated into a biogas plant and spectra were recorded over a period of 500 days using maize silage as feedstock. The evaluation of a spectra time series demonstrated the general applicability of this method in practice and the technical requirements for applying the NIR technique under rough working conditions were further addressed.

Ward et al. (2011a) used an NIR application for the estimation of the total alkalinity in the digester, stating that this method is also suitable for approximate indication of this parameter. These results are notable because carbonate ion, as an inorganic compound, is considered NIR inactive (cited in Shenk et al., 2008). The authors presumed that the NIR model for total alkalinity utilises a large range of feedstock-specific information and, hence, a validation with different feedstock composition was considered necessary for determination of the universal

applicability of the models. Further studies were conducted by Ward et al. (2011b) with the aim of combining different technologies for AD monitoring, including NIR analysis in the liquid phase, and gas chromatography and membrane-inlet mass spectroscopy in the head-space. The NIR spectra were obtained via a reflectance probe and aspects regarding correct positioning of the probe in the slurry stream were further addressed in this study. Work published by Raju et al. (2012) used this diffuse reflectance probe for the estimation of the total ammonia nitrogen in an anaerobic digester run on cattle manure. The digester was spiked with urea for variation of the parameter concentration range and the estimates by PLS models confirmed the ability for sound monitoring of this parameter via NIR.

Overall, this previous research demonstrates the large potential for monitoring agricultural AD with NIR technology. The studies show that the true benefit of this method is justified in its application for screening purposes aimed at detection of process imbalances at early stages. In this regard, analytical speed may be of critical value and can be 'more important than precision which again can be more important than accuracy' (Martens and Næs, 1991). This statement is connected to the assumption that the ability to detect relative changes with time in the concentrations, e.g., of VFA, rather than the correct evaluation of their actual values is the requirement for efficient AD monitoring (see Ahring et al., 1995) and, hence, is the relevant criterion for assessing the applicability of the NIR approach for this purpose.

For the practical handling of a wide range of feedstock materials that are commonly used in agricultural biogas plants, specific calibrations such as those developed for single feedstocks or single digesters with a constant feedstock mixture are of limited use because a great deal of time is required to calibrate each one individually. It follows that, in addition to precision and accuracy (Næs et al., 2002), the evaluation of model robustness versus changes in the feedstock composition and process stages is needed in order to implement this process analytical technology in common practice. In this regard, no clear statement with respect to the robustness of the models can be drawn from the above research. This is because the slurry used for calibration was kept quite constant in each of the different experiments. Furthermore, neither the reported cross-validation nor the splitting of one training data set into calibration and validation sets includes sufficient information for future use of the calibrations at a different time and place (Esbensen and Geladi, 2010).

A first attempt at independent testing of the robustness of NIR calibrations for the estimation of process parameters in slurry based on different feedstocks was reported by Ward et al. (2010). Calibrations for the estimation of acid parameters using pig slurry mixed with maize silage and chicken manure were developed separately and were subsequently used for estimation of VFA of samples from other material. Based on these results obtained from an independent validation, the authors concluded that NIR spectroscopy 'is not suitable for estimation of parameters in a material that was not included in the calibration set'. Results that were more promising were obtained for the estimation of VFA in a material that was included in the calibration set. With these findings, the study nicely reflected the general assumption that the success for calibration development for an NIR approach strongly relies on the extent to which the calibration samples span the spectral phenomena that vary in the target population which will be analysed in the future (Shenk et al., 2008).

2. Objective of thesis

The aim of this work was to investigate the potential for the development of robust NIR applications for the estimation of key process parameters in the digester slurry of agricultural biogas plants. NIR calibrations were developed for the parameters VS, NH_4-N, total inorganic carbon (TIC), total VFA, acetic acid, and propionic acid. In doing so, the study aimed at assessment of the extent to which a model can be used as a starting point for an NIR approach to be used under varying feedstock compositions, process stages, and temperature variations in the digester. This included the investigation of the ability of the NIR technology to be used as a lab method as well as for in situ monitoring at the digester.

In addition to the evaluation of the model robustness, the ability of NIR sensors to indicate short-term AD process dynamics was a further emphasis of this work. Special focus was placed on the time periods directly after feeding as these events are particularly important for control of the AD process. Investigations were also conducted regarding whether changes in the concentration level of key parameters can be indicated with sensitivity sufficient for real-time assessment of the feeding-induced process dynamics.

In detail, the following questions were addressed and are presented in chapters 4 and 5:

1. What is the potential for the development of feedstock-robust PLS models under controlled laboratory conditions? Does an increase in feedstock heterogeneity affect the estimation performance of the model assessed by internal validation including cross-validation and splitting of the training samples into calibration and validation sets? (chpt. 4.1)

2. Can the results obtained from the internal validation procedure be confirmed when testing the models on samples collected at different biogas plants? Furthermore, how much feedstock variability should be included in a calibration developed to be applied under this broad-based application? Does the NIR approach allow for specific estimation of acetic acid and propionic acid under such conditions? (chpt. 4.2)

3. Can the lab application (questions 1 and 2) be transferred to an online process analyser for an in situ monitoring of these parameters? What is the effect of an exchange of the digester content, i.e., a change in the slurry matrix, on the accuracy and precision of the NIR estimates obtained from a spectral time series? (chpt. 4.3)

4. Is it possible to use NIR spectroscopy for the indication of the short-term dynamics of the parameters VS and total VFA in a digester fed with maize silage? Can process imbalances induced by organic overloads be detected in real-time and are the two parameters independently estimated? (chpt. 5.1)

5. Is it possible to update the maize-specific calibration in order to account for variations in the feedstock composition and temperature changes in the digester? Can process imbalances induced by factors other than overfeeding be indicated via an NIR model? (chpt. 5.2)

3. General material and methods part

3.1. Overview of the experiments

Five different experiments are presented and discussed in this work addressing the questions outlined in chapter 2. Questions 1 to 3 were examined by a sequence of three different experiments reported in chapter 4. In the first experiment, offline PLS calibrations were developed for three different training sets whose samples varied with respect to their origin and feedstock material including energy crops and livestock waste from cattle, swine and poultry.

The second experiment describes the external (independent) validation of these lab models. The models were tested on a data set composed of samples removed from 26 biogas plants whose material was not used for calibration. The ability of the models to specifically detect acetic acid and propionic acid was further tested by serial spiking of slurry with both analytes. Spectral analyses for these experiments were performed with a Fourier-transform (FT)-NIR spectrometer (Vector 22/N, Bruker Optics, Ettlingen, Germany) in diffuse reflectance mode.

The third experiment describes the transfer of this lab application to an online process analyser for in situ monitoring on a technical scale. The offline spectra were subjected to spectra transfer using PDS. This spectra transfer was used together with samples recorded online for recalibration of the PLS models. In addition, models were performed solely with the samples measured online. The calibrations from both approaches were validated on a time series of spectral data collected from an independent fermentation run for comparison of the model performance [FT-NIR spectrometer (Matrix-F, Bruker Optics) which was equipped with a measuring head (Q-412, Bruker Optics) enabling spectral recording in diffuse reflectance].

Questions 4 and 5 were addressed in chapter 5 where the spectra recordings were only performed online. For experiment four, a pilot plant digester solely fed with maize silage and operated for eight months at highly variable feeding conditions was used for calibration development. Thus, the NIR calibration stability for long-term indication of VS and total VFA concentration courses was evaluated and special focus was put on the ability of the models to track the digester's short-term analyte concentration after perturbations induced by feeding. Therefore, the AD process was brought to critical conditions by discontinuous feeding, including phases of organic overload, and the calibrations developed were applied to a spectra time series for validation.

In experiment five, it was tested whether a change in the slurry matrix due to inclusion of further feedstock can be accounted for by updating the calibration. Additional calibration samples were generated via fermentation runs with different feedstock material and these samples were used together with the maize-slurry samples for recalibration of the models for VS and total VFA. The calibrations developed were tested on time series spectra for assessment of their ability to follow the overall concentration trend and to indicate short-term feeding-induced changes. PLS models were also developed for the parameters NH_4-N, TIC, acetic acid, and propionic acid. For the experiments four and five, two diode-array spectrometers of similar configuration (X-Three, NIR-Online GmbH, Walldorf, Germany) were used, providing online measurements in diffuse reflectance mode.

3.2. NIR calibration performance parameters

An overview of the statistical performance parameters used in this work to evaluate the NIR calibrations is provided in this chapter. Unless otherwise specified, their definitions are given according to Næs et al. (2002) and Esbensen (2009).

The root mean square error of calibration (*RMSEC*) is a measure for the model fit and is calculated from the calibration samples according to:

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{(N - A - 1)}}$$

where the y values are the actual values, the values of $\hat{y}$ are obtained by testing the calibration equation directly on the calibration data, A is the number of components used, and N is the number of samples tested. The *RMSEC* error is given in the original units of measurement.

Cross-validation is a validation technique solely based on the calibration data. For leave-one-out (full) cross-validation, the first sample is removed from the calibration set and the calibration is performed with the rest of the samples before it is tested on the first sample by comparison of y with $\hat{y}$. Next, the second sample is removed from calibration and this procedure is repeated until all samples have been deleted once. The estimate obtained from this technique is defined as the root mean square error of cross-validation (*RMSECV*), given as:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{y}_{CV,i} - y_i\right)^2}{N}}$$

where $\hat{y}_{CV,i}$ is the estimate for y_i based on the calibration equation with sample i deleted. As an alternative to leave-one-out cross-validation, the *RMSECV* can also be computed by deletion of segments of several samples from the calibration set, which was done in chapter 4.3.

The root mean square error of prediction (*RMSEP*) is a measure of the accuracy of the estimation. Unlike calibrations tested by cross-validation, the computations are conducted for a separate set of samples according to:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{N_p}(\hat{y}_i - y_i)^2}{N_P}}$$

where $\hat{y}_i$ and y_i are the estimated and measured reference values for the test samples, respectively, and N_p is the number of samples used for validation.

The standard error of prediction (*SEP*) expresses the precision of a model, and it is defined as the standard deviation of the estimated residuals for the validation set:

$$SEP = \sqrt{\frac{\sum_{i=1}^{N_p}(\hat{y}_i - y_i - bias)^2}{(N_P - 1)}}$$

where the bias reflects the systematic error, i.e., the average difference between $\hat{y}_i$ and y_i:

$$bias = \sum_{i=1}^{N_p} \frac{(\hat{y}_i - y_i)}{N_P}$$

The relationship between *SEP* and *RMSEP* is given by:

$$RMSEP \approx SEP^2 + bias^2$$

Exact equality is not obtained because *N* - 1 is used in the denominator of *SEP* instead of *N*, which is used for *RMSEP*.

The predicted residual error sum of squares (*PRESS*) is the residual Y variance calculated over the number of validation samples according to:

$$PRESS = \sum_{i=1}^{N} (\hat{y}_i - y_i)^2$$

The *PRESS* can be used for determination of the optimal number of factors assessing whether a new factor represents a significant addition to a model.

The ratio of performance to deviation (*RPD*) is defined as the ratio of the standard deviation of the reference data of the validation samples (SD_{val}) to the *SEP* (Malley et al., 2004):

$$RPD = \frac{SD_{val}}{SEP}$$

According to Williams (2001), the *RPD* can be used to evaluate the analytical efficiency of a calibration (Tab. 3.1). Note that this classification was developed for the evaluation of NIR performance for analysing crops, commodities, food, and feed (Malley et al., 2004).

Table 3.1: Significance of the values of the *RPD* statistics according to Williams (2001).

RPD value	Classification	Application
0.0–2.3	very poor	not recommended
2.4–3.0	poor	very rough screening
3.1–4.9	fair	screening
5.0–6.4	good	quality control
6.5–8.0	very good	process control
8.1+	excellent	any application

An alternative classification (Tab. 3.2) given by Malley et al. (2004) is recommended for environmental samples such as soil, sediments, animal manure, and compost.

Table 3.2: Significance of the values of the *RPD* statistics according to Malley et al. (2004).

RPD value	Classification
2.3–3.0	moderately successful
3.0–4.0	successful
4+	excellent

The coefficient of determination (R^2) describes the proportion of the Y variance in the reference data set that is explained by the NIR model. The value of R^2 can be calculated as the ratio of the explained variation to the overall variation. Alternatively, it can be directly calculated by subtraction of the ratio of the unexplained variation (residuals) to the overall variation from the maximum value of 1 as follows (Kessler, 2007):

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Table 3.3 provides an overview for interpretation of the performance of the NIR model using R^2. Similar to *RPD*, using R^2 values requires consideration of the fact that this parameter is influenced by the range of the Y values (Davies and Fearn, 2006).

Table 3.3: Guidelines for interpretation of R^2 (cited in Malley et al., 2004)

Correlation coef. (R)	R^2	Interpretation
up to ± 0.5	up to 0.25	not usable in NIR calibration
±0.51–0.70	0.26–0.49	poor correlation: reasons should be researched
±0.71–0.80	0.50–0.64	ok for rough screening, >50% of variance in Y accounted for by X
±0.81–0.90	0.66–0.81	ok for screening and some other “approximate” calibrations
±0.91–0.95	0.83–0.90	usable with caution for most applications, including research
±0.96–0.98	0.92–0.96	usable in most applications, including quality assurance
±0.99+	0.98+	usable in any application

The inclination of the regression line between the estimated values via NIR (ordinate) and the measured reference (abscissa) is defined as the slope. A slope below 1.0 indicates that samples at higher values tend to be underestimated whereas samples at lower values are overestimated. For slopes above 1.0, the opposite holds true.

The standard error of differences (*SED*) is a measure of the precision of an analytical method and is calculated according to (Workman Jr, 2008):

$$SED = \left\{\frac{\sum_{i=1}^{N}(y_i - \bar{y})^2}{N}\right\}^{1/2}$$

where y_i = an individual analytical result, $\bar{y}$ = the average analytical result for all replicate values, and N = the total number of samples (not the total number of replicates).

3.3. Reference methods of analysis

Total solids (TS) and volatile solids (VS)

The concentration of TS was measured by oven drying the sample at 105 °C to a constant mass (DIN 38414-S2). The VS concentration was determined by subsequent incineration of the oven-dried sample at 550 °C in a preheated muffle furnace for one hour (DIN 38414-S3).

Ammonium (NH_4-N)

The NH_4-N concentration was determined via distillation (VELP UDK 126 D, VELP Scientifica, Usmate, Italy) using MgO and subsequent titration (Schott TitroLine easy, SI Analytics GmbH, Mainz, Germany) to pH 5.2 with 0.05 M NaOH (DIN 38406-E5).

Total inorganic carbon (TIC)

The titrimetric analysis (with 0.1 M H_2SO_4 to pH 5.0) of the TIC was performed with the centrifuged supernatant (20 min, 15000 min^{-1}) of the sample. The amount of TIC was calculated according to the formula:

$$\mathrm{TIC} = \frac{20\ [mL]}{V\ [mL]} * a * 250$$

where V = volume of the sample and a = content of TIC determined via titration in millilitres of 0.1 M H_2SO_4.

Volatile fatty acids (VFA)

The six single fatty acids from acetic acid to oenanthic acid (including their isoforms), were analysed using a gas chromatograph (GC) (Agilent 6890N, Agilent Technologies, Santa Clara, California, USA). The acetic acid equivalent, defined as total VFA, was calculated via the molar mass of the single fatty acids. An aliquot from the centrifuged (Hettich Mikro 22, GMI Inc., Ramsey, Minnesota, USA) supernatant of the sample was taken (15 min, 14000 min^{-1}) for GC analysis. The detection limits for acetic acid and propionic acid were 0.10 and 0.03 g kg^{-1} fresh matter (FM), respectively. Alternatively, determination of total VFA was also performed via titration (DIN 38409-H7). For this, the untreated sample was distilled (VELP UDK 126 D, VELP Scientifica, Usmate, Italy) and subsequently titrated (Schott TitroLine easy, SI Analytics GmbH, Mainz, Germany) to pH 8.8 using 0.1 M NaOH.

All of the analyses were performed in duplicate. If the difference between these duplicates for one sample exceeded a critical value (3% for TS, 1% for VS, 5% for NH_4-N and TIC, and 10% for the acids), a third measurement was made. Unless otherwise stated, the average value of all of the replicates was used to determine the reference value. All data presented in this work refers to the FM.

3.3.1. Precision of the reference method

The repeatability of an analytical method describes the precision of within-run replicates, i.e., results obtained in the same laboratory by the same operator and with the same apparatus within a short time period (Miller and Miller, 2005). In contrast, the reproducibility describes the precision of between-run replicates and may involve different labs, which implies the involvement of different operators and apparatuses. For this study, the reference analyses were performed in a single laboratory with the same apparatus, but with the involvement of different operators over a time period of several months. According to Murphy (2010), such intermediate precision conditions can best be described as the 'long-term within-laboratory precision' of the method. The standard error of laboratory analysis (*SEL*) of the wet chemistry method was calculated from the duplicate measurements according to (Workman Jr, 2008):

$$SEL = \left\{ \frac{\sum_{i=1}^{N} (y_i - \bar{y})^2}{N} \right\}^{1/2}$$

where y_i = an individual laboratory result, $\bar{y}$ = the mean of the lab results calculated for the replicates, and N = the total number of samples (not the total number of analyses). The precision of the reference method is exemplarily given for the 80 samples used in experiment two (Tab. 3.4). For these samples, two separate analyses were performed in order to investigate the effect of an additional sampling together with a time-displaced second analysis on the precision of the method. A bulk sample was removed from the digesters and was mixed thoroughly before it was poured into two subsamples. The samples of 'subset 1' were frozen for a few weeks prior to their analysis; the samples of 'subset 2' were kept frozen over a several-month period. The two subsamples were both analysed in duplicate and the SEL_1 and SEL_2 were calculated separately. In addition, the four replicates were also used together for calculation of SEL_{1+2}. The results in Table 3.4 indicate that the SEL_1 and SEL_2 show rather similar precision; differences for NH_4-N and TIC were due to deviations in the analysis of a few samples. The SEL_{1+2} which tended to be larger, demonstrated a moderate loss in the precision for the parameters of both the solid (VS) and liquid phase. This was most likely due to a larger random error reflecting the time period between the two analyses. Presumably, it also indicated an additional sampling error that occurred by splitting of the bulk sample into two subsets.

Table 3.4: Precision of the reference analysis. The SEL_1 and SEL_2 were obtained from duplicate analyses of subsamples which originated from the same bulk samples (N = 80). The four replicates were used together for calculation of SEL_{1+2}.

Parameter [g kg^{-1} FM]	SEL_1 *(subset 1)*	SEL_2 *(subset 2)*	SEL_{1+2} *(subsets 1+2)*
VS	2.1	2.1	3.4
NH_4-N	0.06	0.12	0.13
TIC	0.19	0.11	0.30
Total VFA_{tit}	0.21	0.26	0.30
Total VFA_{GC}	0.19	0.22	0.25
Acetic acid	0.14	0.13	0.18
Propionic acid	0.06	0.07	0.10

3.3.2. Effect of sample treatment on wet chemistry results

The spectrally recorded samples were routinely frozen at -18 °C prior to the wet chemistry analyses to ensure that storage for a period of up to several months is possible without physicochemical alteration. However, it has been reported that freezing and subsequent thawing affect the wet chemistry results (Hecht, 2009; Henkelmann et al., 2010). Therefore, prior to the routine analysis performed for this work, the effect of sample storage was investigated for three different treatments. Firstly, samples were removed from the AD process and immediately ('fresh') subjected to wet chemistry. Secondly, subsamples were stored in airtight plastic bottles at 5 °C for one day ('cool') prior to their analysis. Thirdly, subsamples were stored at -18 °C for one day ('frozen') followed by a two-day thawing phase at 5 °C in a water basin before wet chemistry analysis was performed. For the test, a set of 34 samples was removed from lab digesters and large-scale biogas plants representing the sample matrices experienced in this work. A bulk sample of 2 L was removed via the outlet valve of the reactor and was thoroughly mixed before splitting it into three bottles (0.5 L) for generation of the subsamples 'fresh', 'cool', and 'frozen'. This procedure ensured that the subsamples were all from one lot and thus were representative of the bulk sample. Whether the different sample treatments resulted in significantly different results in the wet chemistry analysis was tested by applying the non-parametric Friedman test for paired samples. The test uses χ^2 statistics in order to assess the differences that occur between the total rank values of the different treatments (Miller and Miller, 2005). In case differences were observed among the three treatments, i.e., if the experimental value exceeded the critical value, the Wilcoxon-Wilcox test was subsequently performed for multiple comparisons of the median (Köhler et al., 2007).

The Friedman test indicated significant differences among the three treatments for all parameters (Tab. 3.5). Within-sample comparisons revealed that freezing affected wet chemistry results, whereas no statistically significant differences were observed between the treatments 'fresh' and 'cool'. Figure 3.1 illustrates the results obtained from the reference analysis for each subsample. The frozen samples tended to report higher concentrations of NH_4-N, which may be explained by the disintegration of the plant cells during thawing. A similar observation was made for the total VFA_{tit} and the single acids. Concentration of TIC, as determined via reference analysis, was lower for the subsamples that were subjected to freezing. Overall, the effect of freezing on the wet chemistry results can be considered low if related to the absolute values. It was decided that all samples were to be frozen prior to their analysis in order to avoid including uncertainty in the calibration modelling as a result of variations of the reference values due to different sample treatments (Workman Jr, 2008).

Table 3.5: Results from the Friedman test for paired samples and Wilcoxon-Wilcox test for investigation of the effect of different sample treatments on the reference analysis. Samples with a propionic acid concentration lower than the detection limit were omitted from calculation. Descriptive statistics are shown for the 'fresh' samples. *SD* = standard deviation.

Parameter [g kg^{-1} FM]	N	Range	Mean	Median	*SD*	Friedman (α = 5%)	Wilcoxon-Wilcox (α = 5%)
NH_4-N	34	1.29–3.98	2.27	2.00	0.73	s	frozen ≠ fresh, frozen ≠ cool
TIC	34	4.95–21.30	11.55	10.80	3.30	s	frozen ≠ fresh, frozen ≠ cool
Total VFA_{tit}	34	0.17–15.64	2.03	0.93	2.92	s	frozen ≠ fresh
Acetic acid	34	<0.10–8.40	0.97	0.49	1.52	s	frozen ≠ fresh, frozen ≠ cool
Propionic acid	22	<0.30–4.75	0.73	0.15	1.24	s	frozen ≠ fresh

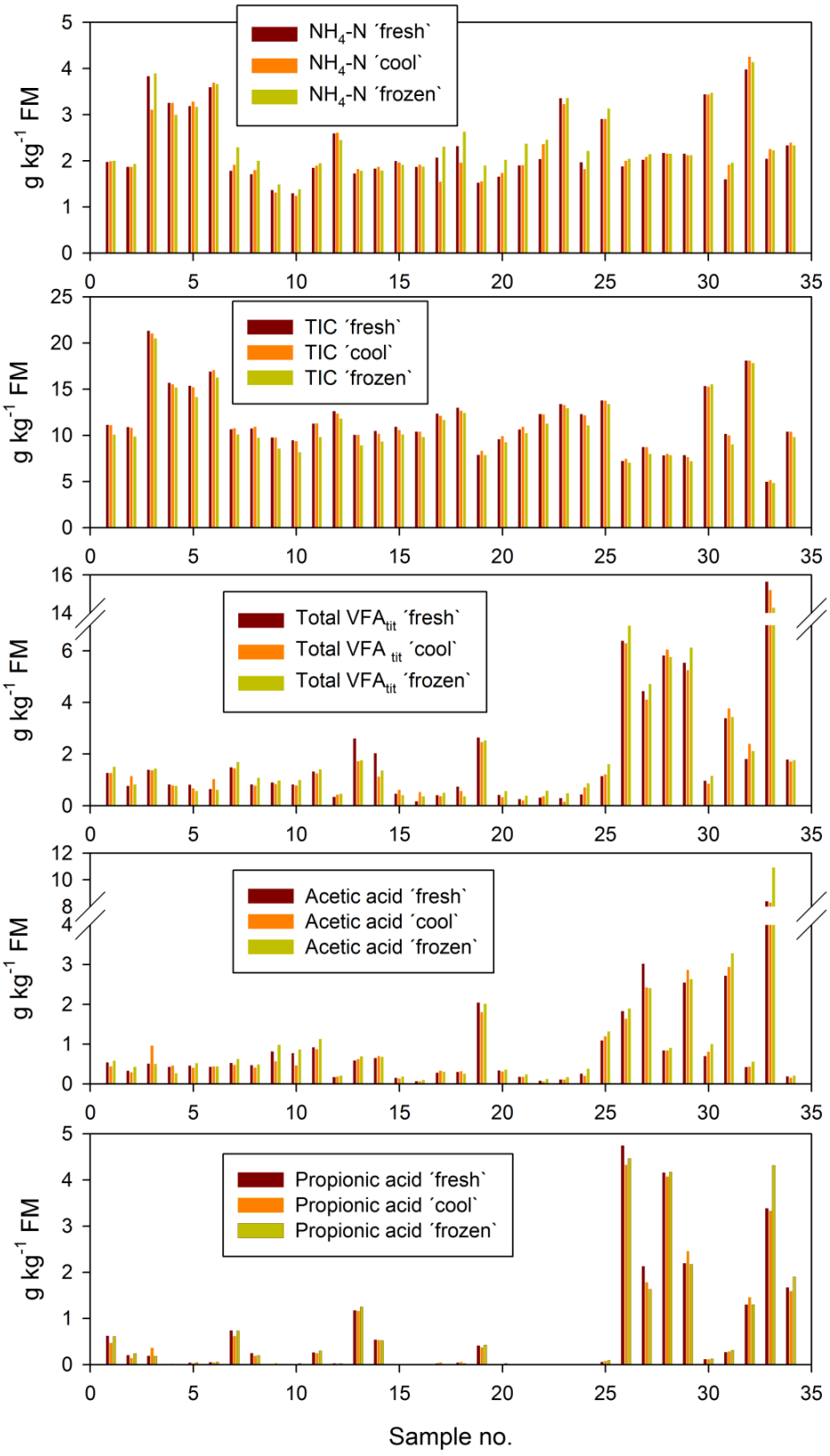

Figure 3.1: Graphical results of the study of the effect of sample treatment on the wet chemical analysis of the volatile parameters. 'Fresh' subsamples were removed from the digester and were analysed immediately, 'cool' subsamples were stored at 5 °C for one day prior to analysis, and 'frozen' subsamples were frozen at -18 °C for one day and subjected to thawing at 5 °C for two days prior to analysis.

During the course of sampling, spectral offline measurements, and wet chemistry analysis, the samples were temporarily exposed to ambient air conditions such as by stirring and decanting. As a consequence, oxygen entry and evaporation processes may alter the chemical characteristics of the samples, thereby affecting the wet chemistry results. This may be the case for analysis of the steam-volatile organic acids, for the NH_4-N owing to ammonia loss (Ndegwa et al., 2008), and for the TIC owing to release of CO_2. If such uncontrolled alterations of the sample occur after its spectral recording, the reference value no longer reflects the slurry at the time of its spectral recording. As a consequence, a mismatch in the X (spectral)-Y (reference) relationship, as one of the major sources of error in NIR spectroscopy (Workman Jr, 2008), will then affect the performance of the calibrations. Therefore, preliminary tests were performed for assessment of whether the sample handling in this work causes relevant changes in the concentration of these parameters. Sample material removed from two digesters was added to open cups (10 cm diameter) and exposed to ambient conditions. For each digester, three subsamples (0.4 L) were generated from a 2-L bulk sample. The first subsample was directly frozen for reference analysis while the other two subsamples were stored in the cups for 30 and 90 min, respectively. In order to simulate routine sample handling, the material was thoroughly stirred on several occasions during these time periods. Finally, these subsamples were frozen and wet chemistry analysis was performed.

Figure 3.2 illustrates the results of the wet chemistry analysis of the different subsamples exposed to ambient air conditions for different periods of time. The material of both digesters reported different concentrations for the tested parameters since the material was removed from the digesters at different process stages and used different feedstocks. For the sample material, the concentrations did not indicate any change over time, i.e., there was no loss over the 90 min time period. Instead, the repeated analyses clustered around the same values within the precision of each method. For the conditions tested, these data suggest that routine sample handling had no practical implication on the reference values, i.e., the error due to X-Y mismatch resulting from sample alteration can be considered negligible for this work.

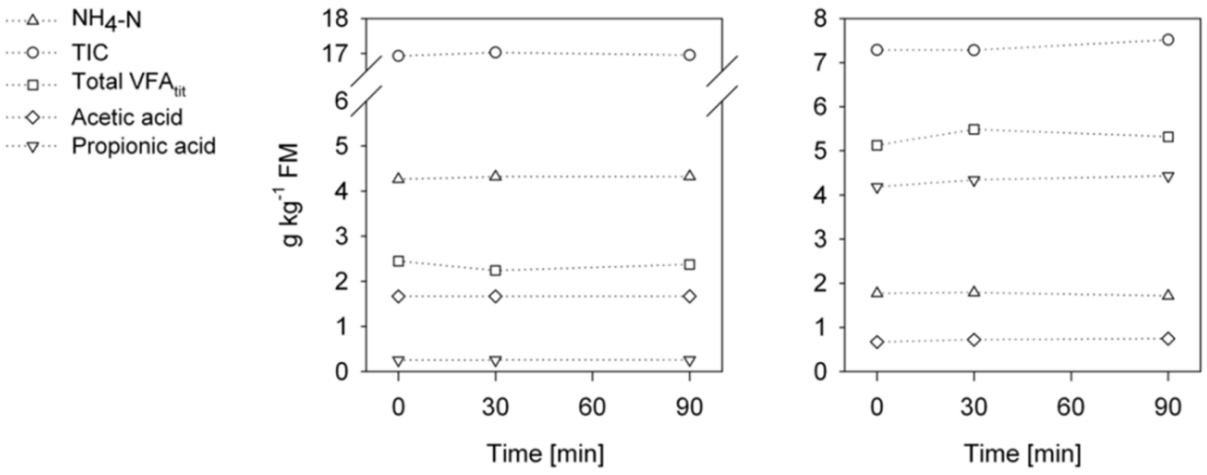

Figure 3.2: Effect of sample exposure to ambient air conditions for controlled time periods on the wet chemistry results of the potentially volatile process parameters. Left: Slurry removed from a digester mainly operated with a grass-clover mixture with high concentrations of NH_4-N and TIC. Right: Slurry removed from a digester mainly fed with maize silage at elevated propionic acid concentrations.

3.3.3. Comparison of total VFA_{tit} with total VFA_{GC}

For the samples collected in this work, determination of the reference value for total VFA was performed by both titration and GC analysis, and it was of interest to compare the results of the two methods. It has been reported that the determination of the total VFA as a sum parameter (e.g., from titration) shows only a low level of correlation with the acetic acid equivalent as obtained via GC analysis (Henkelmann et al., 2010). For the samples generated for the first three experiments (chpt. 4), the total VFA_{tit} result of each sample is plotted against the result of total VFA_{GC} for direct comparison of the analytical methods (Fig. 3.3). Overall, a high correlation was observed between the two methods, which is reflected by an r^2 value of 0.97. Clearly visible is the influence of the concentration level on the analytical results. While the titrimetric method indicated higher acid concentrations for the samples at a lower total VFA concentration, it resulted in lower concentrations for the above-average samples when compared to the GC results. From the plot, it can also be drawn that at higher total VFA concentrations, both methods reported a decrease in their correlation reflected by the higher degree of scattering of the samples around the regression line. This behaviour was associated with a loss in the (absolute) precision of both methods at higher concentrations. This is because routine analysis of the biogas process aims at precise indication of an early stage of acid accumulation and, hence, the methods are optimised to provide high precision primarily in the lower concentration region.

For the first two experiments of this work, calibration development was conducted with the total VFA_{GC} as the reference. For experiments three to five, the total VFA_{tit} was used as the reference for model development instead. The reason for choosing different reference methods within this work was a temporary defect of the GC, which limited prompt availability of the results for total VFA_{GC}. A similar precision for both methods was given by the *SEL* values (Tab. 3.4) and the implication the use of these alternative analyses may have on the total VFA model performance is considered further in the discussion of chapter 4.3.

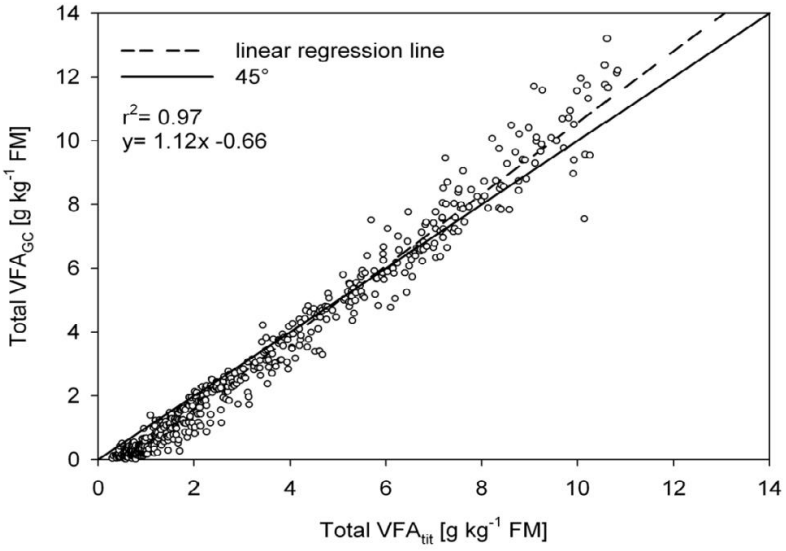

Figure 3.3: Total VFA results obtained from GC analysis plotted against the results obtained from titration (abscissa) for illustration of differences between the analytical methods. The samples obtained from the first three experiments are shown.

4. Development of a feedstock-robust near infrared application

4.1. Calibration development under lab conditions

Abstract

The calibration development for the estimation of the process parameters VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid is reported in this chapter. Spectra from samples in different training sets, varying with respect to their origin and feedstock composition, were assessed using PLS regression accounting for the complex matrices of the digester slurry of agricultural biogas plants. The comparison of the offline calibration results among the training sets revealed that an increase in the heterogeneity of the sample matrices did not result in a relevant performance loss of the NIR models. With a RMSECV of 4.0 and 0.16 g kg^{-1} FM, VS and NH_4-N exhibited the highest potential for estimation via NIR spectroscopy. The strong X-Y relationship for the structurally NIR-inactive TIC (RMSECV = 0.80 g kg^{-1} FM) indicated a satisfactory screening potential. Overall, this outcome emphasizes the high potential for the development of broad-based NIR calibrations under controlled lab conditions. The VFA parameters did not result in good NIR models and the limits for an accurate calibration were also demonstrated. However, there appears to be a realistic potential for a global total VFA model with an estimation error of 0.9 g kg^{-1} FM, which may support the use of NIR-based rapid screening of the dynamics of the acidity level in digester slurry.

4.1.1. Material and methods

4.1.1.1. Sample material and sampling procedure

All samples used for the first experiment were separated into two subsets that were later combined into one data set (Tab. 4.1). Subset A was composed of samples from laboratory digesters located at the Bavarian State Research Center for Agriculture in Freising, Germany. The samples were obtained from 14 flow-through digesters that had been fed with energy crops including silage from maize, grass-clover mixture and rye under mesophilic conditions. The effective volume of the continuously stirred digesters ranged from 28 to 2500 L. The gas production was measured via drum-type gas meters (Ritter GmbH, Bochum, Germany). The gas quality (CO_2, CH_4) was recorded via infrared sensors (Awite GmbH, Langenbach, Germany). For some of the digesters, different process conditions were induced by a temporary but strong variation of the OLR measured as VS of up to 15 kg m^{-3} d^{-1}. Sampling via outlet valves was performed a) under stabile AD conditions b) at an increased VFA level and c) under conditions characterised by high acidification accompanied by a strong decrease in biogas productivity.

During sampling, approximately 2 L of material were removed from the digesters and subsequently poured into a 0.5 L gastight bottle for temporary storage prior to the spectral recordings and subsequent reference analysis. In addition to these original samples, secondary samples were prepared by combining the slurry of different lab digesters. The slurry from

acidified lab digesters was sampled and directly mixed with material from digesters running under stabile AD conditions. This mixing was performed to diversify the range of different process stages for calibration. The mixing was performed in proportions ranging from 10 to 90% by volume, and care was taken not to apply serial dilutions that would result in an artificial collinearity of the chemical constituents. The secondary samples were treated the same way as the original samples, i.e., they were stored in bottles prior to the NIR measurements. The original and secondary samples together were defined as Subset A, which represents the different process conditions that can realistically occur in digesters solely fed with energy crops. Subset B was composed of samples from 12 commercial biogas plants (meso- and thermophilic operation) with a feedstock composition including various energy crops and manure from cattle, swine and poultry and, compared to Subset A, reflected the large range of materials added to agricultural biogas plants. For these plants, the average amount of manure in the feedstock was 32% FM (maximum 80% FM). For the creation of the secondary samples of Subset B, the slurry of different commercial plants was also partially mixed. Additionally, the slurry was also combined with acidified material from eight out of the 14 lab digesters. This combination was found necessary to simulate critical process conditions. By only selecting slurry from the commercial biogas plants, Subset B would have predominantly reflected a stabile AD process. Because it was the aim to develop NIR models for the mesophilic range, all samples were heated in a water basin at 40 °C after bottling for one hour prior to the spectral measurements. The complete procedure from the point of sampling until the spectral scanning required three hours on average.

Table 4.1: Data set characteristics of the samples used for calibration. Original samples: Slurry obtained from a single digester at one specific point in time. Secondary samples: Slurry obtained from different digesters under different process conditions and subsequently mixed in proportions ranging from 10 to 90% by volume. *For Subset B, nine secondary samples were prepared by mixing material from different biogas plants. An additional 42 samples were generated by combining the material from biogas plants with slurry from eight of the 14 lab digesters that were also used for Subset A.

			N		
Data set	Origin	Feedstock material	Original samples	Secondary samples	Total
Subset A	14 lab digesters	MS, GS, WCS, CCM	149	52	201
Subset B	8 lab digesters 12 biogas plants	MS, GS, WCS, CCM, CM, SM, PM	105	9/42*	156
Data Set C (Subsets A + B)	14 lab digesters 12 biogas plants	MS, GS, WCS, CCM, CM, SM, PM	254	103	357

Abbreviations: MS, maize silage; GS, grass silage (including clover); WCS, whole crop silage from rye; CCM, corn cob mix; CM, cattle manure; SM, swine manure; PM, poultry manure.

4.1.1.2. Near infrared spectroscopy

The samples were scanned offline with the Vector 22/N FT-NIR spectrometer equipped with an external integrating sphere (Fig. 4.1). Log (1/R) spectra from the diffuse reflection were recorded using a PbS detector from 12000 to 3700 cm^{-1}. An eccentrically rotating cuvette (90 mm diameter) was used to present the samples to the measuring area (20 mm diameter, 196 mm^2 observed area per turn). The pre-heated sample was poured into the cuvette and scanned in duplicate. The slurry was manually stirred before each scan to avoid sedimentation, and the temperature was measured. No stirring occurred during the spectral measurements. To consider the temperature effects on the spectral absorption (Hageman et al., 2005), the

temperature was randomly changed from 35 to 40 °C among the samples during the data collection. The sample was then returned to the bottle and stored at -18 °C until analysis by wet chemistry was performed. Referencing using a gold-coated reflectance standard was performed every 40 minutes. With 16 cm^{-1} resolution, 138 scans were needed for a constant signal to noise ratio above 1000, corresponding to a recording time of 60 s (six cuvette turns). The duplicate spectra recorded for each sample were averaged for the calibration.

Figure 4.1: FT-NIR spectrometer Vector 22/N (Bruker Optics) used for the lab application. Left: Spectrometer with integration sphere. Right: Sample presentation offline using cuvette measurements.

4.1.1.3. Multivariate data analysis

Prior to the calibration, a PCA was performed both for the X- (spectra) and the Y-data (reference) to determine the most dominant factors (Unscrambler 9.8, Camo, Oslo, Norway). Subsequently, a PLS regression was performed using the software package Opus 6.5 (Bruker Optics). For the Subsets A and B, the preliminary calibrations were determined for the detection of outliers in X and Y by applying the first derivation spectra in the 9000 to 4000 cm^{-1} region. For the six parameters, each calibration was automatically tested by a leave-one-out cross-validation. The final calibrations were determined from an optimisation routine implemented in the software after the removal of the outliers. During the optimisation, various frequency regions and spectral pre-treatments were systematically tested to determine the optimal combination for each parameter. The 9000 to 4000 cm^{-1} range was divided into five segments, whereas the water bands at approximately 6900 and 5150 cm^{-1} were excluded from the calibration because the narrowing of the spectra in these regions indicated a deviation from Beer's law, i.e., it indicated the non-linearity between chemical and spectral readings (Workman Jr, 2008). The pre-treatments included evaluation of the first derivative, MSC, SNV transformation, subtraction of a straight line and subtraction of a constant offset. During the optimisation, the maximum number of PLS components was restricted to 10. The calibrations resulting in the lowest error of estimation for the leave-one-out cross-validation after optimisation were termed Models A_{CV}, B_{CV} and C_{CV} for Subset A, Subset B and the complete Data Set C, respectively (Fig. 4.2). For similar error values, the model with the lower rank was preferred. The optimal number of PLS components was chosen as proposed by Opus based on the impact of each new component on the reduction of the *PRESS* calculated for the cross-validation. Each training set was then separated into two halves, and one was used for the calibration and the other for an internal test set validation. The validation samples were selected using the Kennard-Stone algorithm (Kennard and Stone, 1969) to evenly span the X-space after the calculation of a PCA with 10

PCs. The calibrations were recalculated and the test set was validated to result in Models A_{TS}, B_{TS} and C_{TS}.

For the calibration, the samples should be as evenly distributed over all combinations of variables as possible (Næs et al., 2002). However, for the acidity parameters, the training sets exhibited a highly skewed distribution that reflected the large number of samples obtained at low acid concentrations (Tab. 4.2). Spiking the samples with acids was not an alternative because artificial samples do not necessarily correspond spectrally to the slurry samples in their natural state (Esbensen, 2009). The calibrations for the acid parameters were first calculated by applying all of the samples of each training set because all of the samples were found to be potentially relevant and of different feedstock material and origins and thus representative of the X-space. As an alternative calibration approach, a cluster analysis (Heise and Winzen, 2002) (20 clusters, 20 iterations, Euclidean distance) using those samples of Data Set C with a total VFA concentration below 2 g kg^{-1} FM was applied over the 9000 to 4000 cm^{-1} region of the first derivative spectra (Unscrambler). From each cluster, five samples were randomly selected and were then recombined with the samples at a total VFA level above 2 g kg^{-1} FM. The calibrations for total VFA, acetic acid and propionic acid were calculated using this recombined Data Set C according to the above procedure via Opus, resulting in Models C_{2CV} and C_{2TS}. The low-concentration samples that were excluded from the calibration via clustering were subsequently estimated by Model C_{2CV} to specifically assess the model's behaviour at total VFA concentrations below 2 g kg^{-1} FM.

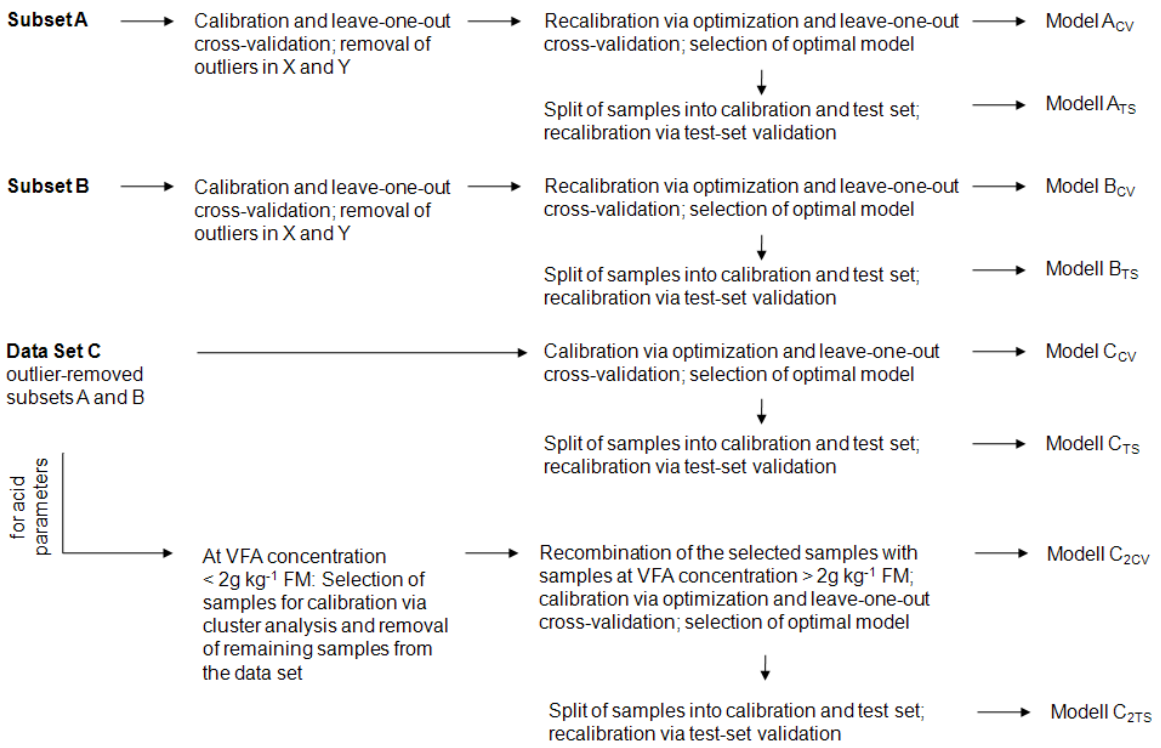

Figure 4.2: Flow chart for the development of the calibration algorithms and their validation. For the models validated via a test set, the same optimisation algorithm defined for the cross-validated models was used. An equal separation of the samples into calibration and test sets was performed by evenly spanning the spectral space using the Kennard-Stone algorithm after the calculation of a PCA with 10 PCs. See text for further details.

4.1.2. Results and discussion

4.1.2.1. Reference data set characteristics

Three samples were removed from the subsets due to an overestimation of the VS concentration that was caused by a rapid sedimentation during the spectral recordings at concentrations <30 g kg^{-1} FM. One sample that exhibited a gross error during the reference analysis for VS was removed. Two samples were defined as outliers because they showed large X- and Y-residuals with a strong influence on the model. Both subsets exhibited a nearly normal distribution for VS and differed slightly in their concentration ranges and mean values (Tab. 4.2). A strong correlation between TS and VS was observed for all sets because both parameters differ only by ash content (Tab. 4.3). The samples with a high concentration of NH_4-N and TIC originated from digesters fed with N-rich feedstock, such as grass silage and poultry manure. For Subset A, a weak bimodal distribution was observed for these two parameters, resulting in a higher standard deviation relative to Subset B. A strong positive correlation between NH_4-N and the TIC was observed. Because collinearity can cause model instability, the TIC calibration for all three sample sets was performed using only the samples in the range of the H_2CO_3 buffer system from pH 5.0 to 8.0 (Hecht, 2009). After the removal of samples with a pH >8.0, the collinearity between NH_4-N and TIC decreased because samples with high values for both parameters were excluded. Concurrently, the negative correlation between the TIC and the acid parameters increased because an increase in the VFA concentration was less compensated for by the buffering effect of the degradation of N-rich substrate (Kissel et al., 2009). The total VFA concentration was up to 10 g kg^{-1} FM indicating acidified conditions far above the potentially critical limit of 3 g kg^{-1} FM (Holm-Nielsen et al., 2007).

Table 4.2: Descriptive statistics of the parameters after the removal of six outliers. The arithmetic mean, median, standard deviation (*SD*) and the *SEL* of the wet chemistry method are presented. For the TIC, the values for the sets are provided before/after removal of samples with pH >8.0. The latter sets were used for the calibration of the TIC. For the acid parameters the low values indicate the limit of detection.

Parameter	Data set	N	Range	Mean	Median	*SD*	*SEL*
VS [g kg^{-1} FM]	Subset A	197	48.2-130.3	81.1	78.0	16.6	
	Subset B	154	35.0-116.6	68.9	67.9	16.8	
	Data Set C	351	35.0-130.3	75.8	74.8	17.7	2.2
NH_4-N [g kg^{-1} FM]	Subset A	197	0.97-5.38	2.64	2.19	1.08	
	Subset B	154	1.20-4.48	2.35	2.31	0.64	
	Data Set C	351	0.97-5.38	2.51	2.25	0.92	0.11
TIC [g kg^{-1} FM]	Subset A	197/112	3.8-23.8/3.8-12.4	12.0/8.5	10.7/8.8	5.0/2.3	
	Subset B	154/104	4.8-24.5/4.8-14.1	11.1/9.3	10.4/8.9	3.8/2.1	
	Data Set C	351/216	3.8-24.5/3.8-14.1	11.6/8.9	10.6/8.8	4.5/2.2	0.15
Total VFA [g kg^{-1} FM]	Subset A	197	<0.13-9.89	2.32	1.12	2.61	
	Subset B	154	<0.13-9.39	2.49	1.51	2.58	
	Data Set C	351	<0.13-9.89	2.40	1.22	2.60	0.22
Acetic acid [g kg^{-1} FM]	Subset A	197	<0.10-5.71	0.98	0.47	1.11	
	Subset B	154	<0.10-6.61	1.55	0.93	1.55	
	Data Set C	351	<0.10-6.61	1.23	0.61	1.35	0.13
Propionic acid [g kg^{-1} FM]	Subset A	197	<0.03-7.24	1.29	0.34	1.81	
	Subset B	154	<0.03-5.74	0.81	0.22	1.15	
	Data Set C	351	<0.03-7.24	1.08	0.29	1.58	0.11

Table 4.3: Parameter cross-correlation. The coefficient of correlation (r) among the TS, VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid parameters. The slash separates r values calculated for the different sets (Subset A/Subset B/Data Set C). For the TIC, r was calculated for the overall sample sets and for the pH-reduced sets (see text). The latter sets were used for the calibration of the TIC.

	TS	VS	NH_4-N	Total VFA	Acetic	Propionic
VS	.96/.98/.97	-				
NH_4-N	.12/.51/.28	.10/.42/.11	-			
Total VFA	-.18/.26/-.06	.05/.35/.11	-.39/-.32/-.36	-		
Acetic	-.05/.16/.01	.01/.26/.05	-.12/-.32/-.21	.78/.93/.83	-	
Propionic	-.21/.36/.06	-.07/.41/.13	-.45/-.22/-.37	.90/.82/.84	.45/.56/.41	-
TIC	.17/.46/.30	.06/.34/.11	.96/.91/.94	-.57/-.56/-.57	-.27/-.55/-.39	-.62/-.38/-.53
TIC (<pH 8.1)	.47/-.32/.04	.39/-.41/.03	.76/.71/.74	-.88/-.77/-.82	-.66/-.68/-.57	-.81/-.67/-.76

Figure 4.3 summarises the PCA results calculated for the standardised reference parameters for Data Set C. According to the PC1-PC2 loading plot (Fig. 4.3b), the major variance (46 %) is explained by PC1 and highlights the differences among the samples with both a low TIC concentration and a high acid concentration. This suggests that the carbonate buffer neutralises the acidity level in the digester because an increased VFA concentration potentially leads to a decrease in the buffering capacity as the bicarbonate becomes protonated and is released as carbon dioxide (Jantsch and Mattiasson, 2004). This plot also confirmed the high positive correlation between the levels of NH_4-N and TIC. The numerous samples exhibiting below-average values for PC1 in the corresponding score plot (Fig. 4.3a) are those samples with a low acid concentration.

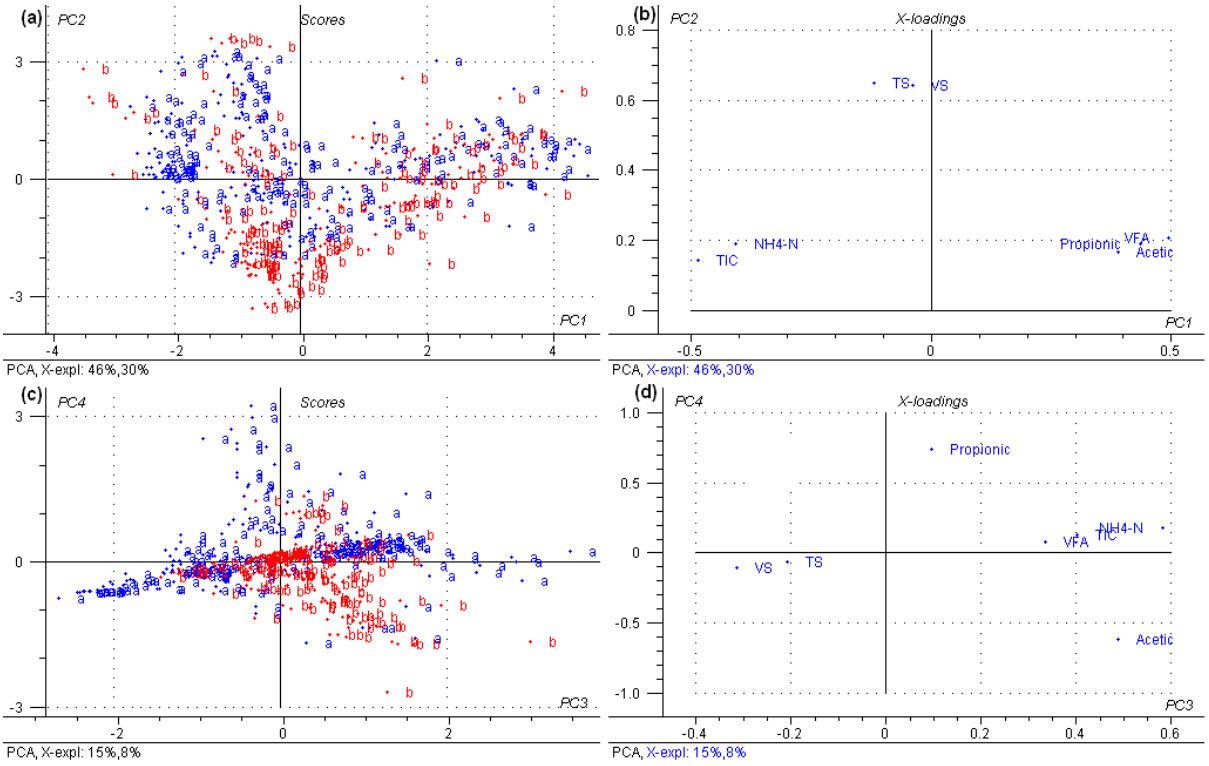

Figure 4.3: PCA of the TS, VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid parameters (standardised). The PCA was calculated for Data Set C, indicating the samples of Subset A (a) and Subset B (b) after the removal of outliers. (a): PC1-PC2 score plot; (b): PC1-PC2 loading plot; (c): PC3-PC4 score plot; (d): PC3-PC4 loading plot.

Information about the TS/VS concentration can be obtained by noting that PC2 in the loading plot explains 30 % of the variance in the data set. PC3 (Fig. 4.3c and 4.3d) reflects 15 % of the overall variation among the parameters, indicating that differences exist between TS/VS and the other parameters. PC4 explains 8 % of the variation, which reflects differences between the samples of high and low acetic acid and propionic acid concentrations.

4.1.2.2. Major spectral characteristics

Figure 4.4 (left graph) depicts the log (1/R) spectra of the samples measured from 12000 to 3700 cm^{-1}. The main visible difference among the raw spectra is the large vertical shift among the samples. This baseline offset illustrates that scattering is frequency dependent and increases at higher frequencies (Kessler, 2007). The water absorption features at approximately 6900 and 5150 cm^{-1} are clearly evident. As shown in the graph on the right, the first derivative can correct for the baseline shifts and amplify minor spectral variations in the spectra.

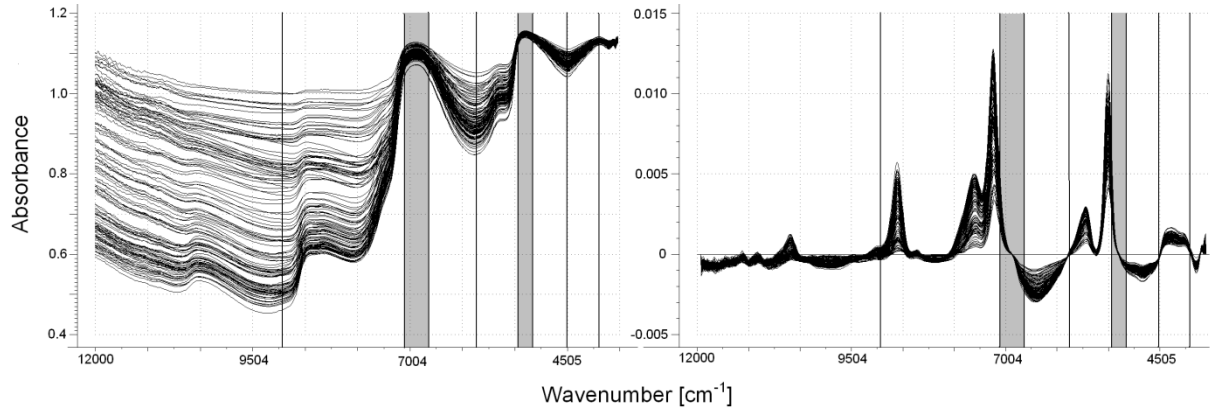

Figure 4.4: Left: Raw log(1/R) NIR spectra in the 12000 to 3700 cm^{-1} region. For PLS regression, the following regions (cm^{-1}) were used for the optimisation as indicated by the vertical lines: 9000-7100, 6700-6000, 6000-5300, 5050-4500, 4500-4000 cm^{-1}. The grey areas indicate those regions excluded from the calibration (see text for further detail). Right: First derivative spectra of the same samples (17 smoothing points, Savitzky Golay, 2nd order polynomial). 33% of the samples of Data Set C were randomly selected for this illustration.

Figure 4.5 illustrates the PC1-PC2 score plot of a PCR performed for TS with Data Set C after the removal of outliers. The spectral data were restricted to the 9000 to 4000 cm^{-1} region (first derivative, mean centred), including the water bands. Although PC1 explains 81 % of the spectral variation, no clear separation is apparent between the two subgroups, indicating that the separation of the samples into Subset A (a) and Subset B (b) based on the feedstock material does not account for the major spectral characteristics of the slurry. PC1 cannot be attributed to the TS concentration or to the further chemical analytes investigated in this study because these properties are associated with higher PCs (not shown here). Presumably, PC1 contains information related to the physical state of the slurry, describing the global trend of the spectra (Tillmann, 1996). PC2 explains 17 % of the spectral variations, and the samples of Subset A generally exhibit score values that are above average, whereas the samples of Subset B exhibit values that are below average. Because PC2 explains 65 % of the variation in the TS concentration, this gradient reflects the overall higher TS concentration for the samples of Subset A compared to those of Subset B.

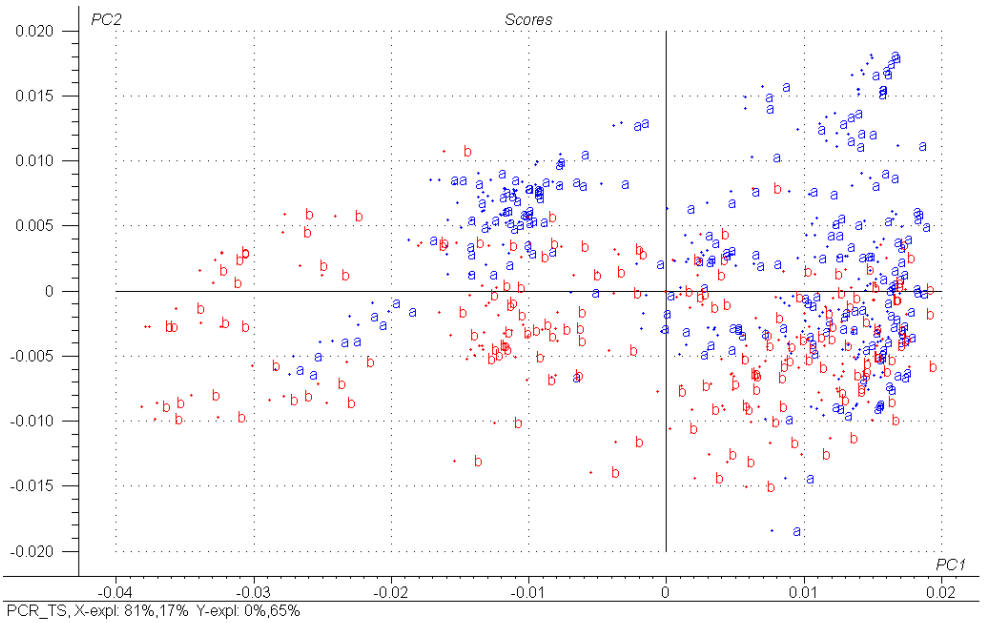

Figure 4.5: PC1-PC2 score plot of a PCR calculated for the TS for the investigation of major X-Y relationships in Data Set C, indicating the samples of Subset A (a) and of Subset B (b) after the removal of outliers applying 9000-4000 cm^{-1} region (first derivative, 17 smoothing points).

4.1.2.3. PLS regression modelling

Volatile solids (VS)

The validation for the VS in Subset A showed a high stability, with only minor fluctuations between the performance of the Models A_{CV} and A_{TS} (Tab. 4.4). This result indicated that the overall number of samples is likely sufficient to describe the differences in the physical and chemical state of this training set with respect to VS; otherwise the test set validation should have resulted in deterioration of the model performance (Conzen, 2005). The results for Subset B and Data Set C showed a similar outcome because the cross-validation and test set validation for these sets were very similar. The models for all three data sets resulted in mean estimation errors of approximately 4 g kg^{-1} FM. This error was slightly higher than that reported by a previous study by Holm-Nielsen et al. (2007) that used slurry from only two biogas plants (*RMSECV* = 2.7 g kg^{-1} FM). The high r^2 above 0.90 that was reported for all training sets supports the finding that an increase in substrate heterogeneity did not negatively affect the model performance.

The estimated versus measured plot in Figure 4.6a illustrates the VS results for the Models A_{CV} and B_{CV}. The *RMSECV* is plotted as a function of the number of PLS components for both models and indicates where the error reaches a minimum (Fig. 4.6d). For VS, the higher PLS components presumably corrected for a number of minor effects, such as temperature variations, non-linear instrument responses or remaining light scattering effects (Martens and Martens, 2001). The number of PLS components that was considered optimal for Subset A (and also for Data Set C) was higher than that for Subset B. For Subset A, the use of original and secondary samples from lab digesters that were primarily fed with maize silage may have contributed to the creation of many average samples in X. Calibration sets that possess a significant amount of redundant information tend to overfit, and because of their accumulation, the mean estimation

error for cross-validation decreases as the number of PLS components increases (Kessler, 2007). Strategies such as the selection of representative calibration samples by cluster analysis, may help to determine the optimal number of components (Isaksson and Næs, 1989; Næs et al., 2002; Kessler, 2007).

Table 4.4: Statistics of the PLS models of the VS, NH_4-N and TIC parameters. The results of the cross-validation (models X_{CV}) and test set validation (models X_{TS}) for Subset A, Subset B and Data Set C are presented. The coefficient of determination was calculated for the calibration (R_c^2) and validation (r_v^2).

	Model	Pretr	Region cm^{-1}	PLS	*RMSEC*	*RMSECV*	*RMSEP*	R_c^2	r_v^2	Bias	Slope	*RPD*
VS [g kg^{-1} FM]	A_{CV}	1st +	6703-5299,	6	3.73	3.97		0.95	0.94	0.03	0.94	4.1
	A_{TS}	Line	4505-3996		4.25		3.93	0.95	0.93	0.33	0.97	4.0
	B_{CV}	Offs	6703-5994,	5	3.73	3.87		0.95	0.94	0.01	0.94	4.3
	B_{TS}		5052-4497		3.54		4.08	0.95	0.94	-0.56	0.96	4.2
	C_{CV}	1st +	6703-5299,	7	3.97	4.14		0.95	0.94	-0.02	0.94	4.2
	C_{TS}	MSC	5052-3996		4.26		4.19	0.94	0.94	0.39	0.92	4.1
NH_4-N [g kg^{-1} FM]	A_{CV}	MSC	5052-3996	4	0.15	0.16		0.98	0.97	-0.01	0.98	7.0
	A_{TS}				0.16		0.14	0.97	0.98	-0.12	0.97	7.8
	B_{CV}	MSC	6001-5646,	3	0.16	0.17		0.93	0.93	-0.01	0.93	3.9
	B_{TS}		5052-3996		0.16		0.17	0.94	0.93	0.03	0.91	3.9
	C_{CV}	MSC	6001-5646,	4	0.16	0.16		0.97	0.96	0.01	0.97	5.7
	C_{TS}		5052-3996		0.17		0.16	0.96	0.97	0.01	1.00	6.1
TIC [g kg^{-1} FM]	A_{CV}	SNV	5400-5300,	7	0.70	0.80		0.91	0.87	0.01	0.90	2.8
	A_{TS}		5050-3900		0.61		0.85	0.94	0.84	0.23	0.92	2.6
	B_{CV}	SNV	5400-5300,	6	0.67	0.74		0.90	0.87	0.02	0.88	2.8
	B_{TS}		5050-3900		0.65		0.68	0.92	0.88	0.07	0.89	2.9
	C_{CV}	SNV	5400-5300,	8	0.67	0.74		0.91	0.89	-0.01	0.90	3.0
	C_{TS}		5050-3900		0.63		0.74	0.92	0.89	0.06	0.85	3.0

Abbreviations: Pretr, pre-treatment; 1st, first derivative (calculated with 17 smoothing points); Line, subtraction of a straight line; Offs, subtraction of a constant offset; MSC, multiplicative scatter correction; SNV, standard normal variate; PLS, number of PLS components; *RMSE*, root mean square error reported for calibration (*C*), cross-validation (*CV*) and test set validation (*P*).

Ammonium (NH_4-N)

An *RMSEP* value of approximately 0.16 g kg^{-1} FM for all three sets indicated the high accuracy for the estimation of NH_4-N under broad-based modelling conditions. The r^2 value was as high as 0.97 for Data Set C, which highlighted the potential of NIR spectroscopy for estimating this parameter and confirmed the results reported elsewhere (Holm-Nielsen et al., 2007; Raju et al., 2012). The relatively low *RPD* value for Subset B compared to the other sets is primarily due to the lower standard deviation reported for this subset (Tab. 4.2). The pre-treatment using MSC resulted in the best model performance and accounted for the baseline offsets and the global spectral trend (Tillmann, 1996). For all three sets, the spectral region between 5052 and 3996 cm^{-1} was particularly important for the calibration (Tab. 4.4). A study of the regression coefficient revealed that the spectral region near 4525 cm^{-1} showed the strongest response to changes in the NH_4-N concentration (not shown here). This is the region in which ammonia in water exhibits a strong band (Workman and Weyer, 2008). The models for Subset B also contained the smallest number of PLS components (Fig. 4.6e), which can presumably be explained with similar reasons as discussed for VS.

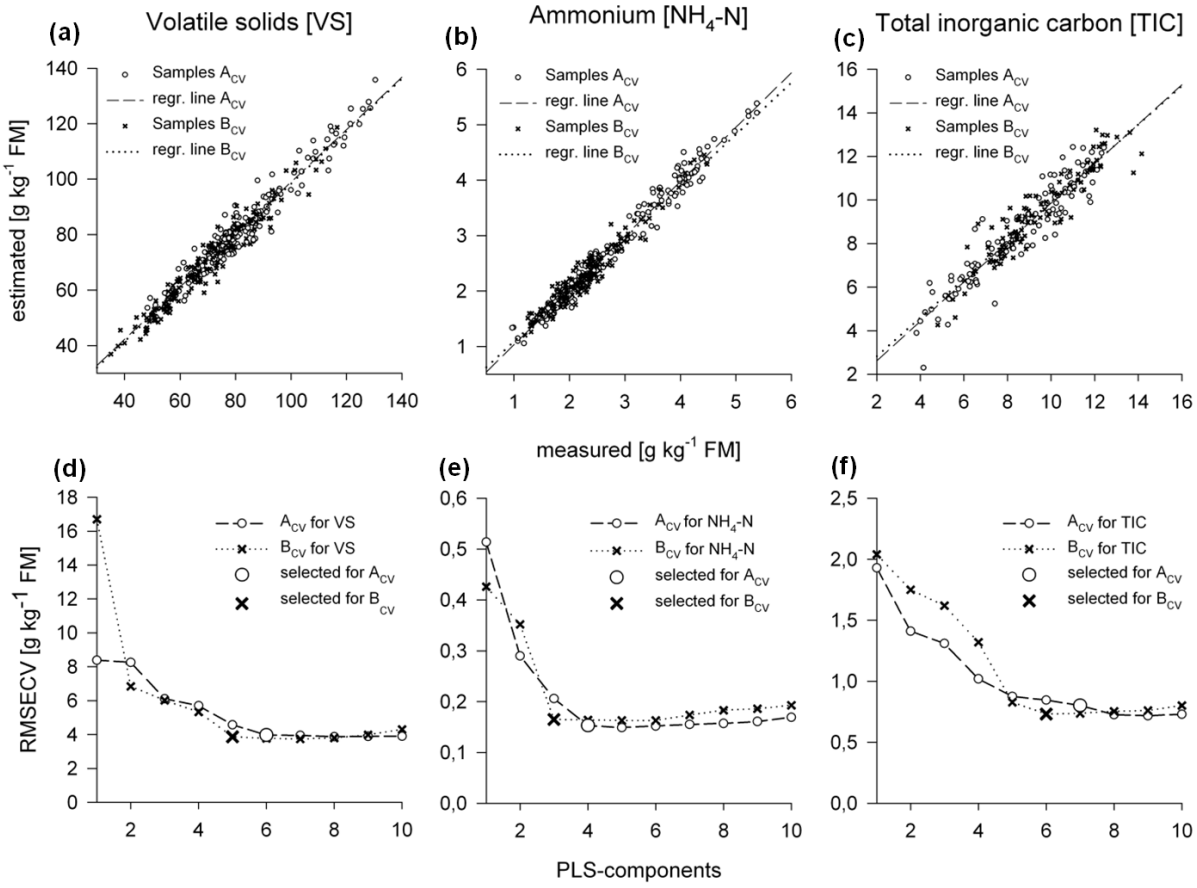

Figure 4.6: NIR estimates of VS, NH_4-N and TIC validated via leave-one-out cross-validation for Subsets A and B. Top: The estimated vs. measured plots related to the A_{CV} and B_{CV} models in Table 4.4. Bottom: The *RMSECV* plotted as a function of the number of PLS components for the A_{CV} and B_{CV} models, indicating where the estimation error reaches its minimum.

Total inorganic carbon (TIC)

Theoretically, there are no absorption bands for mineral components and their ionic forms within the NIR spectrum (Clark et al., 1987; Shenk et al., 2008). However, it has been reported that the carbonate ion constitutes an exception for the mineral components, exhibiting minor absorptivities at 5263, 5000 and 4630 cm^{-1} (Hunt and Salisbury, 1971; Osborne and Fearn, 1986). This ion has been further analysed via NIR spectroscopy in several geological studies, which reported two bands near 4255 and 3920 cm^{-1} (cited in Workman and Weyer, 2008). It was argued that 'the 3920 cm^{-1} band is most likely the sum of an infrared-forbidden symmetric stretch and twice the asymmetric stretch ($v_1 + 2v_3$). The 4255 cm^{-1} band has been assigned as the second overtone of the strong asymmetric stretch ($3\ v_3$).' In contrast to the calibrations developed for the other parameters by applying the optimisation routine of the software (Opus), each training set model for the TIC was developed solely using these spectral regions of interest. Because these band positions were reported to shift depending on the environmental conditions, the region from 5400 to 3900 cm^{-1} was used for this model development (the usable detector limit was 3900 cm^{-1}). The results provided in Table 4.4 exhibit good correlation of the X-Y space inherent in all the three sample sets. With an *RPD* of approximately 3, the internal validation indicated a moderately successful screening potential (compare with Tab. 3.2).

Figure 4.7 depicts the regression coefficients of the TIC calibration for the three data sets after cross-validation. The high similarity among the patterns in the region from 4800 to 4000 cm^{-1} is evident and highlights that for all three sets, the changes in the TIC concentrations were associated with similar frequency regions. However, the reasons for this characteristic spectral response over the large range of process conditions and sample matrices remain unclear. The individual inorganic compounds can potentially also be estimated via complexation with an organic compound (Shenk et al., 2008) or by an effect on some absorption bands, e.g., ´by means of an induced shift in the absorption due to water` (Begley et al., 1984; Osborne and Fearn, 1986). Further investigations are needed to provide a more comprehensive interpretation of these findings.

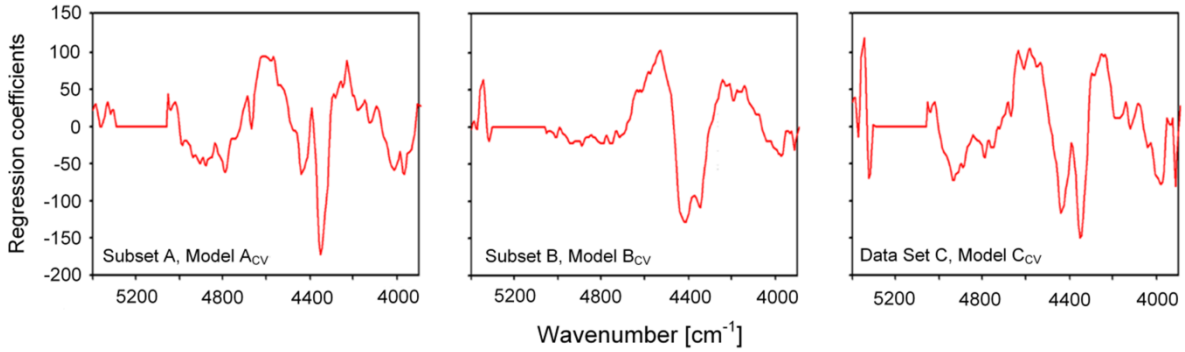

Figure 4.7: Regression coefficients for the calibration of the TIC after a leave-one-out cross-validation. The models were calculated using the 5400-5300 and 5050-3900 cm^{-1} regions; log(1/R) spectra were pre-treated with a SNV transformation. The regression coefficients relate to the models in Table 4.4.

Total volatile fatty acids (VFA)

For total VFA, the cross-validation and test set validation for Subset A provided similar results, demonstrating internal stability (Tab. 4.5). The results for Subset B also indicated that the number of samples sufficiently described the variation in this subset, although the number of samples was smaller than that of Subset A and had a higher matrix variation due to the use of samples containing livestock residues. An *RMSEP* value between 0.8 and 0.9 g kg^{-1} FM for total VFA for Models A_{TS}, B_{TS} and C_{TS} illustrated that the increase in substrate heterogeneity that was indicated by these data did not relevantly affect the model performance (also see Fig. 4.8a). This error is in accordance with an *RMSECV* value of 0.82 g kg^{-1} FM for the titrated total VFA reported from a study that applied online measurements at a digester solely fed with maize silage (Jacobi et al., 2009). Other research resulted in mean estimation errors ranging between 0.2 and 1.6 g kg^{-1} FM, and these errors were obtained from different experimental conditions, validation methods and concentration ranges (Holm-Nielsen et al., 2007; Lomborg et al., 2009).

A reduction in the skewness of the concentration range did not affect the mean estimation error because the total VFA calibration result for Model C_{2CV} (after selecting the representative samples at concentrations below 2 g kg^{-1} FM via a cluster analysis) was similar to those for Model C_{CV} (Tab. 4.5). A graphical illustration of the validation of Model C_{2CV} is provided in Figure 4.9a. The samples that were excluded from the calibration were separately estimated by the model (inset). The *RMSEP* value of 0.82 g kg^{-1} FM reported for these samples illustrated

that a change in the total VFA concentration in the below 2 g kg^{-1} FM region can be estimated as accurately as that for the overall total VFA range.

Because of the skewed distribution in the present work, the estimation error and *RPD* require careful interpretation. In contrast, the r^2 value is a direct measure of the model performance. For a total VFA concentration up to 10 g kg^{-1} FM (Tab. 4.2), an r^2 value near 0.90 indicated a strongly correlated X-Y space in all training sets. To illustrate the effect of a decrease in the overall variance of the analyte concentration on r^2, Model C_{TS} was also test-set validated with the samples at a concentration below 5 g kg^{-1} FM. As a result, r^2 decreased from 0.89 (Tab. 4.5) to 0.72, a value which more realistically reflects the proportion of the variance in the total VFA concentration that is accounted for by the model at an early to intermediate stage of acid accumulation. In contrast, with a value of 0.86 g kg^{-1} FM the *RMSEP* remained constant supporting the assumption that for the model tested, the accuracy of the estimates was independent of the concentration range used for validation.

Table 4.5: Statistics of the PLS models of the total VFA, acetic acid and propionic acid parameters. The results for the cross-validation (models X_{CV}) and test set validation (models X_{TS}) for Subset A, Subset B and Data Set C are presented. For the C_{2CV} models, 111 samples with total VFA concentrations below 2 g kg^{-1} FM were excluded from the calibration after a cluster analysis. The coefficient of determination was calculated for the calibration (R_c^2) and validation (r_v^2).

	Model	Pretr	Region cm^{-1}	PLS	*RMSEC*	*RMSECV*	*RMSEP*	R_c^2	r_v^2	Bias	Slope	*RPD*
Total VFA [g kg^{-1} FM]	A_{CV}	MSC	6001-5299,	9	0.62	0.72		0.94	0.92	-0.01	0.92	3.6
	A_{TS}		5052-3996		0.62		0.81	0.94	0.90	0.08	0.95	3.2
	B_{CV}	Line	6001-5299,	7	0.72	0.83		0.92	0.89	-0.01	0.90	3.1
	B_{TS}		4505-3996		0.65		0.91	0.94	0.87	-0.06	0.90	2.8
	C_{CV}	MSC	6001-5299,	10	0.73	0.81		0.92	0.90	-0.01	0.91	3.2
	C_{TS}		5052-3996		0.67		0.86	0.93	0.89	-0.04	0.88	3.0
	C_{2CV}	MSC	6001-5299,	9	0.76	0.87		0.92	0.89	-0.02	0.90	3.1
	C_{2TS}		5052-3996		0.76		0.83	0.92	0.91	0.13	0.88	3.3
Acetic acid [g kg^{-1} FM]	A_{CV}	MSC	6703-5299,	9	0.53	0.66		0.78	0.63	0.01	0.69	1.6
	A_{TS}		4505-3996		0.48		0.69	0.81	0.63	0.05	0.65	1.6
	B_{CV}	1st +	6001-5299,	6	0.70	0.76		0.80	0.75	-0.01	0.78	2.0
	B_{TS}	MSC	4505-3996		0.61		0.86	0.85	0.69	-0.07	0.75	1.8
	C_{CV}	MSC	6001-5299,	9	0.72	0.78		0.72	0.66	-0.01	0.69	1.7
	C_{TS}		5052-3996		0.66		0.80	0.78	0.63	-0.02	0.68	1.6
	C_{2CV}	MSC	6001-5299,	9	0.73	0.85		0.75	0.66	-0.01	0.70	1.7
	C_{2TS}		5052-3996		0.69		0.90	0.79	0.62	0.16	0.72	1.6
Propionic acid [g kg^{-1} FM]	A_{CV}	Line	6001-5299,	8	0.64	0.71		0.88	0.84	-0.01	0.85	2.5
	A_{TS}		4505-3996		0.66		0.72	0.87	0.85	0.10	0.91	2.6
	B_{CV}	Line	6001-5299,	6	0.56	0.64		0.77	0.69	0.01	0.71	1.8
	B_{TS}		4505-3996		0.56		0.57	0.78	0.76	0.10	0.69	2.1
	C_{CV}	Line	6001-5646,	6	0.68	0.71		0.81	0.79	-0.01	0.80	2.2
	C_{TS}		5052-4520		0.66		0.72	0.82	0.80	-0.01	0.85	2.2
	C_{2CV}	Line	6001-5646,	7	0.74	0.78		0.82	0.79	-0.01	0.81	2.2
	C_{2TS}		4505-4250		0.67		0.83	0.85	0.78	0.01	0.76	2.1

For a full list of abbreviations see Table 4.4

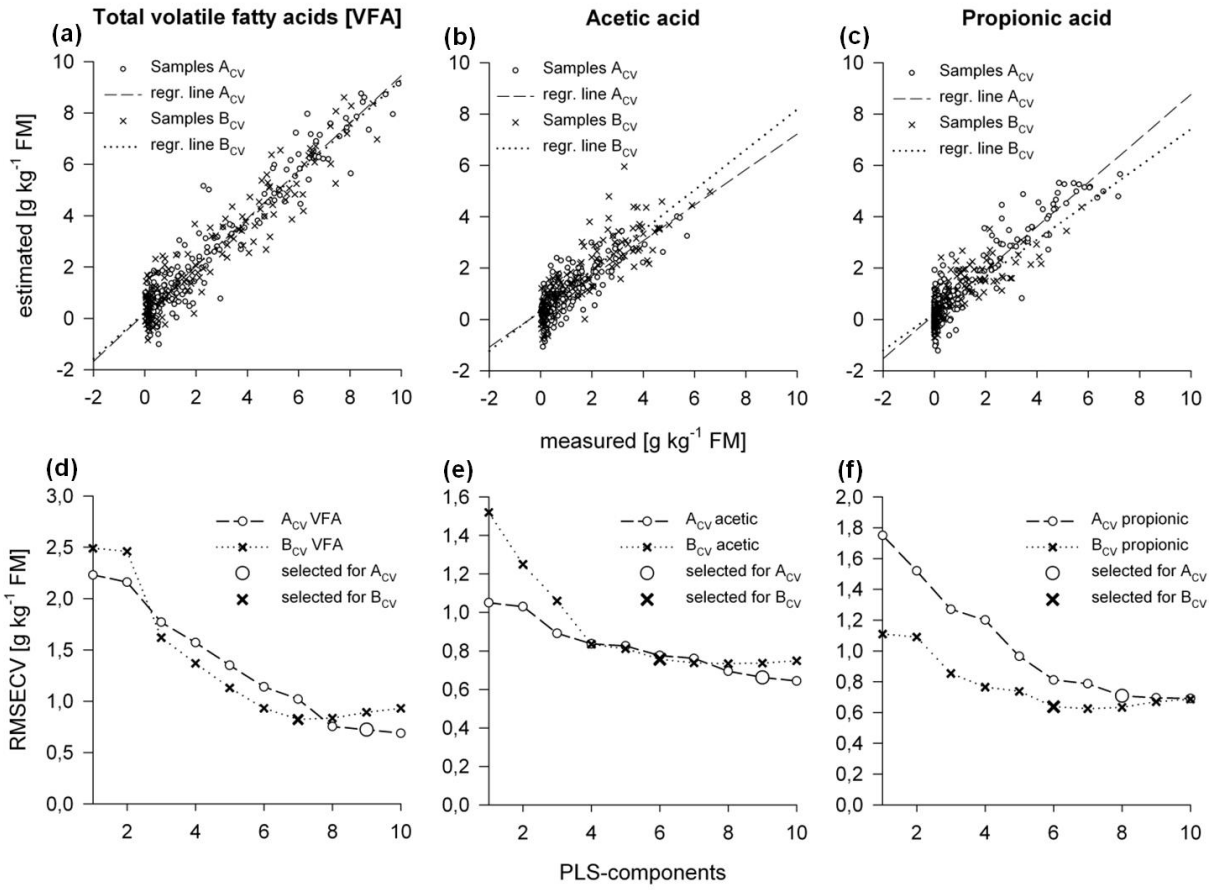

Figure 4.8: NIR estimates of total VFA, acetic acid and propionic acid validated via leave-one-out cross-validation for Subsets A and B. Top: The estimated vs. measured plots related to the A_{CV} and B_{CV} models given in Table 4.5. Bottom: The *RMSECV* plotted as a function of the number of PLS components for the A_{CV} and B_{CV} models, indicating where the estimation error reaches its minimum.

In addition to the validation of the models as illustrated in Figure 4.2, Model A_{CV} was tested on the samples of Subset B. With an *RMSEP* value of 1.1 g kg^{-1} FM, the estimates were less accurate than the results given in Table 4.5. However, with an r^2 of 0.82, a satisfactory X-Y relationship was demonstrated for the validation. This result indicates that for the data presented herein, a calibration whose samples did not include all of the feedstock variation reported for the validation samples may still adequately estimate the concentrations of total VFA. In contrast, samples of Subset A that were estimated via Model B_{CV} reported an *RMSEP* of 2.1 g kg^{-1} FM (r^2 = 0.36) with several low-concentration samples exhibiting large y-residuals, i.e., indicating overestimation (results not shown). One explanation for these contradicting results may be that for Subset B, the quasi ´spiking` with material from acidified lab digesters resulted in collinearity among the analyte concentration and the sample matrix; hence, higher acidity levels are associated with higher amounts of energy crops in the FM of the slurry. Consequently, the PLS algorithm of Model B_{CV} may have failed to accurately estimate the Subset A samples at a low total VFA concentration.

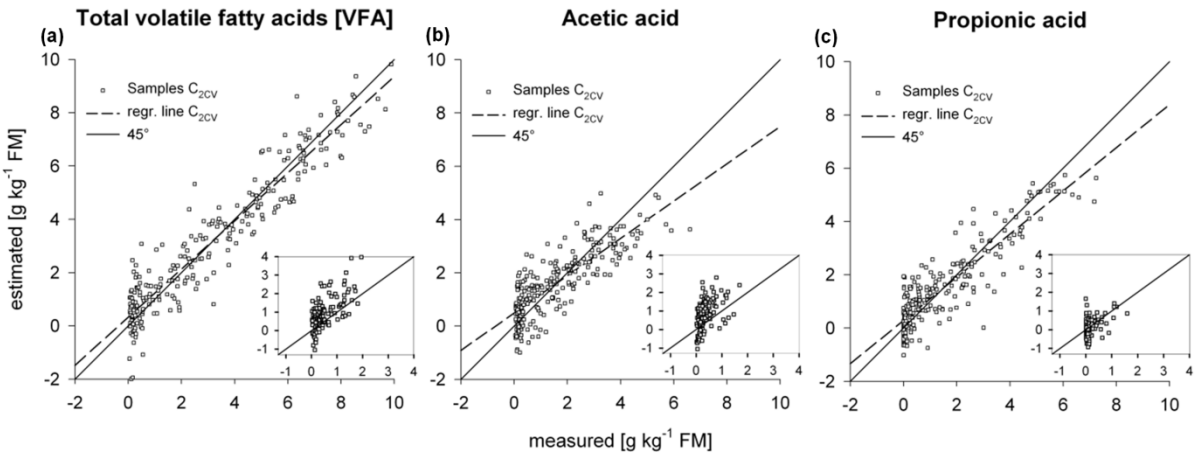

Figure 4.9: NIR estimates of total VFA, acetic acid and propionic acid via leave-one-out cross-validation for Data Set C after a cluster analysis to partly remove low-concentration samples. The estimated vs. measured plots relate to the C_{2CV} models given in Table 4.5. Of the 351 samples of Data Set C, 111 samples with a total VFA concentration below 2 g kg^{-1} FM were excluded from the calibration. These samples were estimated by the C_{2CV} models (results shown in the box, same x-and y-axis) reporting an *RMSEP* of 0.82 (VFA), 0.83 (acetic acid) and 0.47 g^{-1} kg FM (propionic acid).

Acetic and propionic acids

For the model development of the single acids, the correlation between the acids and total VFA should be as small as possible to avoid using spectral information related to changes in the overall VFA concentration. Using this spectral information may lead to erroneous estimates because during AD, the single acid composition may vary depending on the process conditions. Due to the exclusion of many dilute samples from Data Set C after the cluster analysis, the correlation coefficient between total VFA and acetic acid decreased slightly to r = 0.80 and r = 0.81, respectively, for propionic acid (compare with Tab. 4.3). The validation of the models for acetic acid resulted in a similar error of the estimates compared to the total VFA results (Tab. 4.5). As a result of the overall lower range for this single acid, the r^2 and *RPD* values indicated a performance loss. With a slope of approximately 0.7 (Model A_{CV}), the estimates reported a larger systematic over- and underestimation. The *RMSEP* value of 0.83 g kg^{-1} FM indicated that results for the low-concentration samples that were excluded from the calibration by the cluster analysis and estimated by Model C_{2CV} were reported as accurately as those for the test set validation using the overall concentration range. For propionic acid, the models of Subset B indicated the lowest estimation errors for validation (Tab. 4.5). The removal of the many low-concentration samples did not result in a significant change in the estimates because Models C_{CV} and C_{2CV} performed similarly. The interpretation of the satisfactory r^2 value for propionic acid that was reported for the A and C models indicated the need to consider higher overall concentrations up to 7.2 g kg^{-1} FM of samples from lab digesters, which also affects the *RPD*. For this single acid, a concentration as low as 1.5 g kg^{-1} FM may reflect process imbalances (Barredo and Evison, 1991). With an *RMSEP* of 0.47 g kg^{-1} FM reported for the validation of Model C_{2CV} for the samples that were excluded via clustering (Fig. 4.9c inset), the accuracy was observed to be higher than that reported for the overall concentration range. This finding also indicates that, of the two single acids tested, propionic acid allows for a more accurate estimation at low acid concentrations, i.e., an early stage of process imbalance.

4.2. Model validation in the lab

Abstract

The external validation of the NIR models whose development was described in chapter 4.1 is reported. The results, which were obtained using digester slurry from 26 agricultural biogas plants that were fed with various types of energy crops and livestock residues, support the use of NIR spectroscopy for the rapid screening of AD process parameters under open population conditions. With a RMSEP of 5 g kg^{-1} FM for the VS and of 0.2 g kg^{-1} FM for NH_4-N, both parameters were soundly estimated. Total VFA, with a RMSEP of 0.9 g kg^{-1} FM, confirmed the use of NIR analysis for the rapid detection of the accumulation of the total acidity in the slurry. For the tested samples, the alkalinity was found to be a function of the concentrations of NH_4-N and total VFA and, hence, can be as well estimated with NIR spectroscopy with a precision of 1 g kg^{-1} FM. By applying the models to a series of samples that were spiked with propionic acid, the specific estimation of this single acid was demonstrated. Overall, these validation results suggest that a limited number of different feedstock material used for calibration may already reflect those spectral variability necessary for the development of a feedstock-robust application. For VS and NH_4-N, the high repeatability of the NIR measurements ensured that the error in the estimates could be kept small. Repeatability of the NIR estimates for the total VFA was low compared with the reference analysis and may indicate the limitations for specific determination of this analyte by NIR spectroscopy in heterogeneous slurry. The spiking experiments demonstrated that NIR analysis not only supports the screening of the total VFA level in the slurry but, to a limited degree, also allows for the assessment of the composition of the VFA. Whether this supports a robust indication of a trend shift at a low propionic acid concentration, as it is needed to detect an early stage of AD process imbalance, requires further evaluation of the models online.

4.2.1. Material and methods

4.2.1.1. *Collection of sample material for model validation*

The calibration of the six parameters was performed with three training sets, as described in chapter 4.1.1.1. The first set (Subset A, N = 197) was composed of samples from lab digesters that were fed solely with energy crops, particularly, silage of maize, grass-clover mixture and green rye. Subset B (N = 154) mainly contained samples taken from 12 agricultural biogas plants with a more diverse feedstock composition also including manure from cattle, swine and poultry. The third data set used for calibration, Data Set C, was composed of all of the Subsets A and B samples.

The samples used for validation were collected from another 26 biogas plants. Their digesters were operated under the co-fermentation of various energy crops, primarily silage from maize, grass, legumes and green rye with varying proportions of manure from cattle, swine and poultry. The samples were removed from the first and the second digesters (meso- and thermophil operation) in order to consider different AD process stages. The sample handling, including the bottling of the material and the temporary storage at mesophilic conditions prior to the NIR

measurements, was performed as described for the calibration samples (chpt. 4.1.1.1). Sampling was performed during the regular operation of the biogas plants and was repeated up to three times for each digester. A low-acid concentration was reported for most of the samples, reflecting stabile AD processes. In order to assess the model's ability to estimate slurry at elevated acid concentrations, material from different plants was also mixed for the creation of secondary samples. For this purpose, acidified material that was removed from a hydrolytic stage reactor (cow manure together with maize silage) of one biogas plant was combined with the slurry from four digesters running at stabile AD conditions in proportions ranging from 33% to 66% by volume. Of the 80 samples used for validation, 10% of the samples were prepared according to this procedure.

4.2.1.2. *Spectra recordings and reference analysis*

Spectra recordings were performed as previously reported chapter 4.1.1.2. The sample material was scanned offline in the lab using the Vector 22/N FT-NIR spectrometer that was equipped with an integrating sphere. Log (1/R) spectra obtained from diffuse reflection were recorded by a PbS detector in the range of 12000 to 3700 cm^{-1} (16 cm^{-1} resolution, 138 scans). During the spectral measurements, the temperature of the pre-heated samples (water basin ≈ 38°C) randomly varied between 35 and 40°C. The validation samples were scanned in triplicate and subsequently frozen at -18°C before the wet chemistry analysis was performed. For the reference analysis, the samples were analysed in duplicate, and their average value was used for further calculations. The *SEL* for wet chemical methods was calculated separately for this sample set.

4.2.1.3. *PLS calibration and performance testing*

Model development by PLS regression was performed using an optimisation routine that is implemented in the software (Opus 6.5, Bruker Optics) for the testing of various wave-number regions and spectral pre-treatments (chpt. 4.1.1.3). The PLS models for the six parameters were validated by a leave-one-out cross-validation for the selection of the best optimisation algorithm. The calibrations were termed Model A, B and C for Subset A, Subset B or the complete Data Set C, respectively (subscript ´cv` of the model names were removed for simplification; Model C_{2CV} was not considered further). For the calibration of the TIC, the spectral range that was taken for calibration was restricted to the 5400–3900 cm^{-1} region because this range has been reported to potentially contain information that relates to the carbonate ion.

These models were tested on the 80 samples which were additionally collected for their independent validation. The triplicate NIR measurements of each sample were averaged prior to validation. The accuracy of the calibrations was assessed by the *RMSEP*. For external validation of the models, the number of PLS-components was chosen as determined by earlier cross-validation. However, when the *RMSEP* reported a minimum at a lower model complexity, the number of components was reduced accordingly. Whether differences existed in the accuracy of the estimates among the models of the three calibration sets was tested by non-parametric Friedman test for paired samples that calculated the difference (residual) between the estimate and the reference of each sample (Miller and Miller, 2005). The ranking was judged by Wilcoxon-Wilcox test applied for multiple comparisons of the medians (Köhler et al., 2007). The precision of a model was described by the *SEP*. Moreover, the *RPD* was calculated. For the

assessment of the repeatability of the NIR estimates (Dardenne, 2010), while accounting for errors due to the inhomogeneity of the samples (repack error) and errors due to the instrument (Esbensen, 2009), the models were tested on the 240 single recordings for calculation of the *SED* (see chpt. 3.2.).

4.2.1.4. Spiking experiments

It was investigated whether the models allowed for a distinct estimation of acetic acid and propionic acid independent of any collinear behaviour with the total VFA. The digestate of a commercial biogas plant (VS of 63 g kg^{-1} FM) was serially spiked with acetic acid (1 to 10 g kg^{-1} FM, 1 g increments, 1st series) and propionic acid (1 to 10 g kg^{-1} FM, 1 g increments, 2nd series). The same spiking was performed with the slurry of a second plant (VS of 91 g kg^{-1} FM) at increased natural propionic acid concentrations because this material was removed from a disturbed AD process, resulting in the 3rd and 4th dilution series. The 42 spiked samples (including two blank tests) were subsequently used to estimate the concentrations of total VFA, acetic acid and propionic acid. Prior to the spiking experiments, technical service to the spectrometer had resulted in a shift in the instrument response. In order to include this new spectral variation to the calibrations (Bouveresse and Campbell, 2008), the models were not directly applied for the estimation of the spiked samples but were first recalculated using an additional set of samples (n = 63). This set was scanned under these new optical conditions. Therefore, additional fermentation runs with two lab digesters were performed for the collection of these samples. The first digester was solely fed with maize silage and feeding at temporarily high OLR was applied in order to generate different process conditions i.e. to induce variations in the concentration of the parameters. The second digester was inoculated with material from a large scale biogas plant treating a mixture of cattle manure together with energy crops (silage of maize and grass). PLS regressions were performed with these new samples in conjunction with samples of the Data Set C. The models again were generated using the *OPUS* software optimisation routine and were validated by the leave-one-out cross-validation. For each parameter, five models were manually selected to represent different pre-treatments and variable wave-number selections and were tested on the spiked samples. The models with the smallest *RMSEP* are presented herein.

4.2.2. Results and discussion

4.2.2.1. Validation data set characteristics

The descriptive statistics of the process parameters of the validation set (Tab. 4.6) indicated a similar outcome as that observed for the three calibration sets (Tab. 4.2) From the comparison of the means and medians, it was concluded that the data of VS, NH_4-N and TIC were almost normally distributed. The data of the three acidity parameters indicated a strong skew towards the low concentration values, confirming the fact that most biogas plants were operating at stabile AD conditions. The parameter VS was only weakly correlated with the other parameters, whereas NH_4-N and TIC were strongly correlated (Tab. 4.7). The TIC was negatively correlated with the three acidity parameters which were strongly positive correlated to each other.

Table 4.6: Descriptive statistics for VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid of the validation data set (N = 80). The arithmetic mean, median, standard deviation (*SD*) and the *SEL* of the reference analyses are presented.

Parameter [g kg^{-1} FM]	Range	Mean	Median	*SD*	*SEL*
VS	37.6–97.8	69.2	67.5	13.7	2.1
NH_4-N	1.02–5.10	2.19	2.06	0.66	0.06
TIC	3.96–23.19	10.68	10.22	3.22	0.19
Total VFA	<0.13–10.68	2.06	0.61	2.94	0.19
Acetic acid	<0.10–5.86	1.38	0.56	1.69	0.14
Propionic acid	<0.03–3.67	0.47	0.05	0.83	0.06

Table 4.7: Coefficients of correlation (r) among the parameters calculated for the external validation set. The r was tested for significance (1% = **).

Parameter	VS	NH_4-N	TIC	Total VFA	Acetic acid
NH_4-N	0.12				
TIC	0.12	0.84**			
Total VFA	0.32**	-0.08	-0.54**		
Acetic acid	0.33**	-0.10	-0.54**	0.98**	
Propionic acid	0.33**	-0.03	-0.49**	0.96**	0.90**

4.2.2.2. *PLS model validation results*

Volatile solids (VS)

Of the three models validated, Model B indicated the highest accuracy (*RMSEP*) though it was based on the smallest of the three calibration sets (Tab. 4.8, Fig. 4.10a). With five PLS components, the model complexity for cross-validation was confirmed as optimal (see Tab. 4.4). With a *RMSEP* close to 5.0 g kg^{-1} FM, the external validation indicated some loss in the estimation performance compared to the *RMSECV* of approximately 4.0 g kg^{-1} FM for internal validation. Compared with the other two models, it reported the largest share of outliers as computed by Mahalanobis distance (*MD*), indicating that the spectral distance did not necessarily affect the accuracy of the estimates. When Friedman test was applied, no significant differences in the accuracy among the three models could be shown on the ordinal scale (Tab. 4.9). From this it has to be summarized, that the ability to detect VS was not depending on the variability of the calibration samples. This can be expected because NIR spectroscopy frequently excellently performs in the detection of water or total organic matter within different matrices (Roberts et al., 2004), which here most likely determined the model quality. For Model A, the lowest *RMSEP* was obtained with three PLS components, whereas cross-validation reported six factors as optimal (Tab. 4.4). This indicated that the spectral information located at higher factors described a specific X-Y relationship inherent in Subset A rather than reflecting changes in the VS concentration of a broad-based population. With a *SEP* of around 6 g kg^{-1} FM, Model C showed the least precise estimates, indicating that the combination of both subsets did not result in any improvement of the model's estimates. Shown by the *SED* value (Tab. 4.8), repeatability of the estimates for the VS-models was similar to those of the reference analysis as given by the *SEL* value (Tab. 4.6). However, variation of the *SED* value indicated that the specificity of the spectral information used for calibration of this parameter tended to differ among the three models.

Ammonium (NH_4-N)

The external validation of NH_4-N estimates revealed significant differences (Tab. 4.9) in the accuracy among the models even though, all models had a similar prediction error with *RMSEP* values ranging between 0.20 g kg^{-1} (Model C) and 0.24 g kg^{-1} FM (Model B). Only the slope and offset slightly differentiated between Model A and C compared to Model B and may help to understand these significant differences. Moderate increase of the error estimates was observed by comparison of the model results with those from cross-validation. Model A (*RMSEP*: 0.22 g kg^{-1} FM) indicated that, for NH_4-N, the use of a model based solely on energy crops was sufficient to estimate NH_4-N in digester slurry having considerably larger feedstock variation. This accuracy may be because the spectral region that was relevant for the estimation of NH_4-N was hardly affected by the new feedstock variability of the validation samples including manure and only few samples were indicated as outliers (Tab. 4.8). The small spectral range used in the calibration by the software's optimisation routine was believed to have improved the robustness of the calibrations because irrelevant parts, which could introduce noise into the models, were excluded (chpt. 4.1.1.3). With an *SED* of 0.07 g kg^{-1} FM (Model C) high repeatability was demonstrated for the NIR replicates which was close to the *SEL* of the reference method. This repeatability was similar among the three models supporting the assumption of a high specificity of the NH_4-N-signal in the spectra.

Table 4.8: Statistics of the external validation of the PLS models for the estimation of the six parameters. Results for the Models A (Subset A, N = 197), B (Subset B, N = 154) and C (Data Set C, N = 351) are shown. *Data from Table 4.4 and 4.5; number of PLS components may vary.

Parameter [g kg^{-1} FM]	Model	PLS	*RMSECV**	*RMSEP*	Bias	*SEP*	Offset	Slope	*RPD*	*SED*	MD (%)
VS	A	3	6.11	6.14	-3.04	5.37	16.59	0.80	2.5	1.5	7
	B	5	3.87	4.94	0.55	4.94	0.58	0.98	2.7	2.1	17
	C	7	4.14	6.03	-0.99	5.98	-2.15	1.04	2.2	2.8	5
NH_4-N	A	4	0.16	0.22	-0.14	0.22	0.16	0.93	3.0	0.08	5
	B	3	0.17	0.24	0.01	0.24	0.30	0.86	2.8	0.09	5
	C	4	0.16	0.20	0.02	0.20	0.11	0.94	3.3	0.07	4
TIC	A	7	0.80	1.68	1.26	1.13	1.51	0.74	2.8	0.37	28
	B	6	0.74	1.42	0.55	1.32	2.03	0.76	2.4	0.39	12
	C	8	0.74	1.09	0.40	1.02	1.61	0.81	3.1	0.41	9
Total VFA	A	8	0.76	0.91	0.36	0.84	0.06	0.79	3.4	0.73	8
	B	6	0.94	1.16	-0.32	1.12	0.56	0.88	2.6	0.57	11
	C	9	0.85	0.87	-0.04	0.87	0.48	0.79	3.3	0.73	5
Acetic acid	A	8	0.70	0.92	0.18	0.90	0.44	0.54	1.8	0.44	10
	B	6	0.76	0.72	-014	0.70	0.43	0.79	2.4	0.44	9
	C	9	0.78	0.74	0.18	0.72	0.24	0.70	2.3	0.47	5
Propionic acid	A	8	0.71	0.53	-0.22	0.49	0.14	1.16	1.7	0.51	7
	B	6	0.64	0.57	-0.10	0.57	0.14	0.93	1.4	0.40	12
	C	5	0.73	0.47	-0.06	0.47	-0.08	1.32	1.7	0.47	3

Abbreviations: PLS, number of PLS components; *RMSE*, root mean square error reported for cross-validation (*CV*) and external validation (*P*); *SEP*, standard error of precision; *RPD*, ratio of the *SD* of the reference values to the *SEP*; *SED*, standard error of differences, calculated from the replicates of the NIR measurements; *MD*, Mahalanobis distance: comparison of the external validation set to the mean spectrum of the calibration set; if their standardized *MD* was beyond the cut-off value of 3, they were marked as outliers. The data indicate the proportion of outliers within the validation data set.

Table 4.9: Results of testing the absolute difference between the NIR estimate and the reference value of each sample for significance. The non-parametric Friedman test for paired samples was used to check for significant differences between Models A, B and C. The ranking was judged by Wilcoxon-Wilcox test for multiple comparisons.

Parameter	Friedman (α = 5%)	Wilcoxon (α = 1%)
VS	ns	-
NH_4-N	s	A ≠ B; A ≠ C; B ≠ C
TIC	s	A ≠ B; A ≠ C; B ≠ C
Total VFA	s	A ≠ B; B ≠ C
Acetic acid	ns	-
Propionic acid	ns	-

Abbreviations: ns, not significant; s, significant

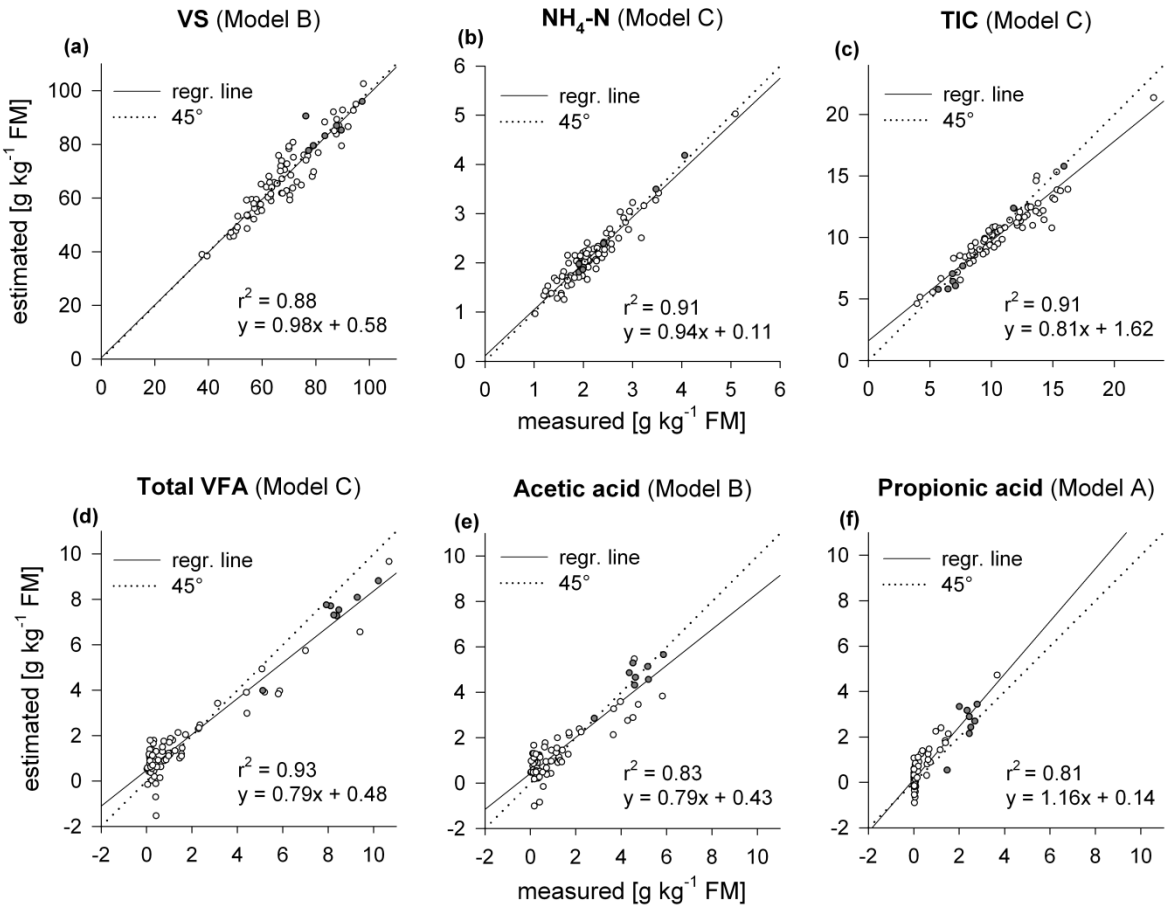

Figure 4.10: External validation results of the NIR estimates for VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid shown for selected models of Table 4.8. The filled symbols indicate samples prepared by mixing slurry with material from a hydrolytic stage.

Total inorganic carbon (TIC)

The precision of the three models, as described by the *SEP* (Tab. 4.8), ranged from 1.02 g kg^{-1} (Model C) to 1.32 g kg^{-1} FM (Model B). For Model A, larger bias was observed. Most likely this was due to the limited ability of Subset A to describe the spectral variation of the validation set as was indicated by the large number of samples judged as outliers (28%). Compared with the models for VS and NH_4-N using Subset A, this demonstrated that the number of outliers was not only dependent on the calibration set themselves but was also affected by the selection of the spectral regions used for calibration of each parameter. As illustrated in Figure 4.10c, accurate

estimation outside of the calibration's concentration range (14 g kg^{-1} FM, see Tab. 4.2) was demonstrated for Model C; the single sample at a high analyte concentration appeared along the regression line. However, more samples at higher TIC levels are needed to confirm the linearity of the model. Overall, these results confirmed the potential of the NIR models for screening the alkalinity of digester slurry, although a loss in performance was reported compared to cross-validation. A study by Ward et al. (2011a), reported similar accuracy obtained from an internal validation. The titrimetric analysis for determination of total alkalinity differed and this limits the direct comparison of the results. For the study presented herein, TIC was analysed by titration with H_2SO_4 to pH 5.0. For discussion about the effect of different titrimetric endpoint pH on the quantification of the alkalinity in digester slurry also see Björnsson et al. (2001) and Hecht et al. (2009).

The effect of the NH_4/NH_3 buffer on alkalinity due to the formation of NH_4HCO_3 at an increased pH was reflected by the positive correlation between the TIC and NH_4-N (Tab. 4.7). The negative correlation between TIC and total VFA corresponded to the fact that the bicarbonate ion is released as carbon dioxide if the VFA concentration increases (Jantsch and Mattiasson, 2004). Whether the satisfying NIR estimates for TIC were due to this collinearity was tested by MLR (Næs et al., 2002). The reference values of the samples of Data Set C (at pH <8.1) were utilized for the calibration of a MLR model that was subsequently tested on the 80 external samples. The estimated vs. measured plot for the MLR model shown in Figure 4.11 indicated that the TIC was described by these two parameters (VFA and NH_4-N) with a *RMSEP* of 1.21 g kg^{-1} FM. As well as NH_4-N, the concentration of total VFA could also be estimated with NIR; therefore, the satisfactory decision to use NIR technology to determine TIC may be rationalised, as its concentration was primarily a function of both analytes. The spectral regions used for the estimation of the TIC indicated an overlap with those regions selected for the NH_4-N and total VFA models, supporting this assumption (Tab. 4.4 and 4.5)

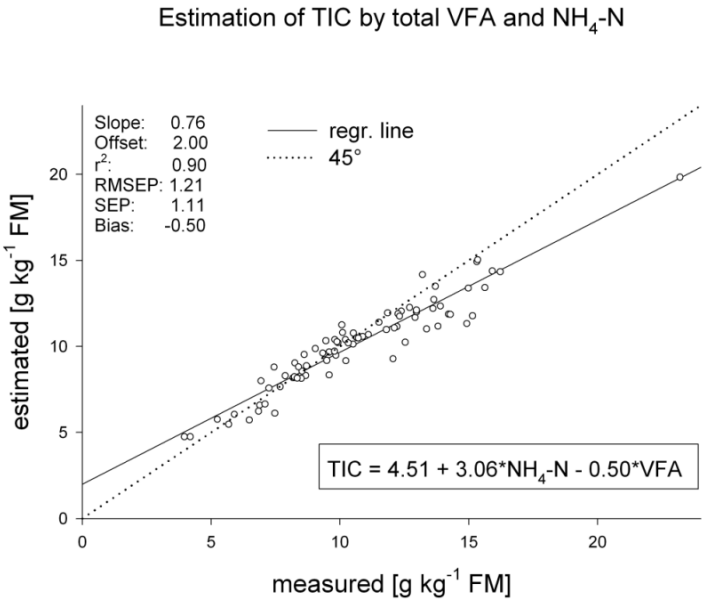

Figure 4.11: Reference (measured) value of TIC compared with the estimates of a MLR with total VFA and NH_4-N as independent variables. Results are shown for the external validation samples. The MLR model was developed with samples of Data Set C. See text for further details.

Volatile fatty acids (VFA)

The external validation of the three total VFA-models reported a *RMSEP* of 0.87 g kg^{-1} (Model C) to 1.16 g kg^{-1} FM (Model B), largely confirming their accuracy, as reported for cross-validation. Figure 4.10d shows the estimated vs. measured plot for total VFA using Model C, separately indicating samples that were produced by mixing. With a slope of the regression line close to 0.8, the model overestimated samples at lower values than average and vice versa. The similar results of Model A and Model C demonstrated that an inclusion of feedstock variation into the model, which was believed to better reflect a broad-based population, did not result in a significant increase in the accuracy of the estimates (Tab. 4.8). Compared to literature values, these error estimates for total VFA were similar to those found for feedstock-specific (Jacobi et al, 2009) or digester-specific (Holm-Nielsen and Esbensen, 2011) calibrations. This finding supports the assumption that an increase in the heterogeneity of the slurry matrices did not necessarily imply a relevant decrease in the accuracy of the estimates. Although this may have led to an increase in the physical inter-sample variation, the analyte's and the interferent's X responses still appeared to be sufficiently different (Martens and Næs, 1991). However, comparison of the *SED* (NIR) with the *SEL* (reference) revealed the comparably low spectral repeatability contributing to the overall error of the NIR model. Presumably, this was due to random noise in the spectra of the validation samples (X_{test}) and this noise may indicate 'an intrinsic limitation in the analytical method' (Martens and Næs, 1991). Although, the estimation error of the regression coefficients and, hence, the uncertainty of y was assumed to be rather small due to the high number of calibration samples (Model C with n = 351), this low repeatability restricted the use of the total VFA models to screening purposes. This was also indicated by an *RPD* value of 3.3 (Tab. 4.8).

The validation results for acetic acid are shown for Model B (Fig. 10e). Similar to total VFA, high-acetic acid samples reported an underestimation which presumably illustrates the least-squares effect of the inverse regression i.e. the shrinking of the estimations towards the mean (Næs et al., 2002). This may also indicate that random noise was included in the calibration (X_{cal}) (Martens and Næs, 1991). The *RPD* value generally decreased, owing to the smaller standard deviation of the reference values of acetic acid that was observed for the validation samples (Tab. 4.6). With a *RMSEP* below 0.6 g kg^{-1} FM, absolute accuracy for the propionic acid models was higher than that for the other two acid parameters. The relative prediction ability was worse because of the narrower range of propionic acid up to 3.6 g kg^{-1} FM. This was also expressed by the poor *RPD* values below 2 (Tab. 4.8). Noteworthy, the *RMSEP* revealed a generally lower error than the *RMSECV* reported for internal validation of the models. The calibration data sets were skewed towards the lower concentration samples and, hence, the calibration was weighted to most closely fit those samples. The few calibration samples at higher propionic acid concentrations (up to 7.3 g kg^{-1} FM, see Tab. 4.2) showed a poorer approximation resulting in higher overall error estimates. For external validation, practically all samples at higher propionic acid concentration were produced by using material of the hydrolytic stage of one plant for mixing. Therefore, the performance indicators of the models (for example the slope) strongly reflected the ability of the calibrations to estimate slurry from this particular source and the conclusions about a broad-based use of the models may be limited. However, it may also be argued that the mixing simulated changes in the digester content over time for example by variation of the feedstock, and the estimates were not affected by these

changes. This was concluded from visual inspection of the estimates of Model A (Fig. 10f) where there was no indication of a disturbance due to variation of the sample matrix. This again would confirm the robustness of the models. For the three calibration sets, acetic acid and propionic acid contributed most to the total VFA concentration and, hence, the data sets showed a high degree of collinearity of each single acid to the total VFA concentration (Tab. 4.3). As a consequence, the single-acid NIR models were possibly closer related to spectral information describing the total VFA concentration rather than those of a single acid. Thus, the possibility for an independent NIR calibration of the single acids was maybe limited and for realistic assessment of the model performance a validation data set without such parallel collinearity of the single acid to total VFA is needed.

Can the NIR models distinguish between acetic acid and propionic acid?

The results of the spiking experiments are listed in Table 4.10; a graphical illustration is provided as estimated vs. measured plots (Fig. 4.12a–4.14a). For each model, the four dilution series are also separately displayed in the estimated vs. sample plots (Fig. 4.12b–4.14b). For total VFA, the samples spiked with acetic and propionic acid indicated an increase in their concentration corresponding to the amount of acid added. The elevated total VFA concentration of one of the blank sample (0**) was closely fitted to the plot (Fig. 4.12a). This higher initial acid level of the slurry used for the 3rd and 4th dilution series resulted in the higher final concentrations (A10** and P10**) reflected by the model's estimates as well. For the acetic acid model, the estimates of the samples spiked with acetic acid also correctly indicated a concentration increase though, underestimation of the samples at higher acetic acid values was observed (Fig. 4.13a). A clear overestimation of the samples spiked with propionic acid at an acetic acid concentration of 0 was apparent. This was supported by Fig. 4.13b, in which increases in acetic acid estimates corresponded to the order of their propionic acid dosage in the samples spiked with propionic acid. Thus, the estimates for acetic acid were not independent from changes in the propionic acid concentration and, hence, the model was of limited use. In contrast, the propionic acid model correctly indicated an increase in the samples spiked with propionic acid and showed no systematic increase in the estimates of the samples spiked with acetic acid (Fig. 4.14b). For acetic acid, the coefficient of correlation (r) with total VFA was as high as $r = 0.92$ for the calibration set. Therefore, the NIR model for acetic acid was possibly closer related to the model for total VFA. On the other hand, collinearity among propionic acid and total VFA was less pronounced ($r = 0.78$). The narrow spectral range (6000-5646 cm^{-1}) used for the propionic acid model, corresponds to the CH_2 and CH_3 group peaks of the acid (Workman and Weyer, 2008) and this further supported the assumed higher selectivity of the propionic acid model. The methyl group of acetic acid also indicates peaks in the first overtone region (2v) and it is concluded that for the propionic acid model, these chemical interferences were accounted for by PLS regression. Whether the reason for the insufficient NIR model for acetic acid foremost was mainly due to this higher collinearity to total VFA or, if also, it can be explained by an overall lower NIR-selectivity of acetic acid in digester slurry, are not yet clear. For NIR analyses, samples spiked with an analyte do not necessarily correspond spectrally to samples in their natural state (Esbensen, 2009). However, because spiking was performed on the samples used for validation while the calibrations were based on samples with naturally formed acids, the results support the selectivity of the models for total VFA and propionic acid.

Table 4.10: Statistics of the spiking experiments. Acetic acid and propionic acid were added to 42 samples (including two blank tests) followed by NIR measurement. Validation was performed with models after recalibration using samples of Data Set C together with new 63 samples accounting for changes in the optics of the instrument.

Parameter [g kg^{-1} FM]	Pretr	Region cm^{-1}	PLS	*RMSEC*	*RMSECV*	*RMSEP*	Bias	*SEP*	Offset	Slope	*RPD*
Total VFA	1st + Line	6703-5299, 5052-3996	9	0.99	1.08	1.19	-0.31	1.16	0.33	0.99	2.9
Acetic acid	Line	6001-5299	8	0.96	1.02	1.94	-0.29	1.94	1.33	0.66	1.9
Propionic acid	Raw	6001-5646	8	0.79	0.82	1.54	0.34	1.52	0.94	0.72	2.7

Abbreviations: Pretr, spectral pre-treatment; 1st, first derivative; Line, subtraction of a straight line, Raw, raw spectra from log(1/R); *RMSEC*, root mean square error of calibration. For a full list of abbreviations see Table 4.8.

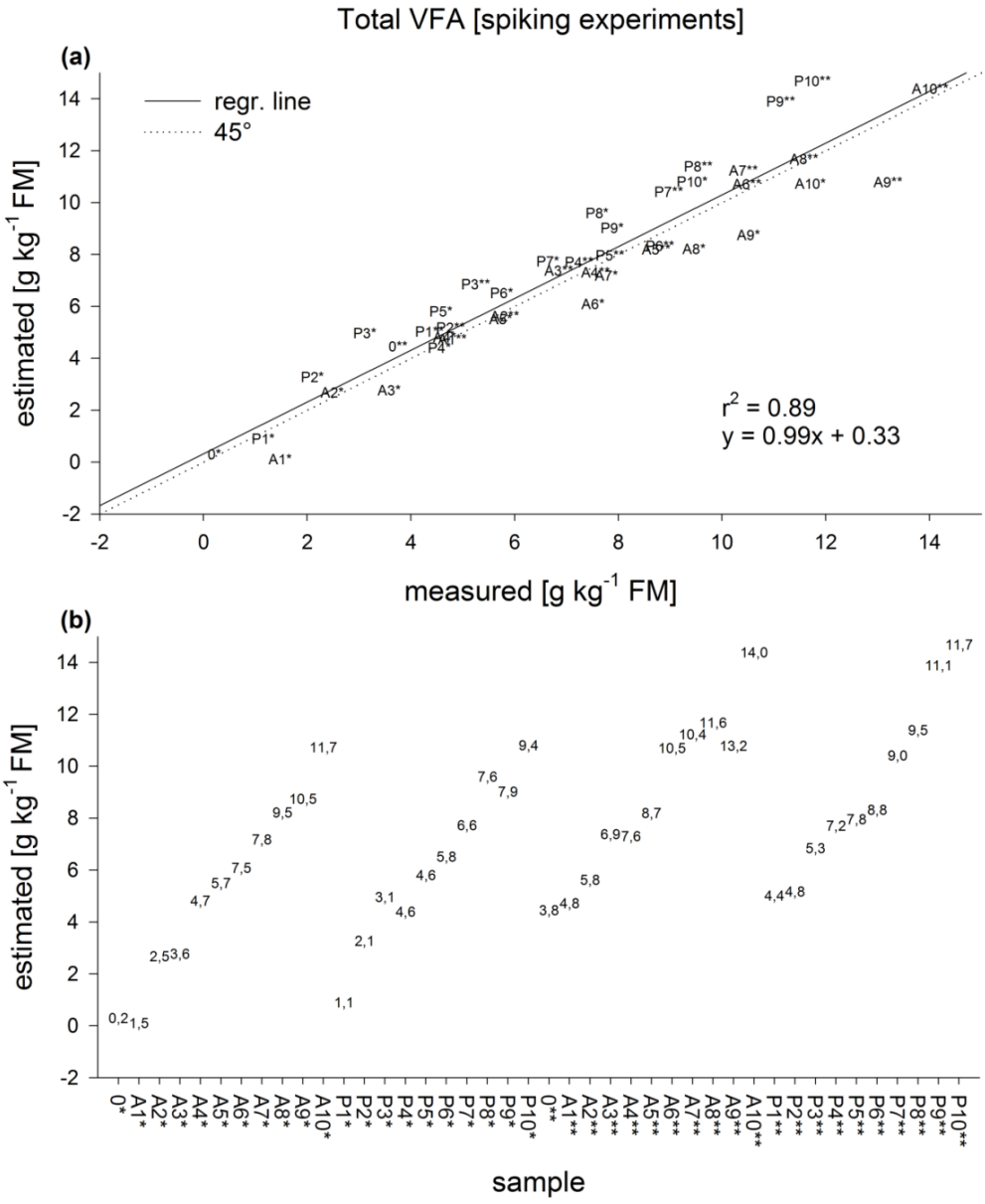

Figure 4.12: NIR-estimated concentration of total VFA in the spiked samples (n = 42) compared to the measured concentration (a) and plotted over each sample (b). For each sample, the added acid (A = acetic acid; P = propionic acid) is given plus the dosage (g kg^{-1} FM) at which it was given (value figure). Asterisk (*) = first biogas plant; double asterisk (**) = second biogas plant. The blank samples are indicated by 0* and 0**. For the estimated vs. sample plot (b), sample values indicate the measured concentration of total VFA.

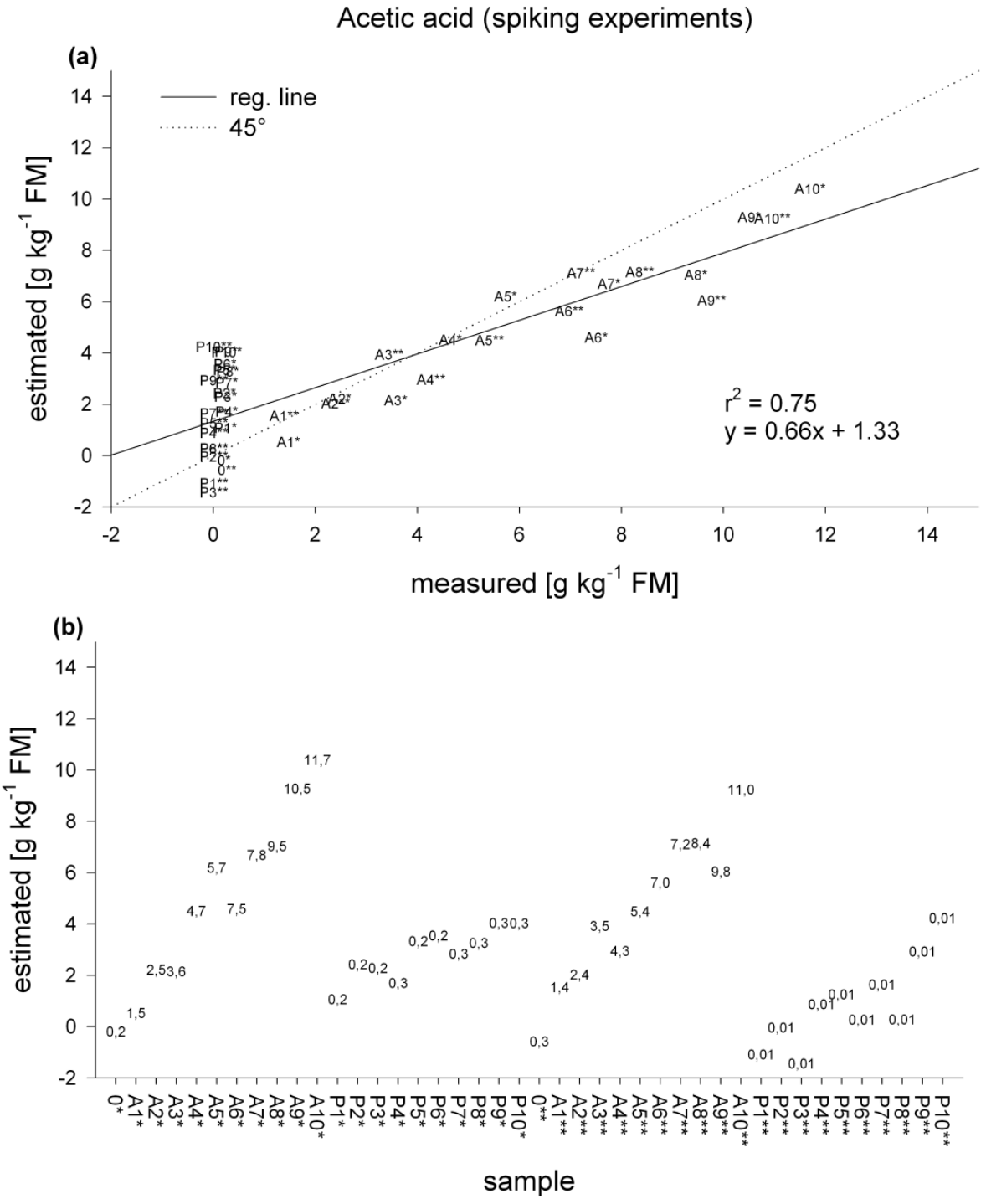

Figure 4.13: NIR-estimated concentration of acetic acid in the spiked samples (n = 42) compared to the measured concentration (a) and plotted over each sample (b). For each sample, the added acid (A = acetic acid; P = propionic acid) is given plus the dosage (g kg^{-1} FM) at which it was given (value figure). Asterisk (*) = first biogas plant; double asterisk (**) = second biogas plant. The blank samples are indicated by 0* and 0**. For the estimated vs. sample plot (b), sample values indicate the measured concentration of acetic acid.

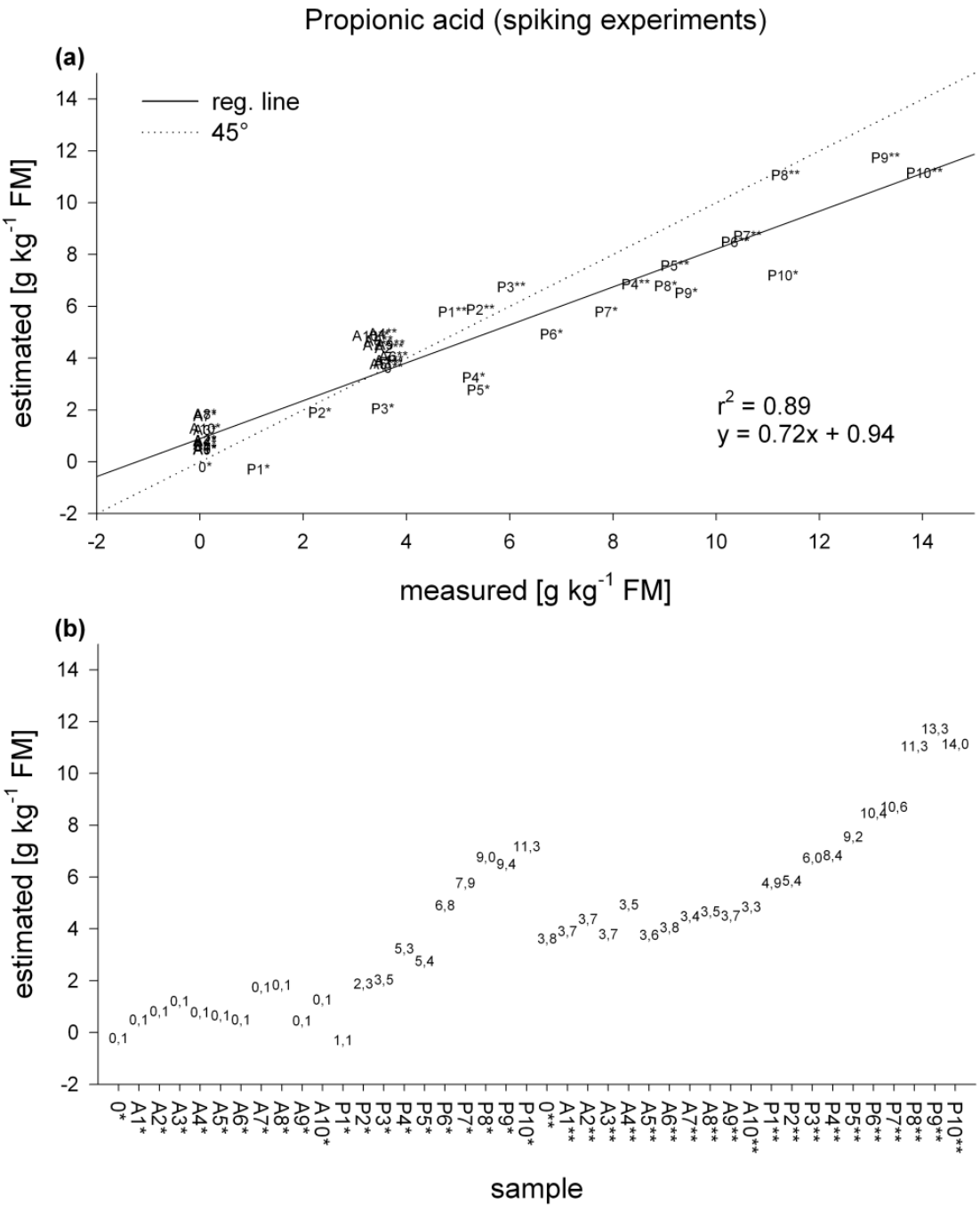

Figure 4.14: NIR-estimated concentration of propionic acid in the spiked samples (n = 42) compared to the measured concentration (a) and plotted over each sample (b). For each sample, the added acid (A = acetic acid; P = propionic acid) is given plus the dosage (g kg^{-1} FM) at which it was given (value figure). Asterisk (*) = first biogas plant; double asterisk (**) = second biogas plant. The blank samples are indicated by 0* and 0**. For the estimated vs. sample plot (b), sample values indicate the measured concentration of propionic acid.

4.3. Spectra transfer for real-time in situ monitoring of anaerobic digestion

Abstract

The operation of an NIR online process analyser for in situ monitoring of anaerobic digestion is reported. The samples measured on the lab instrument, including feedstock from energy crops and livestock residues, were subjected to piecewise direct standardisation for their spectra transfer. This transfer was used in conjunction with samples recorded online for the partial least squares regression of VS, NH_4-N, TIC, and VFA parameters in the fresh matter of a digester slurry. The high potential for in situ monitoring of AD process parameters with varying feedstock mixtures and temperature changes was demonstrated. The validation performed on independent time series spectra confirmed the ability of a robust NIR-based rapid screening of the dynamics of these parameters directly at the digester including precise indication of the total VFA concentration. Thereby, the potential for cross-linking the different spectrometers was also demonstrated. Spectra transfer is an option for expanding the calibration set even when the standard samples are not perfectly representative of the calibration samples that are to be transferred. As a second approach, the benefit of full recalibration of the models with comparably few samples under the new environmental conditions was also demonstrated. These results confirmed the observation that a limited number of feedstock material used for calibration already reflect those spectral variability necessary for the development of feedstock-robust applications. For the practical handling of a wide range of feedstock material that are commonly used in agricultural biogas plants this outcome is crucial since it indicates that it is not necessary to each time perform a digester-specific calibration. No final conclusion can be drawn regarding the advantage of one approach over the other because this would require a more detailed evaluation, such as their long-term stability in different digesters.

4.3.1. Material and methods

4.3.1.1. *Use of samples that were obtained from the lab instrument*

The experiments of the lab application included samples that had been collected from 38 large-scale biogas plants and 14 lab digesters that contained a feedstock composition of energy crops and livestock waste as described in chapters 4.1 and 4.2. The samples (N = 431) were scanned offline with a FT-NIR spectrometer (Vector 22/N, Bruker Optics, Ettlingen, Germany) using diffuse reflection (log(1/R)) in the 12000-3700 cm^{-1} region. The temperatures of the samples were randomly changed from 35-40 °C during the recordings. Not all the samples were used in the present study, so that spectral redundancies in the PLS model could be avoided (Kessler, 2007). Therefore, a cluster analysis (Næs et al., 2002; 30 clusters, 20 iterations, Kendall's Tau) was performed with the lab samples in the 9000-4100 cm^{-1} region of the first derivation spectra (Unscrambler 9.8, Camo, Oslo, Norway). From each cluster, seven samples were selected to evenly span the total VFA range within one cluster. If the number of samples in one cluster was <7, the missing samples were taken from adjacent clusters. This resulted in the selection of 210 samples that were used for further data analysis.

4.3.1.2. *Online spectral recordings*

A standing tank digester stirred at 10-min intervals by a vertically mounted agitator was used for the experiment (Fig. 4.15a). This digester A had a working volume of 2500 L and was operated for 90 days at 35-42 °C using maize silage as feedstock. An OLR measured as VS of up to 9 kg

m^{-3} d^{-1} was maintained for 10 consecutive days, causing an increase in the concentration of total VFA to 20 g kg^{-1} FM. To speed up digester recovery, on day 68, 500 L of the digester content was substituted with material from a biogas plant that was fed with silage from maize and grass; further, 300 L were exchanged with cattle manure. A successive recovery was observed, and the experiment was stopped after stable AD conditions were re-established (total VFA <1 g kg^{-1} FM). The spectral data were measured using an external-loop system (Fig. 4.15b) that was specifically developed to enable online measurements and process sampling (Anamenter, Högemann GmbH, Garrel, Germany). The system was directly connected to the sidewall of the digester. The inlet was located at the bottom of the digester and the outlet halfway up the side of the digester. The slurry was pumped through a metal pipe (50 mm in diameter) for the duration of the experiment to provide continuous loop-circulation. The metal pipe was insulated to maintain the temperature of the slurry. A sapphire window (40 mm in diameter) was fitted to an upstream section of the pipe on which a measuring head (Q-412A, Bruker Optics) was mounted at a distance of 10 cm (measuring spot 10 mm; Fig. 4.15c). The measuring head was connected to an FT-NIR spectrometer (Matrix-F, Bruker Optics) via fibre optics for online recording of the log(1/R) spectra by an InGaAs detector (12400-4000 cm^{-1}). The spectra were averaged for 45 s intervals (144 scans, 16 cm^{-1} resolution) using Opus 6.5 software (Bruker Optics). For sampling (N = 48), the slurry was removed via a ball valve closely situated behind the measuring head during this 45 s interval. The volume of the primary sample was 40 L, from which a 5-L sub-sample was taken before a 0.5-L sample was bottled for wet chemistry analyses.

Digester B (Fig. 4.15d) had a liquid volume of 250 L and was operated for 26 days at 32-48 °C. The inoculum was composed of maize silage, grass and cattle manure. During the experiment, feeding up to an OLR of 10 kg m^{-3} d^{-1} was administered to generate a strong variation in the AD process using silage from maize, grass-clover mixture and wheat straw. The digester content was partly replaced by material from four biogas plants, which were fed with maize silage, clover-grass, green rye and manure from cattle and poultry in different proportions. The same spectrometer was used as for digester A, and the sapphire window was directly integrated into the digester wall. During sampling (N = 49), the primary samples were collected via a ball valve located close to the window. The volume removed during 45 s of sampling was 10 L, from which a 0.5-L reference sample was directly taken.

Digester C (Fig. 4.15e) was used for external validation and had an effective capacity of 150 L. It was stirred by a paddle agitator (140 min^{-1}) during the 26 days of fermentation. An inoculum of a biogas plant fed with maize silage, grass, and poultry manure was taken, and the temperature ranged from 36-46 °C during the operation. Discontinuous feeding with VS at 3 kg m^{-3} d^{-1} was established at the beginning (day 4) and was increased to 6 kg m^{-3} d^{-1} (days 21 and 22). The digester slurry was partly substituted with material from two commercial biogas plants (energy crops and livestock waste) on day 11 (90 L) and day 19 (120 L). The silage and the slurry used for digester C were of a different origin than the material that was used for the lab application and for operating digesters A and B. The measuring head was flanged to a window that was mounted on the digester wall. Spectral recordings were continuously performed and the spectra were averaged over 45-s intervals. After a pause of 135 s the scanning resumed, which resulted in a 3-min measurement cycle. The primary sample that was removed during the 45 s of sampling (N = 42) was 10 L, from which a 0.5-L reference sample was taken.

Figure 4.15: NIR in situ monitoring at three technical scale digesters. (a): Standing tank digester A; (b): ´Anamenter` connected to digester A for providing continuous loop-circulation for spectra acquisition and sampling; (c): Sapphire window and measuring head for spectra acquisition at digester A; (d): Digester B with sapphire window installed directly in the outer wall with sampling valve; (e): NIR measuring head flanged at the sapphire window of digester C which was used for external validation.

4.3.1.3. *Spectral transfer by piecewise direct standardisation*

Triplicate scans were immediately taken offline (at the line) with the material from the 0.5-L reference bottle that was obtained from online-sampling for generating standard samples. The offline spectra were averaged to compute the PDS model. The temperature of the sample that was scanned offline was on average 3 °C lower than that for the online measurement of the respective sample. Technical servicing of the lab spectrometer resulted in a shift in the instrument response on day 20 of the operation of digester A. Therefore, the samples available for calculating the transfer function (the PDS-model) were restricted to those that had been measured on both instruments *prior* to servicing (N = 12). The online process analyser was set as the master device and the lab instrument was defined as the host. The spectra transfer required the offline spectra to be adjusted so that they were similar to the spectra that were measured online. This 'forward transfer' (Bouveresse and Campbell, 2008) was chosen because the process analyser will be used as the standard device in the future. The 9000-4100 cm^{-1} region was used for the transfer, which resulted in 637 points per spectrum from each instrument.

A PCA was performed with the first derivative spectra to visualise the distribution of the 210 lab samples (l) and the 12 standard samples (s) in the score space (Fig. 4.16). For a PDS, the standard samples should cover the domain in which the samples to be transferred will lie as

homogenously as possible (Despagne et al., 2000). From the PC1-PC2 plot, it was obvious that this was not the case because the standard samples did not show any variation over the score interval; instead, they clustered at below average values of PC1. Clustering occurred because all of the samples came from digester A and, hence, solely reflected the physicochemical characteristics of the maize silage slurry. The log(1/R) spectra of the standard samples recorded on the master were plotted against the same spectra recorded on the host to identify the characteristics of the spectral differences among the two spectrometers (Fig. 17). Assuming that there were no differences between the instruments, the absorbance values should have lined up along the intercept line (Despagne et al., 2000). A comparison of the spectra indicated the occurrence of a non-linear effect, which may have been the result of stray light, scattering or detector saturation (Despagne et al., 2000). The master device showed a higher degree of absorption, and this systematic instrument difference was somewhat similar for all 12 samples because they were superimposed on each other.

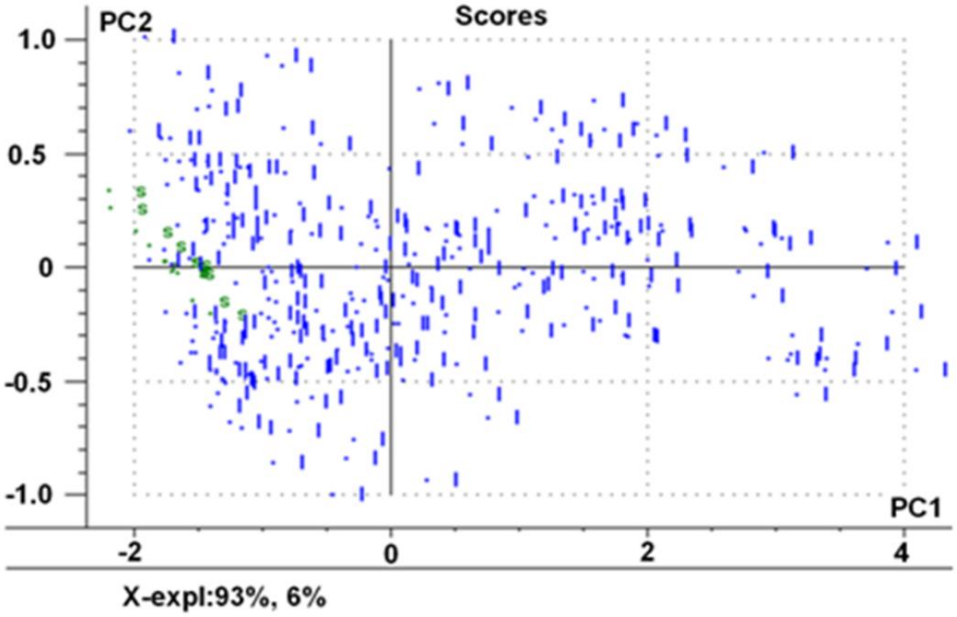

Figure 4.16: PC1-PC2 score plot of a PCA calculated from the first derivative spectra (2nd order polynomial, Savitzky-Golay, 17 smoothing points) of the 210 selected lab samples (l, blue) prior to their PDS transfer together with the 12 standard samples (s, green) measured with the same instrument. The information from wavenumbers 9000 to 4100 cm^{-1} was used.

The differences in the spectra of the standard samples (spectrum on master-spectrum on host) are shown (Fig. 4.18). The 7000 cm^{-1} and 5100 cm^{-1} regions had variations that were mostly due to changes in water absorption. This variance occurred because the data for the sample material were recorded offline and were not identical to the materials that were measured online and because their temperature was different. Additionally, the optical path length of the light that passed through the sample was different for the two measurement modes. It was decided to exclude those three samples from the transfer function with the largest differences based on visual inspection. The final PDS model was built via the Opus software after carrying out column mean-centring of the nine log(1/R) spectra to account for additive differences, for which a moving window size of seven was chosen (Feudale et al., 2002).

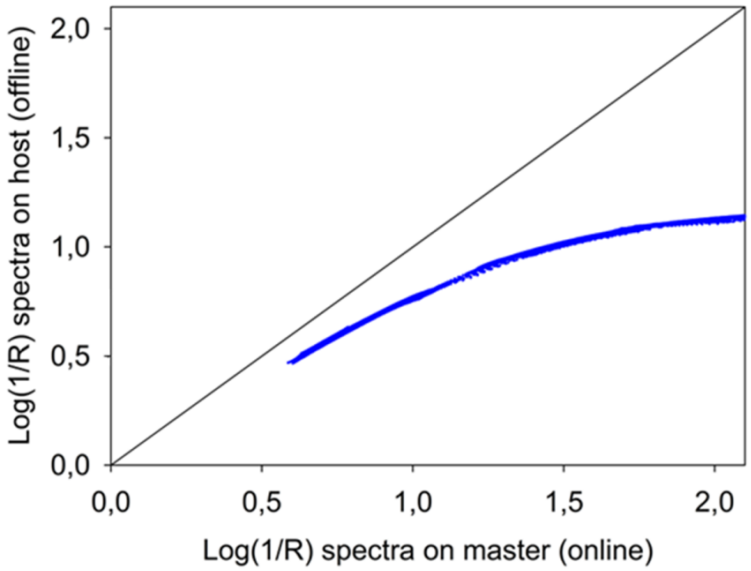

Figure 4.17: Comparison of the instrumental differences before spectra transfer. The log(1/R) values of the 9000 to 4100 cm^{-1} region recorded for the 12 standard samples are plotted.

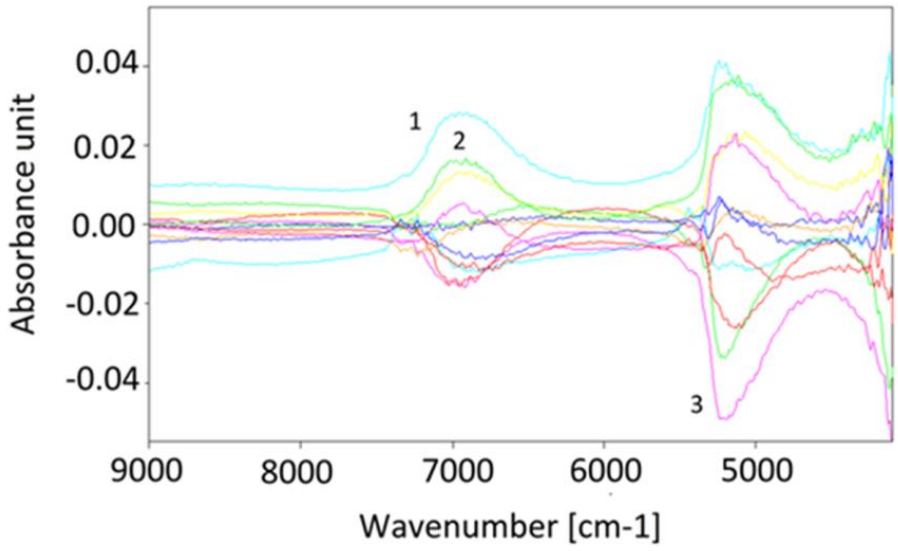

Figure 4.18: Difference spectra (master spectra - host spectra) of the standard samples. Samples 1-3 were removed prior to the computation of the PDS model.

The first derivative spectra of the 9000-4100 cm^{-1} region were used to calculate the PCA for the indication of the samples' distribution in the score space after the spectral transfer, i.e. application of the transfer function to the 210 lab samples (Fig. 4.19). The lab samples (l) and the samples that were obtained from digester A (a) and digester B (b) are displayed together in this PC1-PC2 score plot. The online samples did not fully cover the domain of the transferred lab samples because of the overall lower physicochemical variability of this online data set; an uneven distribution, such as that caused by clustering of the score values, was observed.

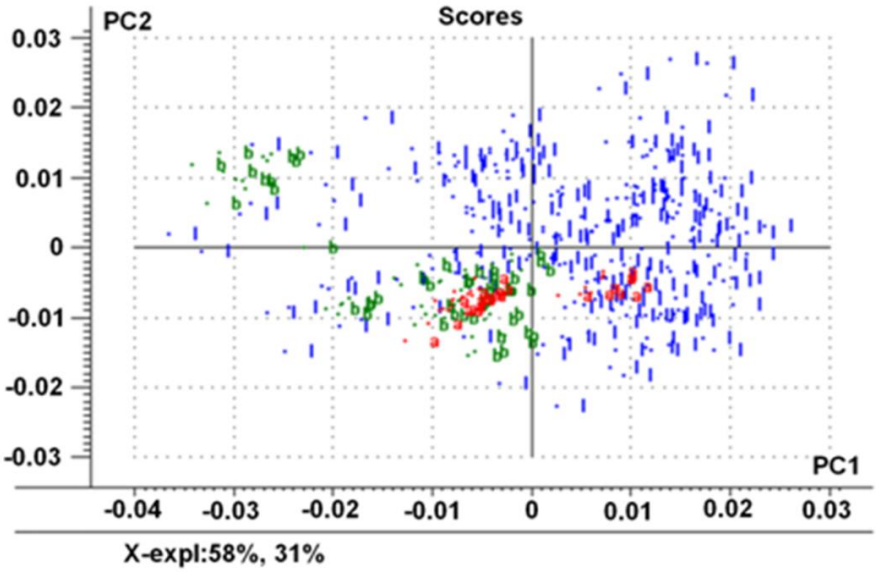

Figure 4.19: PC1-PC2 score plot calculated with the first derivative of the same spectral region for the combined calibration dataset (N = 285) indicating the transferred lab samples (l, blue), the samples from digester A (a, red) and digester B (b, green).

4.3.1.4. PLS model development

PLS regression was performed with the Unscrambler software using the transferred lab samples together with the online samples of digesters A and B (N = 285). Overall, 22 samples that were obtained from digester A were omitted from model development because of their high total VFA concentration >12 g kg^{-1} FM, a value which was defined as upper model boundary for this experiment. No further outlier removal in the X (spectra) or in the Y (reference) space was performed. Calibration was carried out with the log(1/R) spectra of the 9000-4100 cm^{-1} region after pre-treatment by employing the first derivative. These *combined models* were tested via segmented cross-validation with 10 samples that were reserved for calculating the sub-models (Næs et al., 2002). A re-sampling method that was based on the jack-knifing principle was used to select the spectral variables to compute the final model (Martens et al., 2001). The regression coefficients were calculated for each segment to yield information about the variability among the sub-models. The squared difference between the regression coefficient of each sub-model and that of the overall model was calculated for each variable. With an uncertainty limit that corresponded to two standard deviations, the significance of the variables was tested via a t-test (Martens and Martens, 2000; Anon., 2006). For computation of the final PLS model, the jack-knifing procedure was repeated several times. Each time, the non-significant variables from the previous round were omitted until the significant variables remained constant. Individual wavenumber regions were also manually removed to further increase model robustness by excluding noise and non-linearity from the calibration. These extractions included regions with high absorptivity (water bands) and limited detector stability at lower frequencies. This procedure resulted in a maximum of 15 different *combined models* that were calculated for each parameter. In addition, the same calibration procedure was performed with the spectra after carrying out pre-treatment with extended multiplicative scatter correction (EMSC) to remove multiplicative and additive effects (Kessler, 2007). EMSC modeled the physical effects that were related to wavelength-dependent light scatter variations and the mean of all of the spectra was used as reference (Anon., 2006). No further pre-treatments including combinations of pre-

treatments were tested. Analogous to this combined approach using the transferred offline spectra in conjunction with the online spectra, *online models* were developed solely using the samples that were obtained from digesters A and B. Because of the smaller number of samples that were available for calibration (N = 75), leave-one-out cross-validation was applied: each sample was held out to calculate the sub-models (Martens and Næs, 1991). Accuracy of the models was assessed by the *RMSEC* and the *RMSECV*.

4.3.1.5. External validation and time series evaluation

The *combined* and *online models* were both applied to the spectra obtained from the 42 samples that were removed from digester C for independent validation. The accuracy of the models was determined by *RMSEP*. Specifically, the precision was described by the *SEP*, which was corrected for bias, i.e. the systematic difference between the mean of the NIR estimates and that of the reference (Esbensen, 2009). Of the different models generated for each calibration approach, those that best represented the realistic potential of the method are presented. This selection was based on model complexity (number of PLS components), model accuracy and model precision. The models selected in this manner were applied to the spectra of the time series from digester C to evaluate their utility for continuous in-situ monitoring over an extended period. Prior to this, the time series spectra were averaged over a 12-min interval that corresponded to an effective measurement time of 3 min according to the 45 s/135 s cycle.

4.3.2. Results

4.3.2.1. Reference data set characteristics

The descriptive statistics for VS, NH_4-N and TIC indicate a broad concentration range for the combined calibration set, which reflects the variability in the feedstock mixture at different process stages (Tab. 4.11). In comparison, the online calibration set covered a reduced range. Whereas the combined calibration set fully enclosed the concentration range of the validation set, the online calibration set did not contain samples in the upper concentration regions that were reported for the validation set. Hence, external validation of the *online models* of these three parameters included testing the ability of the models to extrapolate values beyond their boundaries. For the three acid parameters, comparisons of means with medians indicated that the calibration data sets were skewed towards low-concentration samples. From a comparison of the values of the total VFA and the single acids, it was obvious that for the validation set, acetic acid contributed the most to the total VFA concentration. For the three data sets, medium to very high correlations were found between NH_4-N and TIC (Tab. 4.12). A high correlation was also observed between the total VFA and single acids; a high correlation among the single acids was only found for the validation set. For the validation samples, the temperature and VS exhibited medium to high correlation with some of the other parameters, which was not observed for the two calibration sets**.** The *SEL* were calculated from the duplicated wet chemical analyses to assess the contribution of the precision of the reference methods to the overall error estimates of the NIR models (Faber and Kowalski, 1997). For each parameter, the *SEL* is indicated separately for the three data sets. An overall similar repeatability was observed when relating the parameters to the absolute values. Moderate differences among the three data sets (e.g. for the total VFA) could be attributed to single samples measured at a comparably low precision.

Table 4.11: Descriptive statistics of the parameters VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid for the combined calibration set, the online calibration set and the validation data set. The arithmetic mean, median and standard deviation (*SD*) and the standard error of the laboratory (*SEL*) from wet chemistry analysis are presented. FM represents fresh matter.

Parameter [g kg^{-1} FM]	Data set	N	Range	Mean	Median	*SD*	*SEL*
VS	Comb. calibration set	285	39.5-128.6	72.4	70.5	15.6	2.1
	Online calibration set	75	46.1-85.4	68.5	67.9	9.41	2.0
	Validation data set	42	65.0-96.3	83.3	86.6	10.4	1.5
NH_4-N	Comb. calibration set	285	0.97-5.38	2.43	2.17	0.82	0.08
	Online calibration set	75	1.74-3.32	2.26	2.05	0.44	0.06
	Validation data set	42	1.90-4.37	3.42	3.69	0.84	0.11
TIC	Comb. calibration set	285	3.83-23.80	11.05	10.24	4.12	0.16
	Online calibration set	75	4.82-15.19	9.44	9.31	2.84	0.11
	Validation data set	42	10.32-20.71	16.58	17.72	3.99	0.21
Total VFA	Comb. calibration set	285	0.30-11.61	3.14	2.04	2.74	0.20
	Online calibration set	75	0.49-11.61	4.46	2.78	3.33	0.21
	Validation data set	42	1.24-7.51	3.64	3.32	1.66	0.32
Acetic acid	Comb. calibration set	285	0.01-7.91	1.68	1.05	1.83	0.13
	Online calibration set	75	0.01-7.91	2.66	1.53	2.40	0.14
	Validation data set	42	0.11.5.96	2.46	2.02	1.64	0.11
Propionic acid	Comb. calibration set	285	0.03-7.17	1.06	0.42	1.45	0.08
	Online calibration set	75	0.03-4.89	1.49	0.74	1.53	0.07
	Validation data set	42	0.03-1.32	0.53	0.48	0.44	0.03

Table 4.12: Statistical results of parameter cross-correlation analyses for the parameters VS, NH_4-N, TIC, total VFA, acetic acid, propionic acid and the temperature of the slurry. Coefficient of correlation (r) between the reference values for the combined calibration data set, the online calibration data set and the validation data set are given.

Combined calibration data set (lab data set and digesters A and B; N = 285)						
Parameter	Temp	VS	NH_4-N	TIC	Total VFA	Acetic
VS	-0.12	1.00				
NH_4-N	-0.05	0.25	1.00			
TIC	-0.12	0.27	0.89	1.00		
Total VFA	0.08	0.06	-0.16	-0.51	1.00	
Acetic acid	0.11	0.06	-0.07	-0.41	0.88	1.00
Propionic acid	0.01	0.01	-0.23	-0.50	0.81	0.48
Online calibration data set (digesters A and B; N = 75)						
Parameter	Temp	VS	NH_4-N	TIC	Total VFA	Acetic
VS	-0.01	1.00				
NH_4-N	0.15	0.41	1.00			
TIC	0.15	0.29	0.62	1.00		
Total VFA	-0.10	0.04	0.03	-0.72	1.00	
Acetic acid	-0.05	0.21	0.16	-0.61	0.91	1.00
Propionic acid	-0.19	-0.21	-0.16	-0.66	0.79	0.49
Validation data set (digester C; N = 42)						
Parameter	Temp	VS	NH_4-N	TIC	Total VFA	Acetic
VS	-0.54	1.00				
NH_4-N	-0.79	0.81	1.00			
TIC	-0.82	0.75	0.98	1.00		
Total VFA	0.20	0.51	0.08	-0.08	1.00	
Acetic acid	0.27	0.46	-0.03	-0.17	0.95	1.00
Propionic acid	0.57	-0.01	-0.44	-0.56	0.78	0.84

4.3.2.2. Preliminary estimations without spectra transfer

Pre-treatment performed on spectra recorded with the master and the host can reduce inter-instrument variations among the spectra while preserving the chemical signals (Despagne et al., 2000). Therefore, *preliminary models* were developed solely with the 210 samples measured on the lab instrument prior to their transfer to test whether pre-treatment could account for instrument differences shown in Figure 4.17. The same calibration procedure used for the combined approach was carried out. The results from the cross-validation and external validation of the six parameters are given in Table 4.13. For the cross-validation, pre-treatment by employing the first derivative and EMSC resulted in similar accuracy indicated by the *RMSECV* values. Independent model validation with the samples obtained from digester C revealed a large bias, except for the EMSC pre-treated model for NH_4-N.This indicates that pre-treatment of the spectra was generally not sufficient to eliminate instrument-related differences; however, a comparison of the *RMSECV* with the *SEP* from external validation showed that the estimates for the EMSC pre-treated models had similar error estimates.

Table 4.13: Estimation results from cross-validation and external validation of the *preliminary models* using the 42 samples from digester C for external validation of the parameters VS, NH_4-N, TIC, total VFA, acetic acid, propionic acid. For model development, the spectra of the 210 lab samples were used prior to their standardisation and were pre-treated by the first derivative (calculated with 17 smoothing points, Savitzky-Golay, 2nd order polynomial) and EMSC. FM = fresh matter; PLS = number of PLS components; EMSC = extended multiplicative scatter correction; *RMSE* = root mean square error of cross-validation (*CV*) and external validation (*P*); *SEP* = standard error of prediction.

Parameter [g kg^{-1} FM]	Pre-treatment	PLS	*RMSECV*	*RMSEP*	*SEP*	Bias
VS	1st derivative	6	5.0	106.9	10.8	106.4
	EMSC	6	5.6	44.1	5.3	-43.8
NH_4-N	1st derivative	3	0.24	3.78	2.36	-2.97
	EMSC	5	0.27	0.36	0.18	0.32
TIC	1st derivative	5	1.25	94.29	8.69	93.01
	EMSC	5	1.30	16.02	1.36	15.96
Total VFA	1st derivative	8	0.93	35.29	2.16	-35.23
	EMSC	10	1.15	16.34	1.19	-16.29
Acetic acid	1st derivative	8	0.94	14.53	5.28	13.56
	EMSC	9	0.93	9.73	1.30	-9.64
Propionic acid	1st derivative	8	0.73	25.45	2.07	-25.37
	EMSC	9	0.80	5.01	0.60	-5.00

4.3.2.3. PLS modelling results after spectra transfer

Figure 4.20 illustrates the results for the cross-validation of VS, NH_4-N and TIC given for the *combined models*. For VS, the *RMSECV* showed a steady decrease until the sixth component (Fig. 4.20a), while for the other two parameters, the first components contributed most to the explained variance of the reference data (Fig. 4.20b and c). The estimated vs. measured plot of the model for NH_4-N (Fig. 4.20e) indicated that the online samples, especially those from digester A, were all clustered at almost the same concentration. The range for the other two parameters was less correlated with the samples' origin, thereby giving a more even distribution; however, higher-concentration regions were also covered by the lab samples (Fig. 4.20d and f).

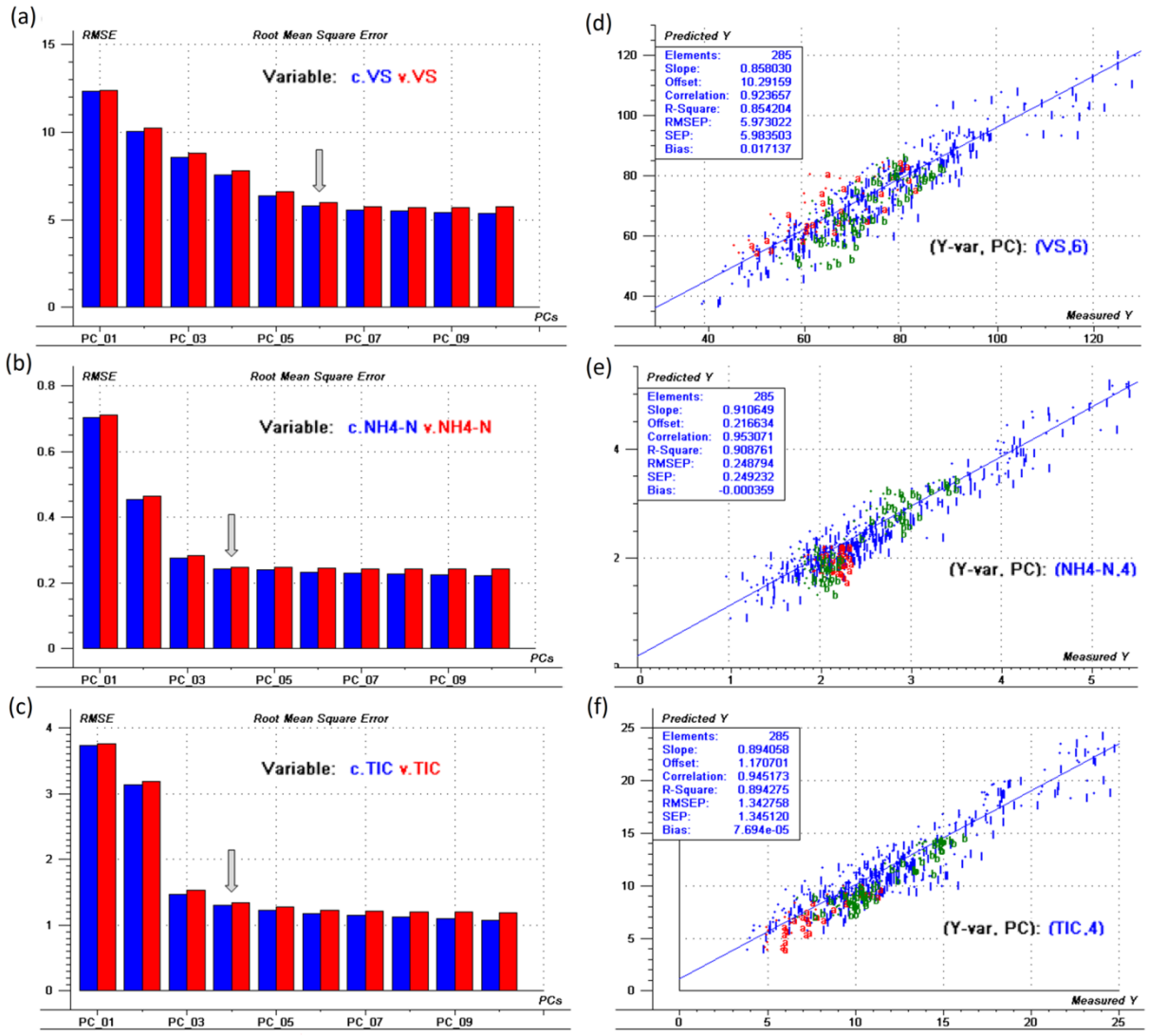

Figure 4.20: Results of the cross-validation for VS, NH_4-N and TIC of the *combined models*. (a-c) Root mean square error is plotted as a function of the number of the partial least square components. The results from the calibration (left bar) and cross-validation (right bar) are given. The arrow indicates the number of the components that were chosen. (d-f) Estimated vs. measured plots for the corresponding models, indicating samples that were obtained from the lab instrument (l, blue) and digester A (a, red) and digester B (b, green) recorded with the online process analyser. The models are the same as those in Table 4.14.

Table 4.14 summarises the results from the cross- and external validations, i.e., the application of the models to the 42 samples that were obtained from digester C. For the first three parameters (VS, NH_4-N, TIC), a comparison of the cross-validation results (*RMSECV*) indicated that the *online models* were more accurate than the *combined models* and that these differences were obtained for similar model complexity. The overall higher accuracy of the *online models* was also confirmed for the external validation as indicated by the *RMSEP* values. Both the models revealed the occurrence of systematic underestimation of the VS concentration and the high precision (*SEP*) of the *combined model*. Systematic overestimation of the *online model* was observed for NH_4-N. An *SEP* equal to the *RMSECV* indicated the high precision of the model obtained for independent validation. For TIC, the accuracy of the independent validation of the *online model* was similar to that of the cross-validation; for the *combined model*, validation of the 42 samples obtained from digester C revealed that the accuracy was higher than that

obtained from cross-validation. The *RPD* was calculated as standard deviation of the reference values of the validation set divided by the *SEP* used for standardising the *SEP* (Yookyung et al., 2007). For VS, an *RPD* value of 3.1 for the *combined model* demonstrated the suitability of the model for screening purposes according to the classification of Williams (2001). The *RPD* values of the two NH_4-N models differed, and a good screening potential was obtained for the *online model*, which was also confirmed for the *online model* of the TIC.

Table 4.14: Statistics of the validation of the partial least square models for the parameters VS, NH_4-N, TIC, total VFA, acetic acid and propionic acid. Results for the *combined models* and the *online models* (selected models) are shown for their cross-validation and their external validation with the 42 samples from digester C. FM = fresh matter; EMSC = extended multiplicative scatter correction; PLS = number of PLS components; *RMSE* = root mean square error of calibration (*C*), cross-validation (*CV*) and external validation (*P*); *SEP* = standard error of prediction; *RPD* = standard deviation of the reference values of the validation set (Tab. 4.11) divided by the *SEP*; *SED* = standard error of differences.

Parameter [g kg-1FM]	Model	PLS	*RMSEC*	*RMSECV*	*RMSEP*	*SEP*	Bias	Offset	Slope	*RPD*	*SED*
VS	Combined (EMSC)	6	6.0	6.0	6.4	3.3	-5.5	-0.76	0.94	3.1	1.1
	Online (1st deriv.)	5	3.0	3.3	5.3	3.9	-3.6	17.7	0.74	2.7	0.6
NH_4-N	Combined (1st deriv.)	4	0.25	0.25	0.38	0.35	0.17	-1.08	1.36	2.4	0.07
	Online (1st deriv.)	4	0.11	0.13	0.29	0.13	0.27	0.35	0.97	6.4	0.02
TIC	Combined (1st deriv.)	4	1.30	1.35	0.94	0.85	-0.42	-2.16	1.10	4.7	0.27
	Online (1st derive.)	4	0.69	0.75	0.66	0.66	-0.03	1.92	0.88	6.0	0.20
Total VFA	Combined (1st deriv.)	6	1.10	1.15	1.45	0.94	-1.11	-0.46	0.82	1.8	0.24
	Online (1st deriv.)	5	0.63	0.71	1.01	0.84	0.60	0.10	0.81	2.0	0.21
Acetic acid	Combined (EMSC)	8	0.89	0.95	0.87	0.87	-0.04	1.09	0.56	1.9	0.22
	Online (1st deriv.)	6	0.66	0.75	0.87	0.68	-0.56	0.23	0.69	2.4	0.17
Propionic acid	Combined (1st deriv.)	7	0.87	0.93	0.63	0.28	-0.56	-0.69	1.22	1.6	0.13
	Online (1st deriv.)	3	0.67	0.74	0.26	0.26	-0.05	0.21	0.58	1.7	0.07

Fig. 4.21 illustrates the cross-validation results for the three acid parameters that were obtained for the *online models*. For the estimated vs. measured plot of the total VFA (Fig. 4.21a), samples removed from digesters A and B were distributed along the concentration range, and the highest values were removed from digester A during periods of strong process imbalance. The models' corresponding regression coefficients are given in Fig. 4.21d, and they indicate that the major spectral information was located in the first overtone region (v2) over the range of 6000-5500 cm^{-1} (Workman and Weyer, 2008). The estimated vs. measured plots for acetic and propionic acids are shown. For the acetic acid model (Fig. 4.21b), the respective concentration of propionic acid of each plotted sample is indicated as 'low', 'med', or 'high'. For the propionic acid model (Fig. 4.21c), the concentration of acetic acid of each sample is indicated in a similar way. The corresponding U-T score plots for the selected components are provided in Fig. 4.21e (acetic acid) and Fig. 4.21f (propionic acid) for the graphical inspection of the operative X-Y

relationship (Esbensen and Geladi, 2010). These plots also show the concentration of the other acid. For the propionic acid model, a group of five samples (circled) with high measured concentrations indicated that underestimation had occurred (Fig. 4.21c); at the same time, their acetic acid concentrations were 'low'. These samples proved to be possible X-Y relation outliers because they ranged far away from the relationship that was defined by the other samples in the corresponding U-T score plot (Fig. 4.21f).

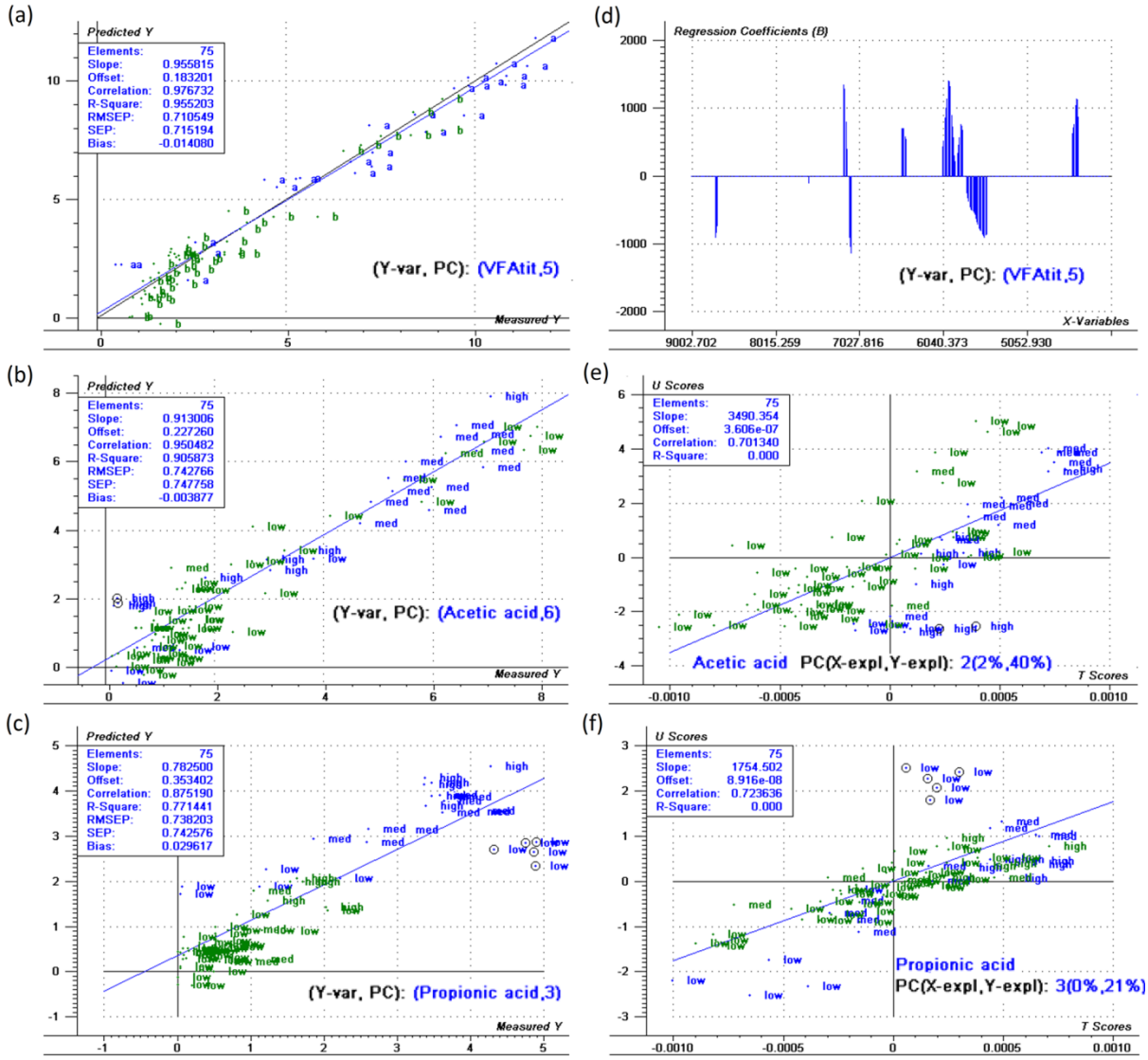

Figure 4.21: Cross-validation results for the *online models*. (a) Estimated vs. measured plot for total VFA, with samples from digester A (a, blue) and digester B (b, green). (b) Estimated vs. measured plot for acetic acid. The amount of propionic acid (in g kg^{-1} fresh matter) of the samples is given as 'low' (0–2), 'med' (2–4), or 'high' (4–6). (c) Estimated vs. measured plot for propionic acid, where the amount of acetic acid in the samples is given as 'low' (0–3), 'med' (3–6), or 'high' (6–9). (d) Regression coefficients of the total VFA model with five components. (e-f) U-T score plot for acetic acid and propionic acid. The U-T score plot for the partial least square component with the highest contribution to the explanation in Y was selected. The models are the same as those in Table 4.14.

With an *RMSECV* close to 0.75 g kg^{-1} FM reported for the *online models*, the accuracy of each of the three acid parameters was higher than that observed for the *combined models*, having values of approximately 1 g kg^{-1} FM or higher (Tab. 4.14). External validation for total VFA

indicated a moderate loss in the accuracy of both models, which was most likely due to systematic error. The *SEP* suggested that the two total VFA models were similarly precise. For acetic acid, the external validation also indicated a small loss in the accuracy of the *online model*. The application of the propionic acid models to the samples that were obtained from digester C showed a decrease in error estimates compared with the cross-validation. Overall, the independent validation of the acid parameters indicated a (very) poor screening ability with *RPD* values ≤2.4 (Williams, 2001). An alternative classification developed by Malley et al. (2004) for the specific assessment of environmental samples such as soil, sediments, animal manure, and compost also considered these models as improper (see Tab. 3.2).

4.3.2.4. *Time series evaluation*

Application of the models for VS, NH_4-N and TIC of the spectral measurements during the 26 days of operation of digester C is shown in Figure 4.22a-c, where the NIR estimates are plotted together with the reference data. Although absolute accuracy was not achieved, the VS of the models correctly followed the concentration trend determined by the reference. Clearly visible are the models' underestimations during the period of higher VS concentration. On day 11, exchange of 60% of the digester content (Fig. 4.22g) resulted in a moderate increase in the VS level, which was indicated by the models. The models also revealed the inhomogeneity directly after the addition of the slurry. The opposite occurred on day 19 when 80% of the digester content was substituted with slurry at a lower VS concentration. Note that, on the first days of the fermentation run, the *combined model* reported daily swings in the estimates. The time series estimates for NH_4-N showed high precision, and visual inspection of the *online model*, in particular, indicated the occurrence of a stable signal. A decrease in the NH_4-N concentration after substituting the digester slurry was indicated by both models. For the *combined model*, systematic overestimation occurred at the beginning of the fermentation run, whereas from day 19 onwards, the estimates indicated that underestimation occurred. In contrast, estimates of the *online model* resulted in a moderate overestimation of NH_4-N throughout the complete observation period. For the TIC, both the models indicated accurate estimations with some systematic errors from day 19 onwards.

The concentration courses for total VFA, and acetic and propionic acids are given in Figure 4.22d-f. Initially, the total VFA concentration was approximately 2 g kg^{-1} FM. An increase in VS feeding to up to 4 kg m^{-3} d^{-1} caused the accumulation of total VFA to climb to 7,5 g kg^{-1} FM (day 19). The NIR estimates of both models followed this trend, indicating the presence of fluctuations over the first 10 days of the measurement period. After substituting the slurry with material removed from a stabile AD process (day 19), the drop in the total VFA concentration was correctly indicated by the models. Although underestimation was observed, the rising total VFA concentration caused by stronger feeding on days 21 and 22 was precisely predicted. The same was true for the flattening of the total VFA increase when feeding was stopped at the end of the observation period. Acetic acid contributed the most to the total VFA concentration and the estimates showed overall similar behavior. Compared with the total VFA estimates, the acetic acid *online model* indicated the occurrence of a more even pattern; it was less noisy and showed only small fluctuations on the first days. A similar trend was observed for propionic acid, and the absolute values ranged from 0 to 1.3 g kg^{-1} FM. Changes within this range were adequately estimated and in particular, the *online model* demonstrated high accuracy.

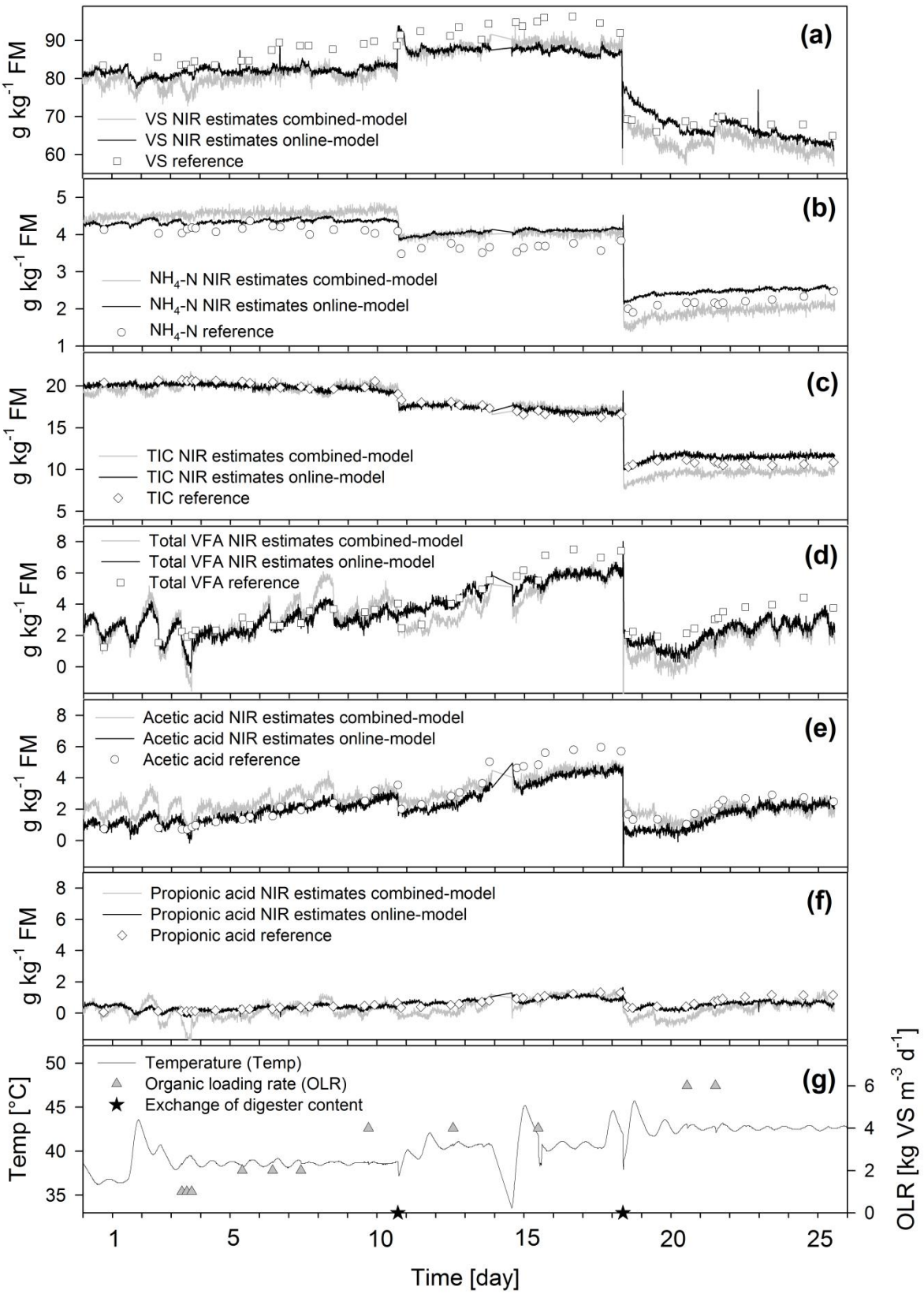

Figure 4.22: (a-f) Results from time series validation of VS, NH_4-N, TIC, total VFA, acetic and propionic acids parameters. The combined and *online models* were validated with the online measurements of digester C. The spectra were averaged over a 12-min interval that corresponded to an effective recording time of three min. A power failure caused a standstill in the measurements on day 15. (g) The temperature of the digester during the fermentation run plotted together with the organic loading rate. The stars indicate the time of exchange of the digester content on day 11 (60%) and day 19 (90%) with material from two biogas plants. Low fill level caused erroneous estimates during exchange on day 19.

The standard error of differences (*SED*) was calculated for the six parameters to compare the repeatability of the NIR estimates over the entire measurement series (day 1-26) according to the following equation (Workman Jr., 2008):

$$SED = \left\{ \frac{\sum (y_1 - y_2)^2}{2N} \right\}^{1/2}$$

Here y_1 = first analysis, y_2 = second analysis and N = total number of samples. For calculating the *SED*, two consecutive NIR estimates were treated as if they were duplicated measurements of one sample (first and second analysis) to exclude the effect of a change in the concentration over time (the time period during the exchange of slurry was omitted). The *SED* values (Tab. 4.14) indicated that the *online models* resulted in a higher precision overall. In particular NH_4-N revealed differences in the repeatability between estimates of the models. For a visual comparison of the analytical precision, the mean-centred NIR estimates are shown specifically for day 17, and they reveal the deviations from their 24-h average value (Fig. 4.23). The distribution shown in the box plots confirmed the overall higher precision of the *online models* compared to that of the *combined models* because the variation in the estimated data was less pronounced. Deviations of the median from zero were caused by an individual trend in the estimates during the day. Of the three acid parameters, propionic acid demonstrated the highest repeatability in absolute terms.

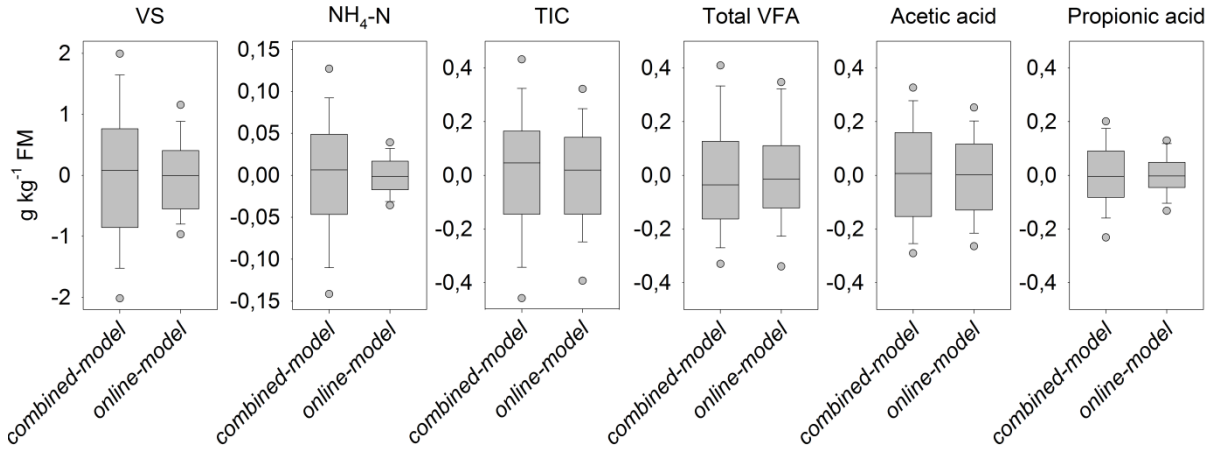

Figure 4.23: Variations in the mean-centred near infrared estimates that were calculated as examples for day 17 (24 h at 12-min intervals) to demonstrate the repeatability of the analysis. Box: line = median, boundaries = 25th and 75th percentiles, whiskers = 10th and 90th percentile, and dots = 5th and 95th percentiles.

4.3.3. Discussion

The representativeness and stability of the standard samples were sufficient for the generation of an appropriate PDS model because transfer of the calibration samples did not result in any relevant loss of extractable chemical information. This conclusion was made by noting that the *RMSECV* values obtained for the *combined models* were similar to the *RMSECV* values of the

preliminary models. The reason for this similarity may be that the different instrument responses featured similar patterns for all the standard samples. For example, no additional independent perturbations which would have negatively affected the spectra transfer were observed for the standard samples (Despagne et al., 2000). The influence of small discontinuities or irregularities in the transferred spectra was presumably reduced because of the a priori weighting of the input variables by PLS modelling (Bjørsvik and Martens, 2008). The use of jack-knifing may have further increased model robustness by excluding noisy frequency regions, including those regions that may have contained deficits, after spectra transfer. These results also suggest that the higher *RMSECV* values that obtained for the *combined models* compared with those of the *online models* could not be explained by the transfer themselves and thus require an alternative explanation. For instance, the considerably larger feedstock variation of the lab samples may have increased the occurrence of spectral interference, thereby resulting in the gradual loss of extractable chemical information. Furthermore, for VS, NH_4-N and TIC, the *combined models*' concentration range was larger, thus increasing the risk of including non-linear relationships between chemical and spectral readings (Næs et al., 2002). In addition, differences among measurement modes and instrument performance also may have affected the model performance. One example might be the assumed higher representation of the online recordings for the heterogeneous slurry compared with sample cup measurements that were performed for the lab samples, e.g., stronger effects from repacking errors (Workman Jr., 2008). Furthermore, the lab instrument showed lower absorptivity, presumably because of an early saturation of the detector. Overall, these factors may have had a cumulative effect, and this may explain the observed differences between the internal accuracy (*RMSECV*) of the two modelling approaches. The results of the *preliminary models* also allow assessment of the effect of the two different reference methods for analysis of the total VFA on the accuracy of the NIR method (chpt. 3.3.3). The accuracy of the best *preliminary model* for total VFA_{tit} (*RMSECV* = 0.93 g kg^{-1} FM; Tab. 4.13) was only slightly lower than the accuracy obtained for the total VFA_{GC} of the models developed from Data Set C in Table 4.5 (*RMSECV* values of 0.81-0.87 g kg^{-1} FM). These similar error values suggest that the use of the two different reference methods did not affect the NIR model performance to a large degree which can be expected since both methods reported a similar precision (*SEL*) as was shown in Table 3.4.

Although the *combined models* indicated a loss in accuracy, as reported from cross-validation, their robustness was expected to be higher than that of the *online models* because of the larger sample pool of the combined calibration set. However, this assumption could not be proven as shown by the results from the external validation that used digester C samples. Although they were based on a comparably small data set and incorporated material from only five different biogas plants, the performance of the *online models* in general was higher for these open population conditions. This indicated that for the data generated in the present study, a comparably small calibration set already provided satisfactory estimations when it was tested on a structurally similar slurry of different origin. For each parameter and training set, selected models are presented in Table 4.14. In general, this procedure may result in selecting a model that shows a good fit for a particular validation data set used in the experiment. However, the model may fail to adequately describe a slurry from a different source. The conclusion that the model meets the requirements for robust application would then be misleading. To reduce the risk of drawing such misleading conclusions from these results, the slurry used for validation

was chosen to represent a range of sample matrices that reflect different feedstock mixtures. The exchange of material from biogas plants operated with different types of feedstock ensured that the models generated in the present study proved their robustness.

The time series estimates illustrated both systematic and random errors of the *combined* and *online models* over time, and the substitution of digester content revealed the effect a change in the sample matrix had on the estimates. The underestimation of both of the VS models at higher concentrations also may have occurred because, the PLS model was balanced to most closely fit the samples in the 60 to 80 g kg^{-1} FM region for both calibration sets (see Fig. 4.20d). Furthermore, the online calibration set did not cover these upper concentration regions, which may have resulted in non-linear X-Y relationships above the model's boundary. The daily swings for the estimates of the *combined models*, which mostly occurred during the first days of the experiment, were also observed for other parameters and could not be rationalised by AD process dynamics. These fluctuations did not coincide with the temperature in the digester during this period, and hence these could not be explained by temperature sensitivity of the models. The reasons for these fluctuations are not yet clear (see Jacobi et al. (2011) who reported similar observations for NIR analyses of the substrate). Presumably, they were caused by the daily rhythm in the air moisture in front of the measuring head. Further study is necessary to test whether technical improvements may correct this effect. This also includes performing the spectrometers' white referencing more frequently and on an automatic basis since for this fermentation run referencing was performed manually once a day.

The repeatability (*SED*) of the consecutive NIR estimates for the six parameters was higher than what was reported for the lab application (Tab. 4.8). With this precision, it seems possible to follow the medium-term (VS and NH_4-N) and the dynamic (TIC and VFA) process changes over time, which thus favours the use of this technique for the fast adaptation of feeding in quasi-real time. The higher precision of the *online models* compared with that of the *combined models* presumably demonstrates the ability of the former models to more specifically extract spectral information related to these parameters. This information on the liquid phase may also be used in conjunction with NIR-generated data on the feeding substrate quality for advanced plant control (Jacobi et al., 2011). The overall similar course of NH_4-N and TIC during the 26-d period required an investigation into whether these two parameters provided different information or could be summarised as one 'latent' variable. The high collinearity between both the parameters was addressed earlier where the TIC was found to be a function of the concentration of NH_4-N and total VFA (chpt. 4.2). The correlation coefficients (r) among the estimates of the TIC and the NH_4-N were calculated for the three phases between the exchanges of digester material to exclude slurry-specific effects. The r was tested at α = 1% for significance (***). Given the results of the *online models*, during the first phase (day 5-11; day 1-4 was omitted due to fluctuations), the estimates of the two parameters were weakly and negatively correlated (r = -0.40***). For the second (day 11-19) and third (day 19-26) phases, different correlations of r = -0.55*** and r = 0.64***, respectively, were obtained. Therefore, the overall downward trend of both of the parameters was caused by the exchange of the digester content with slurry that was low in NH_4-N and TIC. However, the short-term changes in the concentration course of the TIC were also affected by the feeding-induced changes in total VFA (day 11-19). This trend was also correctly indicated via NIR. For the TIC, the smaller

RMSEP of the *combined model* compared with the *RMSECV* was noticeable and presumably caused by a small matrix-specific error observed for the slurry used here. In the long run and when testing the models on different materials, an *RMSEP* value not smaller than that obtained from cross-validation should be expected (see Ward et al. (2011) for comparison of the results).

The time series evaluation of the total VFA concentration reflected the expectations for the process behaviour. When feeding was increased, an increase in the concentration was indicated by the reference values and the NIR estimates correctly followed this trend. This outcome confirmed the results of Jacobi et al. (2009) and Lomborg et al. (2009), which were obtained with a simpler slurry-matrix, also for a feedstock-robust approach. The ability of the NIR method to specifically estimate the total VFA concentration even under varying AD conditions and by using different sample matrices is supported by data in Figure 4.21d, where the regression coefficients of the NIR model corresponded to the acid's CH_2- and CH_3- group peaks in the first overtone region. With respect to the standardised model performance, for total VFA, an *RPD* value of as low as 2.0 (Tab. 4.14) probably undervalues the true potential of the method for the early detection of the accumulation of the acid as concluded from the visual inspection of the time series. Here, even small changes in concentration could be visualised due to the high repeatability of the NIR analyses and at a high temporal resolution. Two aspects may help to explain this discrepancy. In general, the *RPD* is affected by the range in the reference values as well as by the distribution of the samples (Malley et al., 2004). Since most samples were clustered closely around the mean value, an overall small *SD* was the consequence and resulted in a low value for the *RPD* as well. Secondly, since the reference method contained a measurement error (*SEL* of 0.32 g kg^{-1} FM, Tab. 4.11), the obtained prediction error was systematically overestimated, which also affected the *RPD* value. This second aspect is related to the general problem of multivariate calibration in stating that 'the contribution of the measurement error in the reference values to the apparent prediction error estimate is interpreted as an inadequacy of the calibration model rather than an inadequacy of the reference values themselves' (Faber and Kowalski, 1997).

Evaluating the single acids required consideration of their high collinearity with respect to the total VFAs. In general, PLS modelling of inter-correlated parameters runs the risk of including non-specific information in the calibration (Workman Jr., 2008). For the present study, this problem presumably is visualised in Figure 4.21c for the five samples for which the propionic acid concentrations were clearly underestimated. These samples were removed from digester A during the recovery phase, during which the acetic acid concentration decreased rapidly, whereas the propionic acid concentration remained stable. This situation is often observed for AD processes following instabilities since the degradation of propionic acid is thermodynamically unfavourable (Boe, 2006). From this different behaviour of the five calibration samples, which was also observed in the U-T score plot, it is assumed that the *online model* of propionic acid was possibly closely related with the NIR model of the total VFA. The two acids were also highly correlated to the total VFA during external validation. Therefore, the results obtained from the time series evaluation indeed indicated the very high repeatability of the NIR method for the single acids; however, they could not confirm the accuracy of the estimates on an independent basis.

5. The potential for monitoring the short-term process dynamics

5.1. Development of a maize-specific online calibration

Abstract

This experiment reports the development of a NIR spectroscopy online calibration for monitoring the process parameters of VS and total VFA. The objective was to investigate the potential for estimation of the AD process parameter dynamics after feeding events. A recirculation loop was employed to record permanent online measurements during an eight-month experiment using a pilot plant fed with maize silage under mesophilic conditions. Sampling was performed to obtain calibrations and subsequent test-set validations, comparing the NIR spectroscopy estimations to the reference values. The accuracy of the generated calibrations exhibited mean prediction errors (RMSEP) of 3 g kg^{-1} for VS and 0.9 g kg^{-1} for total VFA in the fresh matter. By applying the calibrations to time series spectra, the model accuracy allowed adequate indication of the concentration changes including highly sensitive monitoring of the total VFA short-term dynamics. The independence and high sensitivity in the estimation of both parameters underline that the estimates for VFA were independent of the VS content and the high potential this technology offers for the optimised process operation of biogas digesters. The monitoring of relative VS and VFA changes may be the key to further process control via NIR spectroscopy, as feeding may be adjusted to the current AD process conditions.

5.1.1. Material and methods

5.1.1.1. *Experimental setup and process operation*

The standing tank digester A (see Fig. 4.15a) was also used for this experiment. With a working volume of 2500 L, the digester was operated under mesophilic conditions. For more than six months prior to the experiment, the digester was fed at an OLR, measured in terms of VS, of 2 kg m^{-3} d^{-1} with maize silage. Gas production was measured via a drum-type gas meter (Ritter GmbH, Bochum, Germany) at 10-min intervals. The gas quality (CO_2 and CH_4) was recorded via infrared sensors (Awite GmbH, Langenbach, Germany) every two hours. The digester temperature was regulated by a sensor located on the sidewall of the digester, which automatically controlled the flow of heating water through the closed heating circuit. Changes in temperature (logged at 10-min intervals) over the range 31 to 47 °C were induced during the experiment to include the effects that temperature may have on the spectra (Hageman et al., 2005). The digester was fed with maize silage at an average VS concentration of 33% of FM. Irregular feeding was administered to generate strong variations in the AD process, with an OLR ranging up to 35 kg m^{-3} d^{-1}, which corresponds to a hydraulic retention time of nine days.

During feeding, digester slurry was pumped from the lower part of the digester into a tank where it was mixed with the silage before being repumped into the upper region of the digester using a

spiral pump. When the OLR was increased, the ratio of silage to digester slurry was kept constant by continuous repumping of the mixture into the digester while adding silage into the tank. Five phases of organic overload were induced by strongly increasing the amount of silage that was fed to the digester during the 240 days of the experiment (Fig. 5.1). To allow the digester to recover from each overload, feeding was usually stopped if the methane concentration in the gas dropped below 45% prior to the next feeding event. For the first phase of organic overload, an OLR between 4.4 to 6.4 kg m^{-3} d^{-1} was chosen. For the second, third, and fourth phases of overload, feeding was increased up to 27 kg m^{-3} d^{-1} (day 147), resulting in a VFA concentration of 9 g kg^{-1} FM. When feeding was stopped, a sudden recovery of the digester was observed, as the accumulated acids were quickly degraded. For the fifth overload, an OLR of 35 kg m^{-3} d^{-1} (day 190) was selected, which caused long-term process imbalances accompanied by large amounts of propionic acid.

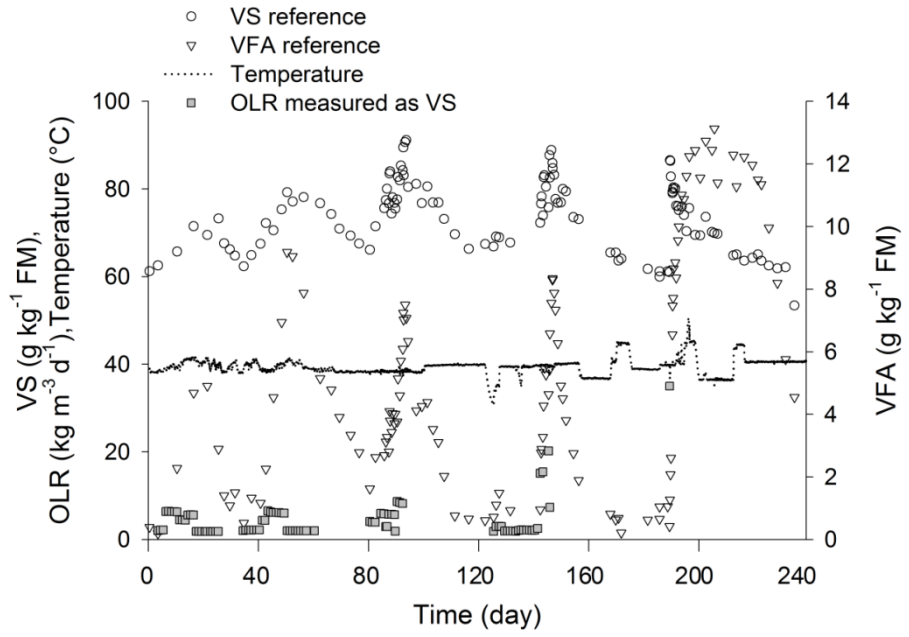

Figure 5.1: Concentrations of VS and total VFA during the five phases of organic overload induced by excess feeding. The OLR refers to the amount of VS fed to the digester during one feeding event.

5.1.1.2. Spectra acquisition and sampling

The spectral data were measured using the same external-loop as described in experiment three (chpt. 4.3.1.2). The system was directly connected to the sidewall of the digester. The inlet was located at the bottom of the digester and the outlet halfway up the side of the digester. The digester content was pumped through a metal pipe (50 mm inner diameter) for the duration of the experiment to provide continuous loop-circulation. The metal pipe of the loop was insulated to maintain the temperature of the slurry. A sapphire window was fitted in an upstream section of the pipe, and a diode-array spectrometer (X-Three, NIR-Online GmbH, Walldorf, Germany) was mounted on this window (Fig. 5.2). Spectra from diffuse reflection were recorded, covering the 400-1750 nm region at a spectral resolution of 10 nm (Si and InGaAs detectors). The measurement area was 13 mm in diameter. The raw spectral data captured at intervals of 15 s were averaged over three minutes for one data point. White referencing was performed automatically at irregular intervals depending on the collection conditions.

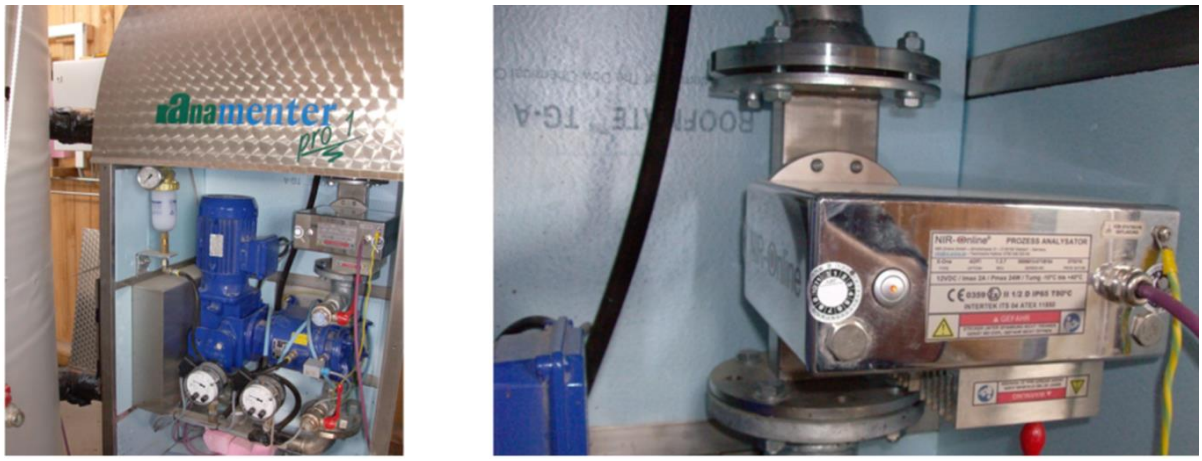

Figure 5.2: Left: 'Anamenter' used at digester A for enabling continous loop-circulation of the slurry. Right: Spectrometer 'X-Three' flaged at the sapphire window at an upstream section of the pipe.

In total, 120 samples were collected. During times of intensive feeding, sampling was conducted directly before and several hours after each feeding event. During times of reduced feeding, sampling was usually performed twice a week. Primary samples were collected via a ball valve in the loop that was situated after the spectrometer. The spectra that were recorded over a sampling time of 20 s were averaged and utilised for subsequent calibration. The volume of the primary sample ranged from 20 to 30 L depending on the viscosity of the slurry. From this, a 5-L sub-sample was collected before a 0.5-L sample was bottled and immediately frozen at -18 °C until the wet chemical analysis. The analyses were performed in duplicate. If the difference between the replicates for one sample was higher than 10%, a third measurement was recorded. The average value of all replicates determined the reference value that was used for calibration. For the samples of experiment four, the *SEL* of the wet chemistry method (see chpt. 3.3.1) was 1.7 g kg^{-1} for VS and 0.19 g kg^{-1} FM for total VFA.

5.1.1.3. Multivariate data analysis

Multivariate data analyses were performed with the log (1/R) spectra of the 1100-1750 nm region using the Unscrambler 9.8 (Camo). For explorative data analysis, a PCR (Næs et al., 2002) was conducted with the complete data set. Each model that was developed for VS and VFA was tested with a leave-one-out cross-validation. During this procedure, each sample in the data set was used for calibration and validation (Martens and Martens, 2000). By looking for unexplained variances of X (spectra) and Y (reference), the data set was probed for outliers. Only those samples that were insufficiently described by the model and at the same time had a significant influence were removed. This resulted in the removal of one sample, which was defined as a Y-outlier for the VFA. The final calibrations for VS and VFA were developed using PLS regression and test-set validation. The data set was divided into two sub-sets, one used for calibration and the other for validation. The samples obtained during the first, second (day 1–83), and fifth (day 159–240) phases of organic overload were used for calibration. The samples from the third and fourth (day 84–158) phases of overload were used as a test-set (Fig. 5.1). Therefore, the calibration and validation sets originated from the same data set and were only independent with respect to time. In total, 65 (64 for VFA) and 55 samples were used for calibration and validation, respectively. PLS regression was performed using the raw spectra.

Additionally, a first derivative was applied to reduce morphological structure in the spectra (Heise and Winzen, 2002). The calibrations with the smallest *RMSEP* (Kessler, 2007) were selected for time series evaluation.

5.1.1.4. Spectra time series

The selected calibrations were applied to the continuously measured spectra during the third and fourth phases of organic overload. The NIR estimates were compared to the 55 reference values obtained for each parameter. Whether the NIR spectroscopy correctly described the long-term trends in the concentration of both parameters during the variable feeding was examined. Subsequently, the suitability of the models to capture short-term changes was tested by focussing on a detailed picture of the VS and VFA dynamics estimated by the PLS model. For this purpose, two different sections characterised by a VS feed varying from 1.9 to 8.7 kg m^{-3} (days 89 to 93) and from 7.5 to 20 kg m^{-3} (days 144 to 148) were selected.

5.1.2. Results and discussion

5.1.2.1. Reference data set characteristics

Descriptive statistics of the calibration and validation data sets are shown in Table 5.1. The VS concentrations in the calibration set varied from 53 to 87 g kg^{-1} FM; the VS concentrations in the validation set slightly exceeded this concentration range. Compared to the validation set, the samples used for calibration were spread over a wider range with respect to VFA, ranging from 0.2 to 13.1 g kg^{-1} FM. The upper value reflects an acidified process, as it is clearly >3 g kg^{-1} FM, a value that is potentially a critical limit for AD in agricultural biogas plants (Holm-Nielsen et al., 2007). As collinearity can reduce model stability, correlations in the calibration data set are of interest. A strong correlation (r = 0.99) between the TS and VS was observed, as they differed only by ash content. Although the VS and VFA concentrations showed similar behaviour during the experiments (feeding resulted in a concentration increase, non-feeding in a decrease), only a small correlation (r = 0.26) was found between the two parameters within the calibration data set. In contrast, for the test-set samples, a high correlation between VS and VFA (r = 0.86) was found. Within the whole data set, a small correlation (r = 0.20) between the two parameters was found, indicating that the parameters vary independently of each other over the complete experimental period.

Table 5.1: Descriptive statistics of the calibration and validation data set for VS and total VFA.

	Calibration data set (N = 65)[1]				Validation data set (N = 55)			
	Range	Mean	Median	*SD*	Range	Mean	Median	*SD*
VS (g kg^{-1} FM)	53.4-86.5	69.7	69.4	7.2	66.3-91.1	78.5	77.7	6.1
VFA (g kg^{-1} FM)[1]	0.2-13.1	5.9	4.8	4.5	0.6-8.3	4.2	4.0	2.1
Temperature (°C)	36.4-47.5	39.9	39.8	2.3	31.2-40.3	38.8	39.4	1.6

[1]For VFA, calculations for the calibration data set were performed after the removal of one sample that was defined as an outlier resulting in n = 64. FM = Fresh matter; *SD* = Standard deviation.

5.1.2.2. PLS calibration results

Figure 5.3 provides an overview of the PLS regression results for VS and VFA based on test-set validations. Both parameters were estimated with acceptable accuracy. A mean estimation error (*RMSEP*) of 3 g kg^{-1} FM for the VS is in accordance with results from another study that reported a *RMSECV* of 2.7 g kg^{-1} FM (Holm-Nielsen et al., 2007). Smaller errors were obtained by Lomborg et al. 2009, which may have been because of a reduction in heterogeneity resulting from the use of ground maize silage in their experiment. The relatively low coefficient of determination (r^2 = 0.76) for VS in the present work compared to Holm-Nielsen et al. (2007) is presumably because of the small overall variation, which ranges from 66 to 91 g kg^{-1} FM. To examine the robustness of the model with respect to temperature effects, samples measured at different temperatures in the calibration set were included (Hageman et al, 2005). Within the calibration set, the temperature varied from 36-47 °C. As larger errors in the estimates were not observed for the test-set samples at lower temperatures, it is assumed that the calibration for VS also accounted for the temperature effects under moderate extrapolation conditions.

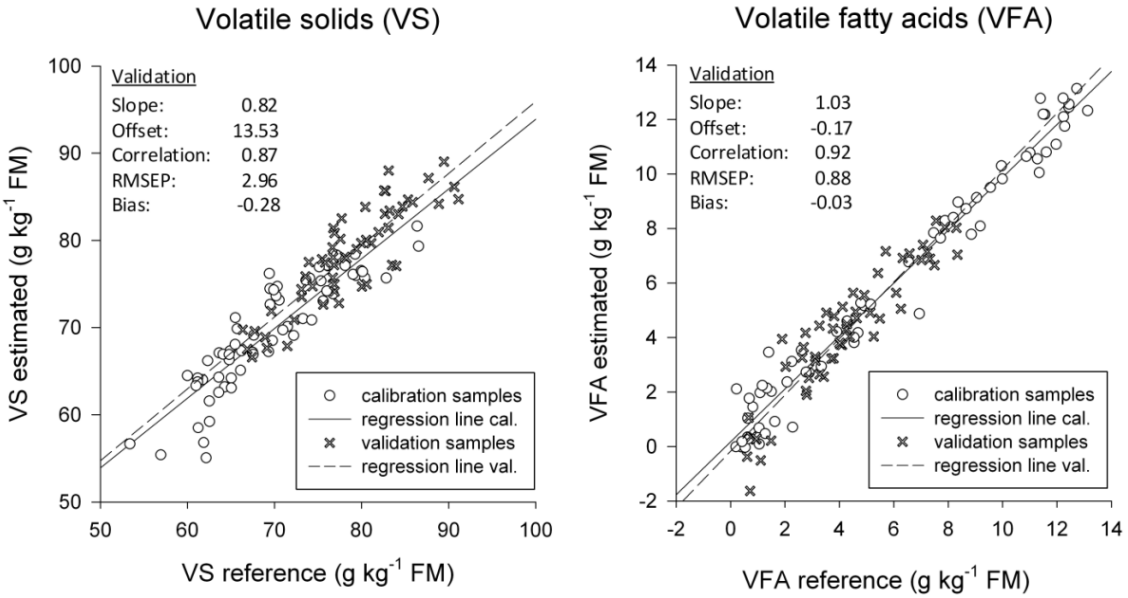

Figure 5.3: PLS regression results of the test-set validation (val). Left: Calibration (cal) for the VS with three PLS components calculated from the raw spectra (1100–1750 nm). Cal N = 65, val N = 55. Right: Calibration results for total VFA with five PLS components calculated from the first derivative (1100–1750 nm, Savitsky-Golay, three smoothing points, 2nd degree polynomial). Cal N = 64 (one outlier removed), val N = 55.

The VFA concentration was estimated with a *RMSEP* of approximately 0.9 g kg^{-1} FM (r^2 = 0.85) using the first derivative. In the literature, the error for this parameter ranges between 0.2 to 1.6 g kg^{-1} FM and has been reported for varying experimental conditions and different validation methods (Holm-Nielsen et al., 2007; Lomborg et al., 2009). Jacobi et al. (2009) performed an online calibration using a digester fed with maize silage and employing an experimental setup comparable to the one reported here. With a *RMSECV* of approximately 0.8 g kg^{-1} FM (r^2 = 0.94), a similar VFA performance was achieved.

5.1.2.3. *Time series evaluation*

Long-term evaluation

The models for the two parameters with the spectral measurements during the third and fourth phases of overload are illustrated in Figure 5.4. Owing to feeding, an increase in VS was observed in the digester, and this is adequately described by the model. Additionally, a steady decrease was observed in the concentration level after feeding was temporarily stopped on day 94. However, at times, the reference value for VS obviously differed from the estimates. The dominant contribution to this difference seems to be associated with the sampling error. On day 103, the reference value showed an increase in VS concentration, a change that seems unlikely, as feeding was stopped on day 94. In contrast, the NIR estimates indicated a smooth downward trend, which presumably describes the continuous biodegradation of the organic matter. Constitutional and distributional heterogeneity of the digester slurry may be considered important factors that can influence the sampling error (Esbensen and Geladi, 2010).

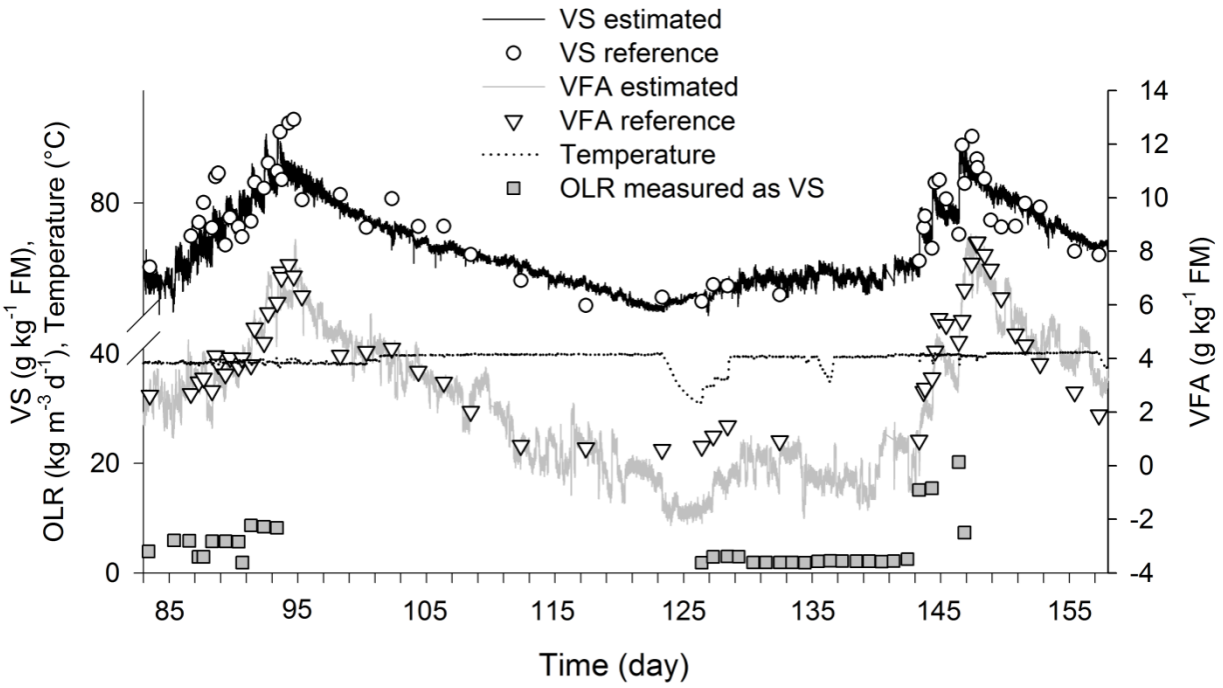

Figure 5.4: Long-term validation for the third and the fourth phases of overload. NIR estimates and reference values for VS and total VFA are shown together with the OLR and the temperature profile in the digester.

As in the case of VS, the estimates for VFA also followed the feeding-induced trend of the reference analyses. The differences between the estimated and experimental values were less pronounced compared to those observed for VS. This is probably because the determination of VFA is less sensitive to the sampling error because the acids are present in the aqueous phase. Starting on day 124, the estimates decreased and approached an implausible value of −2 g kg^{-1} FM, which was quite different from the reference data. At that time, the temperature in the digester had been reduced to approximately 31 °C. After a stepwise increase of the temperature to 40 °C (day 129), the estimated values again resembled the reference values. This drift in the estimates indicates that, in contrast to that of the VS, the model for VFA was sensitive to temperature changes that were not considered during calibration. However, before and after this

period, the estimates for VFA also sometimes fell below zero. This finding most likely reflects the random error of the model at low acid concentrations and may also indicate a weakness in the calibration owing to an offset. During the validation phase, a decrease in temperature coincided with low VFA concentration, and no clear distinction between the temperature effect and model stability at low concentrations could be elucidated. At VFA concentrations below 1 g kg^{-1} FM, Jacobi et al. (2009) also reported different behaviour to that of the estimates, and similarly, Reed et al. (2011) for a variety of substrates below 0.3 g L^{-1}. Hansson et al. (2002) previously suggested that this behaviour is related to the detection limit of the NIR method, which was found to be 0.3 g L^{-1} for propionic acid. For the data presented here, further research is required to assess the calibration stability at the lower model boundary for both temperature and VFA.

Short-term evaluation

The VS and VFA concentrations followed a similar pattern during the full validation phase and so did the NIR estimates. To investigate whether the estimation of VFA was independent of VS concentration changes, Figure 5.5 illustrates a more detailed time course between days 89 to 93 and days 144 to 148. In the first period, six different feedings ranging from 1.9 to 8.7 kg m^{-3} were administered. After each feeding, the estimated VS displayed an abrupt increase. Within less than one hour, the VS reached a new level and then steadily declined, reflecting the conversion of organic matter into biogas. The intensity of the feeding had a direct impact on the estimates of both parameters, and even the second feeding of 1.9 kg m^{-3} on day 91 was reflected in the NIR estimates (Fig. 5.6). As in the case of VS, the VFA estimates also showed a continuous build up in concentration level after feeding. However, when the VS concentration was already declining, the estimated VFA still continued to rise and reached its peak several hours later. This rise then began to drop comparatively fast, but did not reach the level prior to feeding. Therefore, owing to the high overall feeding, a steady accumulation of the acids was observed over this five-day period. In this regard, the NIR estimates nicely reflected the expectations of the process behaviour, both for the quite stable VS parameter and the highly variable VFA parameter. The high degree of independence of the two parameters during the initial hours after feeding is underlined by their consistently low correlation (r), which was calculated at each interval between the two feeding events (Fig. 5.5). Comparison of the estimated values with the reference values for both parameters demonstrates that the relative (precision) as well as the absolute (accuracy) changes in the concentration levels could mostly be captured via NIR spectroscopy. Thus, the accumulation of VFA could be observed even at early stages with 2 to 3 g kg^{-1} FM. The estimates for days 144 to 148 illustrate the reproducibility of the results of this work. Owing to a feeding of up to 20 kg m^{-3} VS, both parameters were affected more during this interval. Unfortunately, a malfunction due to plugging of the loop-circulation pump occurred on day 147, which affected the estimates. By comparing the VS concentration before and after feeding on days 144, 145, and 147, the sharp increase in the NIR estimates and in the reference values is less than the increase calculated from the actual amount of VS fed to the digester. This result is possibly explained by the formation of a floating layer after these heavy feeding events. This layer of fresh silage did not enter the loop system as its inlet was located at the bottom of the digester.

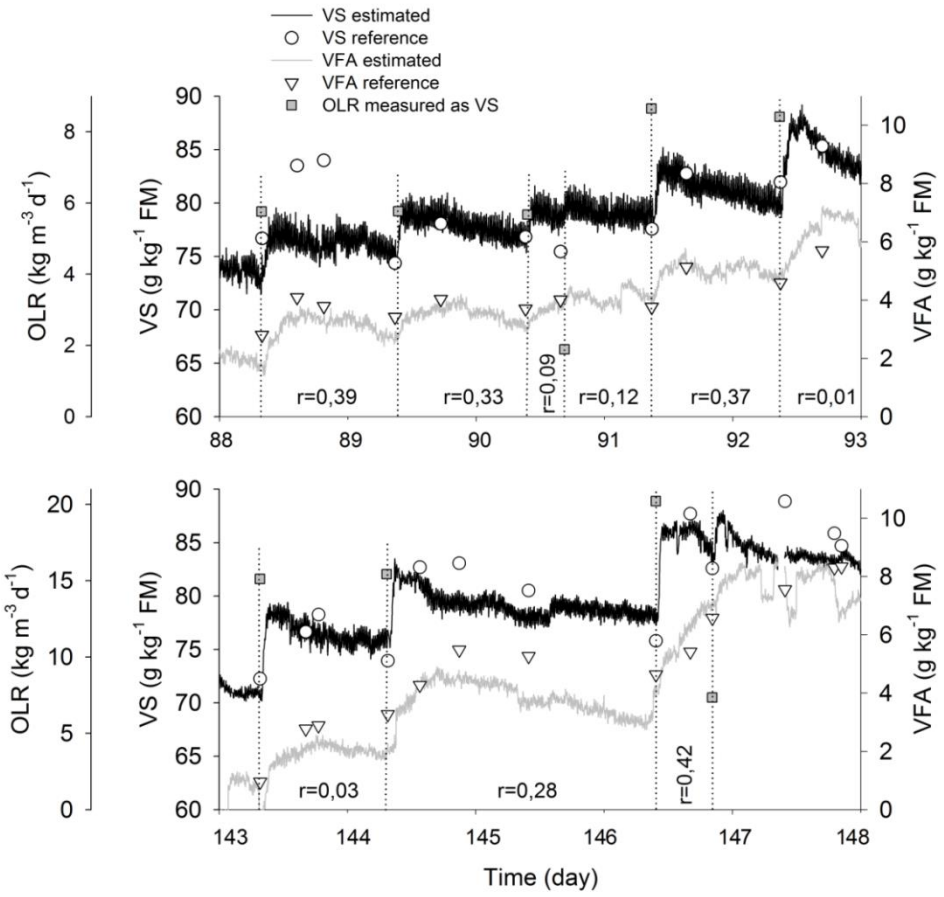

Figure 5.5: Short-term validation: NIR estimates of VS and total VFA for days 89 to 93 and days 144 to 148. Between two consecutive feeding events, the correlation coefficient (r) for the estimated values of VS and VFA were calculated. On day 148, plugging of the external-loop pump resulted in temporarily erroneous measurements.

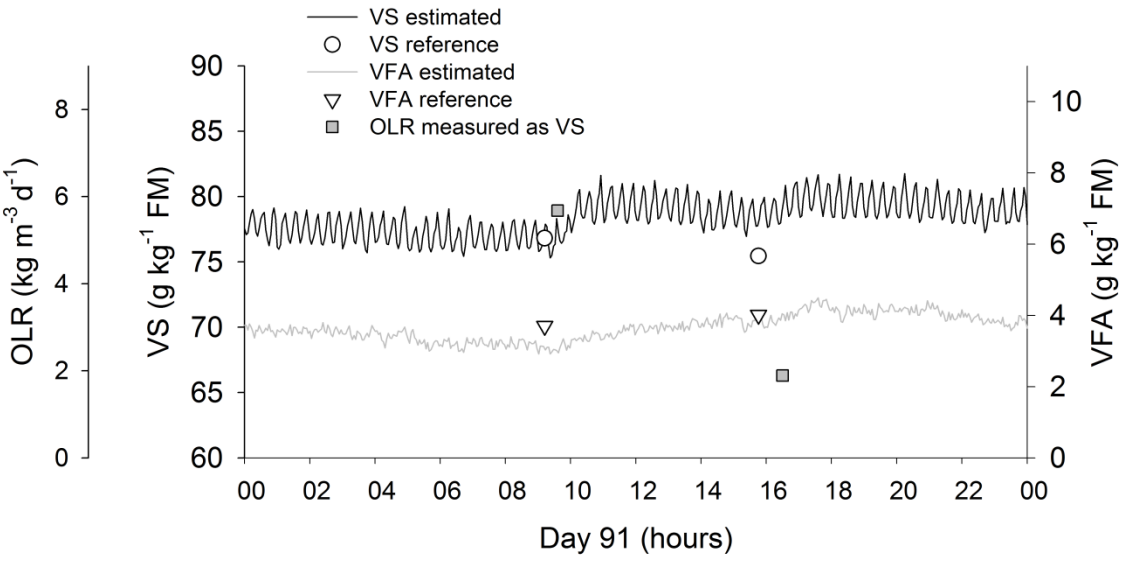

Figure 5.6: NIR estimates for VS and total VFA on day 91 showing two feeding events at different OLR.

5.2. Model expansion by inclusion of further feedstock

Abstract

The data set used in the previous experiment was extended by additional samples for the inclusion of further types of feedstock in calibration modelling. Fermentation runs were performed utilising a mixture of silage from energy crops and livestock manure for feeding. In addition to the recalibration of the models for VS and total VFA, models were also developed for NH_4-N, TIC, acetic acid, and propionic acid and were subsequently tested on a spectral time series obtained from a 45-day fermentation run. Spoiled silage was temporarily used for feeding in order to monitor the effect that the quality of the silage has on the process stability. The indication of long-term and short-term changes of the concentration of AD process parameters was confirmed after recalibration of the maize models with different feedstock. An accumulation of the total VFA can be detected by NIR in real-time, both in the case of excess feeding as well as low feedstock quality, and the dynamics of acetic and propionic acid were estimated independently from each other. Though the acid models did not yield accurate results, they sufficiently indicated the concentration courses over time, which can also be ascribed to the high repeatability of the NIR method. Since the time series estimates were more affected by noise compared to the mono maize application, smoothing of the estimates was necessary.

5.2.1. Material and methods

5.2.1.1. *Fermentation runs for collection of additional samples*

Digester A was operated for another eight weeks for the generation of additional samples for calibration after completion of experiment four. Feeding at an OLR of 2 kg m^{-3} d^{-1} (maize silage) was administered and the temperature in the digester ranged from 38-42 °C. For spectral recordings, the diode-array spectrometer was replaced by a second instrument of the same type (X-Three, NIR-Online GmbH). Spectral recordings and sampling (N = 17) were conducted as described in chapter 5.1.1.2. Prior to experiment four, the digester had been operated for several months, and the samples that had been collected with the first instrument during this earlier time period (N = 21) were also added to the updated data set. Digester B (chpt. 4.3.1.2) was also operated for an eight-week period. Material from a biogas plant fed with silage of maize, grass, and manure from cattle and poultry was used as inoculum. Discontinuous feeding with a grass-legume mixture was applied at an OLR ranging from 1 to 5 kg m^{-3} d^{-1} at 38-46 °C. For spectral recordings, the digester was equipped with the second diode-array spectrometer. During a 20 s sampling event (N = 30), slurry was removed from the outlet valve resulting in a 3-L sample of which a 0.5-L subsample was taken for reference analyses. Digester C (chpt. 4.3.1.2) was operated for 45 days. The inoculum was composed of 50% material taken from digester A which was mixed with slurry of a commercial biogas plant that had been fed with silage of maize and grass together with cattle manure. During the fermentation run, the temperature in the digester varied between 38-51 °C and discontinuous feeding at an OLR ranging from 1 to 8 kg m^{-3} d^{-1} was administered. At the beginning of the experiment (days 1–5), maize silage which had been stored in a plastic box for three weeks at about 25 °C was used as feedstock. This storage

resulted in a loss of the quality of the silage owing to the formation of mould. No feeding was applied between days 5 and 22. From day 22 onwards, a mixture of fresh silage of clover-grass, maize silage, and straw was used as feed. During the fermentation run, the digester material was partly substituted with cattle manure resulting in further diversification of the digester content. Unlike in experiment three (chpt. 4.3), only small amounts were exchanged at a time (max. 10%). On day 41, 20% of the digester content was substituted with acidified material from digester A that had been collected on day 200 of experiment four (see Fig. 5.1). Digester C was equipped with the first diode-array spectrometer and the permanently recorded spectra were averaged for a 3-min interval (see chpt. 5.1.1.2). During a 20 s sampling event (N = 65), slurry was removed from the outlet valve resulting in a 3-L sample of which a 0.5-L subsample was taken for referencing. This additional sample collection from the three digesters together with the samples of experiment four resulted in N = 252 samples available for the updated data set.

5.2.1.2. *Model development and time series validation*

Multivariate data analysis was performed using the SX-Plus software (NIR-Online GmbH). In the first step, the spectra were subjected to a standard pre-treatment implemented in the software in order to account for differences in the instrumental responses. Although the first and the second spectrometer were identically constructed, their instrument wavenumber calibrations and line shapes varied. This is because the elements of a diode array are statistically aligned along the wavenumber range of the spectrometer. This made it necessary to recompute the absorbencies determined by measurement to a fixed scheme in order to allow for combination of data from the two instruments. Therefore, the spectra obtained from the instruments were mathematically corrected for synchronization of their responses and the spectra were recalculated to a vector representing the absorbencies with a spacing of 10 nm. A PCA with six PCs was computed separately for the subsets of each digester using the 1000-1750 nm region of the first derivative spectra for detecting spectral outliers. One sample of digester B was defined as outlier because of an *MD* >4 and was removed from the data set. Subsequent PLS regression for the parameters VS, NH_4-N, TIC, total VFA, acetic and propionic acids was performed for the same spectral region. For the three digesters, the samples were sorted in the chronological order in which they were collected and every third sample was used for validation; the other samples were used for calibration. Hence, the calibration and the validation sets both contained samples from the three digesters. Pre-treatments including first and second derivative and SNV transformation were tested, and the best algorithm was selected on the basis of the *RMSEP*. The number of PLS components was chosen as proposed by the software based on the impact of each new component on the reduction of the *PRESS* calculated for the internal test set validation. The calibrations selected for the six parameters were applied to the continuously measured spectra of the fermentation run of digester C. It was examined whether NIR spectroscopy correctly described the long-term trend in the concentration level of the parameters over this 45-day period. Similar to experiment four (chpt. 5.1), the ability to follow the short-term dynamics was also tested by focusing on a more detailed picture of the time series estimates. Therefore, the time periods of days 1–8 and days 29–36 were investigated further.

5.2.2. Results and discussion

5.2.2.1. Reference data set characteristics

For VS, a concentration range from 53 to above 100 g kg^{-1} FM was observed for the training data set; NH_4-N showed values from 1.2 to 4.4 g kg^{-1} FM (Tab. 5.2). For these parameters, the samples with the highest concentration were removed from digester B; samples with lower concentrations originated from digester A. Alkalinity, as described by the TIC, varied from 5 to 19 g kg^{-1} FM. Samples that tended to have high TIC also reported larger values for NH_4-N. This relationship was less pronounced, as shown by comparison of the correlation coefficient (r) in Table 5.3 with those from the other experiments (Tab. 4.3, 4.7, and 4.12, except for the online calibration set of experiment three, which also showed a low r). With concentrations of up to 16 g kg^{-1} FM, the total VFA range was broader compared to the earlier experiments. The *SEL* was calculated from the duplicated reference analyses (see chpt. 3.3.1) to assess the contribution of the reference error to the overall error estimates of the NIR models (Tillmann, 1996). Overall, a similar precision of the reference method was reached as compared to the results from the previous data sets (compare with Tab. 3.4 and 4.2, 4.6, 4.11 and chpt. 5.1.1.2).

Table 5.2: Descriptive statistics of the reference analyses for the training data set including samples obtained from digesters A, B, and C (N = 252).

Parameters	Range	Mean	Median	*SD*	*SEL*
VS [g kg^{-1} FM]	53.4–103.2	73.3	71.8	9.5	1.7
NH_4-N [g kg^{-1} FM]	1.18–4.43	2.29	2.12	0.60	0.08
TIC [g kg^{-1} FM]	4.70–19.24	9.59	9.58	2.98	0.13
Total VFA [g kg^{-1} FM]	0.19–16.01	4.90	4.11	3.67	0.28
Acetic acid [g kg^{-1} FM]	<0.10–11.06	3.15	1.77	3.32	0.20
Propionic acid [g kg^{-1} FM]	<0.03–4.69	1.87	1.91	1.53	0.12
Temperature [°C]	31.2–51.4	40.4	40.0	2.7	-

For a full list of abbreviations see Table 4.11

Table 5.3: Coefficients of correlation (r) among the parameters VS, NH_4-N, TIC, Total VFA, Acetic and Propionic acid calculated for the training data set.

Parameters	VS	NH_4-N	TIC	Total VFA	Acetic acid	Propionic acid
NH_4-N	0.31	1.00				
TIC	0.25	0.65	1.00			
Total VFA	0.18	0.17	-0.54	1.00		
Acetic acid	0.06	0.15	-0.54	0.93	1.00	
Propionic acid	0.16	0.03	-0.49	0.73	0.49	1.00
Temperature	-0.13	0.02	-0.18	0.17	0.19	0.07

5.2.2.2. PLS calibration results

The PLS-calibration results for the NH_4-N and TIC parameters demonstrated satisfactory estimation performance as indicated by the *RPD* (Tab. 5.4). The overall similar fit of the calibration and validation samples is illustrated by the plots of estimated versus measured values in which the values closely cluster along the regression line (Fig. 5.7). For VS, the *RMSEP* was somewhat higher compared to the *RMSEC*. An *RMSEP* of 3.8 g kg^{-1} FM showed accuracy similar to that reported for the VS results of the first experiment (see Tab. 4.4); the *RPD* was comparably low owing to the smaller *SD* of the validation set. The *RMSEP* for the three acid parameters was similar to the *RMSEC*. For total VFA, most samples fitted well to the regression line; however, a few calibration samples were clearly underestimated at the lowest and highest

concentrations. For acetic acid, samples at lower concentrations (<1 g kg^{-1} FM) deviated significantly from the reference value compared to samples in the medium-to-high concentration range. A similar observation was made for propionic acid where stronger deviations from the expected values also occurred at higher concentrations. As shown in the plot equation, the three acid models were offset, i.e., the samples at low acid concentrations were overestimated. Further, the *RPD* values of the three acid parameters were different due to the different concentration ranges. With an *RPD* of <2, the propionic acid model showed a notably low performance. The *SED* was calculated using the time series estimates of digester C and two consecutive NIR estimates were treated as if they were duplicate measurements of one sample in order to exclude the effect of a change in the concentration over time (see chpt. 4.3.2.4). Comparison with the *SEL* of the reference methods showed a higher precision of the NIR method for VS and NH_4-N; similar repeatability was found for the other four parameters.

Table 5.4: PLS model results for the parameters VS, NH_4-N, TIC, Total VFA, Acetic and Propionic acid after splitting the training data set into subsets used for calibration (2/3) and validation (1/3).

Parameters [g kg^{-1} FM]	PLS components	*RMSEC*	*RMSEP*	Bias	*RPD*	*SED*
VS	7	3.16	3.80	0.18	2.5	0.4
NH_4-N	5	0.14	0.12	0.03	4.6	0.01
TIC	6	0.69	0.77	0.03	3.6	0.16
Total VFA	6	1.14	1.09	0.23	3.3	0.18
Acetic acid	8	1.30	1.36	0.05	2.4	0.15
Propionic acid	7	0.76	0.87	0.03	1.8	0.09

For a full list of abbreviations see Table 4.4

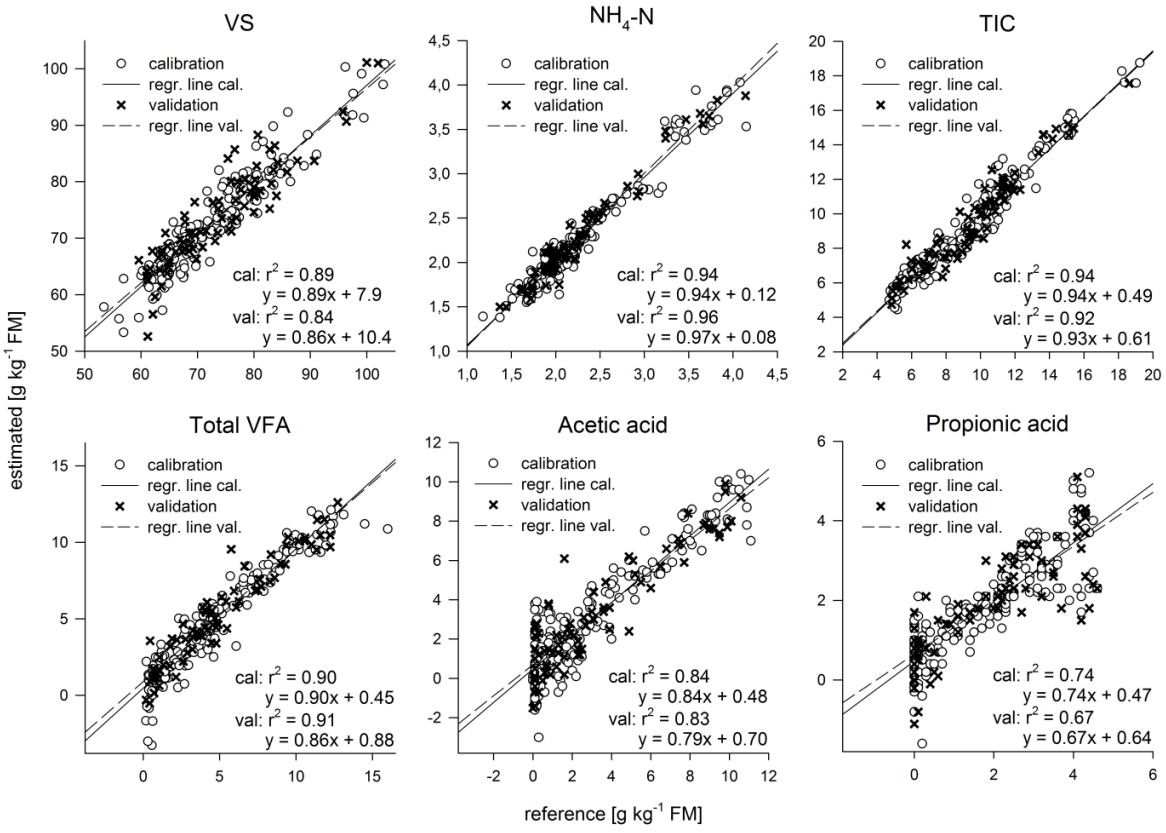

Figure 5.7: PLS-calibration results for the parameters VS, NH_4-N, TIC, Total VFA, Acetic and Propionic acid. The estimated vs. measured plots relate to the models of Table 5.4.

Comparison of the model results for VS and total VFA with those from experiment four (mono maize; Fig. 5.3) revealed the accuracy of the estimates to be slightly lower. For VS, the number of factors of the updated model was higher compared to the maize model. This most likely can be explained by the increase in the spectral variability of the updated training set. Comparison of the results from the PLS model with those from the internal test set validation of the first experiment (Tab. 4.4 and 4.5) reported similar error values, with an exception for acetic acid indicating a loss in accuracy of the estimates. The overall similarity of these results suggests that the use of different training data sets and spectrometer systems had only a moderate impact on the accuracy of the estimates. This also highlights the possibility for the development of robust calibrations because the results were similar under different experimental conditions. Since for the two experiments, a similar 'within-laboratory precision' (*SEL*) was given, the contribution of the reference error to the overall error of the NIR model was of the same magnitude. This further supports the assumption of only a moderate impact of the spectrometer systems and measuring modes on the model performance. The observed differences in the error estimates for the acid parameters may also be affected by sample-set-specific reasons, e.g., differences in the concentration range. For propionic acid, the higher error estimates of the samples at lower and higher concentrations may indicate a non-linear behaviour, and these observations were not made in the earlier experiments. Overall, the PLS-modelling results of experiment five also confirmed the limited potential of an accurate estimation of the acid parameters in the concentration region relevant for monitoring AD, i.e., at an early stage of acid accumulation.

5.2.2.3. *Time series validation*

Long-term trend

Feeding at the beginning of the operation of digester C tended to cause an increase in the VS concentration as was indicated by both the reference values and NIR estimates (Fig. 5.8). When feeding was temporarily stopped on day 5, a decreasing trend was observed. For days 13 and 14, deviation of the estimated values from the reference values was apparent. After feeding was restarted on day 22, the VS concentration course was mostly correctly described by the NIR estimates with an overall increasing tendency. This increase was accelerated at the end of the fermentation run which coincided with the higher OLR of up to 8 kg m^{-3} d^{-1} on day 43. From day 43 onwards, the VS reference values did not follow this trend, but rather clustered near 80 g kg^{-1} FM. The reference values for NH_4-N at the beginning of the fermentation run ranged below 2.5 g kg^{-1} FM with a slight upwards trend until day 12, which was also shown by the NIR estimates. When N-rich material was used for feeding from day 22 onwards, this trend was more pronounced, with the reference values reaching nearly 3 g kg^{-1} FM. The decrease of the NH_4-N concentration at the end of the fermentation run coincided with the increased OLR, i.e., the addition of larger amounts of fresh silage to the digester. Overall, the NIR estimates accurately followed the reference values and only a few samplings indicated moderate mismatches. Visual inspection of the NH_4-N estimates showed a stable behaviour, i.e., a high repeatability, which corresponded to the low *SED* value in Table 5.4. At the beginning of the fermentation run, total VFA indicated a strong increase from around 2 g kg^{-1} to 12 g kg^{-1} FM (day 8), though a moderate OLR was chosen and was stopped on day 5. After substitution of digester content with cattle manure and a pause in feeding, total VFA concentration steadily decreased from day 9

onwards. This decrease was accompanied by a reverse trend observed for the TIC reflecting the recovery of the system after acid accumulation. Similar to total VFA, the results for the estimates of TIC mostly followed the concentration trend of the reference over time. This reverse behaviour of the two parameters was also observed on day 9 and day 22 when digester content was exchanged for manure. Whereas a sudden increase of the TIC estimates was observed, which corresponded to the higher buffer capacity of the added slurry compared to the material in the digester, the total VFA estimates decreased. A reverse trend was shown after the addition of acidified maize slurry on day 41, which was poor in TIC. This time, the TIC estimates showed a concentration drop of around 2 g kg^{-1} FM, whereas the total VFA increased by the same order of magnitude. Since acetic acid contributed most to the total VFA concentration, the concentration course of these two parameters was very similar and this single acid was, for the most part, correctly estimated by the NIR model over the 45-day period. For propionic acid, the concentration course was less dynamic compared to that for acetic acid with values ranging from 0 to 3.5 g kg^{-1} FM. After recovery of the digester from the high acidification dominated by acetic acid, the propionic acid level remained constant at 2 g kg^{-1} FM until the end of the period, which was accurately followed by the NIR estimates. As drawn from visual inspection, the propionic acid signal showed a smoother, i.e., less noisy, behaviour compared to the other two acid parameters and this observation corresponded to the smaller *SED* value of this acid as shown in Table 5.4.

The time series comparison of the long-term behaviours of the six parameters with those of experiment three (Fig. 4.22) showed high similarity, i.e., the trend of the concentrations was more or less correctly indicated by the NIR estimates in both cases. Comparison of the *SED* value for the time series estimation of the two experiments demonstrated an overall similar repeatability for the parameter estimates. For both experiments, NH_4-N indicated a very high repeatability which suggests that the NH_4-N signal in the spectra is highly specific, an assumption already concluded from the results of the validation of the lab models (chpt. 4.2). The two different online process analysers (Matrix-F, Bruker Optics; X-Three, NIR-Online GmbH) used in this work differ in their technical specifications. These overall similar results suggest that instrument-specific aspects only had a small impact on the suitability of the method for online monitoring of AD. These results also suggest that even a limited spectral range covering the overtone regions, such as those used for the spectrometer system in experiments four and five, results in satisfactory NIR models under the conditions examined in this work. However, direct comparison of the time series of experiments three and five is limited since the results were obtained from different validation procedures. For the former experiment, time series validation was done with slurries of different origin than the material used for the calibration. For the latter experiment, slurry of digester C used for validation was also used for model development, which may also explain the absence of any relevant bias reported for experiment five. Note that the aim of this work was not an explicit comparison of the two types of online process analysers. Such comparison would have required that both systems be run in parallel at one digester and that the same samples be used to exclude sample-specific factors influencing the model performance.

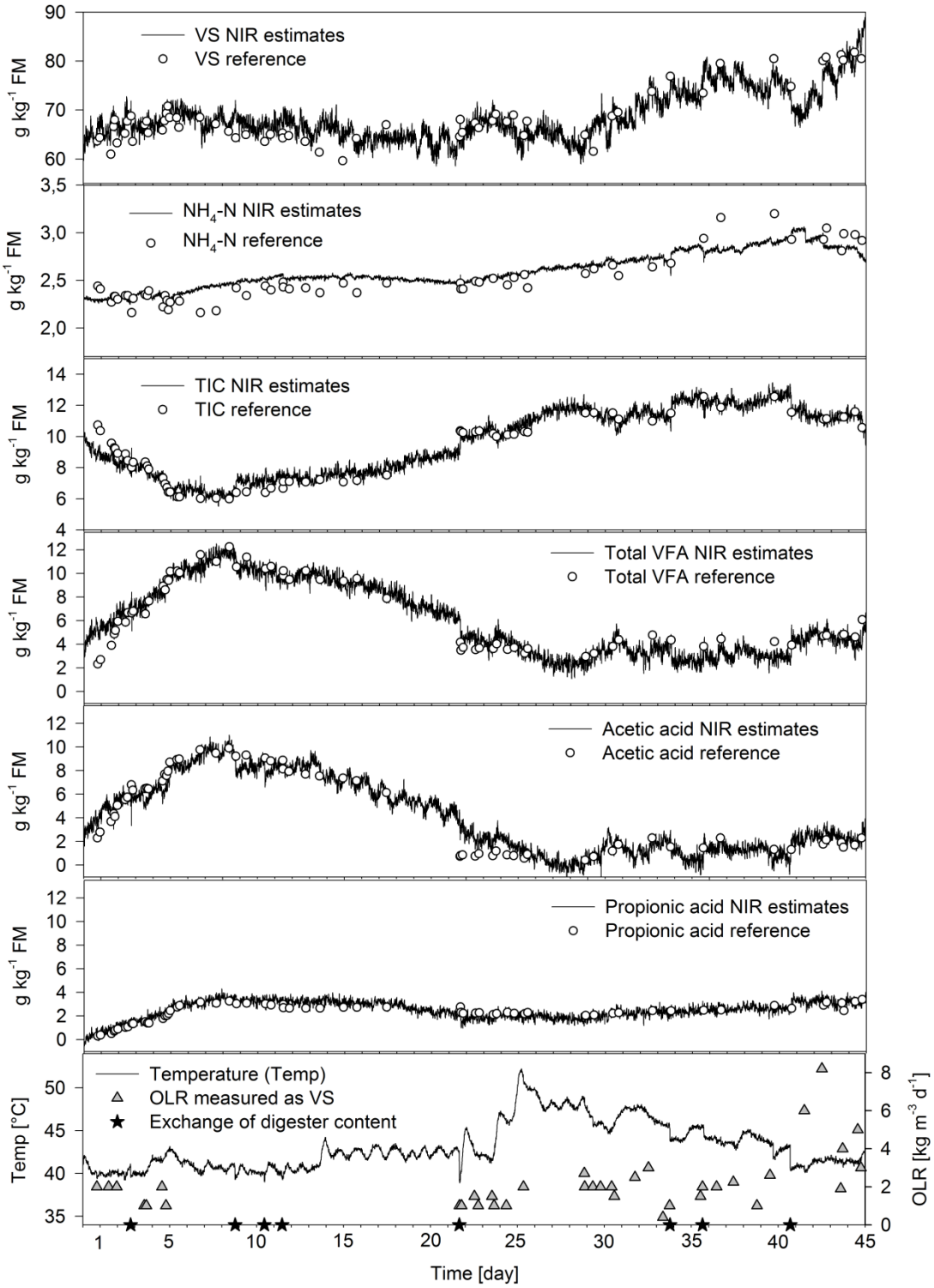

Figure 5.8: Results from time series validation of the six parameters from days 1 to 45 of the online measurements of digester C; the NIR spectra were averaged over a 3-min interval. The temperature of the digester during the fermentation run is plotted together with the OLR. The stars indicate the times at which the digester content was partly exchanged (see text for further details).

Short-term dynamics

The time series validation for days 1–8 (Fig. 5.9) and days 29–36 (Fig. 5.10) illustrate the parameter estimates in more detail. The VS signal responded to the feeding events; the estimates showed an increase in the VS concentration in the digester. However, this VS signal change after feeding at an OLR near 2 kg m^{-3} d^{-1} was masked by noisy structures and, at times, the VS estimates indicated fast drops by a few g kg^{-1} FM, e.g., shown on day 7 and day 30. These drops occurred without manipulation of the digester (note that the sudden VS drop on day 3 coincided with the exchange of digester content with material with a lower VS concentration). For the acid parameters, the continuous increase in their concentration levels from days 1–8, as stated by the reference, was reflected fairly accurately by the estimates. An overestimation at the beginning of the experiment was observed for total VFA and this overestimation coincided with an underestimation of the TIC estimates. The concentration of total VFA and propionic acid showed a rather monotonous upwards trend which was mostly disconnected from feeding, i.e., it did not change between two feeding events. This upwards trend continued after feeding was stopped on day 5. A slight decreasing trend for the acetic acid estimates was observed between two feeding events shown for days 2 and 5. However, the assignment of this observation to feeding is clearly limited by the noise in the estimates. The short-term period between days 29 and 36, i.e., after the recovery of the digester from the strong acid accumulation, was characterised by frequent feeding and the total VFA and acetic acid estimates responded to these events. When feeding was applied twice a day (days 29–31), the acidity level in the digester increased. Although total VFA and acetic acid mostly showed a similar concentration course, on day 31, differences in their estimates were obvious. A decrease in the estimates of total VFA and acetic acid was observed when feeding was paused until day 32. Different from this dynamic concentration course, the propionic acid estimates were near 2 g kg^{-1} FM, indicating that they were not affected by feeding.

Time series validation of day 32 (24 h) visualises the noise level in the estimates in more detail (Fig. 5.11). In addition to the background noise inherent to the VS estimates, sudden changes in the NIR signals in the order of several g kg^{-1} FM are visible (e.g., at around 5 a.m. and from 9 a.m. to 1 p.m.). Such abrupt changes were observed for the other parameters as well, and some occurred at the same time (e.g., the total VFA and acetic acid estimates at around 3 p.m.). The estimates of day 32 were treated with a median filter in order to test the possibility of reducing these structures from the NIR estimates by smoothing. Window sizes of five and 25 were chosen, i.e., each original data point of the spectral time series was replaced by the median of five and 25 neighbouring estimates (the window was centred around the data point). These window sizes corresponded to measuring times of 15 and 75 min according to the 3-min data points. The narrow window size of five resulted in smoothing of the background noise but had no effect on the abrupt changes; this can be expected since the ‘median filter is an effective method that can suppress isolated noise without blurring sharp edges’ (Wang, 2011). A filter with a window size of 25 also reduced these ‘structures’ more effectively. At the same time, this stronger filtering did not blur the effect of the feeding of 2.5 kg m^{-3} d^{-1} on the estimates as shown after the feeding at 6 p.m. By comparing the VS estimates before and after feeding, it is clearly visible that the feeding-induced increase in the NIR estimates is higher than the increase calculated from the actual amount of VS subjected to the digester.

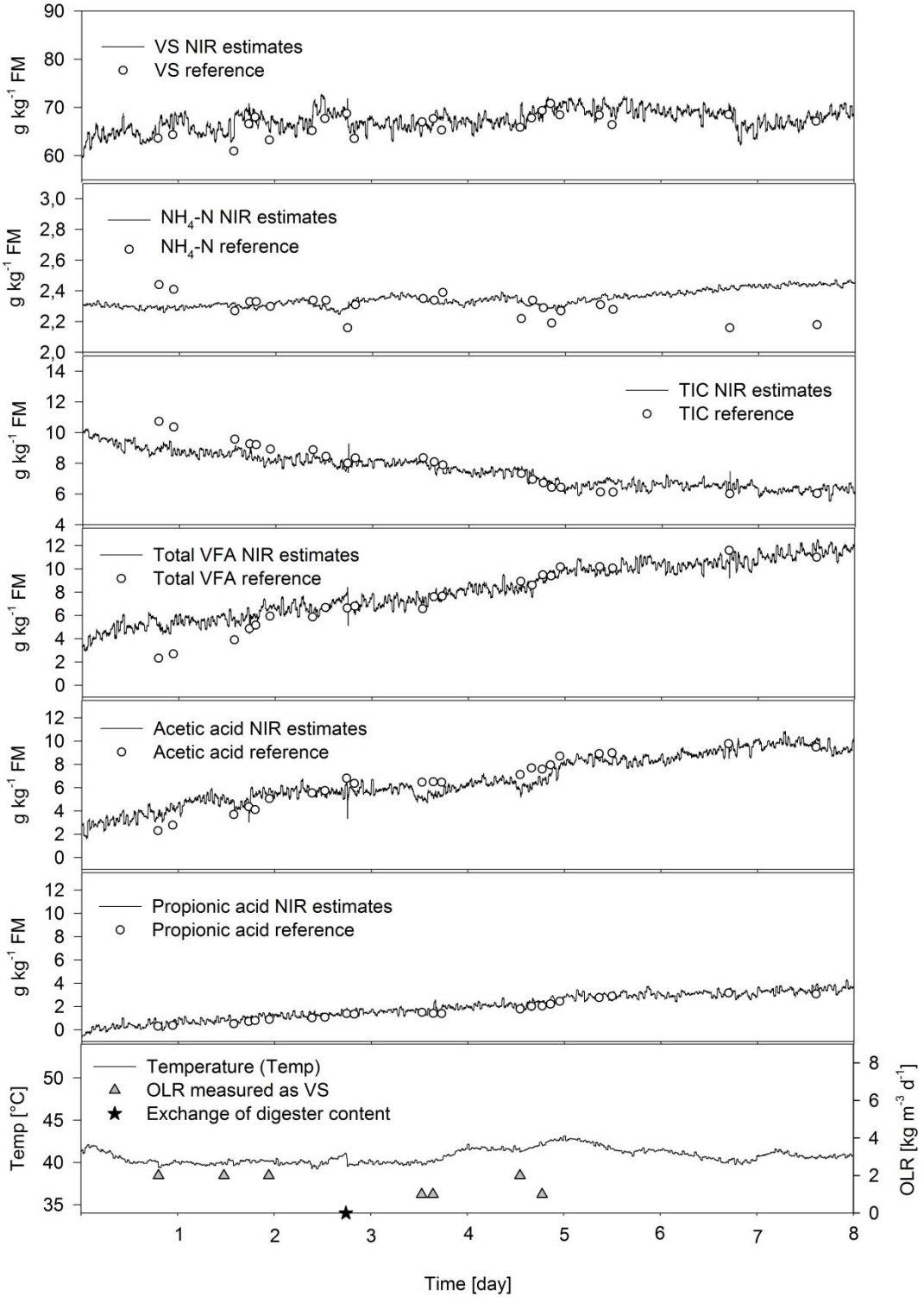

Figure 5.9: Results from time series validation of the six process parameters from days 1 to 8 of the online measurements of digester C; the NIR spectra were averaged over a 3-min interval. The temperature of the digester during the fermentation run is plotted together with the OLR. The star indicate the times at which the digester content was partly exchanged (see text for further details).

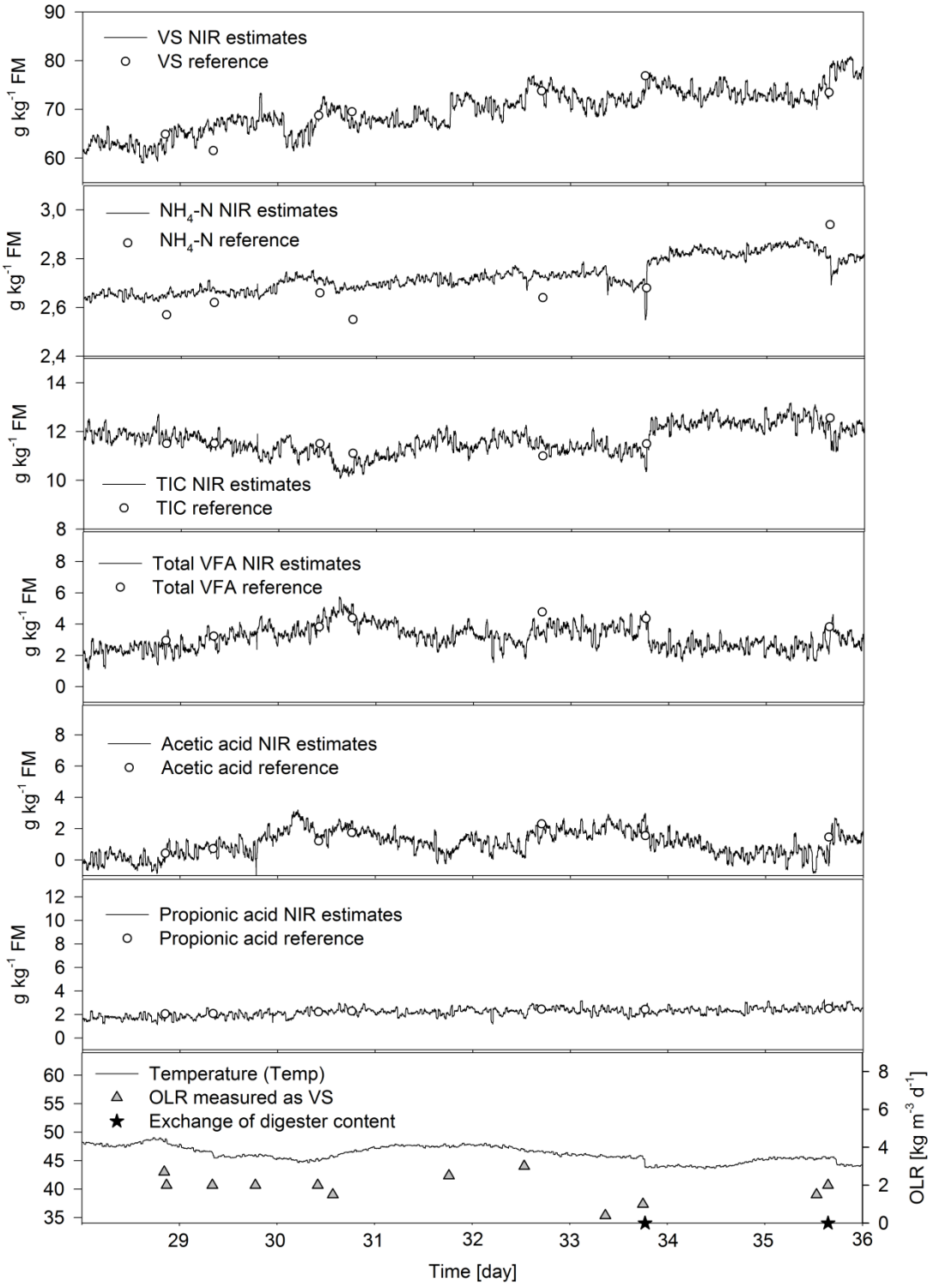

Figure 5.10: Results from time series validation of the six process parameters from days 29 to 36 of the online measurements of digester C; the NIR spectra were averaged over a 3-min interval. The temperature of the digester during the fermentation run is plotted together with the OLR. The stars indicate the times at which the digester content was partly exchanged (see text for further details).

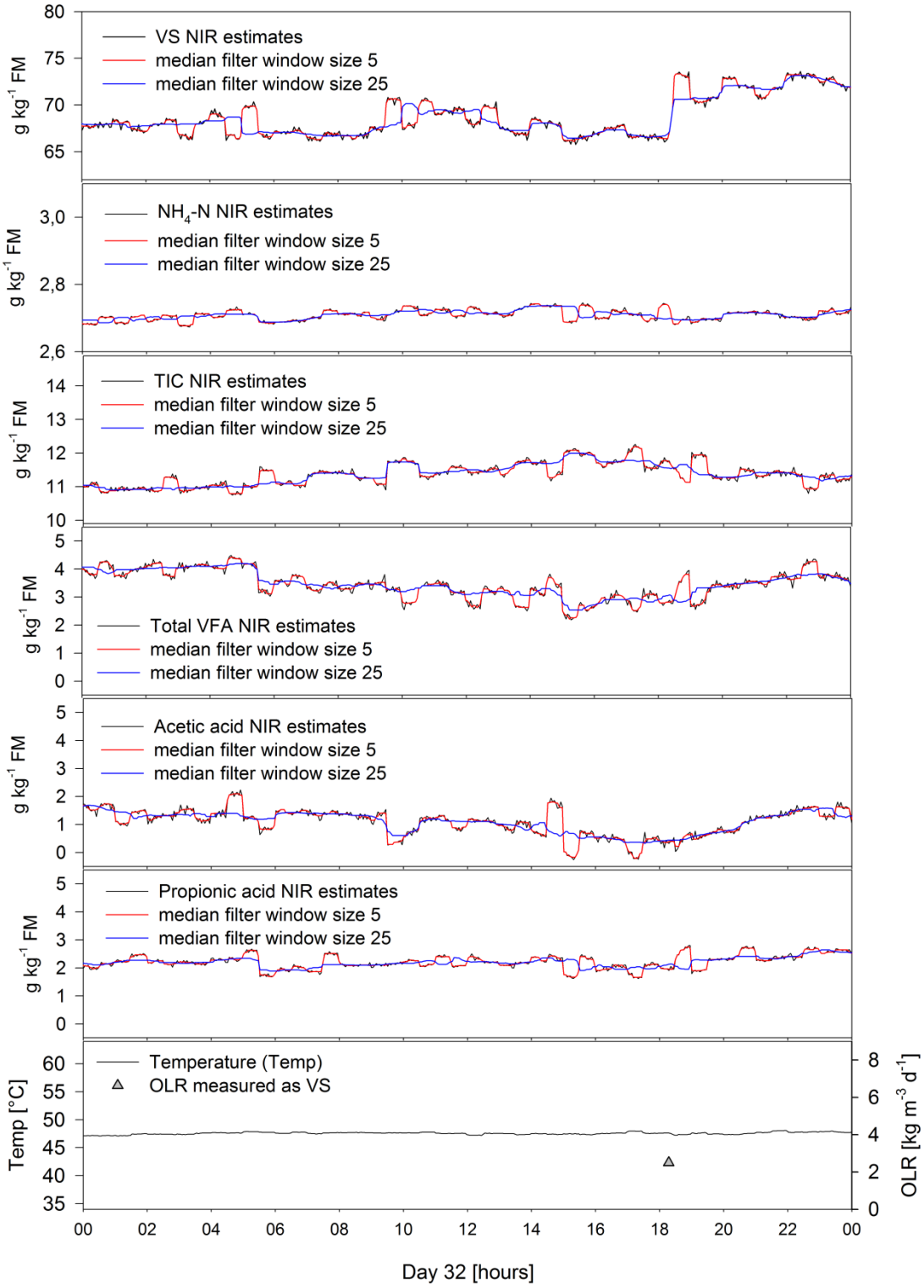

Figure 5.11: Results from time series validation of the six process parameters for day 32 (24 h) of the online measurements of digester C; the NIR spectra were averaged over a 3-min interval. Two median filters with window sizes of five (15 min) and 25 (75 min) were tested for noise reduction.

The results of the selected time periods of experiment five demonstrated that the short-term dynamics of the total VFA can be indicated with the updated model. However, compared to experiment four, overall lower OLR were applied in this experiment and, hence, the effect of feeding on the short-term process dynamics, e.g., an acid increase, was less pronounced. In addition, the concentration course of the parameters were disturbed by noise and sudden changes in the estimates, and these fluctuations did not appear as the random error observed in the time series estimates for VS and total VFA on day 91 of experiment four (see Fig. 5.6). Reasons for these disturbances are not yet clear. Jacobi et al. (2009) also reported sudden shifts to a different level for VFA estimates and ascribed these changes to the long time span between white referencing. For the present study, white referencing was done frequently on an automatic basis; thus, this aspect most likely does not explain the observed shifts that occurred in close succession. Lomborg et al. (2009) also reported a fast downward trend for acetic acid from 11 g kg^{-1} by approximately 3 g kg^{-1} FM, followed by a sudden shift back to the previous level. The authors did not ascribe these changes to an error of the model and instead presumed that these sudden shifts reflected the actual concentration changes of this acid. It is likely that this explanation does not hold true for the shifts observed in the present study. Presumably, these shifts indicate that non-specific information was included in the updated PLS models which caused small changes in the spectra to result in these offsets in the estimates. This was probably caused by the inclusion of different feedstock material in the calibration, which may have caused an increase in the physical inter-sample variation compared to the mono maize slurry. As a consequence, the PLS algorithm was less able to separate the X responses of the analyte and the interferent which resulted in inclusion of noise in the calibration (Martens and Næs, 1991). It appears necessary to further address this aspect of unexplained signal changes which, e.g., may also be due to technical aspects. This also includes inspection of whether the training data set allowed for an adequate pre-treatment of the spectral raw data for synchronisation of the two spectrometer responses. Overall, these disturbances in the estimates are not believed to limit the use of the models for AD monitoring since these shifts did not mask the concentration dynamics to a large degree. A filter can be used to reduce noise-related variations from the estimates; however, the selection of an adequate window size for smoothing is a critical aspect. If the chosen window size is too wide, the estimates will not reflect short-term dynamics of the parameters, e.g., the VFA. For the data presented here, a median filter over an interval of 75 min was suitable for removing noisy structures from the estimates while preserving real changes caused by feeding.

The reason for fluctuations of the VS estimates may be due to temporary inhomogeneity in the digester. The feed was added to the digester from the top and, hence, the VS level in the upper part of the digester was temporarily highest. Spectral recordings took place halfway up the side of the digester in the upper part of the filling level (see Fig. 4.15e). This over proportional increase in the VS estimates directly after feeding (e.g., during the first days) followed by a comparably fast decrease may reflect the retarded spatial distribution of the feed in the digester. Similar to the time series validation of the mono maize models, at times, the VS estimates differed from the reference values. With respect to the mismatch from day 43 onwards, the increasing NIR estimates probably reflected the ‘true’ concentration course of VS more realistically. This is because the increased trend of the estimates nicely coincided with the

stronger feeding and the constant reference values appear implausible. This finding may again be ascribed to sampling errors as discussed earlier for experiment four.

Unlike the mono maize fermentation where process imbalances were induced by excess feeding, the OLR chosen for experiment five was moderate. This suggests that the increase in the acidity level at the beginning of the fermentation run (Fig. 5.9) was not due to the intensity of feeding, but was more likely associated with the quality of the feed; in this case, mould silage. It has been reported that the use of poor-quality silage can negatively affect the activity of the microbial community which may result in accumulation of VFA, e.g., by the entry of toxic compounds such as mycotoxins into the digester (Preißler, 2010). The increase in the total VFA concentration during the first days, which also continued when feeding was stopped, supported the assumption of a strong uncoupling between acid producers and acid consumers in the digester. This was different than described for experiment four where a clear decrease between two feeding events was evident, indicating an active consortium of acid-degrading microorganisms. This observation supports the assumption that the NIR technique can not only be used for detection of instabilities caused by overfeeding, but also for instabilities induced by inhibitory compounds entering the digester. This was visualised by a VFA accumulation together with a parallel decrease of the buffering capacity.

Interpretation of the NIR estimates of the single acids requires consideration of their strong collinearity with the total VFA as previously discussed in more detail in this work. The findings of this experiment are relevant because of the different behaviour of total VFA and acetic acid (more dynamic) and the more stable behaviour of propionic acid throughout the fermentation run of digester C. This observation, which is particularly shown for the period from day 29 to 32, gives rise to the assumption that propionic acid can indeed be independently estimated. In this respect, the significance of the result of experiment five goes beyond those of experiment three (chpt. 4.3) where no such conclusion could be drawn since the concentration courses of acetic and propionic acid were very similar over the complete fermentation run.

6. Major outcomes and suggestions for future work

Different sample sets were used in this work for the calibration of PLS models which were later evaluated with different validation data sets. This approach should prevent misleading conclusions from being drawn from single calibration results that may be influenced by data set specific (spurious) correlations, e.g., among the chemical constituents. Although moderate differences in the performance between the models of the individual experiments were observed, their overall similar results proved the robustness of the relationship between the spectral and chemical readings for VS, NH_4-N, TIC, and total VFA presented in this work. The use of different data sets had only a moderate impact on the performance of the estimates and this finding highlights the possibility for the development of robust calibrations because the results are reproducible under varying conditions. For the single acids, high collinearity with the total VFA made the development of specific acid models a potentially difficult task since the single-acid models were possibly closely related to spectral information describing the total VFA concentration rather than those of a single acid. The negative implication this can have on the model robustness was discussed in this work. Artificial samples, e.g., those spiked with an analyte, are often not useful for development of a NIR model (Shenk et al., 2008). This is because the spectra of such samples usually do not correspond to the complexities of natural samples (Esbensen, 2009). However, the propionic acid model based on calibration samples with naturally formed acids allowed estimation of slurry spiked with this single acid. For the present case, this finding suggests that the spiked samples corresponded spectrally to samples with naturally formed intermediates. Therefore, for future calibrations, it may be promising to also work with spiked samples for calibration in order to avoid collinearity among the acid parameters, e.g., by planned experimental design (Kessler, 2007). This would also enable more even distribution of the samples along the model's concentration range and would meet this basic principle on a calibration model (Næs et al., 2002).

Cross-validation allowed for an objective comparison of the different pre-treatments and X-variable selection alternatives, which was empirically tested via the optimisation routine during the first experiment. The comparison of the offline calibration results among the training sets revealed that an increase in the heterogeneity of the sample matrices did not result in a relevant performance loss of the NIR models. The splitting of the training set into subsets used for calibration and validation demonstrated that the number of samples sufficiently represented the spectral variation within each training set. As discussed by Esbensen and Geladi (2010), these two validation techniques do not allow for a realistic assessment of the 'future prediction potential' of the model. Since the validation samples were essentially drawn from the same data set used for calibration, they do not reflect an independent situation. According to the theory of sampling (TOS), a proper validation must take into account statistical issues as well as chemical and physical sampling error issues (Esbensen and Geladi, 2010). With respect to the latter aspect, realistic assessment of the model performance for future situations requires information of the sampling error, which changes with every new sampling from heterogeneous materials. From this, it follows that only an independent (second) sampling provides information about the conditions that may have changed between sampling the training data set and the future working situation of the model. This independent data set used for validation was composed of material

from agricultural biogas plants that were not used for calibration because the aim was to test the models on an open population. The selection of these samples used for validation was based on the idea that they should be 'similar' with respect to the intended future working conditions of the models. These samples used for the second experiment were obtained from biogas plants located within a 100 km radius of Freising. Hence, they represented a population of biogas plants from a local domain. Because spectral variability of agricultural products can also be related to regional effects (Shenk et al., 2008), it is of interest to further test the models on digester slurry from agricultural biogas plants located in different regions. For instance, one such effect could be the influence of different cultures or animal nutrition on the spectral characteristics of the slurry.

The lab application proved its suitability as a tool for satisfactory screening of the parameters VS, NH_4-N, TIC, and total VFA. However, the poor accuracy of the NIR analysis compared to the reference method may not justify its use for routine analysis in the lab. For an at-line application (i.e., an offline use directly at the biogas facility), alternative quick tests such as the test for alkalinity are available and these methods may be more economically feasible (see Hecht, 2009). However, the benefit of the non-destructive NIR analysis is its potential to determine several parameters simultaneously. Thus, an offline use may be justified if several parameters are to be assessed, e.g., for a pre-screening for the selection of samples to be analysed further by classical methods. In addition, if the matrix-specific error of the slurry is known, the NIR analysis can be corrected for this systematic error, which may improve the significance of the estimates. This especially holds true for parameters with highly specific signals in the NIR region, e.g., as in the case of NH_4-N. From this, it follows that NIR analysis may be particularly suitable for the regular monitoring of digesters with known process history.

Similar to the wet chemical method, the offline NIR estimates only reflect a given moment in time, and, hence, do not cover the process dynamics over an extended period. This principal limitation was the motivation for transfer of this application to an online process analyser and the time series results indicated that an in situ use is the true value of the NIR method for AD. This is because the estimates allow for continuous indication of the concentration courses of key parameters, including their long-term trends and short-term dynamics. Furthermore, such in situ measurement is the prerequisite for using the NIR-generated data for AD control by adjustment of the intensity and frequency of the feeding to the actual process conditions. Such control strategies can be achieved manually or can be based on an automated approach (see references in Boe, 2006). For an automated control, critical values must be set for the process output variables, e.g., the total VFA. The manipulated variable (organic load, feed rate) can then possibly be regulated to achieve the desired value of the output variable (Boe, 2006). These critical values can be based on relative changes rather than on absolute values. This is because 'different anaerobic systems have their own 'normal' levels of VFA, determined by the composition of the substrates' (Ahring et al., 1995; Angelidaki et al., 1993). Since the relative changes are the relevant criteria here, e.g. the VFA increase in between feeding events, matrix-specific errors do not constitute a hindrance for an AD control by NIR. However, the success of the method strongly depends on the stability of the estimates against erroneous drifts and sudden shifts. Therefore, future work has to critically address the suitability of the NIR estimates,

particularly with regard to their ability to provide a stable indication of the short-term dynamics at an early stage of process disturbance.

The results of experiment three confirmed the finding of the lab approach that a limited number of feedstocks used for calibration already reflect the spectral variability necessary for the development of feedstock-robust applications. This finding has a practical implication since it suggests that it is unnecessary to perform a laborious digester-specific calibration each time and that changes in the feedstock material can be accounted for by a single model. However, it would be false to conclude that these models are expected to include all the information necessary for estimation of agricultural digester slurry in the future. In NIR analysis, no model is usually able a priori to include all information that will be necessary in the future and this is especially true if the models are used in an open environment. New sources of variation will occur and these variations have to be included in the calibration database; thus, the calibration model must be updated (Bouveresse and Campbell, 2008). Therefore, the presented models should be understood as a starting point for the development of an NIR approach to be used under varying feedstock compositions, process stages, and temperature variations. The successful transfer of the spectra across the lab instrument and the online process analyser by PDS further shows the potential for cross-linking spectrometers and for sharing applications among standardised instruments (Shenk et al., 2008). Since the two FT-NIR spectrometers used for the spectra transfer had no shifts along their wavenumber axes, it would be of further interest to test the potential of a spectra transfer across different types of instruments, e.g., between an FT-NIR and a diode-array spectrometer.

The models developed during experiments four and five were generated by splitting the training set into subsets used for calibration and validation, and the exact splitting method differed for the two experiments. For experiment four, the samples used for validation of the models were selected from a separate period, i.e., the validation samples were independent from the calibration samples with respect to time. From this, it is concluded that as long as the environmental conditions remain constant, accuracy such as that obtained from validation realistically reflects the future model performance. According to the TOS, this conclusion cannot be drawn from the results of experiment five because the validation samples were not independent. However, the key finding of this experiment, namely that the short-term dynamics of acetic and propionic acid could be monitored independently from each other, is not affected by this limitation. This is because the ability for an in situ monitoring of the (uncorrelated) concentration changes of these analytes over time was the relevant criterion, rather than the verification of the accuracy and precision of the models with respect to the reference values themselves. The same holds true for the demonstration that acid accumulation induced by factors other than overfeeding, in this case the use of poor-quality silage, can also be detected by NIR. This further supports the assumption that there exists a specific relationship between the spectra and the VFA concentration in the slurry. Since the occurrence of these single acids have different meaning for the AD process stability, their independent estimation may also support the development of advanced monitoring and control strategies, e.g., by use of the propionic acid to acetic acid ratio (see references in Boe, 2006).

The technical scale digester used for experiment four was well mixed as demonstrated by the VS estimates, which quickly reached a new level after feeding. This may not be the case for larger

digesters where inhomogeneity due to suboptimal mixing makes representative spectral recording a potentially difficult task. For up-scale of the experiments, i.e., the application of the bypass at a large-scale biogas plant, the correct positioning is an important aspect and is required in order to obtain representative spectral recordings. Further research may also include the parallel use of measuring heads to compare the estimates at different digester positions, which may allow assessment of the efficiency of the agitator. Prior to experiment four, no information was available regarding the sensitivity of the NIR sensor in detection of changes in the total VFA level and a high OLR should ensure a reaction which was expected to be detectable by NIR. Because of the high sensitivity of the sensor for detection of changes in the total VFA level at lower concentrations (2–3 g kg^{-1} FM), feeding for experiment five was adjusted to an OLR that reflected the conditions of one-stage digesters in practice. This ability of the NIR method to detect acid accumulation at low concentrations is an important aspect since it allows the operator to take action at early stages of AD imbalances.

Further improvements for the monitoring of the AD process of agricultural biogas plants seem possible by combination of the NIR technique with other analytical methods. One such combination can be the use of NIR-generated online data in conjunction with measurements of the stable carbon isotope ratio ($\delta^{13}C_{CH4}$ and $\delta^{13}C_{CO2}$) in the gas phase. It was reported that the shifts in the isotope signature may provide information about the actual microbiological activities, which may help to identify the degradation kinetics and specific methanogenic pathways of AD. The isotope dynamics may 'indicate a temporal change in degradation pathways and/or a change in the relative contribution from different carbon fractions within the substrate' (Laukenmann et al., 2010). Use of ^{13}C spectroscopy also enables high resolution since the measurements can be performed online using optical measurement techniques based on laser absorption spectroscopy (Keppler et al., 2010). The combination of real-time information regarding the chemical conditions in the digester (NIR) with information on the actual degradation kinetics (^{13}C) may support the development of advanced strategies for the optimisation of the biogas process. This makes it conceivable not only to derive the optimal feeding strategy with respect to the quantity of feeding (OLR), but also to define suitable feedstock mixtures that are adapted to the actual microbial activities. Such a strategy may also include optimisation towards a feedstock-specific temperature level since the NIR models proved to be temperature robust. Overall, such a combined online approach may support reduction of the retention time due to an increased degradation efficiency which may allow for higher throughput and increased biogas productivity.

Full-load operation of biogas plants requires AD monitoring at fast analytical speed, which is of critical value for sensitive periods, for example, after feeding events, in order to detect imbalances at early stages. In addition, any analytical approach to be used in practice must be robust against changes in the AD process induced by variation of the feedstock and temperature fluctuations in the digester. The results of this work demonstrate that it is possible to develop NIR applications that meet these criteria. The time series results suggest that considerable improvement of AD monitoring is possible using NIR spectroscopy compared to the wet chemistry method. This classical analysis should be continued on an event-related basis, e.g., after variation of the feedstock composition to account for matrix-specific errors and to allow for the inclusion of new spectral information by model updating.

7. References

Ahring, B.K., Sandberg, M., Angelidaki, I., 1995. Volatile fatty acids as indicators of process imbalance in anaerobic digesters. Appl. Microbiol. Biotechnol. 43, 559-565.

Angelidaki, I., Ellegard, L., Ahring, B.K., 1993. A mathematical model for dynamic simulation of anaerobic digestion of complex substrates, focusing on ammonia inhibition. Biotechnol. Bioeng. 42, 159-166.

Anonymous, 2006. The Unscrambler Methods: Principles of data pre-processing. Available at: http://www.camo.com/downloads/U9.6%20pdf%20manual/The%20Unscrambler%20Methods.pdf (accessed on 05 December 2012).

Barredo, M.S., Evison, L.M., 1991. Effect of propionate toxicity on methanogenetic-enriched sludge, Methanobrevibacter smithii, and Methanospirillum hungatti at different pH values. Appl. Environ. Microb. 57, 1764-1769.

Begley, T.H., Lanza, E., Norris, K.H., Hruschka, W.R., 1984. Determination of NaCl in meat by near infrared spectroscopy. J. Agr. Food Chem. 32, 987-987.

Björnsson, L., Murto, M., Jantsch, T.G., Mattiasson, B., 2001. Evaluation of new methods for the monitoring of alkalinity, dissolved hydrogen and the microbial community in anaerobic digestion, Water Res. 35, 2833-2840.

Bjørsvik, H.R., Martens, H., 2008. Data Analysis: Calibration of NIR Instruments by PLS Regression, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 189-205.

Boe, K., 2006. Online monitoring and control of the biogas process. Dissertation. Institute of Environment & Resources Technical University of Denmark. Available at: http://orbit.dtu.dk/fedora/objects/orbit:82736/datastreams/file_5083454/content (accessed on 05 December 2012).

Boe, K., Batstone, D.J., Steyer, J.P., Angelidaki, I., 2010. State indicators for monitoring the anaerobic digestion process. Water Res. 44, 5973-5980.

Bokobza, L., 2002. Origin of near-infrared absorption bands, in: Siesler, H.W., Ozaki, Y., Kawata, S., Heise H.M. (Eds.), Near-Infrared Spectroscopy, Principles, Instruments, Applications, Wiley-VCH, Weinheim, pp.11-41.

Bouveresse, E., Campbell, B., 2008. Transfer of Multivariate Calibration Models Based on Near-Infrared Spectroscopy, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 231-243.

Boysworth, M.K., Booksh, K.S., 2008. Aspects of Multivariate Calibration Applied to Near-Infrared Spectroscopy, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 207-229.

Burns, D.A., Ciurczak, E.W., 2008. Handbook of near-infrared analysis, fourth ed. CRC Press, Boca Raton.

Chen, Y., Cheng, J.J., Creamer, K.S., 2008. Inhibition of anaerobic digestion process: A review. Bioresour. Technol. 99, 4044-4064.

Clark, D.H., Mayland, H.F., Lamb, R.C., 1987. Mineral analysis of forages with near infrared reflectance spectroscopy. Agron. J. 79, 485-490.

Conzen, J.P., 2005. Multivariate Kalibration. Ein praktischer Leitfaden zur Methodenentwicklung in der quantitativen Analytik. Bruker, Ettlingen.

Danzer, K., Horbert., H., Fischbacher., C., Jagemann, K.U., 2001. Chemometrik. Grundlagen und Anwendungen, Springer, Berlin Heidelberg.

Dardenne, P., 2010. Some considerations about NIR spectroscopy: Closing speech at NIR-2009. NIR news 21, 8-10.

Davies, A.M.C., Fearn, T., 2006. Back to Basics: calibration statistics. Spectroscopy Europe 18, 31-32.

Despagne, F., Massart, D.L., Jansen, M., van Daalen, H., 2000. Intersite transfer of industrial calibration models. Anal. Chim. Acta 406, 233-245.

Duckworth, J., 2004. Mathematical Data Preprocessing, in: Roberts, C.A., Workman Jr, J., Reeves III, J.B. (Eds.) Near-Infrared Spectroscopy in Agriculture, American Society of Agronomy, Wisconsin, pp. 115-132.

Esbensen, K.H., 2009. Multivariate Data Analysis – In Practice. An introduction to Multivariate Data Analysis and Experimental Design, fifth ed. Camo Inc., Esbjerg.

Esbensen K.H., Geladi, P., 2010. Principles of proper validation: use and abuse of re-sampling for validation. J. Chemometr. 24, 168-187.

Faber, K., Kowalski, B.R., 1997. Improved prediction error estimates for multivariate calibration by correcting for the measurement error in the reference values. Appl. Spectrosc. 51, 660–665.

Feudale, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D., Ferré, J., 2002. Transfer of multivariate calibration models: a review. Chemom. Intell. Lab. Syst. 64, 181-192.

Forina, M., Lanteri, S., Casale, M., 2007. Multivariate calibration. J Chromatogr, A 1158, 61-93.

Gerardi, M.H., 2003. The microbiology of anaerobic digesters. First ed. Wiley Interscience, New Jersey.

Griffiths, P.R., Dahm, D.J., 2008. Continuum and Discontinuum Theories of Diffuse Reflection, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 21-64.

Hageman, J.A., Westerhuis, J.A., Smilde, A.K., 2005. Temperature robust multivariate calibration: an overview of methods for dealing with temperature influences on near infrared spectra. J. Near Infrared Spec. 13, 53-62.

Hansen, K.H., Angelidaki, I., Ahring, B.K, 1998. Anaerobic digestion of swine manure: inhibition by ammonia. Wat. Res. 32, 5-12.

Hansson, M., Nordberg, Å., Sundh, I., Mathisen, B., 2002. Early warning of disturbances in a laboratory-scale MSW biogas process. Water Sci. Technol. 45, 255-260.

Hecht, M., 2009. Die Bedeutung des Carbonat-Puffersystems für die Stabilität des Gärprozesses landwirtschaftlicher Biogasanlagen. Dissertation, Friedrich-Wilhelms-University, Bonn. Available at: http://hss.ulb.uni-bonn.de/2009/1705/1705.htm (accessed on 05 December 2012).

Heise, H.M., Winzen, R., 2002. Fundamental Chemometric Methods, in: Siesler, H.W., Ozaki, Y., Kawata, S., Heise H.M. (Eds.), Near-Infrared Spectroscopy, Principles, Instruments, Applications, Wiley-VCH, Weinheim, pp.125-162.

Henkelmann, G., Meyer zu Köcker, K., Götz, J., Beck, J., 2010. Schlüsselparameter zur Kontrolle des Gärprozesses und Motivation, Voraussetzung und Möglichkeiten für die Prozessüberwachung. Bayerische Landesanstalt für Landwirtschaft, Freising. Available at: http://www.lfl.bayern.de/publikationen/daten/schriftenreihe/p_40100.pdf (accessed on 05 December 2012).

Holm-Nielsen, J.B., Andree, H., Lindorfer, H., Esbensen, K.H., 2007. Transflexive embedded near infrared monitoring for key process intermediates in anaerobic digestion/biogas production. J. Near Infrared Spectrosc. 15, 123-135.

Holm-Nielsen, J.B., Dahl, C.K., Esbensen, K.H., 2006. Representative sampling for process analytical characterization of heterogeneous bioslurry systems – a reference study of sampling issues in PAT. Chemometr. Intell. Lab. 83, 114-126.

Holm-Nielsen, J.B., Esbensen, K.H., 2011. Monitoring of biogas test plants – a process analytical technology approach. J. Chemometr. 25, 357-365.

Holm-Nielsen, J.B., Lomborg, C.J., Oleskowicz-Popiel, P., Esbensen, K.H., 2008. On-line near infrared monitoring of glycerol-boosted anaerobic digestion processes: Evaluation of process analytical technologies. Biotechnol. Bioeng. 99, 302-313.

Hunt, G.R., Salisbury, J.W., 1971. Visible and near-infrared spectra of minerals and rocks. II. Carbonates. Modern Geology 2, 195-205.

Isaksson, T., Næs, T., 1989. Selection of samples for calibration in near infrared spectroscopy, I. General principles illustrated by example. Appl. Spectrosc. 43, 328-335.

Jacobi, H.F., Moschner, C.R., Hartung, E., 2009. Use of near infrared spectroscopy in monitoring of volatile fatty acids in anaerobic digestion. Water Sci. Technol. 60, 339-346.

Jacobi, H.F., Moschner, C.R., Hartung, E., 2011. Use of near infrared spectroscopy in online-monitoring of feeding substrate quality in anaerobic digestion. Bioresour. Technol. 102, 4688-4698.

Jantsch, T.G., Mattiasson, B., 2004. An automated spectrophotometric system for monitoring buffer capacity in anaerobic digestion processes. Water Res. 38, 3645-3650.

Kawano, S., 2002. Sampling and Sample Presentation, in: Siesler, H.W., Ozaki, Y., Kawata, S., Heise H.M. (Eds.), Near-Infrared Spectroscopy, Principles, Instruments, Applications, Wiley-VCH, Weinheim, pp.115-124.

Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137-148.

Keppler, F., Laukenmann, S., Rinne, J., Heuwinkel, H., Greule, M., Whiticar, M., Lelieveld, J., 2010. Measurement of 13C/12C methane from anaerobic digesters: comparison of continuous isotope ratio mass spectrometry and optical spectrometry. Environ. Sci. Technol. 44, 5067-5073.

Kessler, W., 2007. Multivariate Datenanalyse für die Pharma-, Bio- und Prozessanalytik , Wiley-VCH, Weinheim.

Kissel, R., Bachmaier, H., Effenberger, M,, Gronauer, A., 2009. Auslastungsgrade und Ursachen für Ertragseinbußen an Biogasanlagen unterschiedlicher Konzeptionierung, in: Proceedings of the International Conference on Biogas Science, Bavarian State Research Center for Agriculture, Erding, pp. 227-236.

Köhler, W., Schachtel, G., Voleske, P., 2007. Biostatistik, fourth ed. Springer, Berlin Heidelberg.

Laukenmann, S., Polag, D., Heuwinkel, H., Greule, M., Gronauer, A., Lelieveld, J., Keppler, F., 2010. Identification of methanogenic pathways in anaerobic digesters using stable carbon isotopes. Engineering in Life Sciences, 10, 509–514.

Lehtomäki, A., Huttunen, S., Rintala, J.A., 2007. Laboratory investigations on co- digestion of energy crops and crop residues with cow manure for methane production: Effect of crop to manure ratio. Resour., Conserv. Recycl. 51, 591-609.

Lomborg, C.J., Holm-Nielsen, J.B., Oleskowicz-Popiel, P., Esbensen, K.H., 2009. Near infrared and acoustic chemometrics monitoring of volatile fatty acids and dry matter during co-digestion of manure and maize silage. Bioresour. Technol. 100, 1711-1719.

Malley, D., Martin, P., Ben-Dor, E., 2004. Analysis of Soils, in: Roberts, C.A., Workman Jr, J., Reeves III, J.B. (Eds.) Near-Infrared Spectroscopy in Agriculture, American Society of Agronomy, Wisconsin, pp. 729-784.

Martens, H., Hoy, M., Westad, F., Folkenberg, D., Martens, M., 2001. Analysis of designed experiments by stabilized PLS Regression and jack-knifing. Chemom. Intell. Lab. Syst. 58, 151-170.

Martens, H., Martens, M., 2000. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). Food Qual. Prefer. 11, 5-16.

Martens, H., Martens, M., 2001. Multivariate Analysis of Quality, Wiley, Chichester.

Martens, H., Næs, T., 1991. Multivariate Calibration, updated version, John Wiley and Sons, Chichester.

McClure, W.F., 2008. Analysis Using Fourier-Transforms, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 93-121.

Miller, J.N., Miller, J.C., 2005. Statistics and Chemometrics for Analytical Chemistry, fifth ed. Pearson, Essex.

Murphy, T., 2010. The Intermediate Precision of ASTM International Test Methods. ASTM Standardization News. Available at: http://www.astm.org/SNEWS/ND_2010/datapoints_nd10.html (accessed on 05 July 2012).

Myint, M., Nirmalakhandan, N., Speece, R.E., 2007. Anaerobic fermentation of cattle manure: Modeling of hydrolysis and acidogenesis. Water Res. 41, 323-332.

Næs, T., Isaksson, T., Fearn, T., Davies, T., 2002. Multivariate Calibration and Classification, NIR publications, Chichester.

Ndegwa, P.M., Hristov, A.N., Arogo, J., Sheffield, R.E., 2008. A review of ammonia emission mitigation techniques for concentrated animal feeding operations. Biosystems Engineering. 100, 453-469.

Osborne, B.G., Fearn, T., 1986. Near infrared spectroscopy in food analysis, Wiley, New York.

Pohland, F.G., Ghosh, S., 1971. Developments in anaerobic stabilization of organic wastes – the two-phase concept. Environ. Lett. 1, 255–266.

Preißler, D., 2010. Prozessbiologische Störungen in NawaRo- und Gülleanlagen: Symptome, Ursachen und möglich Lösungsansätze. Biogas Forum Bayern. Available at: www.biogas-forum-bayern.de/publikationen/Prozessbiologische_Storungen_in_NawaRo-Anlagen.pdf (accessed on 05 December 2012).

Pullammanappallil, P.C., Chynoweth, D.P., Lyberatos, G., Svoronos, S.A., 2001. Stable performance of anaerobic digestion in the presence of a high concentration of propionic acid. Bioresour. Technol. 78, 165-169.

Raju, C.S., Løkke, M.M., Sutaryo, S., Ward, A.J., Møller, H.B., 2012. NIR Monitoring of Ammonia in Anearobic Digesters Using a Diffuse Reflectance Probe. Sensors 12, 2340-2350.

Reed, J., Davies, A.N., 2011. Near infrared spectroscopy in bioreactor performance monitoring. Spectroscopy Europe. 3, 22-25.

Reed J., Devlin D., Esteves S.R.R., Dinsdale R., Guwy A.J., 2011. Performance parameter prediction for sewage sludge digesters using reflectance FT-NIR spectroscopy. Water Res. 45, 2463-1472.

Roberts, C.A., Workman Jr, J., Reeves III., J.B., 2004. Near-infrared spectroscopy in agriculture, American Society of Agronomy, Madison.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry 36, 1627-1639.

Shenk, J.S., 2004. Standardization and Calibration Transfer, in: Roberts, C.A., Workman Jr, J., Reeves III, J.B. (Eds.) Near-Infrared Spectroscopy in Agriculture, American Society of Agronomy, Wisconsin, pp. 207-227.

Shenk, J.S., Workman Jr, J., Westerhaus, M.O., 2008. Application of NIR Spectroscopy to Agricultural Products, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 347-386.

Siesler, H.W., 2008. Basic Principles of Near-Infrared Spectroscopy, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 7-19.

Stuart, B., 2004. Infrared spectroscopy: fundamentals and applications, first ed. Willey, Chichester.

Swierenga, H., Wülfert, F., de Nood, O.E., de Weijer, A.P., Smilde, A.K., Buydens, L.M.C., 2000. Development of robust calibration models in near infra-red spectrometric applications. Anal Chim Acta. 411, 121-135.

Switzenbaum M.S., Giraldo-Gomez E., Hickey R.F., 1990. Monitoring of the anaerobic methane fermentation process. Enzyme Microb. Technol., 12, 722-730.

Tillmann, P., 1996. Kalibrationsentwicklung für NIRS-Geräte. Eine Einführung, Cuvillier, Göttingen.

Wang, R., 2011. Computer Image Processing and Analysis (E161). Available at: http://fourier.eng.hmc.edu/e161/lectures/smooth_sharpen/node3.html (accessed on 05 December 2012).

Ward, A.J., Bruni, E., Lykkegaard, M.K., Feilberg, A., Adamsen, A.P.S., Jensen, A.P., Poulsen, A.K., 2011b. Real time monitoring of a biogas digester with gas chromatography, near-infrared spectroscopy, and membrane-inlet mass spectroscopy. Bioresour. Technol. 102, 4098-4103.

Ward, A.J., Hobbs, P.J., Holliman, P.J., Jones, D.L., 2008. Optimization of the anaerobic digestion of agricultural resources. Bioresour. Technol. 99, 7928-7940.

Ward, A.J., Hobbs, P.J., Holliman, P.J., Jones, D.L., 2011a. Evaluation of near infrared spectroscopy and of software sensor methods for determination of total alkalinity in anaerobic digesters. Bioresour. Technol. 102, 4083-4090.

Ward, A.J., Møller, H.B., Raju, C.S., 2010. Using near infrared spectroscopy to predict volatile fatty acids in biogas processes utilizing different feedstocks. Proceedings of the 14th Ramiran International

Conference, Lisbon, Portugal, 12-15 September 2010. Available at: http://www.ramiran.net/ramiran2010/docs/Ramiran2010_0026_final.pdf (accessed on 05 December 2012).

Weiland, P., 2010. Biogas production: current state and perspectives. Appl. Microbiol. Biotechnol. 85, 849-860.

Westerhaus, M., Workman Jr, J, Reeves III, J.B, Mark, H., 2004. Quantitative Analysis, in: Roberts, C.A., Workman Jr, J., Reeves III, J.B. (Eds.) Near-Infrared Spectroscopy in Agriculture, American Society of Agronomy, Wisconsin, pp. 133-174.

Williams, P.C., 2001. Implementation of Near-Infrared Technology, in: Williams, P.C., Norris, K. (Eds.) Near-Infrared Technology in the Agricultural and Food Industries, American Association of Cereal Chemists, St Paul, pp. 145-169.

Workman Jr, J., 2008. NIR Spectroscopy Calibration Basics, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, third ed. CRC Press, Boca Raton, pp. 123-150.

Workman Jr, J., Shenk, J., 2004. Understanding and Using the Near-Infrared Spectrum as an Analytical Method, in: Roberts, C.A., Workman Jr, J., Reeves III, J.B. (Eds.) Near-Infrared Spectroscopy in Agriculture, American Society of Agronomy, Wisconsin, pp. 3-10.

Workman Jr, J., Weyer, L., 2008. Practical Guide to Interpretive Near-Infrared Spectroscopy. CRC Press, Boca Raton.

Yookyung, K., Himmelsbach, D.S., Kays, S.E., 2007. ATR-fourier transform mid-infrared spectroscopy for determination of trans fatty acids in ground cereal products without oil extraction. J. Agric. Food Chem. 55, 4327–4333.