LWT - Food Science and Technology 139 (2021) 110557

Contents lists available at ScienceDirect

LWT

journal homepage: www.elsevier.com/locate/lwt

Multi-scale benchtop ^{1}H NMR spectroscopy for milk analysis

Alper Soyler [a], Sevil Cikrikci [b], Cagri Cavdaroglu [c], Dylan Bouillaud [d], Jonathan Farjon [d], Patrick Giraudeau [d,**], Mecit H. Oztop [a,*]

[a] Department of Food Engineering, Middle East Technical University, Ankara, Turkey
[b] Department of Food Engineering, Konya Food and Agriculture University, Konya, Turkey
[c] Department of Food Engineering, Izmir Institute of Technology, Izmir, Turkey
[d] Université de Nantes, CNRS, CEISAM UMR, 6230, Nantes, France

ARTICLE INFO

Keywords:
Benchtop NMR spectroscopy
Milk
Lactose hydrolysis
Classification
Flow NMR
Artificial neural network (ANN)

ABSTRACT

Benchtop NMR systems offers various advantages such as being easy to use, not requiring constant maintenance and being available at affordable prices. In this study, multiple aspects of benchtop NMR spectroscopy were explored to analyze milk in an industrial context, either regarding the quality of production or regarding the differentiation of the final product. The first part focuses on the production conditions of lactose hydrolysis in milk and quantitative online NMR spectroscopy was adapted to follow lactose hydrolysis in milk in continuous flow mode. The second part focuses on differentiating milk samples having different properties. 36 milk samples from France and Turkey were analysed and glycerol, fat and sugar contents were measured from the NMR spectra. Combination of spectroscopic data with a proposed Artificial Neural Network model enabled to classify milk of different origins and different properties. This study shows that benchtop NMR spectroscopy is a versatile non-destructive control method that can help controlling milk quality both during and after production.

1. Introduction

Milk is a unique product including significant amounts of nutrients such as calcium, proteins and vitamins (Dutra Rosolen, Gennari, Volpato, & Volken De Souza, 2015). It can be considered as the 'magic liquid' due to its rich and various composition and confirmed benefits on the newborns. In this context, accessible analytical tools are crucial to ensure the quality of milk, both to optimize the production processes and after production to ensure its characteristic properties.

The first focus of this paper is about the production stage, on the example of the potential production of lactose-free milk. One of the milk constituents, lactose (*aka milk sugar*) is a disaccharide composed of glucose and galactose that are linked by a β-1,4 glycosidic bond (Dong & Zhong, 2019). However, lactose found in the milk or milk products cause digestion problems in some part of the population. The absence of the digestive enzyme lactase which is also known as 'lactose intolerance' is the reason of this problem. Significant amount of world population suffers from lactose intolerance thus demand for lactose free products is getting increasing attention (Churakova et al., 2019). Lactose free milk is produced by the pre-digestion of the lactose in milk with the addition of the lactase enzyme. Controlling the production of lactose-free milk during its production has a major industrial impact. There exist many different techniques recognized by analytical standard agencies to determine the lactose content of milk, such as polarimetry, mid-infrared detection, gravimetry, differential pH, enzymatic methods detecting either the glucose or galactose moiety of lactose and HPLC (Churakova et al., 2019). Each of these methods has its own advantages and disadvantages, but to our knowledge none of them have the ability to monitor the hydrolysis process online in the continuous mode. Therefore, developing new online tools for controlling this reaction is of major importance for the milk industry.

On the other hand, the differentiation of milk products not only regarding the sugar type (*lactose free or not*) but also according to the other parameters such as process, origin, fat content, etc. requires considerable attention especially considering the regulatory specifications and quality of the milk. In recent years, several studies have focused on discriminating the products coming from different farming systems, animals or feeding strategies, origin etc. (Bergamaschi, Cipolat-Gotet, Cecchinato, Schiavon, & Bittante, 2020). Traditional methods like chromatographic techniques or sensory analyses might

* Corresponding author.
** Corresponding author.
E-mail addresses: patrick.giraudeau@univ-nantes.fr, patrick.giraudeau@univ-nantes.fr (P. Giraudeau), mecit@metu.edu.tr, mecit@metu.edu.tr (M.H. Oztop).

https://doi.org/10.1016/j.lwt.2020.110557
Received 15 July 2020; Received in revised form 30 September 2020; Accepted 8 November 2020
Available online 13 November 2020

require highly skilled operators and they could not be easily adapted to online monitoring. Moreover, they could be time consuming and expensive. Easy, quick to use, minimum chemical requiring methods have always been on the interest of manufacturers for ensuring the quality and authenticity.

Nuclear magnetic resonance (NMR) spectroscopy, is a powerful, fingerprint tool for the simultaneous identification and quantification of compounds in complex mixtures (Monakhova, Kuballa, Leitz, Andlauer, & Lachenmeier, 2012). It is recognized as a valuable tool for food analysis enabling not only the characterization of food matrices in terms of authenticity and quality but also identification of counterfeits (Sobolev et al., 2019). Also, studies of ^{1}H NMR-based metabolomics to get metabolite profiles during storage (Jansson et al., 2014) or to analyze heritability of individual milk metabolites (Buitenhuis et al., 2013) in animal based products have been explored (Bertram, 2018). However, most of the NMR spectroscopy instruments work at high field (^{1}H frequencies >300 MHz) and requires a massive dedicated place with specific installation and trained staff as well as a substantial investment cost and expensive cryogenic fluids for maintenance (Bouillaud, Farjon, Gonçalves, & Giraudeau, 2019). To overcome such limitations, benchtop systems capable of conducting spectroscopy experiments have been developed (Perlo et al., 2005). These systems have been successfully used to monitor authenticity in some kind of foods such as in coffee (Defernez et al., 2017) and in meat (Jakes et al., 2015). Furthermore, thanks to their compact character, NMR spectroscopy can easily be used for online monitoring under flow conditions (Anderssen & McCarney, 2020;Bouillaud, Heredia, et al., 2019; Knox, Parkinson, Stone, & Warren, 2019; Rönnols, Danieli, Freichels, & Aldaeus, 2020; Soyler, Bouillaud, Farjon, Giraudeau, & Oztop, 2020), which makes them ideal candidates for the online control of food production processes. For instance, benchtop NMR spectroscopy has been used to monitor continuous processes such as enzymatic hydrolysis of marine by-products (Anderssen & McCarney, 2020), sucrose hydrolysis (Soyler, Bouillaud, Farjon, Giraudeau, & Oztop, 2020), *in vivo* monitoring of lipid accumulation in microalgae (Bouillaud et al., 2020) and lignin analysis (Rönnols et al., 2020). However it is important to note that implementing NMR spectroscopy in a flow system is challenging and can introduce additional complexity relative to static samples (Knox et al., 2019). The impact of flow rate on signal intensities should be investigated carefully and the parameters should be set accordingly (Soyler, Bouillaud, Farjon, Giraudeau, & Oztop, 2020).

To our knowledge, this will be 1st study where milk will be investigated by using a bench top NMR spectroscopy system from different aspects. We have explored the versatility of benchtop NMR on milk samples for two different purposes. In the first part of the study, benchtop NMR spectroscopy was utilized in conditions mimicking the online control of lactose hydrolysis in milk and lactose hydrolysis was monitored first in a model system, then in a milk sample -with no need for deuterated solvents through the use of a tailored solvent suppression pulse sequence (Gouilleux, Charrier, Akoka, & Giraudeau, 2017). Kinetic analysis of the reaction was also performed. In the second part, 36 milk samples from Turkey and France; having different fat and sugar contents; production styles (conventional vs organic) were analysed by NMR spectroscopy together with multivariate approaches to assess the potential of benchtop NMR as a differentiation tool.

2. Materials and methods

2.1. Materials

Lactose in the form of *D-(+)-Lactose monohydrate* was purchased from Fluka BioChemika, Switzerland. Lactase enzyme (Maxilact® LGi 5000) for lactose hydrolysis was supplied from DSM, (Heerlen, The Netherlands). D-(+)-Glucose, D-(+)-Galactose, imidazole and 3-(Trimethylsilyl) propionic-2,2,3,3-d_4 acid (TSP) were provided from Sigma-Aldrich Co. (St. Louis, MO, USA). HPLC grade water from a water purification system (Nanopure Infinity, Barnstead International, IA) was used for the preparation of the solutions. For the lactose hydrolysis experiments in milk, a UHT whole milk (Lactel, France) bought from a grocery store in France was used.

For the second part of the study, 36 different samples of milk (different cow milks: whole milk, organic whole milk, low-fat milk, lactose-free milk and goat milk) were purchased from local Turkish and French supermarkets. Information on the milk samples are given in Table 1.

2.2. Sample preparations

For online monitoring analysis, hydrolysis reactions were performed both for a model system and on a milk sample. As the model solution, a solution of glucose, galactose and lactose was used. 5% (w/w) glucose, galactose and lactose solutions were prepared separately by dissolving in non-deuterated HPLC grade water in a 50 mL tube. TSP was used as the internal standard.

For milk differentiation experiments, 0.5% (w/w) imidazole was added to the 36 milk samples (Table 1) as the internal reference for NMR measurements since TSP was masked with the peaks of the fats present in the milk. Imidazole had ^{1}H chemical shifts of 7.26 ppm and 8.11 ppm. All the samples were stirred for 5 min for complete dissolution of imidazole in the milk. Afterwards, the samples were transferred into 5 mm NMR tubes and stored in the dark and refrigerated before analysis to prevent spoilage.

2.3. NMR spectroscopy

NMR spectroscopy experiments were performed on a benchtop spectrometer operating at a 43 MHz ^{1}H frequency with a compact permanent magnet based on the Halbach design (Magritek Spinsolve, Wellington, NZ) (E. Danieli, Perlo, Blümich, & Casanova, 2013; Ernesto Danieli, Perlo, Blümich, & Casanova, 2010). The spectrometer was equipped with a gradient coil along the B_0 axis in the transverse plane of the NMR tube that could produce a maximum field gradient of 0.16 T m^{-1} and also had an external lock system which allowed the use of non-deuterated solvents. The experiments were performed at 29 °C since this was the temperature at which the magnet stability was optimal. Since deuterated solvents were not used, an experiment to suppress the water signal was used, based on a WET-180-NOESY (Gouilleux, Charrier, Akoka, & Giraudeau, 2017). This pulse sequence provides an optimal solvent suppression for small molecules on a benchtop spectrometer, leading to a lower and narrower water signal with a clean phase with a minimal impact on nearby peaks.

An inversion recovery sequence was used for measuring longitudinal relaxation times T_1 using an inversion time range of 0.1–10,000 ms with 15 points.

The 1D ^{1}H spectra were obtained with 64 scans for a total experiment time of 6 min for model lactose solution hydrolysis and 7 min for milk hydrolysis and '*milk differentiation*' experiments. The 90° pulse angle was achieved by a pulse length of 6.7 μs at 0 dB. The FIDs were recorded with 16 K points, a dwell time of 200 μs, and a repetition time of 6 s for model lactose solution and 7 s for milk, corresponding to 5 times the longest T_1 in the sample to ensure quantitative conditions. 1D data were processed with MestReNova. To align all the spectra correctly, TSP was used to calibrate the chemical shift axis at 0 ppm for the hydrolysis experiments. For milk spectra, the signal of imidazole at 7.26 ppm was used as the chemical shift reference. All spectra were processed with a 0.2 Hz exponential apodization, an automatic phase correction and an automatic baseline correction via a Whittaker smoother algorithm. Manual phase correction and manual baseline correction were also performed in addition to the automatic corrections. Relevant peak integrals were calculated by integration with the MestReNova software. Deconvolution tools were also evaluated but yielded a slightly lower performance than integration, probably due to the non-ideal line shapes.

Table 1
Milk samples used in the study.

Sample Name	Country	Fat Content	Sugar Content	Milk Source	Production Style	Processing
WMF-1	France	Whole	Regular	Cow	Conventional	UHT
WMF-2	France	Whole	Regular	Cow	Conventional	UHT
WMF-3	France	Whole	Regular	Cow	Conventional	Pasteurized
WMF-4	France	Whole	Regular	Cow	Conventional	Pasteurized
OMF-5	France	Whole	Regular	Cow	Organic	UHT
OMF-6	France	Whole	Regular	Cow	Organic	Pasteurized
OMF-7	France	Whole	Regular	Cow	Organic	Pasteurized
HFMF-8	France	Half	Regular	Cow	Conventional	UHT
OHFMF-9	France	Half	Regular	Cow	Organic	Pasteurized
HFMF-10	France	Half	Regular	Cow	Conventional	Pasteurized
HFMF-11	France	Half	Regular	Cow	Conventional	Pasteurized
HFMF-12	France	Half	Regular	Cow	Conventional	UHT
OHFMF-13	France	Half	Regular	Cow	Organic	UHT
WMF-14	France	Whole	Regular	Cow	Conventional	UHT
OHFMF-15	France	Half	Regular	Cow	Organic	UHT
HFMF-16	France	Half	Regular	Cow	Conventional	UHT
OMF-17	France	Whole	Regular	Cow	Organic	Pasteurized
OHFMF-18	France	Half	Regular	Cow	Organic	UHT
GMF-19	France	Half	Regular	Goat	Conventional	UHT
GMF-20	France	Whole	Regular	Goat	Organic	UHT
GMF-21	France	Whole	Regular	Goat	Organic	UHT
GMF-22	France	Whole	Regular	Goat	Organic	UHT
WMT-23	Turkey	Whole	Regular	Cow	Conventional	UHT
WMT-24	Turkey	Whole	Regular	Cow	Conventional	UHT
WMT-25	Turkey	Whole	Regular	Cow	Conventional	UHT
WMT-26	Turkey	Whole	Regular	Cow	Conventional	UHT
WMT-27	Turkey	Whole	Regular	Cow	Conventional	UHT
WMT-28	Turkey	Whole	Regular	Cow	Conventional	UHT
OMT-29	Turkey	Whole	Regular	Cow	Organic	UHT
GMT-30	Turkey	Whole	Regular	Goat	Conventional	UHT
LFMF-31	France	Half	Lactose-Free	Cow	Conventional	UHT
LFMF-32	France	Half	Lactose-Free	Cow	Conventional	UHT
LFMT-33	Turkey	Half	Lactose-Free	Cow	Conventional	UHT
LFMT-34	Turkey	Half	Lactose-Free	Cow	Conventional	UHT
LFMT-35	Turkey	Half	Lactose-Free	Cow	Conventional	UHT
LFMT-36	Turkey	Half	Lactose-Free	Cow	Conventional	UHT

2.4. Lactose hydrolysis in continuous flow system

For on-line monitoring, the system included a glass flow cell with 5 mm outer diameter, a peristaltic pump (Reglo Digital, Ismatec, Wertheim, Germany) and PEEK tubing. A heating plate (RCT Basic, IKA-Werke GmbH & Co. KG, Staufen, Germany) was added to the system to control the temperature of the hydrolysis reaction.

NMR signal is very sensitive to the flow rates as the signal becomes broader and the resolution decreases with increasing flow rates. The approach followed in a previous study (Soyler, Bouillaud, Farjon, Giraudeau, & Oztop, 2020) was used to determine the optimal flow rate of 0.5 mL/min.

2.5. Differentiation milk samples

Experiments were performed in 5 mm NMR tubes in static mode. For the quantitative analysis, lactose, glycerol and fat contents were determined from the spectra as will be shown later.

Considering the factors given in Table 1 (country (TR/FR), milk type (Goat/Cow), production style (Conventional/Organic), processing (UHT/Pasteurized), fat and sugar content), Analysis of Variance (ANOVA), was conducted on the responses of glycerol; lactose and fat contents to see if any of the responses showed differences with milk type. If significant difference was detected, means were compared by the Tukey test ($p < 0.05$) using MINITAB (Version 19, Coventry, U.K) software.

In addition, to obtain some grouping info for the given data set, an Artificial Neural Networks (ANN) model was developed by using NMR spectra results. Artificial Neural Networks (ANN) model is a nonlinear mathematical model with the capability of developing meaningful relationships between input and output variables through a learning process (Zheng et al., 2011). The function and organization of the human brain gave inspiration to its function and structure (Bila et al., 1999). The mathematical model designed to classify milk samples with respect to their fat and sugar contents, milk source, production and processing type was based on nonlinear ANNs. These intelligent algorithms have been used to link chemical information coming from 36 milk samples with milk type to achieve a potential tool for accurate differentiation and quality control.

ANN model was implemented with the 'neural net package' (Version 1.44.2) in R (Stefan Fritsch, Frauke Guenther, Marvin N. Wright, Marc Suling, & Sebastian M. Mueller, 2019). For training (calibration), 2/3 of the data were used and 1/3 was separated to test the model (external validation). In order to obtain the best score, ANN model was trained using different number of neurons between 2 and 15 in the hidden layers. ANN model yielded possible classifications of the samples and possibility ratios. R^2_{cal} was used to investigate the prediction ability of the trained ANN model and R^2_{pred} for external validation. Error values were stated as root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP). R^2 values should be close to 1 while error values should be low to obtain a robust prediction model (Uncu & Ozen, 2019). Optimal number of learning steps was selected according to the *minimum root means square error cross validation* (RMSECV) value for the calibration dataset. A value for R^2 between 0.50 and 0.65 can be used for discrimination between high and low concentrations. A value for R^2 between 0.66 and 0.81 shows approximate quantitative predictions., R^2 between 0.82 and 0.90 are considered as good prediction and above 0.91 are excellent (Saeys, Mouazen, & Ramon, 2005). Categorical variables were used as binary data, while numeric data was used for normalization and scaling.

3. RESULTS and DISCUSSION

3.1. Lactose hydrolysis in continuous flow system

3.1.1. 1D solvent suppression enhanced NMR spectra

At first, 1D ^{1}H NMR experiments were carried out with a water signal suppression pulse sequence on model solutions of isolated compounds to identify the peaks which could be used to monitor the lactose hydrolysis reactions. Fig. 1 shows the stacked spectra of glucose, galactose and lactose alone. Due to low magnetic field, the spectra of the sugars were highly overlapped. The suppressed water peak was observed at 4.9 ppm. As can be seen in Fig. 1, alpha anomeric proton peaks of glucose, galactose and lactose are overlapped at 5.3 ppm. Although beta anomeric proton peaks of glucose and galactose are masked by the water peak, the beta anomeric proton peak of lactose is seen as a distinctive peak at 4.4 ppm. Despite the fact that it is difficult to integrate only alpha lactose site due to alpha glucose and galactose signal overlap, it is still possible to monitor the lactose hydrolysis with the isolated beta anomeric signal. Moreover, anomeric ratio between alpha and beta anomers is constant during the monitoring therefore it is reliable to follow the beta anomeric site along the course of the hydrolysis.

3.1.2. Online monitoring and kinetic modelling of lactose hydrolysis in the model system

Lactose when hydrolyzed yields glucose and galactose as seen below.

For model lactose solution hydrolysis, a 5% lactose solution was hydrolyzed with 20 μL lactase enzyme. The NMR spectra were obtained at 6 min intervals for 240 min (Fig. 2a). The consumption of lactose can be seen from β-lactose peaks (Fig. 2b).

Concentrations of lactose for each time point were calculated using Equation (1). The number of protons contributing to the signal was 9 for TSP, and 3 for imidazole.

$$C_x = \frac{A_x/N_x}{A_{REF}/N_{REF}} \cdot \frac{MW_x}{MW_{REF}} C_{REF} \quad \text{(Eq.1)}$$

C, and MW were the concentrations in mg/g, the molecular weight. REF referred to the internal reference. A and N, denoted the integral area in a fully relaxed ^{1}H NMR spectrum and the number of hydrogens contributing to the signals. Hydrolysis experiments were repeated three times and the change on the average concentrations of lactose from the experiments were plotted as a function of time in Fig. 3.

Lactose data were fitted to Equation (2) ($R^2 > 0.99$) and the average rate constant was found as $k = 1.66*10^{-2}\ min^{-1}$ at 29 °C.

$$\frac{C - C_\infty}{C_o - C_\infty} = e^{-kt} \quad \text{(Eq. 2)}$$

3.1.3. Lactose hydrolysis in milk

Based on the good repeatability of the model lactose solution hydrolysis, milk was also hydrolyzed with 20 μL of enzyme. Milk is a much more complex fluid than the model lactose solution. The major constituents of milk are water, fat, proteins and lactose. Milk also contains trace amounts of vitamins, minerals, organic acids, enzymes. Therefore, 1D ^{1}H NMR spectrum of milk was quite different than the model lactose solution (Fig. 4a). The NMR spectra were obtained at each 7 min for 280 min. As seen in Fig. 4b, the consumption of lactose was followed using the peak of β-lactose. The average rate constant from the fractional conversion model yield a value of 1.52×10^{-2} min^{-1} at 29 °C which was consistent with the previous study of Panesar (2007) where it was found as 1.56×10^{-2} min^{-1} at 30 °C (Fig. 5).

Results of this part showed that lactose hydrolysis can be controlled online in a continuous mode successfully by using benchtop NMR spectroscopy, thus the method has high potential to be adapted for industrial processes. For lactose free milk production, it should be ensured that residual lactose amounts are within the legal limits for that product category. This can further be confirmed by other high-resolution spectroscopic techniques.

3.1.4. Differentiation of milk samples by benchtop ^{1}H NMR profiling and chemometrics

In this part of the study, 36 milk samples of which the properties have been listed in Table 1 were analysed. A sample spectrum of milk

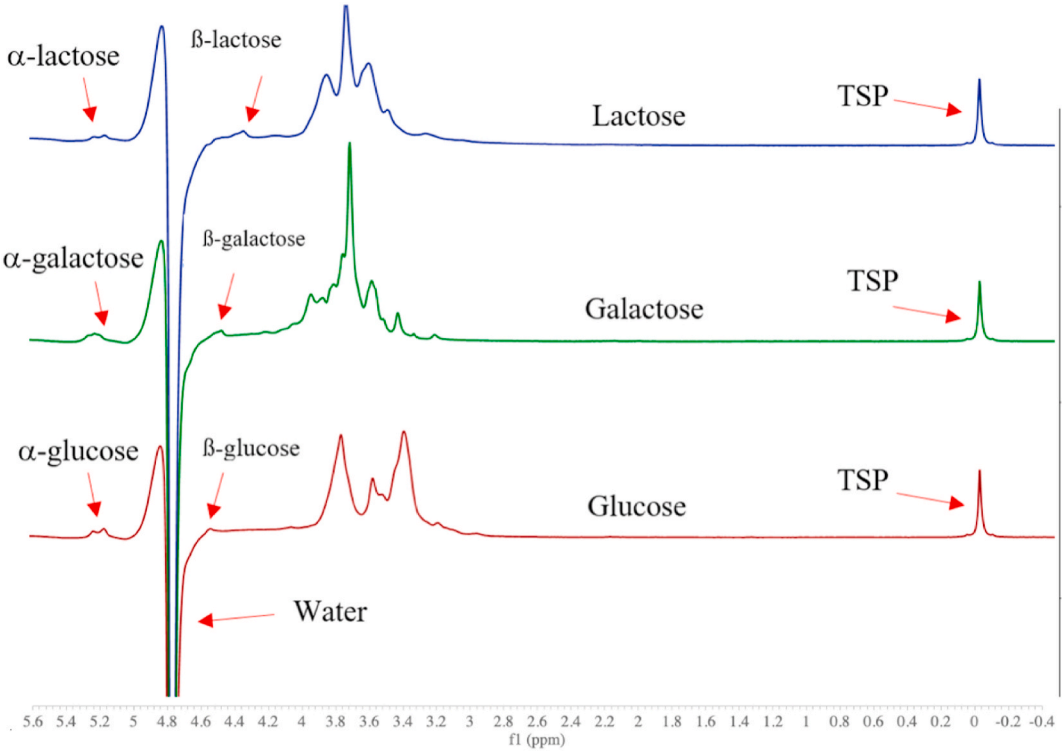

Fig. 1. Stacked spectra of glucose, galactose and lactose.

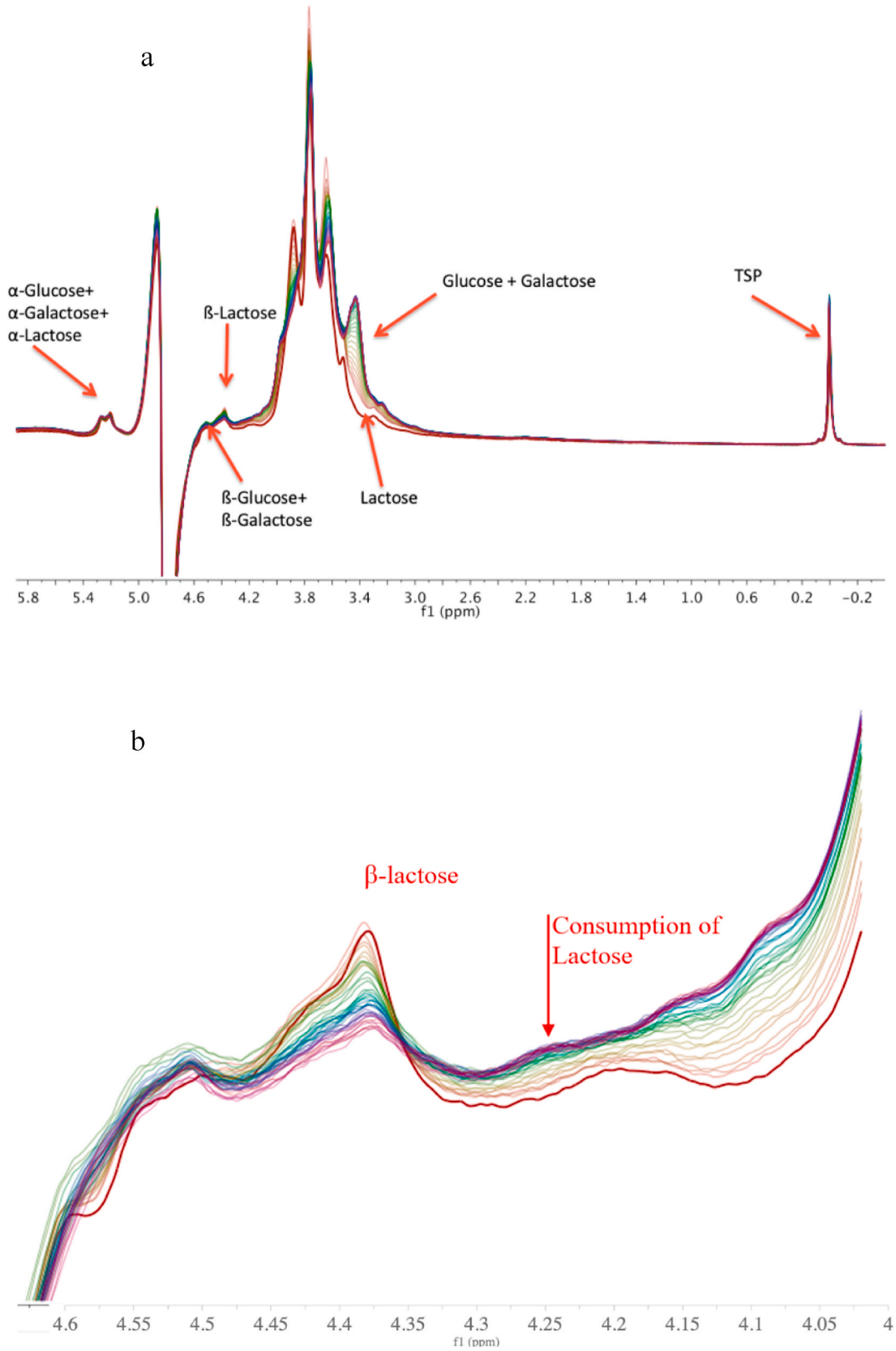

Fig. 2. a) Full superimposed spectra of lactose hydrolysis of model lactose solution **b)** β-lactose region.

was given in Fig. 4a. As seen in the spectrum, there were broad fat peaks between 0.5 and 2.5 ppm. For the differentiation experiment of the samples, integral area of *glycerol, lactose* and *fat* peak were used. The imidazole peak integral area was used for the normalization of other peaks. The normalized integral areas are given in Table 2. Lactose-free milks (LFM(T/F)) were found to have no lactose and half-fat milks (HFM(T/F)) had nearly half fat content compared to whole milk samples.

Glycerol peak was observed next to β-lactose peak between 4.0 and 4.25 ppm. When the glycerol contents were compared, Turkish and

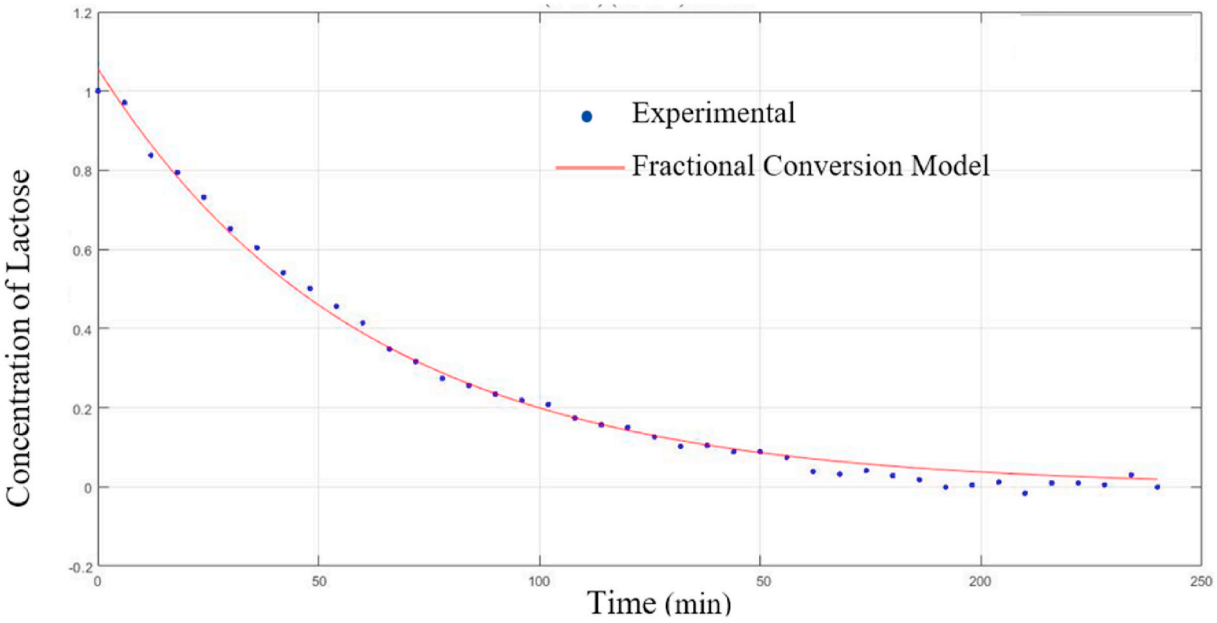

Fig. 3. Hydrolysis of lactose for the model system.

French milks were found to be significantly different from each other; glycerol contents of French milks being higher ($p < 0.05$).

Two reasons could suggest for the presence of glycerol in the milk. Firstly, glycerol is used for the treatment of ketosis which is a metabolic disease caused by the negative energy balance during lactation (Johnson, 1954). And secondly, glycerol is used in the feed as an energy source. Expansion of the biodiesel industry decreased the availability of corn for animal feed and increased its prices. Therefore, glycerol which is also a by-product of the biodiesel industry has started to be used as an energy source for cows (Carvalho, Schmelz-Roberts, White, Doane, & Donkin, 2011). The energy concentration of glycerol (1.98–2.29 Mcal/kg) is almost equal to corn starch (Schröder & Südekum, 1999). And inclusion of that in the diet was eventually observed in the milk.

According to ANOVA results, only the fat and sugar status of the milk were found to be significant for fat and lactose contents ($p < 0.05$). Production type, country and processing were found to be insignificant on lactose and fat content ($p > 0.05$).

3.1.5. *Artificial Neural Network (ANN) model*

As stated before, ANN was used as the multivariate approach for classification of the different milk samples based on the outputs of the NMR spectra. The most commonly used quantitative prediction multivariate technique is *Partial Least Square Regression* (PLSR), which is mostly used to analyze spectroscopic data, especially FTIR. However, ANN was found to show much better differentiative prediction ability as compared to other classification models (Efenberger-Szmechtyk, Nowak, & Kregiel, 2018). This is explained with several reasons. ANNs learn relationship between the input variables and the output values through successive trainings and the non-linear technique allows model to fit the data better. Analysis is noise insensitive which provides accurate prediction in the presence of noisy data (Berrueta, Alonso-Salces, & Heberger, 2007). There has been several studies on foods which show that ANN is superior to PLSR in prediction (Panagou, Mohareb, Argyri, Bessant, & Nychas, 2011; Perai, Moghaddam, Asadpour, Bahrampour, & Mansoori, 2010; Vasquez et al., 2018).

In this work, ANN model has been designed and optimized to discriminate milk samples according to their glycerol, lactose and fat peak areas obtained from the 1D NMR spectra. The mathematical tool employed six categorical parameters (country, fat content, sugar content, milk source, production type and processing type) as variables to accurately classify the milk samples.

RMSECV values of the model were used to determine ruggedness of the model. The breakpoint in the RMSECV *vs*. iteration number (step) plot were identified as the number of iteration (step) parameter for the analyses. Following such an approach prevented overtraining of the model. Since the RMSECV values are usually an integral part of the ANN analysis, they were not saved thus not reported in the model. Feedforward supervised method was used for training ANN model with 5 neurons and 2 neurons in the two hidden layers. Multilayer feedforward neural network diagrams are also given in the supplement. RMSEC value was obtained as 0.017 and R^2_{cal} was 0.997. When external validation data was fitted to model, RMSEP was observed as 0.128 and R^2_{pred} was 0.819 and these values indicated good prediction. Each sample was correctly classified in the cross validation (24/24), while one of the 12 samples was misclassified in the external validation set (Fig. 6). These results showed that glycerol, fat and lactose contents, supplied with information such as production type, processing method, milk source and analysed with ANN, could be used to locate the origin (*geography*) of milk.

ANN was also used to differentiate organic milk samples from conventional milks samples, and pasteurized milk samples from UHT processed milk samples. For classification of organic and conventional samples, RMSEC value was obtained as 0.121 and R^2_{cal} was 0.893. When external validation data was fitted to model, RMSEP was observed as 0.232 and R^2_{pred} was 0.419. Although R^2_{pred} was relatively low, 0.42 is adequate for discrimination between distinct classes. Prediction ability of model according to samples' production type was 24 out of 24 for calibration set and 9 out of 12 for validation set. In the validation set, two organic samples were misclassified as conventional and one sample was misclassified as organic.

Same ANN parameters were used to classify pasteurized milk samples and UHT processed milk samples. RMSEC for calibration set was 0.018 and RMSEP for prediction was 0.153. R^2 for calibration was obtained as 0.999 while R^2 for prediction was 0.851. Each sample in calibration set and pasteurized samples in validation set were classified correctly while two of UHT processed milk samples in validation set were misclassified. R^2 for prediction value showed that model was good enough for prediction.

In summary, the ANN model was successfully used to predict origins of milk samples, production type and process type. For all parameters

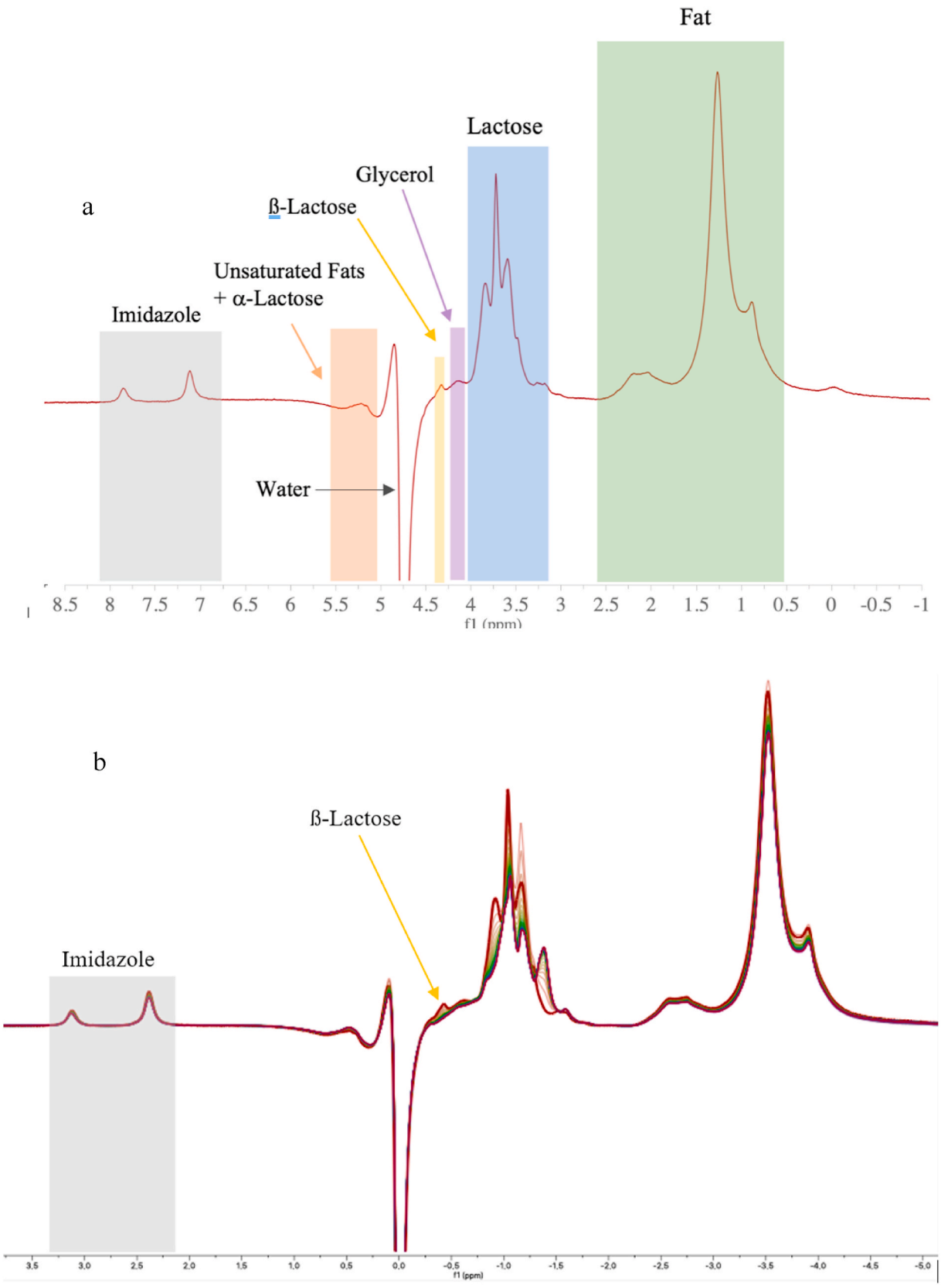

Fig. 4. 1D ^{1}H NMR spectrum of milk with the assignment for its main constituents (a) Superimposed spectra of lactose hydrolysis in milk (b).

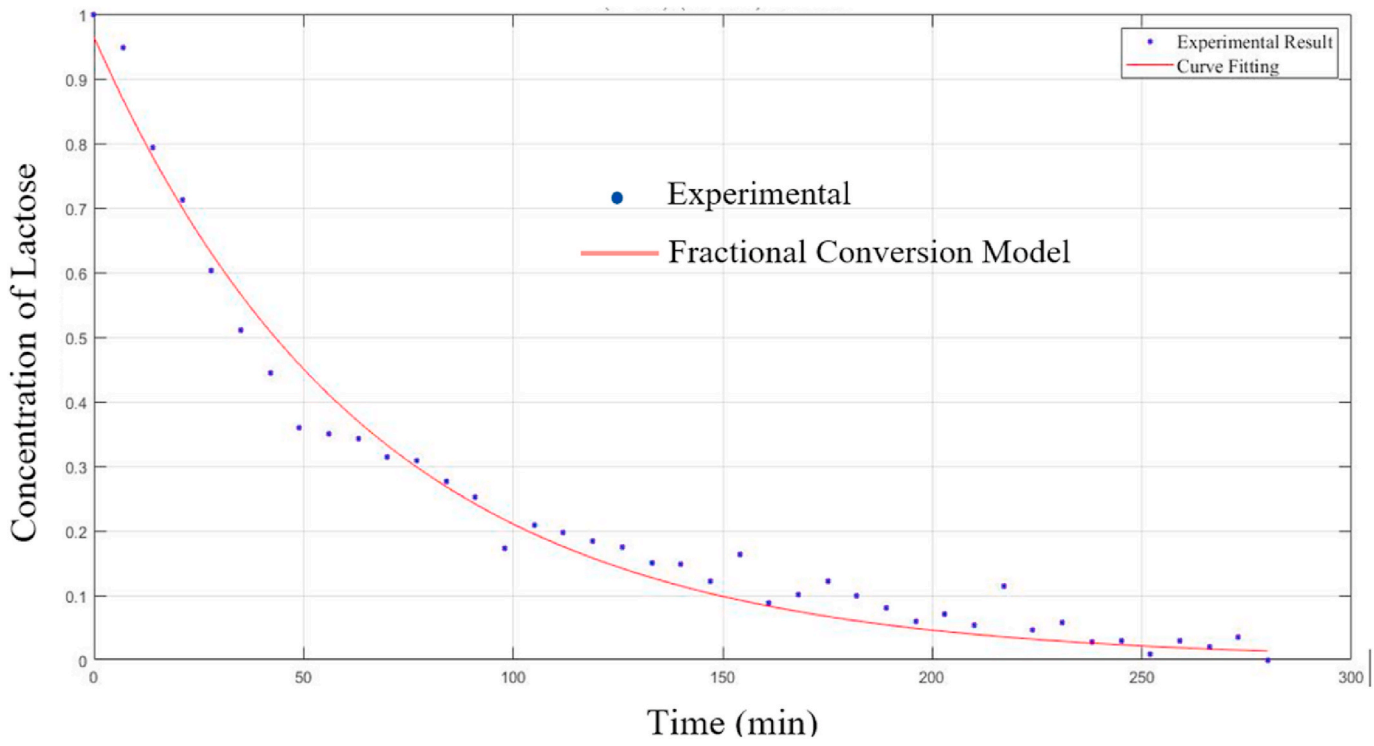

Fig. 5. Hydrolysis of lactose in a real milk sample.

Table 2
Integral areas of glycerol, lactose and fat NMR peaks of analysed milks, normalized to the imidazole signal.

Sample Name	Glycerol	Lactose	Fat	Sample Name	Glycerol	Lactose	Fat
WMF-1	1.126 ± 0.0003	18.023 ± 0.0108	40.585 ± 0.0209	GMF-19	0.723 ± 0.0006	20.1 ± 0.0301	27.899 ± 0.0089
WMF-2	1.204 ± 0.0037	21.068 ± 0.0582	49.632 ± 0.1252	GMF-20	0.749 ± 0.0004	17.719 ± 0.0146	47.585 ± 0.0178
WMF-3	0.845 ± 0.0005	18.416 ± 0.0047	44.177 ± 0.0169	GMF-21	0.841 ± 0.0017	20.372 ± 0.0191	46.935 ± 0.0752
WMF-4	1.022 ± 0.0019	21.37 ± 0.0097	49.785 ± 0.0029	GMF-22	1.006 ± 0.0012	18.727 ± 0.0215	52.168 ± 0.0439
OMF-5	1.006 ± 0.0001	18.631 ± 0.0031	47.615 ± 0.0068	WMT-23	0.863 ± 0.0006	19.08 ± 0.0208	44.018 ± 0.0429
OMF-6	1.043 ± 0.0003	20.927 ± 0.0033	55.941 ± 0.0076	WMT-24	1.007 ± 0.0024	19.483 ± 0.0295	48.226 ± 0.0474
OMF-7	0.576 ± 0.0010	19.352 ± 0.0090	48.897 ± 0.0210	WMT-25	0.279 ± 0.0020	22.412 ± 0.0668	45.022 ± 0.0609
HFMF-8	0.687 ± 0.0008	20.311 ± 0.0015	25.845 ± 0.0284	WMT-26	0.428 ± 0.0010	18.615 ± 0.0257	41.632 ± 0.0496
OHFMF-9	0.275 ± 0.0002	19.461 ± 0.0211	20.38 ± 0.0286	WMT-27	0.772 ± 0.0004	18.911 ± 0.0092	44.755 ± 0.0124
HFMF-10	0.472 ± 0.0004	20.384 ± 0.0047	25.257 ± 0.0235	WMT-28	0.23 ± 0.0006	19.651 ± 0.0275	44.747 ± 0.0299
HFMF-11	0.195 ± 0.0012	22.737 ± 0.0488	24.772 ± 0.0430	OMT-29	0.175 ± 0.0005	22.332 ± 0.0176	46.65 ± 0.0544
HFMF-12	0.55 ± 0.0004	22.638 ± 0.0188	25.964 ± 0.0748	GMT-30	0.163 ± 0.0012	19.955 ± 0.0248	44.893 ± 0.0698
OHFMF-13	0.7 ± 0.0015	23.062 ± 0.0313	28.13 ± 0.0998	LFMF-31	0.436 ± 0.0014	0	24.89 ± 0.0947
WMF-14	0.956 ± 0.0007	21.808 ± 0.0065	57.543 ± 0.0135	LFMF-32	0.123 ± 0.0004	0	24.288 ± 0.0186
OHFMF-15	0.522 ± 0.0009	18.763 ± 0.0165	22.294 ± 0.0364	LFMT-33	0.412 ± 0.0016	0	22.52 ± 0.0412
HFMF-16	0.502 ± 0.0052	18.866 ± 0.04292	22.594 ± 0.1573	LFMT-34	0.297 ± 0.0009	0	23.29 ± 0.01711
OMF-17	0.817 ± 0.0007	22.858 ± 0.0236	57.398 ± 0.0498	LFMT-35	0.373 ± 0.0006	0	20.793 ± 0.0204
OHFMF-18	0.482 ± 0.0005	20.006 ± 0.0202	26.536 ± 0.0403	LFMT-36	0.059 ± 0.0417	0	27.991 ± 0.0841

evaluated in the study, the maximum number of incorrect matches was 2 which was a good indicator for the strong classification ability of the model. These results confirmed the existence of a relationship between the six categorical parameters and three main components in the classification of milks, leading to potential design of an objective evaluating method for classification of diverse milk samples. The model supported potential use of benchtop NMR spectra for differentiating milk samples and provided input for a more comprehensive authentication study.

4. Conclusion

Milk is a very valuable food for the human being. Ensuring its quality and authenticity is significant. In this study, multiple aspects of benchtop NMR spectroscopy have been evaluated to investigate its potential as a multi-scale control tool at different stages of milk production. The first part of the study focused on the online monitoring of lactose hydrolysis in milk. Enzymatic hydrolysis of lactose was successfully followed and a kinetic model was obtained in agreement with previous studies. In the second part, milk samples that have different properties were examined. Glycerol, fat, sugar content of the milks was obtained from the NMR spectra and an Artificial Neural Network (ANN) model was developed to seek for classifications based on the production, processing type, milk source and geographic origin. The developed model supported the potential use of benchtop NMR spectra for differentiating milk samples. Overall, it was shown that even for a complex fluid like milk, benchtop NMR spectroscopy provides valuable information and the technique after a full analytical validation that is out of the scope of this proof-of-concept study has the potential to be used in industry for monitoring

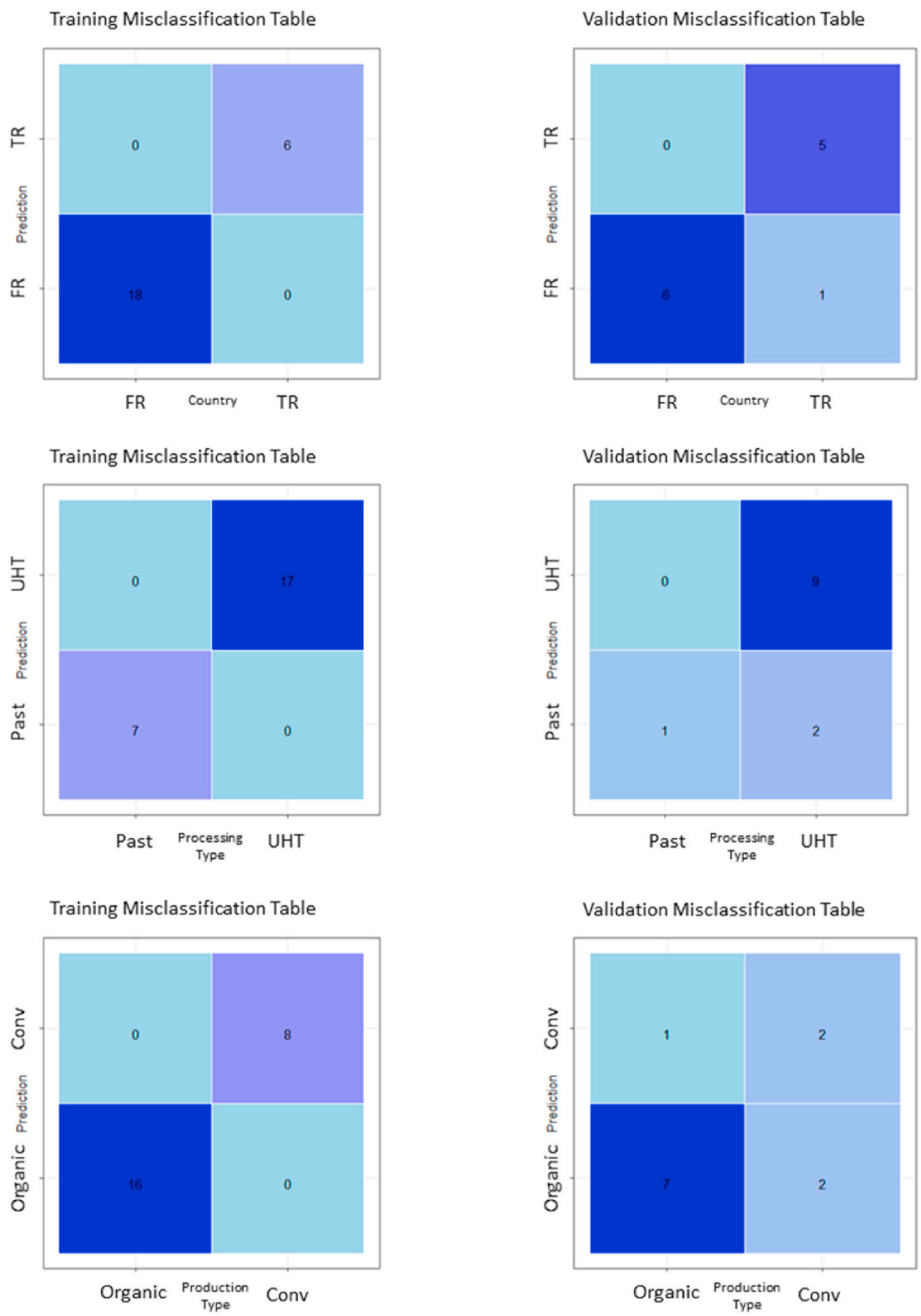

Fig. 6. The confusion matrices for training dataset (left) and test dataset (right).

production, quality control and authenticity determination (*but on larger sample set*).

Author contribution

Alper Soyler: The study is a part of the PhD dissertation of. Farjon and Bouillaud helped Soyler in the experiments. Cikrikci: helped to arrange the data. Cavdaroglu conducted and wrote the statistical analysis. Oztop: was one of the major advisors of the study.. Giraudeau was one of the major advisors of the study.

Funding

Alper Soyler got funding from Short Term Scientific Mission (STSM) program under COST Action CA15209 EURELAX – European Network on NMR Relaxometry (http://www.cost.eu/COST_Actions/ca/CA15209) to conduct experiments at the University of Nantes. Authors from Nantes acknowledge support from the Corsaire metabolomics core facility (Biogenouest). J.F. acknowledges Sandrine Bouchet for a continuous essential assistance. Middle East Technical University, Scientific Research Projects Coordination Unit with grant # BAP-07-02-2017-004-077 is also acknowledged.

Declaration of competing interest

All authors declare no conflict of interest. Publication has been approved by all individual participants.

References

Anderssen, K. E., & McCarney, E. R. (2020). Online monitoring of enzymatic hydrolysis of marine by-products using benchtop nuclear magnetic resonance spectroscopy.

Food Control, 112(September 2019), 107053. https://doi.org/10.1016/j.foodcont.2019.107053

Bergamaschi, M., Cipolat-Gotet, C., Cecchinato, A., Schiavon, S., & Bittante, G. (2020). Chemometric authentication of farming systems of origin of food (milk and ripened cheese) using infrared spectra, fatty acid profiles, flavor fingerprints, and sensory descriptions. *Food Chemistry, 305*, Article 125480. https://doi.org/10.1016/j.foodchem.2019.125480

Berrueta, L. A., Alonso-Salces, R. M., & Heberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A, 1158*(1–2), 196–214. https://doi.org/10.1016/j.chroma.2007.05.024

Bertram, H. C. (2018). NMR-based metabolomics: Quality and authenticity of milk and meat. In *Modern magnetic resonance* (pp. 1–13). Springer.

Bila, S., Harkouss, Y., Ibrahim, M., Rousset, J., N'Goya, E., Baillargeat, D., et al. (1999). Accurate wavelet neural-network-based model for electromagnetic optimization of microwave circuits. In *International journal of RF and microwave computer-aided engineering*. https://doi.org/10.1002/(SICI)1099-047X (199905)9:3<297::AID-MMCE13>3.0.CO;2-W.

Bouillaud, D., Drouin, D., Charrier, B., Jacquemmoz, C., Farjon, J., Giraudeau, P., et al. (2020). Using benchtop NMR spectroscopy as an online non-invasive in vivo lipid sensor for microalgae cultivated in photobioreactors. *Process Biochemistry, 93*, 63–68. https://doi.org/10.1016/j.procbio.2020.03.016. January.

Bouillaud, D., Farjon, J., Gonçalves, O., & Giraudeau, P. (2019). Benchtop NMR for the monitoring of bioprocesses. *Magnetic Resonance in Chemistry, 57*(10), 794–804. https://doi.org/10.1002/mrc.4821

Bouillaud, D., Heredia, V., Castaing-Cordier, T., Drouin, D., Charrier, B., Gonçalves, O., et al. (2019). Benchtop flow NMR spectroscopy as an online device for the in vivo monitoring of lipid accumulation in microalgae. *Algal Research, 43*, 101624. https://doi.org/10.1016/j.algal.2019.101624. July.

Buitenhuis, A. J., Sundekilde, U. K., Poulsen, N. A., Bertram, H. C., Larsen, L. B., & Sørensen, P. (2013). Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk. *Journal of Dairy Science, 96*(5), 3285–3295. https://doi.org/10.3168/jds.2012-5914

Carvalho, E. R., Schmelz-Roberts, N. S., White, H. M., Doane, P. H., & Donkin, S. S. (2011). Replacing corn with glycerol in diets for transition dairy cows. *Journal of Dairy Science, 94*(2), 908–916. https://doi.org/10.3168/jds.2010-3581

Churakova, E., Peri, K., Vis, J. S., Smith, D. W., Beam, J. M., Vijverberg, M. P., et al. (2019). Accurate analysis of residual lactose in low-lactose milk: Comparing a variety of analytical techniques. *International Dairy Journal, 96*, 126–131. https://doi.org/10.1016/j.idairyj.2019.02.020

Danieli, E., Perlo, J., Blümich, B., & Casanova, F. (2010). Small magnets for portable NMR spectrometers. *Angewandte Chemie International Edition, 49*(24), 4133–4135. https://doi.org/10.1002/anie.201000221

Danieli, E., Perlo, J., Blümich, B., & Casanova, F. (2013). Highly stable and finely tuned magnetic fields generated by permanent magnet assemblies. *Physical Review Letters, 110*(18), Article 180801. https://doi.org/10.1103/PhysRevLett.110.180801

Defernez, M., Wren, E., Watson, A. D., Gunning, Y., Colquhoun, I. J., Le Gall, G., et al. (2017). Low-field 1H NMR spectroscopy for distinguishing between arabica and robusta ground roast coffees. *Food Chemistry, 216*, 106–113. https://doi.org/10.1016/j.foodchem.2016.08.028

Dong, L., & Zhong, Q. (2019). Dispersible biopolymer particles loaded with lactase as a potential delivery system to control lactose hydrolysis in milk. *Journal of Agricultural and Food Chemistry, 67*(23), 6559–6568. https://doi.org/10.1021/acs.jafc.9b01546

Dutra Rosolen, M., Gennari, A., Volpato, G., & Volken De Souza, C. F. (2015). *Lactose hydrolysis in milk and dairy whey using microbial β-galactosidases*. Enzyme Research, 2015. https://doi.org/10.1155/2015/806240

Efenberger-Szmechtyk, M., Nowak, A., & Kregiel, D. (2018). Implementation of chemometrics in quality evaluation of food and beverages. *Critical Reviews in Food Science and Nutrition, 58*(10), 1747–1766. https://doi.org/10.1080/10408398.2016.1276883

Fritsch, S., Guenther, F., Wright, M. N., Suling, M., Sebastian, M., & Mueller. (2019). *Package "neuralnet" training of neural networks* (Vols. 1–15).

Gouilleux, B., Charrier, B., Akoka, S., & Giraudeau, P. (2017). Gradient-based solvent suppression methods on a benchtop spectrometer. *Magnetic Resonance in Chemistry, 55*(2), 91–98. https://doi.org/10.1002/mrc.4493

Jakes, W., Gerdova, A., Defernez, M., Watson, A. D., McCallum, C., Limer, E., et al. (2015). Authentication of beef versus horse meat using 60 MHz 1H NMR spectroscopy. *Food Chemistry, 175*, 1–9. https://doi.org/10.1016/j.foodchem.2014.11.110

Jansson, T., Jensen, H. B., Sundekilde, U. K., Clausen, M. R., Eggers, N., Larsen, L. B., et al. (2014). Chemical and proteolysis-derived changes during long-term storage of lactose-hydrolyzed ultrahigh-temperature (UHT) milk. *Journal of Agricultural and Food Chemistry, 62*(46), 11270–11278. https://doi.org/10.1021/jf504104q

Johnson, R. B. (1954). The treatment of ketosis with glycerol and propylene glycol. *Cornell Veterinarian, 44*(1), 6–21.

Knox, S. T., Parkinson, S., Stone, R., & Warren, N. J. (2019). Benchtop flow-NMR for rapid online monitoring of RAFT and free radical polymerisation in batch and continuous reactors. *Polymer Chemistry, 10*(35), 4774–4778. https://doi.org/10.1039/c9py00982e

Monakhova, Y. B., Kuballa, T., Leitz, J., Andlauer, C., & Lachenmeier, D. W. (2012). NMR spectroscopy as a screening tool to validate nutrition labeling of milk, lactose-free milk, and milk substitutes based on soy and grains. *Dairy Science & Technology, 92*(2), 109–120. https://doi.org/10.1007/s13594-011-0050-5

Panagou, E. Z., Mohareb, F. R., Argyri, A. A., Bessant, C. M., & Nychas, G.-J. E. (2011). A comparison of artificial neural networks and partial least squares modelling for the rapid detection of the microbial spoilage of beef fillets based on Fourier transform infrared spectral fingerprints. *Food Microbiology, 28*(4), 782–790. https://doi.org/10.1016/j.fm.2010.05.014. SI.

Perai, A. H., Moghaddam, H. N., Asadpour, S., Bahrampour, J., & Mansoori, G. (2010). A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. *Poultry Science, 89*(7), 1562–1568. https://doi.org/10.3382/ps.2010-00639

Perlo, J., Demas, V., Casanova, F., Meriles, C. A., Reimer, J., Pines, A., et al. (2005). High-resolution NMR spectroscopy with a portable single-sided sensor. https://doi.org/10.1126/science.1108944.

Rönnols, J., Danieli, E., Freichels, H., & Aldaeus, F. (2020). Lignin analysis with benchtop NMR spectroscopy. *Holzforschung, 74*(2), 226–231. https://doi.org/10.1515/hf-2018-0282

Saeys, W., Mouazen, A. M., & Ramon, H. (2005). Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosystems Engineering, 91*(4), 393–402. https://doi.org/10.1016/j.biosystemseng.2005.05.001

Schröder, A., & Südekum, K.-H. (1999). Glycerol as a by-product of biodiesel production in diets for ruminants. In *Proc. 10th Int. Rapeseed Congr.: New Horizons for an Old Crop, 2* p. 241).

Sobolev, A. P., Thomas, F., Donarski, J., Ingallina, C., Circi, S., Cesare Marincola, F., et al. (2019). Use of NMR applications to tackle future food fraud issues. *Trends in Food Science & Technology, 91*, 347–353. https://doi.org/10.1016/j.tifs.2019.07.035. March 2018.

Soyler, A., Bouillaud, D., Farjon, J., Giraudeau, P., & Oztop, M. H. (2020). Real-time benchtop NMR spectroscopy for the online monitoring of sucrose hydrolysis. *Lebensmittel-Wissenschaft & Technologie, 118*, Article 108832. https://doi.org/10.1016/j.lwt.2019.108832

Uncu, O., & Ozen, B. (2019). A comparative study of mid-infrared, UV-Visible and fluorescence spectroscopy in combination with chemometrics for the detection of adulteration of fresh olive oils with old olive oils. *Food Control, 105*, 209–218. https://doi.org/10.1016/j.foodcont.2019.06.013

Vasquez, N., Magan, C., Oblitas, J., Chuquizuta, T., Avila-George, H., & Castro, W. (2018). Comparison between artificial neural network and partial least squares regression models for hardness modeling during the ripening process of Swiss-type cheese using spectral profiles. *Journal of Food Engineering, 219*, 8–15. https://doi.org/10.1016/j.jfoodeng.2017.09.008

Zheng, H., Jiang, L., Lou, H., Hu, Y., Kong, X., & Lu, H. (2011). Application of Artificial Neural Network (ANN) and Partial Least-Squares Regression (PLSR) to predict the changes of anthocyanins, ascorbic acid, total phenols, flavonoids, and antioxidant activity during storage of red bayberry juice based on fractal ana. *Journal of Agricultural and Food Chemistry, 59*(2), 592–600. https://doi.org/10.1021/jf1032476