Chip-size, low cost near infrared sensors for milk analysis with lab grade performance

M. Medhat, M. Said, T. Elarabi, A. Abd El-Fadeel, W. Sorour and B. George

Si-Ware Systems, Heliopolis, Cairo, PO 11361, Egypt
Corresponding Author: www.si-ware.com

Abstract

In 2019, dairy market was valued at more than $ 600 Billion, and it is projected to exceed 1 trillion $ by 2024, making it one of the largest food sectors worldwide. Milk is also considered one of the most foods consumed worldwide. Therefore, milk composition and quality are two crucial factors to be measured. Milk components are measured by various technologies; most common is ultrasonic in addition to spectroscopic techniques like MIR technology. The issue with these two techniques are mainly accuracy in case of ultrasonic and very high price in case of MIR spectroscopy. In this work, the robustness and efficiency of the chip-size, low cost near-infra red sensors NeoSpectra® to accurately determine the milk components are demonstrated. Compared to other competing technologies in the market as well as some benchtop instruments, NeoSpectra is showing a much better performance and exceeding ICAR on farm standards and even close to lab standards.

The prediction models are built using cow milk taken from various regions/herds. 110 raw, fresh cow milk samples were collected from 3 different breeds and 5 different herds. Chemical analysis for fat, protein and lactose is performed using official methods of analysis of AOAC international 19th edition. Milk measurements using NeoSpectra sensors are collected using milk samples in off the shelf glass beakers and milk is filling at least 1 cm from the bottom of the beaker.

As a summary of the results, fat is measured using NeoSpectra sensors and prediction models with a coefficient of determination (R^2 = 0.984) and (RMSE=0.158) and testing results (R^2=0.991, RMSE=0.167, SD (standard deviation) =0.147). Protein is measured with (R^2 = 0.891) and (RMSE=0.078) and testing results (R^2=0.979, RMSE=0.127, SD=0.131). Lactose is measured with (R^2= 0.818) and (RMSE=0.105) and testing results (R^2=0.767, RMSE=0.111, SD= 0.114). These results show that NeoSpectra sensors and prediction models surpassed ICAR's standard for on farm analyzer which requires SD (standard deviation) = 0.25 for fat, protein and are very close to ICAR's standard for lab analyzer which requires SD = 0.1(1, 2).

In conclusion, the NeoSpectra® NIR sensors showed the ability to enable low cost milk analyzers that not only complies with ICAR standard for on-farm milk analyzers, but also exceeds it to a performance level that is very close to ICAR lab analyzer requirements.

Introduction

By 2050 the world population is expected to grow to 9.8 billion according to the UN, with that fast-growing population the demand for securing a continuous food supply at reasonable cost is pushing the farmers to the edge, so far dairy farming industry has done a remarkable progress in terms of selecting and tracking the best breeds, cow genetics improvements, artificial insemination and feed management. Farmers also spend time, effort and money on applying these high standard measures expecting an optimum outcome in terms of reproduction and milk quality and quantity. However, without a reliable proper feedback system the farm stays vulnerable to a revenue leak incidents that might not be observed until it is too late.

Among the required feedbacks for the farm managers are the real-time status of the individual animal health, the quality of milk produced per cow and the cow group performance. The industry is fully aware of the importance of these feedbacks and hence on organizational level the milk recording organizations (MRO) do a periodical farm visits. Milk samples from the milking herd are collected and analyzed. This process takes place every six weeks for the MROs' member farms. Also, on milking equipment manufacturing level, the industry tries to deploy the possible sensory solutions to afford such in-field feedbacks for its customer farms.

The MROs send the farm milk samples to a central lab and get back to the farmer with a results report within 48 hours after sampling. The report mainly involves milk constituents of fat, protein and lactose in addition to somatic-cell count (SCC) analysis. Through such data the MRO can advise the farm about a potential or existing metabolic disorder, feed diets and concentrates for cow groups and also clinical or sub-mastitis in specific cows. However, in such cases the lack of real-time data could cause the farmer to endure late decision losses.

As for automatic milking systems (AMS) manufacturers the endeavors for developing in-line and on-line measurements to achieve such real-time monitoring has been ongoing for a few years. Most of these on-line systems use reagents within its analysis process which makes it an expensive solution and complex for installation in parlors. There are other inline solutions that depend on spectroscopy technology in the near infrared range which makes it reagent-less which counts as an advantageous point, but those systems did not achieve the required accuracy yet.

Speaking of the accuracy of milk constituents' analysis, the ICAR organization have set a defining measurements maximum permissible deviation for in-line and at-line systems to be certified [1, 2]. The ICAR organization did not certify according to the author's knowledge any on-farm milk analyzers yet indicating that the industry with its different parties need to work on fulfilling this gap. A feedback solution based on real-time milk constituents' analysis to be optimum should fulfil ICAR standards, prove cost effectiveness for the farmer to deploy it and be appropriate for working in on-farm or parlor conditions. These three parameters should be realized all along in one solution to be a reliable one for the farmer.

Infrared spectroscopy has been a cornerstone for milk analysis for a long time. In fact, some mid-infrared spectrometer (MID-IRS) analyzers are currently certified according to the ICAR laboratory standards but the MID-IRS is an expensive technology to be deployed at the farm and bulky to fit in field. Nowadays, miniaturizing spectrometers and specifically near-infrared (NIRS) using Micro-electro-mechanical (MEMS) techniques has made the technology more cost affordable and more diminished in size to fit in field use. NIRS miniaturized systems successfully achieved two out of the three parameters required for being a reliable system but in a seek of cost and size reduction the accuracy of measurement is slightly sacrificed, balancing those three parameters at such micro size is a very slick iterative work.

Si-ware Systems Company has developed its own version of miniaturized Fourier Transform Near-Infrared Spectrometer (FT-NIRS) called NeoSpectra. Among the

market of miniature NIRSs, NeoSpectra offers the widest measuring spectral range of 1,350 – 2,500nm. At such spectral range Si-ware has put an effort and resources to examine, iterate and improve its sensor for the dedicated use of milk constituents' analysis to reach the optimum balance of accuracy, speed and cost effectiveness requirements.

Building-up on the study presented in ICAR conference in Prague 2019 showing the performance of different NIRSs versus ICAR standards [3] we present in this work the NeoSpectra measurements, models used, reference methods applied and standard deviation (SD) resulted of raw farm milk analysis for fat, protein, lactose. Then, a conclusion of the results is presented comparing the proposed solution to the available market solutions demonstrating what the industry have accomplished to fill in the gap of presenting a real-time milk analysis reliable feedback solution for dairy farmers.

Near infrared spectroscopy theory

Spectroscopy is the study of physical and/or chemical properties of materials by analyzing their response to light. Knowing that each chemical component has a unique spectral pattern, the analysis of the spectral response of matters tells a lot about their chemical composition and/or concentration. Today, spectrometer instruments can be found in labs and industrial environments for material identification and/or quantification in different application areas. There are many conventional topologies for spectrometry instrumentation including Fourier Transform InfraRed (FT-IR) that offers several performance and cost advantages. (FT-IR) spectroscopy has evolved over several decades because of its necessity for various applications throughout the physical, chemical, and biological domains. Spectroscopy is used for identification of different kinds of materials (qualitative analysis), or in quantifying the amount of materials which is called quantitative analysis. Compared to the other technologies like dispersive instruments, FT-IR devices provide multiple advantages such as the wide spectral range, the multiplex (Felleget) advantage, higher optical throughput, the use of a single photodetector in addition to wavenumber range flexibility and measurement speed [4]. Example applications include medical analysis [5,6], food quality control [7, 8, 9] and soil analysis [10]. Near infrared (NIR) spectral instruments are having great advantage due to the lower cost of its optical components and also good performance compared to the mid infrared (MIR) range [11]. In addition, samples do not require special preparation, which preserves the sample integrity after completing the analysis. The sample can be measured either in transmission mode or in diffuse reflectance mode, where the latter is the method of choice for most solid samples or liquid samples containing scattering sites. A good comparison about different spectroscopy solutions for milk analysis shows the advantages of different ranges and the disadvantages [12].

Since every material has a unique response to light, the analysis of the light interacting with a certain material can reveal a lot of information about its composition. The analysis of light to determine the properties of materials is what spectroscopy is all about. Light interacting with materials is essentially a spectrum of electromagnetic waves with different wavelengths. Fundamentals of bands absorbed by materials are found in Mid-Infrared (MIR) region of light (wavelengths >3,500 nm). The Near Infrared (NIR) region of light is absorbed by the energy bands related to the overtones and combinations of the fundamental bands.

The unique information from the vibrational absorption bands of a molecule is reflected in the NIR spectrum, but some spectral numerical processing and statistical analysis are required to "unlock" this information. The application of statistical methods to the analysis of experimental data is known as chemometrics. NIR spectroscopy has long been used as a material analysis tool. It provides various advantages over other analysis methods including:

- Tests are non-destructive.
- Almost no sample preparation is required.
- Different parameters can be measured simultaneously.
- Ability to obtain results instantly.

Conventional FT-IR spectrometers are bulky, expensive, and sensitive to vibrations limiting their usage to the lab or controlled environments. There is a huge demand for a portable device that can be used for inline or field applications [13]. That is why many new emerging technologies have gained a lot of attention in the past decade in different industries. One of the main enabling technologies is the micro-electro-mechanical system (MEMS) [14, 15, and 16]. There are many MEMS based spectrometers reported in literature based on different technologies, including diffraction gratings [17], micro-mirror devices (DMDs) [18], multimode interference (MMI) interferometers [19], linear variable filters with photo-detector array and tunable Fabry-Pérot filters [20]. These solutions are either limited in spectral range or not scalable enough to meet the growing need of the spectral sensing market. The adoption of NIR analysis instruments has been mostly limited to labs or in-line process analyzers. This limitation was mostly due to the fact that those instruments are very bulky. Recently, portable NIR analyzers have become commercially available. However, their high price tags limited their adoption to limited use cases, or their limited performance limited the application they could enable.

NeoSpectra solutions based on MEMS FTIR technology introduces a real lab performance compared to benchtop devices while at the same time having instance and cost effective solutions for in field applications. They are based on Fourier Transform Infrared (FT-IR) technology that offers a wide spectral range for the best qualification and quantification of materials. NeoSpectra solutions operate in the Near-Infrared (NIR) from 1,350 to 2,500 nm. This wavelength range goes up to the highest point of the NIR and is the widest range versus comparable solutions.

Looking on the milk analysis industry, the most commonly commercially available milk analysis solutions are either based on wet chemistry, Ultrasound technologies, or Mid Infrared Spectroscopy Technology.

Wet chemistry provides highly accurate results. However, it requires analytical lab professionals to perform lengthy analysis processes. Hence, they can only be performed in labs and it does not enable quick analysis checks. Ultrasound technology comes at relatively low cost and is relatively easy to use, making them more adoptable across the supply chain but their results are not reliable. The accuracy of the results is highly dependent on sample composition and conditions. Mid IR technologies enable highly accurate analysis and quickly. However, their costs and size limited their adoption to labs and large stakeholders in the supply chain, leaving the rest of the supply unchecked.

In the last decade new technologies that allow miniaturizing NIR based solutions have been emerging. While these solutions are quite promising in potentially enabling ubiquitous, easy to use and quick milk analyzers, achieving the accuracy level that is acceptable across the milk supply chain has been a major challenge. To date, to the best of the author's knowledge, there has been no commercially available solution based on NIR technology that satisfies ICAR standards for in-farm or in-lab use.

This raises the questions: How to bring the advantages of the NIR spectroscopy to the field for instant, non-destructive, easy to use, and cost effective milk analysis solution. NeoSpectra spectral sensing solutions presents a great answer for this question. The technology combines the advantages of MEMS technology, FTIR spectroscopy and

artificial intelligence to enable a chip-sized, cost effective, and scalable solutions. The solutions offer performance comparable to laboratory-based spectrometers, and at the same time with instant and cost-effective performance.

Experimental results

In order to demonstrate the capability of the technology in milk analysis, NeoSpectra solutions are used in testing different milk samples in the following experiment. In order to make sure that different variations of milk samples were covered, milk was lactated from different cow types, different locations, and different farms that use different farm management systems. In numbers: 110 cows from 3 different breeds and 5 different herds were lactated to provide 110 milk samples from four different villages in two different states. Upon lactation, milk samples were stored in an ice-box until it reached the lab and stored at -20C in the lab. Before collecting measurements from stored milk samples, the milk samples are pre-warmed at 30C for 20-30 minutes. To homogenize the samples, gentle stirring (10 times clockwise and 10 anti-clockwise) is performed before each measurement. The steps for developing a model that can be used in predicting the different milk parameters is shown in Figure 1. After milk collection the milk samples are splitted in to two identical groups. First group is used to build the reference database using the standard chemical analysis methods while the other group is used to measure the spectral reference measurements using NeoSpectra solutions.

The spectral data and their corresponding reference values can be used to build mathematical data models or train AI algorithms. Those models and algorithms can be used to predict the chemical composition of samples with unknown reference values. The following steps are required in order to build the analysis models that predicts the chemical composition of milk samples:

1. Collecting various milk samples. The collected samples should cover the different variations that are expected to be found in the real use-case. The number of samples to be collected usually vary depending on the expected variations in the sample to be analyzed. Typically, a sample set of around 100 samples is good enough to have a robust analysis model.
2. Measuring reference values for the parameters of interest using reliable reference methods.
3. Measuring the spectral data of the collected samples.
4. Splitting the collected data sets to 2 parts: A calibration set that is used to build the chemometrics model, and a validation set that is used to test how well the model

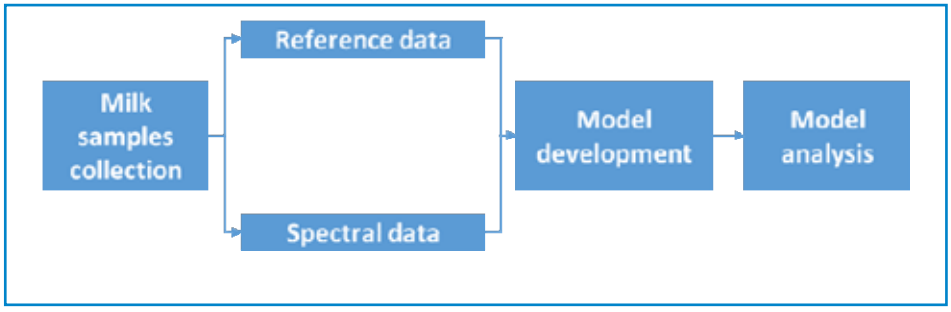


Figure 1. Model development steps.

can predict the values of the parameters for samples that were not part of the training data.

5. Data pre-treatment: Data are preprocessed to eliminate any variabilities that may affect the accuracy of the models to be developed. This includes removing the effect of any sample presentation irregularities and normalizing the different samples. [21]
6. Developing AI models using the calibration set.
7. Analyzing results by comparing the predicted values with the pre-known reference value. A study of the prediction error and bias evaluates the model performance.

Reference database building

In order to have high confidence in the prediction of NIR, it is important to make sure that accurate methods were used to collect the reference value used in building the analysis models. In other words, garbage input data to the modeling process would result in garbage output. Therefore, according to Procedure 1 of Section 12 of ICAR Guidelines - Protocol for the Evaluation of Milk Analysers for ICAR Approval the reference methods used should comply to standardized methods of at least one of the international guidelines (ISO, IDF, AOAC). In this study, chemical methods used to measure milk components, which complies with AOAC (Association of Official Analytical Chemists) official methods No. 990.20 (2012), as follows:

Total Protein Analysis – Kjeldahl

The total milk proteins were measured chemically according to Kjeldahl method. Briefly, solutions of 15.00 g K_2SO_4, 1 ml $CuSO_4$ catalyst solution and 5ml of warm milk (38°C) were mixed and weighed and immediately place in digestion flask. 25 ml H_2SO_4 were added, and digestion was conducted over heating device. Acid digest was cooled to room temperature. After digest, the mixture was cooled to room temperature, and 300 ml H_2O was added to flask and mixed by swirling. For distillation, the condenser water was turned on and 50 ml H_3BO_3 solution was added with indicator to graduated 500 ml Erlenmeyer titration flask and flask placed under condenser trip. 75 ml 50% NaOH was added to the diluted digest down sidewall of Kjeldahl flask with no agitation. Immediately flask was connected to distillation bulb on condenser. The ammonia formed is distilled into a boric acid solution containing the indicators bromocresol green and methyl red (≥ 200 ml total volume). H_3BO_3 receiving solution was titrated with standard 0.1 N HCL solution until the first trace of a pink color. In order to identify the amount of the total protein, the following equation was used.

Moles of HCl = moles of NH_3 = moles of N in the sample.

A reagent blank should be run to subtract reagent nitrogen from the sample nitrogen.

$$\%N = \frac{NHCl \times \text{Corrected acid volume}}{\text{g of sample} \times 1000} \times 14 \times 100$$

where:

NHCl = normality of HCl, in mol/1000 ml. Corrected acid vol. = (ml std. acid for sample) - (ml std. acid for blank). 14 = atomic weight of nitrogen. A factor is used to convert percent N to percent crude protein = 6.38. % Protein = % N * 6.38.

Total fat analysis – Gerber method

The total fat analysis is performed using the Gerber method. The procedure is as follow : Total of 10 ml of H_2SO_4 at 15–21 ℃ were added into a Gerber milk bottle, and using a Gerber pipette, 11 ml milk sample 11 ml were accurately added into the Gerber bottle, and 1ml of isoamyl alcohol was added to the bottle. The mixture was shaken and centrifuged briefly for 5 minutes. The mixture was then placed in a water bath at 60–63 ℃ and the fat content was measured from the graduations on the bottle neck. To get an accurate measure of the fat content, the previous steps were repeated three times for the same milk sample and the fat content was determined by taking the average of the three measurements.

Lactose content determination - Fehling's Test

The lactose content is measured using the Fehling test. The procedure is as follow: 25 ml of milk was transferred to a standard 250 mL flask and 50 ml of distilled water was added in addition to 5 ml of ferrocyanide potassium solution and 5 ml zinc acetate (with flipping after each addition), and filtrated. 10 ml of Fehling's solution A ($CuSO_4.5H_2O$) was added into a conical flask followed by Fehling's solution B (sodium potassium tartrate solution) and 25 ml of the filtrate was added into the conical flask containing solutions A and B and heated to boil for two minutes. Then, three drops of methylene blue indicator were added into the boiling solution and filtrated drop wise at intervals of 10 seconds until the blue color of the methylene blue indicator disappears. The volume of filtrate used at the end point of the reaction (red color) was measured and Lactose % was calculated from *Lane-Eynone Table*.

Spectra data collection

The spectral data collection procedure should follow the same procedure that shall be used in the prediction of real use-case measurements. Therefore, it is important to keep in consideration that the lab setup used in the experiment shall enable an easy to use, stable, and repeatable sample handling and measurement procedure.

The collection of spectral data consists of two main steps:

- Background measurements: This step is important to calibrate the spectrometer to a reference spectral response and minimize any contribution from the instrument on the output spectrum. To do so, a disc of a white reference material "Spectralon" that has a flat spectral response over the NIR range is placed on top of the NeoSpectra sensor, and a background measurement is collected. This step, despite its importance, doesn't necessarily have to be done every time a new spectrum is collected, but should be done in the beginning of the measurements session or whenever any variations in the setup or the environment in which the experiment is taking place.
- Sample measurement: The samples are poured in a beaker that has a thin, flat bottom surface. An off the shelf beaker is used in this experiment which is commercially available and the material of the beaker should be transparent in the NIR range. This makes sample handling simple and an easy procedure which does not need any sample preparation. The user should only make sure that the sample has a height or thickness larger than 10 mm inside the beaker. This is to ensure that the sample thickness is larger than the penetration depth of NIR light in the milk. The beaker is then simply placed on top of the NeoSpectra sensor and a measurement is acquired. The selected scan time for each scan is 5 s.

The output spectrum consists of the ratio of sample measurement and background measurement. For each sample, 3 measurements are collected and the position of the beaker is changed after each measurement.

Data analysis and modeling

The collected spectral data for all samples were tabulated with their corresponding reference values. Data was split into two groups:

- 90% samples for calibration set.
- 10% samples for validation set.

In order to ensure the representativeness of the data used to build and validate the analysis models, the selection of the data sets was made in a way that ensures that reference values of the different parameters are well distributed across the value ranges.

To check the accuracy of the prediction models a comparison is performed between the predicted values with the reference values. This is done by plotting the reference value against the predicted value. Ideally, this plot should be an identity line. In practice, the predicted values are fitted to a regression line and model is characterized by using different statistical parameters

- **Coefficient of determination or R^2**: This parameter describes how well the predicted values fit the regression line. An ideal model has R^2 of 1. A model with R^2 of 0.75 means that the fit describes 75% of the variability of the target value being predicted [21, 22].
- **Root Mean Square Error of Prediction or RMSEP**: This parameter measures the average accuracy of the prediction. It is considered that 2 times the RMSEP represents a 95% confidence interval for the real value. For instance, if the fat prediction model has an RMSEP value of 0.1%, and the predicted value of fat content is 4%, then there is a 95% chance that the reference value of fat content for this sample is between 3.8% - 4.2%.
- **Bias**: This parameter represents the average difference between predicted values and reference values. Ideally, the value of the bias is 0. Higher values mean that the model tends to overestimate the composition, and lower values mean that model tend to underestimate the composition of the material.
- **Standard error of prediction or SEP**: This parameter measures the precision of prediction which means the difference between different measurements for the same sample. When Bias tends to 0, the SEP tends to have the same value as the RMSEP.

Modeling results

First, it is important to make sure that range of values cover the ranges to be expected in the field. In the calibration set, fat content ranged from 1.03 to 4.93, protein content ranged from 2.89 to 4.11, and lactose content ranged from 4.18 to 5.31. In the validation set, fat content ranged from 1.03 to 4.52, protein content ranged from 2.89 to 3.96, and lactose content ranged from 4.36 to 5.06. The reference values of both the calibration and validation sets were scattered across the range of the corresponding milk constituent to have a good representation of the population of the study in both sets. The figures below show the histogram of the fat, protein and lactose values of the samples collected in this study.

Regression models based on non-linear neural networks architecture [23] were built for each parameter and were tested using the validation set. Dimensionality reduction was first carried out on the spectral data to build the model based on the independent variables that best represent the dataset.

- For the fat content, the calibration model shown in figure (3) (a) achieved R^2 of 0.98 and RMSE of 0.158. The model was then validated using the validation set where it scored the following: R^2=0.99 and SEP=0.147. this is shown in figure (3) (b).
- For the protein content, the calibration model shown in figure (4) (a) achieved R^2 of 0.89 and RMSE of 0.078. The model was then validated using the validation set where it scored the following: R^2 =0.98 and SEP=0.131, this is shown in figure (4) (b).
- For the lactose content, the calibration model achieved R^2 of 0.82 and RMSE of 0.105. The model was then validated using the validation set where it scored the following: R^2 =0.77 and SEP=0.114, this is shown in figure (5) (b).
- Models of the three milk constituents show high correlation with the chemical references. Results show a high generalization to the validation set assuring the regression model robustness to new samples.

Table 1 summarizes the different results achieved for the three parameters. Biases of the three modeled parameters are all close to zero showing no shifts in predictions from actual values. SEPs of the three modeled parameters are below 0.15.

In order to judge on how those results are, they are compared to other results reported in literature [3, 24] for experiments conducted using miniaturized NIR technologies for in-field use. The results are also compared to ICAR standard for in-field and in-lab use as shown in figure (6). It can be seen that results generated using NeoSpectra are the only ones exceeding the ICAR standards for in-field use for all parameters, and their performance is very close to those of in-lab standards.

Table 1. Fat, protein, and lactose modeling results summary.

	Calibration		Validation			
	R2	RMSE	R2	RMSE	SEP	Bias
Fat	0.984	0.158	0.991	0.167	0.147	0.096
Protein	0.891	0.078	0.979	0.127	0.131	-0.033
Lactose	0.818	0.105	0.767	0.111	0.114	0.033

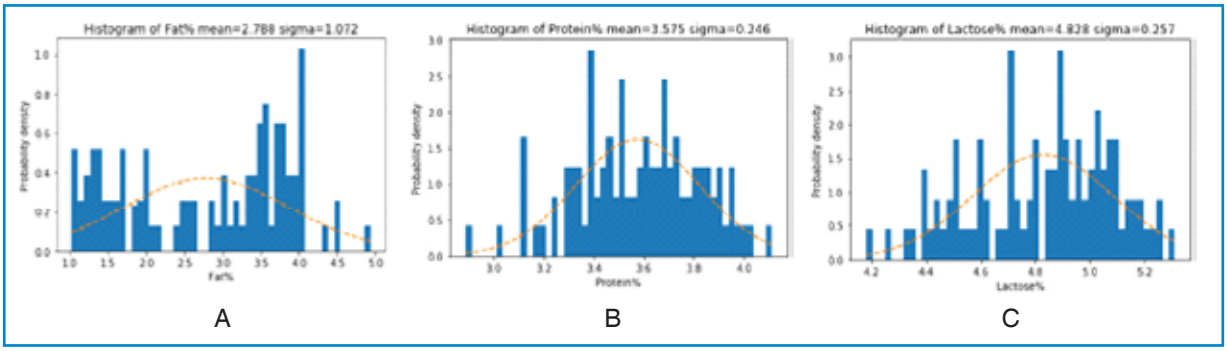
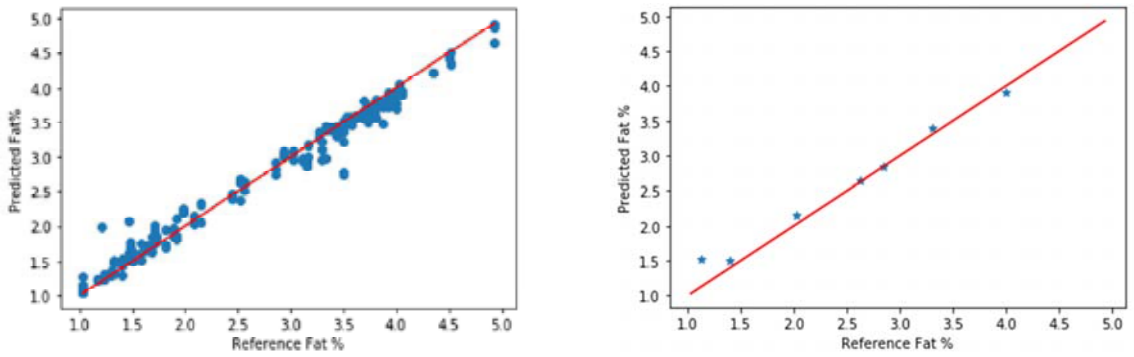
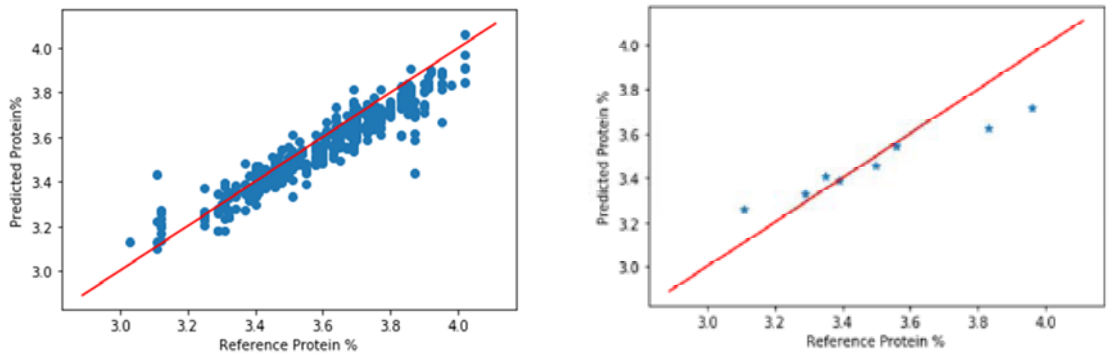


Figure 2. (a) Histogram for fat samples (b) Histogram for protein samples (c) Histogram for lactose samples



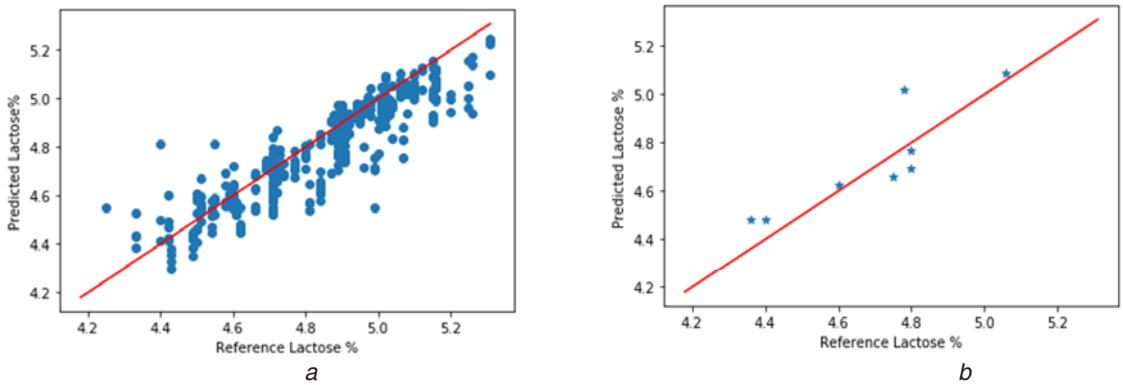
a b

Figure 3. (a) Fat Calibration model with R2 of 0.984 and RMSE of 0.158. (b) Fat Validation results with R^2=0.991, RMSE = 0.167, Bias=0.096 and SEP=0.147



a b

Figure 4. (a) Protein calibration model with R^2 of 0.891 and RMSE of 0.078. (b) Protein validation results with R^2 = 0.979, RMSE = 0.127, Bias = -0.033 and SEP=0.131



a b

Figure 5. (a) Lactose calibration model with R^2 of 0.818 and RMSE of 0.105. (b) Lactose validation results with R^2 =0.767, RMSE=0.111, Bias = 0.033 and SEP=0.114

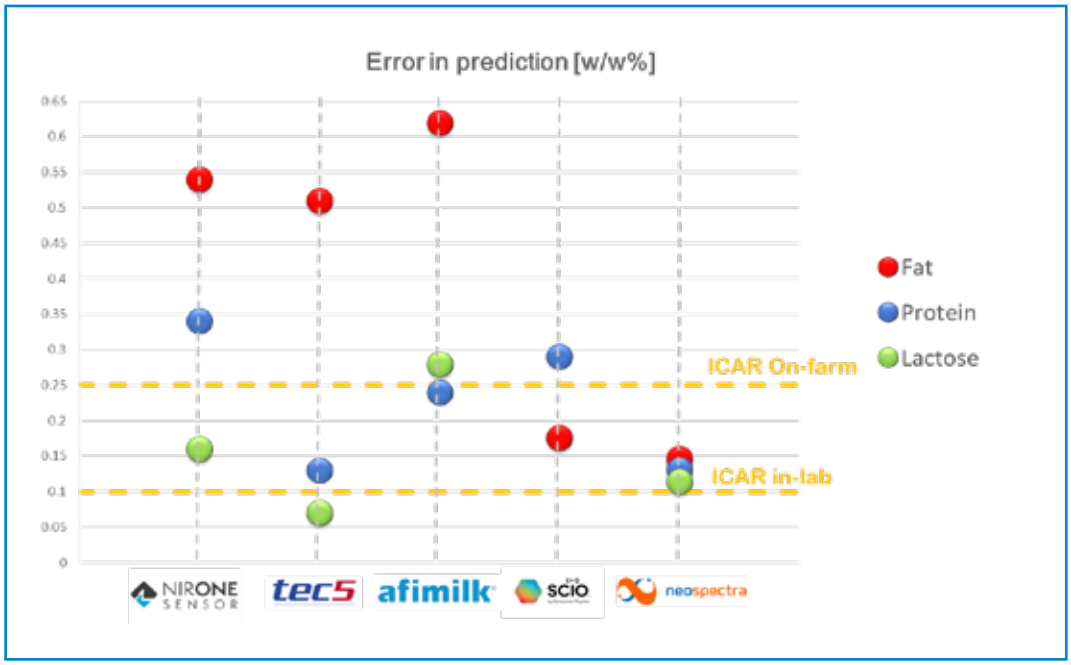


Figure 6. NeoSpectra achieved results compared to different technologies reported in literature and ICAR standards.

Conclusion

The results presented in this work indicates that NeoSpectra NIR spectral sensor has the potential to become the first commercial milk analysis solution that complies with ICAR standard for in-use applications and at the same time democratizes the adoption of milk analysis across the whole milk supply chain. Unlike other conventional and emerging technologies for milk analysis, NeoSpectra is a solution that uniquely combines a set of features that makes it possible to enable cost-effective, quick, ubiquitous, easy to use, and accurate milk analysis. Further experimentations are also conducted to assess the ability of NIR analysis to detect milk adulteration and somatic cell count (SCC). Combining milk accurate in-field milk composition analysis, with adulteration detection and SCC in one solution can disrupt the way we qualify, produce, and trade milk across the whole milk supply chain from cow to cup.

List of references

1. ICAR. 2017a. Guidelines for Testing, Approval and Checking of Milk Recording Devices.

2. ICAR. 2017b. Guidelines for On-Line Milk Analysis.

3. S. Uusitalo, B. Aernouts, J. Sumen, E. Hietala, M. Utriainen, L. Frondelius, S. Kajava, and M. Pastell , Comparison of milk analysis performance between NIR laboratory analyser and miniaturised NIR MEMS sensors, Proceedings ICAR Conference 2019, Prague

4. P. R. Griffiths and J. A. de Haseth, Fourier Transform Infrared Spectrometry, 2nd edition, Hoboken, NJ: John Wiley and Sons, 2007.

5. Yang, H., Yang, S., Kong, J., Dong, A. and Yu, S. Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. Nat. Protoc. 10, 382 (2015).

6. Martin, F. L. *et al.* Distinguishing cell types or populations based on the computational analysis of their infrared spectra. Nat. Protoc. 5, 1748 (2010).

7. de Voort, F. R., Sedman, J., Emo, G. and Ismail, A. A. A rapid FTIR quality control method for fat and moisture determination in butter. Food Res. Int. 25, 193–198 (1992).

8. Shiroma, C. and Rodriguez-Saona, L. Application of NIR and MIR spectroscopy in quality control of potato chips. J. Food Compost. Anal. 22, 596–605 (2009).

9. Das, A. J., Wahi, A., Kothari, I. and Raskar, R. Ultra-portable, wireless smartphone spectrometer for rapid, non-destructive testing of fruit ripeness. Sci. Rep. 6, 32504 (2016).

10. Madejova, J. FTIR techniques in clay mineral studies. Vib. Spectrosc. 31, 1–10 (2003).

11. Erfan, M. *et al.* On-chip micro-electro-mechanical system Fourier transform infrared (MEMS FT-IR) spectrometer-based gas sensing. Appl. Spectrosc. *https://doi.org/10.1177/0003702816638295* (2016).

12. Gastélum-Barrios, G. M. Soto-Zarazúa, A. Escamilla-García, M. Toledano-Ayala, G. Macías-Bobadilla, and D. Jauregui-Vazquez, "Optical Methods Based on Ultraviolet, Visible, and Near-Infrared Spectra to Estimate Fat and Protein in Raw Milk: A Review," Sensors, vol. 20, no. 12, p. 3356, Jun. 2020.

13. Crocombe, R. A. Portable spectroscopy. Appl. Spectrosc. 72, 1701-1751 (2018).

14. B. Saadany, H. Omran, M. Medhat, F. Marty, D. Khalil, T. Bourouina. "MEMS Tunable Michelson Interferometer with Robust Beam Splitting Architecture". Paper presented at: IEEE/LEOS International Conference on Optical MEMS and Nanophotonics 2009. Clearwater, FL; August 17–20 2009.

15. Y. Nada, M. Medhat, M. Nagi, F. Marty, B. Saadany, T. Bourouina. "Mechanical Displacement Multiplier: 250 mm Stable Travel Range MEMS Actuator Using Frictionless Simple Compliant Structures". Paper presented at: IEEE International Conference on MEMS 2012. Paris; January 29–February 2 2012.

16. Yasser M. Sabry, Khaled Hassan, Momen Anwar, Mohamed H. Alharon, Mostafa Medhat, George A. Adib, Rich Dumont, Bassam Saadany, Diaa Khalil, "Ultra-compact MEMS FTIR spectrometer", Proceedings Volume 10210, Next-Generation Spectroscopic Technologies X; 102100H (2017).

17. Pügner, T., Knobbe, J. and Grüger, H. Near-infrared grating spectrometer for mobile phone applications. Appl. Spectrosc. 70, 734–745 (2016).

18. Pedro Gelabert, Eric Pruett, Gavin Perrella, Sreeram Subramanian, and Aravind Lakshminarayanan "DLP NIRscan Nano: an ultra-mobile DLP-based near-infrared Bluetooth spectrometer", Proc. SPIE 9761, Emerging Digital Micromirror Device Based Systems and Applications VIII, 97610B (15 March 2016).

19. B. Mortada, M. Erfan, M. Medhat, Y. M. Sabry, B. Saadany, and D. Khalil, "Wideband Optical MEMS Interferometer Enabled by Multimode Interference Waveguides," J. Light. Technol., 2016.

20. Sabry, Y. M. *et al*. In-plane deeply-etched optical MEMS notch filter with high speed tunability. J. Opt. 17, 125703 (2015).

21. Naes, T., A User-friendly Guide to Multivariate Calibration and Classification, NIR Publications, (2004).

22. Allen B. Downey. Think Stats - Probability and Statistics for Programmers , 2011

23. Hopfield, J. J. (1982). "Neural networks and physical systems with emergent collective computational abilities". Proc. Natl. Acad. Sci. U.S.A. 79 (8): 2554–2558.

24. Riu J, Gorla G, Chakif D, Boqué R, Giussani B. Rapid Analysis of Milk Using Low-Cost Pocket-Size NIR Spectrometers and Multivariate Analysis. Foods. 2020; 9(8):1090. *https://doi.org/10.3390/foods9081090*