engineering proceedings

MDPI

Proceeding

RGB LED sensor for fat quantification in milk

Marina Salukova [1*], Anastasiia Surkova [2], Yana Shmakova [1], Natalya Samokhina [1], Julia Kostyuchenko [2], Alina Parshina [3], Ildar Ibatullin [3] and Andrey Bogomolov [2]

[1] High Biotechnology School, Samara State Technical University, Samara, Russia;
[2] Department of Analytical and Physical Chemistry, Samara State Technical University, Samara, Russia;
[3] Department of Radio Engineering Devices, Samara State Technical University, Samara, Russia
* maria.barash17@gmail.com

Abstract: In this study, a portable desktop analyzer for determination of fat content in milk is introduced. A prototype of the sensor consists of three light emitting diodes (red, green, and blue) as a light source. The transmitted light is detected by a photoresistor and continuous voltage measurements provided by the microcontroller, is recorded by a computer. The resulting univariate and multivariate models show that the developed analytical device is capable to determinate fat content in raw and homogenized milk with sufficient accuracy.

Keywords: visible spectroscopy; optical multisensor systems; RGB LED sensor; milk analysis; fat content; multivariate data analysis

1. Introduction

The quality of milk is determined by the quantitative content of its nutrients. The amount of fat in milk is an indicator of various cow diseases, such as ketosis [1] or acidosis [2]. Therefore, the fat content must be monitored carefully during milk collection and throughout all stages of dairy production.

There are various methods for determining the fat content in milk. Chemical methods are commonly applied, e.g. the Gerber method [3] and the Röse-Gottlieb method [4]. Optical spectroscopy in the mid infrared region [5] has gained popularity among the instrumental methods of milk analysis. However, despite their high analytical accuracy, these methods are complex, labor-intensive, and expensive. The development of new simple and inexpensive analyzers for fast and accurate analysis is an urgent task in the dairy industry.

Visible and short-wave near infrared (Vis/SW-NIR) region is the most promising for the milk analysis [6–9], because it is possible to create easy-to-use and cost-effective devices, functioning in this spectral region and, consisting of light emitting diodes (LED), optical fibers, 3D-printed parts, and photodetectors [10–14]. The possibility of using 400–1100 nm region for the determination of fat and protein content has been shown in our previous work [9].

In the present study, a novel RGB sensor composed of three LEDs, photoresistor, and the Arduino Nano microcontroller was developed to quantify fat in milk [15]. The proposed prototype was applied for analysis of homogenized and raw milk. A linear regression model, projection on latent structures (PLS) and multiple linear regression (MLR), were tested and compared to assess the sensor performance. This simplified technology has shown an acceptable determination accuracy of fat content in milk and can be further used by small dairy farms for milk quality control.

2. Materials and Methods

The developed LED sensor consists of several modules and operates in an automatic

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Lastname

Published: date

mode provided by the Arduino Nano microcontroller. Red (R), green (G), and blue (B) LEDs are used as a light source. A photoresistor registers the signal from the optical sensor. The initial analytical signal is the voltage drop generated at the photoresistor output. The data is then transferred to a computer using the serial port of the microcontroller via a USB cable for the further signal processing.

Two calibration sets of milk samples were prepared to test the capability of the developed analyzer to determine the fat content in milk. The first one was prepared by the mixing two packages of normalized cow's milk with different fat content, bought from a supermarket. The full dataset consists of 10 samples with varying fat content from 1.5 to 3.2% with 0.19% step. The second set consisted of twenty-nine raw milk obtained from the Samara Oblast Agro-Industrial Complex Support Center (Samara, Russia). The fat content in these samples ranged from 3.1 to 6.01%. A tablet of Broad Spectrum Microtabs II preservative was added to the milk samples to prevent spoilage, giving the samples a pink color. Reference analysis of fat content in raw milk was performed on a CombiFossTM7 (FOSS Electric, Denmark).

The sensor prototype works as follows. A milk sample with a volume of 3 mL is placed in a glass cuvette with an optical path of 10 mm. The cuvette is inserted into the slot, covered with a lid from above and the examined sample is alternately illuminated with three LEDs. A USB-cable is used to connect the sensor to a computer and a serial port monitor is displayed. The detected data is digitized, filtered by averaging and tabulated to the port every 100 ms. The power of the light passing through the cuvette with the sample is proportional to the value of the output data. The detected signals reflect alternating pulses of LEDs irradiation at different wavelengths with the emission maxima of 631 nm for the red, 531 nm for the green, and 466 nm for the blue channels.

To test reproducibility of the measurement, a new portion of milk from each sample was taken and measured three times. Each measurement took about 10 seconds, because the developed analyzer produces the readings continuously, and thus, 10 measurements were taken and averaged. Next, three measurements were also averaged. Therefore, one sample is represented by one measurement in the regression models.

To relate the analytical signal response of LED sensor to the fat content, the calibration models were trained using a linear regression for individual sensor channels, partial least-squares (PLS) regression [16] and multiple linear regression (MLR) [17] for combination of three channels. The performance of PLS and MLR models was estimated by leave-one-out cross-validation (CV). Root mean-square error (RMSE) of calibration (RMSEC) and cross-validation (RMSECV) as well as the coefficients of determination (R2) were used to compare the model performances.

3. Results

In accordance with the proposed experimental setup (Figure 1), the light from three LEDs alternately illuminates the milk sample and the transmitted light is registered with a single photometric detector. The voltage from the detector is then read using the microcontroller and transmitted to the computer.

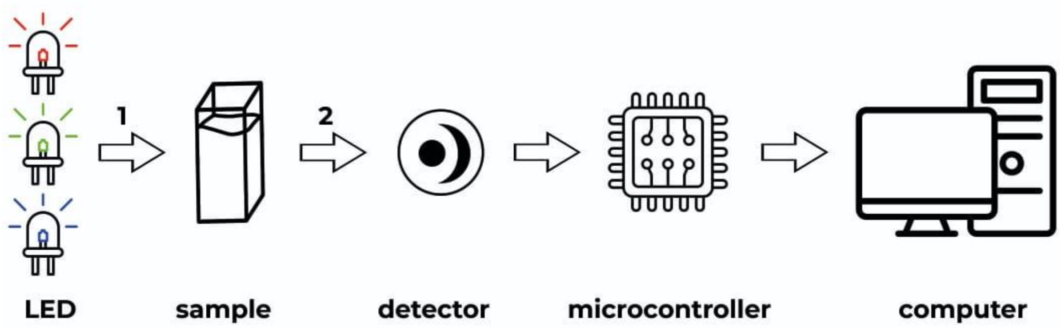

Figure 1. Block diagram of the RGB sensor prototype, where (1) designates the emitted light; (2) – transmitted light.

The responses of the RGB LED sensor for the homogenized milk dataset are shown in Figure 2. It can be seen that the samples in Figure 2 lie presumably in the order of increasing fat content for all three channels, although the distribution is more uniform for the red channel.

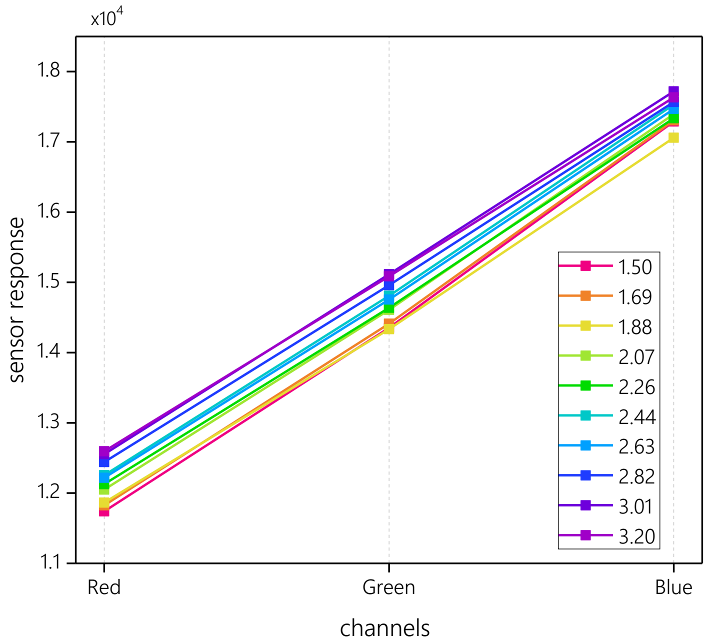

Figure 2. The channel responses of LED sensor for dataset of homogenized milk, colored in rainbow gradient defending on fat content (there is a legend of colors on the right side of the plot).

Firstly, the univariate regression models were built to find the relationship between the response of each channel and fat content. The linear regression equation relating the sensor response and the fat content for the red channel is $y = 522.87x+10938$ with $R^2 = 0.98$, for the green channel is $y = 489.6x+13556$ with $R^2 = 0.94$ and for the blue channel is $y = 286.62x+16756$ with $R^2 = 0.71$. Thus, the red channel is the most sensitive to milk fat content, the green channel performs slightly worse result and the blue channel is the least informative.

The next step of the study was to use multivariate regression methods, such as PLS and MLR, to relate the response for all three channels to fat content. The difference of size distribution of colloidal milk particles, especially fat globules, influences the spectra, as it was shown in our previous work [9]. The combination of different wavelengths provides to the effect and allows taking this effect into account and obtaining adequate models. The combination of the three channels leads to significantly improvement of PLS prediction accuracy compared to univariate models: RMSECV is 0.08% with R2 = 0.98 (Figure 3A). The validation statistics of MLR (RMSECV = 0.10% and R2 = 0.97) is very similar to PLS regression and, thus, the algorithm can also be successfully used to analyze OMS data that confirms our previous statements [18].

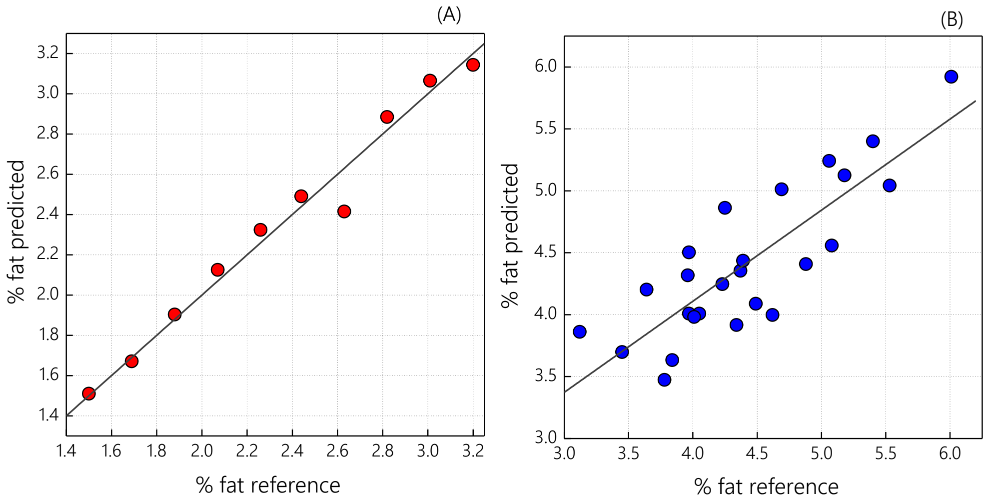

Figure 3. PLS-predicted versus reference plot for fat content in: (A) homogenized milk samples and (B) raw milk samples. Red (A) and blue circles (B) indicate cross-validation results.

A calibration set of raw milk samples was measured using the developed RGB sensor, and four evident outliers were removed from 29 milk samples at the preliminary modelling stage. The PLS statistics (Figure 3B) for three sensor channels is also better (R^2 CV = 0.72, RMSECV= 0.37), as in the case of homogenized milk, than individual sensor channel results. The worse PLS statistics of raw milk could be related to the large variation in fat globule size in natural raw milk, which increases the wavelength dependence of light scattering [9]. The presence of a preservative that gave the samples a pink color presented an additional complexity. Besides, the content of other milk components such as protein and lactose also varied. At this high level of complexity, a much larger number of samples are required to train a calibration model, but building a global model was not the goal of the present study.

4. Conclusions

The portable desktop analyzer for determination of fat content in milk has been developed using cost-effective components: LEDs, a photoresistor, an Arduino Nano microcontroller and a few other easily available components. Homogenized and raw milk sample sets were used to test the performance of the developed sensor within the task of fat content determination. The univariate models relating the response of each channel and fat content showed that the red channel was the most informative. The combination of the three channels gives a significantly better PLS statistics. MLR can also be successfully used to analyze multisensor data, when the number of samples is smaller than the number of variables (wavelengths).

The prototype was deliberately constructed for the most practicable visible light region. The RGB analysis can be based using any digital camera images [19] even smartphone-based. Further improvements of the sensor prototype include optimization of the number and central wavelengths of LED light sources and expansion of the spectral region towards the NIR and mid infrared range.

Author Contributions: Conceptualization, A.B., and I.I.; methodology, I.I., A.S.; formal analysis, A.S., A.P., and J.K.; investigation, A.S., A.P, M.S., Y.S., N.S., and J.K.; data curation, A.S. and A.B.; writing—original draft preparation, M.S., A.S., A.B., and I.I.; visualization, A.S. and M.S.; supervision, A.B., I.I., and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the Ministry of Science and Higher Education of the Russian Federation (theme No. FSSE-2023-0003) as part of the state task of the Samara State Technical University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Samara Region Agricultural Support Center for providing research milk samples.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Heuer, C.; Schukken, Y. H.; Dobbelaar, P. Postpartum body condition score and results from the first test day milk as predictors of disease, fertility, yield, and culling in commercial dairy herds. *J. Dairy Sci.* **1999**, *82*, 295–304. https://doi.org/10.3168/jds.S0022-0302(99)75236-7
2. Plaizier, J. C.; Krause, D. O.; Gozho, G. N.; McBride B. W. Subacute ruminal acidosis in dairy cows: The physiological causes, incidence and consequences. *Vet. J.* **2008**, *176*, 21–31. https://doi.org/10.1016/j.tvjl.2007.12.016
3. ISO 11870:2009(en) Milk and milk products — Determination of fat content — General guidance on the use of butyrometric methods. Available online: https://www.iso.org/obp/ui/#iso:std:iso:11870:ed-2:v1:en (accessed on 15 June 2023).
4. ISO 8381:2000 Milk-based infant foods — Determination of fat content — Gravimetric method. Available online: https://www.iso.org/ru/standard/27796.html?browse=tc (accessed on 15 June 2023).
5. ISO 9622:2013|IDF 141:2013 Milk and liquid milk products — Guidelines for the application of mid-infrared spectrometry. Available online: https://www.iso.org/standard/56874.html (accessed on 15 June 2023).
6. Tsenkova, R.; Atanassova, S.; Ozaki, Y.; Toyoda, K.; Itoh, K. Near-infrared spectroscopy for biomonitoring: influence of somatic cell count on cow's milk composition analysis. *Int. Dairy J.* **2001**, *11*, 779–783. https://doi.org/10.1016/S0958-6946(01)00110-8
7. Šašić, S.; Ozaki, Y. Short-wave near-infrared spectroscopy of biological fluids. 1. Quantitative analysis of fat, protein, and lactose in raw milk by partial least-squares regression and band assignment. *Anal. Chem.* **2001**, *73*, 64–71. https://doi.org/10.1021/ac000469c
8. Kalinin, A.V.; Krasheninnikov, V.N. Portable milk product quality analyzer based on spectrometry in the near IR range. *J. Appl. Spectrosc.* **2008**, *75*, 288–294. https://doi.org/10.1007/s10812-008-9042-5
9. Bogomolov, A.; Melenteva, A. Scatter-based quantitative spectroscopic analysis of milk fat and total protein in the region 400-1100 nm in the presence of fat globule size variability. *Chemom. Intell. Lab. Syst.* **2013**, *126*, 129–139. https://doi.org/10.1016/j.chemolab.2013.02.00
10. Prasanth, P.; Viswan, G.; Bennaceur, K. Development of a low-cost portable spectrophotometer for milk quality analysis. *Mater. Today* **2021**, *46*, 4863-4868. https://doi.org/10.1016/j.matpr.2020.10.327
11. Angrasari, F.; Arifin, A.; Abdullah, B. Fabrication of milk fat sensor based on plastic optical fiber. *J. Phys. Conf. Ser.* **2019**, *1341*, 082038. https://doi.org/10.1088/1742-6596/1341/8/082038
12. Ragni, L.; Iaccheri, E.; Cevoli, C.; Berardinelli, A. Spectral-sensitive pulsed photometry to predict the fat content of commercialized milk. *J. Food Eng.* **2016**, *171*, 95–101. https://doi.org/10.1016/j.jfoodeng.2015.10.017
13. Yang, B.; Guo, W.; Liang, W.; Zhou, Y.; Zhu, X. Design and evaluation of a miniature milk quality detection system based on UV/Vis spectroscopy. *J. Food Compos. Anal.* **2022**, *106*, 104341. https://doi.org/10.1016/j.jfca.2021.104341
14. Yang, B.; Huang, X.; Yan, X.; Zhu, X.; Guo, W. A cost-effective on-site milk analyzer based on multispectral sensor. Comput. Electron. Agric. 2020, 179, 105823. https://doi.org/10.1016/j.compag.2020.105823
15. Surkova, A.; Shmakova, Y.; Salukova,M.; Samokhina, N.; Kostyuchenko, J.; Parshina, A.; Ibatullin, I.; Artyushenko, V.; Bogomolov, A. LED-Based Desktop Analyzer for Fat Content Determination in Milk. *Sensors* **2023**, *23*, 6861. https://doi.org/10.3390/s23156861
16. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. Chemom. Intell. Lab. Syst. **2001**, *58*, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1
17. Esbensen, K.H. Multivariate Data Analysis—In Practice, 5th ed.; CAMO Process AS: Oslo, Norway, 2001.
18. Surkova, A.; Bogomolov, A.; Legin, A.; Kirsanov, D. Calibration transfer for LED-based optical multisensor systems. *ACS Sens.* **2020**, *5*, 2587–2595. https://dx.doi.org/10.1021/acssensors.0c01018
19. Kucheryavskiy, S.; Melenteva, A.; Bogomolov, A. Determination of fat and total protein content in milk using conventional digital imaging. *Talanta* **2014**, *121*, 144–152. http://dx.doi.org/10.1016/j.talanta.2013.12.055