Optical Multisensor System Based on Lanthanide(III) Complexes as Near-Infrared Light Sources for Analysis of Milk

Anastasiia Surkova 1,2, Andrey Bogomolov 2,*, Aleksandra Paderina 1, Viktoria Khistiaeva 1, O, Ekaterina Boichenko 3, Elena Grachova 1 and Dmitry Kirsanov 1,3

1 Institute of Chemistry, St Petersburg University, Universitetskii pr. 26, 198504 St. Petersburg, Russia; melenteva-anastasija@rambler.ru (A.S.); ksaniasha@list.ru (A.P.); v.v.khist@gmail.com (V.K.); e.grachova@spbu.ru (E.G.); d.kirsanov@gmail.com (D.K.)
2 Department of Analytical and Physical Chemistry, Samara State Technical University, Molodogvardeyskaya Street 244, 443100 Samara, Russia
3 Center of Chemical Engineering, ITMO University, Kronverksky pr. 49, 197101 St. Petersburg, Russia; ekaterina.boichenko@inbox.ru
* Correspondence: a.bogomolov@mail.ru

Abstract: Optical multisensor systems are easy-to-use and inexpensive analytical devices. In this work, we propose an optical multisensor system based on the luminescence of Nd(III) and Yb(III) complexes in the near-infrared (NIR) spectral region. The observed emission bands play the role of secondary light sources for further analysis of milk—for the determination of fat content and for the recognition of adulteration. The samples for analysis were prepared by putting a drop of milk upon a thin glass covering the powdered mixture of lanthanide complexes, which were excited by a light-emitting diode (LED) in the ultraviolet region (the maximum intensity at 365 nm). The diffuse-reflectance spectra of samples were acquired in the short-wave NIR range 750–1100 nm using a portable NIR spectrometer. The developed optical system was tested using two sets of milk samples with varying concentration levels of fat and added urea. The obtained spectral data were analyzed using a number of multivariate prediction and classification methods of chemometrics and the results were statistically compared. The regression and classification model performances achieved in this proof-of-concept study illustrate the feasibility of the optical multisensor analysis based on luminescent light sources in the short-wave NIR range, in particular, for their application in the dairy.

Keywords: NIR spectroscopy; optical multisensor system; lanthanide(III) complexes; milk analysis; urea; chemometrics

1. Introduction

The development of simple, portable, and inexpensive analytical devices based on the principles of optical spectroscopy is an emerging trend in analytical chemistry. Special attention is drawn to optical multisensor systems (OMS) [1]. An OMS is a spectral analytical device composed of several optical sensors with pronounced cross-sensitivity, which implies the use of multivariate data analysis (chemometrics) for spectral data processing. One of the main advantages of OMS is a possibility to optimize the number and composition of optical sensors for a particular analytical task, which is especially important for practical industrial applications.

The nature of optical sensors or channels forming an OMS can be different. Light-emitting diodes (LEDs) are often used as light sources in OMS design [2]. The main advantages of LEDs are the high brightness, long lifetime, broad choice of wavelengths and bandwidths, and low power consumption [3]. LEDs in the visible and short-wave near-infrared (NIR) regions have been widely suggested for field analysis in different areas: e.g., in ecological monitoring of water [4,5], determination and quantification of sheep
cheese whey in water [6], detection of apple freshness and quality [7], and estimation of sugar content in citrus [8].

In our previous studies, we proposed a new type of light source for OMS development: luminescent complexes of various metals or molecular emitters. We showed that cyclometalated Ir(III) complexes [9] and Cu(I)-based complexes [10] could be successfully applied as a multi-band light source in the spectral range of 450–800 nm. Each luminescent complex yields a specific emission spectrum in the visible region; also, the brightness of emission bands is controllable. Moreover, easy-to-synthesize Cu(I)-based complexes are a cheaper alternative to cyclometalated Ir(III) complexes. These advantages of molecular emitters as light sources give more options for OMS design and optimization in terms of required analytical characteristics and cost-effectiveness. Two OMS prototypes were developed and tested on a set of model aqueous mixtures containing three metal ions: Co(II), Ni(II), and Cu(II) [9]. Then, a real-practical application of the device was demonstrated for the determination of fluoride and phosphate ions in surface and tap waters [10].

The purpose of this study was to extend the suggested concept of molecular emitter-based OMS beyond the visible spectral range and to involve the NIR region, which could be beneficial for numerous practical analytical tasks. For this, we developed an OMS prototype with UV-excited lanthanide(III) complexes as light sources. The advantages of lanthanide(III) complexes are the intra-configuration \( f-f \) transitions, which lead to a sharp luminescence profile that is individual and independent of the coordination sphere composition [11–13]. Long luminescence lifetime and large Stokes shifts are also practically useful features of Ln(III) complexes photonics. The list of unique properties of some lanthanide(III) ions, namely ytterbium(III) and neodymium(III), also includes photoluminescence in the near-infrared (NIR) region (980 nm for Yb(III) and 1060 nm for Nd(III), respectively). Other advantages of lanthanides are given in the Supplementary Material. In our study, the luminescence of the complexes was excited by a LED-based flashlight; the emission spectra were acquired by a portable NIR spectrometer. The developed OMS prototype was used for the quantification of fat and urea in milk in two sets of real milk samples. Several methods of multivariate calibration and classification were also applied to process the obtained spectral data. The performance of the resulting regression and classification models was evaluated and compared.

2. Experimental Section

2.1. Complexes

Lanthanide(III) complexes \([\text{Nd(tta)}_3(\text{dppn})]\) (1) and \([\text{Yb(tta)}_3(\text{dppn})]\) (2); Htta = thenoyltrifluoroacetone; dppn = 3,6-di(2-pyridyl)pyridazine; were synthetized. The synthesis and complete characterization of the Ln(III) complexes have been reported in [14]. The structures of the complexes are shown in Figure 1. All reagents and solvents were purchased from Merck, Alfa Aesar, Fluka, and Vekton and were used without any additional purification. Bidistilled water was used for the sample preparation.

The normalized photoemission spectra of Ln(III) complexes in the solid state are shown in Figure 2. The UV-Vis spectra of 1 and 2 complexes are shown in Figure S1 (in Supplementary Material).

2.2. Samples

Normalized cow milk with different fat content was purchased from a local supermarket. To demonstrate a practical application of the developed OMS, four calibration datasets were prepared.

To reveal the effect of different fat content on the spectra, milk samples with 2.5% and 3.3% fat content were mixed in different proportions. A full series consists of five milk samples with the following percentage of fat content: 2.5, 2.7, 2.9, 3.1, and 3.3%. Additionally, a calibration set of 11 milk samples was prepared by mixing milk from two milk packages with 1.5% and 3.2% fat content (in 0.17% concentration increments). These two calibration sets are further denoted as “F1-series” and “F2-series”.
10 mL of milk (with a fat content of 2.5%). The concentrations of urea in calibration mixtures were 1700, 1900, and 2100 mg/L. Since the urea levels in milk are regulated, the concentration range covered both the recommended levels of urea in milk and the levels, which exceed the permissible range.

To prove a concept of urea quantification in milk by the optical multisensor device, a set of 12 calibration samples, further denoted as “U1-series”, was prepared by adding 100–2100 mg/L (with a 200 mg/L step) of urea in milk (with a fat content of 2.5%). Thus, the concentrations of urea in milk samples were 0, 100, 300, 500, 700, 900, 1100, 1300, 1500, 1700, 1900, and 2100 mg/L. Since the urea levels in milk are regulated, the concentration range covered both the recommended levels of urea in milk and the levels, which exceed the permissible range.

To classify adulterated and non-adulterated milk samples, a set of ten calibration samples (denoted as “U2-series”) was prepared by adding from 0 to 0.09 g of urea into 10 mL of milk (with a fat content of 2.5%). The concentrations of urea in calibration mixtures were 0, 100, 300, 500, 700, 1000, 3000, 5000, 7000, and 9000 mg/L. The samples with urea levels below or equal to 700 mg/L [15] were labeled as non-adulterated; the samples with higher urea levels were considered as adulterated, correspondingly. Thus, the number of non-adulterated and adulterated samples was equal.

A 40 µL drop of each calibration sample was placed on a coverslip and dried to constant mass on the laboratory heater. In the case of calibration samples used for training of classification models this procedure was repeated twice to ensure the reproducibility of the drop. Thus, two series of calibration samples were obtained. This simplified and cost-effective sample preparation technique assumes the significant variability of reflective properties of the dried milk drop, caused by the difference in layer thickness, morphological and chemical heterogeneity of the surface, and the presence of air bubbles. Therefore, the presence of several outliers in the measured spectral datasets was anticipated.

Figure 1. Schematic representation of the lanthanide(III) complexes.

Figure 2. Solid-state emission of complexes 1 and 2, λ_{exct} = 365 nm, room temperature.
2.3. OMS Construction

To obtain a light source in the wavelength range of 750–1100 nm, 10 mg of the unit of 1 and 5 mg of the unit of 2 complexes in the solid state were ground in a mortar to a homogeneous state and mixed together. The developed OMS (Figure 3) consisted of a mixture of the lanthanide(III) complexes, used as a light source in the short-wave NIR region, and STS-NIR Miniature Spectrometer (Ocean Optics, Largo, FL, USA) as a detector. To excite the emission of the complexes, a UV flashlight with an emission maximum of 365 nm (Alonefire SV003, Hong Kong, China) was used, which was the closest standard light source to the absorption maximum in UV-Vis spectra of the Ln(III) complexes [14]. The actual emission parameters could be negligibly different from the technical specification provided by a vendor. Therefore, the flashlight used in this study had a peak maximum at 367 nm and a full width at half maximum of 14 nm (in the interval 362–376 nm) [10]. The emission peak had a Gaussian-like shape with some “tail” towards longer wavelengths. Since the rechargeable lithium battery of the flashlight was unstable and discharged over time, the UV flashlight was connected to a voltage regulator. This optical setup is further denoted as “OMS prototype-1”.

The homogeneous mixture of 1 and 2 powders was placed between two coverslips and fixed on a solid substrate in a way that the detector could receive the light passed through a certain area of the coverslip. A coverslip with the dried drop of a milk sample was placed above the coverslip with the Ln(III) complexes. The UV flashlight excited the emission of the lanthanide complexes. The emitted light passed through the sample and went to the spectrometer. Spectra were acquired in the range of 750–1100 nm with a spectral resolution of 0.46 nm and integration time of 3 s using Ocean Optics software. The raw intensities at the specified wavelength range or at the individually selected wavelengths (Table 1) were used for the analysis. The signal from the Nd(III) complex in the 1050–1100 nm range could be distorted because of the NIR spectrometer limit of wavelengths. To ensure reproducibility, each sample was measured three times at three different points of a milk drop.

Figure 3. Cont.
An alternative optical setup was created to test the possibility of measurements in transmission units. This setup consisted of another UV flashlight with an emission maximum of 375 nm (UV-5 Detector, Beijing, China) and a diode-array spectrometer TIDAS E by J&M Analytik AG (Essingen, Germany). The UV flashlight was placed under the coverslip containing powder of complexes. The coverslip with a milk sample was located between the luminescent complexes and the detector (Figure S2 in Supplementary Material). This optical setup is further denoted as “OMS prototype-2”. The spectra were acquired in absorbance units (Equation (1)) using a NIR fiber probe (art photonics GmbH, Berlin, Germany):

$$A = -\log \frac{l}{l_0}$$  \hspace{1cm} (1)

where \(l_0\) and \(l\) are the raw intensities of incident light and light after its interaction with the sample, respectively.

The spectrum of a standard mixture of complexes was used as a reference measurement. The spectra were acquired in the range of 750–1100 nm with a spectral resolution of 1 nm and integration time of 1 s. The analysis was carried out at room temperature (24 ± 1 °C). The F2-series (Section 2.2) of milk samples was used to evaluate the performance of the OMS prototype-2 by training a calibration model based on the registered spectral intensities. Each sample was measured five times and the spectra were averaged be-

---

**Table 1. PLS and MLR regression results for prediction of fat content and urea concentrations.**

| Dataset       | Wavelength Range (nm) | LV | Calibration | CV |
|---------------|------------------------|----|-------------|----|
|               |                        |    | RMSE        | R² | RMSE | R²  |
| F1-series, %  | 850–1100               | 2  | 0.04        | 0.98 | 0.09 | 0.94 |
|               | 877, 975, 999, 1021    | 2  | 0.06        | 0.96 | 0.13 | 0.87 |
| U1-series, mg/L | 850–1100              | 1  | 212.21      | 0.91 | 307.56 | 0.84 |
|               | 877, 975, 999, 1021    | 1  | 293.01      | 0.81 | 358.53 | 0.79 |
| F1-series, %  | 975, 999               | -  | 0.05        | 0.96 | 0.10 | 0.88 |
| U1-series, mg/L | 877, 975              | -  | 275.65      | 0.84 | 342.65 | 0.75 |

\(a\) in 2.5–3.3% concentration range, \(b\) in 0–2100 mg/L concentration range, \(c\) three outliers.
fore the modeling. The standard deviation for each sample in five replicated measurements did not exceed 0.004 (absorbance units).

2.4. Data Analysis

In this work, the assessment of fat and urea content in milk and detection of “adulterated” milk was addressed. A multivariate regression model is a common way to predict the quantity of a target analyte in a multicomponent medium. We used one of the well-known regression algorithms—projection on latent structures (PLS), widely applied in optical spectroscopy [16]. This algorithm is trained on a dataset of $X$-variables (spectral intensities) and a corresponding $y$-variable ($a$ priori known concentrations of fat or urea) optimizing the regression coefficients in a linear equation, which relates $X$ and $y$. The optimized coefficients can be used for predicting the value of a $y$-variable from a previously unseen set of $X$-variables. Leave-one-out cross-validation (CV) on averaged datasets was applied to estimate the PLS model performance.

When the number of $X$-variables is small (e.g., several measuring channels in our case), it could be beneficial to use multiple linear regression (MLR) because of its simplicity and low computational time [17]. In MLR, the vector of regression coefficients $b$ is calculated directly from $X$- and $y$-variables (Equation (2)), skipping the projection step:

$$b = \left( X^T X \right)^{-1} X^T y$$

(2)

The model performance is assessed by two standard metrics—the correlation coefficient $R^2$ and by the root mean-square error (Equation (3)) of calibration (RMSEC) or cross-validation (RMSECV):

$$RMSE = \sqrt{\frac{1}{k} \sum_{i=1}^{k} (\hat{y}_i - y_i)^2}$$

(3)

where $y_i$ and $\hat{y}_i$ are known and predicted values, and $k$—is the number of samples in the validation set (for CV $k$ is the number of samples).

RMSECV shows the error of the model predictions in the range of concentrations being studied, while $R^2$ determines the linear correlation between the points in the predicted versus measured plot.

For the third task—classification of “adulterated” and “non-adulterated” samples of milk—partial least-squares discriminant analysis (PLS-DA) was used [18]. It is a popular method of classification based on the decomposition of a matrix with $X$-variables onto latent variables (LV), keeping the covariance between the $X$-variables and their binary category variable. In this study, we used the urea levels as labels dividing all milk samples into two categories: “normal” (urea level in the range of 0–700 mg/L) and “adulterated” milk (urea level in the range of 1000–9000 mg/L). Three measurements per sample for the classification model were not averaged and segmented CV with the segments formed by the sample was applied. The validation segment was represented by six measurements: one milk sample on two coverslips, where each coverslip was measured three times.

We used both raw and preprocessed spectral datasets as input for chemometric models. The spectral data were preprocessed by standard normal variate correction (SNV). The calculations were carried out in a web-based chemometrics software TPT-cloud www.tptcloud.com (accessed on 1 November 2021) by Mestrelab Research (Santiago de Compostela, Spain) and Global Modelling (Aalen, Germany) and the Unscrambler 11.0 (CAMO, Oslo, Norway) software package.

3. Results and Discussion

The Ln(III) complexes demonstrate metal-centered photoluminescence under UV excitation with typical sharp peaks in the NIR range, which correspond to $f$-$f$ transitions
of Nd(III) and Yb(III) ions [11,12,19] (Figure 2). A smoothed spectrum of the complexes’ mixture is shown in Figure 4A. The highest intensity of the spectrum is observed at the 850–1050 nm region with two emission maxima at 877 and 975 nm.

Figure 4. (A) The smoothed emission (Savitzky-Golay filter with first-order polynomial, the window width of 15 points) spectrum of 1 and 2 complexes’ mixture excited by the UV flashlight; (B) attenuated emission spectra of the milk samples with varying fat content; (C) attenuated emission spectra of the milk samples with varying urea content.

In the previous works [9,10], the optical multisensor systems were composed of Ir(III) [9] or Cu(I) [10] complexes emitting light in the visible spectral range. The spectra in this optical setup were measured in absolute intensity units due to the lack of a good reference sample. It was shown that the normalized emission intensity is appropriate for the quantification of metal ions in model aqueous solutions [9] and for the determination of fluoride and phosphate content in real surface and tap waters [10].

Absorbance units are widely used in the infrared spectroscopy for quantitative analysis, because of their expected linear dependence from the concentration of absorbing components. Therefore, to choose the best design for an OMS, preliminary measurements
were made both in raw intensity and absorbance units. A mixture of 1 and 2 complexes themselves was proposed to be used as a standard sample (reference measurement).

The luminescence of 1 and 2 complexes was relatively low. Therefore, the measurements in liquid milk samples could be complicated. Thus, it was suggested to dry a drop of the milk sample, and placed on a thin cover glass. In this case, we detected a stable spectral signal of light passed through the sample layer from the light source.

In diffuse reflectance or transmittance spectroscopy, the full reproducibility of measurements is not generally expected, because of the so-called scatter effect that may change the quantity of detected portion of light unpredictably. Two strategies are generally applied to overcome this issue. One of them is based on using multiple measurements of the same sample at different points, after mixing, etc. The replicates can be then averaged or used together to train the model in the presence of an additional variability. Another approach applies mathematical normalization of spectra (i.e., scatter correction methods), so that the data analysis is used on spectral shape differences, rather than absolute intensities. In our work, we have found that the first approach with averaging multiple measurements shows the best results.

Further, the developed OMS prototypes were applied for quantification of fat (OMS prototype-1 and OMS prototype-2) and urea (OMS prototype-1) in milk and for classification of non-adulterated and adulterated milk samples (OMS prototype-1).

3.1. Fat Quantification in Milk

Fat is one of the main parameters that define milk quality, and its quantitative determination is of great importance in the dairy industry [20]. The highly accurate chemical methods of fat quantification (for example, by Gerber (IDF 105–2008) and by Röse-Gottlieb (IDF Standard 1D, 1996)) are labor- and time-consuming. Flexible and cost-effective NIR spectroscopy has been widely studied for milk analysis in recent decades [21–24] as an alternative to the traditional chemical methods. Bogomolov et al. [25] reported that the spectral region of 400–1100 nm was also suitable for the quantification of fat and protein in raw and homogenized milk. The analysis in this region is based on the light scattering by colloidal particles of the milk components.

We assessed the performance of the developed OMS prototype-1 in the quantification of fat using the set of milk samples with varying fat content (F1-series). The measurement of fat in milk in the NIR region is mainly based on light absorption by milk components. Spectra of three milk samples with different fat content are shown in Figure 4B. The spectra are arranged in descending order of concentration in the 850–1050 nm region. A PLS regression model for fat content determination was trained on the spectra in the 850–1100 nm region, where the highest intensity of the Ln(III) complexes was observed. Despite the technical simplicity of the OMS, the resulting PLS model achieved relatively high prediction accuracy: RMSECV was 0.09% and $R^2 = 0.94$ (Figure 5A, Table 1). Since the light sources were almost discrete, and only a few peaks were observed (Figure 3A), one can build a regression model using a number of selected wavelengths instead of the whole spectra. The PLS model was based on four wavelengths: 877, 975, 999, and 1021 nm, where the complexes had the maximum emission intensity. The resulting PLS model had an accuracy that was comparable to the accuracy of the full-spectrum model (RMSECV is 0.13% with $R^2 = 0.87$). The use of an MLR algorithm for two variables (975 and 999 nm) demonstrated a similar RMSECV (0.10%), and $R^2 = 0.88$. Since the loss of accuracy was insignificant, an MLR model can be also effectively used for fat quantification.

To check the possibility of analysis in transmission mode, the calibration set of the F2-series was measured in absorbance units using OMS prototype-2. Ideally, emission signals in a raw sample and reference spectra ($I$ and $I_0$ in Equation (1)) should be equal, so that the resulting absorbance spectra show only the difference related to the absorption due to the presence of the analyte. However, the absorbance NIR spectra in this dataset, shown in Figure S3 (in Supplementary Material), contain the emission peaks of the luminescent complexes. Presumably, the detected luminescence cannot be fully compensated by the
reference measurement because of the presence of multiple physical effects affecting the light distribution within the samples. These effects may include a scattering of both exciting and emitted light, the fluorescence of the milk components, and alteration of the sample morphology due to the presence of added components. As a result, the complex emission structure remains visible in absorbance spectra. Considering that expected fat absorption is weak, it can hardly be visually observed in the background of larger peaks of under-compensated emission. Although the difference in spectral intensities of the samples having the minimal and maximal fat content is noticeable (up to 0.035 absorbance units), the intermediate spectra were close to each other.

![Figure 5](image-url) Predicted versus measured plots for quantification of fat (A) and urea (B) for the full-spectrum PLS regression model (850–1100 nm). Blue circles (A) and red squares (B) indicate cross-validation results.

The PLS model prediction accuracies were slightly lower than those for the PLS model based on the raw intensities \( (LV = 1) \): RMSECV is 0.20% with \( R^2 = 0.87 \) (Figure S4 in Supplementary Material) and RMSECV is 0.23% with \( R^2 = 0.83 \) (Figure S5 in Supplementary Material) for the data with averaged and non-averaged repeated measurements, respectively. Worse prediction accuracy for the data obtained in transmission mode compared to the raw-intensity models can be explained by several reasons. Firstly, the lanthanide complexes are weak light sources, which makes the logarithmic value of \( (I/I_0) \) in Equation (1) prone to a significantly strong noise that starts to dominate in spectra. Secondly, weak absorbance signals in the spectra are overlapped with other more intensive phenomena observed in the UV-irradiated milk drop, including residual emission and noise. Thirdly, the linearity of the absorbance dependence on concentration is not generally expected in a strongly scattering medium of studied samples for both transmission or diffuse reflection measurement modes. A combination of the above-mentioned factors makes the data presented in absorbance units less informative (from a useful signal-to-noise ratio point of view) than the raw intensity spectra. Therefore, the raw intensities were further preferred in this study.

3.2. Detection of Urea-Adulterated Milk

Another practical application, addressed in this study, was related to the determination of the urea content in milk. The Kjeldahl method (IDF 20B:1993) is commonly used for detecting the protein content in milk. This method estimates total nitrogen content, but does not determine its provenance from protein or non-protein sources [26]. The nitrogen-rich urea is known to be an illegal additive used to falsely increase the amount of determined protein in diluted milk. Urea concentrations in natural milk are in the range of 180–400 mg/L with a regulated upper limit of 700 mg/L [15]. Therefore, a higher amount of urea in milk may indicate adulteration. The standard laboratory methods of urea detection, such as Raman spectroscopy, ion chromatography, or liquid chromatography-mass
spectrometry are expensive and require highly trained personnel. The search for new cost-effective alternatives is of great interest: electrochemical [27,28] biosensors, voltammetric electronic tongue [29], dielectric spectroscopy [30], and NIR hyperspectral imaging [31] have been proposed to detect milk adulteration. In [32], a portable optical device in the Vis/NIR/fluorescence region was designed for the detection of fraud in skim milk powder.

Urea has a weak absorption band in the visible and short-wave NIR region (Figure 6; ten spectra were measured on a solid sample of dry urea powder (PanReac, Spain) and averaged over all repetitions), where the second overtone of –NH2 symmetric stretching is observed [33]. Thus, the developed OMS can be potentially applied for the quantification of urea in milk.

![NIR spectrum of dry urea powder](image)

*Figure 6. NIR spectrum of dry urea powder (averaged over 10 repetitions; Spectralon™ was used as a reference standard).*

To prove a concept of urea quantification in milk by the designed OMS prototype-1, 12 samples with different amounts of urea (U1-series, from 0 to 2100 mg/L) were analyzed. The three most characteristic spectra of the milk samples without urea, with 1000 mg/L, and 9000 mg/L (U2-series) are presented in Figure 4C. Despite the dominating emission signals in the spectra, the effect related to the presence of urea can be clearly observed. Several multivariate modeling algorithms were used to assess the feasibility of quantitative or semi-quantitative (classification-based) determination of urea in the milk samples by OMS.

A PLS regression method was used for urea quantification, as discussed in Section 2.4. The resulting PLS statistics are presented in Table 1. All PLS regression models for urea quantification were simple, i.e., required only one latent variable (LV). The first regression model was trained using the whole spectra (850–1100 nm), where the main signals from the emitters were observed. Another PLS model was built using only four (877, 975, 999, and 1021 nm) variables, corresponding to the emission maxima. The best performance was achieved for the full-spectrum PLS model: RMSECV = 307.56 mg/L with $R^2 = 0.84$ (Figure 5B). Nevertheless, the PLS models, trained on four wavelengths, demonstrated comparable accuracy (Table 1). Further, the same variables (877 and 975 nm) were selected for building the classical MLR calibration. The performance of MLR for two variables was superior to the PLS results for four variables (RMSECV = 342.65 mg/L instead of 358.53 mg/L). There are a number of papers on the spectroscopic determination of urea in milk [31,32,34,35], however, a direct comparison of the prediction accuracies is difficult because of the different concentration ranges and validation methods employed for calibration in different publications.
The general sensitivity of the spectra to the presence of urea in milk was relatively low, and the performance of the regression models was moderate. Nevertheless, the OMS approach can be applied for semi-quantitative detection of urea, i.e., classification of milk samples with “acceptable” and “excessive” urea content. To build PLS-DA classification models, the calibration dataset of the U2-series was split into two categories (classes)—“normal” (urea concentration below or equal to 700 mg/L) and “adulterated” (urea above 700 mg/L), which corresponds to the categorical variable values of 0 and 1, respectively.

PLS-DA model with the decision boundary parameter of 0.5 demonstrated appropriate analytical performance in the recognition of “adulterated” samples, as follows from both calibration and cross-validation statistics in Table 2.

| Table 2. PLS-DA results a. |
|---------------------------|
| Calibration | CV |
| Spec, % | Sens, % | Ac, % | Spec, % | Sens, % | Ac, % |
| 100 | 100 | 100 | 88 | 90 | 89 |

a 850–1100 nm range, preprocessing—SNV.

4. Conclusions

Experimental results, obtained in this pilot study, demonstrate the potential of luminescent signals of rare-earth element complexes in the NIR region within practical analytical tasks. We performed the quantitative determination of fat and urea in milk, and the classification of normal and adulterated milk samples based on the area level. Although a complete separation of two classes of urea content was not achieved, we verified the rough estimation of milk adulteration by urea addition with the developed OMS prototype. Also, we studied the effect of spectral data reduction on the performance of classification and regression multivariate models. The insignificant loss of accuracy (0.13% instead of 0.09% for fat and 359 mg/L instead of 308 mg/L for urea) provides an option for further simplification of OMS design. The presented concept is promising for the development of optical multisensor systems applicable for the assessment of food quality or other analytical challenges in the food industry. Although the use of the luminescent complexes as light sources was demonstrated for Yb(III) and Nd(III) complexes, this approach can be potentially extended towards other similar complexes. Assessment of different complexes, technical setups, and analytical applications is a subject of our further research.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/chemosensors10070288/s1, Figure S1: Solid-state emission of complexes 1 and 2 in UV-Vis region, $\lambda_{\text{exct}} = 365$ nm, room temperature; Figure S2: The photo of the OMS prototype-2; Figure S3: The raw NIR spectra of the milk samples (F2-series) with varying fat content in absorbance mode (red gradient indicates an increase in the concentration of fat); Figure S4: Predicted versus measured plot for quantification of fat (F2-series) for the full-spectrum PLS regression model (850–1100 nm) in absorbance mode. Blue circles indicate full cross-validation results; Figure S5: Predicted versus measured plot for quantification of fat (F2-series) for the full-spectrum PLS regression model (850–1100 nm) in absorbance mode. Blue circles represent an average of five repeats and the error bars represent the standard deviation.

Author Contributions: A.S.: Methodology, Investigation, Formal analysis, Writing—original draft. A.B.: Conceptualization, Writing—review & editing. A.P.: Investigation. V.K.: Investigation. E.B.: Investigation, Data analysis, Writing—review & editing. E.G.: Supervision, Writing—review & editing. D.K.: Supervision, Conceptualization, Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: A.S. acknowledges the financial support from the President’s grant #MK-2192.2021.4 (construction and optimization of optical multisensor systems, data analysis). A.B. acknowledges the financial support from the Ministry of Education and Science of the Russian Federation within the framework of state task No. 0778-2020-0005 (conceptualization, experimental design). Mestrelab Research is acknowledged for providing TPT-cloud software supported by the project “Laboratorio
Dixital do Futuro (FutureLab)” (IN854A 2019/16) funded by the Axencia Galega de Innovación (GAIN), through the Programa Industrias do Futuro 4.0 (terceira convocatoria) and co-financed by the European Regional Development Fund (ERDF) under the Feder Galicia 2014-2020 operational program. It also has the support of the Consellería de Economía, Emprego e Industria.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** There are no conflict to declare.

**References**

1. Bogomolov, A. Developing multisensory approach to the optical spectral analysis. *Sensors* **2021**, *21*, 3541. [CrossRef] [PubMed]
2. Yeh, P.; Yeh, N.; Lee, C.-H.; Ding, T.-J. Applications of LEDs in optical sensors and chemical sensing device for detection of biochemicals, heavy metals, and environmental nutrients. *Renew. Sustain. Energy. Rev.* **2017**, *75*, 461–468. [CrossRef]
3. Kovac, J.; Petermai, L.; Lengyel, O. Advanced light emitting diodes structures for optoelectronic applications. *Thin Solid Films* **2003**, *433*, 22–26. [CrossRef]
4. Xiong, Y.; Wang, C.; Tao, T.; Duan, M.; Tan, J.; Wub, J.; Wang, D. Fabrication of a miniaturized capillary waveguide integrated fiber-optic sensor for fluoride determination. *Analyst* **2016**, *10*, 3041–3049. [CrossRef]
5. Sarwar, M.; Leichner, J.; Naja, G.M.; Li, C.-Z. Smart-phone, paper-based fluorescent sensor for ultra-low inorganic phosphate detection in environmental samples. *Microsyst. Nanoeng.* **2019**, *5*, 56. [CrossRef]
6. Lastra-Mejias, M.; Villa-Martinez, A.; Izquierdo, M.; Aroca-Santos, R.; Cancilla, J.C.; Torrecilla, J.S. Combination of LEDs and cognitive modeling to quantify sheep cheese whey in watercourses. *Talanta* **2019**, *203*, 290–296. [CrossRef]
7. Gao, F.; Dong, Y.; Xiao, W.; Yin, B.; Yan, C.; He, S. LED-induced fluorescence spectroscopy technique for apple freshness and quality detection. *Postharvest Biol. Technol.* **2016**, *119*, 27–32. [CrossRef]
8. Fu, X.; Wang, X.; Rao, X. An LED-based spectrally tuneable light source for visible and near-infrared spectroscopy analysis: A case study for sugar content estimation of citrus. *Biosyst. Eng.* **2017**, *163*, 87–93. [CrossRef]
9. Gitlina, A.Y.; Surkova, A.; Ivonina, M.V.; Sizov, V.V.; Petrovskii, S.; Starova, G.L.; Koshevoy, I.O.; Grachova, E.V.; Kirsanov, D.O. Cyclometalated Ir(III) complexes as tuneable multiband light sources for optical multisensor systems: Feasibility study. *Dyes Pigments* **2020**, *180*, 108428. [CrossRef]
10. Surkova, A.A.; Paderina, A.V.; Legin, A.V.; Grachova, E.V.; Kirsanov, D.O. Cu(I)-based molecular emitters for quantification of fluoride and phosphate in surface waters. *Measurement* **2021**, *184*, 109976. [CrossRef]
11. Bunzli, J.C.G. Lanthanide luminescence for biomedical analyses and imaging. *Chem. Rev.* **2010**, *110*, 2729–2755. [CrossRef] [PubMed]
12. Bunzli, J.C.G.; Piguet, C. Taking advantage of luminescent lanthanide ions. *Chem. Soc. Rev.* **2005**, *34*, 1048–1077. [CrossRef]
13. Faulkner, S.; Pope, S.J.A.; Burton-Pye, B.P. Lanthanide complexes for luminescence imaging applications. *Appl. Spectrosc.* **2005**, *59*, 1–31. [CrossRef]
14. Khistiaeva, V.V.; Melnikov, A.S.; Slavova, S.O.; Sizov, V.V.; Starova, G.L.; Koshevoy, I.O.; Grachova, E.V. Heteroleptic β-diketonate Ln(III) complexes decorated with pyridyl substituted pyridazine ligands: Synthesis, structure and luminescence properties. *Inorg. Chem. Front.* **2018**, *5*, 3015. [CrossRef]
15. Kasture, M.; JadHAV, S.; Fouad, H.; Gosavi, S. Detection of melamine and urea in milk and milk products using graphene/goldnano-composite. *Sens. Lett.* **2015**, *13*, 471–474. [CrossRef]
16. Sjiström, M.; Wold, S.; Lindberg, W.; Martens, H. A multivariate calibration problem in analytical chemistry solved by partial least squares models in latent variables. *Anal. Chim. Acta* **1983**, *150*, 61–70. [CrossRef]
17. Ebens, K.H. Multivariate Data Analysis—In Practice, 5th ed.; CAMO Process AS: Oslo, Norway, 2001.
18. Lee, L.C.; Liong, C.Y.; Jemain, A.A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analyst* **2018**, *143*, 3526–3539. [CrossRef]
19. Bunzli, J.C.G.; Eliseeva, S.V. Lanthanide NIR luminescence for telecommunications, bioanalyses and solar energy conversion. *J. Rare Earths* **2010**, *28*, 824–842. [CrossRef]
20. 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. [CrossRef]
21. Tsenkova, R.; Atanassova, S.; Itoh, K.; Ozaki, Y.; Toyoda, K. Near-infrared spectroscopy for bio-monitoring: Cow milk composition measurement in a spectral region from 1100 to 2400 nanometers. *J. Anim. Sci.* **2000**, *78*, 515–522. [CrossRef]
22. Šašić, S.; Ozaki, Y. Short-wave near-infrared spectroscopy of biological fluids. Quantitative analysis of fat, protein, and lactose in raw milk by partial least-squares regression and band assignment. *Anal. Chem.* **2001**, *73*, 64–71. [CrossRef] [PubMed]
23. Kalinin, A.V.; Krashelininikov, V.N. Portable milk product quality analyzer based on spectrometry in the near IR range. *J. Appl. Spectrosc.* **2008**, *75*, 288–294. [CrossRef]
24. Aernouts, B.; Polshin, E.; Lammetyn, J.; Saey, W. Visible and near-infrared spectroscopic analysis of raw milk for cow health monitoring: Reflectance or transmittance. *J. Dairy Sci.* **2011**, *94*, 5315–5329. [CrossRef] [PubMed]
25. 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. [CrossRef]

26. Junsomboon, J.; Jakmunee, J. Flow injection conductometric system with gas diffusion separation for the determination of Kjeldahl nitrogen in milk and chicken meat. Anal. Chim. Acta 2008, 627, 232–238. [CrossRef]

27. Ezhilan, M.; Gumpu, M.B.; Ramachandra, B.L.; Nesakumar, N.; Babu, K.J.; Krishnan, U.M.; Rayappan, J.B.B. Design and development of electrochemical biosensor for the simultaneous detection of melamine and urea in adulterated milk samples. Sens. Actuators B Chem. 2017, 238, 1283–1292. [CrossRef]

28. Trivedi, U.B.; Lakshminarayanan, D.; Kothari, I.L.; Patel, N.G.; Kapused, H.N.; Makhija, K.K.; Patel, P.B.; Panchale, C.J. Potentiometric biosensor for urea determination in milk. Sens. Actuators B Chem. 2009, 140, 260–266. [CrossRef]

29. Bueno, L.; de Araujo, W.R.; Salles, M.O.; Kussuda, M.Y.; Paixão, T.R.L.C. Voltammetric electronic tongue for discrimination of milk adulterated with urea, formaldehyde and melamine. Chemosensors 2014, 2, 251–266. [CrossRef]

30. Zhao, K.; Liu, Y.; Zhang, Q. Dielectric behavior of adulterated milk with urea and water. J. Mol. Liq. 2019, 273, 37–44. [CrossRef]

31. Forchetti, D.A.P.; Popp, R.J. Use of NIR hyperspectral imaging and multivariate curve resolution (MCR) for detection and quantification of adulterants in milk powder. LWT-Food Sci. Technol. 2017, 76, 337–343. [CrossRef]

32. Müller-Maatsch, J.; Alewijn, M.; Wijitven, M.; Weesepoel, Y. Detecting fraudulent additions in skimmed milk powder using a portable, hyphenated, optical multi-sensor approach in combination with one-class classification. Food Control 2021, 121, 107744. [CrossRef]

33. Siesler, H.W.; Kawata, S.; Heise, H.M.; Ozaki, Y. Near-Infrared Spectroscopy: Principles, Instruments, Applications, 1st ed.; Wiley-VCH: Weinheim, Germany, 2002.

34. Choudhary, S.; Joshi, A. Development of an embedded system for real-time milk spoilage monitoring and adulteration detection. Int. Dairy J. 2021, 127, 105207. [CrossRef]

35. Kawasaki, M.; Kawamura, S.; Tsukahara, M.; Morita, S.; Komiya, M.; Natsuga, M. Near-infrared spectroscopic sensing system for on-line milk quality assessment in a milking robot. Comput. Electron. Agric. 2008, 63, 22–27. [CrossRef]