

## Article

# Integration of Sensor Fusion to Enhance Quality Assessment of White Brine Cheeses

Zlatin Zlatev <sup>1,\*</sup> , Tatjana Spahiu <sup>2</sup> , Ira Taneva <sup>1</sup>, Milen Dimov <sup>1</sup> and Miroslav Vasilev <sup>1</sup>

<sup>1</sup> Faculty of Technics and Technologies, Trakia University, 38 Graf Ignatiev Str., 8602 Yambol, Bulgaria; ira.dimitrova@trakia-uni.bg (I.T.); ilen.dimov@trakia-uni.bg (M.D.); miroslav.vasilev@trakia-uni.bg (M.V.)

<sup>2</sup> Department of Textile and Fashion, Polytechnic University of Tirana, 1019 Tirana, Albania; tspahiu@fim.edu.al

\* Correspondence: zlatin.zlatev@trakia-uni.bg

**Abstract:** The article examines the main characteristics of white brine cheeses from different manufacturers and changes in their quality indicators. These characteristics include the active acidity, electrical conductivity, total dissolved solids, oxidation–reduction potential, and organoleptic assessment. In this context, the connection to biomimetics lies in the approach of integrating multiple sensory modalities, similar to how biological systems often use multiple senses to perceive and understand their environment. For this purpose, spectral, ultrasonic, and gas characteristics were used, from which informative indices were extracted, united at a later stage in a vector of features. Based on the classification, it was found that the optical characteristics of cheeses from different manufacturers overlap, thus making it possible to predict the main indicators for each type of cheese. The results show that the use of a multimodal approach combining features from different sensors contributes to a better understanding of the variations in cheese properties, while improving the predictive abilities of the created models. The obtained results give a clear idea of the quality of the cheese, thus enabling adequate decisions to be made during the production process.

**Keywords:** multimodal sensory analysis; cheese quality assessment; non-destructive analysis; biomimetics-inspired approach; data integration and algorithm adaptation



**Citation:** Zlatev, Z.; Spahiu, T.; Taneva, I.; Dimov, M.; Vasilev, M. Integration of Sensor Fusion to Enhance Quality Assessment of White Brine Cheeses. *Dairy* **2024**, *5*, 249–270. <https://doi.org/10.3390/dairy5020021>

Academic Editor: Ali Adnan Hayaloglu

Received: 7 April 2024  
Revised: 23 April 2024  
Accepted: 28 April 2024  
Published: 30 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Milk, lactic acid products, and cheese are foods with proven health effects [1]. This is due to their complete composition and content that includes a number of biologically active components: proteins, fats, minerals, trace elements, organic acids, a high content of calcium, conjugated linoleic acid, vitamins, valuable lactic acid bacteria, and others.

The growing demand for different types of cheeses makes manufacturers look for new methods for their production, changing the conditions of coagulation and ripening and using different additives or acids to obtain cheeses with a certain texture and taste [2].

In international markets, producers compete by offering an ever-widening variety of types of cheeses. The quality of cheese, defined by parameters such as appearance, taste, texture, functionality, and nutritional value, is an essential aspect for both consumers and producers.

In Bulgaria, a number of producers of white brine cheese use standardized raw materials and processes for its production. The producers adhere to strictly coordinated technological processes, allowing minimal permissible deviations in them, and here also in the quality of the cheese. This leads to the search for new methods to evaluate the quality of the finished product, such as color images and spectral characteristics, through which we recognize the existing differences in the same product category (white brined cheese), even though they are produced according to the same standard.

Two main factors related to the main raw materials used for cheese production and its ripening contribute to the expectation of variation in the basic properties obtained from

color images and spectral characteristics of white brine cheese when produced to the same standard but from different producers [3].

One factor is the milk that can be taken from a cow, sheep, or goat, which can vary from dairy to dairy. This difference may contribute to discrepancies in the visual and spectral characteristics of the final products [4].

The other factor leading to variations in cheese properties is the “ripening” process, its duration, and the temperature regime [5].

These factors can influence the quality of the finished product and be captured by color images and spectral characteristics. Despite these differences in cheese production, the demands of the market impose a relatively constant quality, regardless of its producer. In order to meet this requirement, the use of online monitoring and evaluation technologies is required, allowing adjustments to be made in the technological processes continuously to maintain the preset quality standards.

Due to the high protein content of cheese, it is categorized as a high-risk product when stored under uncontrolled temperature conditions or under-ripened. Therefore, the provision of reliable laboratory conditions, the use of precise chemical reagents, and the involvement of qualified experts in the examination of each sample are imperative. Despite the expertise of these professionals, subjective variations in the evaluation of cheese color indicators continue to exist [6].

In the context of preserving cheese quality, the determination of internal defects becomes paramount [7]. To ensure the successful preservation of cheese products, the development of non-destructive monitoring techniques is necessary. These techniques help to track changes in the quality indicators of products during their storage, thus contributing to their overall control.

As a result of the analysis of the available literary sources, it can be concluded that there is a need to improve and expand the techniques for evaluating the quality of cheese. While medical-imaging methods such as X-rays and magnetic resonance imaging offer accurate results, their high cost and complexity require the development of more affordable and user-friendly devices. The integration of gas and optical sensors provides a comprehensive approach enabling the detection of volatile compounds for freshness and ripeness, together with the evaluation of visual characteristics. In addition, the non-destructive nature of ultrasound is highlighted, as it allows for a sensitive assessment of fat and moisture content without altering the structure of the cheese, which is crucial for maintaining the integrity of the product during production.

In the field of spectroscopy, there is a need to improve the application of various techniques, such as NIR, VIS, FTIR, hyperspectral, multispectral, and dielectric spectroscopy, in the evaluation of cheese quality. Future research directions emphasize the importance of refining the predictive models, evaluating the performance of the portable spectrometer in practical environments, investigating the impact of spectrum broadening, and expanding the ability to predict additional chemical components. The incorporation of advanced chemometric approaches, such as Bayesian methods and machine learning, is recommended to improve accuracy, especially for challenging features. Rigorous validation and testing strategies for industrial applications are recommended, emphasizing the holistic and diverse approach needed to increase the reliability of VIS-NIR spectroscopy in predicting different components of cheese composition.

The combination of sensor and data fusion in cheese quality assessment synergistically improves the accuracy and comprehensiveness of the assessment. By integrating information from different sensors, such as gas, ultrasonic, and optical sensors, the approach ensures the determination of key quality characteristics, including freshness, ripeness, color, and texture. This comprehensive analysis ensures a more reliable assessment by compensating for the limitations of individual sensors and offering cross-validation. The adaptability of the approach to diverse parameters and the holistic understanding it provides make the fusion of sensors and data a powerful tool to improve the overall evaluation of cheese quality in a multifaceted way.

An integrated approach using ultrasonic, spectral, and gas characteristics not only provides objective data on the composition and structure of cheese but also successfully replaces or complements traditional organoleptic analysis. Ultrasonic features provide a sensitive and accurate analysis of fat and moisture content, while preserving product integrity. Spectral features further highlight this multi-layered approach, providing information on the cheese's texture, structure, and chemical composition. Gas characteristics, in turn, play a key role in aroma and flavor analysis through the detection of various gases that contribute to the cheese's characteristic qualities. This modern and technologically advanced approach provides a comprehensive and objective evaluation of the product, which ultimately helps achieve the desired quality and taste characteristics of the cheese.

The aim of the present work is to predict the main characteristics of white brine cheeses from different producers, characterizing changes in their quality indicators based on an integrated approach using ultrasonic, spectral, and gas characteristics. The focus is on the importance of the implementation of various sensory inputs for a deeper understanding of cheese properties by highlighting the limitations of independent evaluations of the use of single characteristics for a quality analysis. The introduction of a novel approach that integrates ultrasonic, spectral, and gas characteristics would play a key role in overcoming these limitations and enhancing cheese quality assessment. By gathering data taken from diverse sensors, this research aims to improve predictive models, foster adaptability in decision-making processes, and, at the end, enhance cheese production outcomes.

## 2. Material and Methods

The research was conducted in laboratory conditions at a temperature of  $10 \pm 2$  °C and a relative air humidity of  $70 \pm 3\%$  RH. The number of observations is 120 per product from each manufacturer. The spectral, gas, and ultrasonic characteristics of the cheese were investigated. The processing of the data was conducted by a method for their fusion. Classification procedures and regression analysis were also used.

### 2.1. Source Data for the Cheese Used

The research focuses on samples of white brine cheese manufactured by various producers, all adhering to the same standard (according to the standard BNS 14-2010 "Bulgarian cheese"). The samples were purchased from the commercial network of the city of Yambol, Bulgaria.

Table 1 shows data recorded on the package labels for the cheese used by three manufacturers. Although all producers use cow's milk, sourdough, yeast, salt, and calcium dichloride, there are subtle differences in additional components, such as rennet enzyme and citric acid. The nutritional content also varies between cheeses, with differences in fat content, carbohydrate composition, protein levels, salt content, and energy value. In particular, Manufacturer 2 (M2) stood out with significantly higher levels of saturated fat and carbohydrates compared to the other manufacturers. These differences offer valuable insight into the variety of cheeses available on the market catering to different dietary preferences and requirements.

**Table 1.** Cheese ingredients from different manufactures.

Manufacturer (M)	M1	M2	M3	Manufacturer (M)	M1	M2	M3
Fat, %	21.5	22	22.5	Proteins, %	18	17	13
Fats, of which saturated fat, %	13.2	15	13.2	Salt, %	3.5	3.5	3.5
Carbohydrates, %	0.5	4.2	0.7	Water content, %	54	-	-
Carbohydrates, of which sugars, %	0.5	4.2	0.7	Energy value, kJ/268 kcal	1121.3	1131	1070

## 2.2. Determination of pH, EC, TDS, and ORP

The samples were meticulously prepared for measurement following the rigorous methodology outlined in AACC 02-52.01 Hydrogen-Ion Activity (pH)—Electrometric Method. Following this protocol, distilled water was heated to 70 °C, and each cheese sample was dissolved in distilled water at a precise ratio of 1:10 (5 g of raw material per 50 mL of distilled water). The mixture underwent periodic stirring until a homogeneous solution was achieved, and then it was cooled to ambient temperature. To ensure accuracy and reliability, three consecutive measurements were taken for each characteristic, and their mean values, along with standard deviations, were calculated. The technical means used for these measurements are presented in Table 2.

**Table 2.** Technical means used in this study.

Technical Means	Apparatus	Producer
Temperature	Digital thermometer	Shanghai Vihua V&A instrument CO., Ltd., Shanghai, China
Material weight	Pocket Scale MH-200	ZheZhong Weighing Apparatus Factory, Yongkang City, China
Active acidity	pH meter PH-108	Hangzhou Lohand Biological Co., Ltd., Jiubao Town, China
Electrical conductivity	Conductivity Meter AP-2	HM Digital, Inc., Culver City, CA, USA
Total dissolved solids	Measuring instrument TDS-3	HM Digital, Inc., Culver City, CA, USA
Oxidation–reduction potential	Measuring instrument ORP-2069	Shanghai Longway Optical Instruments Co., Ltd., Shanghai, China

## 2.3. Organoleptic Analysis of Cheese

The organoleptic evaluation included serves as an additional component aimed at providing a more comprehensive view of the product. Collective evaluations were formed in the organoleptic analysis of white brine cheese based on the individual evaluations of experts in the field. The studied indicators and assessment requirements are based on the BDS 14:2010 standard. The tasting evaluation is made on a 9-point scale (1–9) with step 1 (from 1, does not correspond to the indicator, to 9, fully meets the requirements). All the results presented in the following table have statistically significant differences at  $p < 0.05$  (Table 3).

**Table 3.** Results from organoleptic assessment of cheese.

Manufacturer Characteristic	M1	M2	M3	Mean
Appearance of the pieces	7 ± 1	7 ± 1.3	8 ± 1	7.33 ± 0.58
Consistency	8 ± 1	8 ± 1	7 ± 1.1	8 ± 1
Cut surface, construction, and color	8 ± 1	8 ± 1	7 ± 1.3	8.33 ± 1.16
Taste and aroma	7 ± 1.2	7 ± 1.5	8 ± 1	7.33 ± 0.58

## 2.4. Obtaining Spectral Characteristics and Indices

The spectral characteristics were obtained according to the methodology presented in Dineva et al. [8].

A Huawei P10 mobile device video sensor (Huawei Technologies Co., Ltd., Shenzhen, China) was used.

Color correction was performed using a 24-field color scale, namely the Danes Picta BST11 color chart (Danes-Picta, Praha, Czech Republic).

The obtained values from the RGB color model were transformed to the XYZ model in reflectance spectra in the visible (VIS) range, covering 390–730 nm, respectively. The matrices used to convert color components into reflectance spectra in the visible range have an observer angle of 2° (LMS 2°, CIE 2006). The illuminance data are in accordance

with the D65 standard (average daylight with the UV component, 6500 K). Additionally, the conversion functions between the RGB and XYZ models in the 380–780 nm range include transformation matrices, taking into account certain observer (2°) and illuminance (D65) conditions.

Since cheese is a specific product, there are no precisely defined spectral indices for it. These should be defined specifically for cheese.

Ju et al. [9] and Mendiguren et al. [10] defined the basic spectral indices used in the study. These indices are not calculated at fixed spectral wavelengths. To be used, it is important to select six informative wavelengths specific to the product under investigation. The general form of these indices ( $SI$ ) is as follows:

$$SI_1 = \frac{R_{\lambda_2} - R_{\lambda_1}}{R_{\lambda_2} + R_{\lambda_1}} \quad (1)$$

$$SI_2 = \frac{R_{\lambda_1} - R_{\lambda_2}}{R_{\lambda_1} + R_{\lambda_2}} \quad (2)$$

$$SI_3 = \frac{R_{\lambda_1}}{R_{\lambda_2}} \quad (3)$$

$$SI_4 = R_{\lambda_1} - R_{\lambda_2} \quad (4)$$

$$SI_5 = 2.5 \frac{R_{\lambda_1} - R_{\lambda_2}}{R_{\lambda_2} + 6R_{\lambda_1} - 7.5R_{\lambda_3}} \quad (5)$$

$$SI_6 = \frac{R_{\lambda_2} - R_{\lambda_5}}{R_{\lambda_2} + R_{\lambda_5}} \quad (6)$$

$$SI_7 = \frac{R_{\lambda_2} - R_{\lambda_6}}{R_{\lambda_2} + R_{\lambda_6}} \quad (7)$$

$$SI_8 = \frac{R_{\lambda_2}}{R_{\lambda_5}} \quad (8)$$

$$SI_9 = \frac{R_{\lambda_4} - R_{\lambda_1}}{R_{\lambda_4} + R_{\lambda_1} - R_{\lambda_3}} \quad (9)$$

where  $\lambda$  (nm) is the spectral wavelength, and  $R$  is the reflectance at a specific wavelength.

### 2.5. Obtaining Data from Gas Sensors

Data from gas sensors were acquired by a system that consists of a sensor module and a personal computer, as presented in Baycheva et al. [11]. The sensor module is based on a single-board Mega computer (INHAOS Technology Co., Ltd., Dongguan, China). It uses four MQ-xx series metal oxide sensors (Zhengzhou Winsen Electronics Technology Co., Ltd., Zhengzhou, China). The sensors can detect MQ-3, alcohol compounds and benzene; MQ-4, methane, propane, and butane; MQ-6, propane, butane, and LPG; and MQ-135, ammonium compounds and sulfides.

In the present work, data from gas sensors are processed with a Kalman filter. Combining them is realized by the central limit theorem and the Fraser–Potter fixed-interval smoother.

The signals from the gas sensors are precisely filtered by a Kalman filter and fed to a combining unit. A combined synthetic signal is obtained at the output of the software sensor.

Another 12 statistical features were obtained. They were calculated according to Matz et al. [12] and Zhang et al. [13]. The features describing the combined data from the gas sensors ( $GI$ ) have the following form:

$$GI_1 = \frac{1}{N} \sum_{i=1}^N x_i \quad (10)$$

$$GI_2 = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \text{mean})^2} \quad (11)$$

$$GI_3 = N \frac{\sum_{i=1}^N (x_i - \text{mean})^4}{\left(\sum_{i=1}^N (x_i - \text{mean})^2\right)^2} \quad (12)$$

$$GI_4 = \frac{\sum_{i=1}^N (x_i - \text{mean})^3}{(N-1)SD^3} \quad (13)$$

$$GI_5 = \sum_{i=1}^N x_i^2 \quad (14)$$

$$GI_6 = \sum_{i=1}^N x_i \log_2 x_i \quad (15)$$

$$GI_7 = \frac{SD}{\text{mean}} \quad (16)$$

$$GI_8 = \sum_{i=1}^N \frac{x_i}{i} \quad (17)$$

$$GI_9 = \sum_{i=1}^N ix_i \quad (18)$$

$$GI_{10} = \sum_{i=1}^N \frac{(i-\bar{i})(1-\text{mean})x_i}{SD \cdot \sigma_i} \quad (19)$$

$$GI_{11} = \sum_{i=1}^N i^2 x_i \quad (20)$$

$$GI_{12} = \sum_{i=1}^N ix_i \quad (21)$$

where  $N$  denotes the number of reports in one combined characteristic from the gas sensors, and  $x$  is the amplitude of the signal in the eighth report.

## 2.6. Obtaining Ultrasonic Characteristics

A system presented in [14] was used to obtain ultrasonic characteristics. This system is characterized in that all data acquisition and basic processing operations are performed by the single-board microcomputer. The system consists of an ultrasonic sensor, humidity and temperature sensors, a removable stand, and a basis. The compensation of the ultrasound signal by humidity and temperature was performed according to the methodology presented by Ilarionov et al. [15]. A total of 12 features were obtained. They were calculated according to Matz et al. [12] and Zhang et al. [13].

The features describing the ultrasonic characteristics ( $UI$ ) of cheese have the following form:

$$UI_1 = \frac{1}{N} \sum_{i=1}^N x_i \quad (22)$$

$$UI_2 = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (23)$$

$$UI_3 = N \frac{\sum_{i=1}^N (x_i - \bar{x})^4}{\left(\sum_{i=1}^N (x_i - \bar{x})^2\right)^2} \quad (24)$$

$$UI_4 = \frac{\sum_{i=1}^N (x_i - \bar{x})^3}{(N-1)SD^3} \quad (25)$$

$$UI_5 = \sum_{i=1}^N x_i^2 \quad (26)$$

$$UI_6 = \sum_{i=1}^N x_i \log_2 x_i \quad (27)$$

$$UI_7 = \frac{SD(x)}{\bar{x}} \quad (28)$$

$$UI_8 = \sum_{i=1}^N \frac{x_i}{i} \quad (29)$$

$$UI_9 = \sum_{i=1}^N ix_i \quad (30)$$

$$UI_{10} = \sum_{i=1}^N \frac{(i - \bar{i})(1 - \bar{x})x_i}{SD \cdot \sigma_i} \quad (31)$$

$$UI_{11} = \sum_{i=1}^N i^2 x_i \quad (32)$$

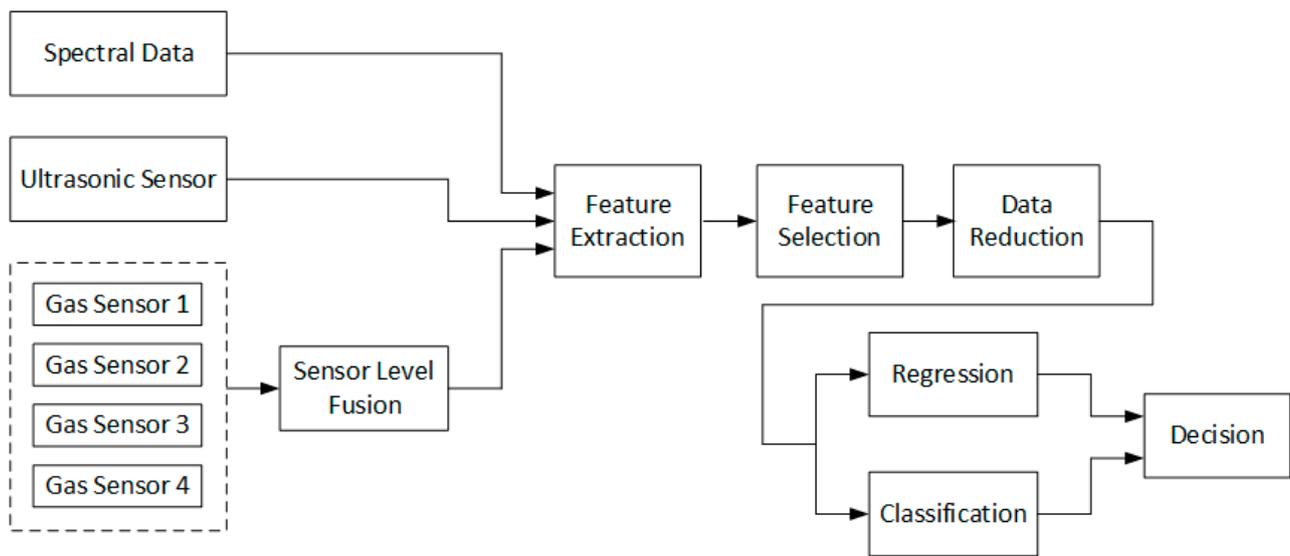
$$UI_{12} = \sum_{i=1}^N ix_i \quad (33)$$

where  $N$ —the number of reports for one ultrasound characteristic;  $X$ —the amplitude of the ultrasound signal in the eighth report; and  $SD$ —the standard deviation.

### 2.7. Fusion of Sensor Data

An early feature fusion method was used, as detailed in Pereira et al. [16]. It is a methodology more commonly used in artificial intelligence, machine learning, and pattern recognition tasks. In this approach, features extracted from different levels are merged or combined at an early stage. This enables the integration of low- and high-level features, capturing both detailed and abstract information simultaneously. By combining features, classifiers and regression models provide more comprehensive representations of the input data, potentially leading to improved performance in tasks such as data image classification, object detection, and semantic segmentation.

Figure 1 shows a flow diagram of the approach used to combine data from the spectral, ultrasonic, and gas characteristics of cheese. The outputs of spectral and ultrasonic characteristics are used directly; meanwhile, those of the four gas sensors are combined into a common characteristic. From the three types of characteristics—ultrasonic, spectral, and gas—are extracted features in the form of indices. The most informative features are selected and used to create vectors. These vectors are reduced, and classification and regression are performed. The output from these methods is used in making a decision about the condition of the product—in this case, cheese.



**Figure 1.** Early feature fusion steps.

### 2.8. Classification Methods

Classification is an important aspect of machine learning, where the goal is to determine the class or category to which an object or dataset belongs. Through the different classification procedures, an answer can be given to the question of whether the data describing the characteristics of the examined cheeses are similar or significantly different. For this purpose, the following classifiers were used:

**Bayesian classifier.** The naïve Bayesian Classifier (NBC) is a statistical classification algorithm based on a Bayesian probabilistic model that allows for the determination of the probability of an event occurring when some data about it are known. This method is “naïve” in relation to the assumption that object attributes are statistically independent, i.e., that the presence or value of one attribute does not provide information about other attributes. Conditional independence between individual attributes allows the method to process large volumes of data efficiently.

**Discriminant analysis.** The discriminant analysis is a data classification method that uses a grouping variable, as explained by Kirilova et al. [17] and Nachev et al. [18]. The procedure can be implemented through the use of linear or non-linear partition functions. The non-linear variant of the method (e.g., QDA—quadratic discriminant analysis) is considered more suitable for large datasets. This is due to the lower bias and larger variation of QDA. On the other hand, the linear variant (LDA—linear discriminant analysis) may be preferred for smaller datasets with higher clustering and lower variance.

The following discriminant functions were used in the discriminant analysis:

**Linear:** a linear partition function applicable to data with a multivariate normal density of each group and a common covariance estimate.

**Quadratic:** a function using covariance to group the data. The separation of the groups is by a non-linear function—in most cases, of the second degree.

**Support vector machines method.** The support vector machines (SVMs) method serves to represent the training data in an  $n$ -dimensional space in order to achieve linear separability. Items from the training sample are associated with one of two classes. The data are transformed into a new domain where the resulting model has the ability to dispose of them so that there is a clear distinction between classes. With SVM, high-dimensional hyperplanes are created that serve classification purposes. The efficiency of the method is achieved when there is a large distance between the two classes of the training data. In the presence of a larger distance, a smaller classification error is observed. Thus, SVM seeks to create an optimal separation between classes that ensures the reliable and accurate classification of new data.

If the training data can be linearly separated, the support vector method finds two boundary planes that do not pass through any data points. After the linear algorithm of the SVM method, those using non-linear kernel (kernel) dividing functions [11] were created. This allows the hyperplane to be projected with the maximum distance between the two classes in the transformed feature space. The present study used the following partition functions:

**Linear:** A term applied to data with a multivariate normal density of each group and a common covariance estimate.

**Quadratic:** A quadratic function that sorts the multivariate normal density data by calculating the covariance and grouping them.

**RBF (Radial Basis Function):** The function is represented by radial basis elements. SVM uses the RBF kernel to transform the data into a higher-dimensional space, which can facilitate solving complex classification and regression tasks.

### 2.9. Evaluation of the Performance of Classification Procedures

Basic, actual, and total classification errors were calculated for the number of classes:

$$\text{Main error } e_i = \frac{FN_i}{TP_i + FN_i} \cdot 100, \% \quad (34)$$

$$\text{Actual error } g_i = \frac{FP_i}{TP_i + FP_i} \cdot 100, \% \quad (35)$$

$$\text{Total error } e_0 = \frac{\sum_{i=1}^m FN_i}{\sum_{i=1}^m TP_i + \sum_{i=1}^m FN_i} \cdot 100, \% \quad (36)$$

The basic error is a measure of how much data from class  $i$  is misclassified into other classes, where  $FN$  is the number of data from class  $i$  misassigned to other classes, and  $TP$  is the number of correctly classified data from class  $i$ . The actual error indicates the relative proportion of data from other classes incorrectly assigned by the classifier to a given class  $i$ , where  $FP$  is the number of data from other classes associated with class  $i$ . The total error shows the misclassified data relative to all the data in the sample.

### 2.10. Regression Methods

Regression prognostic models based on chemometric techniques were applied to quantitatively analyze the cheese data. Partial least squares regression (PLSR) and principal component regression (PCR) were used. Through these techniques, new regression factors are created that concentrate information from the entire spectrum of the data used. A key aspect of PCR is obtaining the so-called principal components, which serve as new predictors in the regression model. This allows the method to be used when a significant number of variables are present and are highly correlated. In the PLSR method, a linear combination of predictors is constructed that maximizes the correlation with the target variable. The aim is to find new, more informative features that are highly correlated with the target variable, thus improving the predictive ability and interpretability of the model.

In addition to the regression models obtained by the mentioned chemometric techniques, a second-order prediction model was also used, which is more often applied to describe the change in quality indicators of cheese and, in general, of products of biological origin. The model looks like this:

$$z = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 \quad (37)$$

where  $z$  is the dependent variable,  $x$  and  $y$  are the independent variables, and  $b$  denotes the coefficients of the model. The evaluation of the model is realized by the coefficient of determination, Fisher's test,  $p$ -test, and standard error. An analysis of the residuals of the resulting model was carried out.

### 2.11. Validation of the Resulting Models

In creating reliable and applicable regression models describing the cheese data with sufficient accuracy, an important step is their validation. It is an integral part of the modeling process and is a key stage for detecting and solving potential problems. The validation process ensures that the built model is well calibrated so that it can successfully handle a variety of data, including those that were not used during training. In the validation process, a separate dataset comprising 30% of the total group is used to test the performance of the regression models.

The results are presented by the distribution of the measured and predicted values of the sought technological quantity, showing how these values are distributed relative to the appropriate regression line. This analysis visualizes the degree of agreement between the model predictions and the actual measurements, providing a clear insight into the accuracy of the regression model. When predicting the quality indicators of the cheese, the determination of the relationship between the real and predicted values of the technological parameters was carried out by means of the coefficient of determination ( $R^2$ ), mean square error ( $MSE$ ), root mean square error ( $RMSE$ ), mean absolute error ( $MAE$ ), and standard error ( $SE$ ). These errors are calculated using the following formulas:

$$MSE \quad MSE = \frac{1}{n} \sum_{i=1}^n (y_m - y_p)^2 \quad (38)$$

$$RMSE \quad RMSE = \sqrt{MSE} \quad (39)$$

$$MAE \quad MAE = \frac{1}{n} \sum_{i=1}^n |y_m - y_p| \quad (40)$$

$$SE \quad SE = \sqrt{\frac{\sum_{i=1}^n (y_m - y_p)^2}{n - 2}} \quad (41)$$

where  $n$  is the number of data,  $y_m$  is the actual measured value, and  $y_p$  is the predicted value.

The coefficient of determination ( $R^2$ ) represents the proportion of the total variation in the predicted values that can be explained by the actual measured values. A high value of the indicator indicates a greater fit between the model and the real data. To determine a more accurate relationship between actual and predicted data, a thorough evaluation using additional criteria is necessary.

**MSE**—This criterion calculates the arithmetic mean of the squares of the differences between the actual and predicted values. A lower MSE value indicates less bias and better model precision. **RMSE** determines the average size of the error between the predicted and actual values. The smaller the **RMSE**, the more accurate the validated regression model. **MAE** is a measure of the average magnitude of the errors without regard to their direction. **MAE** treats all errors equally unlike **RMSE** where large errors are reflected. **SE** is a measure of how likely the mean of the actual measured data is to deviate from the actual mean of the actual measured data. The smaller the standard error, the more accurate the validated regression model.

To perform the necessary analyses and calculation procedures, we used the products MATLAB version 2017b (MathWorks Inc., Natick, MA, USA), Stat Soft Statistica version 12 (TIBCO Software Inc., Palo Alto, CA, USA), and MS Excel version 2016 (Microsoft Corp., Albuquerque, NM, USA).

## 3. Results and Discussion

### 3.1. Results of Technological Measurements of Cheese

Table 4 shows results for the technological characteristics of cheese from three manufacturers. The results reveal several observations regarding the technological characteristics of the cheese from the three manufacturers. First, variations in pH levels across manufacturers indicate potential differences in acidity that can affect flavor, texture, and microbial activity during cheese production and ripening. Manufacturer 2 (M2) stands out with

the lowest pH value, suggesting a potentially sharper taste compared to the other manufacturers. Second, differences in electrical conductivity (EC) and total dissolved solids (TDS) suggest differences in mineral content and salt concentration among cheese samples. Manufacturer 3 (M3) shows the lowest EC and TDS values, indicating potentially different processes or compositions that influence the mineral profile of the cheese. Finally, variations in the oxidation–reduction potential (ORP) suggest differences in the oxidation state of the cheese, which can affect its shelf life and susceptibility to spoilage. In particular, Manufacturer 2 (M2) presents the highest ORP value, suggesting a potentially more stable product in terms of oxidation.

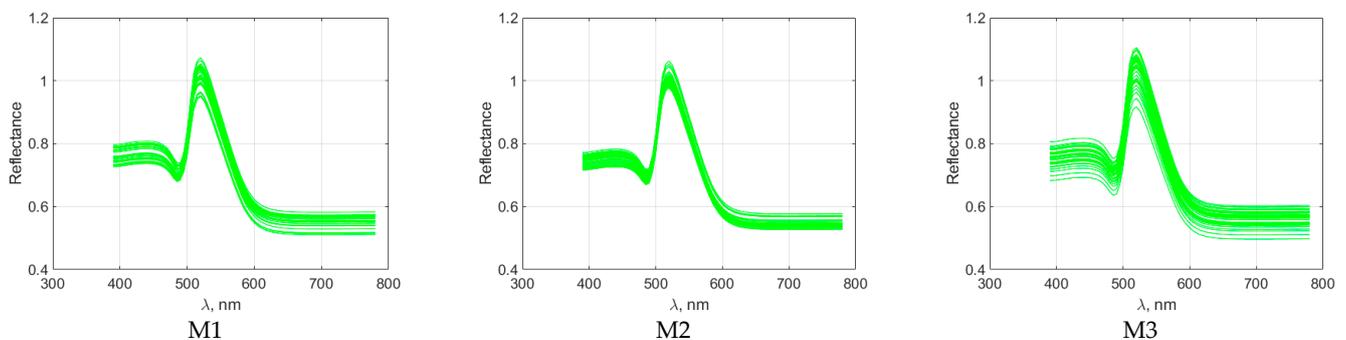
**Table 4.** Technological characteristics of cheese. All data have statistically significant difference at  $p < 0.05$ .

Manufacturer Characteristic	M1	M2	M3
pH	$6.7 \pm 0.4$	$6.5 \pm 0.4$	$6.6 \pm 0.4$
EC, $\mu\text{S}/\text{cm}$	$2356 \pm 149$	$2310 \pm 114$	$1996 \pm 177$
TDS, ppm	$450 \pm 51$	$445 \pm 71$	$452 \pm 63$
ORP, mV	$202 \pm 52$	$212 \pm 41$	$210 \pm 33$

M, manufacturer; pH, active acidity; EC, electrical conductivity; TDS, total dissolved solids; ORP, oxidation–reduction potential.

### 3.2. Results of Determination of Spectral Characteristics and Indices

Figure 2 shows VIS spectral reflectance characteristics for cheese from three manufacturers. In the investigated spectral range of 380–780 nm, the highest reflectance values are observed for manufacturer M3, compared to the other two manufacturers. The lowest reflectance values are observed for manufacturer M2. These differences are due to differences in the fat content and whether the milk contains beta-carotene, which depends on the diet of the animals from which the milk is obtained.



**Figure 2.** Spectral characteristics of cheese from three manufacturers.

The results of a selection of informative spectral wavelengths by the RReliefF method show that the most informative wavelengths are observed in the ranges 500–550 nm and 620–690 nm, and the least in the range 410–490 nm. Six informative wavelengths were determined and are presented in Table 5. They cover three wavelengths for the red spectrum of visible light, one for blue and two for green. Since cheese is close to white in color, this is why wavelengths from the entire visible light spectrum are informative.

**Table 5.** Informative spectral wavelengths.

Wavelength	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$
Value, nm	520	530	470	685	675	645

After substituting the informative wavelengths in the formulas for calculating spectral indices for cheese, they take the following form:

$$SI_1 = \frac{R_{530} - R_{520}}{R_{530} + R_{520}} \quad (42)$$

$$SI_2 = \frac{R_{520} - R_{530}}{R_{520} + R_{530}} \quad (43)$$

$$SI_3 = \frac{R_{520}}{R_{530}} \quad (44)$$

$$SI_4 = R_{520} - R_{530} \quad (45)$$

$$SI_5 = 2.5 \frac{R_{520} - R_{530}}{R_{530} + 6R_{520} - 7.5R_{470}} \quad (46)$$

$$SI_6 = \frac{R_{530} - R_{675}}{R_{530} + R_{675}} \quad (47)$$

$$SI_7 = \frac{R_{530} - R_{645}}{R_{530} + R_{645}} \quad (48)$$

$$SI_8 = \frac{R_{530}}{R_{675}} \quad (49)$$

$$SI_9 = \frac{R_{685} - R_{520}}{R_{685} + R_{520} - R_{470}} \quad (50)$$

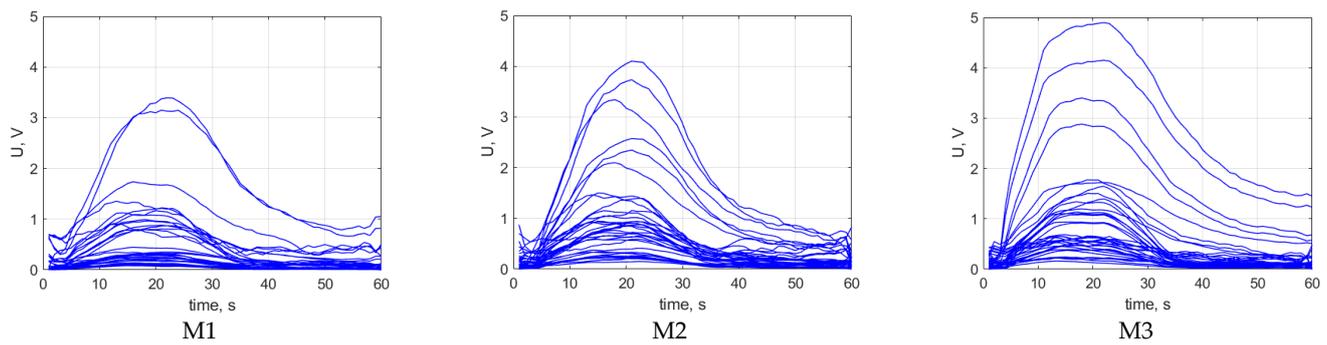
Table 6 shows the mean and standard deviation of spectral index values for cheese from three manufacturers. It can be seen that, for the first two manufacturers, the first three indices have close values. Sufficient differences were observed for the remaining six indices. All data have statistically significant difference at  $p < 0.05$

**Table 6.** Mean and standard deviation of spectral indices of cheese.

Manufacturer (M) Index	M1	M2	M3
SI1	0.0197 ± 0.0001	0.0197 ± 0.0001	0.0196 ± 0
SI2	0.0197 ± 0.0001	0.0197 ± 0.0001	0.0196 ± 0
SI3	1.0401 ± 0.0002	1.0401 ± 0.0001	1.04 ± 0.0001
SI4	0.0393 ± 0.0012	0.0388 ± 0.0008	0.04 ± 0.0017
SI5	0.0672 ± 0.0106	0.0645 ± 0.0048	0.0609 ± 0.0031
SI6	0.2804 ± 0.003	0.2796 ± 0.0021	0.2787 ± 0.0013
SI7	0.2783 ± 0.0029	0.2776 ± 0.002	0.2767 ± 0.0012
SI8	1.7794 ± 0.0115	1.7764 ± 0.0081	1.7729 ± 0.0049
SI9	0.5705 ± 0.0264	0.5646 ± 0.0144	0.555 ± 0.0094

### 3.3. Results of Obtaining Characteristics and Indices from Gas Sensors

Figure 3 shows the characteristics obtained from cheese gas sensors from three manufacturers. The differences in characteristics are based on the compounds: acetic acid, butanoic acid, dimethyl trisulfide, methional, hexanal, (E)-2-nonenal, acetoin, 1-octen-3-one,  $\delta$ -dodecalactone, furanol, hexanoic acid, heptanal, and ethyl caproate, which are formed during its ripening.



**Figure 3.** Gas characteristics for cheese from three manufacturers.

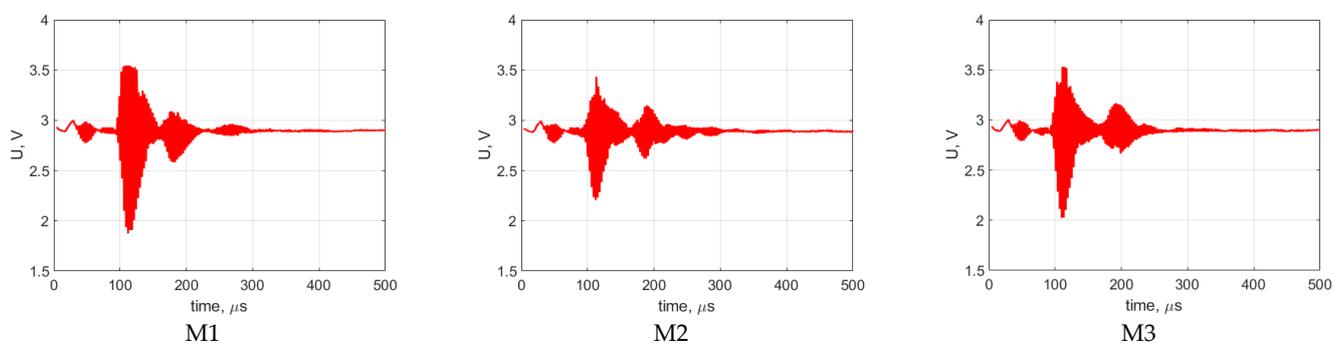
Table 7 plots the mean and standard deviation of index values for cheese from three manufacturers obtained from gas sensor data. All data have statistically significant difference at  $p < 0.05$

**Table 7.** Mean and standard deviation of indices of cheese obtained from gas sensors data.

Manufacturer (M) Index	M1	M2	M3
GI1	$0.32 \pm 0.4$	$0.6 \pm 0.3$	$0.58 \pm 0.06$
GI2	$0.2 \pm 0.5$	$0.38 \pm 0.15$	$0.35 \pm 0.04$
GI3	$1.72 \pm 0.15$	$1.79 \pm 0.19$	$1.76 \pm 0.15$
GI4	$0.5 \pm 0.13$	$0.53 \pm 0.09$	$0.4 \pm 0.17$
GI5	$1012.18 \pm 258.71$	$2261.71 \pm 383.16$	$2721.3 \pm 669.39$
GI6	$21.12 \pm 2.11$	$30.64 \pm 3.42$	$25.49 \pm 12.2$
GI7	$0.66 \pm 0.18$	$0.65 \pm 0.16$	$0.66 \pm 0.02$
GI8	$1.31 \pm 0.16$	$2.22 \pm 1.72$	$2.19 \pm 1.08$
GI9	$483.41 \pm 68.25$	$903.75 \pm 83.82$	$866.7 \pm 107.81$
GI10	$23.85 \pm 4.21$	$13.79 \pm 1.53$	$17.82 \pm 9.89$
GI11	$15,852.15 \pm 2316.75$	$29,359.01 \pm 2712.54$	$28,212.73 \pm 3731.72$
GI12	$483.41 \pm 68.25$	$903.75 \pm 83.82$	$866.7 \pm 107.83$

### 3.4. Results of Obtaining Ultrasonic Characteristics

The differences in characteristics of typical ultrasonic characteristics of cheese from three manufacturers are presented in Figure 4. They are based on the moisture content and fatty acid composition of the studied products.



**Figure 4.** Ultrasonic characteristics of cheese from three manufacturers.

Table 8 shows the informative features for the three manufacturers. Thirteen features were selected for manufacturer M1. With this manufacturer, the indices of gas and ultrasonic

sensors prevail. For manufacturer M2, 14 features were selected. At M2, the spectral indices prevail. For manufacturer M3, again 14 features were selected. Spectral and ultrasonic indices mainly prevail.

**Table 8.** Informative features for three manufacturers (Ms).

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
F	SI1	SI2	SI3	SI4	SI5	SI6	SI7	SI8	SI9	GI1	GI2	GI3	GI4	GI5	GI6	GI7	GI8
M1	0.35	0.40	0.34	0.36	0.42	0.55	0.65	0.62	0.55	0.60	0.77	0.52	0.66	0.45	0.39	0.18	0.68
M2	0.96	0.82	0.77	0.58	0.63	0.61	0.74	0.63	0.56	0.53	0.57	0.60	0.58	0.62	0.66	0.50	0.74
M3	0.80	0.91	0.91	0.92	0.66	0.51	0.47	0.66	0.83	0.83	0.58	0.38	0.10	0.35	0.40	0.34	0.32
No.	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	-
F	GI9	GI10	GI11	GI12	UI1	UI2	UI3	UI4	UI5	UI6	UI7	UI8	UI9	UI10	UI11	UI12	-
M1	0.12	0.05	0.09	0.33	0.25	0.59	0.73	0.74	0.84	0.76	0.87	0.70	0.49	0.35	0.50	0.61	-
M2	0.52	0.17	0.09	0.10	0.47	0.41	0.11	0.20	0.23	0.63	0.30	0.29	0.35	0.61	0.40	0.66	-
M3	0.33	0.45	0.70	0.30	0.49	0.41	0.49	0.43	0.64	0.67	0.65	0.64	0.65	0.23	0.23	0.20	-

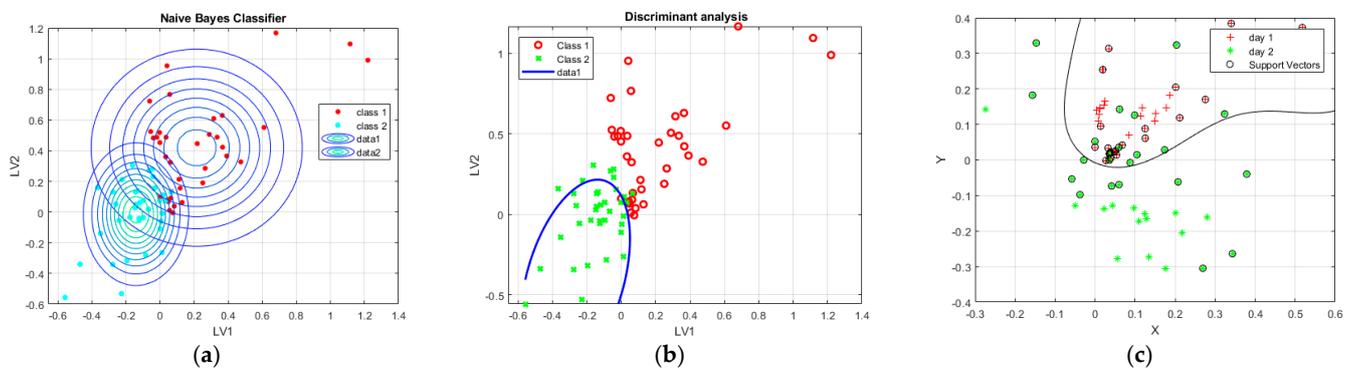
M, manufacturer; F, feature.

A vector of features that are informative for the three manufacturers is selected. It contains six spectral, two gas, and five ultrasonic features. This vector has the following form:

$$FV = [SI1 SI2 SI3 SI5 SI7 SI8 GI1 GI8 UI5 UI6 UI7 UI8 UI12] \tag{51}$$

### 3.5. Classification Results of Cheese from Three Manufacturers

Figure 5 shows examples of classification by latent variables for producers M1 and M2.



**Figure 5.** Classification examples between M1 and M2: (a) Naive Bayesian classifier, (b) discriminant analysis with quadratic function, and (c) SVM with quadratic function.

Table 9 shows the results of Naive Bayes classifier. Values closer to 1 indicate a high error rate, suggesting that a significant proportion of instances from one manufacturer are being misclassified as belonging to another manufacturer or category. For example, if the error rate for “ei” between M1 and M2 is 0.9, it suggests that 90% of instances from M1 are misclassified as “ei” when they actually belong to M2. Values closer to 0 indicate a low error rate, suggesting that fewer instances are being misclassified. For example, if the error rate for “gi” between M2 and M3 is 0.1, it indicates that only 10% of instances from M2 are misclassified as “gi” when they actually belong to M3.

**Table 9.** Results from Naive Bayesian classifier.

Data Reduction Method	Manufacturer Error	M1–M2	M1–M3	M2–M3
PC	$e_i$	0.28	0.17	0.22
	$g_i$	0.16	0.06	0.13
	$e_0$	0.26	0.18	0.19
LV	$e_i$	0.00	0.03	0.17
	$g_i$	0.18	0.10	0.25
	$e_0$	0.13	0.08	0.22

M, manufacturer; PC, principal component; LV, latent variable.

Table 10 shows the results of discriminant analysis. Across comparisons between M1 and M2, M1 and M3, and M2 and M3, varying degrees of misclassification are evident. For linear discriminant analysis, particularly high error rates are observed for class “ $e_i$ ” and “ $g_i$ ”, indicating substantial misclassifications, while error rates for class “ $e_0$ ” are slightly lower but still considerable. Quadratic discriminant analysis generally yields lower error rates for class “ $e_i$ ” but moderate-to-high rates for classes “ $g_i$ ” and “ $e_0$ ”. These findings suggest that, while quadratic discriminant analysis generally performs better for class “ $e_i$ ”, it struggles with distinguishing classes “ $g_i$ ” and “ $e_0$ ” across different manufacturer pairs, highlighting potential challenges in accurately classifying certain categories of data.

**Table 10.** Results from discriminant analysis.

Data Reduction Method		PC			LV		
Separation Function	Manufacturer Error	M1–M2	M1–M3	M2–M3	M1–M2	M1–M3	M2–M3
Linear	$e_i$	0.80	0.80	0.20	0.60	0.60	0.20
	$g_i$	0.80	0.80	0.50	0.67	0.67	0.56
	$e_0$	0.84	0.84	0.76	0.89	0.89	0.77
Quadratic	$e_i$	0.00	0.00	0.27	0.00	0.00	0.00
	$g_i$	0.58	0.58	0.35	0.42	0.42	0.32
	$e_0$	0.22	0.11	0.39	0.07	0.07	0.15

M, manufacturer; PC, principal component; LV, latent variable.

Table 11 plots the SVM results. Across comparisons between M1–M2, M1–M3, and M2–M3, distinct patterns emerge. For linear SVM, moderate error rates are observed across all classes (“ $e_i$ ”, “ $g_i$ ”, and “ $e_0$ ”), with slightly higher rates for class “ $e_0$ ”. Quadratic SVM generally yields lower error rates for classes “ $e_i$ ” and “ $g_i$ ”, as is especially evident in the M1–M3 comparison, but with higher rates for class “ $e_0$ ”. Notably, for the quadratic kernel, error rates for class “ $g_i$ ” are exceptionally low, reaching 0 in certain cases. RBF SVM presents similar trends to the quadratic kernel, with generally low error rates for classes “ $e_i$ ” and “ $g_i$ ”, particularly notable in the M1–M3 comparison, but with slightly higher rates for class “ $e_0$ ”. For latent variables (LVs), linear SVM shows low error rates across all classes and comparisons, while quadratic and RBF SVMs consistently exhibit even lower error rates, emphasizing the effectiveness of SVM, especially with the quadratic kernel, in accurately classifying latent variables.

The classification results show that the data for the three manufacturers overlap. This is evidenced by the relatively high values of classification errors, which, in most cases, exceed 10%. But, in the following analyses, the data from the three manufacturers are processed together.

**Table 11.** Results from SVM.

Separation Function	Data Reduction Method	PC			LV		
	Manufacturer Error	M1–M2	M1–M3	M2–M3	M1–M2	M1–M3	M2–M3
Linear	$e_i$	0.33	0.22	0.25	0.06	0.03	0.08
	$g_i$	0.14	0.10	0.13	0.06	0.08	0.25
	$e_0$	0.46	0.27	0.32	0.08	0.08	0.08
Quadratic	$e_i$	0.31	0.19	0.25	0.06	0.03	0.03
	$g_i$	0.00	0.09	0.04	0.03	0.03	0.03
	$e_0$	0.33	0.22	0.27	0.07	0.07	0.08
RBF	$e_i$	0.31	0.19	0.25	0.06	0.03	0.03
	$g_i$	0.00	0.09	0.04	0.03	0.03	0.03
	$e_0$	0.33	0.22	0.27	0.07	0.07	0.08

M, manufacturer; PC, principal component; LV, latent variable.

### 3.6. Prediction of Basic Characteristics of Cheese

PCR and PLSR methods were used to verify to what extent the optical, gas, and ultrasonic characteristics of cheese could be used to predict its main characteristics that determine its quality. The used criteria, the coefficient of determination, and both types of errors give a sufficient idea of how possible this prediction is. In total, 30% of the data were used in this analysis. Table 12 shows the results of the preliminary analysis of the possibility of predicting the main characteristics of cheese by the PCR and PLSR methods.

**Table 12.** Amendment of cheese main characteristics on PCR and PLSR results.

Data Reduction Method	PCR			PLSR		
Criteria Parameter	SSE	R <sup>2</sup>	RMSE	SSE	R <sup>2</sup>	RMSE
pH	0.37	0.85	0.19	0.15	0.96	0.04
EC	0.79	0.66	8.65	0.12	0.95	0.33
TDS	0.77	0.52	8.49	0.56	0.93	0.23
ORP	0.39	0.58	3.62	0.29	0.89	0.17

The results show that the latent variables obtained by the PLSR method are suitable for predicting the main characteristics of cheese. This is because their use results in low error values and high coefficient-of-determination values compared to the PCR results.

### 3.7. Building Regression Predictive Models for Key Cheese Characteristics

Using two latent variables describing over 95% of the experimental data, regression predictive models were created for key cheese characteristics. The following regression models were obtained:

$$\text{pH} = f(\text{LV}_1, \text{LV}_2) \quad \text{pH} = 6.03 + 7.7\text{LV}_1 - 0.97\text{LV}_2 + 61.17\text{LV}_1^2 + 17.53\text{LV}_1\text{LV}_2 \quad (52)$$

$$\text{EC} = f(\text{LV}_1, \text{LV}_2) \quad \text{EC} = 2106 + 2621\text{LV}_1 - 609\text{LV}_2 + 12126\text{LV}_1^2 \quad (53)$$

$$\text{TDS} = f(\text{LV}_1, \text{LV}_2) \quad \text{TDS} = 442 + 188\text{LV}_1 - 79\text{LV}_2 + 613\text{LV}_1^2 + 104\text{LV}_2^2 \quad (54)$$

$$\text{ORP} = f(\text{LV}_1, \text{LV}_2) \quad \text{ORP} = 201 + 110\text{LV}_1 - 16\text{LV}_2 + 795\text{LV}_1^2 + 157\text{LV}_1\text{LV}_2 \quad (55)$$

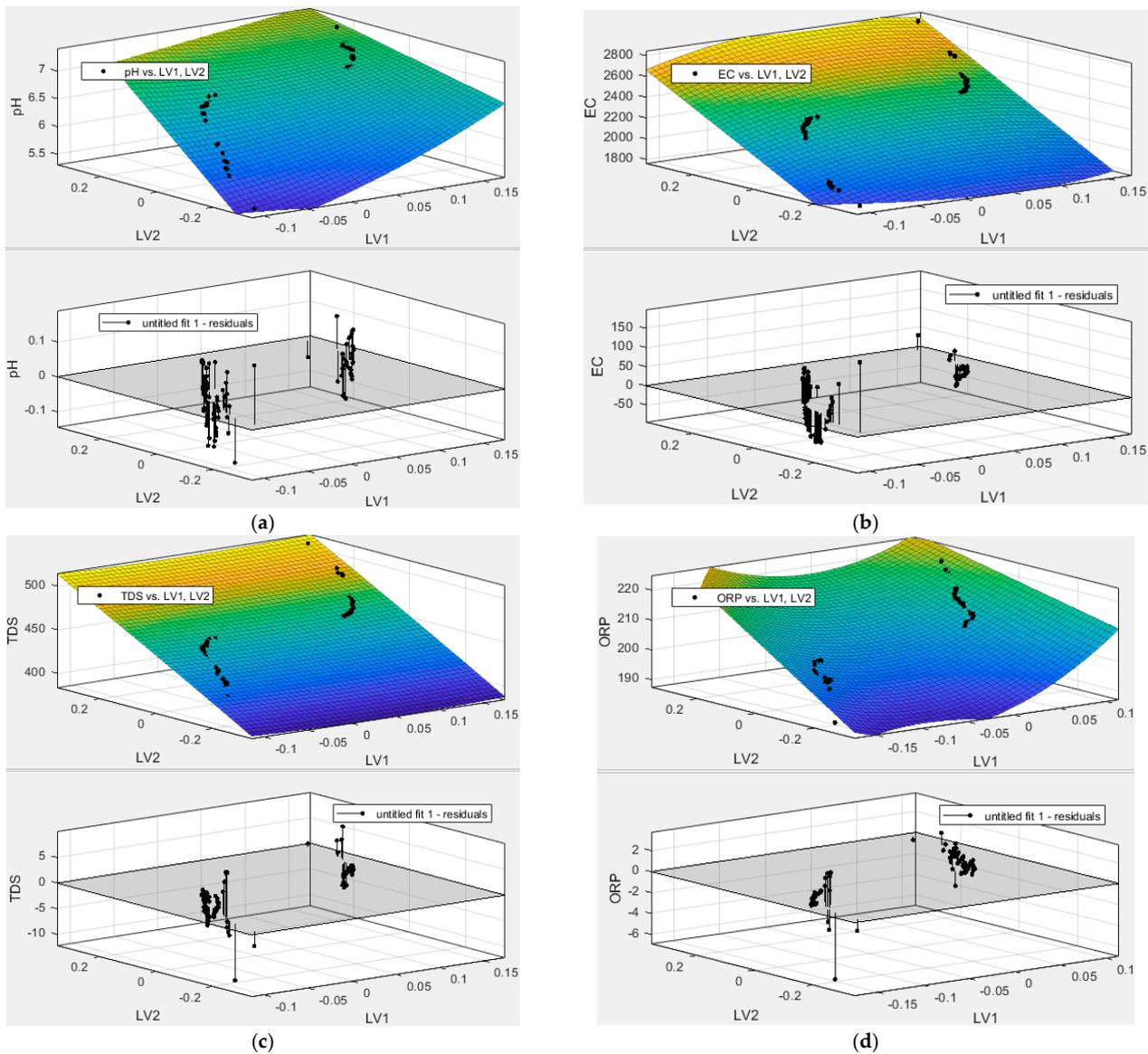
Table 13 shows the values of the criteria for evaluating the regression models. From the table, it can be seen that the conductivity model has the best predictive ability, while the resulting R<sup>2</sup> for TDS indicates the lowest prediction rates compared to all others. According to Fisher's criterion,  $F \gg F_{cr}$ , the significance level is  $p \ll \alpha$ . The standard error

values are the largest for the *TDS* prediction model and the least significant for the pH prediction model.

**Table 13.** Assessment of regression models.

Model	R <sup>2</sup>	F	Fcr	p-Value	SE
pH = f(LV <sub>1</sub> , LV <sub>2</sub> )	0.91	F(4, 103) = 274.85	2.46	0.00	0.12
EC = f(LV <sub>1</sub> , LV <sub>2</sub> )	0.93	F(3, 104) = 484.58	2.69	0.00	0.51
TDS = f(LV <sub>1</sub> , LV <sub>2</sub> )	0.89	F(4, 103) = 181.05	2.46	0.00	0.63
ORP = f(LV <sub>1</sub> , LV <sub>2</sub> )	0.92	F(4, 103) = 304.64	2.46	0.00	0.17

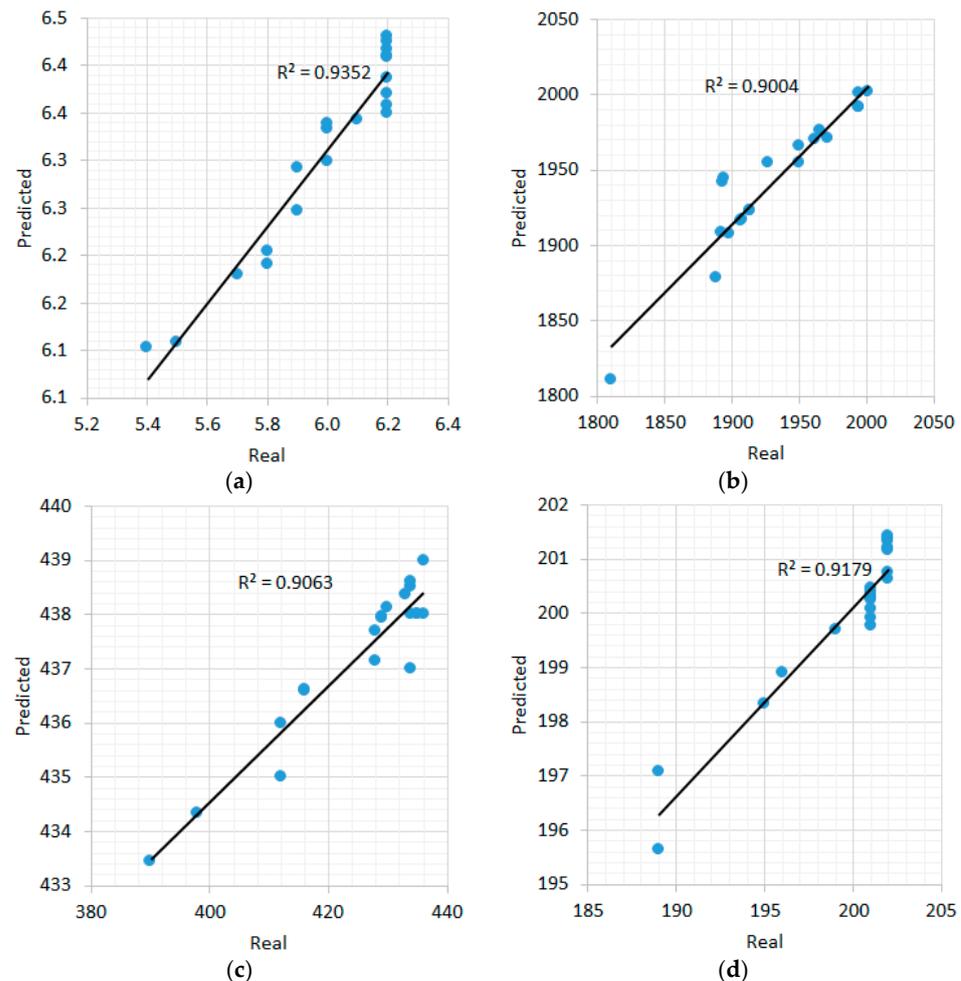
Figure 6 shows an overview of the obtained models and their residuals. From the location of the residuals around the normal probability plane, it can be concluded that the conditions of the regression analysis are met for all four models.



**Figure 6.** Regression models—general view: (a) pH = f(LV<sub>1</sub>, LV<sub>2</sub>), (b) EC = f(LV<sub>1</sub>, LV<sub>2</sub>), (c) TDS = f(LV<sub>1</sub>, LV<sub>2</sub>), and (d) ORP = f(LV<sub>1</sub>, LV<sub>2</sub>).

The validation process involves comparing the predictive ability of regression models under identical measurement conditions.

Figure 7 shows the distribution of the measured and predicted values of the technological quantities relative to the appropriate regression line. The distribution of the data is similar to that of the normal probability plot of the resulting models. When predicting pH, TDS, and ORP, a discrepancy between actual and predicted values occurs at very small and very large values of these characteristics. When predicting the EC of the cheese, a slight scatter of the distribution of the measured and predicted values was observed for the entire range, relative to the regression line.



**Figure 7.** Distribution of actual and predicted cheese data: (a) pH, (b) EC, (c) TDS, and (d) ORP.

The specified data deviations can be considered acceptable. This is proven by the error values obtained.

Table 14 presents the errors obtained when validating the obtained regression models for the main characteristics of cheese. The coefficient of determination has a high value above 0.8, but not a sufficient criterion for the validation of the obtained models. The validation errors obtained are low, with the exception of MSE at ORP (0.71) and SE at pH with a value of 0.59. In the first case, a high value indicates that the predictions made by the model have a larger scatter from the true values than the predictions made by the other models. Higher values in the SE for pH are an indication of greater uncertainty in the model predictions. This may be the result of not fitting the model well enough to the data or the presence of unaccounted-for important factors. It can be concluded that the obtained errors are acceptable, and it follows that the regression models can be used to predict the main characteristics of cheese.

**Table 14.** Errors between actual and predicted values of main cheese characteristics.

Criterion Characteristic	MSE	RMSE	MAE	SE	R <sup>2</sup>
pH	0.12	0.35	0.31	0.59	0.94
EC	0.35	0.19	0.13	0.38	0.90
TDS	0.28	0.17	0.12	0.37	0.91
ORP	0.71	0.27	0.17	0.14	0.92

#### 4. Discussion

Due to the overlap of the optical data from the three types of investigated cheeses, established by the high values of separation errors during classification, general models were created for their separate properties, characterizing the quality indicators of the products. This complements the studies of Bittante et al. [19] regarding the effective use of optical techniques in cheese analysis and also improved on the studies of Eskildsen et al. [20], showing that the main characteristics of the cheese have a significant influence on the changes in the optical characteristics of the product. In this development, this was confirmed by the measurements of the active acidity and electrical conductivity of the product.

Combining the data is a prerequisite for creating an objective model for a given property that captures the general trends and variations in the data for individual cheeses. Using data from three types of sensors, the studies of Sherveglieri et al. [21] used only gas and optical sensors. In this development, by adding data from ultrasonic sensors, the accuracy of predicting the main characteristics of cheese can be increased to over 95%.

Complementing the results of Meza et al. [22], the results obtained in the preliminary prediction of the main characteristics of cheese by chemometric techniques presented significantly higher values of the coefficient of determination and lower errors when using the PLSR method. This can be explained by the fact that PLSR is an approach that effectively deals with the multicollinearity found in the data characterizing cheese quality indicators. This aspect is essential to avoid the problem of “overcomplexity” in models that often occurs when using the PCR method. Therefore, the regression predictive models created are based on latent variable data obtained from spectral, ultrasonic, and gas characteristics. Based on these, the obtained regression predictive models showed predictive ability, as represented by the coefficient of determination R<sup>2</sup>, ranging from 0.89 to 0.93 for TDS and EC, respectively.

The results obtained in this work share some conceptual similarities with biomimetics in terms of data synthesis, model optimization, and the pursuit of efficiency. This fulfills the recommendations of Ju et al. [9] and Falchi et al. [6] regarding the use and processing of data from biomimetics-based sensors and sensor interfaces.

Despite the effectiveness of integrating ultrasonic, spectral, and gas characteristics in assessing cheese quality, there are limitations of independent sensor usage. Data combinations taken from different sensor archives give a broader range of predictive features, which are important in regard to changes in cheese composition and characteristics from different producers. This multimodal approach enhances the precision, reliability, and adaptability of quality assessment models, enabling more informed decision-making in cheese production processes.

#### 5. Conclusions

The integration of ultrasonic, spectral, and gas characteristics to assess cheese quality represents a biomimetics-inspired approach, mimicking nature’s multimodal sensory capabilities and emphasizing adaptability and resilience in decision-making processes. The present research is focused on adapting algorithms and developing instrumentation for a non-destructive, automated quality assessment of white brine cheese based on an integrated approach using ultrasonic, spectral, and gas characteristics. In this way, it complements and improves the approaches and solutions applied up to now in the field of non-contact

cheese analysis for the purpose of its quality. The independent use of characteristics can limit the possibilities for adequate quality of cheese, especially if they are weakly sensitive to changes in composition and hence in its quality. In the study, it was found that the use of ultrasonic, spectral, and gas sensors alone gives a limited number of informative indices for quality analysis of this product which are effective only for a specific cheese producer. Combining these methods provides an opportunity to create a wider range of prognostic features that are sensitive to changes in the composition and characteristics of cheese from different producers. This combination of data also allows greater precision and reliability in the assessment of cheese quality, greater adaptability of the model to different producers and production conditions, and, last but not least, improves the possibilities for appropriate decision-making. The results presented in the work represent an effective and balanced approach for predicting the quality characteristics of cheese, taking into account the many factors that can influence their values. The present study integrated ultrasonic, spectral, and gas characteristics to evaluate cheese quality, but there may be additional data sources or sensory modalities that could provide information to increase the accuracy of the regression models. Exploring and integrating new data sources can improve the accuracy and reliability of cheese quality assessment models. The development of a real-time monitoring system based on the joint use of optical, gas, and ultrasonic techniques could be beneficial to cheese producers. This system can continuously monitor cheese quality during production, allowing for timely corrections and interventions to maintain or improve quality. Also, the study may have applications beyond the evaluation of cheese quality. Adaptation of the methodology to other food products or industrial processes may broaden its impact and utility.

This research underlines the importance of the use of multimodal sensory integration to improve the cheese quality assessment. By integrating ultrasonic, spectral, and gas characteristics, this study provides an effective approach to predict cheese quality among various production ways. Continuous monitoring of systems based on the integrated sensory techniques for real-time quality assessment during cheese production is presented as an important task to be performed by every producer. Moreover, the study suggests the potential application of this methodology not only for cheese evaluation, but also for its implementation on other food products or industrial processes. In this way, it broadens its impact and utility.

**Author Contributions:** Conceptualization, Z.Z. and T.S.; methodology, Z.Z.; software, Z.Z. and M.V.; validation, I.T., M.D. and Z.Z.; formal analysis, M.V.; investigation, Z.Z., I.T. and M.V.; resources—T.S. and M.D.; data curation, Z.Z. and M.V.; writing—original draft preparation, Z.Z.; writing—review and editing—Z.Z., I.T., T.S. and M.V.; visualization—Z.Z.; supervision, Z.Z. and T.S.; project administration, Z.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data is contained within the article.

**Acknowledgments:** This work was administrative and technically supported by the Bulgarian national program “Development of scientific research and innovation at Trakia University in the service of health and sustainable well-being”—BG-RRP-2.004-006-C02.

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

## Abbreviations

AACC American Association of Cereal Chemists  
ANN Artificial Neural Network

BNS	Bulgarian National Standard
EC	electrical conductivity
FFT	Fast Fourier Transformation
FTIR	Fourier-Transformed Infra-Red
GO	Graphene Oxide
ISO	International Organization for Standardization
LDA	linear discriminant analysis
MAE	mean absolute error
MIR	Middle Infra-Red
MLR	Multiple Linear Regression
MO	metal oxide
MSE	Mean Squared Error
NBC	Naive Bayesian classifier
NIR	Near Infra-Red
ORP	oxidation–reduction potential
PCR	principal component regression
PDO	Protected Designation of Origin
PEG	Polyethylene Glycol
pH	active acidity
PLS	partial least squares
PLSR	partial least squares regression
QDA	quadratic discriminant analysis
RBF	Radial Basis Function
RGB	Red, Green, and Blue
RMSE	Root Mean Squared Error
SE	standard error
SPR	Surface Plasmon Resonance
SVM	support vector machines
TCA	Texture Compensation Analysis
TDS	Totally Dissolved Solids
VIS	visible
VOCs	Volatile Organic Compounds

## References

1. Al-Imarah, F.; Al-Kanaan, G.; Jassim, M. Determination of certain chemical elements in cheese as dairy product from local producer southern marshes of Iraq. *Int. J. Appl. Sci. Curr. Future Res. Trends* **2023**, *18*, 152–159.
2. Vasilev, M. Classification of yellow cheese in storage period by nonlinear discriminant analysis and color features. *Innov. Entrep.* **2016**, *4*, 28–37.
3. Dimitrov, E.; Mladenov, M. Complex assessment of the quality of dairy products by analyzing color images and spectral characteristics. *J. ICCST* **2012**, *1*, 60–64.
4. Genet, B.; Molina, G.; Wätjen, A.; Barone, G.; Albersten, K.; Ahrné, L.; Hansen, E.; Bang-Berthelsen, C. Hybrid Cheeses—Supplementation of Cheese with Plant-Based Ingredients for a Tasty, Nutritious and Sustainable Food Transition. *Fermentation* **2023**, *9*, 667. [[CrossRef](#)]
5. Sasmazer, R.; Korukluoglu, M.; Ginoyan, R.; Platova, G. White cheese texture profile at different storage temperatures. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *1052*, 012061. [[CrossRef](#)]
6. Falchi, L.; Pistón, M.; Casarotto, G.; Carro, S.; Cajarville, C.; Iaquina, F. Development of a Green Methodology for the Determination of Artisanal Danbo Cheese Quality Parameters. *Braz. J. Anal. Chem.* **2023**, *10*, 30–41. [[CrossRef](#)]
7. Aguirre-Londoño, J.; Aristizabal-Ferreira, V.A.; Castro-Narváez, S.; Ramírez-Navas, J. Conductimetry: A rapid alternative technique for chlorides determination in cheese. *Univ. Sci. (Pontif. Univ. Javer. Fac. Cienc.)* **2019**, *24*, 307–322. [[CrossRef](#)]
8. Dineva, S.; Veleva-Doneva, P.; Zlatev, Z. Urban Environmental Quality Assessment by Spectral Characteristics of Mulberry (*Morus L.*) Leaves. *Environments* **2021**, *8*, 87. [[CrossRef](#)]
9. Ju, Y.; Pan, J.; Wang, X.; Zhang, H. Detection of *Bursaphelenchus xylophilus* infection in *Pinus massoniana* from hyperspectral data. *Nematology* **2014**, *16*, 1197–1207. [[CrossRef](#)]
10. Mendiguren, G.; Martín, M.; Nieto, H.; Pacheco-Labrador, J.; Jurdao, S. Seasonal variation in vegetation water content estimated from proximal sensing and MODIS time series in a Mediterranean Fluxnet site. *Biogeosciences* **2015**, *12*, 5523–5535. [[CrossRef](#)]

11. Baycheva, S.; Dimov, M.; Zlatev, Z. Classification and prediction of essential oils composition by combined data of MOS gas sensors. *Appl. Res. Technol. Educ.* **2021**, *9*, 94–108.
12. Matz, V.; Kreidl, M.; Šmíd, R. Classification of ultrasonic signals. In Proceedings of the 8th International Conference of the Slovenian Society for Non-Destructive Testing «Application of Contemporary Non-Destructive Testing in Engineering», Portorož, Slovenia, 1–3 September 2005; pp. 27–33.
13. Zhang, X.; Cui, J.; Wang, W.; Lin, C. A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm. *Sensors* **2017**, *17*, 1474. [[CrossRef](#)] [[PubMed](#)]
14. Zlatev, Z.; Stoyanova, V. Influence of Measuring Distance on the Prediction of Quality Characteristics of Cheese by Ultrasonic Sensor. *J. Inform. Innov. Technol.* **2021**, *1*, 35–41.
15. Ilarionov, R.; Simeonov, I.; Kilifarev, H.; Yordanov, S.; Shopov, N.; Ibrishimov, H. *Ultrasonic Device for Non-Contact Studying of Materials International Conference on Computer Systems and Technologies—CompSysTech'10*; Association for Computing Machinery: New York, NY, USA, 2010; pp. 38–45.
16. Pereira, L.; Salazar, A.; Vergara, L. A Comparative Study on Recent Automatic Data Fusion Methods. *Computers* **2024**, *13*, 13. [[CrossRef](#)]
17. Kirilova, E. Synthesis of color character classifiers for the detection of Fusarium Moniliforme molds on maize grains. *Sci. Work. Ruse Univ.* **2012**, *51*, 168–175.
18. Nachev, V.; Titova, T.; Kosturkov, R. Instrumental Data Fusion for Food Analysis Application. *IFAC Pap. Online* **2019**, *52*, 58–63. [[CrossRef](#)]
19. Bittante, G.; Patel, N.; Cecchinato, A.; Berzaghi, P. Invited review: A comprehensive review of visible and near-infrared spectroscopy for predicting the chemical composition of cheese. *J. Dairy Sci.* **2022**, *105*, 1817–1836. [[CrossRef](#)] [[PubMed](#)]
20. Eskildsen, C.; Sanden, K.; Wubshet, S.; Andersen, P.; Øyaas, J.; Wold, J. Estimating dry matter and fat content in blocks of Swiss cheese during production using on-line near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2019**, *27*, 293–301. [[CrossRef](#)]
21. Sherveglieri, V.; Bhandari, M.; Carmona, E.; Betto, G.; Soprani, M.; Malla, R.; Sberveglieri, G. Spectrocolorimetry and nanowire gas sensor device S3 for the analysis of Parmigiano Reggiano cheese ripening. In Proceedings of the ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Montreal, QC, Canada, 28–31 May 2017; pp. 1–3.
22. Meza, L.; Aleman, R.; Marcia, J.; Yadav, A.; Castillo, M. Performance Evaluation of Fiber Near-Infrared (NIR) Optic Probes for Quality Control of Curd Hardness in Cheese Produced by Spray-Dried Milk. *Spectrosc. J.* **2023**, *1*, 152–162. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.