High-precision Discrimination of Maize Silage Based on Olfactory Visualization Technology Integrated with Chemometrics Analysis

Yan Tao, Haiqing Tian,* Kai Zhao, Yang Yu, Li'na Guo, Genhao Liu, and Xin Bai

Rapid, reliable and non-destructive detection of the quality of maize silage is essential to high-efficiency animal husbandry and food safety. In this study, the colorimetric sensor array (CSA) integrated with chemometric methods is innovatively proposed for qualitative discrimination of maize silage. First, 12 color-sensitive dyes were selected to fabricate colorimetric sensor arrays to be used as artificial olfactory sensors for obtaining odor fingerprints of maize silage. Machine vision algorithms were utilized to extract the color features, and principal component analysis was applied to reduce the dimensionality of the obtained data. Finally, the PCA results were input variables to develop different qualitative discrimination models. These models involve support vector machines (SVM), extreme learning machine (ELM), and random forest (RF). The analysis results show the 100% correct identification rate for independent samples. The general results sufficiently reveal that olfactory visualization technology integrated with chemometrics analysis has promising applications for high-precision discrimination of maize silage.

DOI: 10.15376/biores.19.2.3597-3613

Keywords: Maize silage; Olfactory visualization; Colorimetric sensors array; Qualitative discrimination

*Contact information: College of Mechanical and Electrical Engineering, Inner Mongolia Agricultural University, Hohhot, PR China; *Corresponding authors. E-mail: hqtian@126.com*

INTRODUCTION

Maize silage has been one of the most widespread feeds in ruminant diets in most regions worldwide (Garcia-Chavez *et al*. 2022). Since it is productive, nutritious and storable, it is generally utilized as the primary roughage ingredient in total mixed rations (Tharangani *et al*. 2021). However, various factors, such as transportation and storage, can lead to the deterioration of maize silage during production. It not only degrades the nutritional value and production performance of ruminants but even jeopardizes the health of consumers (Khan *et al.* 2015; Borreani *et al.* 2018; Serva *et al*. 2023).

Currently, the major quality evaluation methods for maize silage are the sensory evaluation method and laboratory detection analysis method. The sensory evaluation method is intuitive and convenient, but the evaluation results are highly subjective, which may have certain bias (Tuorila 2015). Laboratory detection analysis method generally relies on the accurate determination of pH, organic acid, and ammoniacal nitrogen, *etc*., in maize silage to evaluate its quality (Ávila and Carvalho 2020), such as gas chromatography-mass spectrometry (GC-MS), Kjeldahl nitrogen determination, and high-performance liquid chromatography (HPLC) (Kayacelebi *et al.* 2014; Dou *et al.* 2022).

Although these methods can provide accurate and reliable results, they have the drawbacks of being destructive, requiring a long time, and high in cost. In addition, the electronic nose technique, although it can detect the odor components of the object to be measured and provide a new direction for silage corn feed quality assessment, the electronic nose instrument is expensive and the technique mainly relies on the weak van der Waals force to capture the odor, which is susceptible to environmental factors such as temperature and humidity, greatly affecting the detection accuracy(Wojnowski et al. 2017). Therefore, it is essential to develop a fast, reliable and non-destructive method for silage quality assessment to materialize efficient animal husbandry and food security.

The olfactory visualization technology is a novel electronic nose technology established by mimicking the mammalian olfactory system. This technology is implemented through the cross-reaction between the gas-sensitive colorimetric dyes on the colorimetric sensor array (CSA) and the volatile organic compounds of the substance for measurement. The sensor array image after the reaction carries the information of the analyte, and qualitative and quantitative analyses of the analyte can be performed in accordance with the color change of the colorimetric dyes before and after the reaction (Janzen *et al.* 2006). Each colorimetric dye can react with different volatile gases. The color change caused by the same amount of different volatile gases and different amounts of the same volatile gases are not the same. Thus, it is possible to distinguish and detect a variety of chemical substances (Suslick *et al.* 2004). Metalloporphyrin and pH indicator are the commonest gas-sensitive colorimetric dyes, but they have different color-changing principles. Metalloporphyrin has an open axially connected ligand site, and when it is bound to the ligand, the spectrum is shifted, showing a strong coloring effect, whereas the color of the pH indicator varies only with the change of acidity and alkalinity. The performance of colorimetric sensor array (CSA) depends on the cross response of the sensor elements and the discriminative ability of the fingerprint model, which is independent of the traditional lock-and-key model of specific receptor-analyte interactions. Compared with the traditional electronic nose technology, the olfactory visualization technology has the advantages of high selectivity, high sensitivity, and low detection limit, which is more suitable for the production requirements of on-line and large-scale detection in modern animal husbandry industry. Nowadays, the technology has been utilized in numerous complicated scenarios, such as food quality assessment, plant disease diagnosis, and toxic gas detection (Jiang *et al.* 2019; Xu *et al.* 2019, 2022a,b; Liu *et al.* 2020; Wang *et al.* 2021; Ouyang *et al.* 2023). However, little research has been done on the application of this technology towards maize silage quality detection.

This study aimed to verify the feasibility of applying olfactory visualization sensor technology for qualitative discrimination of maize silage. The primary work arrangements of this study were as follows. (1) The colorimetric sensor array was fabricated by spot sampling 12 color-sensitive dyes selected from the pre-experiment on the substrate material. (2) The post-reaction odor fingerprints were acquired and computer vision algorithms were used to extract the 108 variables from the different images. (3) PCA was used to extract feature variables by compression and visualize the spatial distribution of different quality maize silage samples. (4) Three non-linear chemometric methods were utilized to develop the qualitative discrimination model of maize silage, and independent datasets were used to validate the optimized models.

EXPERIMENTAL

Sample Preparation

Maize silage (Yandan Silage 202) was purchased from a local farm (Hohhot, Inner Mongolia, China) in three batches of 126 kg. Each batch was subjected to testing immediately upon receipt to ensure the quality of samples. The maize silage samples were placed in polyethylene boxes of 3 kg ± 100 g each at ambient temperature (18 to 22°C). Every two boxes of maize silage were taken as a group, and three sampling points were taken from each group for the olfactory visualization experiment, which resulted in 48 samples per batch, and a total of 144 samples were prepared.

During storage and feeding of silage corn feeds, it is possible that there may be secondary fermentation caused by periodic opening of silo bins (intermittent invasion of oxygen) or permanent destruction of the storage environment (continuous invasion of oxygen), so that two different treatments of laboratory-prepared maize silage samples were carried out on the collected samples, as follows:

(1) Continuous aerobic exposure treatment: To prevent water evaporation, the polyethylene boxes were covered with aluminum foil with small holes of 10 mm diameter distributed equally. During the seven days of storage, aerobic exposures of 0D, 1D, 2D, 3D, 4D, 5D, 6D, and 7D were undertaken, which were performed at ambient temperature (18-22°C).

(2) Intermittent aerobic exposure treatment: In comparison to the previous treatment, the only difference of this method is that the maize silage was placed in a polyethylene box covered with perforated aluminum foil for one hour of aerobic exposure and kept closed for the rest of the time.

Given the exploration of the deterioration pattern of corn silage by the pre-experiment, the first two batches of samples were subject to continuous aerobic exposure treatment, and the third batch of samples was subjected to intermittent aerobic exposure treatment, resulting in relatively balanced samples of each maize silage grade.

Quality Grade Assessment

The pH value is one of the key indicators for assessing maize silage quality, reflecting the acidic and alkaline nature of the maize silage. An appropriate pH range serves to maintain its stability and promote the fermentation activities of beneficial microorganisms, which influence the feeding rate and enzyme activities of animals (Kung *et al*. 2018). Compared to other indicators, pH has the advantages of the visibility and rapidity of the measurement, thus providing a simple and effective means of monitoring for stocking management. According to the Technical Specification for Grading Whole Plant Maize Silage (DB11/T 1759-2020), among the fresh maize silage with pH $\leq$ 3.9, the medium maize silage with pH ranges from 3.9 to 4.4, and the deteriorated maize silage with pH > 4.4.

The pH of maize silage was measured using a high-precision pH meter (PHS-3CB, Shanghai Yueping Scientific Instrument Co., Ltd., Shanghai, China) with a resolution of 0.01 and an accuracy of ±0.01. Maize silage in the polyethylene box was evenly divided into three parts, and the silage from each part was mixed well and reduced using the quadrature method until the sample was reduced to 10.0 g (accurate to 0.1). The weighed maize silage was placed in a 500 mL glass beaker, then 90 mL of distilled water was added (10 g of sample/90 mL of distilled water was the common ratio). Later, the glass beaker was covered with cling film. After sufficient immersion in distilled water for 0.5 h, the maize silage extract was prepared by filtering through 4 layers of gauze. Finally, the pH value of the extracts was measured using the pH meter. Each measurement was repeated in three trials. The average of the three measurements was adopted as the final pH value. This experimental procedure was performed at ambient temperature (18 to 22 °C).

Fabrication of Colorimetric Sensors Array

For acquiring the volatile odor fingerprint information of the different qualified maize silage feeds, this study used an olfactory visualization sensor system developed based on colorimetric sensors to capture the volatile odor of the maize silage samples. Color-sensitive materials used to fabricate colorimetric sensor arrays are essential to the components of the olfactory visualization sensor system. Thus, the selected color-sensitive materials must meet the following requirements: (1) The color-sensitive material must have a functional group that reacts chemically and strongly with the measured substance, allowing for strong mutual interactions; (2) A significant color variation must occur when the color-sensitive material responds to the measured substance; (3) Cross-sensitivity among color-sensitive materials, with overall responsiveness to the measured substance, which can form odor characteristic fingerprint and facilitate qualitative and quantitative detection.

In accordance with the above requirements, after initial experiments, the volatile gas components of maize silage were detected by solid-phase microextraction (SPME) and gas chromatography-mass spectrometry (GC-MS) during the secondary fermentation of maize silage, including acids, esters, phenols, terpenes, alkanes, heterocyclic, aldehydes, alcohols, *etc.*, with phenols and acids showing the most significant changes in the fermentation process. To precisely capture these key odor molecules, metalloporphyrin complexes containing different active metal ions (*e.g.*, zinc, copper, iron, and cobalt) were selected for the design of this assay, and eight kinds of porphyrin (Sigma-Aldrich, USA) were screened and identified for the constructed colorimetric sensor arrays. Moreover, four pH indicators (Sino-pharm, Shanghai, China) with different chromogenic ranges were introduced to complement the sensor array to improve the sensitivity and broad-spectrum of detection as weakly acidic and weakly basic volatiles may be present in the forage. Table 1 lists 12 color-sensitive dyes employed to fabricate olfactory sensor arrays. For preserving color information accuracy, the color-sensitive material needs to be placed on a pure white substrate. Moreover, the substrate material is hydrophobic for protection against ambient humidity. Therefore, a C2 reverse-phase silica gel plate was selected as the substrate material.

The detailed steps for making sensors were as follows: (1) cut the 4 × 4 cm rectangular substrates out of C2 reversed-phase silica gel plates; (2) dissolve 10 mg of porphyrin fully in 5 mL of dichloromethane and dissolve 10 mg of pH reagent fully in 5 mL of anhydrous ethanol; (3) close the prepared solutions hermetically and sonicated for 15 min to achieve the 2 mg/mL solution, and then preserved in the dark; (4) use 100 × 0.3 mm microcapillary to extract approximate 1 μL dissolved solution, then immobilize it on a C2 reverse-phase silica gel plate, fabricating the 4 × 3 colorimetric sensor array. The sensors were comprised of well-arranged 4 × 3 sensor elements, each of which was sized at ϕ3 mm, and (4) the constructed sensor arrays were stabilized in a fume hood for 30 min and then stored individually in plastic packing bags for further experiments.

Table 1. Twelve Color-sensitive Dyes on the Colorimetric Sensor Array

Number	Name
1	2,3,7,8,12,13,17,18-Octaethyl-21h,23h-porphine
2	5,10,15,20-Tetrakis (4-(methoxycarbonyl)phenyl)porphirinato]iron (III) chloride
3	5,10,15,20-Tetrakis (4-methoxyphenyl)-21,22-dihydroporphyrin
4	5,10,15,20-Tetrakis (4-methoxyphenyl)-21h,23h-porphyrine cobalt(ii)
5	5,10,15,20-Tetraphenyl-21h,23h-porphyrin
6	5,10,15,20-Tetraphenyl-21,22-dihydroporphyrin zinc
7	5,10,15,20-Tetraphenyl-21h,23h-porphine copper
8	5,10,15,20-Tetraphenyl-21h,23h-porphine iron (iii) chloride
9	Bromocresol Green
10	Bromothymol Blue
11	Neutral Red
12	Cresol Red

Acquisition of Sensor Array Images and Pre-Processing

Figure 1 shows the data acquisition and processing of the olfactory visualization sensor system. The pre-reaction image of the sensor array was scanned using an HP Scanjet 2600f1 flatbed scanner at a resolution of 1200 dpi. In this study, 50 ± 0.1 g of maize silage feed samples were placed in a beaker with a diameter of 60 mm. The colorimetric sensor array was secured to the back side of the plastic wrap while the front side was facing the maize silage samples. The beaker was hermetically closed with plastic wrap, and the reaction was carried out at ambient temperature for 25 min. Then, the sensor array was scanned again after the response to acquire the post-reaction image data.

The sensor image pre-processing procedures were as follows. (1) The colorimetric sensor array image was sequentially preprocessed with median filtering, binarization, and morphological processing by OpenCV to acquire the center of each sensor element, and the region of interest of each sensor element was segmented by combining the flood-fill algorithm. (2) The next step was to calculate the average gray scale value for the components of the sensor element at three color spaces (RGB, HSV, Lab). (3) The mean values of the grey scale of the reacted and original preprocessed images were subtracted to gain their color differences values. Each color-sensitive point contains 9 variables: three components in RGB color space ($\triangle R$, $\triangle G$, and $\triangle B$), three components in HSV color space ($\triangle H$, $\triangle S$, and $\triangle V$), and three components in CIELab color space ($\triangle L$, $\triangle a$, and $\triangle b$), for a total of 108 variables from 12 dyes in an array (12 dyes × 9 color components).

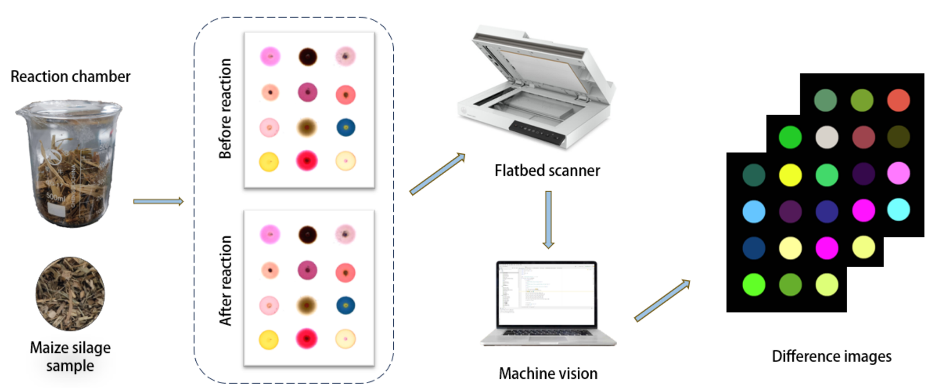

Fig. 1. Data acquisition and processing process of the olfactory visualization sensor system for the determination of maize silage quality

Multivariate Data Analysis

Principal component analysis (PCA) is an unsupervised algorithm that is widely used for dimensionality elimination and initial classification of multivariate data (Jolliffe and Cadima 2016). The fundamental principle is to apply fewer eigenvectors as a means to explain most of the variation of the original data. By analyzing and calculating the internal structural relationships of the correlation matrix of the original data, a series of new eigenvectors that are not related to each other is generated. These principal components are ordered by the magnitude of the variance of the original data. Thus, the preceding principal components contain the most information about the variance in the original data, while the latter principal components contain less information about the variance. In this study, PCA was utilized to further data compression and mining of original features (108 variables) and to visualize the distribution of silage maize samples with different quality grades (Gallo *et al*. 2016).

Extreme Learning Machine (ELM) is a feed-forward neural network algorithm for real-time tasks with excellent learning efficiency and superior generalization performance (Guo *et al*. 2019) Its fundamental principle is to improve the efficiency of model training by randomly generating the connection weights ω and threshold β1 between the input layer and the hidden layer in order to improve the efficiency of the model training. Since it eliminates adjusting the weights during the training process, it can significantly reduce the training time. Compared to the traditional optimality theory, ELM has the unique characteristic that it avoids the process of iteratively adjusting the weights to seek the optimal solution during the training process. Instead, it only requires setting the output matrix of the hidden layer K to acquire the unique optimal solution. The simplified training approach allows the ELM to concentrate only on calculating the output weights during the learning process, which enables more efficient acquisition and adaptation to the data. In this study, the optimum parameter combination of K and PCs in the ELM model was determined by the classification accuracy of the training set by applying a five-fold cross-validation strategy and grid search method. Their value ranged from 10 to 100 (the interval is 10) and 1 to 22 (the interval is 1), respectively.

Support vector machine (SVM) is a supervised learning algorithm for classification and regression with excellent generalization performance and robustness (Bouboulis *et al*. 2015). Its fundamental principle is to train the model by finding the optimal hyperplane of the data and obtaining the maximum interval at the support vectors on the hyperplane for efficient classification and regression of data. The optimization objective of SVM is to find a hyperplane with a maximum interval, which means that it is more cautious in choosing the decision boundary. Therefore, it is less susceptible to the influence of local minima and more adaptable to higher dimensional spaces and nonlinear relationships. In this study, the combination of support vector machine and radial basis kernel function (RBF) was utilized to assess different qualities of maize silage (Virmani and Pandey 2023). For selecting the optimum penalty factor c and kernel function parameter g, it was optimized by using five-fold cross-validation and grid search methods while varying the number of input principal components, and the best PCs and parameters c, g were selected depending on the maximum discrimination rate in the training set. The number of PCs ranging from 1 to 22 (the interval is 1), accompanied by c and g, take values from 2^{-10}, $2^{-9.5}$, …, 2^{-10}, $2^{-9.5}$.

Random Forest (RF) is an integrated learning method consisting of multiple decision trees, which is particularly appropriate for dealing with high-dimensional data and problems with complex relationships (More and Rana 2022). Its fundamental principle is to randomly select a part of the samples to construct a decision tree by bootstrap sampling.

Each decision tree characterizes and predicts the data, and the classification result is finally achieved by voting on each decision tree. Therefore, the number of decision trees M is crucial to the RF model performance. The voting mechanism of random forests reduces the influence of individual trees and contributes to reducing the risk of overfitting. In addition, the voting mechanism of random forests reduces the influence of individual trees and contributes to reducing the risk of overfitting. Accordingly, RF usually has the advantages of high precision, excellent stability, and superior overfitting resistance. In this study, the optimum parameter combination of M and PCs in the RF model was determined by the classification accuracy of the training set by applying a five-fold cross-validation strategy and grid search method. Their value ranging from 50 to 1000 (the interval is 50) and 1 to 22(the interval is 1), respectively.

Model Evolution Criteria

The confusion matrix can intuitively assess the qualitative discrimination performance of the target model (Ouyang *et al*. 2023). Accuracy, precision, and sensitivity are three critical indicators in the assessment matrix. Accuracy is an overall indicator that reflects the proportion of samples correctly categorized by the model. High accuracy denotes that the model has excellent classification ability. The accuracy is determined by Eq. 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where TP refers to the number of positive category samples that the model correctly predicts as positive categories; TN refers to the number of negative category samples that the model correctly predicts as negative categories; FP refers to the number of negative category samples that the model incorrectly predicts as positive categories; and FN refers to the number of positive category samples that the model correctly predicts as negative categories. This study used the test set to construct confusion matrices for each model to determine the classification effect.

Software

The data analysis and visualization were implemented using OriginPro 2018 (OriginLab Corporation, Northampton, USA). Extraction of color feature differences were implemented in Python 3.10 (JetBrains, Prague, Czech Republic), and the chemometric methods, including PCA, ELM, SVM, and RF, were implemented in MatlabR2021a (Mathworks, Natick, USA). All above data processing were run on Windows 11.

RESULTS AND DISCUSSION

Maize Silage Grades and Samples Division

Figure 2 demonstrates the changes in pH over time of maize silage under aerobic exposure conditions at ambient temperature. As the duration of aerobic exposure increased, the pH value of the continuous aerobic exposure treatment group and the intermittent aerobic exposure treatment group both exhibited an increasing tendency. The continuous aerobic exposure treatment group experienced a significant increase in pH value compared to the moderate rise of the intermittent aerobic exposure treatment group. During continuous aerobic exposure, the pH value remained stable during the initial phase and increased slightly on the second day. This indicates that the feed begins to oxidize and ferment on the second day as it begins to deteriorate. Starting with the third day, the temperature of the maize silage samples increased slightly, accompanied by a significant increase in pH, from 4.40 to 6.60. With time, the temperature rose significantly, up to a maximum of 48 °C. Simultaneously, the odor of the maize silage sample changed from a slightly sweet or fruity odor in the initial period of aerobic exposure to a putrid, fishy, or decaying odor, and the texture turned stickier (Karnatam *et al*. 2023). This is considered as a consequence of *Clostridium* contamination. *Clostridium* is a genus of gas-producing anaerobic bacteria that can survive in an anoxic environment and produce gas by decomposing organic matter in the feed, resulting in a change in the odor of the feed (Zhang *et al*. 2018). From the fifth to the seventh day, the molds started to grow and the pH steadily increased, fluctuating between 6.70 and 7.60. At this time, maize silage had deteriorated entirely and could not be utilized for ruminant feeding. By contrast, the pH of the intermittent aerobic exposure treatment group gradually rose from 3.66 to 4.28 during the seven days. The quality of maize silage was still within acceptable grade. The results indicated that maize silage subjected to continuous aerobic exposure conditions was more susceptible to deterioration. Changes in pH values mainly originated from changes in lactic acid content. During the initial fermentation stages of maize silage, lactic acid was the primary fermentation product, keeping the feed in an acidic environment and helping to inhibit the reproduction of harmful microorganisms. However, the oxidation reaction converts lactic acid to acetic acid, increasing the pH of maize silage. The changes lead to the loss of a favorable acidic environment for the forage, which further fosters the growth of unfavorable microorganisms and leads to the ultimate deterioration of the maize silage (Wilkinson and Davies 2013).

The pH of samples ranged from 3.62 to 7.60. The quality of maize silage forage can be divided into three stages under continuous aerobic exposure treatment: fresh phase for 0 to 1 days, medium phase for the second day, and deteriorated phase for 3 to 7 days. As for the intermittent aerobic exposure treatments, it is divided into two stages: fresh phase for 0 to 3 days and medium phase for 4 to 7 days. Therefore, 48 samples of fresh, 36 samples of medium, and 60 samples of deterioration were collected. Subsequently, the Kennard-Stone (KS) method was utilized to classify the 144 samples. Three quarters (108) of the samples were used as the training set to develop the qualitative discrimination model, and the remaining quarter (36) formed the prediction set for validating the performance of the developed model (Wu *et al*. 2015).

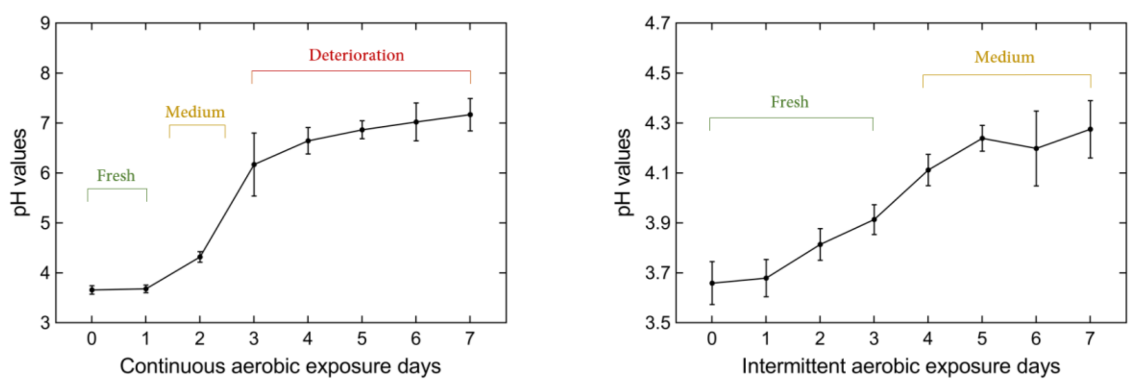

Fig. 2. Changes in pH of maize silage under different aerobic exposure treatments during 7 days. Error bars represent ± standard deviation.

Response Results of Colorimetric Sensor Array

Figure 3 shows the images of the characteristic differences among the maize silage samples of different qualities in different color spaces. The study captured color variations by integrating these three color spaces to deeply analyze and differentiate the quality levels of maize silage samples. RGB space focuses on the spectral components of colors, HSV emphasizes on hue, saturation, and luminance, and Lab space, based on human vision's perception of color differences, which is more sensitive to color contrasts and luminance variations. Each of these color spaces contributes a deep understanding of color variation from a unique perspective, jointly depicting images that accurately map the characteristics of different qualities of maize silage. Twelve color-sensitive points present color variations, and different quality grades of maize silage have their unique images, which can be distinguished by the naked eye. This is primarily attributed to the different composition and content of volatile components in different quality maize silage samples. The colorimetric sensors array is sensitive to the volatile organic compounds in deteriorated maize silage, which results in the volatile components reacting differently with the individual color-sensitive points on the sensor array, thus showing different color changes. Therefore, the combination of 108 color feature variables from different images with appropriate chemometric methods has the potential to differentiate maize silage quality.

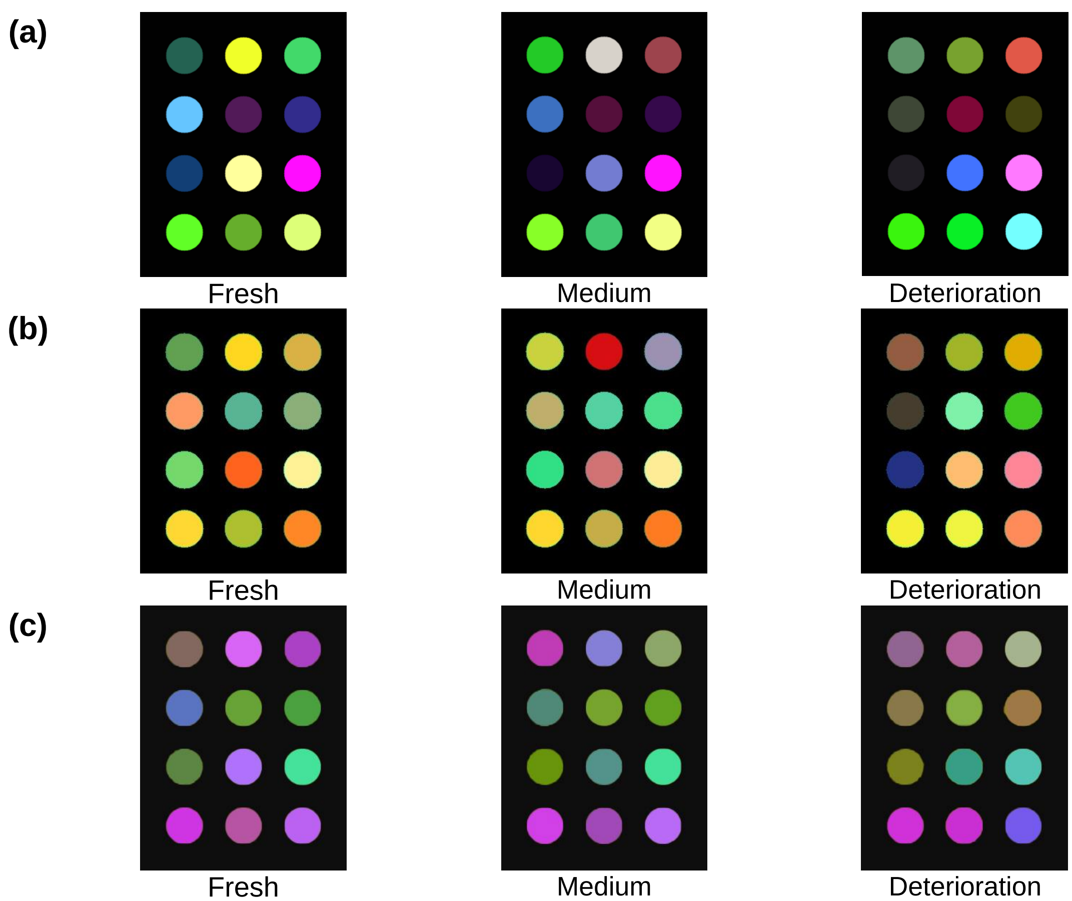

Fig. 3. The color differences image of sensor array with different grades of maize silage samples during 7 days. (a) RGB, (b) HSV, (c) Lab

Results of PCA

Figure 3 demonstrates that the responses of the sensor elements between adjacent aerobic exposures were relatively similar, which could be attributed to the cross-response of some of the sensor elements to the same VOCs or the same response results for different VOCs. In addition, this study selected three color space color feature differences as odor fingerprint information, and the feature difference images varied in different color spaces. All the situations mentioned above may lead to the redundancy of information between the variables extracted from the olfactory sensors. Thus, PCA was used to extract the dominant feature variables from the 108 variables. Figure 4 depicts the cumulative variance contribution of the top 22 PCs obtained after PCA processing of sensor data. The figure demonstrates that the first three principal components contributed to the original feature data of 20.5%, 10.8% and 9.5%, respectively. Moreover, the first 22 principal components contributed to a cumulative variance contribution sum of 90.01%, far exceeding the standard of 85%, which explains most of the valid information in the original feature data.

PCA produced satisfactory results in data reconstruction and extraction of valuable information. However, from Fig. 5, given the insufficient variance contribution and overlap of the principal components, the top three principal components were not appropriate for recognizing the different qualities of maize silage. Therefore, the top 22 PCs were selected as input feature variables. On that basis, three nonlinear ELM, SVM, and RF qualitative discriminant models were developed for qualitative discrimination of maize silage.

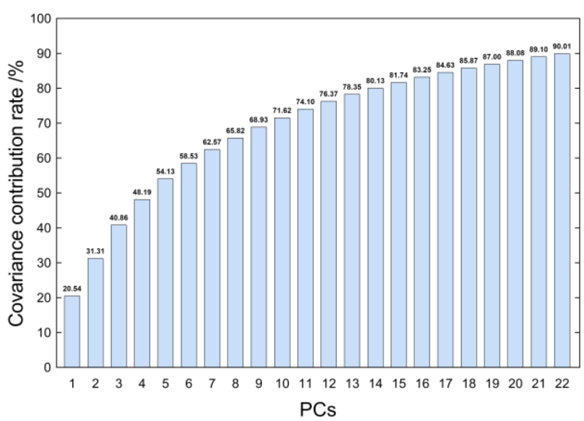

Fig. 4. The cumulative variance contribution of the top 22 PCs obtained by PCA processing of sensor data

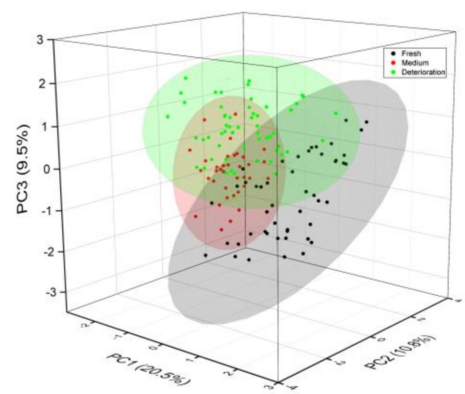

Fig. 5. Three-dimensional principal component scores of maize silage samples with different grades

Results of ELM Model

Because the connection weights ω and thresholds β1 from the input layer to the hidden layer are generated randomly, the output matrix of the hidden layer, K, is required to be set in advance. Thus, K dictates the complexity of the ELM model. Reasonable K can not only maintain the model learning ability but also avoid overfitting and improving the generalization performance. Figure 6 depicts the discrimination rating of the different ELM classification models with different principal component PCs and K values. With the increase of the K and PCs, the discrimination rating of the ELM dramatically improved, then stabilized range from 88% to 100%. However, the model discriminant gradually decreased as the K-value and principal component PCs increased. This is attributed to the overfitting of the training data as a result of increasing the number of K values and PCs, which makes the model more complicated. The model was over-adapted to the noise and details in the training set, all of which cannot be generalized to novel data, which leads to performance degradation on the test set.

According to the results of the five-fold cross-validation and grid search, there were 12 parameter combinations with a 100% discrimination rating, in which the K are all in the range of 10 to 50. Among them, four groups had a K value of 20, at which point the model discrimination performance was much better than the others. When K = 20 and PCs = 10, the ELM model gave the best results for prediction sets. The optimized model achieved an identification percentage of 94.44% in the prediction set, with 2 of the 36 samples misclassified. The detailed classification information of the prediction set is further exhibited in Fig. 8(a). The model correctly classified 91.67%, 100%, and 93.33% samples into corresponding categories, respectively. Herein, there were fresh and deteriorated samples, each misidentified as medium and misclassified samples were placed in the neighboring grades. Thus, ELM exhibited a good ability to discriminate the quality of maize silage.

Table 2. Discrimination Rates of SVM Models with Different PCs and Best Parameter c, g

PCs	Parameter Optimization		Discrimination Rate (%)	
	c	g	Training set	Prediction set
1	67.5926	181.0293	64.81	66.6667
2	8.7220	0.3536	77.78	77.78
3	0.7171	0.5000	97.22	83.33
4	2.8384	2.8284	98.15	88.89
5	1.0000	0.7071	97.22	88.89
6	8.7220	0.7071	97.22	94.44
7	1.4142	1.0000	97.22	91.67
8	0.3536	0.5000	99.07	97.22
9	0.5000	0.2500	99.07	94.44
10	0.7071	0.2500	100.00	100.00
11	1.4142	0.0884	99.07	97.22
12	2.8384	0.2500	100.00	100.00
13	0.5000	0.2500	99.07	97.22
14	1.0000	0.2500	98.15	94.44
15	1.0000	0.1250	99.07	94.44
16	1.4142	0.1250	100.00	97.22
17	2.0000	0.0625	97.22	97.22
18	1.0000	0.0442	99.07	97.22
19	0.5000	0.1768	99.07	100.00
20	4.0000	0.1250	100.00	97.22
21	2.8384	0.0884	100.00	97.22
22	2.8384	0.1250	99.07	94.44

PCs: number of principal components; c: the penalty coefficient of SVM; g: kernel parameter of radial basis function (RBF)

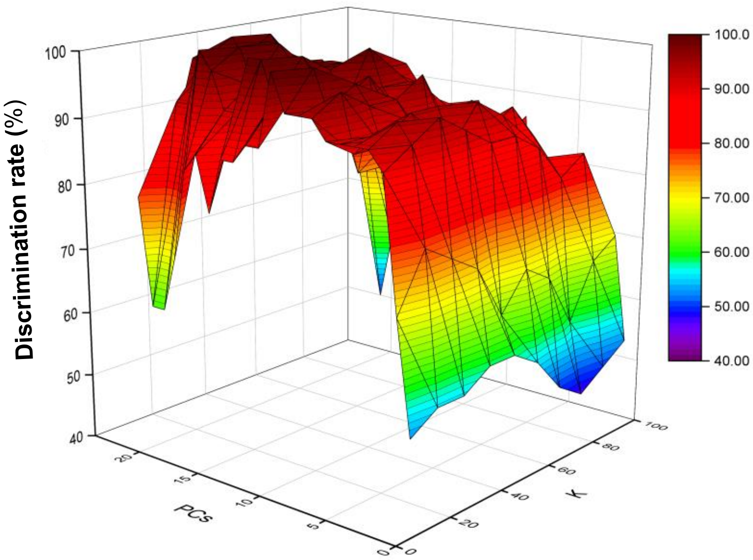

Fig. 6. Discrimination rate of ELM models with different K values and PCs in the training set

Results of SVM Model

The parameters c, g, and PCs are considered to be the most critical parameters in SVM, which significantly influence the performance. Table 2 demonstrates the results of simultaneously optimizing the SVM with three parameters, utilizing the five-fold cross-validation and grid search strategy. The performance of SVM increased gradually with increasing PCs from 1 to 10. When c= 0.7071, g = 0.2500, and PCs = 10, the SVM obtained the best discrimination percentage of 100.00%. However, with the increasing of PCs, the performance of the SVM showed a certain fluctuation, at discrimination percentages of 98 to 100. This phenomenon is caused by the introduction of too many PCs into the model, leading to interference and invalid information for the model. The 36 independent samples in the prediction set were imported into the optimized model, and a superior result was acquired with a discrimination of 100%. The confusion matrix for the prediction set is illustrated in Fig. 8(b), which achieved clear discrimination among the three quality grades. Therefore, the olfactory visualization technique integrated with the SVM model demonstrated an excellent application for the qualitative identification of maize silage.

Results of RF Model

Generally, the higher number of decision trees in RF, the better the performance of the model, but it can lead to an increase in computational cost as well. This study set the initial M as 50 and increased it at intervals from 50 to 1000. Figure 7 depicts the discriminant rate of different RF classification model for different PCs and M values. The superb result of the RF in the training model was 99.50%. The discrimination fluctuated with the increase of M and PCs but was remarkably stable compared to ELM, with fluctuations ranging from 98% to 100%. The parameter-optimized model was applied to the independent 36 samples in the prediction set, and a satisfactory recognition percentage of 97.22% was obtained. Figure 8(c) presents the confusion matrix of the RF about the identification samples in the prediction set. There were superior sensitivity scores of 91.67%, 100%, and 100% for fresh, medium, and deteriorated grades, merely one fresh sample was misidentified as the medium. Therefore, there is also trustworthiness for applying RF to the qualitative identification of maize silage.

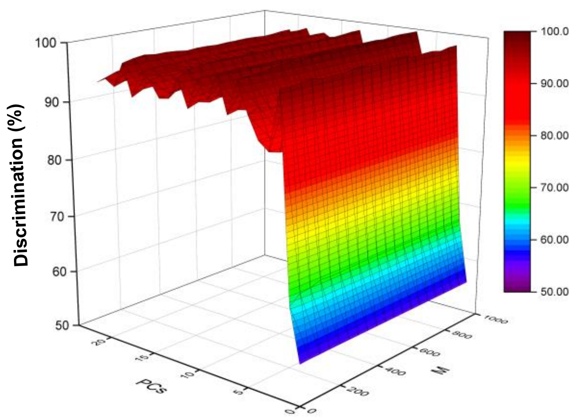

Fig. 7. Discrimination rate of RF models with different M values and PCs in the training set

Comparison of Different Classification Models Performance

Olfactory visualization sensors and chemometric methods were used to characterize different quality grades of maize silage. The olfactory sensor platform was constructed using a 4 × 3 multichannel with 108 variables. Moreover, the three qualitative models and optimization strategies were employed to establish the classification models. The discrimination results of ELM, SVM, and RF modes in recognizing maize silage quality are demonstrated in Table 3. Among the three models, SVM achieved the 100% correct identification rate for independent samples, which is considered to be the most desirable model. The reason that the olfactory visualization system combined with the SVM classification model can achieve such perfect recognition results can be summarized as follows.

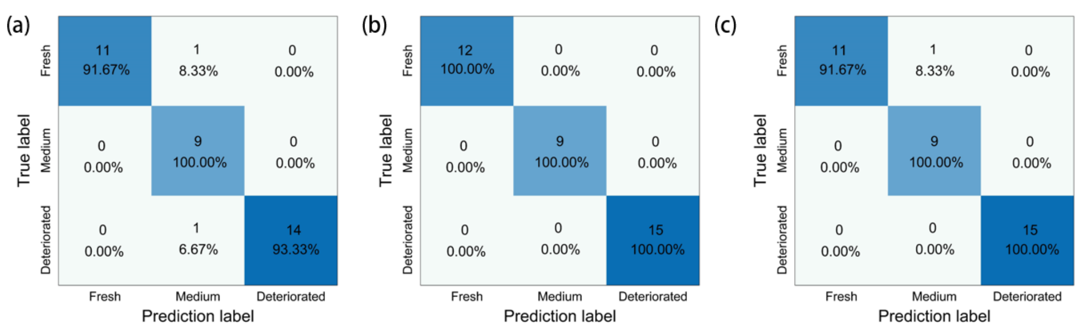

Fig. 8. The confusion matrixes of discrimination rates for three qualitative models in the predicting set. (a) ELM; (b) SVM; (c) RF.

First, concerning aspects of olfactory sensor systems, during aerobic exposure to maize silage, organic substances in the maize silage are oxidized, accompanied by abnormal activity of microbial metabolism, and these variations contribute to different levels of volatile organic compounds in different qualities of maize silage. Although these small changes are difficult to recognize by humans, the differences can be captured by olfactory sensors fabricated from this study. Appropriate porphyrins and pH indicators have strong interactions with various VOCs and have a broad range of responses to different concentrations of VOCs. These allowed the olfactory visualization sensors to recognize differences in odor among different maize silage qualities. In addition, the 4 × 3 sensor array was able to acquire VOCs information from 12 channels, which made it possible to identify and differentiate different odors and reduce the impact of cross-response. Therefore, the colorimetric sensors array demonstrated a unique advantage for detecting maize silage.

Table 3. Results and Comparison of the Best Extreme Learning Machine (ELM) Model, Support Vector Machine (SVM) and Random Forest (RF) Model

Model	PCs	Other parameters	Discrimination rate (%)	
			Training set (%)	Prediction set (%)
ELM	10	K = 20	100.00	94.44
SVM	10	c = 0.7071, g = 0.2500	100.00	100.00
RF	7	M = 600	99.50	97.22

PCs: number of principal components; K: number of neurons of the hidden layer in the ELM; c: the penalty coefficient of SVM; g: kernel parameter of radial basis function (RBF); M: number of decision trees in the RF

From the perspective of chemometric methods, the olfactory sensor array contained 108 variables, which varied significantly from each other. Hence, PCA played an influential function in dimensionality reduction and principal component extraction. Notably, the top three PCs were relatively minor, accounting for a mere 40.8% of the total. It implies that each sample has a unique position in the coordinate system of the principal components, emphasizing that each variable is essential. Consequently, the top 22 principal components were selected to explain the variability of the data better. Among the three qualitative discrimination models, the discrimination abilities of each model after the parameter optimization were credible. However, compared to the training set, the discrimination rates of ELM and RF in predicting independent samples were reduced by 5.56% and 2.28%, respectively. This difference can be attributed to several key reasons. First, SVM has inherent advantages for handling linearly differentiable data or high-dimensional data after kernel transformations, while RF and ELM may not be efficient enough in handling such data structures. Second, SVM can effectively avoid overfitting by precisely tuning the regularization parameters and kernel function, ensuring that the model performs well on sightless data. In contrast, the other two models, while also capable of handling complex datasets, might not be as flexible as SVM in adapting to data simplified by PCA. Furthermore, the dimensionality-decreased data from PCA simplifies the amount of information and cuts down the noise, providing clearer inputs for SVM. Although all of the models were able to deal with the same downscaled data, SVM showed itself to be more effective in capturing feature dimensions critical to classification decisions through careful parameter tuning. This does not mean that SVM will perform best on all reduced dimensionality datasets, but rather that SVM happens to have an advantage relative to the particular data and task of the moment. Consequently, SVM was shown to be more appropriate for the qualitative discrimination of maize silage quality.

CONCLUSIONS

1. This study verified that integrating colorimetric visualization sensor technology with appropriate chemometrics can enable high-precision quality discrimination of maize silage.

2. The study fabricated a novel olfactory visualization sensor system using 12 color-sensitive materials, which were used to form different odor fingerprints by sensitively cross-reacting the sensors with indicative volatile compounds released from maize silage feed during aerobic exposure.
3. Rapid, reliable, and non-destructive discrimination of maize silage based on these fingerprints was integrated with chemometric methods (PCA, ELM, SVM, RF). The SVM approach was shown to be superior to the other models, indicating a preferable generalization ability.
4. The results can provide high-precision technical support for quality identification of maize silage, having promising prospects for practical applications.

ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (32071893); the Science and Technology Program Project of Inner Mongolia Autonomous Region (2022YFDZ0024); the Natural Science Foundation of Inner Mongolia Autonomous Region (2023MS03044); the Postgraduate Research Innovation Program of Inner Mongolia Autonomous Region (DC2300002028); the Innovation Training Program for College Students of Inner Mongolia Autonomous Region (202310129026).

REFERENCES CITED

Ávila, C. L. S., and Carvalho, B. F. (2020). "Silage fermentation—updates focusing on the performance of micro-organisms," *Journal of Applied Microbiology* 128, 966-984. DOI: 10.1111/jam.14450

Borreani, G., Tabacco, E., Schmidt, R. J., Holmes, B. J., and Muck, R. E. (2018). "Silage review: Factors affecting dry matter and quality losses in silages," *Journal of Dairy Science* 101, 3952-3979. DOI: 10.3168/jds.2017-13837

Bouboulis, P., Theodoridis, S., Mavroforakis, C., and Dalla, L. (2015). "Complex support vector machines for regression and quaternary classification," *IEEE Transactions on Neural Networks and Learning Systems* 26, 1260-1274. DOI: 10.1109/TNNLS.2014.2336679

Dou, W.-T., Han, H.-H., Sedgwick, A. C., Zhu, G.-B., Zang, Y., Yang, X.-R., Yoon, J., James, T. D., Li, J., and He, X.-P. (2022). "Fluorescent probes for the detection of disease-associated biomarkers," *Science Bulletin* 67, 853-878. DOI: 10.1016/j.scib.2022.01.014

Gallo, A., Giuberti, G., Bruschi, S., Fortunati, P., and Masoero, F. (2016). "Use of principal factor analysis to generate a corn silage fermentative quality index to rank well- or poorly preserved forages," *Journal of the Science of Food and Agriculture* 96, 1686-1696. DOI: 10.1002/jsfa.7272

Garcia-Chavez, I., Meraz-Romero, E., Castelan-Ortega, O., Zaragoza-Esparza, J., Avalos, J. O., Jimenez, L. E. R., and Gonzalez-Ronquillo, M. (2022). "Corn silage, a systematic review of the quality and yield in different regions around the world," *Revista Corpoica-Ciencia y Tecnologia Agropecuaria* 23, article e2547. DOI: 10.21930/rcta.vol23_num3_art:2547

Guo, T., Tan, X., and Zhang, L. (2019). "Hierarchical pruning discriminative extreme learning machine," in: *Proceedings of ELM-2017*, Vol. 10, pp. 230-239, Springer International Publishing, Cham. DOI: 10.1007/978-3-030-01520-6_21

Janzen, M. C., Ponder, J. B., Bailey, D. P., Ingison, C. K., and Suslick, K. S. (2006). "Colorimetric sensor arrays for volatile organic compounds," *Analytical Chemistry* 78, 3591-3600. DOI: 10.1021/ac052111s

Jiang, H., Xu, W., and Chen, Q. (2019). "Evaluating aroma quality of black tea by an olfactory visualization system: Selection of feature sensor using particle swarm optimization," *Food Research International* 126, article 108605. DOI: 10.1016/j.foodres.2019.108605

Jolliffe, I. T., and Cadima, J. (2016). "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374, article 20150202. DOI: 10.1098/rsta.2015.0202

Karnatam, K. S., Mythri, B., Un Nisa, W., Sharma, H., Meena, T. K., Rana, P., Vikal, Y., Gowda, M., Dhillon, B. S., and Sandhu, S. (2023). "Silage maize as a potent candidate for sustainable animal husbandry development—Perspectives and strategies for genetic enhancement," *Frontiers in Genetics* 14, article 1150132. DOI: 10.3389/fgene.2023.1150132

Kayacelebi, A.A., Beckmann, B., Gutzki, F.-M., Jordan, J., and Tsikas, D. (2014). "GC–MS and GC–MS/MS measurement of the cardiovascular risk factor homoarginine in biological samples," *Amino Acids* 46, 2205-2217. DOI: 10.1007/s00726-014-1774-3

Khan, N. A., Yu, P., Ali, M., Cone, J. W., and Hendriks, W. H. (2015). "Nutritive value of maize silage in relation to dairy cow performance and milk quality," *Journal of the Science of Food and Agriculture* 95, 238-252. DOI: 10.1002/jsfa.6703

Kung, L., Shaver, R. D., Grant, R. J., and Schmidt, R. J. (2018). "Silage review: Interpretation of chemical, microbial, and organoleptic components of silages," *Journal of Dairy Science* 101, 4020-4033. DOI: 10.3168/jds.2017-13909

Liu, T., Jiang, H., and Chen, Q. (2020). "Qualitative identification of rice actual storage period using olfactory visualization technique combined with chemometrics analysis," *Microchemical Journal* 159, article 105339. DOI: 10.1016/j.microc.2020.105339

More, A. S., and Rana, D. P. (2022). "Performance enrichment through parameter tuning of random forest classification for imbalanced data applications," *Materials Today: Proceedings* 56, 3585-3593. DOI: 10.1016/j.matpr.2021.12.020

Ouyang, Q., Rong, Y., Wu, J., Wang, Z., Lin, H., and Chen, Q. (2023). "Application of colorimetric sensor array combined with visible near-infrared spectroscopy for the matcha classification," *Food Chemistry* 420, and 136078. DOI: 10.1016/j.foodchem.2023.136078

Serva, L., Andrighetto, I., Segato, S., Marchesini, G., Chinello, M., and Magrin, L. (2023). "Assessment of maize silage quality under different pre-ensiling conditions," *Data* 8, article 117. DOI: 10.3390/data8070117

Suslick, K. S., Rakow, N. A., and Sen, A. (2004). "Colorimetric sensor arrays for molecular recognition," *Tetrahedron* 60, 11133-11138. DOI: 10.1016/j.tet.2004.09.007

Tharangani, R. M. H., Yakun, C., Zhao, L. S., Ma, L., Liu, H. L., Su, S. L., Shan, L., Yang, Z. N., Kononoff, P. J., Weiss, W. P., *et al*. (2021). "Corn silage quality index: An index combining milk yield, silage nutritional and fermentation parameters," *Animal Feed Science and Technology* 273, article 114817. DOI: 10.1016/j.anifeedsci.2021.114817

Tuorila, H. (2015). "From sensory evaluation to sensory and consumer research of food: An autobiographical perspective," *Food Quality and Preference* 40, 255-262. Presented at the 10th Pangborn Sensory Science Symposium. DOI: 10.1016/j.foodqual.2014.05.006

Virmani, D., and Pandey, H. (2023). "Comparative analysis on effect of different SVM kernel functions for classification," in: *International Conference on Innovative Computing and Communications*, Vol. 492, pp. 657-670. Springer Nature Singapore, Singapore. DOI: 10.1007/978-981-19-3679-1_56

Wang, J., Jiang, H., and Chen, Q. (2021). "High-precision recognition of wheat mildew degree based on colorimetric sensor technique combined with multivariate analysis," *Microchemical Journal* 168, article 106468. DOI: 10.1016/j.microc.2021.106468

Wilkinson, J. M., and Davies, D. R. (2013). "The aerobic stability of silage: key findings and recent developments," *Grass and Forage Science* 68, 1-19. DOI: 10.1111/j.1365-2494.2012.00891.x

Wojnowski, W., Majchrzak, T., Dymerski, T., Gębicki, J., and Namieśnik, J. (2017). "Electronic noses: Powerful tools in meat quality assessment," *Meat Science* 131, 119-131. DOI: 10.1016/j.meatsci.2017.04.240

Wu, Z., Xu, E., Long, J., Wang, F., Xu, X., Jin, Z., and Jiao, A. (2015). "Measurement of fermentation parameters of Chinese rice wine using Raman spectroscopy combined with linear and non-linear regression methods," *Food Control* 56, 95-102. DOI: 10.1016/j.foodcont.2015.03.015

Xu, W., Jiang, H., Liu, T., He, Y., and Chen, Q. (2019). "Qualitative discrimination of yeast fermentation stages based on an olfactory visualization sensor system integrated with a pattern recognition algorithm," *Analytical Methods* 11, 3294-3300. DOI: 10.1039/C9AY00760A

Xu, W., He, Y., Li, J., Deng, Y., Xu, E., Feng, J., Ding, T., Liu, D., and Wang, W. (2022a). "Non-destructive determination of beef freshness based on colorimetric sensor array and multivariate analysis," *Sensors and Actuators B: Chemical* 369, article 132282. DOI: 10.1016/j.snb.2022.132282

Xu, W., He, Y., Li, J., Deng, Y., Zhou, J., Xu, E., Ding, T., Wang, W., and Liu, D. (2022b). "Olfactory visualization sensor system based on colorimetric sensor array and chemometric methods for high precision assessing beef freshness," *Meat Science* 194, article 108950. DOI: 10.1016/j.meatsci.2022.108950

Zhang, H., Wu, J., Gao, L., Yu, J., Yuan, X., Zhu, W., Wang, X., and Cui, Z. (2018). "Aerobic deterioration of corn stalk silage and its effect on methane production and microbial community dynamics in anaerobic digestion," *Bioresource Technology* 250, 828-837. DOI: 10.1016/j.biortech.2017.09.149

Article submitted: November 16, 2023; Peer review completed: March 13, 2024; Revised version received: March 19, 2024; Accepted: April 9, 2024; Published: April 18, 2024.
DOI: 10.15376/biores.19.2.3597-3613