



Research article

A fresh-cut papaya freshness prediction model based on partial least squares regression and support vector machine regression

Liyan Rong, Yajing Wang, Yanqun Wang, Donghua Jiang, Jinrong Bai, Zhaoxia Wu*, Lu Li, Tianyu Wang, Hui Tan**

College of Food Science, Shenyang Agricultural University, Shenyang, Liaoning 110866, China

ARTICLE INFO

Keywords:

Fresh-cut papaya
Freshness classification
Electronic tongue
Electronic nose
Model regression
Predictive analysis

ABSTRACT

This study investigated the physicochemical and flavor quality changes in fresh-cut papaya that was stored at 4 °C. Multivariate statistical analysis was used to evaluate the freshness of fresh-cut papaya. Aerobic plate counts were selected as a predictor of freshness of fresh-cut papaya, and a prediction model for freshness was established using partial least squares regression (PLSR), and support vector machine regression (SVMR) algorithms. Freshness of fresh-cut papaya could be well distinguished based on physicochemical and flavor quality analyses. The aerobic plate counts, as a predictor of freshness of fresh-cut papaya, significantly correlated with storage time. The SVMR model had a higher prediction accuracy than the PLSR model. Combining flavor quality with multivariate statistical analysis can be effectively used for evaluating the freshness of fresh-cut papaya.

1. Introduction

Papaya pulp has a high nutritional and medicinal value. It contains a variety of nutrients and mineral elements, making it a popular fruit for fresh consumer market consumption. Owing to consumers' busy lifestyles and increased health consciousness, fresh-cut fruit is highly sought after for convenience and health benefits. However, microorganisms can more easily invade fresh-cut papaya, since the original protective structure of the papaya is disrupted. Abiotic stresses, such as oxidative stress, accelerate the internal biochemical processes of papaya metabolism. These actions increase the susceptibility of fresh-cut papaya to decay and deterioration [1]. Therefore, ensuring the safety of fresh-cut papaya is particularly important.

The methods traditionally used to determine the shelf life of fresh-cut fruit rely on the physical, chemical, microbiological, and sensory evaluations of basic indicators [2]. However, most physicochemical methods require skilled personnel to operate, are extremely time-consuming and labor-intensive, and are not applicable to fresh-cut food with a short shelf life [3]. The electronic tongue (E-tongue) and electronic nose (E-nose) effectively simulate the human senses of taste and smell, respectively, and have already proved to be good substitutes for methods of detecting food taste and odor, with obvious advantages such as good reproducibility and objectivity [4]. The E-tongue selectively adsorbs different flavors of samples through biofilms on different sensors and generates different electrical signals into data for analysis based on different chemical information [5]. E-nose is an electronic sensing device detects odors through the interaction of various gas sensors with odor molecules to produce responsive sensor electrical signals, which

* Corresponding author. Shenyang Agricultural University, 120 Dongling Rd., Shenyang, 110866, China.

** Corresponding author.

E-mail addresses: wuzhaoxia@syau.edu.cn (Z. Wu), thth229@gmail.com (H. Tan).

<https://doi.org/10.1016/j.heliyon.2024.e30255>

Received 21 November 2023; Received in revised form 19 April 2024; Accepted 23 April 2024

Available online 26 April 2024

2405-8440/© 2024 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

is converted into processed data. They can quickly reflect the overall quality information of the samples, thereby identifying and classifying the samples.

Storing fresh-cut papaya reduces the microbial resistance of the papaya, gradually softens the flesh, changes its color and flavor, and limits consumer purchases. The microbiological infestation limit of fresh-cut fruit is 10^5 CFU/g (GB 29921-2021); therefore, a reliable shelf-life prediction model for fresh-cut papaya using aerobic plate counts as an indicator is particularly important. Partial least squares regression prediction (PLSR), which is the most commonly used analytical method for meteorological modeling, has been widely used in prediction models for the ripeness of apples [6], bananas [3], and croakers [7], etc. Support vector machine regression prediction (SVMR) is popular owing to its ability to efficiently train small datasets and achieve structural risk minimization [8], with applications in the prediction of the freshness of blueberries [9], mushrooms [10], chickens [11], and horse mackerel [12].

Currently, only a few methods have been reported for the rapid assessment of the freshness of fresh-cut fruit. Therefore, an immediately rational method for assessing the freshness of fresh-cut fruit and establishing the criteria for evaluating their quality is useful. This study aimed to use the PLSR and SVMR models of the E-tongue and E-nose, with the total aerobic plate counts as a freshness evaluation factor, to provide a reference for the assessment and prediction of freshness of fresh-cut papaya.

2. Materials and methods

2.1. Fruit materials

Undamaged Xiangmi red-fleshed papaya of similar weight was purchased from Fruit Wholesale Market located in Shenyang, Liaoning, China. The papaya originated in Hainan, China. After washing the intact papaya with room temperature ($20\text{ }^{\circ}\text{C}$) tap water to remove skin dust and wiping off the water with clean paper, the papaya was peeled and cut into cubes of approximately $3.0 \times 3.0 \times 3.0$ cm, which were placed in a box. A polyethylene-polyethylene terephthalate (PE-PET) laminate film was wrapped around the top of the boxes using a sealing machine and stored in a refrigerator at $4 \pm 1\text{ }^{\circ}\text{C}$.

2.2. Physicochemical quality measurement

Fresh-cut papaya samples were measured daily during eight days storage, with three technical replicates and three biological replicates for each sample (nine replicate experiments).

2.2.1. Flesh firmness

The firmness of fresh-cut papaya was determined using a texture analyzer (CT3 10 K; Brookfield, WI, USA). The test was performed with a 2 mm diameter TA39 cylindrical flat probe with a compression level of approximately 30 %, a test rate of 1.0 mm/s and a trigger force of 0.1 N.

2.2.2. Weight loss

During storage, loss of mass caused the fresh-cut papaya to shrink. The mass of the fresh-cut papaya was measured using a laboratory analytical balance (JJ500, Jiangsu, China) and the weight loss rate was calculated using equation (1):

$$\Delta W(\%) = \frac{W_0 - W_1}{W_0} \times 100\% \quad (1)$$

Where ΔW is the weight loss rate of fresh-cut papaya (%), W_0 is the initial mass of fresh-cut papaya before storage (g), W_1 is the mass of fresh-cut papaya after storage (g).

2.2.3. Titratable acid

The titratable acid (TAC) content of fresh-cut papaya was measured using the acid-base titration method. Fresh-cut papaya homogenate was taken in a triangular flask, distilled water was added to it, which was then placed in a water bath at $80\text{ }^{\circ}\text{C}$, then taken filtrate after cooling. Using phenolphthalein indicator, the filtrate was titrated with 0.10 mol L^{-1} NaOH for 30s without fading. The TAC equation (Eq. 2) is as follows:

$$\text{TAC}(\%) = \frac{V * 0.10 * V_1 * 0.070}{m * V_0} \times 100\% \quad (2)$$

Where V is the volume of the sample extract (ml), V_1 is the volume of the NaOH solution consumed (ml), m is the weight of the fresh-cut papaya (g), V_0 is the measured sampling volume.

2.2.4. Soluble sugar

The soluble sugar (SS) content of fresh-cut papaya was determined using the anthrone-sulphuric acid method [13].

2.2.5. Color difference

The color difference of fresh-cut papaya was measured using a colorimeter (Konica Minolta CR-400, Tokyo, Japan) with a pulsed xenon lamp as the light source maintained at 90° perpendicular to the samples during use. First, the colorimeter was calibrated by

pressing the measurement key after pointing the probe at the center of the white calibration plate. The probe was then aimed at the fresh-cut papaya, and color data were obtained by pressing the measurement key. The data were recorded as L^* (brightness), a^* (red-green), and b^* (yellow-blue). The final formula for the total color difference (ΔE) value of fresh-cut papaya was calculated using equation (3):

$$\Delta E = \sqrt{(L^* - L)^2 + (a^* - a)^2 + (b^* - b)^2} \quad (3)$$

Where L^* , a^* , b^* are the color parameters of fresh-cut papaya during storage and L , a , b are the initial color parameters of fresh-cut papaya before storage.

2.3. E-tongue analysis

According to the method suggested by Ref. [14], the flavor of fresh-cut papaya was analyzed using an E-tongue (Taste-Sensing System SA-402B, Tokyo, Japan), with slight modifications. The fresh-cut papaya samples were pounded and homogenized; the supernatant was filtered after centrifugation at 8000 r/min, 4 °C for 15 min. The pure juice of the filtered papaya was diluted with distilled water at a ratio of 1:2 to obtain the E-tongue test solution. The SA-402B E-tongue taste analysis system consists of six test sensors and three reference electrodes. They were activated 24 h prior to sample measurement. The test sensors measured the sourness (CA0), bitterness (C00), astringency (AE1), umami (AAE), saltiness (CT0), and sweetness (GL1). The taste of fresh-cut papaya samples was digitally determined using the six sensors mentioned above.

2.4. Electronic nose analysis

The aroma of fruit is essential for consumers' sensory acceptance. The characteristic aroma of the fresh-cut papaya was measured using a PEN3 E-nose (Insent Company, Atsugi City, Japan). The E-nose experiments were performed using direct headspace aspiration according to the method of [15]. The E-nose contains 10 different metallic oxide sensors respond to aromatic components such as benzene (W1C), nitrogen oxides (W5S), ammonia and aromatic components (W3C), hydrogen (W6S), alkane aromatic components (W5C), short chain alkanes (W1S), anode sulfides (W1W), alcohols, ethers, aldehydes and ketones (W2S), aromatic components and organic sulfides (W2W), and long-chain alkanes (W3S), to form the sensor array.

2.5. Qualitative analysis

Multivariate analysis was performed on the E-tongue and E-nose data of fresh-cut papaya. Hierarchical cluster analysis (HCA) and principal component analysis (PCA) were applied to qualitatively classify freshness of fresh-cut papaya. With the help of HCA, the dataset was divided into different levels according to the degree of similarity between variables; all samples were clustered into one category, and the results were displayed as a tree-shaped clustering diagram.

PCA is an effective dimensionality reduction method transforms multiple variables into a composite variable that represents as much of the original information as possible, that is, the principal components. PCA plots indicate the position of each sample in the principal components, and classify the samples in a basic manner [16].

2.6. Quantitative regression

A regression prediction model for the physicochemical quality indicators of fresh-cut papaya was developed based on sensor data from the E-tongue and E-nose. PLSR is one of the most commonly used methods for multiple linear regression modeling in chemometric modeling analysis; it creates a linear regression model by simulating predictors and observable variables and projecting them into a new space [17]. PLSR has the advantages of PCA and multiple linear regression, and is always applied to establish high-performance models in case of a high linear correlation between variables [18]. This time, Leave-one-out cross-validation method was used for cross-validation and the NIPALS iterative algorithm was applied to further implement the PLSR modeling.

SVMR is a classical nonlinear supervised learning modeling approach, and is an important branch that supports vector machine. It uses kernel functions to map low-dimensional data in a high-dimensional feature space, find linearly separable hyperplanes in the high-dimensional feature space, and map the hyperplanes in the high-dimensional space back to the low-dimensional space to explain nonlinear relationships [19]. This time, 5-fold cross-validation method was adopted for model validation, and selected Radial Basis Function (RBF) as the kernel function to find the optimal combination of penalty factor C and kernel parameter γ for modeling by using a grid-based search method [20].

2.7. Data set and model evaluation

During storage, 72 samples (8 storage days * 9 sets per day) were randomly selected as modeling samples, and 24 samples (8 storage days * 3 sets per day) were used as test samples to classify the freshness of fresh-cut papaya by applying physicochemical quality data, E-tongue, and E-nose sensor data, and to develop a regression prediction model for the freshness of fresh-cut papaya.

Leave-one-out validation (PLSR) and 5-fold cross-validation method (SVMR) of cross-validation methods were applied to validate the built model. The model evaluation parameters for the quantitative prediction were the model coefficient of determination (R_c^2),

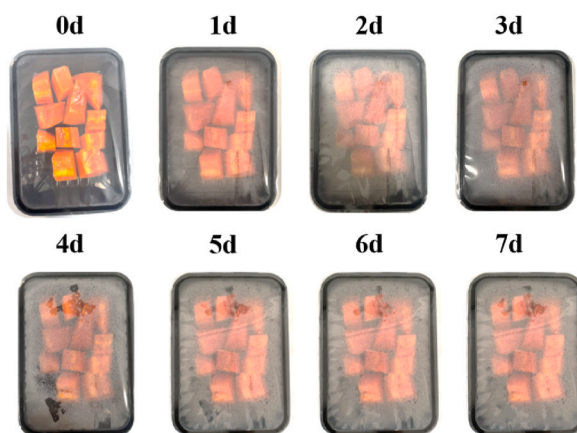


Fig. 1. Evolution of fresh-cut papaya during eight days of storage.

Table 1
Changes in the physicochemical quality of fresh-cut papaya during storage.

| Storage time (d) | Flesh firmness (N) | Weight loss (%) | TAC (%) | SS (%) |
|------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 0 | 0.74 ± 0.07 ^a | 0.00 ± 0.00 ^g | 0.13 ± 0.00 ^a | 2.29 ± 0.03 ^h |
| 1 | 0.57 ± 0.05 ^b | 0.01 ± 0.00 ^g | 0.12 ± 0.00 ^b | 2.38 ± 0.02 ^g |
| 2 | 0.42 ± 0.03 ^c | 0.06 ± 0.01 ^f | 0.10 ± 0.00 ^c | 2.50 ± 0.01 ^f |
| 3 | 0.35 ± 0.03 ^d | 0.12 ± 0.01 ^e | 0.10 ± 0.00 ^c | 2.77 ± 0.01 ^e |
| 4 | 0.30 ± 0.02 ^e | 0.16 ± 0.01 ^d | 0.09 ± 0.01 ^d | 3.10 ± 0.04 ^d |
| 5 | 0.23 ± 0.02 ^f | 0.25 ± 0.02 ^c | 0.08 ± 0.00 ^d | 3.27 ± 0.01 ^c |
| 6 | 0.20 ± 0.02 ^f | 0.36 ± 0.04 ^b | 0.08 ± 0.01 ^d | 3.41 ± 0.01 ^b |
| 7 | 0.13 ± 0.02 ^g | 0.43 ± 0.02 ^a | 0.07 ± 0.01 ^e | 3.83 ± 0.01 ^a |

Different letters in the same column indicate significant differences ($p < 0.05$).

root mean square error of calibration (RMSEC), prediction coefficient of determination (R_p^2), root mean square error of prediction (RMSEP), and residual prediction deviation (RPD) [21]. The specific calculation equations (Eq. 4, 5, 6) are as follows:

$$R_c^2, R_p^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - y_m)^2} \quad (4)$$

$$RMSEC, RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2} \quad (5)$$

$$RPD = \frac{SD}{RMSEP} = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - y_m)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2}} \quad (6)$$

Where n is the number of samples in the calibration or the prediction; y_i and \hat{y} represent the actual and predicted values of the i -th sample; y_m is the mean value of the actual values in the calibration or the prediction; SD is the standard deviation of the physicochemical indexes in the prediction.

2.8. Statistical analysis

Physicochemical indices are expressed as the mean ± standard deviation obtained in the nine experiments. One-way analysis of variance (ANOVA) was performed using SPSS 25 (SPSS Inc., Chicago, IL, USA). Duncan's multiple range test was conducted to determine the significance of the data at a significance level of $p < 0.05$. HCA and PCA were performed using Origin 2021 (Origin Lab, Northampton, MA, USA), and PLSR prediction models were constructed using Unscrambler X 10.4 (CAMO AS, Trondheim, Norway).

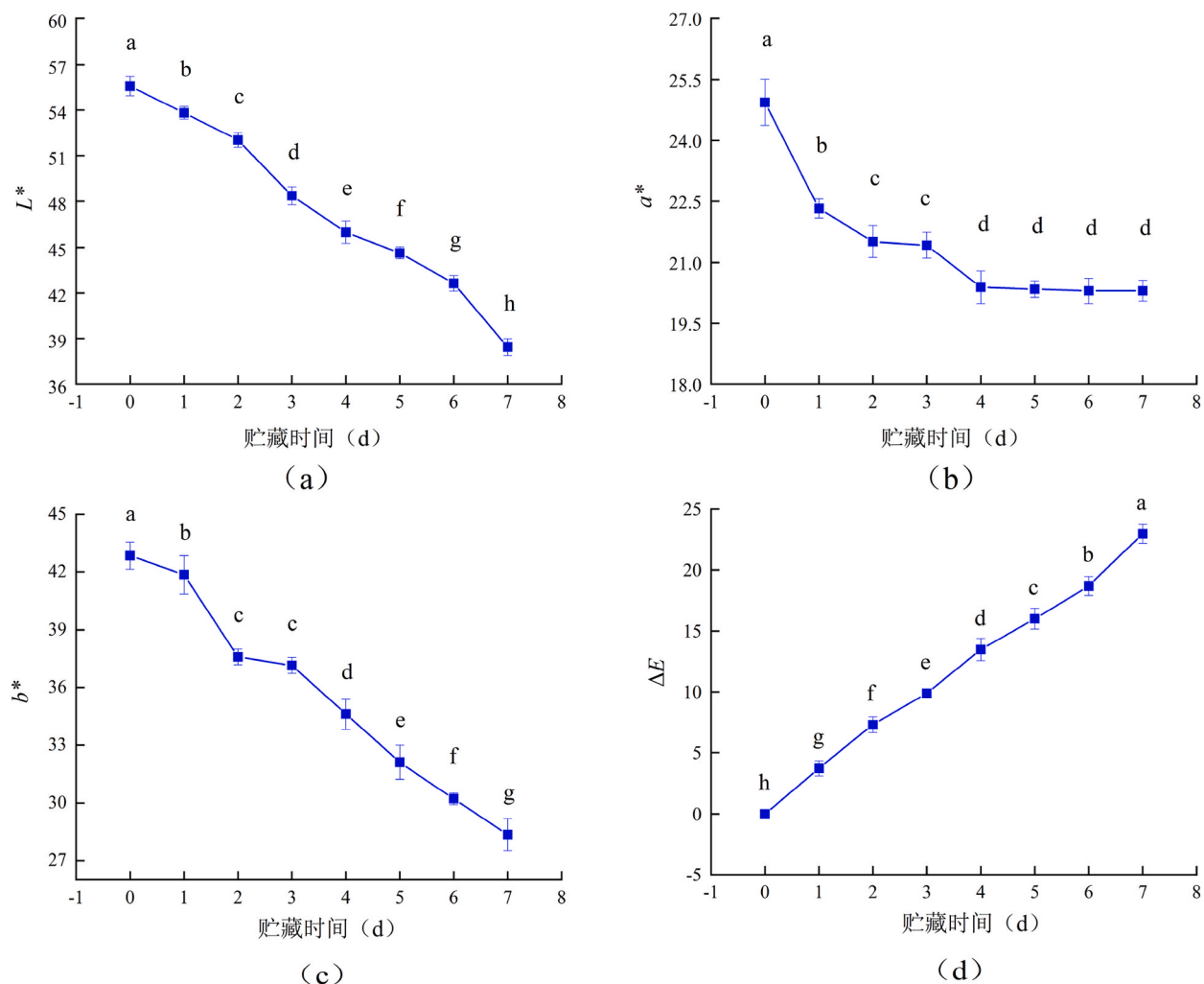


Fig. 2. Changes in L^* (A), a^* (B), b^* (C), and ΔE (D) of fresh-cut papaya during storage. Different letters indicate significant differences ($p < 0.05$).

3. Results and discussion

3.1. Physicochemical quality analysis of fresh-cut papaya

3.1.1. Firmness and weight loss

The evolution of fresh-cut papaya during 8 days of storage was shown in Fig. 1, and the specific variations in firmness, weight loss, TAC, SS, aerobic plate counts, and color difference (ΔE) were depicted in Table 1. Moisture loss is the primary reason for the change in firmness in the fruit [22], Table 1 showed that the firmness of fresh-cut papaya decreased with increase in storage time, from 0.74 ± 0.07 N to 0.13 ± 0.02 N; the result aligned with the finding of Sekarina's study [23]. For fresh-cut products, absence of peel protection and a larger surface area exposure result in greater weight loss. Weight loss usually happens owing to water loss or juice leakage, and significant variations during storage are attributed mainly to respiration and transpiration of freshly-cut papaya [24].

3.1.2. TAC and SS

TAC and SS are important factors affecting the sensory characteristics of fruit. The TAC content demonstrated a decreasing trend during storage; SS content demonstrated a significant increase. This can be attributed to the degradation of polysaccharides and conversion of organic acids, leading to an increase in sweetness, and decrease in acidity [6].

3.1.3. Color difference

The colorimeter provides a numerical representation of the color that is difficult to distinguish with the naked eye in response to the magnitude of the color difference. As shown in Fig. 2, the L^* (Fig. 2A), a^* (Fig. 2B), and b^* (Fig. 2C) values of fresh-cut papaya decreased continuously during storage, indicating the fresh-cut papaya flesh color gradually darkened, the redness value gradually

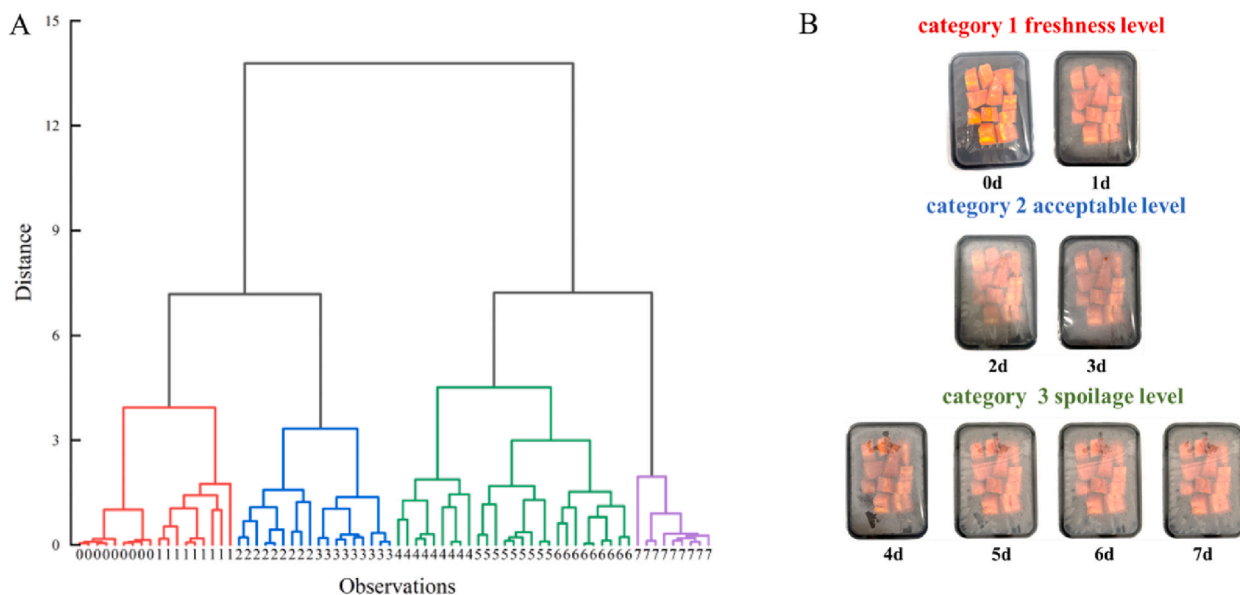


Fig. 3. HCA analysis of fresh-cut papaya (A) and photographs (B) of physicochemical indicators during storage.

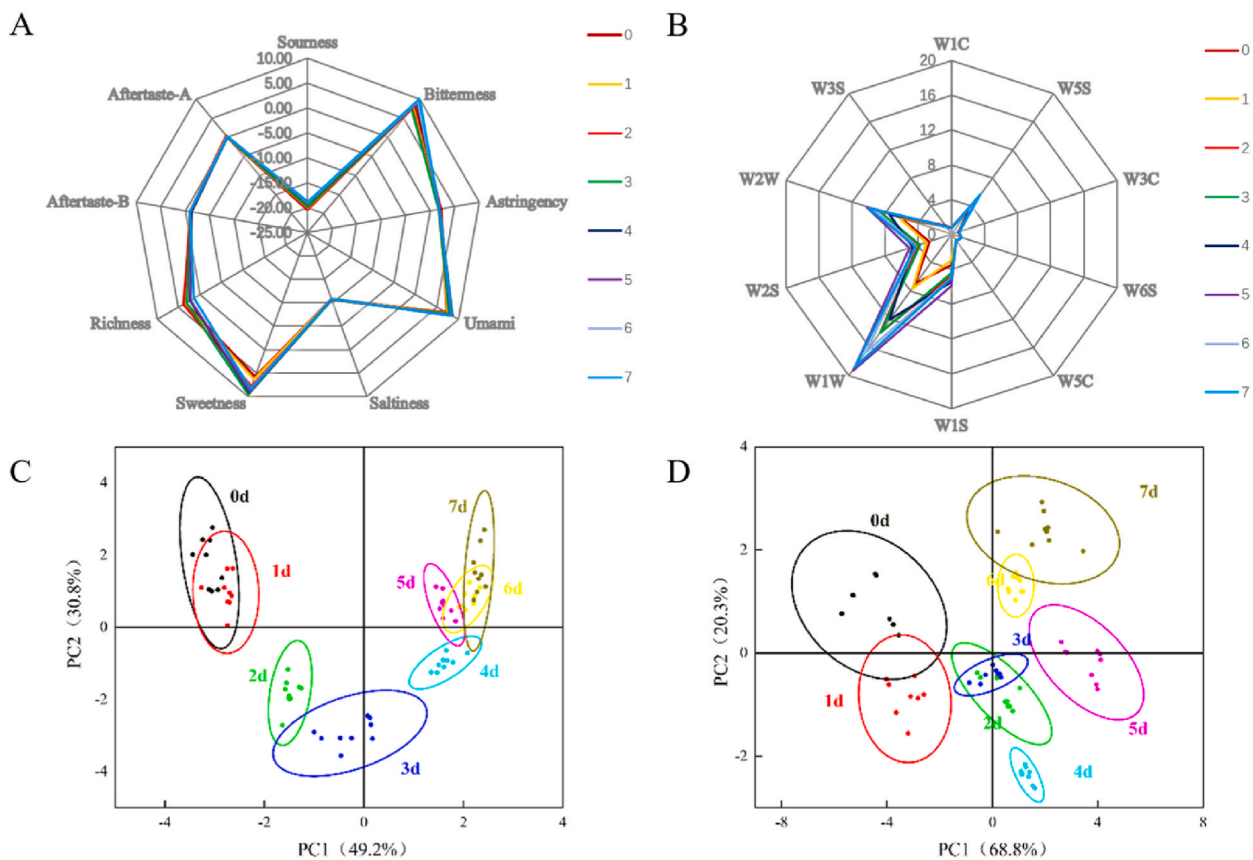


Fig. 4. Changes in E-tongue (A), and E-nose (B) response values of fresh-cut papaya during storage. PCA plot of E-tongue (C) and E-nose (D) of fresh-cut papaya during storage.

decreased, and the yellow color kept fading. These changes illustrate the gradual loss of freshness of fresh-cut papaya. The ΔE value (Fig. 2D) of fresh-cut papaya increased significantly daily, reaching 22.96 ± 0.77 by the end of the storage period. It was reported by Wei [25] that $\Delta E > 2$ was a significant color difference, which indicated that the color of fresh-cut papaya was easily distinguished by the color change visible to the naked eye every day during storage.

3.1.4. Aerobic plate counts and sensory scores

In a previous study [26], we analyzed the aerobic plate counts and sensory scores of fresh-cut papaya during storage. A preliminary determination of freshness of fresh-cut papaya was performed on day 3 of storage using a combination of the two.

3.2. Classification of fresh-cut papaya freshness on the basis of physicochemical quality

The physicochemical quality indicators of fresh-cut papaya mentioned above were analyzed as parameters for HCA by applying the clustering method of the longest distance and Euclidean distance measures, and the results were illustrated in Fig. 3A. HCA is represented as a dendrogram, where the distance between the observations indicates the similarity between different observations [27]. The dendrogram revealed that the storage period of fresh-cut papaya was divided into four major categories depending on the length of storage: 0 and 1 days as the first category, 2 and 3 days as the second category; 4, 5, and 6 days as the third category, and 7 day as the fourth category. Of the 72 samples used in the experiment, all 72 were correctly classified; the HCA was 100.0 % correct. By combining the acceptable limits of the total aerobic plate counts and sensory score limits for fresh-cut papaya, we classified the freshness of fresh-cut papaya into three levels (Fig. 3B). That is, category 1 freshness level (days 0 and 1), category (days 2 and 3), and category 3 spoilage level (days 4, 5, 6, and 7).

3.3. Flavor analysis of fresh-cut papaya

3.3.1. Electronic tongue analysis

The E-tongue simulates the human oral taste organ using a taste-intelligent bionic system to detect the taste of a sample. The taste variation of fresh-cut papaya at 4 °C in the storage period was shown in Fig. 4A. The radar plot can indirectly reflect the variability of the taste of fresh-cut papaya during storage, where the response value of the CT0 sensor was around -10 below 6, and almost constant during storage, indicating that fresh-cut papaya was tasteless in terms of salinity [28]. The GL1 sensor showed the greatest variability, indicating that the sweetness of fresh-cut papaya changed significantly during storage, followed by AAE, C00, CA0, and AE1. At the early storage stage of fresh-cut papaya, the sweetness showed an upward trend, and the bitterness displayed a downward trend; at the late storage stage, a significant increase in bitterness was observed, and the sweetness decreased, which may be attributed to the reduction of internal active substances and the oxidative accumulation of phenolic substances in papaya [6]. Umami and richness were aligned with the trend of sweetness, whereas astringency, aftertaste-A, aftertaste-B, and bitterness showed the same trend. The sourness of fresh-cut papaya showed an increasing trend throughout the storage process, which was consistent with the results of [28], and may be attributed to the accumulation of organic acids. To conclude, the E-tongue sensor response values of fresh-cut papaya changed significantly with storage time, indicating that the taste of fresh-cut papaya changed regularly during storage.

3.3.2. Electronic nose analysis

The response values of the E-nose sensors depended on the chemical composition of the volatiles [29]. Fresh-cut papaya has a strong aroma, to which the E-nose sensor responds significantly, as depicted in Fig. 4B; the W1W, W2W, W5S, W1S and W2S sensors were mainly sensitive to response by fresh-cut papaya. With an increase in storage time, the response values of all five sensors showed a gradual increase, among which the W1W sensor showed maximum change, indicating that the inorganic sulfide content of fresh-cut papaya changed significantly during storage, followed by W2W for organic sulfide and aromatic components, whereas the rest did not increase significantly. The response values of the WIC, W3C, W6S, W5C, and W3S sensors were approximately 1 throughout the test with no significant changes. To conclude, the response values of the E-nose sensors of fresh-cut papaya increased with an increase in storage time, and the response values were close to each other on storage days 0 and 1; the response values increased significantly on storage day 2; the changes were stable at storage days 3 and 4; the response values increased significantly on storage day 5; and the changes were stable at storage days 6 and 7, which indicated that the changes in the aroma of fresh-cut papaya during storage had obvious regularity.

3.4. Classification of fresh-cut papaya freshness based on flavor analysis

3.4.1. Classification of fresh-cut papaya freshness using electronic tongue

The E-tongue response values for fresh-cut papaya at different storage periods were analyzed using PCA to determine the difference in taste. PCA is an efficacious dimensionality reduction method that retains the majority of information from raw data and assigns each observation to a specific group [16]. Fig. 4 showed the two-dimensional PCA plots indicating incomplete separation between the different storage stages of fresh-cut papaya. Different colored circles represented samples at different storage stages. As shown in Fig. 4C, the plot of principal component (PC) scores, the first two PCs, described the eight different categories of samples well, with PC1 and PC2 explaining 49.2 % and 30.8 % of the variance, respectively, and accounting for 80.0 % of the total variance. However, there were certain overlapping areas in this plot, where the data were clustered together: days 0 and 1; days 2, and 3; days 4, 5, 6, and 7. Based on the acceptable limit of the aerobic plate counts of fresh-cut fruit, the PCA plot enabled explicit distinction of the following

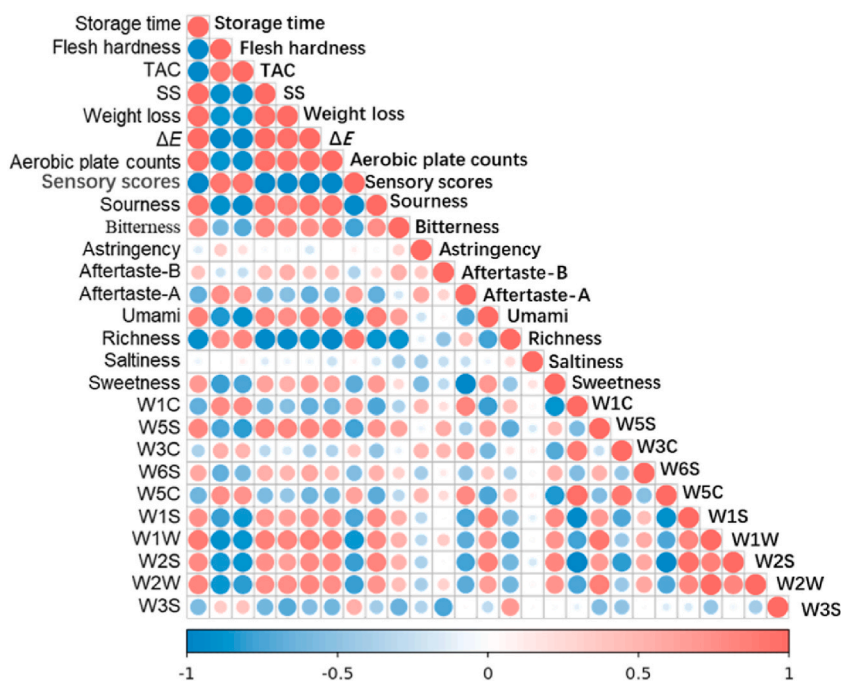


Fig. 5. Correlation analysis of physicochemical quality of fresh-cut papaya, e-tongue, and e-nose responses ($p < 0.05$). The sizes of the circles in the figure are proportional to the correlation; the larger the area, the higher the correlation. The color of the circles indicates a positive or negative correlation, with red depicting a positive and blue depicting a negative correlation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

three categories of fresh-cut papaya storage periods: category 1 freshness level (days 0 and 1); category 2 acceptable level (days 2, and 3); and, category 3 spoilage level (days 4, 5, 6, and 7), which was consistent with the above trend of freshness classification based on the physicochemical quality indicators of fresh-cut papaya.

3.4.2. Classification of fresh-cut papaya freshness based on E-nose

The PCA results of E-nose showed two PC that explained 89.1 % of the total variance during storage of fresh-cut papaya at 4 °C, with 68.8 % for PC1 and 20.3 % for PC2. From Fig. 4D, it can be seen that the response values of samples on day 0 and day 1 were close to each other and were significantly distinct on days 2, 3, 4, 5, 6, and 7, as revealed by PC1; while samples on days 5, 6 and 7 were significantly distinct as revealed by PC2. This indicated that the volatile compounds in fresh-cut papaya changed significantly during the different storage periods. Fresh-cut papaya samples stored for different durations occupied relatively independent spaces in the principal component space (Fig. 3D), indicating that fresh-cut papaya samples stored for different durations could be clearly distinguished. The storage periods of fresh-cut papaya can be roughly divided into five groups: days 0 and 1; days 2, 3 and 4; days 5, 6 and 7. As per the analysis of aerobic plate counts, the freshness levels of the five groups of fresh-cut papaya corresponded to fresh, sub-fresh, acceptable, primary spoilage, and spoilage levels.

3.5. Freshness prediction model construction

3.5.1. Correlation analysis

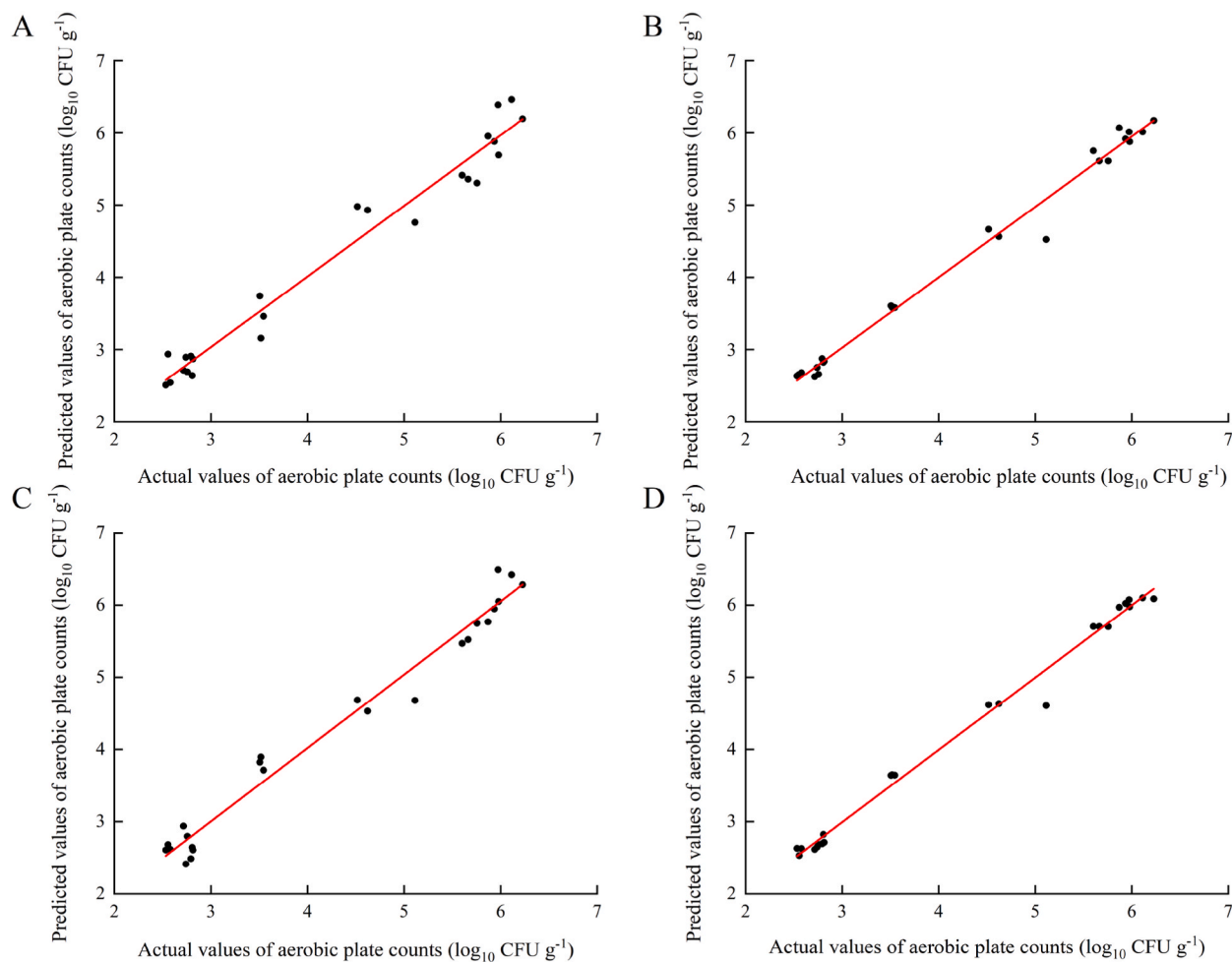
The flesh texture, color, TAC, SS, aerobic plate counts, sensory scores, and flavor of fresh-cut papaya changed significantly during storage. Correlation analysis was performed to evaluate their effects on the shelf life of fresh-cut papaya. As shown in Fig. 5, the storage time of fresh-cut papaya showed a highly significant negative correlation with flesh firmness, TAC, sensory scores, and a highly significant positive correlation with SS, weight loss, aerobic plate counts, and ΔE . It was revealed that they were all closely related to the storage time of fresh-cut papaya, and a reasonable identification of the freshness of fresh-cut papaya can be achieved using their prediction. As the total aerobic plate counts is the quality indicator with clear limit values and an important parameter in determining food safety, the total aerobic plate counts of fresh-cut papaya was selected as a predictor for the freshness model.

Among the response values of the E-tongue, sourness, umami, bitterness, and sweetness were positively related to storage time and aerobic plate counts of fresh-cut papaya, whereas richness and aftertaste-A was negatively related to them. The response values of the 10 E-nose sensors had good correlation coefficients with respect to storage time and aerobic plate counts of fresh-cut papaya; W1C, W5C, W3S, and W3C correlated negatively with the storage time and aerobic plate counts of fresh-cut papaya, and W2S, W1S, W6S, W1W, W5S, and W2W were positively correlated. The correlation results indicated that the taste and aroma of fresh-cut papaya changed as storage time increased, with a stronger correlation with the aerobic plate counts. Therefore, the freshness prediction model

Table 2

Aerobic plate counts prediction results based on fresh-cut papaya E-tongue dataset and E-nose dataset.

| Dataset | Model | Calibration | | Validation | | RPD |
|----------|-------|-------------|--------|------------|--------|-------|
| | | R_c^2 | RMSEC | R_p^2 | RMSEP | |
| E-tongue | PLSR | 0.9655 | 0.2634 | 0.9682 | 0.2554 | 5.72 |
| | SVMR | 0.9919 | 0.1300 | 0.9887 | 0.1523 | 9.60 |
| E-nose | PLSR | 0.9617 | 0.2777 | 0.9746 | 0.2284 | 6.40 |
| | SVMR | 0.9867 | 0.1639 | 0.9912 | 0.1343 | 10.89 |

**Fig. 6.** Actual and predicted values of PLSR (A, C) and SVMR (B, D) models for aerobic plate counts of fresh-cut papaya during storage based on E-tongue (A, B) and E-nose (C, D) dataset.

based on flavor and aerobic plate counts was reasonable.

3.5.2. Freshness prediction model

PLSR and SVMR models were used to predict the aerobic plate counts of fresh-cut papaya, and the prediction performances of different models were compared to test the feasibility of model construction [14]. The larger the models R_c^2 and R_p^2 , the less RMSEC and RMSEP, the higher was the model accuracy and predictive power [30]. Furthermore, $RPD > 3.0$ indicates an excellent prediction accuracy of the model [31].

The PLSR model was applied to the E-tongue and E-nose datasets, with modeling results listed in Table 2. Of the two datasets, the model based on E-tongue dataset possessed higher R_c^2 , R_p^2 , and lower RMSEC and RMSEP, illustrating the PLSR model was more suitable for prediction of E-tongue dataset. However, both RPDs were greater than 3.0 with little gap, indicating the PLSR model exhibited good prediction accuracy in both E-tongue and E-nose datasets.

Meanwhile, SVMR model was applied to E-tongue and E-nose datasets. The RBF was chosen as the kernel function, and optimal result of C and γ were found by using grid search method. $\log_{10}C$ and $\log_{10}(\gamma)$ were in the range of $[-2, 2]$, and the optimal parameters were selected by using the 5-fold cross-validation method [20]. The optimal parameter pairs (C , γ) for the E-tongue and E-nose datasets were obtained after cross-validation as (4.462, 0.215) and (35.938, 0.074), respectively. The data in Table 2 show that the SVMR model has good prediction accuracy in both the e-tongue and e-nose datasets. The SVMR prediction performance is optimal when comparing the two models. The final optimal parameter pairs (C , γ) of (4.462, 0.074) and (100, 0.074) were obtained for the E-tongue and E-nose datasets, respectively. Table 2 data revealed the SVMR model provided good prediction accuracy in both the e-tongue and e-nose datasets.

Fig. 6 visualized the predicted and actual aerobic plate counts of fresh-cut papaya during storage. Comparing the two models, SVMR model (Fig. 6B and D) surpassed PLSR model (Fig. 6A and C) in predicting the total aerobic plate counts of fresh-cut papaya during storage. Hence, the SVMR model can well predict the total aerobic plate counts of fresh-cut papaya during storage, thereby evaluating its freshness, and ensuring consumer safety.

4. Conclusions

The freshness of fresh-cut papaya correlated highly with the physicochemical quality indicators and changes in flavor quality during storage. HCA was used to synthesize several single quality indicators related to freshness to make a holistic judgment of freshness, with an accuracy of 100.0%. PCA was performed to determine the freshness of fresh-cut papaya at different storage periods, and the results of freshness classification were consistent with those based on physicochemical quality indicators. Correlation analysis showed a significant correlation between the E-tongue and E-nose sensor response values and aerobic plate counts, with clear edible limits for aerobic plate counts, making their prediction a reasonable response to the freshness of fresh-cut papaya. By comparing the RPD of the two models, the PLSR model was only suitable for predicting freshness based on the E-tongue dataset, whereas the SVMR model presented high prediction accuracy for both the E-tongue and E-nose datasets. These results demonstrated that SVMR model can be effectively applied for prediction of freshness of fresh-cut papaya. However, the application of freshness prediction for different packaging methods, storage conditions, and fruit requires further investigation.

Data availability statement

No data associated in this work has been deposited into a publicly available repository.

CRediT authorship contribution statement

Liyan Rong: Writing – original draft, Data curation, Conceptualization. **Yajing Wang:** Investigation. **Yanqun Wang:** Methodology. **Donghua Jiang:** Methodology. **Jinrong Bai:** Resources. **Zhaoxia Wu:** Writing – review & editing, Supervision. **Lu Li:** Funding acquisition. **Tianyu Wang:** Validation. **Hui Tan:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Application and Foundation Research Project of Liaoning Province (No. 2022JH2/101300177).

References

- [1] A.S. More, C.S. Ranadheera, Z. Fang, et al., Biomarkers associated with quality and safety of fresh-cut produce, *Food Biosci.* 34 (2020) 100524.
- [2] M. Mohd Ali, N. Hashim, S.K. Bejo, et al., Innovative non-destructive technologies for quality monitoring of pineapples: recent advances and applications, *Trends Food Sci. Technol.* 133 (2023) 176–188.
- [3] X. Chu, P. Miao, K. Zhang, et al., Green banana maturity classification and quality evaluation using hyperspectral imaging [J/OL] 12 (4) (2022), <https://doi.org/10.3390/agriculture12040530>.
- [4] L. Chen, F. Ning, L. Zhao, et al., Quality assessment of royal jelly based on physicochemical properties and flavor profiles using HS-SPME-GC/MS combined with electronic nose and electronic tongue analyses, *Food Chem.* 403 (2023) 134392.
- [5] J.X. Leon-Medina, M. Anaya, D.A. Tibaduiza, Yogurt classification using an electronic tongue system and machine learning techniques, *Intelligent Systems with Applications* 16 (2022) 200143.
- [6] Y. Zhang, D. Zhu, X. Ren, et al., Quality changes and shelf-life prediction model of postharvest apples using partial least squares and artificial neural network analysis, *Food Chem.* 394 (2022) 133526.
- [7] Y. Shao, Y. Shi, K. Wang, et al., Detection of small yellow croaker freshness by hyperspectral imaging, *J. Food Compos. Anal.* 115 (2023) 104980.
- [8] B.S. Luka, T.K. Yuguda, M. Adnoui, et al., Drying temperature-dependent profile of bioactive compounds and prediction of antioxidant capacity of cashew apple pomace using coupled Gaussian Process Regression and Support Vector Regression (GPR–SVR) model, *Heliyon* 8 (9) (2022) e10461.
- [9] W. Huang, X. Wang, J. Zhang, et al., Improvement of blueberry freshness prediction based on machine learning and multi-source sensing in the cold chain logistics, *Food Control* 145 (2023) 109496.
- [10] A. Mukherjee, T. Sarkar, K. Chatterjee, et al., Development of artificial vision system for quality assessment of oyster mushrooms, *Food Anal. Methods* 15 (6) (2022) 1663–1676.

- [11] S.D. Astuti, M.H. Tamimi, A.A.S. Pradhana, et al., Gas sensor array to classify the chicken meat with *E. coli* contaminant by using random forest and support vector machine, *Biosens. Bioelectron.* X 9 (2021) 100083.
- [12] H. Li, Y. Wang, J. Zhang, et al., Prediction of the freshness of horse mackerel (*Trachurus japonicus*) using E-nose, E-tongue, and colorimeter based on biochemical indexes analyzed during frozen storage of whole fish, *Food Chem.* 402 (2023) 134325.
- [13] Y. Tang, J. Ren, C. Liu, et al., Genetic characteristics and QTL analysis of the soluble sugar content in ripe tomato fruits, *Sci. Hortic.* 276 (2021) 109785.
- [14] D. Zhu, X. Ren, L. Wei, et al., Collaborative analysis on difference of apple fruits flavour using electronic nose and electronic tongue, *Sci. Hortic.* 260 (2020) 108879.
- [15] D. Zhu, Y. Shen, L. Wei, et al., Effect of particle size on the stability and flavor of cloudy apple juice, *Food Chem.* 328 (2020) 126967.
- [16] S. Panebianco, P. Mazzoleni, G. Barone, et al., Feasibility study of tomato fruit characterization by fast XRF analysis for quality assessment and food traceability, *Food Chem.* 383 (2022) 132364.
- [17] D. Saha, T. Senthilkumar, S. Sharma, et al., Application of near-infrared hyperspectral imaging coupled with chemometrics for rapid and non-destructive prediction of protein content in single chickpea seed, *J. Food Compos. Anal.* 115 (2023) 104938.
- [18] S. Yu, X. Huang, L. Wang, et al., Qualitative and quantitative assessment of flavor quality of Chinese soybean paste using multiple sensor technologies combined with chemometrics and a data fusion strategy, *Food Chem.* 405 (2023) 134859.
- [19] H. Song, F. Li, P. Guang, et al., Detection of aflatoxin B1 in peanut oil using attenuated total reflection fourier transform infrared spectroscopy combined with partial least squares discriminant analysis and support vector machine models, *J. Food Protect.* 84 (8) (2021) 1315–1320.
- [20] A. Tan, Y. Zhao, K. Sivashanmugan, et al., Quantitative TLC-SERS detection of histamine in seafood with support vector machine analysis, *Food Control* 103 (2019) 111–118.
- [21] S.R. Karunathilaka, B.J. Yakes, S.H. Choi, et al., Comparison of the performance of partial least squares and support vector regressions for predicting fatty acids and fatty acid classes in marine oil dietary supplements by using vibrational spectroscopic data, *J. Food Protect.* 83 (5) (2020) 881–889.
- [22] S. Dos Passos Braga, M. Magnani, M.S. Madruga, et al., Characterization of edible coatings formulated with chitosan and *Mentha* essential oils and their use to preserve papaya (*Carica papaya* L.), *Innovat. Food Sci. Emerg. Technol.* 65 (2020) 102472.
- [23] A.S. Sekarina, Supriyadi, H.S.H. Munawaroh, et al., Effects of edible coatings of chitosan - fish skin gelatine containing black tea extract on quality of minimally processed papaya during refrigerated storage, *Carbohydrate Polymer Technologies and Applications* 5 (2023) 100287.
- [24] N. Tabassum, M.A. Khan, Modified atmosphere packaging of fresh-cut papaya using alginate based edible coating: quality evaluation and shelf life study, *Sci. Hortic.* 259 (2020) 108853.
- [25] X. Wei, T. Verma, S. Irmak, et al., Effect of storage on microbial reductions after gaseous chlorine dioxide treatment of black peppercorns, cumin seeds, and dried basil leaves, *Food Control* 148 (2023) 109627.
- [26] L. Rong, T. Zhang, Y. Ma, et al., An intelligent label using sodium carboxymethyl cellulose and carrageenan for monitoring the freshness of fresh-cut papaya, *Food Control* 145 (2023) 109420.
- [27] Q. Jiang, M. Zhang, A.S. Mujumdar, et al., Non-destructive quality determination of frozen food using NIR spectroscopy-based machine learning and predictive modelling, *J. Food Eng.* 343 (2023) 111374.
- [28] M. Li, X. Li, C. Han, et al., UV-C treatment maintains quality and enhances antioxidant capacity of fresh-cut strawberries, *Postharvest Biol. Technol.* 156 (2019).
- [29] L.I. Junxing, M. Aiqing, Z. Gangjun, et al., Assessment of the 'taro-like' aroma of pumpkin fruit (*Cucurbita moschata* D.) via E-nose, GC-MS and GC-O analysis, *Food Chem. X* 15 (2022) 100435.
- [30] Y. Li, B. Ma, C. Li, et al., Accurate prediction of soluble solid content in dried Hami jujube using SWIR hyperspectral imaging with comparative analysis of models, *Comput. Electron. Agric.* 193 (2022) 106655.
- [31] B.K. Panda, G. Mishra, W.A. Ramirez, et al., Rancidity and moisture estimation in shelled almond kernels using NIR hyperspectral imaging and chemometric analysis, *J. Food Eng.* 318 (2022) 110889.