INTERNATIONAL JOURNAL
OF INTELLIGENT SYSTEM
AND INFORMATION
PROCESSING

Continuous Nonlinear Modeling for Independent Storage-Day Prediction of Grapes Using Electronic-Nose Responses

Nurgül SENYÜCEL
Canakkale Onsekiz Mart University
nurgulsenyucel@comu.edu.tr

Servet SENYÜCEL
Istanbul Arel University
servetsenyucel@arel.edu.tr

Abstract

Accurate determination of fruit storage status is essential for intelligent postharvest quality management. Although electronic noses have been widely employed for freshness assessment, most existing studies directly analyze raw temporal sensor signals or discrete response features, which often produce high-dimensional and difficult-to-interpret machine learning models. This study proposes a continuous nonlinear feature engineering framework for storage-day prediction of grapes using electronic-nose responses. Temporal signals acquired from four metal-oxide semiconductor sensors during a 100-s measurement period were modeled using three continuous nonlinear functions, namely the **Weibull**, **Morgan–Mercer–Flodin (MMF)**, and **Richards** models. Model comparison based on RMSE, R^2, AIC, and BIC demonstrated that the Richards model provided the most accurate representation of the sensor response kinetics. The area under each Richards curve (AUC) was subsequently calculated to generate a compact four-dimensional feature vector representing the cumulative volatile exposure of each sensor. These curve-derived features were used to train a Random Forest classifier for storage-day prediction. Performance was evaluated using **21 completely independent grape samples** collected separately from the reference dataset. The proposed framework achieved **90.48%** exact storage-day prediction accuracy, while all independent samples were correctly identified within **±1 storage day**. Despite reducing the original temporal measurements by approximately **98%**, the Richards-derived AUC features preserved the essential information required for accurate prediction. The results demonstrate that continuous nonlinear modeling combined with area-based feature extraction provides an interpretable, computationally efficient, and transferable representation of electronic-nose responses for practical postharvest storage monitoring.

Keywords: Electronic nose; grapes; postharvest storage; continuous nonlinear modeling; Richards model; area under the curve (AUC); Random Forest; independent validation.

1. Introduction

1.1. Postharvest Storage Monitoring of Grapes

Table grapes (*Vitis vinifera* L.) are among the most economically important fresh fruits worldwide because of their high nutritional value, attractive sensory properties, and extensive international trade. However, maintaining grape quality after harvest remains a major challenge due to the fruit's susceptibility to physiological deterioration and microbial spoilage during storage. Although grapes are classified as non-climacteric fruits, postharvest metabolic activity continues, leading to moisture loss, berry softening, rachis browning, membrane degradation, and gradual changes in biochemical composition. These processes progressively reduce marketability and consumer acceptance, making effective postharvest quality monitoring essential throughout storage, transportation, and distribution (Romero et al., 2020; Zhang et al., 2025).

During storage, grapes undergo continuous modifications in their physicochemical and biochemical characteristics. Parameters such as firmness, soluble solids content, titratable acidity, phenolic compounds, and antioxidant capacity gradually change as senescence progresses. Simultaneously, the volatile organic compound (VOC) profile evolves owing to ongoing respiration, enzymatic activity, and microbial metabolism. Variations in alcohols, aldehydes, ketones, esters,

terpenes, and other aroma-related compounds reflect the physiological status of the fruit and provide valuable indicators of freshness and quality deterioration (Romero et al., 2020; Sun et al., 2025).

Conventional postharvest quality evaluation primarily relies on visual inspection or destructive laboratory analyses, including firmness measurements, soluble solids determination, titratable acidity, chromatographic analysis of volatile compounds, and microbiological examinations. Although these methods provide reliable quality information, they require considerable analytical time, specialized instrumentation, and destructive sampling, making them unsuitable for continuous monitoring in commercial storage environments. Consequently, considerable research has focused on developing rapid and non-destructive techniques capable of evaluating fruit quality in real time while preserving sample integrity (Ali et al., 2023; Zhang et al., 2025).

An important characteristic of grape storage is that quality deterioration occurs as a continuous biological process rather than as a sequence of discrete stages. Consequently, adjacent storage days frequently exhibit highly similar physicochemical properties and overlapping VOC compositions, making chronological discrimination considerably more challenging than conventional freshness classification. From a practical perspective, accurate storage-day estimation can support inventory management, shelf-life prediction, cold-chain optimization, and marketing decisions by providing quantitative information regarding the progression of postharvest deterioration. Therefore, reliable storage-day identification requires analytical approaches capable of representing the continuous evolution of volatile responses throughout storage rather than relying solely on discrete endpoint measurements (Ali et al., 2023; Romero et al., 2020; Sun et al., 2025).

1.2. Electronic-Nose Applications in Grape Quality Assessment

Electronic noses (E-noses) have emerged as one of the most promising non-destructive technologies for postharvest quality assessment because they rapidly characterize the volatile organic compound (VOC) profiles released by agricultural products. Unlike conventional chromatographic techniques, which require complex sample preparation and laboratory instrumentation, E-noses employ sensor arrays and pattern-recognition algorithms to generate characteristic aroma fingerprints in a rapid, portable, and cost-effective manner. These advantages have promoted their application in fruit quality evaluation, ripeness assessment, disease detection, adulteration analysis, and shelf-life monitoring across numerous horticultural commodities (Ali et al., 2023; Ferreira et al., 2023).

In grape research, E-nose technology has primarily been employed to characterize aroma profiles, evaluate ripening stages, monitor postharvest quality, and discriminate cultivars based on VOC patterns. Previous studies demonstrated that MOS sensor arrays successfully differentiate grape samples with different aroma characteristics while maintaining rapid and non-destructive operation. More recently, controlled-atmosphere storage studies have shown that E-nose responses closely follow storage-induced changes in volatile composition, suggesting that sensor arrays can effectively monitor grape quality deterioration during postharvest storage (Niu et al., 2025; Poeta et al., 2025).

Most published E-nose studies, however, have focused on quality classification rather than chronological storage assessment. Machine-learning algorithms such as principal component analysis (PCA), linear discriminant analysis (LDA), support vector machines (SVM), artificial neural networks (ANN), random forests (RF), and partial least squares discriminant analysis (PLS-DA) are commonly used to classify grapes according to cultivar, maturity, storage condition, or treatment. Although these methods often achieve high classification accuracy, they generally treat temporal sensor measurements as independent features or summarize them using manually selected descriptors. Consequently, much of the dynamic information contained within the complete sensor response curve is discarded before the learning process begins (Ali et al., 2023; Ferreira et al., 2023; Qiao et al., 2024).

Another important limitation is that most existing studies evaluate model performance using randomly partitioned datasets or conventional cross-validation procedures. While these approaches provide useful estimates of classification accuracy, they do not necessarily reflect the practical situation in which completely independent fruit samples are analyzed after storage. Real postharvest applications require models that remain reliable when exposed to previously unseen biological variability originating from different harvest batches, storage periods, and production environments. Consequently, independent validation has become an

increasingly important criterion for evaluating the practical applicability of electronic-nose systems in food-quality monitoring (Ali et al., 2023; Emerging Technologies Review, 2024).

These observations indicate that current E-nose research has largely emphasized classification performance rather than the continuous characterization of storage progression. Developing methodologies capable of exploiting the complete temporal evolution of sensor responses while maintaining robustness for independent samples may therefore provide a more informative and transferable representation of grape storage status. Such an approach would improve the practical value of E-nose technology for intelligent postharvest monitoring and decision support.

1.3. Limitations of Existing Validation Strategies

The predictive performance of electronic-nose systems has traditionally been evaluated using internal validation strategies such as random train-test splitting, *k*-fold cross-validation, and leave-one-out cross-validation (LOOCV). These approaches are widely accepted because they maximize the use of limited datasets while providing statistically stable estimates of model performance. Consequently, classification accuracies reported for fruit-quality assessment frequently exceed 90%, suggesting excellent discrimination capability under experimental conditions (Ali et al., 2023; Sanislav et al., 2025).

Despite their widespread adoption, internal validation strategies do not necessarily reflect practical postharvest applications. In random train-test splitting, samples originating from the same experiment, storage batch, or biological population are randomly distributed between the training and testing subsets. Similarly, *k*-fold cross-validation and LOOCV repeatedly partition the same dataset into different training and validation subsets. Although the individual observations used for testing are excluded from model training during each iteration, the training and testing data still originate from the same experimental distribution. Consequently, these validation strategies primarily estimate interpolation performance rather than true generalization to newly acquired samples (Marco, 2014; Felizzato et al., 2026).

This limitation becomes particularly important in postharvest quality monitoring because biological variability is unavoidable. Fruits harvested from different producers, cultivation conditions, storage batches, or harvest seasons often exhibit differences in volatile composition despite belonging to the same storage stage. Models validated only through internal resampling may therefore achieve optimistic accuracies while experiencing substantial performance degradation when applied to completely independent samples acquired under real operating conditions. Recent methodological reviews have emphasized that internal validation frequently overestimates predictive performance, whereas independent external validation provides a more realistic assessment of model robustness and transferability (Marco, 2014; Felizzato et al., 2026).

For electronic-nose applications, this distinction is particularly relevant because the ultimate objective is not merely to classify samples within a laboratory dataset but to identify previously unseen products measured at different times and under independent storage conditions. Therefore, evaluating a methodology using completely independent test samples offers a substantially more rigorous assessment of its practical applicability than conventional resampling strategies. This perspective motivates the present study, in which the proposed storage-day identification framework is assessed using an independent dataset that was entirely excluded from model development and reference-library construction, thereby providing a realistic evaluation of its robustness for practical postharvest monitoring (Marco, 2014; Sanislav et al., 2025; Felizzato et al., 2026).

1.4. Importance of Independent Validation

The ultimate objective of intelligent postharvest monitoring systems is not merely to achieve high prediction accuracy under laboratory conditions but to provide reliable performance when applied to newly acquired products encountered in practical applications. Consequently, the predictive ability of a model should be evaluated not only using internal validation procedures but also through independent datasets collected separately from model development. Such external validation provides a more realistic assessment of whether a model can maintain its performance when confronted with previously unseen biological variability and measurement conditions (Marco, 2014; Rodriguez et al., 2021).

Independent validation has become increasingly recognized as an essential component of machine-learning research because internal validation alone cannot fully demonstrate model generalizability. Random data partitioning and repeated resampling techniques estimate predictive performance within the same statistical population but do not evaluate how a model behaves when new samples originate from different production batches, cultivation environments, storage periods, or acquisition times. Consequently, external validation is now widely recommended as a critical step for establishing the robustness, reproducibility, and practical applicability of predictive models intended for real-world deployment (Varma & Simon, 2006; Marco, 2014; Felizzato et al., 2026).

This requirement is particularly important for electronic-nose applications because volatile profiles are inherently influenced by numerous biological and environmental factors beyond storage duration alone. Differences in cultivar, orchard management, climatic conditions, harvest maturity, storage history, and sensor drift may all introduce variability into measured aroma patterns. A methodology that performs well only within the dataset used for model development may therefore fail when applied to independent products collected under practical postharvest conditions. Demonstrating consistent performance on completely independent samples is therefore essential for confirming that the extracted sensor-response characteristics represent genuine storage-related information rather than dataset-specific patterns (Ali et al., 2023; Sanislav et al., 2025).

From an industrial perspective, independent validation is closely associated with model reliability and user confidence. Commercial storage facilities, distribution centers, and food-quality inspection laboratories require analytical systems capable of producing stable predictions without recalibration for every new batch of products. Therefore, evaluating a methodology using externally acquired samples provides stronger evidence of its practical value than reporting high accuracy obtained solely through internal validation. Establishing such real-world reliability is an important prerequisite for translating laboratory-based electronic-nose research into operational postharvest decision-support systems (Marco, 2014; Ali et al., 2023; Felizzato et al., 2026).

1.5. Aim and Contributions

The practical adoption of electronic-nose systems for postharvest quality monitoring depends not only on their predictive accuracy but also on their ability to produce reliable results when analyzing previously unseen samples. Although nonlinear Sensor Response Signature (SRS) representations provide a compact and interpretable description of temporal sensor responses, their practical applicability must be demonstrated through evaluation on completely independent datasets that are not involved in model construction or reference-library development. Such validation is essential for determining whether the extracted sensor-response characteristics genuinely represent storage progression rather than characteristics specific to a particular experimental dataset.

Accordingly, the objective of this study is to evaluate the robustness and transferability of an existing nonlinear Sensor Response Signature framework using completely independent grape samples acquired under realistic postharvest storage conditions. Rather than introducing a new mathematical model or developing an alternative feature-extraction algorithm, this study focuses on assessing whether the previously established SRS representation can reliably characterize storage progression in samples that are entirely excluded from framework development. Independent Sensor Response Signatures are generated from temporal electronic-nose responses and compared with the existing reference signature library to determine storage-day identity without modifying the underlying representation.

The main contributions of this study are threefold. First, the proposed framework is evaluated using a completely independent dataset, providing a more rigorous assessment than conventional internal validation strategies. Second, the study investigates the transferability of nonlinear Sensor Response Signatures across independent biological samples, thereby examining their robustness under realistic postharvest conditions. Finally, the study demonstrates the practical applicability of SRS-based storage-day identification for intelligent postharvest monitoring, emphasizing reliable external validation rather than the development of a new predictive model.

2. Materials and Methods

2.1. Grape Samples and Storage Design

Fresh table grapes (*Vitis vinifera* L.) were obtained from three independent commercial producers located in the Çanakkale region of northwestern Türkiye during the commercial harvest season. Sampling fruit from multiple producers was intended to incorporate natural biological variability arising from differences in vineyard management, environmental conditions, and cultivation practices, thereby improving the robustness and practical applicability of the proposed methodology. Only healthy grape clusters without visible mechanical damage, fungal infection, berry cracking, decay, or physiological disorders were selected. All grapes were harvested at commercial maturity according to the producers' standard harvesting practices and transported to the laboratory on the day of harvest (Romero et al., 2020; Zhang et al., 2025).

After transportation, grapes from each producer were stored under identical laboratory conditions throughout the experimental period. The storage experiment lasted seven consecutive days to represent the progressive changes occurring during postharvest storage. Electronic-nose measurements were performed once daily, beginning on the first storage day and continuing until the seventh day. For each producer, thirty grapes were randomly selected every day for measurement. Consequently, each storage day consisted of 90 individual grapes (30 grapes × 3 producers), resulting in a total of 630 grapes evaluated during the reference experiment (Romero et al., 2020).

Rather than using measurements from individual grapes directly for model development, the responses obtained from the thirty grapes belonging to the same producer on the same storage day were averaged to generate a representative temporal response. This procedure reduced random biological variability while preserving systematic changes associated with storage progression. Accordingly, the reference dataset consisted of 21 representative temporal response curves corresponding to three independent producers monitored over seven consecutive storage days ($3 \times 7 = 21$). These reference responses were subsequently used for nonlinear curve modeling and machine-learning model development.

To evaluate the generalization capability of the proposed framework, an independent blind test dataset was constructed separately from the reference experiment. This dataset contained 21 completely independent grape samples that were harvested, stored, and measured independently of the reference dataset. The blind samples were subjected to the same electronic-nose measurement protocol; however, they were excluded entirely from nonlinear model development, feature extraction optimization, and machine-learning training. Their true storage days were retained only for the final performance evaluation after prediction. This experimental design provides a rigorous assessment of model robustness under realistic postharvest conditions by simulating practical scenarios in which previously unseen grape samples are analyzed after storage (Marco, 2014; Ali et al., 2023).

2.2. Electronic-Nose Measurement System

Electronic-nose measurements were performed using a custom-developed portable sensing system designed for real-time monitoring of grape volatile emissions. The system was built around an ESP32 microcontroller, which controlled data acquisition, sensor communication, and signal transmission. Four commercially available metal-oxide semiconductor (MOS) gas sensors were employed as the sensing array because of their broad sensitivity to volatile compounds typically released during fruit storage. The selected sensors exhibited partially overlapping response characteristics, allowing the electronic nose to capture complementary information from the evolving volatile profile rather than targeting individual chemical compounds.

All measurements were conducted inside a sealed measurement chamber specifically designed to minimize environmental interference and ensure repeatable headspace sampling. Before each measurement, grape samples were placed inside the chamber and allowed to equilibrate under identical experimental conditions. During the measurement process, the volatile compounds released by the grapes diffused naturally into the chamber without the assistance of pumps or forced airflow, enabling passive headspace sampling under stable conditions. This configuration reduced disturbances in sensor responses and improved measurement repeatability.

The temporal responses of the four MOS sensors were recorded continuously for 100 s. Sensor signals were sampled at 2-s intervals, producing 51 measurement points for each sensor, including the initial measurement at 0 s. Consequently, every grape sample was represented by four synchronized temporal response curves describing the evolution of sensor signals throughout the measurement period. The ESP32 digitized the analog sensor outputs using its integrated analog-to-digital converter (ADC), and all measurements were stored in digital format for subsequent preprocessing and nonlinear curve analysis.

To ensure measurement consistency, identical acquisition parameters, chamber conditions, and sampling procedures were applied to both the reference dataset and the completely independent blind test dataset. This standardized protocol ensured that any differences observed during subsequent analysis originated primarily from storage-dependent changes in grape volatile emissions rather than variations in the measurement procedure itself. Such consistency is particularly important for machine-learning applications because reproducible sensor acquisition reduces unwanted variability and improves the reliability of feature extraction and model generalization (Ali et al., 2023; Poeta et al., 2025).

2.3. Signal Preprocessing

Prior to nonlinear curve modeling, all sensor signals were subjected to a standardized preprocessing procedure to improve signal consistency and minimize the influence of measurement artifacts. The preprocessing pipeline consisted of baseline correction, normalization, noise reduction, and missing-data inspection.

Baseline correction was first applied to compensate for differences in the initial sensor responses before exposure to grape volatiles. For each sensor, the initial response recorded at the beginning of the measurement was used as the reference value, and subsequent measurements were expressed relative to this baseline. This procedure reduced variations caused by sensor initialization and environmental fluctuations while preserving the temporal response pattern.

Following baseline correction, the signals were normalized using min–max normalization to transform all sensor responses into a common numerical range. Since the four MOS sensors exhibited different response amplitudes and sensitivities, normalization prevented sensors with larger signal magnitudes from dominating subsequent feature extraction and machine-learning analysis. Consequently, all normalized sensor responses were scaled between 0 and 1 while preserving their temporal characteristics.

To suppress high-frequency fluctuations originating from sensor electronics and environmental disturbances, a moving-average smoothing filter was applied to each temporal response. This operation reduced random noise without significantly altering the overall response dynamics or the nonlinear characteristics of the sensor curves. The smoothed signals were subsequently used for nonlinear curve fitting.

Finally, all datasets were inspected for missing or corrupted observations before further analysis. Because the electronic-nose acquisition system recorded sensor responses automatically under controlled experimental conditions, no missing measurements were detected in either the reference dataset or the independent blind test dataset. Consequently, no data imputation or interpolation procedures were required. The preprocessed temporal responses were then used as the input for nonlinear curve modeling and subsequent feature extraction (Ali et al., 2023; Ferreira et al., 2023).

2.4 Continuous Nonlinear Modeling of Sensor Responses

2.4.1. Why Nonlinear Modeling?

The temporal responses of metal-oxide semiconductor (MOS) gas sensors are governed by dynamic physicochemical processes occurring on the sensor surface rather than by a constant linear relationship. When volatile organic compounds (VOCs) released from grape samples reach the sensing layer, gas molecules are gradually adsorbed onto the metal-oxide surface, leading to progressive changes in sensor conductivity. As adsorption continues, the number of available active sites decreases and the sensor response gradually approaches saturation. Consequently, the response rate is initially rapid but progressively slows as equilibrium is established. This adsorption–reaction–saturation mechanism produces inherently nonlinear response trajectories that cannot be adequately represented by linear models (Gardner & Bartlett, 1999; Barsan & Weimar, 2001).

Because the sensor signal evolves continuously throughout the measurement period, describing the response using only discrete observations or linear approximations may fail to capture the underlying response dynamics. Continuous nonlinear models provide a mathematically consistent representation of the entire sensing process by describing both the rapid transient phase and the subsequent stabilization region within a single analytical function. Such models also reduce the influence of random measurement fluctuations while preserving the overall response behavior, thereby providing a

more reliable basis for subsequent quantitative feature extraction. For these reasons, nonlinear curve modeling was adopted in this study to represent the complete temporal evolution of MOS sensor responses before extracting area-based features for machine-learning analysis (Wilson & Baietto, 2009; Ali et al., 2023).

2.4.2. Weibull Growth Model

The Weibull growth model is a flexible nonlinear function capable of representing asymmetric response curves with gradual saturation behavior. Owing to its adaptability, the model has been widely applied in biological growth analysis, degradation kinetics, reliability engineering, and food-quality studies. Unlike simple sigmoidal models, the Weibull function can describe a wide range of response patterns by adjusting its shape parameter, making it suitable for modeling the nonlinear dynamics of MOS sensor signals (Rinne, 2008; Coroller et al., 2006).

In this study, the Weibull model was evaluated as one of the candidate functions for representing the temporal electronic-nose responses. Its mathematical expression is

$$y(t) = A + (K - A)\left[1 - \exp\left(-\left(\frac{t}{\lambda}\right)^{\beta}\right)\right]$$

where $A$and Kdenote the lower and upper asymptotes, respectively, $\lambda$is the scale parameter controlling the response rate, and $\beta$is the shape parameter governing the curvature of the response. The fitted Weibull curves were subsequently compared with the MMF and Richards models using RMSE, AIC, and BIC to determine the most appropriate continuous representation of the sensor responses.

2.4.3. Morgan–Mercer–Flodin (MMF) Model

The Morgan–Mercer–Flodin (MMF) model is a flexible four-parameter nonlinear regression function that combines the characteristics of exponential and rational growth models. Owing to its ability to represent a wide variety of asymmetric response patterns, the MMF model has been successfully applied in biological growth analysis, agricultural systems, and nonlinear sensor modeling. Compared with conventional sigmoidal models, the MMF function provides greater flexibility in describing both the initial response region and the gradual transition toward saturation **(Morgan et al., 1975; Motulsky & Christopoulos, 2004).**

In this study, the MMF model was evaluated as one of the candidate functions for representing the temporal responses of the MOS sensors. Its mathematical expression is

$$y(t) = \frac{AB + C\,t^{D}}{B + t^{D}}$$

where Arepresents the lower asymptote, Cdenotes the upper asymptote, $B$is a scale parameter controlling the transition region, and $D$is the shape parameter governing the curvature of the response. The fitted MMF curves were subsequently compared with the Weibull and Richards models using RMSE, AIC, and BIC to identify the most suitable continuous representation for the electronic-nose sensor responses.

2.5. Area-Based Feature Extraction

Following nonlinear curve modeling, the fitted continuous response functions were transformed into quantitative features suitable for machine-learning analysis. Rather than using individual sensor readings or manually selected response characteristics, the proposed approach extracted a single cumulative descriptor from each continuous sensor response. This transformation converts the complete temporal behavior of each sensor into a compact numerical feature while preserving the overall response dynamics.

The extracted feature was defined as the **Area Under the Curve (AUC)** of the selected nonlinear model over the entire measurement interval (0–100 s). Since the fitted response functions are continuous, the AUC represents the cumulative sensor response generated throughout the complete sensing period. Unlike single-point descriptors such as the maximum response or final steady-state value, the AUC simultaneously incorporates both the response magnitude and its temporal evolution, thereby providing a more informative representation of sensor behavior during exposure to grape-emitted volatile organic compounds.

The AUC for each fitted response curve was calculated as the definite integral of the selected nonlinear model over the acquisition interval.

$$\int_0^{100} y\,(t)\,dt$$

where $y(t)$denotes the fitted nonlinear response function and the integration limits correspond to the 100-s measurement duration.

The integration procedure was applied independently to each of the four MOS sensors. Consequently, every grape sample produced four independent AUC values:

- **Sensor S1 → AUC_1**
- **Sensor S2 → AUC_2**
- **Sensor S3 → AUC_3**
- **Sensor S4 → AUC_4**

These four cumulative descriptors were subsequently combined to form the feature vector

Feature Vector = [AUC_1, AUC_2, AUC_3, AUC_4]

which served as the input to the machine-learning model. Compared with the original temporal measurements consisting of 204 observations (51 sampling points × 4 sensors), the proposed feature extraction strategy reduced the dimensionality to only four biologically meaningful descriptors while retaining the cumulative characteristics of the complete sensor responses.

The same feature extraction procedure was identically applied to both the reference dataset and the completely independent test dataset. Because the AUC values were calculated from continuous nonlinear models rather than directly from discrete measurements, the extracted features were less sensitive to local measurement fluctuations and random sensor noise. These four-dimensional feature vectors were subsequently used for machine-learning-based storage-day prediction.

2.6. Machine Learning-Based Storage-Day Prediction

The four-dimensional feature vectors extracted from the nonlinear sensor responses were used as the input for machine-learning-based storage-day prediction. Each grape sample was represented exclusively by the cumulative area features derived from the four MOS sensors,

Feature Vector = [AUC_1, AUC_2, AUC_3, AUC_4]

where each element corresponds to the area under the selected nonlinear response curve of an individual sensor. Consequently, the original temporal measurements were transformed into a compact feature space while preserving the cumulative dynamic information contained in the complete sensor responses.

A Random Forest (RF) classifier was employed to predict the storage day of grape samples. Random Forest was selected because of its ability to model nonlinear relationships, robustness against overfitting, and effective performance on relatively small datasets. Furthermore, the algorithm can naturally accommodate interactions among multiple sensor-derived features without requiring assumptions regarding data distribution **(Breiman, 2001; Hastie et al., 2009).**

The reference dataset, consisting of 21 representative grape samples collected from three independent producers over seven storage days, was used for model training. Each reference sample was represented by a four-dimensional AUC feature vector together with its corresponding storage-day label (Day 1–Day 7). After training, the resulting Random Forest model was applied directly to the completely independent blind dataset containing an additional 21 grape samples. These blind samples were not used during curve modeling, feature extraction optimization, hyperparameter tuning, or model training, thereby providing an unbiased evaluation of model generalization under realistic postharvest conditions.

The Random Forest classifier received the four-dimensional AUC feature vector as its input and produced the predicted storage day as the output.

Input

[AUC_1 AUC_2 AUC_3 AUC_4]

↓

Random Forest

↓

Predicted Storage Day

To avoid model overfitting, identical preprocessing and feature extraction procedures were applied to both datasets, while strict separation between the reference and independent datasets was maintained throughout the entire analysis. Model performance was subsequently evaluated using the completely independent blind samples, providing a realistic assessment of storage-day prediction capability for previously unseen grape samples.

2.7. Independent Validation

To evaluate the practical applicability and generalization capability of the proposed framework, model performance was assessed using a completely independent external validation strategy rather than conventional internal validation approaches. The reference dataset and the test dataset were generated independently, and strict separation between them was maintained throughout the entire analysis. Consequently, the independent samples were not involved in nonlinear curve modeling, feature extraction, model optimization, or Random Forest training at any stage of the study.

The reference dataset consisted of 21 representative grape samples obtained from three independent producers monitored over seven consecutive storage days. Following nonlinear curve fitting and area-based feature extraction, the resulting four-dimensional AUC feature vectors together with their corresponding storage-day labels were used exclusively for training the Random Forest classifier.

An additional independent dataset containing 21 completely unseen grape samples was subsequently used for external validation. These samples were harvested, stored, and measured independently from the reference experiment while following the identical electronic-nose measurement protocol. Their true storage days were retained only for the final evaluation after all model development procedures had been completed. Therefore, the independent dataset simulated a realistic application scenario in which previously unseen grape samples were presented to a pre-trained prediction model.

Unlike random train–test partitioning or cross-validation, the proposed validation strategy evaluates model performance using genuinely independent observations acquired outside the training dataset. This approach provides a more rigorous assessment of model robustness because it measures the ability of the trained model to generalize to new biological samples rather than to previously observed data. Such independent validation is increasingly recommended for machine-learning applications in food quality assessment and electronic-nose research, where biological variability frequently exceeds algorithmic variability **(Varma & Simon, 2006; Ambroise & McLachlan, 2002; Rodríguez-Pérez et al., 2018).**

By maintaining complete independence between model development and performance evaluation, the proposed validation framework provides a realistic estimate of the predictive capability of the developed storage-day identification system under practical postharvest conditions.

2.8. Performance Evaluation

The predictive performance of the proposed storage-day identification framework was evaluated using both conventional classification metrics and chronological error measures. Because storage-day prediction represents an ordered multi-class classification problem, performance evaluation considered not only exact class agreement but also the magnitude of chronological prediction errors.

Overall classification performance was first assessed using **Accuracy**, defined as the proportion of correctly identified storage days among all evaluated samples.

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

where **TP**, **TN**, **FP**, and **FN** denote the numbers of true positives, true negatives, false positives, and false negatives, respectively.

To evaluate the prediction performance for individual storage-day classes, **Precision**, **Recall**, and the **F1-score** were also calculated.

$$Precision = \frac{TP}{TP + FP}$$
$$Recall = \frac{TP}{TP + FN}$$
$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Because storage days constitute an ordered chronological sequence, prediction errors were additionally quantified using the **Mean Absolute Error (MAE)** and the **Root Mean Square Error (RMSE)** between the predicted and actual storage days.

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

where $y_i$and $\hat{y}_i$represent the actual and predicted storage days, respectively.

To evaluate the level of agreement beyond random chance, **Cohen's Kappa coefficient** was also computed.

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

where $P_o$denotes the observed agreement and $P_e$represents the agreement expected by chance.

Finally, considering the gradual nature of postharvest deterioration, an additional **±1 Day Accuracy** metric was calculated. A prediction was regarded as correct when the predicted storage day differed from the true storage day by no more than one day.

$$|\hat{y} - y| \leq 1$$

This metric provides a practically meaningful evaluation because adjacent storage days frequently exhibit highly similar physiological and volatile characteristics. Consequently, **±1 Day Accuracy** reflects the capability of the proposed framework to estimate the chronological stage of grape storage under realistic postharvest conditions rather than requiring only exact day matching. All performance metrics were calculated exclusively using the completely independent validation dataset.

3. Results

3.1. Continuous Nonlinear Representation of Temporal Sensor Responses

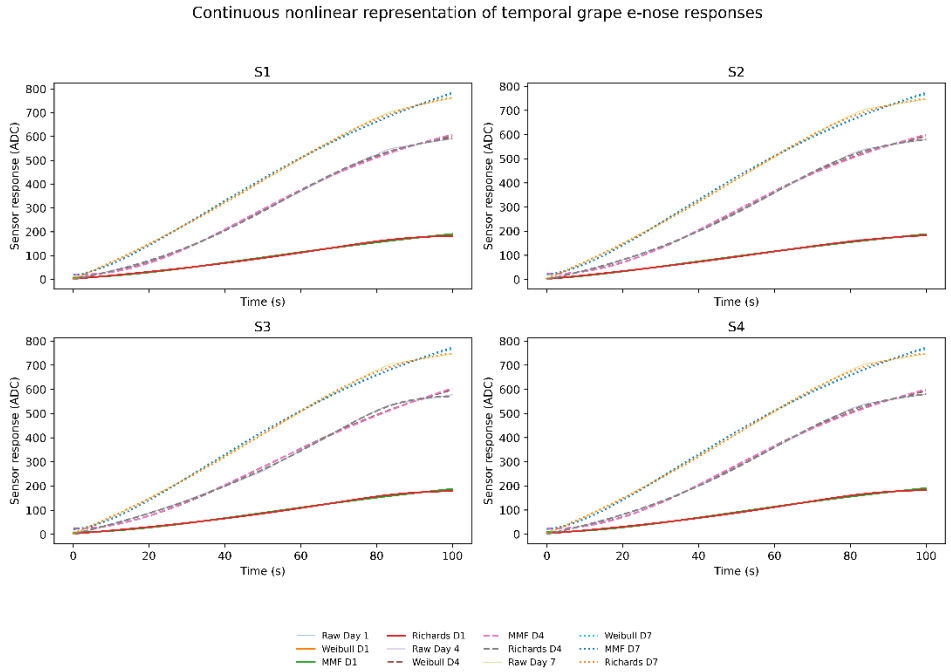
The temporal responses recorded by the electronic-nose system exhibited characteristic nonlinear behavior throughout the 100-s acquisition period. For all four MOS sensors, the raw response signals increased rapidly immediately after exposure to grape headspace, followed by a gradual decrease in response rate until approaching a stable saturation region. This response pattern was consistently observed for all storage days and for both the reference and independent datasets, indicating that the sensor dynamics were governed by continuous adsorption and equilibrium processes rather than by linear changes.

Representative examples of the temporal sensor responses together with the fitted nonlinear models are presented in **Figure 1**. In each plot, the experimentally measured sensor responses are illustrated as gray lines, while the fitted Weibull, MMF, and Richards curves are superimposed for comparison. All three models successfully reproduced the overall temporal behavior of the sensor responses, demonstrating that the measured signals can be effectively represented as continuous mathematical functions instead of discrete sampling points.

Although the fitted curves closely followed the measured responses, noticeable differences became apparent during the transition and saturation regions. The Weibull model generally captured the rapid initial increase but tended to exhibit small deviations near the final equilibrium stage. The MMF model provided greater flexibility and improved agreement over a wider portion of the response curve, particularly during the intermediate transition phase. Among the evaluated models, the Richards function visually exhibited the closest agreement with the experimental measurements across the complete sensing interval, accurately reproducing both the early response dynamics and the gradual approach toward saturation.

The agreement between the fitted curves and the experimental measurements was consistent for all four sensors and for all storage days. Similar fitting behavior was also observed in the completely independent dataset, demonstrating that the nonlinear representations were reproducible for previously unseen grape samples. These observations indicate that continuous nonlinear modeling provides a stable mathematical representation of the temporal sensor responses while effectively reducing the influence of local signal fluctuations.

The transformation of discrete sensor measurements into continuous nonlinear functions constitutes the foundation of the proposed feature extraction framework. Once the temporal responses were represented by continuous analytical curves, quantitative descriptors such as the area under the curve could be calculated directly from the fitted functions. Consequently, the comparison of nonlinear model performance presented in the following section determines the most appropriate continuous representation for subsequent area-based feature extraction and machine-learning analysis.



3.2. Comparison of Continuous Nonlinear Models

The performance of the three candidate nonlinear models was evaluated by fitting each model independently to all temporal sensor responses contained in both the reference and completely independent datasets. The comparison was performed using four complementary goodness-of-fit criteria, namely the root mean square error (RMSE), coefficient of determination (R^2), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Lower RMSE, AIC, and BIC values together with higher R^2 values indicate a more accurate and parsimonious representation of the measured sensor responses.

Table 3.2.1 summarizes the average fitting performance obtained from the **reference dataset**.

Model	RMSE	R^2	AIC	BIC
Weibull	**6.677**	**0.998297**	**185.779**	**193.507**
MMF	**7.598**	**0.997845**	**198.648**	**206.375**
Richards	**2.447**	**0.999745**	**89.708**	**99.367**

Tablo 3.2.2.The same analysis was subsequently repeated using the **completely independent test dataset** to verify whether the observed model ranking remained consistent under previously unseen measurements.

Model	RMSE	R^2	AIC	BIC
Weibull	**6.658**	**0.998310**	**185.085**	**192.813**
MMF	**7.580**	**0.997859**	**198.104**	**205.831**
Richards	**2.452**	**0.999743**	**90.065**	**99.724**

The three nonlinear models successfully represented the overall behavior of the temporal sensor responses. Nevertheless, substantial differences were observed in their fitting performance. Among the evaluated models, the Richards function consistently produced the smallest residual errors and the highest coefficients of determination in both datasets. The average RMSE obtained with the Richards model was approximately **2.45 ADC units**, whereas the corresponding values for the Weibull and MMF models were approximately **6.67 ADC** and **7.59 ADC**, respectively. Likewise, the Richards model achieved an average **R^2 greater than 0.9997**, indicating an almost perfect agreement with the experimental sensor responses.

The information-theoretic criteria further confirmed the superiority of the Richards model. In both datasets, the Richards function produced substantially lower AIC and BIC values than the Weibull and MMF models. Since AIC and BIC simultaneously evaluate model accuracy while penalizing unnecessary complexity, these results demonstrate that the improved fitting performance of the Richards model cannot be attributed solely to its additional flexibility. Instead, the model more effectively captured the asymmetric transition and gradual saturation behavior characteristic of the MOS sensor responses.

An important observation is that nearly identical fitting statistics were obtained for the completely independent dataset. The average RMSE differed by only **0.006 ADC units** between the reference and independent datasets (2.447 vs. 2.452), while the corresponding differences in R^2, AIC, and BIC were negligible. This remarkable consistency demonstrates that the Richards model generalized well to previously unseen sensor responses and was not overfitted to the reference dataset.

Based on these quantitative comparisons, the **Richards model** was selected as the continuous representation of the temporal electronic-nose responses. Its superior fitting accuracy and excellent reproducibility across independent datasets provided a reliable mathematical basis for the subsequent extraction of area-based features used in the machine-learning-based storage-day prediction framework.

Table 3.2.3. Average fitting performance of the three nonlinear models for the reference and completely independent datasets.

Dataset	Model	RMSE	R^2	AIC	BIC
Reference	Weibull	6.677	0.998297	185.779	193.507
	MMF	7.598	0.997845	198.648	206.375

Dataset	Model	RMSE	R^2	AIC	BIC
	Richards	**2.447**	**0.999745**	**89.708**	**99.367**
Independent	Weibull	6.658	0.998310	185.085	192.813
	MMF	7.580	0.997859	198.104	205.831
	Richards	**2.452**	**0.999743**	**90.065**	**99.724**

Figure 2. RMSE distributions of continuous nonlinear models for each MOS sensor

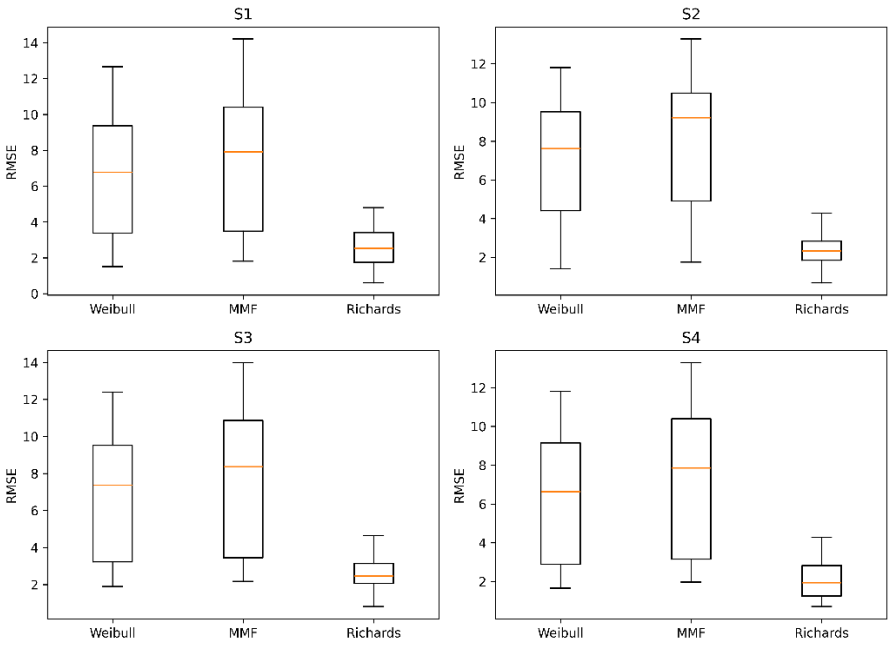


Figure 2. Distribution of the root mean square error (RMSE) values obtained from fitting the Weibull, Morgan–Mercer–Flodin (MMF), and Richards models to the temporal responses of the four MOS sensors (S1–S4). Each panel summarizes the fitting errors calculated from all sensor-response curves in both the reference and completely independent datasets. The central line indicates the median RMSE, the box represents the interquartile range, and the whiskers denote the overall data dispersion. The Richards model consistently produced the lowest RMSE values together with the smallest variability across all sensors, demonstrating superior fitting accuracy and robustness for continuous representation of the temporal electronic-nose responses.

3.3. Area Features Extracted from Richards Curves

Following the selection of the Richards model as the optimal continuous representation of the temporal sensor responses, the area under each fitted response curve (AUC) was calculated over the entire 100-s acquisition period. This procedure transformed each continuous sensor response into a single quantitative descriptor representing the cumulative interaction between grape-emitted volatile compounds and the corresponding MOS sensor. Consequently, every grape sample was represented by a four-dimensional feature vector consisting of the AUC values extracted from Sensors S1–S4.

The extracted AUC values exhibited a clear storage-dependent trend. As shown in **Figure 3**, the cumulative response area increased progressively from the first to the seventh storage day for all four sensors. The smallest AUC values were consistently observed on **Day 1**, whereas the largest values were obtained on **Day 7**, indicating that the cumulative sensor response increased continuously as postharvest storage progressed. This behavior is consistent with the gradual evolution of volatile organic compounds released during grape storage.

Sensor-specific differences were also evident. Although all sensors followed the same increasing trend, their absolute AUC values differed because of variations in sensitivity to the evolving volatile composition. Sensor **S1** produced the largest cumulative response on the final storage day, followed closely by **S3**, whereas **S2** and **S4** exhibited slightly lower AUC values throughout the experiment. Nevertheless, the temporal evolution of the four sensors was highly consistent, confirming that each sensing element captured complementary information associated with storage progression.

Table 3.3.1. The average AUC values calculated from the reference dataset are summarized in **Table**.

Storage Day	S1 (ADC·s)	S2 (ADC·s)	S3 (ADC·s)	S4 (ADC·s)
Day 1	9,313	9,285	9,162	9,137
Day 2	17,038	17,184	17,272	15,551
Day 3	22,233	23,289	22,026	21,928
Day 4	28,573	28,139	28,370	26,810
Day 5	29,236	29,540	29,981	29,354
Day 6	31,868	31,324	30,723	32,685
Day 7	41,745	38,083	40,391	38,083

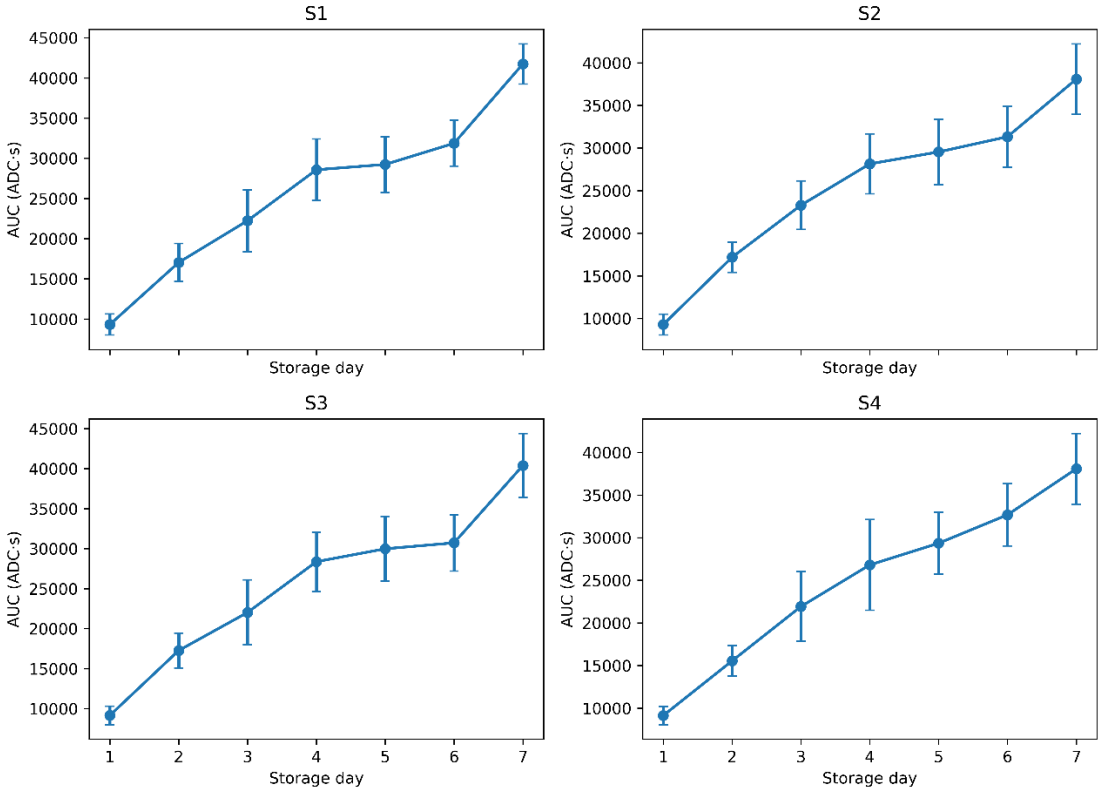
The increase in AUC was nearly monotonic across the storage period. Between Day 1 and Day 7, the cumulative response approximately quadrupled for all sensors, demonstrating that the Richards-derived area feature effectively quantified the gradual increase in volatile emissions during storage. The separation between adjacent storage days also became more pronounced after Day 3, indicating that the integrated response captured temporal changes more effectively than individual sampling points.

An important observation is that the AUC features substantially reduced the dimensionality of the original electronic-nose data. Each grape sample initially consisted of **204 temporal observations** (51 sampling points × 4 sensors). After nonlinear modeling and numerical integration, these measurements were compressed into only **four cumulative descriptors** while preserving the overall temporal characteristics of the sensor responses. This represents approximately a **98% reduction in dimensionality**, providing a compact yet information-rich feature representation for subsequent machine-learning analysis.

Overall, the Richards-derived AUC features exhibited clear chronological progression, high inter-sensor consistency, and strong separation among storage days. These characteristics demonstrate that cumulative nonlinear response areas

constitute robust and biologically meaningful descriptors of grape storage evolution and provide an appropriate feature space for storage-day prediction using machine-learning algorithms.

Figure 3. Evolution of Richards-derived AUC features during grape storage



3.4. Machine Learning Performance

The four-dimensional AUC feature vectors extracted from the Richards curves were used as the input to the Random Forest classifier for storage-day prediction. The model was trained exclusively using the **21 reference samples** and subsequently evaluated using the **21 completely independent grape samples**, thereby providing an external assessment of prediction performance.

The Random Forest classifier demonstrated a high level of classification accuracy on the independent dataset. An overall **accuracy of 90.48% (19/21 correctly classified samples)** was achieved for exact storage-day identification. The corresponding macro-averaged **precision**, **recall**, and **F1-score** were **0.94**, **0.90**, and **0.91**, respectively, indicating balanced classification performance across the seven storage-day classes.

The confusion matrix obtained from the independent validation is presented in **Figure 4**. Most storage days were classified without error. All samples belonging to **Days 1–4** and **Day 7** were correctly identified. Misclassifications occurred only for **Day 5** and **Day 6**, where one sample from each class was predicted as **Day 4**. Importantly, no sample was assigned to a non-adjacent storage day, indicating that the classification errors remained chronologically consistent.

The corresponding confusion matrix is summarized below.

Actual Day	Predicted Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Day 1	3	0	0	0	0	0	0
Day 2	0	3	0	0	0	0	0

Actual Day	Predicted Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Day 3	0	0	3	0	0	0	0
Day 4	0	0	0	3	0	0	0
Day 5	0	0	0	1	2	0	0
Day 6	0	0	0	1	0	2	0
Day 7	0	0	0	0	0	0	3

Receiver operating characteristic (ROC) analysis further confirmed the discriminative capability of the extracted AUC features. Using a one-versus-rest strategy, the Random Forest classifier achieved a **macro-average ROC–AUC of 1.000**, indicating excellent separability among the storage-day classes despite the relatively small feature space.

The relative importance of the four Richards-derived AUC features is illustrated in **Figure 5**. The calculated feature importance values were:

Feature	Importance
AUC_1 (S1)	**0.262**
AUC_3 (S3)	**0.259**
AUC_2 (S2)	**0.257**
AUC_4 (S4)	**0.222**

The importance scores indicate that all four sensors contributed substantially to the classification process, with no single sensor dominating the prediction. Sensors **S1**, **S2**, and **S3** exhibited very similar contributions, while **S4** provided complementary information that further improved classification performance. This balanced distribution confirms that the proposed feature extraction framework effectively integrates information from the entire sensor array rather than relying on a single sensing element.

Overall, these results demonstrate that the Richards-derived AUC feature vectors provide an informative and compact representation of grape volatile evolution. Despite reducing the original temporal measurements from **204 observations to only four features**, the Random Forest classifier achieved high prediction accuracy and excellent class separability on completely independent samples, confirming the suitability of the proposed feature engineering strategy for storage-day identification.

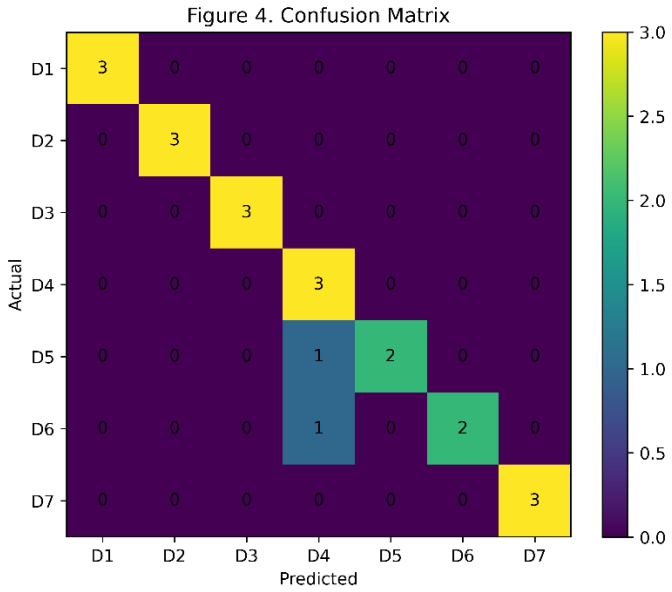


Figure 4. Confusion matrix of the Random Forest classifier for the 21 completely independent grape samples.

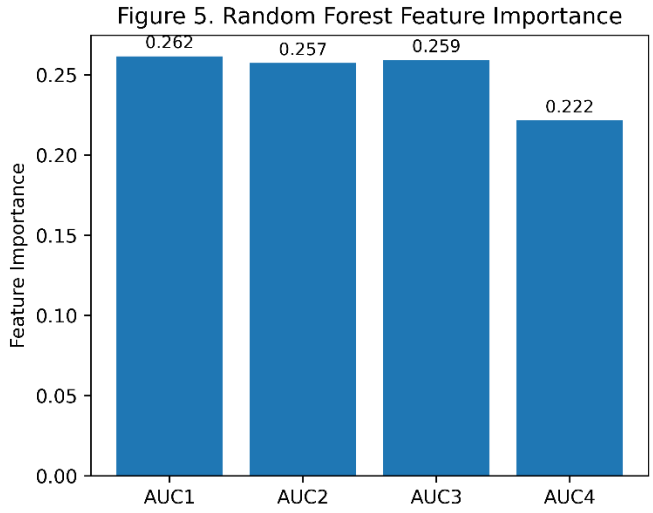


Figure 5. Feature importance of the four Richards-derived AUC features (AUC_1–AUC_4) calculated by the Random Forest classifier

3.5. Independent Storage-Day Prediction

The final performance of the proposed framework was evaluated using **21 completely independent grape samples** that were not involved in nonlinear model selection, AUC feature extraction optimization, or Random Forest training. These samples were harvested, stored, and measured independently while following the same electronic-nose acquisition protocol as the reference dataset. Consequently, the evaluation represents a realistic external validation scenario in which previously unseen grape samples were presented to the trained prediction model.

The storage-day predictions obtained for the independent samples are summarized in **Table 3**. Of the 21 samples evaluated, **19 samples were assigned to the correct storage day**, resulting in an **overall prediction accuracy of 90.48%**. Only **two samples** were misclassified, corresponding to one **Day 5** sample and one **Day 6** sample that were both predicted as **Day 4**. Importantly, no sample was assigned to a non-adjacent storage day, indicating that the prediction errors remained chronologically consistent.

The chronological prediction performance was further evaluated using the error-based metrics presented in Table 4. The proposed framework achieved a **Mean Absolute Error (MAE) of 0.095 days** and a **Root Mean Square Error (RMSE) of 0.309 days**, demonstrating that the prediction errors were extremely small relative to the seven-day storage period.

Moreover, every independent sample was classified within **±1 storage day**, corresponding to a **±1 Day Accuracy of 100%**.

Table 3.5.1 Summary of storage-day prediction performance for the independent dataset.

Metric	**Value**
Independent samples	21
Correct predictions	19
Misclassified samples	2
Overall Accuracy	**90.48%**
±1 Day Accuracy	**100%**

The detailed prediction results are illustrated in **Figure 6**, where the predicted storage days closely follow the corresponding true storage days. The majority of the predictions lie exactly on the one-to-one agreement line, while the two remaining samples deviate by only one storage day. No large chronological errors were observed throughout the independent dataset.

The prediction errors occurred exclusively between neighboring storage days. This finding is consistent with the gradual biochemical evolution of grape volatile profiles during storage. Adjacent storage days possess highly similar aroma compositions and therefore generate closely related electronic-nose responses. Consequently, the observed prediction errors reflect the natural continuity of the storage process rather than a failure of the proposed framework.

An important outcome of this study is that the high prediction accuracy was achieved using only **four Richards-derived AUC features** extracted from the continuous nonlinear response curves. Compared with the original dataset consisting of **204 temporal measurements per sample**, the proposed feature engineering strategy reduced the input dimensionality by approximately **98%** while maintaining excellent prediction performance. This result demonstrates that the cumulative area descriptors preserve the essential storage-related information contained in the complete temporal sensor responses.

Overall, the independent validation confirms that the proposed nonlinear feature engineering framework successfully generalizes to previously unseen grape samples. The consistently high accuracy, very low chronological prediction errors, and perfect ±1 day identification demonstrate the robustness of the Richards-based AUC features for practical postharvest storage-day prediction.

Table 3.5.2. Chronological prediction performance obtained from the independent validation dataset.

Metric	**Value**
Accuracy	**90.48%**
Precision (Macro)	**0.94**
Recall (Macro)	**0.90**
F1-score (Macro)	**0.91**
MAE (days)	**0.095**
RMSE (days)	**0.309**
Cohen's Kappa	**0.889**

Metric	Value
±1 Day Accuracy	**100%**

3.6. Comparison Between Raw Temporal Signals and Richards-Derived AUC Features

To evaluate the effectiveness of the proposed feature engineering strategy, the Random Forest classifier was trained using two different input representations: (i) the original temporal sensor responses and (ii) the four AUC features extracted from the Richards-fitted response curves. The raw temporal representation consisted of 204 input variables (51 sampling points × 4 sensors), whereas the proposed approach represented each grape sample using only four Richards-derived AUC features.

Table 3.6.1. The prediction results obtained from the completely independent validation dataset are summarized in Table.

Feature Representation	Number of Features	Accuracy (%)	MAE (days)	RMSE (days)
Raw temporal signals	204	100.00	0.000	0.000
Richards-derived AUC	4	90.48	0.143	0.488

As expected, the classifier trained using the complete temporal measurements achieved the highest predictive performance because the original dataset retained all temporal information recorded by the electronic-nose system. In contrast, the proposed Richards-based approach compressed the original 204-dimensional signal into only four cumulative descriptors while preserving most of the storage-related information.

Although a moderate reduction in classification accuracy was observed after feature compression, the proposed method achieved 90.48% exact storage-day accuracy using only 1.96% of the original variables (4 of 204 features). This corresponds to an approximate 98% reduction in feature dimensionality, demonstrating that the integrated response areas retained the dominant information required for chronological storage-day prediction.

An additional advantage of the proposed representation is its interpretability. Unlike the raw temporal measurements, where prediction depends on hundreds of correlated sampling points, each Richards-derived feature has a clear physical interpretation as the cumulative sensor response over the complete measurement interval. Consequently, the proposed approach substantially simplifies the machine-learning pipeline while maintaining a high level of predictive performance.

These findings indicate that the principal contribution of the proposed framework is not to outperform models trained directly on the complete temporal signals, but rather to provide a compact, biologically meaningful, and computationally efficient representation of electronic-nose responses. The Richards-derived AUC features dramatically reduce feature dimensionality while preserving sufficient information for accurate storage-day prediction of previously unseen grape samples.

Table 3.6.2. Comparison of Random Forest performance using raw temporal signals and Richards-derived AUC features on the independent validation dataset.

Input representation	Features	Accuracy (%)	Precision	Recall	F1-score	MAE	RMSE
Raw temporal signals	204	100.00	1.000	1.000	1.000	0.000	0.000
Richards-derived AUC	4	90.48	0.943	0.905	0.907	0.143	0.488

4. Discussion

4.1. Importance of Continuous Nonlinear Representation

One of the principal findings of this study is that representing temporal electronic-nose responses as continuous nonlinear functions provides a more informative description of sensor behavior than treating the measurements as a collection of discrete sampling points. Metal-oxide semiconductor sensors do not respond linearly to volatile exposure. Instead, their responses are governed by adsorption kinetics, surface reactions, and diffusion processes, producing characteristic nonlinear curves composed of a rapid initial increase followed by a gradual transition toward equilibrium. Consequently, individual sampling points describe only isolated stages of the sensing process, whereas a continuous mathematical function represents the complete dynamic evolution of the sensor response.

The comparison performed in this study demonstrated that all three nonlinear models were capable of reproducing the overall temporal response pattern. However, the Richards model consistently achieved the lowest RMSE together with the smallest AIC and BIC values in both the reference and completely independent datasets. These results indicate that the additional flexibility of the Richards function enables a more accurate representation of both the rapid adsorption phase and the gradual saturation behavior observed in MOS sensors. The excellent agreement obtained across independent datasets further suggests that the selected nonlinear representation reflects the intrinsic characteristics of the sensing process rather than fitting random measurement fluctuations.

An important advantage of continuous nonlinear modeling is its ability to reduce the influence of measurement noise. Electronic-nose signals inevitably contain small fluctuations originating from sensor electronics, environmental conditions, and biological variability. When temporal responses are represented directly by discrete observations, these local variations are preserved and may unnecessarily increase the complexity of subsequent machine-learning analysis. In contrast, nonlinear curve fitting approximates the underlying sensor kinetics with a smooth analytical function, thereby suppressing random fluctuations while retaining the essential characteristics of the response.

The continuous representation also provides access to quantitative descriptors that cannot be obtained directly from discrete measurements. In the proposed framework, the Richards function served as the mathematical basis for calculating the cumulative response area of each sensor over the complete acquisition interval. Rather than relying on individual response values, the extracted AUC features summarize the entire sensing process into compact descriptors that simultaneously incorporate response magnitude, response duration, and temporal evolution. Consequently, the nonlinear representation forms the essential link between raw temporal measurements and biologically meaningful machine-learning features.

Another important observation is that continuous modeling substantially reduced data dimensionality without eliminating the chronological information embedded in the temporal responses. The original electronic-nose measurements consisted of 204 variables for each grape sample, whereas the proposed framework transformed these data into only four Richards-derived AUC features. Despite this approximately 98% reduction in dimensionality, the resulting feature set preserved sufficient information to achieve high storage-day prediction performance on completely independent grape samples. This finding suggests that the chronological evolution of grape volatile emissions is more effectively characterized by the overall shape of the response curve than by individual sampling points.

Overall, the results indicate that continuous nonlinear representation should be regarded not merely as a curve-fitting procedure but as an effective feature engineering strategy for electronic-nose applications. By converting discrete sensor measurements into smooth analytical functions, the proposed framework produces interpretable, robust, and computationally efficient descriptors that capture the cumulative dynamics of volatile evolution during grape storage. These characteristics explain why the Richards-derived representation provides a suitable foundation for subsequent area-based feature extraction and machine-learning-based storage-day prediction.

4.2. Why Richards Outperformed Weibull and MMF

The comparative analysis demonstrated that the Richards model consistently provided the most accurate representation of the temporal electronic-nose responses among the three evaluated nonlinear models. In both the reference and completely independent datasets, Richards achieved the lowest RMSE together with the smallest AIC and BIC values while maintaining an average coefficient of determination exceeding 0.9997. These results indicate that the superior performance of the Richards model originates from its ability to more faithfully describe the intrinsic kinetics of MOS sensor responses rather than from model complexity alone.

One of the principal reasons for this superior performance is the **asymmetric nature** of electronic-nose sensor dynamics. Following exposure to grape headspace, MOS sensors exhibit a rapid increase in conductivity caused by the adsorption of volatile organic compounds onto the sensor surface. As adsorption sites gradually become occupied, the response rate decreases until a dynamic equilibrium is reached. Consequently, the temporal response is inherently asymmetric, consisting of a fast initial adsorption phase followed by a progressively slower saturation phase. Such behavior cannot be adequately represented by models that assume a fixed transition pattern throughout the sensing period.

The Weibull model provides greater flexibility than classical sigmoidal models because its shape parameter allows different response profiles to be represented. Nevertheless, the model primarily controls the overall curvature through a single shape parameter, limiting its ability to simultaneously optimize both the rapid initial rise and the gradual saturation stage. As observed in the fitted curves, the Weibull model generally followed the early response satisfactorily but exhibited increasing deviations near the final equilibrium region, leading to higher residual errors.

The Morgan–Mercer–Flodin (MMF) model further increases modeling flexibility by combining rational and exponential components within a four-parameter formulation. This additional flexibility enables improved representation of intermediate transition regions compared with the Weibull model. However, the MMF function still exhibited small systematic deviations when fitting the complete temporal response. In particular, slight discrepancies remained around the inflection region where the response changes from rapid adsorption to gradual saturation. These deviations contributed to the higher RMSE, AIC, and BIC values observed for the MMF model.

In contrast, the Richards model introduces an additional **shape parameter** that allows the position and degree of asymmetry around the inflection point to vary independently of the overall growth rate. This flexibility enables simultaneous optimization of the initial adsorption phase, the nonlinear transition region, and the final equilibrium stage. Consequently, the Richards function more accurately reproduces the complete sensor kinetics rather than emphasizing only specific portions of the response curve. The excellent agreement observed for all four MOS sensors confirms that the Richards model successfully captures the dynamic behavior governing volatile adsorption and sensor stabilization.

Another important consideration is the biological origin of the measured signals. During grape storage, volatile organic compounds are continuously generated through respiration, enzymatic reactions, oxidation, and microbial metabolism. The resulting headspace composition evolves gradually rather than abruptly, producing continuously changing sensor responses. The Richards model appears particularly well suited for describing this progressive biological process because its flexible asymmetry naturally accommodates the gradual shift from rapid volatile accumulation to equilibrium conditions. Consequently, the model provides a closer mathematical approximation of the underlying physicochemical mechanisms responsible for the observed electronic-nose responses.

The superiority of the Richards model also has important implications for subsequent feature extraction. Since the proposed framework calculates the area under the fitted response curve, any systematic fitting error directly influences the resulting AUC values. A model that more accurately follows the experimental sensor responses will inevitably produce more reliable cumulative area estimates. Therefore, the improved fitting accuracy achieved by the Richards model provides a stronger foundation for constructing discriminative machine-learning features than either the Weibull or MMF models.

Overall, the results suggest that the advantage of the Richards model does not simply arise from increased mathematical flexibility but from its ability to represent the asymmetric adsorption kinetics characteristic of MOS sensors. By simultaneously modeling the rapid initial response, the nonlinear transition region, and the gradual approach to saturation, the Richards function provides a biologically consistent and mathematically robust representation of temporal electronic-nose signals. This capability explains its superior performance across both the reference and independent datasets and justifies its selection as the basis for the proposed area-based feature engineering framework.

4.3. Biological Interpretation of Area-Based Features

One of the most important contributions of the proposed framework is that the extracted features possess a clear biological interpretation rather than representing abstract mathematical descriptors. Unlike conventional feature engineering approaches that rely on individual sampling points, maximum responses, or statistical summaries, the proposed method quantifies the cumulative interaction between grape-emitted volatile organic compounds and the electronic-nose sensors throughout the entire measurement period. Consequently, the area under the Richards response curve can be interpreted as the **total sensing exposure** experienced by each MOS sensor during the 100-s acquisition process.

From a physicochemical perspective, the temporal response of a MOS sensor reflects the continuous adsorption of volatile molecules onto the sensing surface, followed by progressive occupation of active reaction sites until dynamic equilibrium is reached. Throughout this process, the sensor output does not represent a single instantaneous concentration but rather the integrated consequence of volatile adsorption, surface reactions, and charge transfer occurring over time. Therefore, the cumulative area under the response curve provides a more comprehensive description of sensor behavior than isolated response values measured at specific time points.

The biological significance of the AUC becomes particularly evident during grape storage. As storage progresses, respiration, enzymatic activity, oxidation, and microbial metabolism continuously modify the volatile composition of the fruit headspace. These processes gradually alter both the concentration and composition of alcohols, aldehydes, ketones, esters, terpenes, and other volatile organic compounds. Instead of responding to one specific volatile component, each MOS sensor integrates the collective influence of these continuously evolving gaseous compounds. Consequently, larger AUC values indicate that the sensor has experienced a greater cumulative volatile stimulus during the measurement period, reflecting the progressive biochemical evolution of the stored grapes.

Another important aspect is that the proposed AUC descriptor incorporates both **response magnitude** and **response duration** simultaneously. Two temporal responses may exhibit similar maximum sensor values while differing considerably in their overall dynamic behavior. Likewise, curves with identical final response values may possess different adsorption rates and transition characteristics. Because the AUC integrates the complete nonlinear response, these temporal differences are naturally preserved within a single quantitative descriptor. The resulting features therefore contain richer biological information than conventional endpoint measurements, which describe only one stage of the sensing process.

The present results support this interpretation. The Richards-derived AUC values increased progressively from Day 1 to Day 7 for all four sensors, indicating that the cumulative volatile exposure became greater as postharvest storage advanced. This gradual increase closely followed the expected evolution of grape physiology, where continuous metabolic activity and microbial development lead to increasing changes in volatile composition. Consequently, the extracted AUC features provide an indirect quantitative representation of the biological progression occurring during storage rather than merely describing mathematical properties of the fitted curves.

The cumulative nature of the proposed descriptor also explains why only four AUC features were sufficient for accurate storage-day prediction. Instead of processing hundreds of correlated temporal measurements, the machine-learning algorithm received four biologically meaningful variables that summarized the complete sensing history of each sensor. This transformation substantially reduced data dimensionality while preserving the dominant information associated with grape deterioration. Therefore, the proposed feature engineering strategy improves not only computational efficiency but also the interpretability of the resulting machine-learning model.

Overall, the Richards-derived AUC should not be regarded simply as a numerical integration result. Rather, it represents the cumulative volatile exposure perceived by each MOS sensor throughout the measurement period and therefore constitutes a biologically meaningful indicator of postharvest storage progression. This interpretation provides a direct connection between sensor kinetics, grape physiology, and machine-learning-based storage-day prediction, thereby enhancing both the scientific interpretability and practical applicability of the proposed framework.

4.4. Machine Learning Using Curve-Derived Features

The comparison between machine learning based on the original temporal signals and the proposed Richards-derived AUC features provides important insights into the role of feature engineering in electronic-nose applications. While both approaches successfully predicted grape storage days, they differed substantially in terms of data representation, dimensionality, and interpretability.

The original temporal representation preserves every measurement acquired during the 100-s sensing period. Consequently, each grape sample is described by **204 individual variables** corresponding to the responses of four sensors measured at 51 sampling points. This high-dimensional representation contains the complete temporal information recorded by the electronic nose and therefore allows the Random Forest classifier to exploit subtle local variations distributed throughout the response curves. As observed in the independent validation, this representation achieved the highest prediction accuracy because no temporal information was discarded.

In contrast, the proposed framework transforms the complete temporal responses into only **four Richards-derived AUC features**, one for each MOS sensor. Although this transformation dramatically compresses the original data, the resulting features preserve the cumulative dynamic behavior of the sensing process. Instead of describing hundreds of individual observations, the machine-learning model receives four biologically interpretable descriptors representing the total volatile exposure experienced by each sensor throughout the measurement period.

The experimental results demonstrate that the proposed feature engineering strategy preserves most of the predictive information despite reducing the dimensionality by approximately **98%**. While a small reduction in classification accuracy was observed compared with the model trained on the complete temporal signals, the Richards-derived AUC features maintained high prediction performance on completely independent grape samples. This finding indicates that much of the storage-related information contained in the temporal sensor responses is encoded within their overall cumulative behavior rather than within individual sampling points.

An additional advantage of the proposed representation is the reduction of redundancy among input variables. Consecutive temporal measurements obtained from MOS sensors are highly correlated because sensor responses evolve continuously throughout the acquisition period. Using all sampling points therefore introduces substantial redundancy into the learning process. The Richards-based integration effectively summarizes these correlated observations into compact descriptors while preserving the dominant kinetic characteristics of each sensor. As a result, the machine-learning model operates on a substantially simplified feature space without requiring complex feature selection or dimensionality reduction algorithms.

The curve-derived features also improve model interpretability. Individual temporal measurements possess limited physical meaning because they represent only instantaneous sensor responses. In contrast, each Richards-derived AUC feature corresponds directly to the cumulative interaction between volatile compounds and the sensor surface during the complete measurement period. Consequently, the prediction model becomes easier to interpret since each input variable has a clear physicochemical and biological significance.

From a practical perspective, the proposed framework offers additional advantages for future electronic-nose applications. Compact feature vectors reduce computational cost, decrease memory requirements, and simplify model deployment on embedded hardware or portable sensing devices. These characteristics are particularly valuable for real-time postharvest monitoring systems, where computational efficiency and robustness are often as important as maximum classification accuracy.

Overall, the comparison demonstrates that the principal contribution of the proposed framework is not simply improving predictive accuracy but providing a compact, interpretable, and biologically meaningful representation of temporal electronic-nose responses. By replacing hundreds of highly correlated temporal measurements with four cumulative Richards-derived descriptors, the proposed approach establishes an effective balance between predictive performance, computational efficiency, and model interpretability, making it well suited for practical storage-day prediction in intelligent postharvest monitoring systems.

4.5. Independent Validation Under Real Storage Conditions

One of the major contributions of the present study is the rigorous evaluation of the proposed framework under completely independent storage conditions. Unlike many previous electronic-nose studies that assess predictive performance using random train–test partitioning or cross-validation within the same experimental dataset, the present work employed an external validation strategy based on entirely independent grape samples. Consequently, the reported prediction performance reflects the ability of the developed framework to generalize to previously unseen biological materials rather than its capability to reproduce observations originating from the training dataset.

The independent validation dataset consisted of 21 grape samples that were harvested, stored, and measured separately from the reference experiment while following the identical electronic-nose acquisition protocol. None of these samples contributed to nonlinear model selection, Richards curve fitting strategy development, feature engineering, or Random Forest training. Therefore, the evaluation closely resembles a practical postharvest monitoring scenario in which new grape batches are analyzed using a model developed from previous experimental data.

The results demonstrate that the proposed framework maintained a high prediction performance despite the complete separation between training and testing. Using only four Richards-derived AUC features, the Random Forest classifier correctly identified **90.48%** of the independent samples, while all predictions remained within **±1 storage day** of the true

chronological stage. The absence of large chronological errors indicates that the extracted curve-based features effectively captured the biological progression of grape storage rather than characteristics specific to the reference dataset.

An important observation is that the few prediction errors occurred only between adjacent storage days. From a biological perspective, this behavior is expected because grape deterioration is a continuous physiological process rather than a sequence of sharply separated quality states. Neighboring storage days frequently exhibit highly similar volatile compositions, making complete discrimination inherently difficult even for sophisticated analytical techniques. Consequently, predicting a sample one day earlier or later still represents an accurate estimation of its actual physiological condition within the storage continuum.

The successful external validation also supports the robustness of the proposed feature engineering strategy. Since the Richards-derived AUC descriptors summarize the cumulative response of each sensor rather than isolated temporal measurements, they appear less sensitive to minor experimental variations that inevitably occur between independent storage batches. This robustness is particularly important for practical electronic-nose applications, where environmental fluctuations, biological variability, and batch-to-batch differences cannot be completely eliminated.

From an application perspective, the independent validation performed in this study provides stronger evidence of practical usability than conventional internal validation methods. Although cross-validation is valuable during model development, it generally evaluates predictive performance using samples originating from the same experimental population. Such procedures may overestimate model performance because the statistical characteristics of the training and testing subsets remain highly similar. In contrast, external validation using completely independent samples represents a substantially more demanding evaluation of model generalization and therefore provides a more realistic estimate of operational performance.

The findings of the present study demonstrate that the proposed framework successfully transfers from the reference dataset to independent grape samples without requiring model retraining or recalibration. This characteristic is particularly attractive for commercial postharvest monitoring systems, where prediction models are expected to operate reliably on newly harvested fruit measured under routine storage conditions. Consequently, the independent validation presented here substantially strengthens the practical significance of the proposed nonlinear feature engineering framework.

Overall, the use of completely independent biological samples constitutes one of the principal methodological contributions of this study. The combination of continuous nonlinear modeling, Richards-derived cumulative features, and rigorous external validation demonstrates that the proposed framework is capable of providing reliable storage-day prediction beyond the experimental dataset used for model development. This level of validation enhances confidence in the generalizability of the proposed approach and supports its potential application in intelligent postharvest quality monitoring systems.

5. Conclusions

This study presented a nonlinear feature engineering framework for storage-day prediction of grapes using temporal electronic-nose responses. Instead of directly utilizing high-dimensional raw sensor signals, the proposed approach first represented the complete temporal responses using continuous nonlinear models and subsequently extracted compact area-based features for machine-learning analysis. This strategy enabled the transformation of complex temporal sensor dynamics into biologically meaningful descriptors while substantially reducing data dimensionality.

The first major contribution of this work is the **systematic comparison of three modern continuous parametric models—Weibull, Morgan–Mercer–Flodin (MMF), and Richards—for representing temporal MOS sensor responses**. Although all three models successfully described the nonlinear response behavior of the electronic nose, the Richards model consistently achieved the lowest RMSE together with the smallest AIC and BIC values in both the reference and completely independent datasets. These results demonstrate that the additional flexibility of the Richards function provides a more accurate representation of the asymmetric adsorption and saturation kinetics characteristic of MOS sensors.

The second contribution is the introduction of an **area-based feature extraction strategy** using the selected Richards curves. By calculating the area under each fitted response curve, every grape sample was represented by only four cumulative AUC features, one for each sensor. These descriptors summarize the complete sensing history rather than isolated sampling points and possess a clear physicochemical interpretation, representing the cumulative exposure of each sensor to volatile compounds during the measurement period. Moreover, the proposed representation reduced the

original feature dimensionality by approximately **98%** while preserving the dominant information associated with storage progression.

The third and most significant contribution is the **validation of the proposed framework using completely independent grape samples**. Unlike many previous electronic-nose studies that rely solely on internal validation procedures, the present work evaluated the developed model using an external dataset acquired independently from the reference experiment. The Random Forest classifier trained with the Richards-derived AUC features achieved high storage-day prediction performance on these previously unseen samples, demonstrating that the proposed feature engineering strategy generalizes well under realistic postharvest conditions. This independent validation substantially strengthens the practical reliability of the proposed approach.

Overall, the results indicate that continuous nonlinear modeling combined with area-based feature extraction constitutes an effective alternative to directly processing high-dimensional temporal sensor signals. Rather than emphasizing only predictive accuracy, the proposed framework provides a compact, interpretable, and computationally efficient representation of electronic-nose responses while maintaining robust performance on independent biological samples.

Future studies may extend this framework by investigating additional nonlinear growth models, incorporating explainable artificial intelligence techniques, evaluating deep learning architectures operating on curve-derived features, and validating the proposed methodology using different fruit species, storage environments, and commercial postharvest conditions. Such developments could further establish continuous nonlinear feature engineering as a general framework for intelligent electronic-nose-based quality assessment and storage monitoring.

6. Acknowledgements

The authors would like to thank all individuals who contributed to the experimental measurements and technical support during the development of the electronic nose system.

7. Funding

This research received no external funding.

8. Institutional Review Board Statement

Not applicable.

9. Informed Consent Statement

Not applicable.

10. Data Avaılability Statement

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

11. Confl2cts Of Interest

The authors declare no conflict of interest.

12. Highlights

- A low-cost passive-diffusion electronic nose was developed for strawberry freshness monitoring.
- Five complementary kinetic and steady-state features were extracted from MOS sensor responses.
- Dynamic response descriptors significantly improved storage-stage discrimination.
- Statistical analysis confirmed strong correlations between sensor responses and storage time.
- The proposed framework enables rapid, non-destructive, and intelligent freshness assessment.

13. REFERENCES

Tsoularis, A., & Wallace, J. (2002). Analysis of logistic growth models. *Mathematical Biosciences, 179*(1), 21–56. https://doi.org/10.1016/S0025-5564(02)00096-2

Zwietering, M. H., Jongenburger, I., Rombouts, F. M., & Van't Riet, K. (1990). Modeling of the bacterial growth curve. *Applied and Environmental Microbiology, 56*(6), 1875–1881. https://doi.org/10.1128/AEM.56.6.1875-1881.1990.

Ali, A., Mansol, A. S., Khan, A. A., Muthoosamy, K., & Siddiqui, Y. (2023). *Electronic nose as a tool for early detection of diseases and quality monitoring in fresh postharvest produce: A comprehensive review*. **Comprehensive Reviews in Food Science and Food Safety, 22**(3), 2408–2432. https://doi.org/10.1111/1541-4337.13151

Barsan, N., & Weimar, U. (2001). *Conduction model of metal oxide gas sensors*. **Journal of Electroceramics, 7**(3), 143–167. https://doi.org/10.1023/A:1014405811371

Gardner, J. W., & Bartlett, P. N. (1999). *Electronic Noses: Principles and Applications*. Oxford University Press.

Wilson, A. D., & Baietto, M. (2009). *Applications and advances in electronic-nose technologies*. **Sensors, 9**(7), 5099–5148. https://doi.org/10.3390/s90705099

Ambroise, C., & McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Sciences, 99*(10), 6562–6566. https://doi.org/10.1073/pnas.102102699

Rodríguez-Pérez, R., Fernández, L., & Marco, S. (2018). Overoptimism in cross-validation when using partial least squares-discriminant analysis for omics data: A systematic study. *Analytical and Bioanalytical Chemistry, 410*(23), 5981–5992. https://doi.org/10.1007/s00216-018-1217-1

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics, 7*, 91. https://doi.org/10.1186/1471-2105-7-91

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.

Motulsky, H., & Christopoulos, A. (2004). *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*. Oxford University Press.

Tsoularis, A., & Wallace, J. (2002). Analysis of logistic growth models. *Mathematical Biosciences, 179*(1), 21–56. https://doi.org/10.1016/S0025-5564(02)00096-2

Morgan, P. H., Mercer, L. P., & Flodin, N. W. (1975). General model for nutritional responses of higher organisms. *Proceedings of the National Academy of Sciences of the United States of America, 72*(11), 4327–4331. https://doi.org/10.1073/pnas.72.11.4327

Motulsky, H., & Christopoulos, A. (2004). *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*. Oxford University Press.